

Stima delle mappature di valutazione degli scacchi multiplatforma con regressione modale

Elo+Chess Note di ricerca

29 maggio 2026

Abstract

I giocatori spesso chiedono come si confrontano le valutazioni su Lichess.org con le valutazioni su Chess.com. La domanda è praticamente importante per Elo+Chess perché è il suo punto di riferimento le curve sono costruite da un campione stratificato molto ampio prelevato dal completo Lichess.org cronologie di gioco mensili, mentre gli utenti possono portare cronologie da entrambe sito. Questa nota descrive l'attuale metodo utilizzato per la stima mappature specifiche del tipo di gioco dalle classificazioni Lichess alle classificazioni Chess.com. Il metodo utilizza corrispondenze candidate con lo stesso nome utente, controllo temporale esatto o quasi esatto filtri, requisiti minimi di cronologia del gioco e una procedura di regressione modale progettato per ridurre i pregiudizi derivanti da false corrispondenze con lo stesso nome utente. La corrente la mappatura della produzione è:

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

Le stime sono preliminari e verranno aggiornate man mano che il campione abbinato cresce, soprattutto per partite rapide di durata superiore a 10 minuti.

Mappatura della produzione attuale

Per i lettori che vogliono prima il risultato, le attuali equazioni di produzione sono:

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

Queste linee sono montate su Lichess con rating 650–1600 e non devono essere estrapolate come conversioni universali a tutti i livelli di abilità.

1 Obiettivo

L'obiettivo è mappare una valutazione su Lichess.org sulla scala di valutazione comparabile su Chess.com per la stessa ampia categoria di controllo del tempo. Stimiamo un separato mappatura per ciascuna categoria utilizzata da Elo+Chess:

- 1+0 punto elenco,

- 3+0 lampo,
- 10+0 rapido,
- partite rapide di durata superiore a 10 minuti.

La linea adattata viene utilizzata solo come livello di conversione della scala. Ciò non implica questo i due pool di valutazione hanno meccanismi di valutazione, pool di giocatori o comportamento di selezione del gioco.

2 Provenienza dei dati

Il lato Lichess dell'analisi è costruito a partire dal pubblico completo Lichess archivi di giochi con classificazione standard da gennaio 2025 a marzo 2026. Questo è importante perché il campione Lichess non è un ritaglio di convenienza o un campione ristretto di profili visibili: è costruito dai file mensili completi della cronologia di gioco rilasciato da Lichess per la ricerca e l'uso pubblico.

Lichess pubblica queste esportazioni di database su <https://database.lichess.org/>. La pagina del database afferma che le esportazioni sono rilasciate con una "licenza CC0" e consente esplicitamente agli utenti di "usarli per la ricerca" e "pubblicazione".¹ Questo file open-data la postura è fondamentale per la progettazione del lato benchmark di Elo+Chess.

Sul lato Chess.com, vengono controllati solo gli account candidati con lo stesso nome utente gli archivi mensili e le categorie di giochi necessari per stimare la mappatura. Noi porre particolare enfasi sull'abbinamento delle cronologie di gioco all'interno dello stesso calendario mese quando possibile. La corrispondenza del mese di calendario riduce il rischio che a la valutazione del giocatore è cambiata sostanzialmente tra Lichess e Chess.com misurazioni.

3 Costruzione della partita dei candidati

La sfida principale è che nessuna delle due piattaforme presenta una multipiattaforma universale identificatore del giocatore. Iniziamo quindi con i candidati con lo stesso nome utente: se a il nome utente appare sia su Lichess.org che su Chess.com, quella coppia viene trattata come corrispondenza del candidato. Questo è intenzionalmente rumoroso. Alcuni account con lo stesso nome sono i stessa persona, mentre altri sono giocatori non imparentati che usano lo stesso maniglia.

Per ridurre il rumore evitabile, l'attuale pipeline si applica in modo specifico al tipo di gioco filtri:

1. Gli utenti di Lichess vengono campionati da set di dati di cronologia di gioco stratificati per fascia di valutazione e tipo di controllo temporale.
2. Gli account candidati Chess.com vengono controllati per i corrispondenti categoria di controllo del tempo.
3. Le coppie di candidati vengono conservate solo quando entrambe le parti hanno sufficiente rilevanza storia del gioco.
4. Per bullet, blitz e 10 minuti rapidi, la cronologia minima conservata è 5 partite per parte.

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

5. Per rapide di durata superiore a 10 minuti, la cronologia minima conservata è 10 partite per parte.
6. L'adattamento della regressione riportato qui è limitato alle valutazioni Lichess da Da 650 a 1600, corrispondente alla regione di classificazione più rilevante per il Elo+Chess principale rapporti di riferimento.

L'esempio corrente enfatizza le corrispondenze con lo stesso nome utente dello stesso mese di calendario. Quando a l'account candidato dispone di giochi sufficienti nella categoria di controllo del tempo target in lo stesso mese dei dati di origine Lichess, la valutazione mensile corrispondente è preferito per il montaggio. I dati delle partite multiplatforma precedenti vengono conservati dove si trovano aumentare il supporto del campione, ma l'obiettivo metodologico è specifico del tipo di gioco confronto nello stesso mese.

4 Perché la media o la mediana possono essere distorte

Per un intervallo di classificazione Lichess fisso, il set di candidati con lo stesso nome utente è una miscela:

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

dove x è la classificazione Lichess, y è la classificazione Chess.com, f_{true} is the distribution of Chess.com ratings for true corrispondenze multiplatforma e f_{false} è la distribuzione generata da account Chess.com non correlati con lo stesso nome utente.

Il componente di falsa corrispondenza non è centrato sull'equivalente Chess.com corretto per la benna Lichess. Tende ad assomigliare all'ampia popolazione Chess.com campionati da nomi utente e filtri di attività. Di conseguenza, la media condizionale è spinto verso la distribuzione globale dei giocatori attivi:

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

Se $\pi < 1$, generalmente non è uguale al giocatore abbinato desiderato aspettativa. Lo stesso problema riguarda la mediana ogni volta che si verifica una falsa corrispondenza componente contribuisce con una massa sufficiente per spostare il 50° percentile.

La direzione della polarizzazione dipende dal bucket. Nei bucket Lichess bassi, falso le corrispondenze possono aumentare la media se lo sono account Chess.com attivi casuali più forte dei veri utenti abbinati. Nei secchi alti, la stessa contaminazione può spingere la media verso il basso. Questo effetto di compressione è esattamente ciò che vogliamo evitare quando si crea una conversione in scala.

5 Motivazione modale

L'approccio modale si basa su un presupposto più debole rispetto a quello basato sulla media approccio. Non è necessario che tutte le coppie di nomi utente identici siano corrette. Invece noi presupporre che all'interno di un intervallo di valutazione Lichess sufficientemente ristretto, corrisponda il vero formano il cluster locale più concentrato nello spazio di rating Chess.com. Falso le corrispondenze possono essere numerose, ma dovrebbero essere distribuite più diffusamente perché non sono legati alla fascia di rating Lichess.

Per ciascun bucket Lichess, stimiamo la modalità del rating Chess.com distribuzione utilizzando una stima della densità del kernel. Viene quindi trattato il punto modale come l'equivalente Chess.com più plausibile per quel bucket Lichess. Questo fa la procedura è meno sensibile alle code lunghe e meno sensibile ai disallineamenti account con lo stesso nome campionati dalla popolazione Chess.com più ampia.

Questa non è un'affermazione che la modalità sia perfetta. È una robustezza pragmatica scelta per un problema rumoroso di risoluzione di entità. Il metodo funziona meglio quando è vero le partite rimangono una pluralità vicino alla valutazione multiplatforma corretta e false le partite stesse non sono concentrate a un valore fuorviante.

6 Procedura di adattamento

L'adattamento di produzione corrente utilizza la seguente procedura separatamente per ciascuno tipo di gioco:

1. Conserva le coppie di candidati con lo stesso nome utente che soddisfano il conteggio dei giochi soglie.
2. Limita l'adattamento a $650 \leq R_{ZXQPROT3ZXQ} \leq 1600$.
3. Suddivide le classificazioni Lichess in bucket da 100 punti.
4. Per ogni bucket con almeno 30 utenti, stimare la modale Chess.com valutazione con una stima della densità del kernel gaussiano utilizzando la valutazione della larghezza di banda 75 punti.
5. Rappresenta ciascun bucket in base al punto medio e alla modalità Chess.com stimata.
6. Adatta una linea ordinaria dei minimi quadrati attraverso il modale a livello di secchio punti:

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

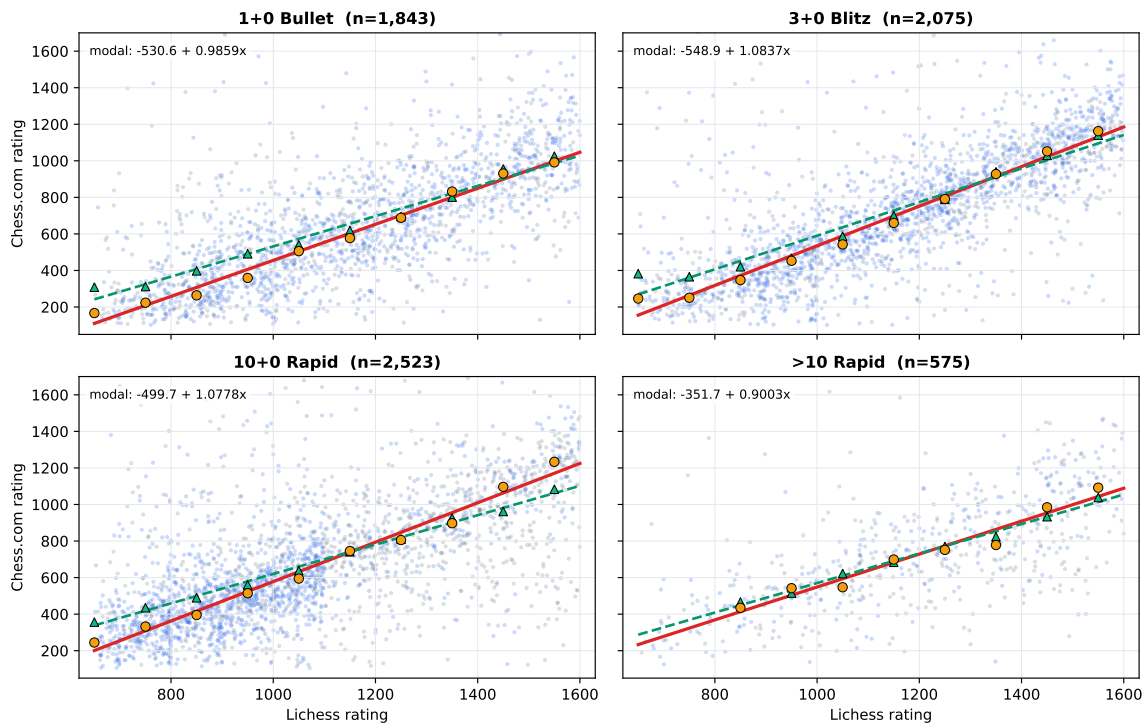
Solo i punti modali a livello di bucket entrano nella regressione lineare finale. Individuale le coppie con lo stesso nome utente influenzano la curva attraverso il loro contributo al stima della densità all'interno del bucket.

7 Stime attuali

La tabella 1 riporta le costanti e le pendenze attuali. Questi sono i valori attualmente collegati a Elo+Chess dopo il 29 maggio 2026 aggiornamento.

La stima rapida di durata superiore a 10 minuti ha materialmente meno supporto della stima rapida altre tre categorie. È quindi più probabile che cambi come aggiuntivo vengono raccolti gli utenti abbinati.

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: Corrispondenze del candidato con lo stesso nome utente e curve di mappatura adattate. Punti grezzi vengono mantenuti gli utenti candidati dopo i filtri della cronologia del gioco. I punti arancioni sono i Valutazioni modali Chess.com all'interno del bucket utilizzate per l'adattamento alla produzione. La linea rossa è la retta di regressione modale. Vengono visualizzati i triangoli verdi e la linea verde tratteggiata l'alternativa della media temporale per il confronto; questi sono inclusi per illustrare come i riepiloghi basati sulla media possono essere estratti da account con lo stesso nome non corrispondenti e osservazioni della coda.

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: L'attuale lineare modale si adatta alle classificazioni Lichess 650–1600. "In." range" conta gli utenti trattenuti all'interno dell'intervallo Lichess montato. Righe di base provengono dalla tabella multipiattaforma esistente a controllo esatto; arrivano le righe dello stesso mese dal campione mensile con lo stesso nome utente di marzo 2026.

8 Interpretazione

Le equazioni adattate vengono interpretate al meglio come conversioni locali per i principianti utenti precocemente avanzati, non come identità di rating universali. Ad esempio, il La stima rapida di 10 minuti è

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

Con un punteggio Lichess di 1200, si ottiene un equivalente Chess.com di circa 793. Con un punteggio Lichess di 1500, dà circa 1117.

Le stime non dovrebbero essere estrapolate molto al di fuori dell'intervallo adattato. A valutazioni molto elevate, composizione del pool di giocatori, comportamento dell'account e piattaforma i meccanismi di valutazione possono differire in modi che non vengono catturati da un singolo lineare mappatura adattata alla fascia di rating inferiore e media.

9 Limitazioni e piano di aggiornamento

L'attuale procedura è volutamente conservativa, ma presenta numerose limitazioni rimangono:

- La corrispondenza dello stesso nome utente è rumorosa e non dimostra l'identità.
- La frazione di corrispondenza reale può variare in base al gruppo di valutazione e al tipo di gioco.
- Le classificazioni Chess.com e Lichess non sono sempre misurate esattamente sullo stesso giorno per tutte le coppie di candidati, sebbene il campione dello stesso mese lo riduca problema.
- Il campione rapido di durata superiore a 10 minuti è attualmente più piccolo del bullet, blitz e campioni rapidi di 10 minuti.
- Il modello lineare è un'approssimazione pratica su 650–1600, non a modello strutturale di entrambi i sistemi di rating.

Il prossimo miglioramento pianificato è continuare a campionare utenti più rapidi e più a lungo, soggetto all'utilizzo conservativo di API e ai termini della piattaforma. Una volta che il campione abbinato è più ampio e più equilibrato tra gli intervalli di rating, dovrebbero esserlo anche gli adattamenti modali rivalutato e le costanti di produzione aggiornate nuovamente.

10 Manutenzione e aggiornamenti

Questo documento riflette la più recente mappatura della classificazione multiplatforma Elo+Chess aggiornamento al 29 maggio 2026. Le stime saranno riviste come aggiuntive vengono raccolti gli utenti abbinati nello stesso mese di calendario, in particolare per partite rapide di durata superiore a 10 minuti.

La versione pubblica mantenuta di questo rapporto è disponibile all'indirizzo:

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