



ELO-BASIERTE POTENZIALANALYSE 2024 **METHODENPAPIER**

Korrespondenz

Geschäftsstelle der PotAS-Kommission
Graurheindorfer Straße 198/6
53117 Bonn

✉ info@potas.de

☎ +49 (0) 228 99640 9004

Inhalt

Einleitung	1
Einführung in Gracenotes Podium	2
Elo-Rating-System (ERS)	2
Weitere Algorithmen	5
PotAS ⁺⁶	7
Disziplinspezifische Entwicklungsraten	7
Aktueller Elo-Wert der Athletinnen & Athleten	8
Vergleich mit der Weltspitze	9
Transfer ins Attributesystem	10
Team-Events	11
Mannschaftssportarten	11
Diskussion	11

Einleitung

Die PotAS-Kommission analysiert die Potenziale der vom DOSB vorgegebenen Disziplingruppen der olympischen Spitzenverbände anhand von transparenten, sportwissenschaftlich und sportfachlich leistungsrelevanten Kriterien. Sie liefert damit eine objektive und unabhängige Entscheidungsgrundlage für die nachfolgenden Schritte des Förderverfahrens, d. h. den Strukturgesprächen mit den Spitzenverbänden und der Sitzung der Förderkommission. Das Attributesystem der Potenzialanalyse umfasst derzeit eine Dreiteilung aus den Säulen ERFOLG, KADERPOTENZIAL und STRUKTUR. Während in den Säulen ERFOLG und STRUKTUR etablierte Bewertungsverfahren vorliegen, befindet sich die Säule KADERPOTENZIAL seit der Einführung von PotAS in einer steten Entwicklung. Ziel dieser Säule ist es, Athletinnen und Athleten mit Erfolgspotenzial (Top-3/Top-8) für die nächsten Olympischen Winter- und Sommerspiele mit einer bestimmten Wahrscheinlichkeit zu identifizieren.

Die analytische Systematik in der Säule KADERPOTENZIAL eröffnete schon in der Vergangenheit immer wieder Diskussionen darüber, wie zukünftige Medaillen- oder Finalplatzpotenziale überhaupt stichhaltig erfasst werden können. In der

Wintersportanalyse 2018 sowie Sommersportanalyse 2019-2021 waren die Verbände selbst in der Verantwortung, unter Beteiligung der Verbandsvertretenden des DOSB entsprechende Listen mit Einschätzungen über ihre Athletinnen und Athleten bereitzustellen, die zur Einhaltung der Datenschutz-Grundverordnung (DSGVO) jedoch anonymisiert wurden. Daraus ergaben sich gleich mehrere Limitationen, weil seitens der PotAS-Kommission weder die angezeigten Potenziale noch die Entscheidungsprozesse der Verbände überprüfbar waren. Vor allem die Entscheidungsprozesse warfen Fragen bezüglich ihrer Objektivität, Transparenz und Datenqualität auf, zumal die verbandspezifischen Potenzialeinschätzungen im strategischen Kontext der Fördervergabe zu bewerten sind. Die PotAS-Kommission entschied sich daher im Vorfeld der Wintersportanalyse 2022 für eine neue Systematik, die die Zuhilfenahme eines umfangreichen Datenpakets des Technologieunternehmens Gracenote beinhaltet.

Nach aktuellem Stand besteht die Säule KADERPOTENZIAL aus einem einzigen Hauptattribut, dem Hauptattribut 4 (HA 4). HA 4 enthält zwei Fragen, die beide mithilfe von Gracenote bewertet werden (vgl. [Abb. 1](#)). Die nachfolgenden Ausführungen geben einen Überblick darüber, welche Daten verwendet und wie sie für die Potenzialanalysen verarbeitet werden.

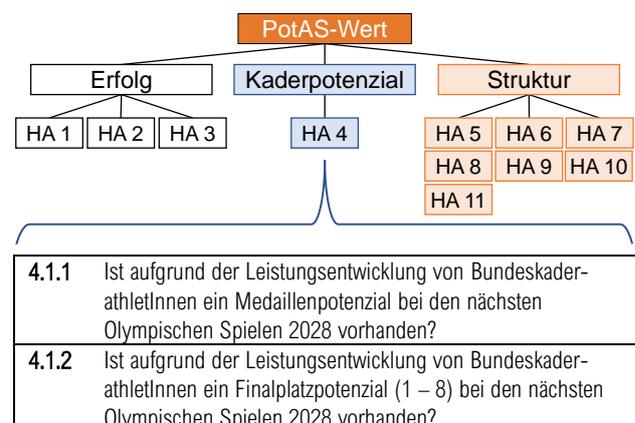


Abb. 1 Der Aufbau des Potenzialanalysesystems für den Sommersport 2024 und die im Hauptattribut 4 enthaltenen Fragen.

Einführung in Gracenotes Podium

Gracenote Inc. ist ein unter dem Dach der Nielsen Company (Nielsen Holdings plc, London, UK) agierendes Unternehmen, welches nach eigenen Angaben eine der weltweit größten Sportdatenbanken betreibt¹. Mit dem Produkt „Gracenote Podium“ werden alle relevanten Sportergebnisse olympischer und nicht-olympischer Sportarten tagesaktuell erfasst. Die NutzerInnen haben Zugriff auf Daten zu Individual- und Mannschaftssportarten und können diese je nach Bedarf systematisch filtern und zu weiteren Analysezielen exportieren. Das der PotAS-Kommission zur Verfügung stehende Datenpaket umfasst alle Platzierungen von Athletinnen und Athleten der olympischen Sommer- und Wintersportarten seit 2004. Darin enthalten sind die Ergebnisse der Olympischen Spiele, Welt- und Kontinentalmeisterschaften (inkl. Weltcups), Jugend-events sowie weitere, wie z. B. der Nations League oder olympischer Qualifikationsturniere. Neben den jeweiligen Platzierungen in den einzelnen Wettbewerben hat Gracenote zudem das sog. Elo-Rating-System (ERS) in die Datenbank integriert. Dieses ermöglicht die Bewertung des Leistungsniveaus von Athletinnen und Athleten im internationalen Kontext und bildet die Basis des von der PotAS-Kommission entwickelten Modells zur Potenzialeinschätzung innerhalb der deutschen olympischen Spitzenverbände.

Elo-Rating-System (ERS)

Das allgemeine Ziel von Rating-Systemen im Sport ist, die Stärke bzw. Leistungsfähigkeit von Kontrahentinnen und Kontrahenten zu vergleichen. Im günstigsten Fall liefern solche Systeme nicht nur ordinale, sondern intervallskalierte Daten, die Individuen oder ganze Teams in ein messbares Verhältnis setzen. Daraus können beispielsweise Informationen über den wahrscheinlichsten Ausgang von Wettkämpfen gewonnen oder Entscheidungshilfen generiert werden, wenn es um die Nominierung aussichtsreicher Athletinnen, Athleten oder Teams geht.

Nicht nur in Sport und Wissenschaft, sondern auch im Online-Gaming haben sich mittlerweile solche

selbst-adjustierenden Rating-Systeme durchgesetzt, die eine hohe prognostische Validität für Wettbewerbe in naher Zukunft aufweisen und beispielsweise regressions-basierten Modellen überlegen zu sein scheinen²⁻⁷. Bekannte Beispiele sind das von Microsoft entwickelte TrueSkill™^{8,9} oder das Glicko-System¹⁰, die beide auf bayesscher Statistik (Wahrscheinlichkeitsverteilungen) beruhen.

Das ERS ist eines der ersten selbst-adjustierenden Systeme, die weit verbreitete Anwendung fanden, und gilt als Vorbild für unzählige Nachfolger. Das Prinzip des ERS wurde in seiner ursprünglichen Form erstmals 1961 von Arpad E. Elo publiziert, fand jedoch schon 1960 Anwendung in der United States Chess Federation¹¹. 1970 wurde es von der World Chess Federation adaptiert¹². Die Grundidee des ERS ist, dass leistungsstarke Personen schwächere Personen nie ausnahmslos übertreffen. Vielmehr erzielen alle Athletinnen und Athleten durchschnittliche Leistungen auf einem bestimmten Niveau, von dem sie mittelfristig und in Abhängigkeit von Tagesform, Trainingsstatus etc. annähernd normalverteilt abweichen¹². Große Abweichungen treten demnach seltener auf als kleine. Bei einer ausreichenden Zahl von Wettkämpfen kann es folglich vorkommen, dass schwache Personen mit guter Tagesform stärkere Personen mit schlechter Tagesform besiegen. [Abbildung 2a](#) skizziert die Wahrscheinlichkeit für ein solches Ereignis anhand der relativen Häufigkeiten, mit denen beispielhaft zwei Athleten eine bestimmte Leistung erbringen. Die Häufigkeit, mit der Athlet B gegen A gewinnt, wird durch die blau gefüllte Fläche beschrieben. In allen anderen Fällen gewinnt Athlet A (rot gefüllte Fläche).

Eine weitere Annahme ist, dass es in Bezug auf das Leistungsniveau eine transitive Rangfolge innerhalb einer Grundgesamtheit von Athletinnen und Athleten gibt. Ist also Athlet A leistungsstärker als Athlet B und Athlet B leistungsstärker als Athlet C, dann ist A ebenfalls leistungsstärker als C. Abgebildet wird dieser Leistungsunterschied anhand numerischer Werte (\cong Elo-Werte), die das Leistungsniveau einer Person relativ zu dem der Kontrahentinnen oder Kontrahenten darstellen und die Berechnung von

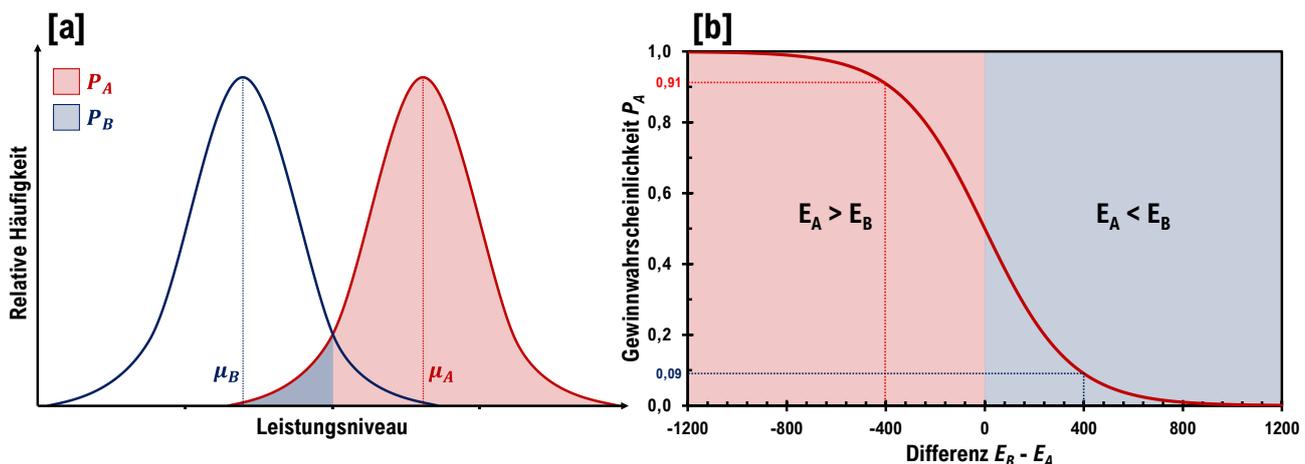


Abb. 2 [a] Annahme einer normalverteilten Abweichung des Leistungsniveaus (Elo-Wert) vom mittleren Leistungsniveau (μ) zweier Athleten A (μ_A) und B (μ_B). Kleinere Abweichungen treten demnach deutlich häufiger auf als große. Die gefüllten Flächen stellen die Gewinnwahrscheinlichkeiten (P) dar, mit der Athlet A (P_A , rote Fläche) und Athlet B (P_B , blaue Fläche) gewinnen, wenn sie gegeneinander antreten. [b] Gewinnwahrscheinlichkeit für Athlet A in Abhängigkeit von der Differenz zwischen den Elo-Werten von A und B. Ist der Elo-Wert von Athlet A (E_A) 400 Punkte höher als der von Athlet B (E_B), so ist die Gewinnwahrscheinlichkeit für A zehnmal höher als für B (~91 % vs. ~9 %). Sind beide Athleten gleich stark, liegt die Gewinnwahrscheinlichkeit für beide bei exakt 50 %¹².

Gewinnwahrscheinlichkeiten zulassen (vgl. Abb. 2b). Wenn zwei Personen also gegeneinander antreten, berechnet das ERS anhand der individuellen Elo-Werte die Wahrscheinlichkeit, mit der die beiden den Wettkampf gewinnen. Die Person, die die Begegnung tatsächlich gewinnt, erhält Elo-Punkte dazu, während die Person, die verliert, Elo-Punkte einbüßt. Die Anzahl der Punkte, die gewonnen oder verloren werden kann, hängt von der Differenz der Elo-Werte beider Personen und dem Ausgang der Begegnung ab. Im Ergebnis lassen sich tagesaktuelle, disziplinspezifische Ranglisten über die Stärke von Athletinnen und Athleten einer Sportart, individuelle Entwicklungsverläufe über die Zeit sowie Prognosen für unmittelbar bevorstehende Sportereignisse ableiten.

Das wichtigste Merkmal eines ERS ist, dass es sich selbstständig anpasst. Wenn eine Person mit einem niedrigen Elo-Wert eine Person mit einem hohen Elo-Wert besiegt, gewinnt sie mehr Elo-Punkte hinzu, als wenn sie eine Person mit einer ähnlich niedrigen Wertung besiegt hätte. Das heißt, wenn AthletInnen mit einer niedrigeren Wertung regelmäßig AthletInnen mit einer höheren Wertung besiegen, wird deren eigene Wertung zunehmend steigen und ihr wahres Leistungsniveau widerspiegeln.

Die konzeptionellen Vorteile eines ERS gegenüber anderen statistischen Modellen sind:

- Ihre Dynamik: Das Elo-Rating aller Athletinnen oder Athleten wird nach jedem Wettkampf auf eine leicht interpretierbare Weise aktualisiert.
- Ihre implizite Gewichtung jüngerer Wettbewerbe: Elo-Ratings geben den letzten Wettkämpfen implizit mehr Gewicht und vermeiden die Notwendigkeit explizit auswählen zu müssen, wie viele Daten aus der Vergangenheit in welcher Gewichtung verwendet werden sollen (wie z. B. beim exponentiell gewichteten, gleitenden Mittelwert).
- Sie bieten eine allgemeine Methode zur Verfolgung von Leistungsveränderungen, ohne ein bestimmtes Modell (und konkrete Zubringergrößen für dieses Modell) vorauszusetzen.¹³

Da es sich hierbei um ein ergebnisbasiertes System handelt, das einzig die Platzierung als Kriterium berücksichtigt, wird es ferner nicht von externen Wettkampffaktoren wie Witterungs- oder Materialbedingungen beeinflusst. Limitierend ist hingegen, dass die realen Leistungsunterschiede (z. B. Tordifferenz, Zeitvorsprung), die zu Sieg oder Niederlage führen, nicht in das klassische Elo-Modell einfließen. Es existieren jedoch Modelle, bei denen dies bereits gelungen ist^{6,7}. Die spezifische Umsetzung und Hinzunahme weiterer Faktoren können sich von System zu System unterscheiden. Gracenote berücksichtigt beispielsweise die Qualität der Begegnung bzw. des

Startfeldes oder die Bedeutung des Wettbewerbes (s. u.).

Die Mathematik jedes ERS beruht im Wesentlichen auf zwei Basisformeln. Formel 1 beschreibt die Gewinnwahrscheinlichkeit (P) für eine Person A, die gegen eine Person B antritt.

$$P_A = \frac{1}{1 + c \frac{E_B - E_A}{d}} \quad (1)$$

E_B und E_A entsprechen den Elo-Werten der Personen, c ist ein Skalierungsfaktor, der sowohl im ursprünglichen ERS als auch bei Gracenote auf $c = 10$ festgesetzt ist. d ist eine Konstante, die historisch gewachsen ist und in vielen Systemen, wie auch Gracenote, bei $d = 400$ liegt. Eine Differenz zwischen zwei Athletinnen oder Athleten von d Elo-Punkten bewirkt eine zehnmal höhere Gewinnwahrscheinlichkeit für die stärkere Person.

Besitzt ein Athlet A bspw. 2.500 Elo-Punkte und ein Athlet B 2.100 Elo-Punkte, so ergibt sich eine Gewinnwahrscheinlichkeit für A entsprechend der Formel 2:

$$P_A = \frac{1}{1 + 10 \frac{2.100 - 2.500}{400}} = \frac{1}{1 + 10^{-1}} \quad (2)$$

$$= \frac{1}{1 + 0,1} \approx 0,91$$

Nachdem die Athleten A und B nun gegeneinander angetreten sind, steht ein tatsächlich eingetretenes Ergebnis fest. Zur Berechnung des neuen Elo-Wertes E' für Athlet A wird die zweite Basisformel herangezogen:

$$E'_A = E_A + k \cdot (S_A - P_A) \quad (3)$$

S_A beschreibt das tatsächliche Ergebnis, d. h., ob Athlet A gewonnen ($S = 1$), verloren ($S = 0$) oder ein Unentschieden erzielt hat ($S = 0,5$). k ist ein elementarer Gewichtungsfaktor, der darüber entscheidet, wie stark die erbrachte Leistung honoriert bzw. bestraft wird. Ein größeres k wird i. d. R. vor allem dann gewählt, wenn die Häufigkeit von Wettbewerben bzw. Begegnungen innerhalb einer Sportart gering ist oder das Leistungsniveau innerhalb der betrachteten

Grundgesamtheit stark schwankt. Denn ein höheres k gibt der letzten erbrachten Leistung implizit mehr Gewicht und führt zu größeren Veränderungen in der Elo-Rangfolge. Der k -Faktor von Gracenote setzt sich zum einen aus dem sog. Competition-Factor (CF) zusammen, der sowohl die Güte des Wettbewerbs als auch die Größe und das Leistungsniveau des dort startenden Teilnehmerfeldes berücksichtigt. Zum anderen besteht er aus einem Teil, der sich je nach Sportart unterscheidet und, wie zuvor beschrieben, von der Anzahl an Ereignissen in der jeweiligen Sportart sowie der Leistungsvariabilität abhängt.¹⁴

Um das vorangegangene Beispiel fortzuführen, seien für Athlet A eine Gewinnwahrscheinlichkeit von $P_A = 91\%$ und $k = 32$ gegeben. A verliert die Begegnung gegen B, woraus $S_A = 0$ und anschließendes Ergebnis folgt:

$$E'_A = 2.500 + 32 \cdot (0 - 0,91) \approx 2.471 \quad (4)$$

Für Athlet B gilt gleichzeitig:

$$E'_B = 2.100 + 32 \cdot (1 - 0,09) \approx 2.129 \quad (5)$$

Athlet A verliert 29 Elo-Punkte und Athlet B gewinnt 29 Elo-Punkte.

Die beschriebene Systematik bezieht sich zunächst auf Begegnungen im Format Eins-gegen-Eins. Adaptionen für das Format Alle-gegen-Alle sind zahlreich. Gracenote übernimmt hierfür die erste Basisformel (Formel 1) und ersetzt den gegnerischen Elo-Wert durch den Mittelwert (\bar{E}_i) aller Wettkampfteilnehmerinnen und -teilnehmer:

$$P_{A,i} = \frac{1}{1 + c \frac{\bar{E}_i - E_A}{d}} \quad (6)$$

Wobei P nicht mehr der Gewinnwahrscheinlichkeit entspricht, sondern der benötigten relativen Anzahl der zu besiegenden Gegnerinnen oder Gegner, um weder Elo-Punkte zu verlieren noch zu gewinnen. Treten beispielsweise vier Athleten mit $E_A = 2.300$, $E_B = 2.400$, $E_C = 2.100$ und $E_D = 2.000$ Elo-Punkten gegeneinander an, so ist der Mittelwert $\bar{E}_i = 2.200$. Daraus ergibt sich für $P_{A,i}$:

$$P_{A,i} = \frac{1}{1 + 10^{\frac{2.200 - 2.300}{400}}} \quad (7)$$

$$= 0,64$$

Athlet A muss demnach 64 % aller Gegner besiegen, um weder Elo-Punkte zu verlieren noch zu gewinnen. Will Athlet A im gegebenen Beispiel Elo-Punkte gewinnen, so muss er mindestens zwei der drei Gegner besiegen ($\approx 67\%$).

Zur Berechnung des neuen Elo-Wertes wird die zweite Basisformel (Formel 3) unverändert übernommen. Das tatsächliche Ergebnis, also der relative Anteil der Gegner, die Athlet A tatsächlich besiegt hat (S_A), liegt zwischen $S = 0$ (letzter Platz) und $S = 1$ (erster Platz), abhängig von der Anzahl an Teilnehmer n sowie der Platzierung von A (ϵ_A):

$$S_{A,i} = \frac{n - \epsilon_A}{n - 1} \quad (8)$$

Belegt Athlet A Platz 1 und wird $k = 32$ angenommen, so folgt daraus für den neuen Elo-Wert:

$$E'_A = E_A + k \cdot \left(\frac{n - \epsilon_A}{n - 1} - P_{A,i} \right)$$

$$= 2.300 + 32 \cdot \left(\frac{4 - 1}{4 - 1} - 0,64 \right) \quad (9)$$

$$\approx 2.312$$

Abschließend ist nochmals zu betonen, dass die hier dargestellten Formeln nur die Grundlage des ERS von Gracenote bilden und sämtliche Rechenbeispiele lediglich Approximationen des deutlich komplexeren Algorithmus sind. Gracenote integriert eine Bandbreite weiterer Funktionen, die sich in jeder Sportart unterscheiden können und das ERS im Sinne einer höchstmöglichen prognostischen Validität ergänzen.

Weitere Algorithmen

Gracenote nutzt weitere Algorithmen, die zu einer präziseren Modellierung des ERS beitragen sollen. Einer resultiert im bereits genannten Competition-Factor, der wiederum unmittelbar Einfluss auf den k-Faktor nimmt.

Während die exakte Formel zum Firmengeheimnis von Gracenote gehört, wurde das Prinzip in persönlicher Kommunikation wie folgt beschrieben: Bei der Kalkulation des CF werden die besten Teilnehmerinnen und Teilnehmer des Wettbewerbs mit den besten der aktuellen Elo-Rangliste verglichen. Die Anzahl an Vergleichspersonen hängt dabei von der Anzahl an Wettkampfteilnehmerinnen bzw. -teilnehmern ab. Je weniger Teilnehmerinnen und Teilnehmer, desto höher der relative Anteil der einbezogenen Vergleichspersonen. Damit wird sichergestellt, dass die Größe des Startfeldes ein Faktor ist und gleichzeitig die besten Teilnehmerinnen und Teilnehmer den größten Anteil am CF haben. Der CF steigt entsprechend der Qualität und Größe des Startfeldes. Abschließend wird die Bedeutung des Wettbewerbes einbezogen. Somit haben Olympische Spiele oder Weltmeisterschaften i. d. R. einen größeren Einfluss als Kontinentalmeisterschaften oder Jugendevents.

Des Weiteren fließt derzeit in der Leichtathletik, im Schwimmen, Kunstturnen und Gewichtheben eine sog. „Four-Year Rolling Personal Season-Best“, d. h. die mittlere Bestleistung der letzten vier Jahre in die Berechnung neuer Elo-Werte ein. Gracenote integrierte diesen Faktor vor dem Hintergrund, dass die besten Athletinnen und Athleten der Welt ihre Vorbereitung häufig ausschließlich auf die bedeutendsten Wettbewerbe ausrichten und nur dort die Höchstleistung zeigen. In einem rein Elo-basierten Modell würden diese Personen im Zeitraum zwischen den wichtigsten Sportereignissen unterschätzt werden. Die Gewichtung von Wettkampfergebnis und „Rolling Personal Best“ unterscheidet sich in diesen Sportarten in einem unbekanntem Ausmaß.

Da bei mittel- bis langfristig ausbleibender Wettbewerbsteilnahme die Unsicherheit über das aktuelle Leistungsniveau von Athletinnen und Athleten steigt, werden Elo-Punkte als „Inactivity-Penalty“ (IP) für jeden Kalendertag vom Elo-Wert abgezogen, an dem sie keinen Wettkampf bestritten haben. Formel 10 beschreibt die Systematik.

$$IP = \frac{(E - BE)}{4 \cdot 365,25} \cdot t \quad (10)$$

E steht auch hier für den letzten Elo-Wert einer Athletin bzw. eines Athleten, BE steht für den Basis-Elo-Wert, den alle Personen bei Neueinstieg in das System erhalten und liegt bei BE = 2.000. t stellt die Anzahl an Tagen dar, die die betrachtete Person bereits inaktiv ist. So folgt für einen Athleten A mit zuletzt 2.500 Elo-Punkten und 90 Tagen Inaktivität:

$$\begin{aligned} E'_A &= E_A - IP \\ &= 2.500 - \frac{(2.500 - 2000)}{4 \cdot 365,25} \cdot 90 \\ &\approx 2.500 - 30,80 \\ &\approx 2.468 \end{aligned} \quad (11)$$

Bei saisonalen Sportarten, insbesondere den Wintersportarten, ist aufgrund des IP ein Abfall der Elo-Werte in der Gesamtheit der Athletinnen und Athleten einer Sportart im Verlaufe des Sommers zu beobachten. Sämtliche Athletinnen und Athleten nähern sich dadurch zu Anfang der nächsten Wintersaison geringfügig einander an (vgl. [Abb. 3](#)).

Personen, die neu ins System integriert werden, starten i. d. R. bei einem Basis-Wert von 2.000 Elo-

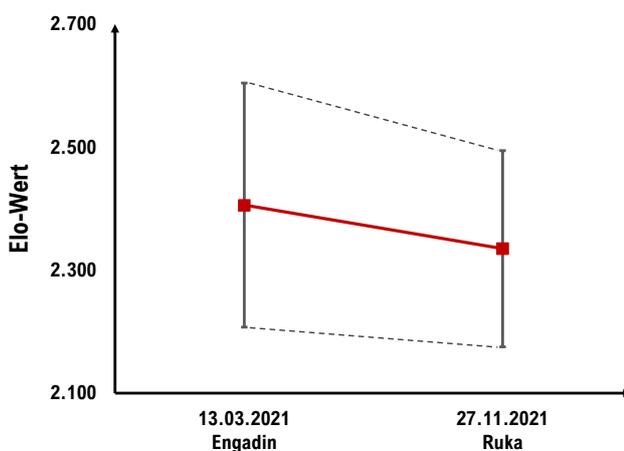


Abb. 3 Mittlere Elo-Werte (\pm SD) derjenigen Skilangläuferinnen (10 km klassisch), die am Ende der Saison 20/21 am letzten Weltcup in Engadin sowie am ersten Weltcup der Saison 21/22 in Ruka teilgenommen haben. Es ist aufgrund des Inactivity-Penalties ein leichtes Absinken der Werte und eine Annäherung der Athletinnen untereinander erkennbar, die sich in einer geringeren Standardabweichung widerspiegelt.

Punkten. Gracernote berücksichtigt jedoch auch, dass Athletinnen und Athleten in mehreren Disziplinen bzw. Gewichtsklassen starten können und ggf. schon Dateneinträge haben. Die Datenbank stellt daher Verbindungen zwischen Disziplinen her, die erkennbar miteinander korrelieren. Ist ein Eisschnellläufer beispielweise bislang ausschließlich in der 1.000-m-Disziplin gestartet und hat dort Elo-Punkte sammeln können, so nimmt er diese anteilig mit, wenn er erstmals in der 1.500-m-Disziplin startet. Auf diese Weise erreicht Gracernote bereits frühzeitig eine präzisere Annäherung an das tatsächliche Leistungsniveau, als wenn der Athlet beim Basis-Wert beginnen würde.

Mannschaftssportarten sowie Team-Events innerhalb der Individualsportarten werden ebenfalls von Gracernote berücksichtigt, jedoch mathematisch unterschiedlich behandelt. Für Mannschaftssportarten wie Basketball oder Feldhockey berechnet Gracernote Nationen-Werte, die von einer wechselnden Konstellation der Nationalspielerinnen und -spieler innerhalb der Teams unberührt bleiben; vergleichbar mit dem Prinzip der FIFA-Weltrangliste ¹⁵.

Die Team-Events des Individualsports werden seitens Gracernote hingegen auf Ebene der teilnehmenden Athletinnen und Athleten verarbeitet. Alle Mitglieder eines Teams verfügen über individuelle Elo-Werte, deren arithmetisches Mittel die Team-Wertung darstellt. Ändert sich demnach die Teamzusammensetzung, ändert sich auch die Team-Wertung und Gewinnwahrscheinlichkeit. Ausnahmen bilden bspw. der Ruder-Achter oder der Gruppenwettbewerb der Rhythmischen Sportgymnastik, wo sowohl Team- als auch Individualwertungen unabhängig voneinander geführt werden. Nach einem Wettkampf erhält jedes Teammitglied Elo-Punkte in Abhängigkeit von ihrer/seiner individuellen Ausgangswertung. So gewinnen stärkere Teammitglieder im Verhältnis zu schwächeren Teammitgliedern weniger Punkte für einen Sieg hinzu, verlieren aber mehr im Falle einer Niederlage.

PotAS^G

Unter PotAS^G (PotAS plus Gracenote) ist eine Prozessierung des Datenpakets „Podium Rankings“ zu verstehen, die über das firmenseitige Angebot an Analysetools hinausgeht. Hintergrund der Weiterverarbeitung ist die eingeschränkte Möglichkeit, anhand des ERS Potenzialeinschätzungen für die Zukunft vorzunehmen. Die Datenbank bietet zwar eine Reihe von Optionen, den aktuellen Leistungsstand von Athletinnen und Athleten in den internationalen Kontext einzuordnen, liefert jedoch keine Anhaltspunkte für mögliche Entwicklungsverläufe in den Folgejahren. Wie einleitend beschrieben, hat PotAS die Aufgabe, Potenziale für die folgenden Olympischen Spiele zu identifizieren und sie in vierjähriger Voraussicht systematisch zu beziffern. Die Herausforderung besteht entsprechend nicht nur darin, eine in die Zukunft gerichtete Datenanalyse vorzunehmen, sondern auch operationalisierte Kriterien anzulegen, die eine individualisierte Einschätzung des aktuellen Medaillen-/Finalplatzpotenzials zulassen. Daraus resultiert die nachfolgend beschriebene Verfahrensweise, die aus drei wesentlichen Komponenten besteht:

- Jährliche Entwicklungsraten innerhalb einer betrachteten Disziplin
- Aktueller Elo-Wert der betrachteten Person
- Niveau der Weltspitze

Disziplinspezifische Entwicklungsraten

Gracenotes Datenpaket enthält historische Informationen über Namen, Geschlecht, Alter, Sportart, Disziplin, Wettbewerb, Platzierung und Elo-Wert von Athletinnen und Athleten, zurückgehend bis in das Jahr 2004. Allein in den olympischen Sportarten umfasst es mehr als 125.000 Wettkämpfe von über 400.000 Personen. Dieser Umfang ermöglicht es, unterschiedlichste sportart- und disziplinspezifische Verteilungen abzuleiten.

Beim Sommersportverfahren 2024 wird das disziplinspezifische Entwicklungspotenzial von Athletinnen und Athleten ermittelt, indem aus der ebenso disziplinspezifischen Grundgesamt von

Athletinnen und Athleten (\cong alle Personen, die seit 2010 in derselben olympischen Disziplin gestartet sind) statistische Verteilungen der jährlichen Elo-Wertentwicklung berechnet werden (vgl. Abb. 4a). Diese Verteilungen lassen sich bspw. in Form eines Violinen-Diagramms darstellen (vgl. Abb. 4b). Das Diagramm beschreibt, welche jährlichen Zuwächse oder Minderungen der individuellen Elo-Werte bislang auftraten und wie häufig sie zu beobachten waren. In der abgebildeten Disziplin haben männliche Athleten seit 2010 jährlich zwischen -165 und +180 Punkte verloren oder hinzugewonnen. Solch große Veränderungen des Elo-Wertes sind jedoch eher selten. 50 % aller Athleten dieser Disziplin verändern sich um ungefähr -50 bis +50 Punkte pro Jahr. Würde eine Person bspw. einen Zuwachs von 100 Elo-Punkten erreichen, würde sie sich damit besser entwickeln, als es bislang in 90 % der Vergleichsfälle eingetreten ist.

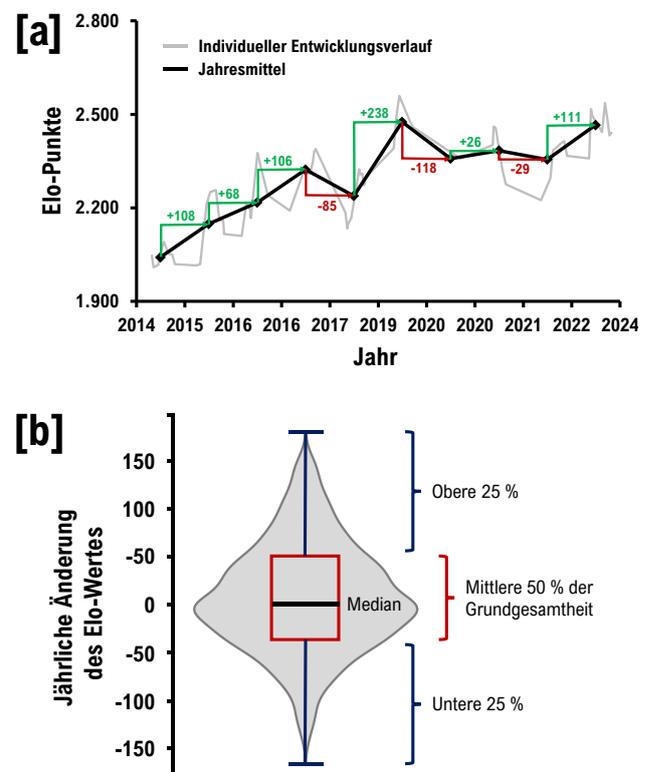


Abb. 4 [a] Berechnung der Entwicklungsraten je Athletin und Athlet anhand der mittleren Elo-Werte pro Kalenderjahr. [b] Violinen-Diagramm über die Verteilung der jährlichen Entwicklungsraten männlicher Athleten einer beispielhaften olympischen Disziplin (N = 2.514). 50 % aller Athleten erreichen jährlich Zuwächse/Minderungen des Elo-Wertes von ca. -50 bis +50 Punkten. Nach Ausreißer-Korrektur erzielen die besten Athleten einen Zuwachs von maximal 180 Punkten, die schlechtesten verlieren bis zu 165 Punkte.

Veränderung des PotAS-Modells seit der Wintersportanalyse 2022

Ausgang des PotAS-Modells, das 2022 für die Wintersportverbände zum Einsatz kam, war die Entwicklung der sportlichen Leistung in Abhängigkeit vom Alter. Grundannahme war dabei, dass jede Sportart eine eigene altersdemografische Struktur hat (Hochleistungsalter) und dass alle Athletinnen und Athleten ein altersabhängiges Entwicklungspotenzial aufweisen. Um dies abzubilden, wurden die Veränderungen des Elo-Werts sowie die daraus resultierenden Entwicklungspotenziale jeder Person mit Bezug zum kalendarischen Alter ermittelt.

In der Sommersportanalyse 2024 wird auf das Alter als Einflussfaktor verzichtet. Denn eine faire Einschätzung des Entwicklungspotenzials setzt voraus, dass in allen Altersklassen eine ausreichende Anzahl von Vergleichsdaten verfügbar ist. In den Randbereichen ist dies für besonders junge oder ältere Personen jedoch nicht garantiert. Zudem konnte in den wenigsten Sportarten statistisch dargelegt werden, dass das Alter einen signifikanten Einfluss auf die Entwicklungsraten des Elo-Ratings hat, auch wenn die maximale sportliche Leistungsfähigkeit in vielen Sportarten nachweisbar vom Alter abhängt²³⁻²⁵. Weiterhin wird die verfügbare Menge an Referenzdaten durch ein altersunabhängiges Verfahren erhöht und so eine belastbarere Modellbildung gewährleistet. Denn vielfach fehlen öffentlich einsehbare Altersangaben, vor allem von jungen Athletinnen und Athleten, die in ein altersabhängiges Modell nicht einbezogen werden können.

Um für die über vier Jahre vorausblickende Potenzialanalyse einen statistisch robusten Rahmen zu schaffen, werden die Daten vorab anhand des 1,5-fachen Interquartilabstandes (\cong 1,5-fache Länge der roten Box) von Ausreißern bereinigt. In Einzelfällen kommt es demnach vor, dass Zuwächse oder Minderungen des Elo-Ratings höher ausfallen, als es im Modell berücksichtigt wird.

Setzt man nun voraus, dass sich alle Athletinnen und Athleten innerhalb der Grenzen solcher statistischen Verteilungen entwickeln, lassen sich Bereiche definieren, in denen individuelle Elo-Werte zukünftig liegen können.

Aktueller Elo-Wert der Athletinnen & Athleten

Um bei einem Event wie den Olympischen Spielen eine realistische Chance auf eine Platzierung unter den Top-3 oder auch Top-8 zu haben, ist es notwendig, dass Athletinnen und Athleten zum Zeitpunkt der Potenzialeinschätzung entweder bereits ihre Leistungsfähigkeit unter Beweis gestellt haben oder aber eine überdurchschnittlich positive Leistungsentwicklung zeigen. Prognosen anhand von Regressionsmodellen sind insofern fehleranfällig, da sie von der Anzahl und dem Muster früherer Datenpunkte abhängen. Bereits ein bis zwei schlechte Saisons können bei linearen oder polynomischen Regressionen zu einem prognostischen Abwärtstrend führen, obwohl oszillierende Phasen in der Leistungsentwicklung von Athletinnen und Athleten keineswegs untypisch und, aufgrund einer an den Olympischen Spielen orientierten Periodisierung, oft sogar einzuplanen sind. Die PotAS-Kommission überprüft daher unabhängig von früheren Verlaufsmustern, welchen Elo-Wert bzw. welche Zuwächse erzielt werden müssen, um in den nächsten vier Jahren Spitzenniveau zu erreichen. Einziger Ausgangspunkt ist der aktuelle Elo-Wert (Saisonmittelwert) der betrachteten Person. Von dort aus wird simuliert, welche Elo-Werte die Person in den kommenden vier Jahren erreichen kann, wenn sie sich innerhalb des zuvor beschriebenen Verteilungsmusters entwickelt, das für die Simulation von Minimum bis Maximum in 99 weitere Perzentile unterteilt wird (vgl. Abb. 5).

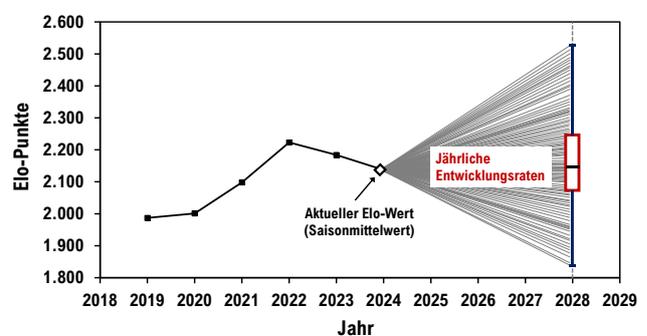


Abb. 5 Simulation der möglichen Elo-Wertentwicklung mithilfe der Entwicklungsraten, die aus der Grundgesamtheit von Athletinnen bzw. Athleten jener Disziplin abgeleitet wurden. Dazu werden die Entwicklungsraten vom Minimum bis Maximum in 99 Perzentile unterteilt und über einen Zeitraum von vier Jahren auf den aktuellen Elo-Wert addiert.

Für die simulierten Elo-Werte E_{Sim} gilt:

$$E_{Sim} = E + 4 \cdot r_{Perc} \quad (12)$$

Wobei E dem aktuellen Elo-Wert und r_{Perc} der Entwicklungsrate am jeweiligen Perzentil entspricht.

Vergleich mit der Weltspitze

Die dritte entscheidende Komponente ist das Leistungsniveau der Weltspitze. Auch dieses lässt sich anhand von Elo-Werten abbilden und dient als zu erreichender Maßstab für Erfolgsaussichten bei den nächsten Olympischen Spielen.

Die Podium-Datenbank verfügt über disziplinspezifische, Elo-basierte Weltranglisten, in die alle Athletinnen und Athleten eingegliedert sind. Zur Bestimmung der disziplinspezifischen Maßstäbe bzw. Grenzwerte wird ermittelt, wie hoch die Elo-Werte der weltweit besten drei bzw. acht Athletinnen oder Athleten in den letzten Jahren waren. Die Weltranglisten werden dafür zu jeweils drei Zeitpunkten exportiert:

1. Stand der Weltrangliste unmittelbar nach dem Zielwettkampf von 2022,
2. Stand der Weltrangliste unmittelbar nach dem Zielwettkampf von 2023,
3. Stand der Weltrangliste unmittelbar nach den Olympischen Spielen 2024.

Die Zielwettkämpfe sind vom DOSB vorgegeben und mit den Spitzenverbänden abgestimmt. Sie entsprechen in aller Regel den jeweiligen Saisonhöhepunkten der Sportarten (Welt- oder Europameisterschaften).

Diejenigen Athletinnen und Athleten, die nach den genannten Saisonhöhepunkten zu den besten drei bzw. acht der Welt gehörten, werden als Kohorten zusammengefasst. Aus jenen zwei Kohorten (Top-3-Kohorte und Top-8-Kohorte) wird jeweils ein Grenzwert berechnet, der dem 10. Perzentil aller enthaltenen Elo-Werte entspricht. Das 10. Perzentil ist als Minimalwert zu verstehen, der bei den nächsten Olympischen Spielen erreicht werden muss, um auf dem Leistungsniveau der Weltspitze zu agieren, abzüglich möglicher Ausreißer (vgl. [Abb. 6](#)).

AthletIn	Jahr	Platzierung	Elo-Wert
AthletIn A	2022	1	2.590
AthletIn B	2023	1	2.586
AthletIn C	2022	2	2.538
AthletIn A	2024	1	2.498
AthletIn D	2024	2	2.476
AthletIn C	2023	2	2.466
AthletIn E	2022	3	2.426
10. Perzentil	AthletIn E	2024	2.409
~2.401 Elo-P.	AthletIn D	2023	2.372

Abb. 6 Beispielhafte Berechnung eines Top-3-Grenzwertes, der jeweils aus den weltweit besten drei Athletinnen der letzten 3 Jahre einer Disziplin abgeleitet wird. Der Grenzwert entspricht dem 10. Perzentil dieser Kohorte. 90 % aller Top-3-Athletinnen hatten in diesem Beispiel also einen Elo-Wert von mindestens 2.401 Punkten. Diese Marke gilt als zu erreichender Minimalwert.

Die zwei Grenzwerte werden abschließend über die simulierten Entwicklungsverläufe der Athletinnen und Athleten gelegt. Je größer der Abstand zu den Grenzwerten ist und je größer dadurch die Entwicklungssprünge sein müssen, desto geringer fällt die Einschätzung des Potenzials aus. Reicht bereits eine Entwicklung bspw. entlang des 48. Perzentils, so liegt die betrachtete Person schon nahe oder über dem entsprechenden Grenzwert und hat großes Potenzial. Liegt sie noch weit vom Grenzwert entfernt und muss sich bspw. entlang des 80. Perzentils entwickeln – sie muss also vier Jahre in Folge einen Zuwachs erreichen, der durchschnittlich größer ist als bei 80 % der Grundgesamtheit – so fällt das Potenzial kleiner aus (vgl. [Abb. 7](#)).

Die Bezifferung des Potenzials (Pot) wird entsprechend der Formel 13 durchgeführt, wobei *Perc* dem Perzentil entspricht, das zum Erreichen des jeweiligen Grenzwertes notwendig ist.

$$Pot = 100 - Perc \quad (13)$$

Auch wenn theoretisch Erfüllungsquoten von 100 % erreicht werden können, kommt dies in dem beschriebenen Modell sehr selten vor. Denn eine solche Erfüllungsquote würde bedeuten, dass eine Athletin oder ein Athlet bereits heute so weit oberhalb der Grenzwerte liegt, dass selbst größte Einbußen im Elo-Wert über vier Jahre in Folge nicht dazu führen, dass die Person unter den avisierten Grenzwert fällt. Großes Potenzial haben Personen im PotAS⁶-Modell daher

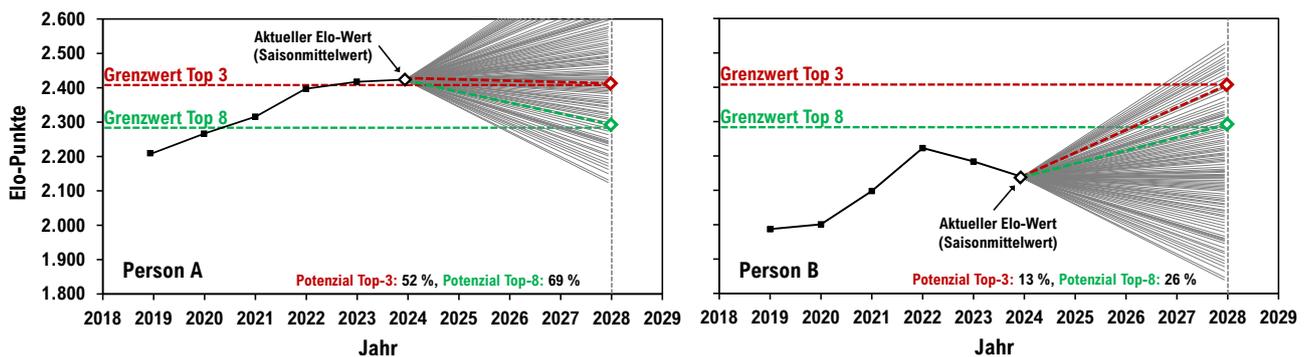


Abb. 7 Beispielhafte Potenzialanalysen für zwei Personen mit unterschiedlichem Elo-Ausgangswert. Person A ist bereits in der Weltspitze etabliert und muss das Leistungsniveau lediglich halten. Person B muss große Entwicklungssprünge machen. Um den Top-3-Grenzwert zu erreichen, sind vier Jahre lang Zuwächse notwendig, die größer sind als bei 87 % der Grundgesamtheit. Das Top-3-Potenzial wird entsprechend mit 13 % beziffert.

bereits mit Erfüllungsquoten von über 40 %. Selbst 25 % weisen zum Analysezeitpunkt schon auf eine internationale Konkurrenzfähigkeit hin. Die Erfüllungsquoten der Athletinnen und Athleten, die bei der Potenzialanalyse 2024 Berücksichtigung finden, werden daher schätzungsweise zwischen 0 und 70 % liegen.

Transfer ins Attributesystem

Im Rahmen der PotAS-Analysen sind die olympischen Disziplinen durch den DOSB in Disziplingruppen eingeteilt und nach Geschlecht getrennt worden. Jede Disziplingruppe lässt sich einem deutschen Spitzenverband zuordnen. „Wurf/Stoß Frauen“ und „Wurf/Stoß Männer“ beinhalten beispielsweise die Disziplinen Kugelstoßen, Diskuswerfen, Hammerwerfen und Speerwerfen, die wiederum alle in die Verantwortlichkeit des Deutschen Leichtathletik-Verbandes fallen. Ein anderes Beispiel ist der Deutsche Ringer-Bund, dem die drei Disziplingruppen „Freistil Frauen“, „Freistil Männer“ und „Griechisch-Römisch Männer“ zugeordnet sind. Die Disziplingruppen entsprechen also den Stilrichtungen des Ringens und deren Disziplinen entsprechen den verschiedenen Gewichtsklassen, die bei den Olympischen Spielen gelten. Die Bewertung der Fragen nach dem Kaderpotenzial findet auf Ebene dieser Disziplinen statt. Anschließend wird daraus eine mittlere Erfüllungsquote für die gesamte Disziplingruppe generiert.

Als Exempel sei eine Disziplingruppe „DG Frauen“ gegeben. Sie beinhaltet die Disziplinen X, Y und Z. Die Fragen 4.1.1 (Medaillenpotenzial) und 4.1.2 (Finalplatzpotenzial) werden jeweils auf alle der drei Disziplinen angewendet. Die Bewertungen werden als Erfüllungsquoten (0-100 %) dargestellt.

Die erste Frage würde lauten: Wie hoch ist das Medaillenpotenzial in Disziplin X der DG Frauen? Mithilfe der beschriebenen Methodik wird das Potenzial jeder Athletin eingeschätzt, die der zuständige Spitzenverband für Disziplin X gelistet hat. Die Erfüllungsquoten der zwei besten Athletinnen werden gemittelt und ergeben das Potenzial für Disziplin X. Die PotAS-Kommission entschied sich für die Bewertung von zwei Athletinnen bzw. Athleten je Disziplin, damit besonders solche Disziplingruppen gut abschneiden, bei denen der sportliche Erfolg nicht nur von einer Person abhängt und dadurch schnell gefährdet wäre, wenn diese Person sich verletzt oder ihr Karriereende vollzieht. Nachdem auch die Disziplinen Y und Z bewertet wurden, wird aus den drei Erfüllungsquoten ein Mittelwert für die Disziplingruppe gebildet. Erreicht die DG Frauen bei Disziplin X 39 %, bei Y 13 % und bei Z 33 %, so liegt die mittlere Erfüllungsquote für diese Disziplingruppe bei ca. 28 %. Dieser Wert entspricht der Bewertung für die Frage 4.1.1. (vgl. Abb. 8).

Die gleiche Systematik wird für die Frage 4.1.2 angewendet. Hier wird jedoch nicht das Medaillenpotenzial (Platz 1-3), sondern das Finalplatzpotenzial (Platz 1-8) geprüft. Die Erfüllungsquoten sind daher immer mindestens so hoch

Disziplingruppe „DG Frauen“						
Disziplin	Medaillenpotenzial (Top 3)			Finalplatzpotenzial (Top 8)		
	1. Athl.	2. Athl.	Quote	1. Athl.	2. Athl.	Quote
X	44 %	34 %	39 %	54 %	42 %	48 %
Y	26 %	0 %	13 %	32 %	4 %	18 %
Z	52 %	13 %	33 %	69 %	26 %	48 %
Ø Erfüllungsquote			28 %	38 %		

Hauptattribut 4 (gesamt)	
Ø Erfüllungsquote	33 %

Abb. 8 Bewertungssystematik für die Fragen 4.1.1 und 4.1.2 nach dem Medaillen- bzw. Finalplatzpotenzial einer Disziplingruppe. Hier als Exempel gezeigt ist die Disziplingruppe DG Frauen. In jeder Disziplin werden die jeweils zwei besten Athletinnen berücksichtigt. Der Mittelwert aus den Erfüllungsquoten beider Fragen ergibt die Gesamtquote für das Hauptattribut 4.

wie die in Frage 4.1.1 oder höher. Der Mittelwert aus beiden Fragen dient abschließend als Gesamtergebnis für das Hauptattribut 4 und somit für die Säule KADERPOTENZIAL.

Team-Events

In Anlehnung an die zuvor beschriebene Systematik werden auch die Team-Disziplinen des Individualsports wie z.B. Staffeln bewertet. Zunächst werden die möglichen individuellen Entwicklungsverläufe via Gracenote abgebildet. Da zum Zeitpunkt der Potenzialanalysen nicht bekannt ist, wer künftig in welcher Konstellation in einer Team-Disziplin antritt, werden diejenigen Personen herangezogen, die aktuell die größten Erfüllungsquoten aufweisen, sodass daraus das größtmögliche Potenzial resultiert. Deren Erfüllungsquoten werden dann im Sinne eines Team-Werts gemittelt.

Der Team-Wert wird äquivalent zu den Individualdisziplinen mit den besten Teams der letzten vier Jahre verglichen. Im Bewertungsverfahren werden bei Teamgrößen von mehr als zwei Personen allerdings nicht zwei Teams mit Potenzial abverlangt, um eine größtmögliche Erfüllungsquote zu erreichen. In diesen Fällen genügt ein einziges Team.

Mannschaftssportarten

Bekannte Herausforderungen, wie die stete Variation der Kaderzusammensetzung und die für

Spilsport typischen Zufallsfaktoren, stellen schwer zu lösende Anforderungen an ein zukunftsorientiertes Bewertungssystem für die Mannschaftssportarten. Im Rahmen der Überarbeitung von PotAS für das Sommersportverfahren 2024 ist die Integration aktueller Weltranglisten diskutiert worden. Die Ranglisten der internationalen Fachverbände sind konzeptionell jedoch so unterschiedlich, dass kein fairer Vergleich zwischen den deutschen Spitzenverbänden gewährleistet wäre. Gracenotes Elo-System ist hingegen innerhalb der Mannschaftssportarten kongruent, liefert historische Daten zu jeder aktiven Nationalmannschaft und bleibt somit die bevorzugte Datenquelle.

Durch den Wegfall des Altersfaktors können die Analysen für den Mannschaftssport nach dem gleichen Verfahren durchgeführt werden, wie es im Individualsport zum Einsatz kommt. Jede Nationalmannschaft wird hierbei, unabhängig von der Kaderzusammensetzung, als Individuum betrachtet. Von jeder Nationalmannschaft werden gleichermaßen jährliche Entwicklungsraten abgeleitet, die zu einem Verteilungsmuster beitragen. Auch die Grenzwerte ergeben sich aus den besten drei bzw. acht Mannschaften der letzten drei Jahre.

Diskussion

Durch die wachsende Datenverfügbarkeit und Fortschritte in der Analytik sind statistische Prognosen im Leistungssport heutzutage mehr denn je gefragt. Direkte Zubringergrößen sportlicher Leistung und Einflussfaktoren, die bis in das Athletinnen- und Athletenumfeld hineinreichen, sind oft gut identifiziert. Über kurzfristige Zeiträume hinausgehende, präzise Ergebnisvorhersagen lassen sich bislang aber dennoch kaum treffen, da die Anzahl unwägbarer Ereignisse und Entscheidungen auf lange Sicht zu hoch ist. In Abhängigkeit von der Sportart bilden selbst KI-gestützte Modelle nur kurze Zeiträume mit unterschiedlich hoher Präzision ab¹⁶. Prognosen für Zeiträume von bis zu vier Jahren bleiben eine Herausforderung.

Für das Verständnis des hier beschriebenen PotAS-Verfahrens ist es daher umso wichtiger, dass die Potenzialanalysen als eben solche aufgefasst und nicht mit statistischen Medaillenvorhersagen verwechselt oder gleichgesetzt werden. PotAS prognostiziert keine Karriereverläufe, sondern bewertet das Entwicklungspotenzial auf Grundlage von Elo-Ratings, die aus derzeitigen Wettkampfleistungen resultieren. Die disziplinspezifischen Querschnitte indizieren, welche Leistungssprünge möglich wären, und die Grenzwerte zeigen auf, welche Leistungssprünge notwendig erscheinen, um potenziell eine Top-3- bzw. Top-8-Platzierung erreichen zu können. Aus allen drei Komponenten ergibt sich ein empirisch begründetes Gesamtbild. Durch die Zuhilfenahme des ERS von Gracenote wird erreicht, dass die Entscheidungen sowohl eine empirische Grundlage haben als auch eine weitgehende Gleichbehandlung der Spitzenverbände mit internationalem Bezug gewährleisten.

Die PotAS-Kommission verkennt dabei nicht, dass Elo-Ratings im Allgemeinen und auch speziell das System von Gracenote vereinzelt Limitationen unterliegen. Die Podium-Datenbank wird bspw. durch die in ihr enthaltenen oder auch durch fehlende Wettkämpfe begrenzt. Dies gilt jedoch nicht grundsätzlich, sondern in Abhängigkeit davon, wie das Teilnehmerinnen- oder Teilnehmerfeld dort konstituiert

ist. Vor allem Wettkämpfe im Nachwuchsbereich beeinflussen die Rankings im oberen Bereich oft kaum oder gar nicht, tragen aber im unteren Bereich zu einer besseren Leistungsdifferenzierung bei. Da die Bedeutung von frühen Erfolgen im Nachwuchsbereich für spätere Erfolgsaussichten noch immer kontrovers diskutiert wird¹⁷⁻²², bleiben aktuelle Erfolge im Seniorinnen- und Seniorenalter die besten Prädiktoren einer Potenzialanalyse. Perspektivisch erscheint es sinnvoll, die Analysen der Kaderpotenziale über kürzere Zeiträume als vier Jahre durchzuführen. Nur so können stete Veränderungen der Leistungsentwicklung sowie junge Talente, die aus dem Nachwuchs in den Spitzensport nachrücken, angemessen, d. h. zu einem möglichst frühen Zeitpunkt berücksichtigt werden.

Eine interne, retrospektive Überprüfung des PotAS^{+G}-Verfahrens anhand der Olympischen Spiele von Tokio zeigt dennoch, dass die Potenziale vier Jahr zuvor – 2017 – differenziert abgebildet werden können. In die Überprüfung wurden alle Ergebnisse aus den Individualdisziplinen einbezogen (N = 7.119), darunter 947 Finalplatzierungen (Platz 4-8) und 634 Medaillenplatzierungen (Platz 1-3). 53 % der Athletinnen und Athleten, die 2021 eine Medaille gewannen, erhielten 2017 eine Top-3-Potenzialeinschätzung zwischen 25 und 100 %. Von denjenigen, die 2021 eine schlechtere Platzierung als

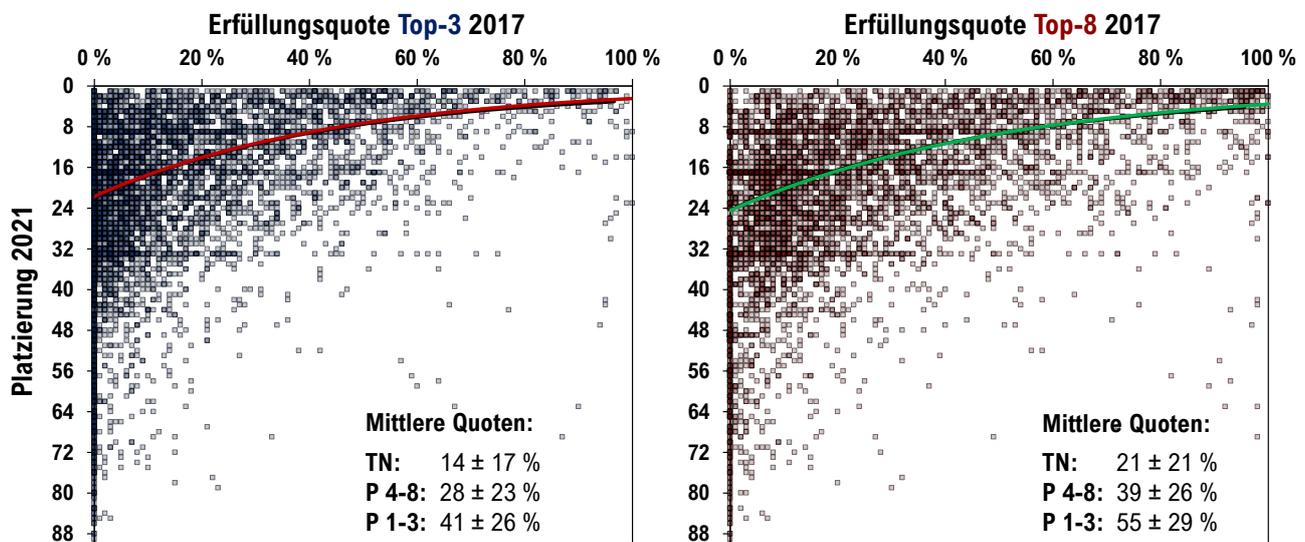


Abb. 9 Retrospektive Analyse des PotAS^{+G}-Verfahrens anhand der Platzierungen bei den Olympischen Spielen 2021 von Tokio. Die x-Achse beschreibt die Potenzialeinschätzung, die die Athletinnen und Athleten laut dem Modell 2017 bekommen hätten. Auf der y-Achse ist die Platzierung 2021 aufgetragen. Berücksichtigt wurden nur Sportlerinnen und Sportler der Individualdisziplinen. TN = Platzierung > Platz 8; P 4-8 = Platz 4-8; P 1-3 = Platz 1-3

Platz 8 erzielten, hatten 85 % eine Top-3-Potenzialeinschätzung von unter 25 %.

Abbildung 9 lässt erkennen, dass die Unsicherheit über die zukünftige Leistungsfähigkeit mit steigender Erfüllungsquote sinkt, denn mit steigender Erfüllungsquote nimmt die Streuung der Platzierungen immer weiter ab. Die Potenzialeinschätzungen stellen somit sichtbar einen Erfolgsindikator dar.

Über die retrospektive Analyse hinaus hat es sich die PotAS-Kommission auch weiterhin zur Aufgabe gemacht, die Anwendung der in diesem Papier beschriebenen Methodik für jede olympische Disziplin kritisch zu prüfen und im Einzelfall robuste Alternativen zu entwickeln. Der Dialog mit den Spitzenverbänden war und ist fester Bestandteil im laufenden und zukünftigen Bewertungsprozessen.

Literatur

- Gracenote Inc. Global Sports Data. <https://www.nielsen.com/solutions/content-metadata/global-sports-services/#use-cases> (2024).
- Zanco, D. G. de P., Szczecinski, L., Kuhn, E. V. & Seara, R. A comprehensive analysis of the Elo rating algorithm: Stochastic model, convergence characteristics, design guidelines, and experimental results. (2022) doi:10.48550/arxiv.2212.12015.
- Aldous, D. Elo Ratings and the Sports Model: A Neglected Topic in Applied Probability? *Statistical Science* **32**, 616–629 (2017).
- Abdi, S., Khosravi, H. & Sadiq, S. Modelling Learners in Adaptive Educational Systems: A Multivariate Glicko-Based Approach. in *LAK21: 11th International Learning Analytics and Knowledge Conference* 497–503 (Association for Computing Machinery, New York, NY, USA, 2021). doi:10.1145/3448139.3448189.
- Pelánek, R. Applications of the Elo rating system in adaptive educational systems. *Comput Educ* **98**, 169–179 (2016).
- Nikolakaki, S. M. et al. Competitive Balance in Team Sports Games. in *2020 IEEE Conference on Games (CoG)* 526–533 (IEEE, 2020). doi:10.1109/CoG47356.2020.9231859.
- Angelini, G., Candila, V. & De Angelis, L. Weighted Elo rating for tennis match predictions. *Eur J Oper Res* **297**, 120–132 (2022).
- Herbrich, R., Minka, T. & Graepel, T. TrueSkill™: A Bayesian skill rating system. *Adv Neural Inf Process Syst* **20**, 569–576 (2006).
- Minka, T., Cleven, R. & Zaykov, Y. *TrueSkill 2: An Improved Bayesian Skill Rating System*. <https://www.microsoft.com/en-us/research/publication/trueskill-2-improved-bayesian-skill-rating-system/> (2018).
- Glickman, M. E. *The Glicko System*. <https://www.glicko.net/glicko/glicko.pdf> (1995).
- Elo, A. E. The USCF Rating System - A Scientific Achievement. *Chess Life* **16**, 160–161 (1961).
- Elo, A. E. *The Rating of Chessplayers Past and Present*. (Arco Pub, New York, 1978).
- Aldous, D. Elo Ratings and the Sports Model: A Neglected Topic in Applied Probability? *Statistical Science* **32**, 616–629 (2017).
- Gracenote Inc. *E-Mail-Korrespondenz Mit Gracenote Im Januar 2023*. (2023).
- FIFA. Revision of the FIFA / Coca-Cola World Ranking. <https://digitalhub.fifa.com/m/f99da4f73212220/original/edbm045h0udbwkqew35a-pdf.pdf> (2021).
- Horvat, T. & Job, J. The use of machine learning in sport outcome prediction: A review. *WIREs Data Mining and Knowledge Discovery* **10**, (2020).
- Barth, M., Güllich, A., Macnamara, B. N. & Hambrick, D. Z. Predictors of Junior Versus Senior Elite Performance are Opposite: A Systematic Review and Meta-Analysis of Participation Patterns. *Sports Medicine* **52**, 1399–1416 (2022).
- Barth, M., Güllich, A., Macnamara, B. N. & Hambrick, D. Z. Quantifying the Extent to Which Junior Performance Predicts Senior Performance in Olympic Sports: A Systematic Review and Meta-analysis. *Sports Medicine* **54**, 95–104 (2024).
- Li, P., Bosscher, V. De, Pion, J., Weissensteiner, J. R. & Vertonghen, J. Is international junior success a reliable predictor for international senior success in elite combat sports? *Eur J Sport Sci* **18**, 550–559 (2018).
- Brouwers, J., Bosscher, V. De & Sotiriadou, P. An examination of the importance of performances in youth and junior competition as an indicator of later success in tennis. *Sport Management Review* **15**, 461–475 (2012).
- Mostaert, M., Vansteenkiste, P., Pion, J., Deconinck, F. J. A. & Lenoir, M. The importance of performance in youth competitions as an indicator of future success in cycling. *Eur J Sport Sci* **22**, 481–490 (2022).
- Boccia, G., Cardinale, M. & Brustio, P. R. Performance progression of elite jumpers: Early performances do not predict later success. *Scand J Med Sci Sports* **31**, 132–139 (2021).
- Allen, S. V. & Hopkins, W. G. Age of Peak Competitive Performance of Elite Athletes: A Systematic Review. *Sports Medicine* **45**, 1431–1441 (2015).
- Tanaka, H. & Toussaint, J.-F. Editorial: Growth, peaking, and aging of competitive athletes. *Front Physiol* **14**, (2023).
- Boccia, G. et al. Career Performance Trajectories in Track and Field Jumping Events from Youth to Senior Success: The Importance of Learning and Development. *PLoS One* **12**, e0170744 (2017).