

Schaakbeoordelingstoewijzingen voor meerdere platforms schatten met modale regressie

Elo+Chess Onderzoeksnotities

29 mei 2026

Abstract

Spelers vragen vaak hoe de beoordelingen op Lichess.org zich verhouden tot de beoordelingen op Chess.com. De vraag is praktisch belangrijk voor Elo+Chess vanwege de benchmark curven zijn opgebouwd uit een zeer grote gestratificeerde steekproef uit het geheel Lichess.org maandelijkse spelgeschiedenissen, terwijl gebruikers geschiedenissen van beide kunnen meenemen plaats. In deze toelichting wordt de huidige schattingsmethode beschreven speltypespecifieke toewijzingen van Lichess-beoordelingen tot Chess.com-beoordelingen. De methode maakt gebruik van kandidaten met dezelfde gebruikersnaam en exacte of bijna exacte tijdcontrole filters, minimale spelgeschiedenisvereisten en een modale regressieprocedure ontworpen om vooringenomenheid door valse overeenkomsten met dezelfde gebruikersnaam te verminderen. De huidige productie mapping is:

$$\hat{R}_{\text{Chess.com}} = \alpha + \beta R_{\text{Lichess}}.$$

De schattingen zijn voorlopig en zullen worden bijgewerkt naarmate de gematchte steekproef groeit. vooral voor snelle games die langer dan 10 minuten duren.

Huidige productie in kaart brengen

Voor lezers die eerst het resultaat willen, zijn de huidige productie vergelijkingen:

$$\begin{aligned}\hat{R}_{\text{Chess.com,bullet}} &= -530.60 + 0.9859R_{\text{Lichess}}, \\ \hat{R}_{\text{Chess.com,blitz}} &= -548.88 + 1.0837R_{\text{Lichess}}, \\ \hat{R}_{\text{Chess.com,10minrapid}} &= -499.68 + 1.0778R_{\text{Lichess}}, \\ \hat{R}_{\text{Chess.com,>10rapid}} &= -351.72 + 0.9003R_{\text{Lichess}}.\end{aligned}$$

Deze leidingen worden gemonteerd over Lichess-classificaties 650–1600 en mogen niet worden geëxtrapoleerd als universele conversies naar alle vaardigheidsniveaus.

1 Objectief

Het doel is om een beoordeling op Lichess.org toe te wijzen aan de vergelijkbare beoordelings-schaal Chess.com voor dezelfde brede tijdcontrolecategorie. Wij schatten een aparte mapping voor elke categorie gebruikt door Elo+Chess:

- 1+0 kogel,

- 3+0 blitz,
- 10+0 snel,
- snelle spellen langer dan 10 minuten.

De gepaste lijn wordt alleen gebruikt als schaalconversie laag. Dat impliceert niet de twee ratingpools hebben identieke ratingmechanismen, spelerspools, of spelselectiegedrag.

2 Gegevens herkomst

De Lichess-kant van de analyse is opgebouwd uit de volledige openbare Lichess game-archieven met standaardclassificatie van januari 2025 tot en met maart 2026. Dit is belangrijk omdat het Lichess-monster geen gemaksschraapsel of een smal monster is zichtbare profielen: het is opgebouwd uit de volledige maandelijkse gamegeschiedenisbestanden vrijgegeven door Lichess voor onderzoek en openbaar gebruik.

Lichess publiceert deze database-exports op <https://database.lichess.org/>. Op de databasepagina staat dat de exports zijn vrijgegeven onder een “CC0-licentie” en staat gebruikers expliciet toe om ze te gebruiken voor onderzoek en “publicatie”.¹ Deze open data houding staat centraal in het ontwerp van de benchmarkkant van de Elo+Chess.

Aan de Chess.com-kant worden alleen kandidaten met dezelfde gebruikersnaam gecontroleerd de maandelijkse archieven en spelcategorieën die nodig zijn om de mapping te schatten. Wij leg speciale nadruk op het matchen van spelgeschiedenissen binnen dezelfde kalender maand waar mogelijk. Het matchen van kalendermaanden verkleint het risico dat a De beoordeling van de speler veranderde aanzienlijk tussen de Lichess en Chess.com metingen.

3 Kandidaat-matchconstructie

De grootste uitdaging is dat geen van beide platforms een universeel platformafhankelijk platform blootlegt identificatie van de speler. We beginnen daarom met kandidaten met dezelfde gebruikersnaam: als a gebruikersnaam verschijnt op zowel Lichess.org als Chess.com, dat paar wordt behandeld als een kandidaat wedstrijd. Dit maakt opzettelijk lawaai. Sommige accounts met dezelfde naam zijn de dezelfde persoon, terwijl anderen niet-verwante spelers zijn die toevalig hetzelfde gebruiken handvat.

Om vermijdbare ruis te verminderen, wordt in de huidige pijplijn speltype-specifiek toegepast filters:

1. Lichess-gebruikers worden bemonsterd uit gestratificeerde gamegeschiedenis-datasets van ratingbucket en tijdcontroletype.
2. Chess.com kandidaat-accounts worden gecontroleerd op de bijbehorende categorie tijdcontrole.
3. Kandidaatparen worden alleen behouden als beide partijen voldoende relevantie hebben spelgeschiedenis.

¹<https://database.lichess.org/>, accessed May 29, 2026. The quoted phrases are from the ZXQPROT3ZXQ database page.

4. Voor kogels, blitz en 10 minuten snel is de minimaal bewaarde geschiedenis gelijk 5 spellen per kant.
5. Voor snelle ritten langer dan 10 minuten is de minimaal bewaarde geschiedenis 10 spellen per kant.
6. De hier gerapporteerde regressie-fit is beperkt tot Lichess-beoordelingen van 650 tot 1600, passend bij de beoordelingsregio die het meest relevant is voor de belangrijkste Elo+Chess benchmarkrapporten.

In het huidige voorbeeld wordt de nadruk gelegd op overeenkomsten met dezelfde kalendermaand en dezelfde gebruikersnaam. Wanneer een kandidaat-account bevat voldoende games in de doeltijdcontrolecategorie dezelfde maand als de Lichess-brongegevens, die overeenkomen met de maandelijksse beoordeling voorkeur voor montage. Oudere platformafhankelijke wedstrijdgegevens worden behouden waar ze zijn de steekproefondersteuning vergroten, maar het methodologische doel is specifiek voor het speltype vergelijking in dezelfde maand.

4 Waarom gemiddelde of mediaan vertekend kunnen zijn

Voor een vaste Lichess-beoordelingsbucket is de kandidaten-set met dezelfde gebruikersnaam een combinatie:

$$f(y | x) = \pi f_{\text{true}}(y | x) + (1 - \pi) f_{\text{false}}(y),$$

waarbij x de Lichess-beoordeling is, y de Chess.com-beoordeling is, f_{true} is the distribution of Chess.com ratings for true cross-platform overeenkomsten, en f_{false} is de gegenereerde distributie door niet-gerelateerde Chess.com-accounts met dezelfde gebruikersnaam.

De false-match-component is niet gecentreerd op het juiste Chess.com-equivalent voor de Lichess-bak. Het lijkt op de brede Chess.com-populatie bemonsterd op gebruikersnamen en activiteitenfilters. Het resultaat is dat het voorwaardelijke gemiddelde gelijk is aan getrokken in de richting van de wereldwijde distributie van actieve spelers:

$$E[y | x] = \pi E_{\text{true}}[y | x] + (1 - \pi) E_{\text{false}}[y].$$

Als $\pi < 1$, is dit over het algemeen niet gelijk aan de gewenste gematchte speler verwachting. Hetzelfde probleem is van invloed op de mediaan wanneer er sprake is van een valse overeenkomst component draagt voldoende massa bij om het 50e percentiel te verplaatsen.

De richting van de bias is afhankelijk van de emmer. In lage Lichess-bakken: false wedstrijden kunnen het gemiddelde omhoog trekken als er willekeurig actieve Chess.com-accounts zijn sterker dan de echte gematchte gebruikers. In hoge emmers dezelfde vervuiling kan het gemiddelde naar beneden trekken. Dit compressie-effect is precies wat we willen vermijden bij het bouwen van een schaalconversie.

5 Modale motivatie

De modale benadering is gebaseerd op een zwakkere aanname dan de gemiddelde benadering aanpak. We vereisen niet dat elk paar met dezelfde gebruikersnaam correct is. In plaats daarvan wij neem aan dat er binnen een voldoende beperkte Lichess-beoordelingsklasse echte overeenkomsten zijn vormen het meest geconcentreerde lokale cluster in de Chess.com-beoordelingsruimte.

Vals Er kunnen talrijke overeenkomsten zijn, maar ze zouden diffuser moeten worden verspreid ze zijn niet gebonden aan de Lichess-ratingbucket.

Voor elke Lichess-bak schatten we de modus van de Chess.com-beoordeling distributie met behulp van een schatting van de kerneldichtheid. Vervolgens wordt het modale punt behandeld als het meest plausibele Chess.com-equivalent voor die Lichess-bak. Dit maakt de procedure is minder gevoelig voor lange staarten en minder gevoelig voor niet-overeenkomende staarten accounts met dezelfde naam, verzameld uit de bredere Chess.com-populatie.

Dit betekent niet dat de modus perfect is. Het is een pragmatische robuustheid keuze voor een luidruchtig entiteitsresolutieprobleem. De methode werkt het beste als deze waar is overeenkomsten blijven een veelvoud in de buurt van de juiste platformafhankelijke beoordeling en onwaar wedstrijden zijn zelf niet geconcentreerd op een misleidende waarde.

6 Montageprocedure

De huidige productiefit gebruikt voor elk afzonderlijk de volgende procedure speltype:

1. Behoud kandidaat-paren met dezelfde gebruikersnaam die voldoen aan de gametelling drempels.
2. Beperk de pasvorm tot $650 \leq R_{ZXQPROT3ZXQ} \leq 1600$.
3. Verdeel de Lichess-beoordelingen in buckets van 100 punten.
4. Maak voor elke bucket met minimaal 30 gebruikers een schatting van de modale Chess.com beoordeling met een schatting van de Gauss-kerneldichtheid met behulp van bandbreedte 75-beoordeling punten.
5. Geef elke bak weer aan de hand van het middelpunt en de geschatte Chess.com-modus.
6. Plaats een gewone lijn met de kleinste kwadraten door het modale niveau op emmerniveau punten:

$$\hat{R}_{\text{Chess.com}} = \alpha + \beta R_{\text{Lichess}}.$$

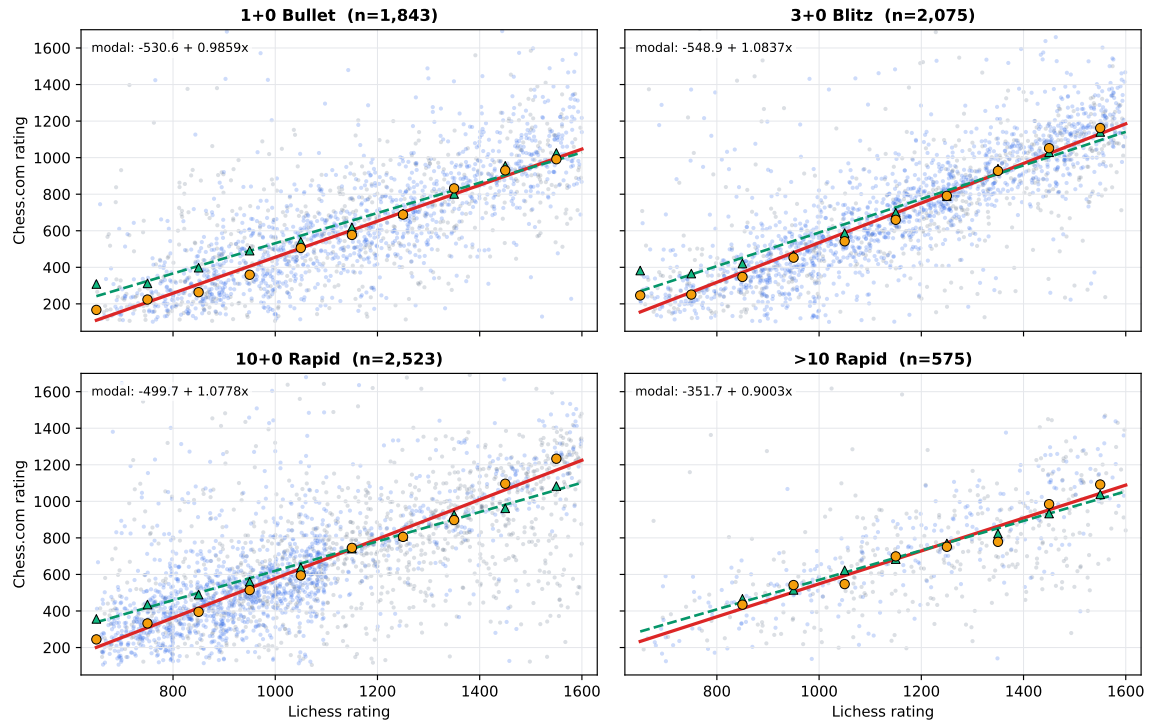
Alleen modale punten op bucketniveau komen in de uiteindelijke lineaire regressie terecht. Individueel paren met dezelfde gebruikersnaam beïnvloeden de curve door hun bijdrage aan de schatting van de dichtheid binnen de emmer.

7 Huidige schattingen

Tabel 1 rapporteert de huidige aangepaste constanten en hellingen. Dit zijn de waarden die momenteel in de Elo+Chess zijn opgenomen na 29 mei 2026 bijwerken.

De snelle schatting van langer dan 10 minuten heeft aanzienlijk minder steun dan de andere drie categorieën. Het is daarom het meest waarschijnlijk dat het als extra wordt gewijzigd gematchte gebruikers worden verzameld.

Cross-platform same-username rating matches and modal mapping fits



Raw points are same-username candidate matches after game-history filters. Red line is the production modal fit; green dashed line shows the bucket-mean alternative for comparison.

● Matched users ● Bucket mode ▲ Bucket mean — Modal fit - - - Mean-bucket fit

Figure 1: Kandidaatmatches met dezelfde gebruikersnaam en passende mappingcurven. Ruwe punten zijn behouden kandidaat-gebruikers na filters in de gamegeschiedenis. Oranje punten zijn de modale Chess.com-beoordelingen binnen de bucket gebruikt voor de productiepassing. De rode lijn is de modale regressielijn. Groene driehoekjes en de gestreepte groene lijn zijn zichtbaar het emmer-gemiddelde alternatief ter vergelijking; deze zijn ter illustratie opgenomen hoe op gemiddelden gebaseerde samenvattingen kunnen worden gemaakt door niet-overeenkomende accounts met dezelfde naam, en staart observaties.

Game type	Total users	Baseline	Same-month	In range	α	β
1+0 bullet	2,999	1,474	1,525	1,843	-530.60	0.9859
3+0 blitz	3,101	1,306	1,795	2,075	-548.88	1.0837
10+0 rapid	3,084	1,509	1,575	2,523	-499.68	1.0778
> 10 rapid	750	445	305	575	-351.72	0.9003

Table 1: De huidige modale lineaire past over de Lichess-beoordelingen 650–1600. “ Binnen bereik” telt het aantal gebruikers dat binnen het aangepaste Lichess-interval blijft. Basislijnrijen afkomstig zijn van de bestaande platformafhankelijke tafel met exacte controle; rijen van dezelfde maand komen uit de maandelijkse steekproef met dezelfde gebruikersnaam van maart 2026.

8 Interpretatie

De aangepaste vergelijkingen kunnen het beste worden geïnterpreteerd als lokale conversies voor beginners vroeg-gevorderde gebruikers, niet als universele beoordelingsidentiteiten. Bijvoorbeeld de Een snelle schatting van 10 minuten is dat wel

$$\hat{R}_{\text{Chess.com}} = -499.68 + 1.0778R_{\text{Lichess}}.$$

Bij een Lichess-beoordeling van 1200 levert dit een Chess.com-equivalent op van ongeveer 793. Bij een Lichess-rating van 1500 geeft dit ongeveer 1117.

De schattingen mogen niet ver buiten het gepaste interval worden geëxtrapoleerd. Bij zeer hoge beoordelingen, samenstelling van de spelerspool, accountgedrag en platform De beoordelingsmechanismen kunnen verschillen op manieren die niet door één enkele lineaire methode worden vastgelegd mapping aangepast aan het lagere en middelste beoordelingsbereik.

9 Beperkingen en updateplan

De huidige procedure is opzettelijk conservatief, maar heeft verschillende beperkingen blijven:

- Het matchen van dezelfde gebruikersnaam is luidruchtig en bewijst geen identiteit.
- De true-match-fractie kan variëren per ratingbucket en speltype.
- Chess.com- en Lichess-beoordelingen worden niet altijd op exact dezelfde manier gemeten dag voor alle kandidaatparen, hoewel de steekproef van dezelfde maand dit vermindert kwestie.
- Het snelle monster van langer dan 10 minuten is momenteel kleiner dan het kogel-, blitz- en snelle monsters van 10 minuten.
- Het lineaire model is een praktische benadering over 650–1600, niet a structureel model van beide beoordelingssystemen.

De volgende geplande verbetering is het blijven bemonsteren van langere snelle gebruikers, onderhevig aan conservatieve API-gebruiks- en platformvoorwaarden. Zodra het overeenkomende monster groter en evenwichtiger is over de beoordelingssegmenten, zouden de modale aanpassingen moeten zijn opnieuw geschat en de productieconstanten opnieuw bijgewerkt.

10 Onderhoud en updates

Dit document weerspiegelt de meest recente cross-platform beoordelingstoewijzing van de Elo+Chess update vanaf 29 mei 2026. De schattingen zullen als aanvullend worden herzien overeenkomende gebruikers met dezelfde kalendermaand worden verzameld, vooral voor snelle spellen van langer dan 10 minuten.

De onderhouden openbare versie van dit rapport is beschikbaar op:

<https://www.elopluschess.com/static/docs/cross-platform-elo-mapping-modal-method.pdf>