

Vytvoření stratifikovaných srovnávacích křivek Lichess pro Elo+Chess

Elo+Chess Poznámky k výzkumu

31. května 2026

Abstract

Srovnávací zprávy Elo+Chess jsou sestaveny z velkých stratifikovaných vzorků Hry Lichess.org, přičemž srovnávací vzorek použitý v této poznámce obsahuje 483 974 kulových her 1+0, 479 974 bleskových her 3+0, 925 636 rychlých her 10+0, a 660 540 delších rychlých her. Pro každý hlavní typ hry se kanál rozšiřuje každá hra do stavů desky před tahem a po tahu počítá stovky střední a hráčské proměnné, agreguje tyto proměnné do koučovací metriky a poté zprůměruje tyto metriky podle bucketu Lichess Elo. The vzorek je záměrně rozvrstvený spíše než úměrný surovému hráči pool: účelem je odhadnout stabilní křivky chování na každé úrovni hodnocení, včetně kbelíků, které by jinak byly řídké, nereprodukovat přirozené rozdělení hodnocení. Výsledkem je rodina empirických referenčních křivek: for u každé metriky se můžeme zeptat, jak často hráči na různých úrovních Elo ukazují daný zvyk nebo princip ve skutečných hrách. Tato poznámka popisuje metodu, rozsah výpočet a několik příklady úvodního principu zahrnující rošádu, vývoj menších kusů, opakované tahy úvodní figurky a raný pohyb královny. Důraz je metodologické: Elo+Chess nepředpokládá, že šachové principy jsou univerzální lineární napříč všemi úrovněmi dovedností. Místo toho jsou referenční křivky kontrolovány pomocí rozsah hodnocení a produkt se zaměřuje na region pro začátečníky až brzy pokročilé kde jsou vztahy nejsilnější a nejlépe interpretovatelné.

1 Účel

Cílem systému benchmarků je převést široké šachové rady do měřitelné veličiny. Prohlášení jako “rozvíjejte své drobné kousky” je užitečné, ale stává se akčnějším, když můžeme říci, kolik menších kusů a hráč se obvykle vyvíjí tahem 10, jak je to ve srovnání s vrstevníky ve stejném hodnotící segment a jak se tento zvyk mění v celém ratingovém spektru.

Elo+Chess proto vytváří metrickou knihovnu ze skutečných her. Každá metrika je vypočítaná pro každý řádek hry hráče, shrnuté podle typu hry a Elo bucketu a použita jako srovnávací křivka v přehledu. Zpráva pak může umístit uživatele na stejnou křivku a ukazují, zda je uživatel nad, pod nebo blízko chování pozorované mezi srovnatelnými hráči.

2 Zdrojová data a stratifikace

Zdrojem benchmarků je veřejná databáze her Lichess.org. Používá Elo+Chess spíše velké stratifikované vzorky podle typu hry a hodnocení než malé vzorek pohodlí. Současné artefakty za běhu výroby pokrývají čtyři hlavní kategorie her:

- kulka 1+0,

- blesk 3+0,
- 10+0 rychle,
- rychlé hry delší než 10 minut.

Cílem stratifikace je šířka napříč úrovněmi hodnocení. Proporcionální náhoda vzorek by byl účinný pro popis celkového počtu hráčů, ale byl by být neefektivní pro vytváření referenčních křivek. Nejběžnější hodnotící kbelíky by ve vzorku dominovaly, zatímcočasné lopaty s nízkou podporou by jich měly příliš málo hry pro stabilní metrické průměry. Elo+Chess tedy vzorky podle průměrné hry Elo vědro. V současné konstrukci čistého vzorku jsou surové hry na prvním místě filtrovány na platné hodnocené hry se známými hráči, známá hodnocení, platné výsledky, a známé časové kontroly. Hry jsou poté filtrovány na hráče s dostatečným počtem her v této rychlosti a dostatečně stabilních hodnoceních, aby představovaly koherentní úroveň dovedností: pro každého hráče, Elo+Chess vypočítává rychlostní rozpětí hráče ($\max(\text{Elo}) - \min(\text{Elo})$) přes zdrojový fond a uchovává hry pouze v případě, že oba hráči mají nakonfigurovaný alespoň minimální počet her a ne více než 300 bodů Elo driftu v rámci rychlosti. Nemyslí se driftový filtr k odstranění normálního zlepšení nebo odchylky; snižuje případy, kdy hráč hry by míchaly materiálně různé úrovně dovedností, jako jsou prozatímní účty, vracející se hráči nebo účty, které se během několika kýblů rychle pohybovaly vzorkované období. Nakonec jsou způsobilé hry seřazeny v rámci každého průměru – Elo bucket podle deterministického hashe id a semena hry a omezeného počtu hry jsou zachovány z každého kbelíku.

Díky tomuto designu jsou referenční křivky mnohem užitečnější. Každý kbelík Elo přispívá dostatečným množstvím pozorování k odhadu průměrné hodnoty každé metriky, a vzácné kbelíky s vysokým nebo nízkým hodnocením nejsou přehlušeny přirozeným tvarem distribuce aktivních hráčů. Srovnávací křivky používané Elo+Chess jsou poté se omezí na rozsah nejvíce relevantní pro daný produkt, zhruba pro začátečníky prostřednictvím rané pokročilé hry. Ve zprávách na stupnici Lichess je tento primární rozsah je 600–1600. Když uživatel zobrazí sestavu na stupnici Chess.com, odpovídající segmenty hodnocení jsou mapovány na stejný podkladový Lichess benchmarkové segmenty pomocí samostatné metody mapování hodnocení napříč platformami popsané v doprovodném dokumentu: [Estimating Cross-Platform Chess Rating Mappings with Modální regrese](#).

3 Distribuce hodnocení aktivních hráčů

Pokud se podíváme na rozložení hodnocení Elo přehrávače Lichess.org ke konci března 2026, podle hlavních typů her vidíme následující počty a percentily. Pro toto zobrazení distribuce je každý hráč započítán jednou za přesný typ hry, pomocí nejnovějšího pozorovaného hodnocení daného hráče v odpovídajícím raw řádky s kontrolou času poté, co v této přesné hře vyžadovaly alespoň pět her typ.¹

Tato distribuce by neměla být zaměňována se stratifikovaným srovnávacím vzorkem. Distribuce popisuje aktivní hráčský fond Lichess pozorovaný v rawu skenovat. Srovnávací vzorek je pak záměrně stratifikován tak, aby Elo bucket mít dostatečnou podporu pro stabilní metrické křivky.

Chcete-li namapovat z hodnocení Lichess.org Elo pro konkrétní typ hry na ekvivalentní hodnocení Chess.com, Elo+Chess používá modal specifický pro typ hry regresní rovnice odhadnuté

¹We apply a five-game minimum because lightly active accounts disproportionately remain at the 1500 starting/provisional anchor, creating an artificial spike in all-player latest-rating distributions. Figures 1–4 zobrazit distribuce s tímto filtrem a bez něj.

Game type	Players	Median	75th pct.	80th pct.	% ≤ 1600	Exact 1500	1500 share
1+0 bullet	1,221,146	1,479	1,858	1,951	58.9%	4,756	0.4%
3+0 blitz	1,144,045	1,542	1,845	1,912	54.9%	2,315	0.2%
10+0 rapid	1,679,401	1,381	1,691	1,765	68.3%	3,472	0.2%
>10 rapid	668,255	1,400	1,686	1,753	68.2%	7,969	1.2%

Table 1: Poslední pozorovaná distribuce hodnocení aktivního hráče Lichess podle přesné typ hry poté, co vyžaduje alespoň pět her v tomto přesném typu hry.

v doprovodu [Elo+Chess rating-mapping paper](#):

$$\begin{aligned}
\hat{R}_{\text{Chess.com,bullet}} &= -531.71 + 0.9871 R_{\text{Lichess}}, \\
\hat{R}_{\text{Chess.com,blitz}} &= -551.54 + 1.0853 R_{\text{Lichess}}, \\
\hat{R}_{\text{Chess.com,10minrapid}} &= -495.04 + 1.0741 R_{\text{Lichess}}, \\
\hat{R}_{\text{Chess.com,>10rapid}} &= -369.02 + 0.9374 R_{\text{Lichess}}.
\end{aligned}$$

Tyto rovnice mapují hodnoty hodnocení, nikoli percentilové pořadí. Percentilní pořadí jsou definované uvnitř konkrétního hráčského fondu a přehrávač Chess.com a Lichess bazény mají různé tvary. V tabulce 2 pro 10minutový rychlý typ hry, první sloupec poskytuje malý vzorek možných Hodnoty hodnocení Chess.com Elo. Druhý sloupec uvádí odpovídající Percentilní hodnocení Chess.com zobrazené na Chess.com k 31. květnu 2026. A 71,7 hodnocení na Lichess.org za 10minutový rapid odpovídá hodnocení Lichess Elo 1,644² a tato hodnota je uvedena v třetí sloupec. Pokud použijeme mapovací rovnici Chess.com–Lichess, 1 644 na Lichess je ekvivalentní asi 1 271 na Chess.com. Za předpokladu odhadovaného mapovací rovnice je přiměřená, mezera +522 mezi hodnocením Chess.com v sloupec 1 a mapované hodnocení ve sloupci 4 lze vysvětlit velkým číslem příležitostných hráčů v hráčském poolu Chess.com nafukování Chess.com percentilové pořadí. Jinými slovy, být v 72. percentilu na Lichess.org je větší úspěch než být v 72. percentilu na Chess.com, protože percentilové pozice se nemapují z jednoho systému do druhého. Skutečné Elo hodnoty lze mapovat pomocí rovnic odhadnutých ze skóre stejného hráče napříč dva systémy.

Chess.com rating	Chess.com pct.	Same-pct. Lichess	Mapped equiv.	Gap
749	71.7%	1,644	1,271	+522
1,012	86.2%	1,871	1,515	+503
1,358	95.4%	2,095	1,755	+397
2,383	99.9%	2,557	2,251	-132

Table 2: Ilustrativní srovnání mezi 10minutovým rychlým percentilem Chess.com bodů a stejných percentilových bodů v aktivním rapidu Lichess 10+0 distribuce po vyžadování alespoň pěti her 10+0. “Mapovaný Chess.com ekv.” platí aktuální Elo+Chess Rychlý převod 10+0 na stupnici hodnocení na stejné procentuální hodnocení Lichess. The Řady nejvyššího ocasu jsou zahrnuty, aby ukázaly rozdíl v měřítku, ale jsou vybaveny ratingová linie je nejspolehlivější v hlavním koučovacím rozsahu.

²Calculated from the cumulative distribution function of the ZXQPROT3ZXQ 10+0 rapid latest-rating sample after requiring at least five games in that exact game type; see Table 1 and Figure 3.

Důvodem mezer v tabulce 2 je mapování odhaduje ekvivalentní hodnocení herní síly, zatímco percentil popisuje pozici v rámci vlastní populace jedné platformy. Hráč na 71,7 procenta Chess.com proto není automaticky 71,7-percentilním hráčem na Lichess, a naopak. Pozorovaná čísla jsou v souladu s veřejností Chess.com rapid pool s mnohem větším množstvím méně hodnocených nebo příležitostných hráčů, zatímco aktivní přesné řízení Lichess pool je koncentrovanější mezi zkušenými repetitory hráčů. Klíčovým bodem není, že jedna percentilová škála je „správná“ a „správná“. Jiný je „špatný“; jde o to, že procentuální hodnocení není přenosné mezi platformami pokud základní rozdělení hráčů nemá stejný tvar.

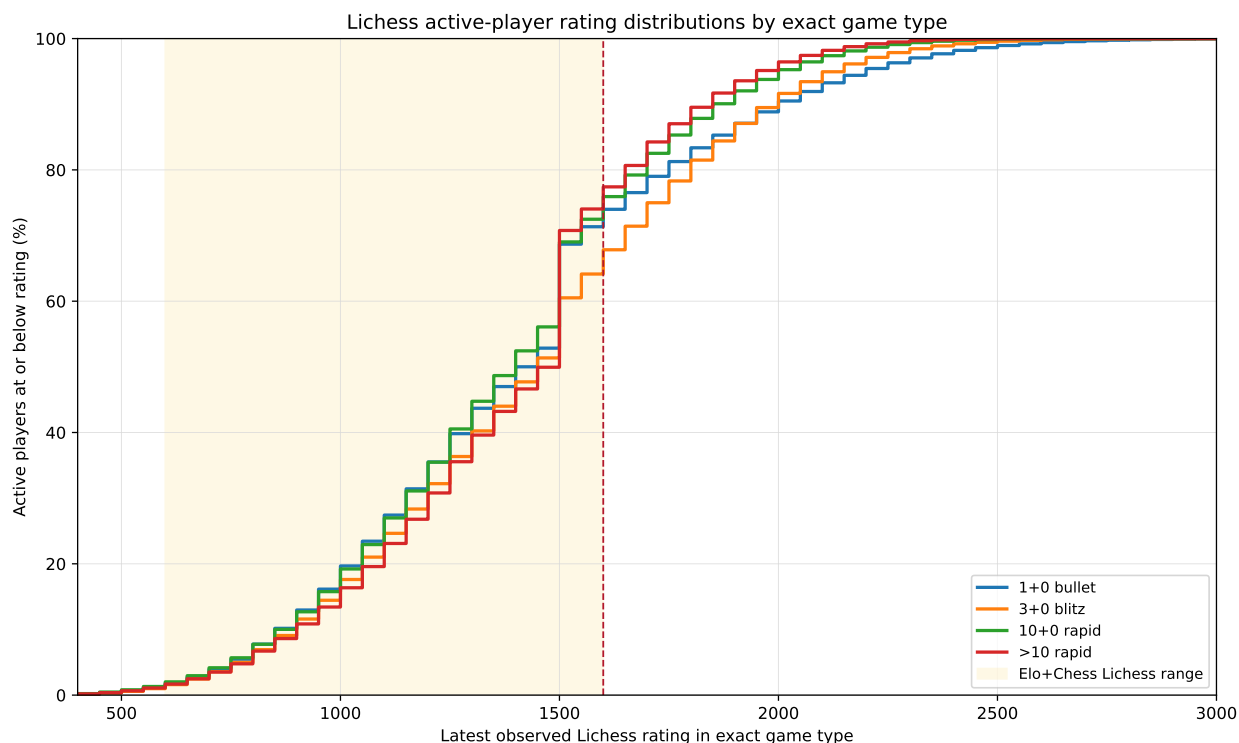


Figure 1: Kumulativní rozdělení hodnocení aktivních hráčů podle přesného typu hry. The žlutý pruh označuje řadu 600–1600 Lichess používanou hlavním Elo+Chess koučovací měřítko.

Note on the 1500 spike. The visible jump near 1500 je v datech spíše přítomno než vytvořeno vyhlazováním. V nejnovějším pozorovaná distribuce hodnocení, přesných 1500 hodnocení představuje 13,1 % z 1+0 bullet hráči, 5,6 17,6

počáteční nebo dočasné ukotvení v systému hodnocení.

Přesný bod 1500 je z velké části artefaktem málo aktivních účtů. Komu zkontrolujte toto, po požadavku jsme přepočítali stejné distribuce nejnovějšího hodnocení aby každý hráč měl alespoň pět her v přesném typu hry. Špice skoro zmizí pod tímto filtrem: přesných 1500 hodnocení klesne z 5,6–17,6 % z hráčů v distribuci pro všechny hráče na 0,2 – 1,2 % v pěti a více hrách distribuce uvedená v tabulce 1.

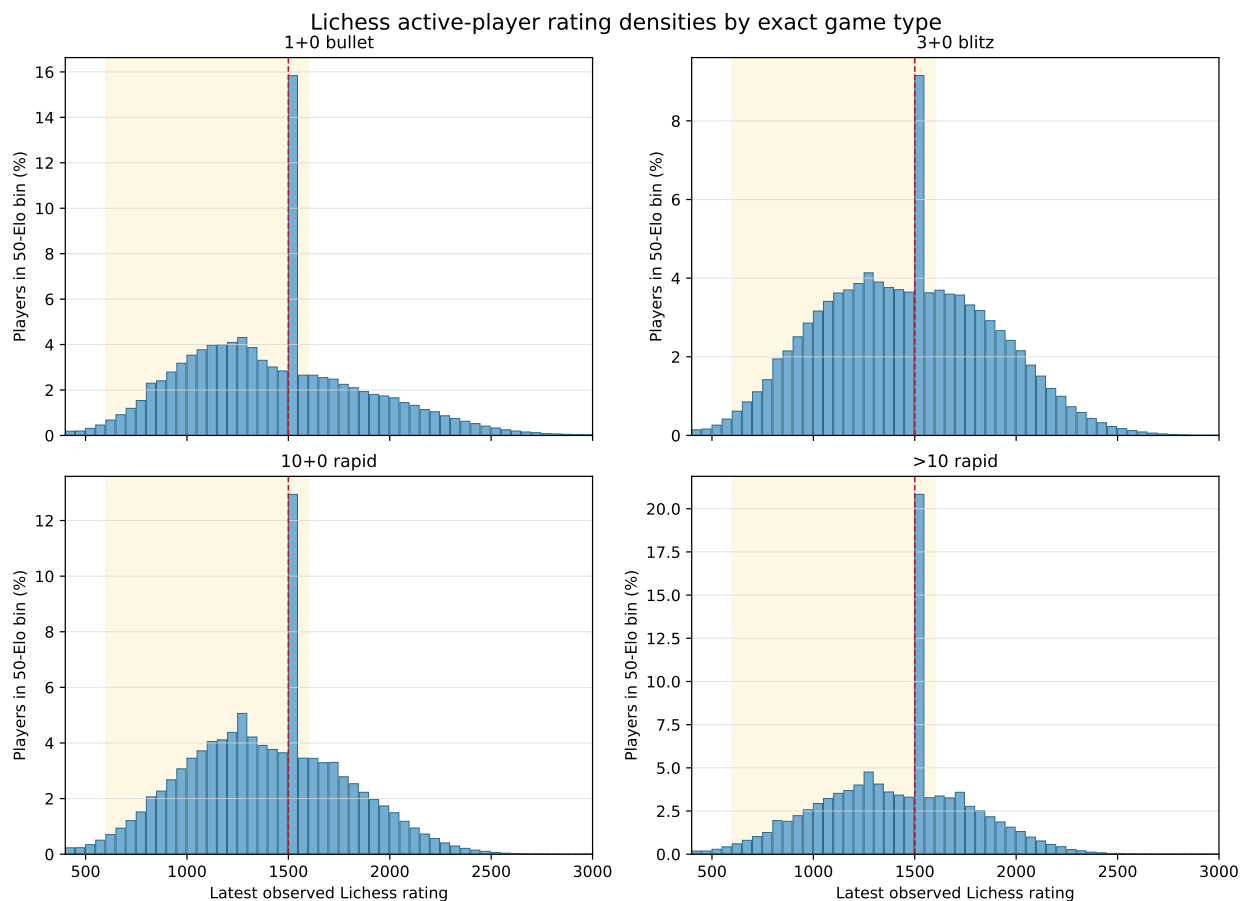


Figure 2: Zobrazení hustoty stejného rozdělení hodnocení aktivních hráčů. The přerušovaná čára označuje 1500 Lichess, kde mají nezpracované distribuce s nejnovějším hodnocením velký přesný hmotnostní bod. Žlutý pruh označuje 600–1600 Lichess rozsah používaný hlavními metrickými sestavami Elo+Chess. Hrot je nejvíce vidět v hustota všech hráčů, protože mnoho málo aktivních účtů zůstává přesně na úrovni 1500; Obrázek 4 opakuje graf hustoty po vyžadování alespoň pěti her v přesném typu hry, což do značné míry odstraňuje tento artefakt.

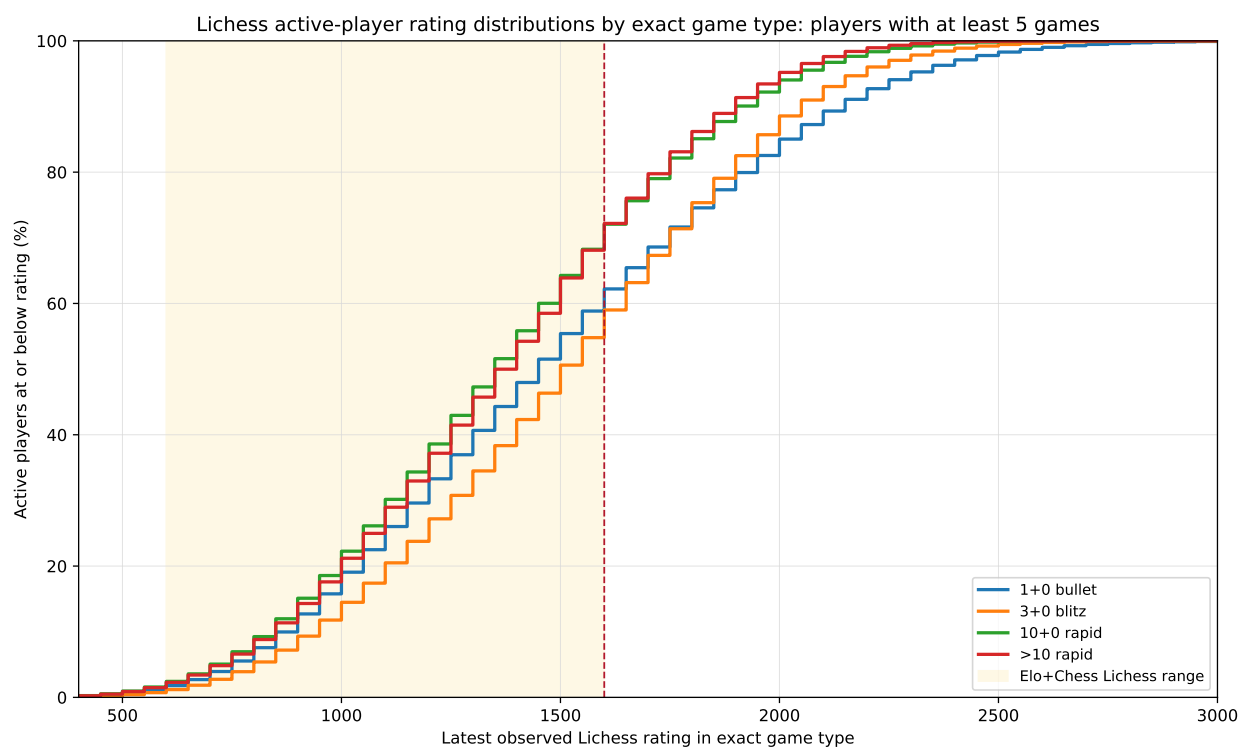


Figure 3: Kumulativní rozdělení hodnocení aktivních hráčů po vyžadování at alespoň pět her v přesném typu hry. Skok 1500 je značně snížen.

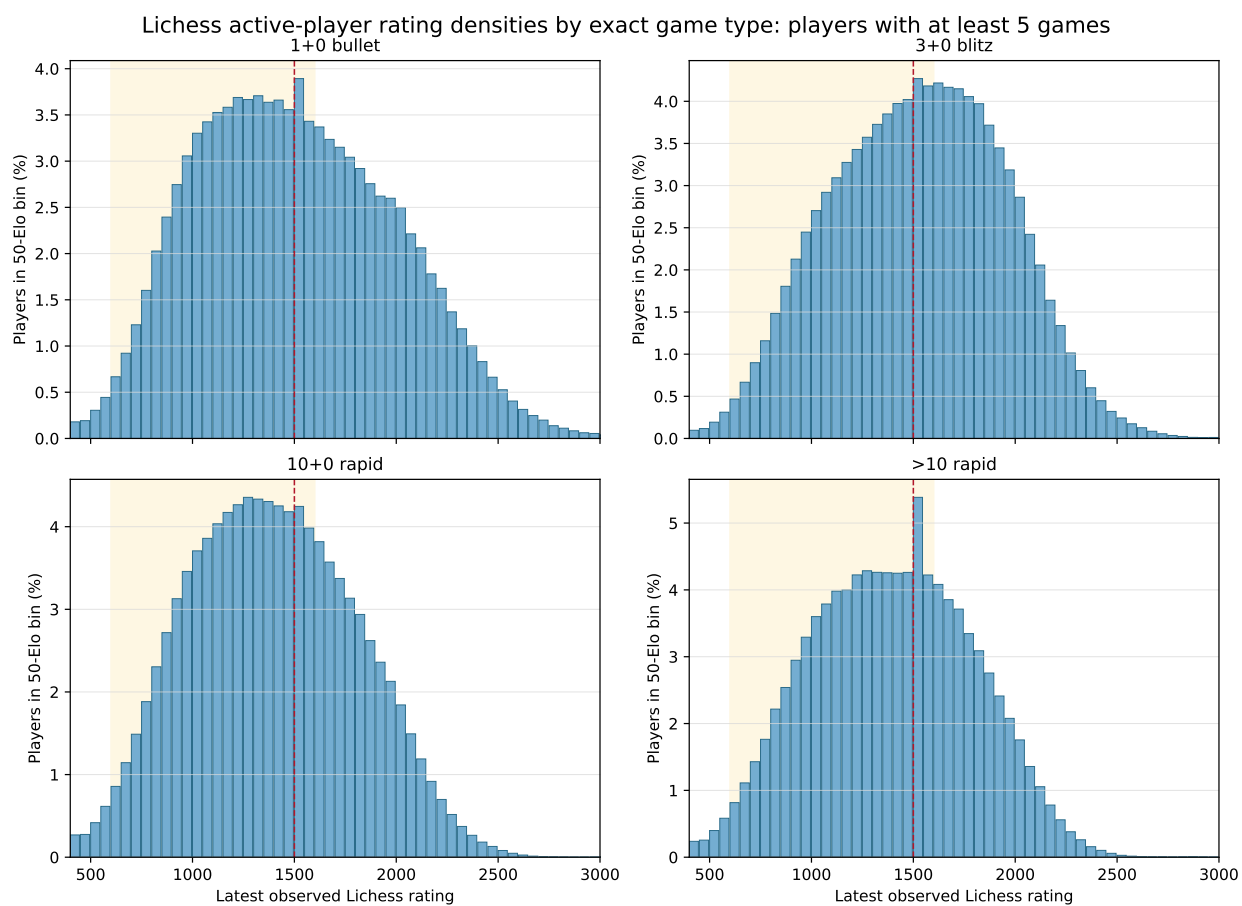


Figure 4: Zobrazení hustoty po vyžadování alespoň pěti her v přesné hře typu. Tento pohled potvrzuje, že velký 1500 hmotnostní bod v all-player distribuce je většinou řízena málo aktivními účty. Značky žlutého pruhu rozsah 600–1600 Lichess používaný v hlavních přehledech metrik Elo+Chess.

4 Měřítka aktuálních artefaktů benchmarku

Výpočet je záměrně velký. Nejedná se o oděr profilu, ale o malý sada stažených PGN nebo ručně sestavená sbírka instruktivních příkladů. Počáteční povolení Lichess pokrývalo kompletní měsíční soubory od ledna 2025 do března 2026. Tabulka nestratifikovaných skenů obsahuje 1 382 178 087 her a 2 764 356 174 řádků hodnocení hráčských her. V těchto řádcích je 6 623 438 jedinečných uživatelská jména hráčů se objeví alespoň jednou.

Speed	Games in raw scan	Player-game rows	Unique players
Bullet	519,375,423	1,038,750,846	3,077,635
Blitz	839,823,863	1,679,647,726	5,987,634
Rapid	22,978,801	45,957,602	1,535,865
All scanned speeds	1,382,178,087	2,764,356,174	6,623,438

Table 3: Nestratifikované měřítko skenování Lichess před vzorkováním. Skenování pokrývá kompletní měsíční soubory herní historie od ledna 2025 do března 2026.

Artefakty produkčního benchmarku jsou mnohem menší než surové skenování, protože jsou vybrány pro metrickou konstrukci. Jsou stále velké: měřítko fáze je expanze a agregace stavu přesunu přes stovky tisíce her na typ hry.

Game type	Games	Player-game rows	Unique players	Move states	Feature columns
1+0 bullet	483,974	967,948	119,732	22,985,272	277
3+0 blitz	479,974	959,948	129,398	28,138,904	277
10+0 rapid	925,636	1,851,272	229,235	57,554,339	277
>10 rapid	660,540	1,321,080	133,661	38,580,214	277

Table 4: Měřítka stratifikovaných artefaktů benchmarku po zarovnání k běžné schéma funkce runtime s 277 sloupci. A “řada hráč-hra” znamená jednu hru z pohledu jednoho hráče, takže většina her vytvořit dvě řady hráčů a her.

Jen pro bleskový 3+0 používá aktuální produkční benchmark 479 974 her zahrnující 129 398 unikátních hráčů a 28 138 904 různých řad tahu. The zarovnaná tabulka funkcí bleskového hráče a hry obsahuje 277 odvozených proměnných. The artefakty rychlého vytváření funkcí jsou ještě širší ve fázi přesunu: úplné Tabulka stavů rychlého pohybu použita při konstrukci prvku obsahuje 168 sloupce stavu přesunu před agregací do funkce hráče s 277 sloupci stůl. Artefakty frontendu runtime benchmarku pro všechny čtyři typy her jsou poté zarovnány se stejným schématem herních funkcí s 277 sloupci, včetně schématu rozšířené takticko-taxonomické unikové proměnné. Z těchto tabulek funkcí, report vybírá a zobrazuje koučovací metriky, které jsou nejlépe interpretovatelné.

Tyto počty jsou důležité pro interpretaci výsledků. Referenční křivky nejsou odhady z několika stovek ručně zkontrolovaných her. Místo toho oni pocházejí z desítek milionů hodnocených správních států a více než čtyř milionů hráčských herních řad napříč čtyřmi aktuálními produkčními kategoriemi her.

5 Přehled potrubí

Srovnávací stavební potrubí má šest hlavních kroků:

1. Vyberte hodnocené hry Lichess v cílovém typu hry a hodnocení.
2. Analyzujte každý tah hry po tahu a uložte tah před a po tahu palubní stát.
3. Vypočítejte mezilehlé proměnné stavu pohybu z hrací desky, pohybu, z pohledu hráče a pohledu soupeře.
4. Agregujte proměnné stavu pohybu do proměnných funkcí hráče ve hře.
5. Převedte proměnné hráče a hry na metriky přehledu s přehledem směr výkladu: vyšší je obvykle lepší, nižší obvykle lepší nebo popisný.
6. Zprůměrujte každou metriku podle segmentu Elo a výsledek ručně zkontrolujte vztah.

Tato struktura je důležitá, protože mnoho metrik přehledů není přímo viditelné v záhlaví PGN. Vyžadují rekonstrukci polohy, určení které figurky se pohybovaly od kterých počátečních polí, porovnání před tahem a materiál a struktura po přesunu a poté agregace výsledku z přesunutou úroveň na úroveň hráče-hry.

6 Mezilehlé proměnné

Mezilehlé proměnné jsou navrženy tak, aby popisovaly, co se na desce změnilo a proč je to důležité z pohledu hráče sestav. Příklady:

- číslo tahu, číslo tahu hráče, strana k tahu, barva tahu a hráč barva;
- ze čtverce, do čtverce, typ pohyblivého kusu, typ zachyceného kusu a zda šlo o zajištění, kontrolu, hrad, povýšení nebo kamaráda;
- materiálová bilance před a po přesunu;
- zda vedlejší figurka opustila své počáteční pole o dané číslo tahu;
- zda se tatáž figurka pohybovala po sobě jdoucích hráčských tahů v otevírací;
- zda se dáma pohnula před tahem 10;
- strana rošády, číslo tahu rošády a zda došlo k propojení věží po rošádování;
- volné figury, závěsné figury, zdvojení pěšci, izolování pěšci, prošel pěšci a zaostalí pěšci;
- stav věžového souboru, otevřené soubory, napůl otevřené soubory a královská bezpečnost vlastnosti;
- taktické příležitosti, provedená taktika, taktika tlaku materiálu, zaručené následné zisky a realizované zisky.

Přesná mezisada se liší podle fáze sestavení. Databáze zpráv za běhu uchovávají pouze sloupce potřebné k rychlému poskytování přehledů při úplném sestavení funkcí artefakty zahrnují širší tabulky pohybových stavů a záplaty funkcí používané ke konstrukci poslední řady herních prvků.

7 Příklady metriky principu otevření

Principy otevírání jsou užitečné příklady, protože je pro uživatele snadné pochopit a snadno se připojit k výpočtům pohybu-stavu.

7.1 Minor Pieces Developed by Move 10

Tato metrika počítá, kolik ze čtyř hráčských rytířů a střelců zbývá jejich domácí pole 10. tahem hráče. Mezivýpočet používá barvu pohybu, typ pohyblivé figurky, počáteční pole a pohyb hráče číslo. Pro bílého jsou domácí čtverce b1, g1, c1 a f1. Pro Blacka jsou b8, g8, c8 a f8. Pro každou hru hráče, Elo+Chess počítá odlišný domov pole mezi rytíři a střelci, která se pohybovala tahem hráče 10.

Toto je vyšší-je-obvykle-lepší metrika. Netvrdí, že každý nezletilý figurka se musí vždy pohybovat tahem 10, ale mezi vyvíjejícími se hráči empiricky vztah je jasný: hráči, kteří dosáhnou vyšších Elo kbelíků, mají tendenci se rozvíjet dříve menší kusy.

7.2 Repeated Piece Moves in the Opening

Tato metrika místo toho počítá další rané tahy strávené opětovným pohybem stejné figurky o vývoji něčeho nového. Mezivýpočet používá předchozí cílové pole, aktuální pole počátku a číslo tahu hráče. Pokud stejná figurka se pohnula v předchozím tahu hráče a znovu se přesune uvnitř prvních 10 tahů hráče se počet opakovaných tahů zvyšuje.

Toto je nižší-je-obvykle-lepší metrika. Měří promarněné otevírací teploty v a jednoduchý způsob. Některé opakované pohyby jsou takticky odůvodněné, ale v měřítku bucket curve vystihuje obecnou tendenci: silnější hráči u cíle úrovně stráví méně raných tahů opakovaným pohybem stejné figurky.

7.3 Queen Touched Before Move 10

Tato metrika počítá tahy dámy před tahem hráče 10. Střední výpočet používá typ pohyblivé figurky a číslo tahu hráče. Pokud je pohyblivý kus je královna a číslo tahu hráče je maximálně 10, počet se zvyšuje.

To je také nižší - je - obvykle - lepší pro cílový rozsah Elo+Chess. Brzy tahy královny mohou vyhrát materiální nebo vynutit ústupky, ale mnoho rozvíjejících se hráčů přivést královnu ven příliš brzy, ztratit tempo kvůli útokům na královnu a zdržet se vývoj menších kusů a bezpečnost krále.

7.4 Castling and King Safety

Metriky rošády používají rošádové praporky, stranu rošády, číslo rošádového tahu a stav desky po rošádu. Mezi příklady patří, zda hráč hradil, zda hráč hradil tahem 10, zda hráč hradil krále nebo královna, zda byly věže po rošádu spojeny a zda soubory věží byly po rošádu otevřené nebo polootevřené.

Všechny tyto metriky nejsou interpretovány stejným způsobem. Dokončení rošády a rošáda tahem 10 jsou obecně vyšší – obvykle – lepší v cílovém rozsahu. Rošáda Queenside nebo rošáda do otevřených souborů může být více závislá na kontextu. To je důvod, proč Elo+Chess jednoduše nezakóduje morální pravidlo; staví kbelík křivky a kontroluje, které vztahy jsou dostatečně stabilní pro koučování.

8 Křivky lopaty a ruční kontrola

Po výpočtu metrik hry hráče je Elo+Chess seskupí podle Lichess Elo bucket a vypočítá průměry bucket. Průměr kbelíku je hrubý empirický srovnávací bod pro danou metriku. Trendová čára je pak použita ve zprávě k přiřadíte uživatelským hodnotám známky ve stylu Elo.

Křivky jsou před použitím ručně kontrolovány. Metrika je užitečnější když prostředky skupiny vykazují soudržný vztah v celém rozsahu hodnocení obsluhované produktem. Metrika je méně užitečná, když je hlučná, plochá, vysoká nemonotónní nebo řízené speciálními případy, které se uživateli těžko vysvětlují.

Tento krok ruční kontroly je zvláště důležitý nad svrškem Elo+Chess přerušení. Silnější hráči vědí, kdy mají porušit zásady otevírání a tak to dělají efektivně. Nad rozsahem pro začátečníky až brzy pokročilé, vztah mezi jednoduchými návyky a hodnocením se často zplošťuje nebo mění tvar. Produkt proto zdůrazňuje rozsah, kde jsou empirické vztahy silné a pedagogicky užitečné.

9 Příklad křivky otevření celého rozsahu

Obrázek 5 ukazuje čtyři křivky bleskového otevření 3+0. Šedé body ukazují širší rozsah hodnocení Lichess. Modré body a modrý trend řádek ukazuje řadu 600–1600 Lichess používanou hlavním koučováním Elo+Chess benchmarky. Přerušovaná červená čára označuje horní hranici.

Stejný vzor je viditelný v plném rozsahu 10+0 rapid a long-rapid artefakty. Ústřední modelovací rozhodnutí se nemění: použijte empiricky stabilní rozsah pro tvrzení koučování a vyvarujte se předstírání, že jde o jednoduchou návykovou křivku zůstává stejně informativní navždy.

Stejný vzor se číselně zobrazí v tabulce 5. Svahy jsou hlášena jako změna metrické hodnoty na 100 bodů Elo. Rošáda, menší kus vývoj, opakovaný pohyb figur a raný pohyb královny, to vše je jasné pohyb přes 600–1600. Nad 1600 jsou v těchto sjezdovky mnohem menší příklady, což je v souladu s myšlenkou, že hráči s vyšším hodnocením se rozbíjejí základní principy selektivněji a úspěšněji. Druhý mechanismus je nasycení. U některých zvyků dosahuje metrika praktického stropu nebo sladkosti místo: silní hráči se už tento zvyk naučili, takže není důvod křivka, aby se stále prudce pohybovala nahoru. Více také není vždy lepší po ten bod. Data proto oba zplošťují, protože silnější hráči se mohou zlomit pravidla v konkrétních pozicích a protože mnoho základních návyků již dosáhlo jejich užitečný provozní rozsah.

10 Proč na horní hranici záleží

Cílový uživatel Elo+Chess je začátečník až pokročilý hráč. Pro tyto hráčů, základní návyky jsou silně spojeny s hodnocením: vyvíjet figurky, vyhnout se opakovaným raným tahům figur, vyhnout se zbytečným raným tahům dám, hrad včas spojte věže, spravujte materiál a vyvarujte se ponechání volných kusů nebo visící.

Při vyšším hodnocení se stejné návyky na povrchu stávají méně diagnostickými. A silný hráč může odložit rošádu, protože centrum je zavřené, posuňte královnu dříve protože pozice obsahuje konkrétní taktiku, nebo přesuňte stejnou figurku dvakrát protože vyhrává tempo nebo si vynucuje strukturální ústupek. Stále ten syrový zvyk má význam, ale vztah mezi zvykem a hodnocením se prohlubuje podmíněný. To je důvod, proč Elo+Chess používá horní cutoff pro své hlavní koučování nároky a proč jsou uživatelé nad tímto rozsahem varováni, že úzký vztah mezi principiálními návyky a hodnocením se rozpadá.

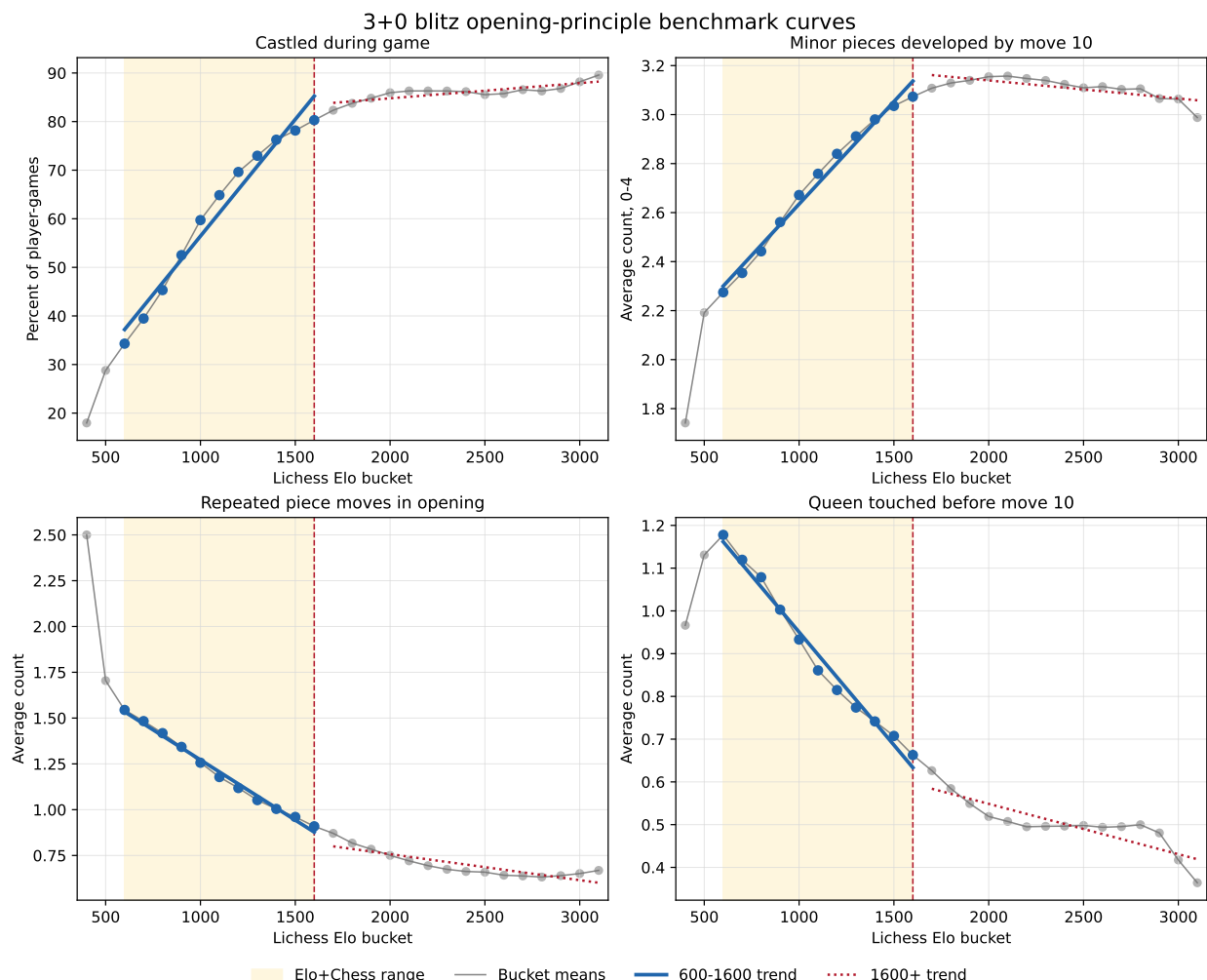


Figure 5: Vybrané benchmarkové křivky principu otevření pro bleskový blesk 3+0. The vztahy jsou nejsilnější a nejužitečnější v řadě koučování Elo+Chess. Nad horní hranicí se několik vztahů podstatně zplošťuje.

11 Dvoustranné taktické metriky a procentuální cíle

Některé metriky mají opačný tvar než příklady s principem otevírání. Vidličky jsou užitečným příkladem, protože událost je ze své podstaty oboustranná. Hráč, který najde a převede užitečnější forky než kolegové dělá něco dobrého, ale populační poměr převedených forků klesá s hodnocením, protože je silnější protivníci umožňují méně příležitostí rozvětvení a brání se taktickým hrozbám aktivněji.

Obrázek 8 ukazuje tento efekt. Obě konvertované vidlice by hráč hlášení a konvertované vidličky soupeřem jsou méně časté jako hodnocení stoupá. To neznamená, že provádění forků je špatné. Znamená to syrovou událost rychlost je částečně řízena chybovostí soupeře a texturou hry. Pro metriky takto Elo+Chess proto používá percentily peer-history uvnitř hráčský vlastní Elo bucket, spíše než spoléhat se pouze na globální bucket-mean trend.

Chcete-li vytvořit tyto percentilové cíle, Elo+Chess vzorkuje hráče podle typu hry a Elo bucket, poté použije nedávné kvalifikační hry pro každého hráče, který je vybrán. The aktuální cílová sada

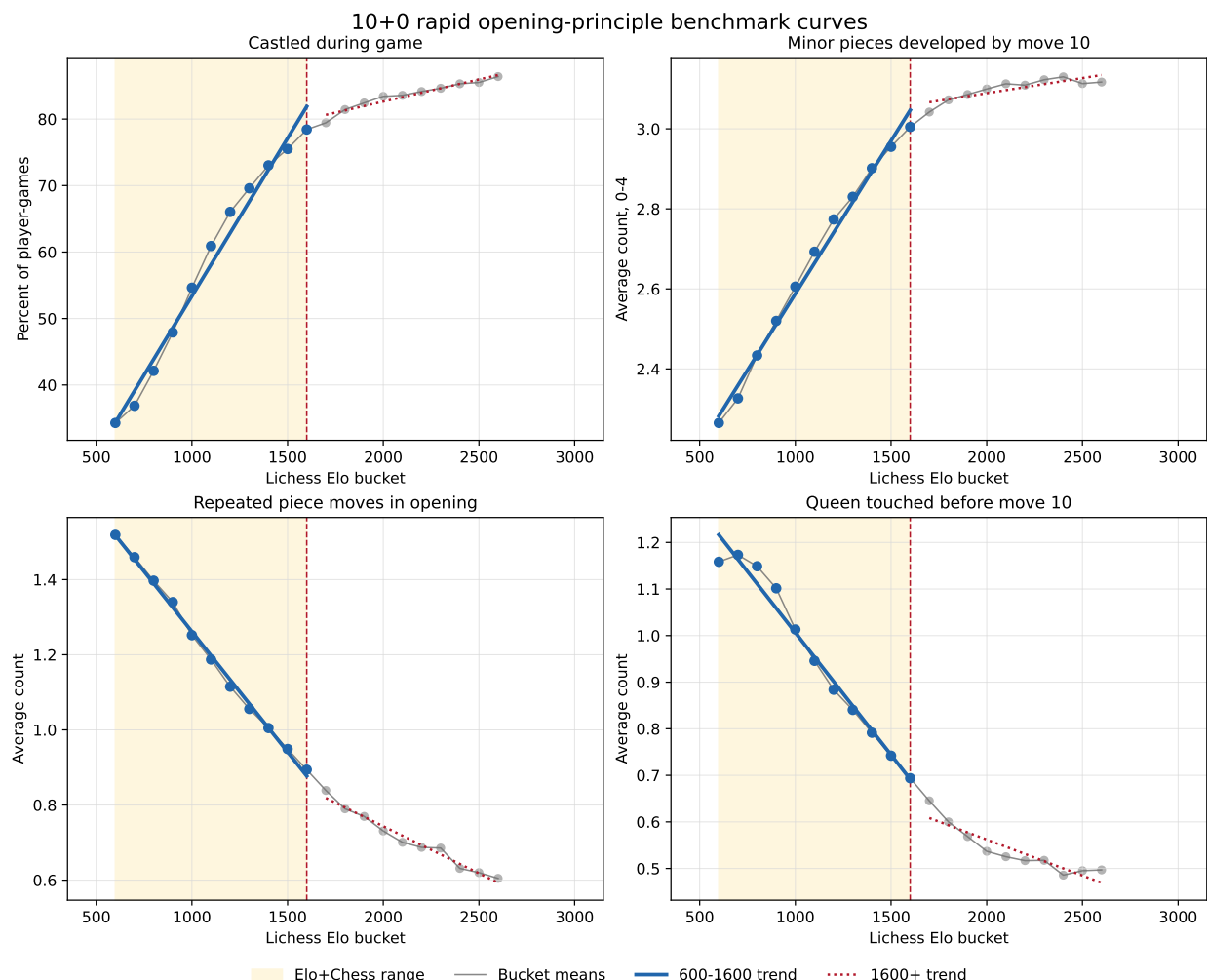


Figure 6: Vybrané benchmarkové křivky principu otevření pro rychlé hry 10+0. The vztahy jsou nejsilnější a nejužitečnější v řadě koučování Elo+Chess. Nad horní hranicí se několik vztahů podstatně zplošťuje.

používá 11 kbelíků od 600 do 1600, 200 hráčů na kbelík a až 50 posledních kvalifikačních her na hráče pro každou produkci typ hry. To dává 2 200 vzorků hráčů na typ hry a přibližně 110 000 hráčských herních řádků na typ hry pro percentil v rámci bloku distribuce.

Obrázek 9 znázorňuje výsledný cíl distribuce pro konvertovaný fork rate hráče se vzorkem v bleskovém 3+0. Každá krabice je distribuce v rámci skupiny mezi vybranými hráči, nikoli průměrná hodnota celé skupiny jednotlivé hry. V uživatelském rozhraní sestavy je hodnota uživatele porovnána s relevantní box pro uživatelem namapovaný Elo kbelík. Z toho vycházejí výroky jako např. zda je konvertovaný fork poměr hráče vysoký nebo nízký vzhledem k ostatním na stejné úrovni, i když celkový počet událostí v populaci klesá s hodnocením.

12 Náklady na replikaci

Metoda je v zásadě reprodukovatelná, ale není lehká. Přestavba referenční křivky vyžadují:

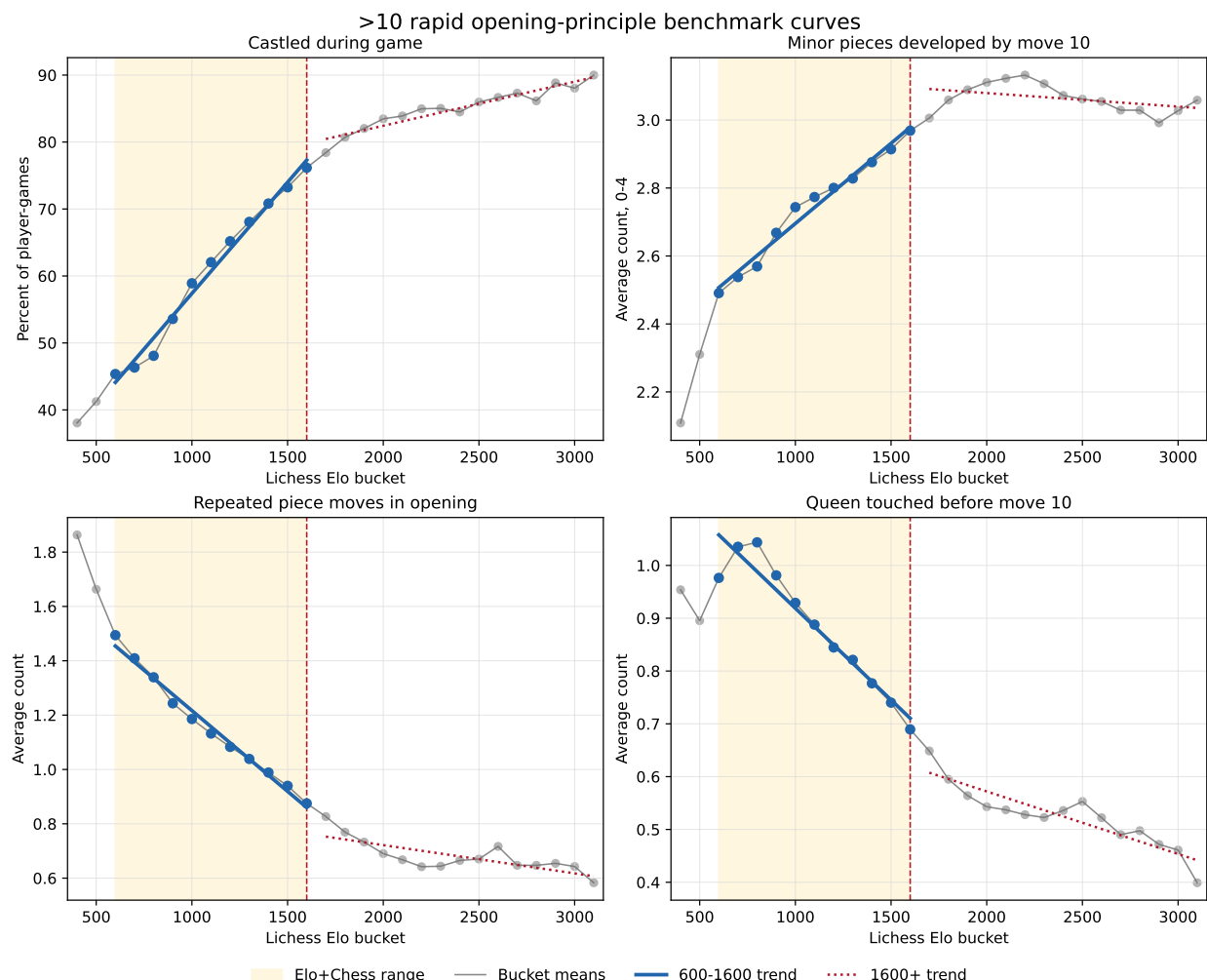


Figure 7: Vybrané benchmarkové křivky s principem otevření pro delší rychlé hry.

- přístup k velkým souborům Lichess s měsíční historií her;
- stratifikované filtrování podle typu hry, skupiny hodnocení a způsobilosti pravidla;
- zákonný tahový analyzátor a hodnotitel správních orgánů schopný expandovat desítky milionů stěhovacích států;
- logika funkce pro otevření, materiál, struktura pěšce, bezpečnost krále, věž aktivita, taktika, řízení času a výsledek hry;
- agregace ze stavu přesunu do herních funkcí hráče;
- souhrny na úrovni segmentu pro každou metriku přehledu;
- ruční kontrola každé srovnávací křivky kandidáta před jejím použitím trenérskou zprávou.

Pro bleskový blesk 3+0 to znamená téměř půl milionu her, téměř jeden milion řádků hráče a hry a více než 28 milionů řádků stavu pohybu. Přes čtyři benchmarkové typy her, aktuální benchmarkové artefakty obsahují více než 2,55 milionů her, více než 5,10 milionu herních řádků hráče a

Game type	Metric	600–1600 slope per 100 Elo	1600+ slope per 100 Elo
1+0 bullet	Castled during game	4.137	0.550
	Minor pieces developed by move 10	0.063	0.009
	Repeated piece moves in opening	-0.051	-0.026
	Queen touched before move 10	-0.040	-0.011
3+0 blitz	Castled during game	4.804	0.313
	Minor pieces developed by move 10	0.084	-0.007
	Repeated piece moves in opening	-0.066	-0.014
	Queen touched before move 10	-0.053	-0.012
10+0 rapid	Castled during game	4.753	0.665
	Minor pieces developed by move 10	0.076	0.008
	Repeated piece moves in opening	-0.064	-0.025
	Queen touched before move 10	-0.052	-0.015
>10 minute rapid	Castled during game	3.320	0.659
	Minor pieces developed by move 10	0.047	-0.004
	Repeated piece moves in opening	-0.059	-0.010
	Queen touched before move 10	-0.035	-0.012

Table 5: Přibližné sklony křivky lopaty pro čtyři příklady principu otevírání metriky napříč typy produkčních her.

Game type	Elo buckets	Players per bucket	Games per player	Player samples
1+0 bullet	11	200	49–50	2200
3+0 blitz	11	200	27–50	2200
10+0 rapid	11	200	37–50	2200
>10 minute rapid	11	200	47–50	2200

Table 6: Současné procentuální cílové sady cílů. Toto jsou vzorky historie hráče používané k odhadu rozdělení v rámci segmentu pro percentilové metriky.

více než 147 milionů řádků stavu pohybu. Stupnice je ústřední pro hodnotu benchmark: umožňuje reportu ukázat uživatelům, co se skutečně děje ve hrách hráči na jejich úrovni, než aby se spoléhali pouze na obecná šachová hesla.

13 Omezení

Referenční křivky jsou empirické popisy, nikoli kauzální důkaz. Hráč nezíská hodnocení pouze tím, že odpovídá průměrnému průměru. Metriky jsou také nedokonalé shrnutí složitých pozic: raný tah dámy může být vynikající v jedné poloze a bezohledně v jiné. Účelem křivek je identifikovat opakující se návyky a porovnat je s chováním vrstevníků, nikoli s nahradit výpočet nebo koučovací úsudek.

Křivky jsou také křivky v měřítku Lichess. Když uživatelé přinesou hru Chess.com historie, Elo+Chess mapuje relevantní segment hodnocení na benchmark Lichess měřítko pomocí samostatného multiplatformního mapovacího modelu. Ta mapovací vrstva je popsán v doprovodné výzkumné poznámce Elo+Chess k Lichess-to-Chess.com konverze hodnocení: [Estimating Cross-Platform Chess Rating Mappings with Modální regrese](#). Toto mapování by nemělo být chápáno jako konzervující percentil hodnot. Zachovává odhadovanou ratingovou ekvivalenci mezi těmito dvěma ratingy systémy; percentilové hodnocení zůstává specifické pro platformu, protože aktivní hráč distribuce jsou různé.

Two-sided converted-fork rates by Elo bucket

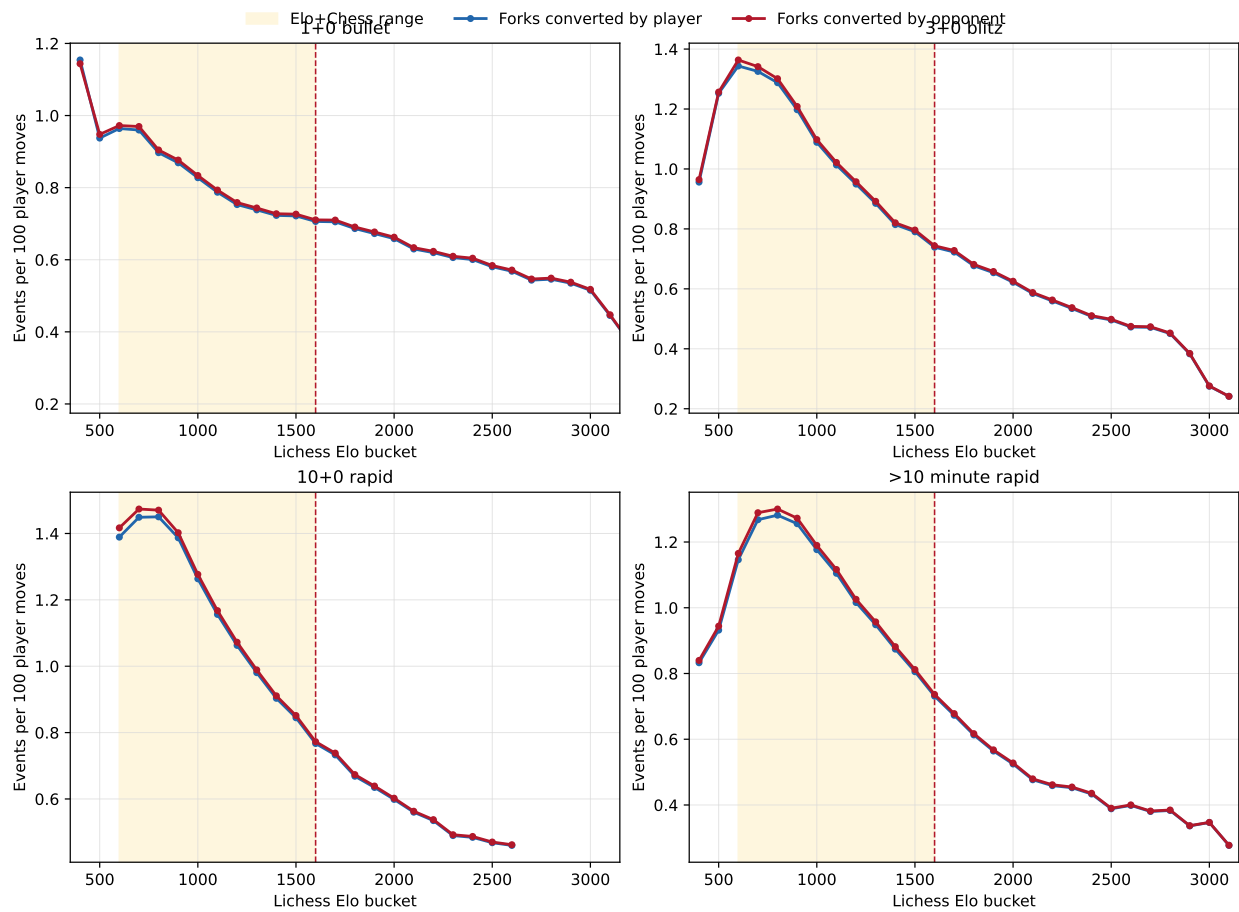


Figure 8: Oboustranné konvertované vidlicové sazby od Elo lopaty. Převedené události vidlice pád na vyšší hodnocení, protože obě strany brání lépe a méně čisté vidlice konverze jsou k dispozici.

14 Grafika výrobní zprávy

Produkční zpráva přemění artefakty benchmarku na panely na metrické úrovni které lze číst bez znalosti celého kanálu sestavení. Obrázky 10 and 11 ukažte tři příklady z živého návrhu sestavy Elo+Chess.

Panely s principem otevírání na obrázku 10 použijte běžnou prezentaci benchmark-křivky. Každá karta začíná na rodina metriky a název metriky, pak uvádí krátké pravidlo výkladu. The velký stupeň v pravém horním rohu je záměrně oddělen od surového metrická hodnota: odpovídá na uživatelskou otázku “jak se tento zvyk hodnotí oproti benchmarku?”, zatímco řádky níže zachovávají naměřené hodnoty. Čtyři řádky souhrnu zobrazují životnost, posledních 10, výhry a prohry. Toto dělá panelová diagnostická spíše než pouze dekorativní. Pokud Lifetime a Last 10 souhlasí, návyk je pravděpodobně stabilní. Pokud se výhry a prohry liší, metrika může být spíše závislé na výsledku nebo vázané na stav hry než na jednoduchý zvyk.

Oblast grafu používá omezené tmavé pozadí, zlatou trendovou linii a barevné značky pro čtyři plátky. V anotaci trendu je výslovně uvedeno, zda referenční vztah stoupá nebo klesá s Elo. To je důležité, protože některé metriky jsou vyšší-je-lépe, jako je vývoj menších kusů jiné jsou nižší-je-

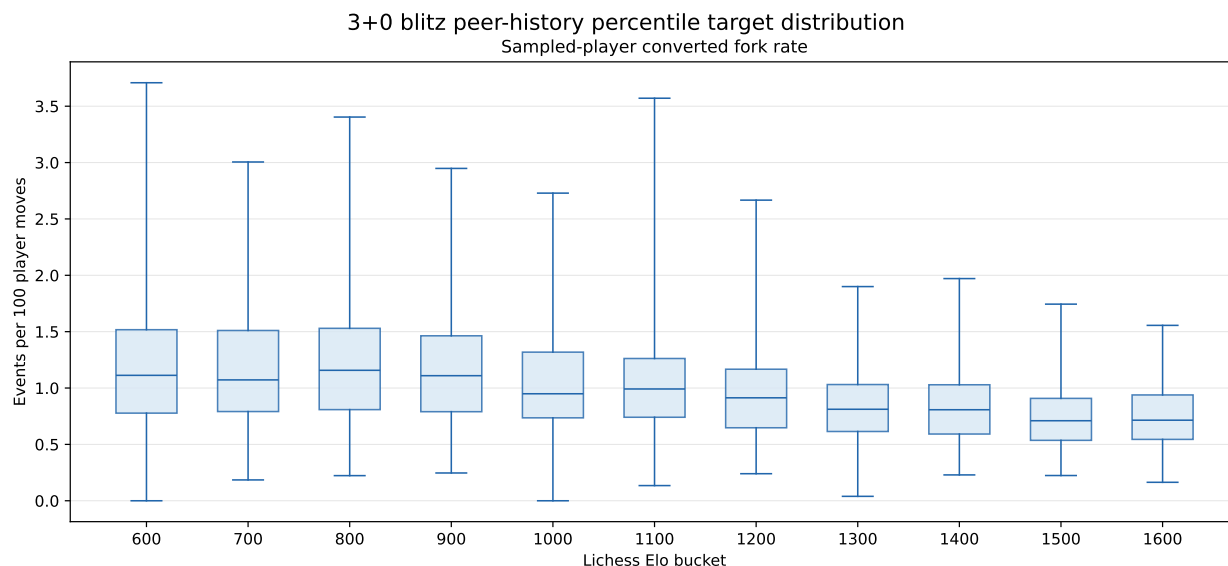


Figure 9: Příklad 3+0 blesková cílová distribuce peer-history pro hráče se vzorkem přepočtená rychlost vidlice. Percentily jsou počítány v rámci Elo segmentu, takže sestava porovnává hráče s vrstevníky, kteří čelí v podstatě podobné taktické příležitosti sazby.

lepší, jako je hnutí rané královny. Karta tedy nežádá uživatele, aby odvodil směr pouze ze svahu. finále Řádky „Co to znamená“ a „Tip koučování“ překládají výsledek benchmarku do konkrétní šachové instrukce při zachování empirického kontextu.

Obrázek 11 ukazuje percentilový styl prezentace používaná pro metriky, jejichž hrubý poměr není dobře reprezentován a jediný celosvětový trend. Vítězné forky jsou taktickým příkladem: míra populace počet převedených vidliček se mění s dovednostmi hráče i chybovostí soupeře. The karta proto porovnává hodnotu uživatele s historií peer hráče v zmapovaný kbelík Elo. Box-plot zobrazuje peer distribuci, včetně mezikvartilní rozsah, medián a vousy. Značka uživatele je vykreslena na stejné měřítko a označené jak hrubou hodnotou, tak procentuálním hodnocením. Sousední srovnávací segmenty zůstávají viditelné ve ztlumené podobě, takže uživatel může vidět, že peer bucket je místní, nikoli univerzální měřítko.

Tento návrh umožňuje auditovat percentilové metriky. Uživatel vidí hrubou hodnotu, percentil, peer bucket použitý pro srovnání a přeložený Mapování lopaty Chess.com na Lichess. Stejná životnost, posledních 10, výhry a Řádky ztrát jsou zachovány, takže skóre taktického percentilu lze stále kontrolovat pro efekty aktuálnosti a závislost na výsledku.

15 Závěr

Referenční zprávy Elo+Chess jsou sestaveny z velkého výpočtu stavu pohybu stratifikované vzorky Lichess. Každá metrika začíná jako soubor konkrétního stavu desky a pohyb-stavové proměnné, se stává prvkem hry hráče a je pak zprůměrován pomocí Elo bucketu k vytvoření empirické referenční křivky. Princip otevírání příklady ukazují, proč je tento přístup užitečný: jednoduché šachové návyky mají silné, viditelné vztahy s hodnocením v cílovém rozmezí, zatímco ty vztahy se nad horní hranicí často zplošťují nebo se stávají více podmíněnými. To je hlavní důvod, proč Elo+Chess zaměřuje svá tvrzení o koučování na rating rozsah, kde je data podporují nejzřetelněji.

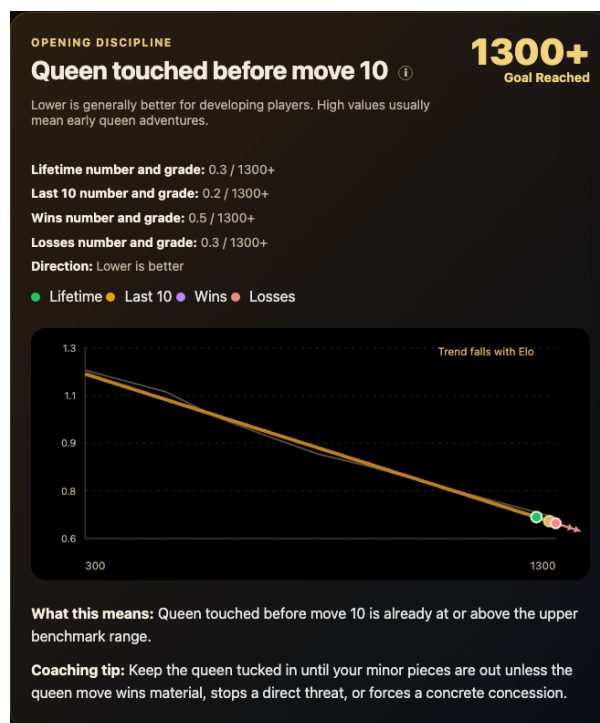
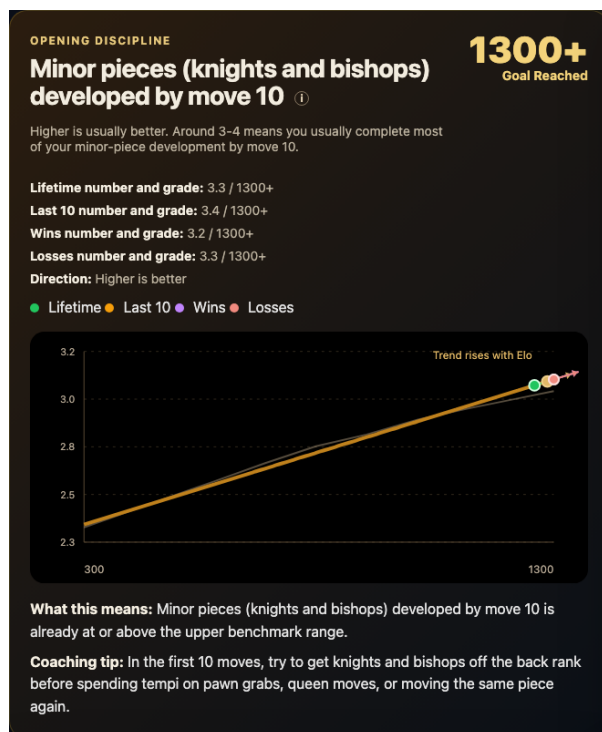


Figure 10: Výroba panelů na principu otevírání. Stejná vizuální gramatika zpracovává a metrika vyšší je lepší, vývoj menšího počtu kusů a metrika nižší je lepší, hnutí rané královny. Oba panely zobrazují životnost, posledních 10, výhry a ztráty takže zpráva může oddělit stabilní návyky od nedávných nebo závislých na výsledcích vzory.

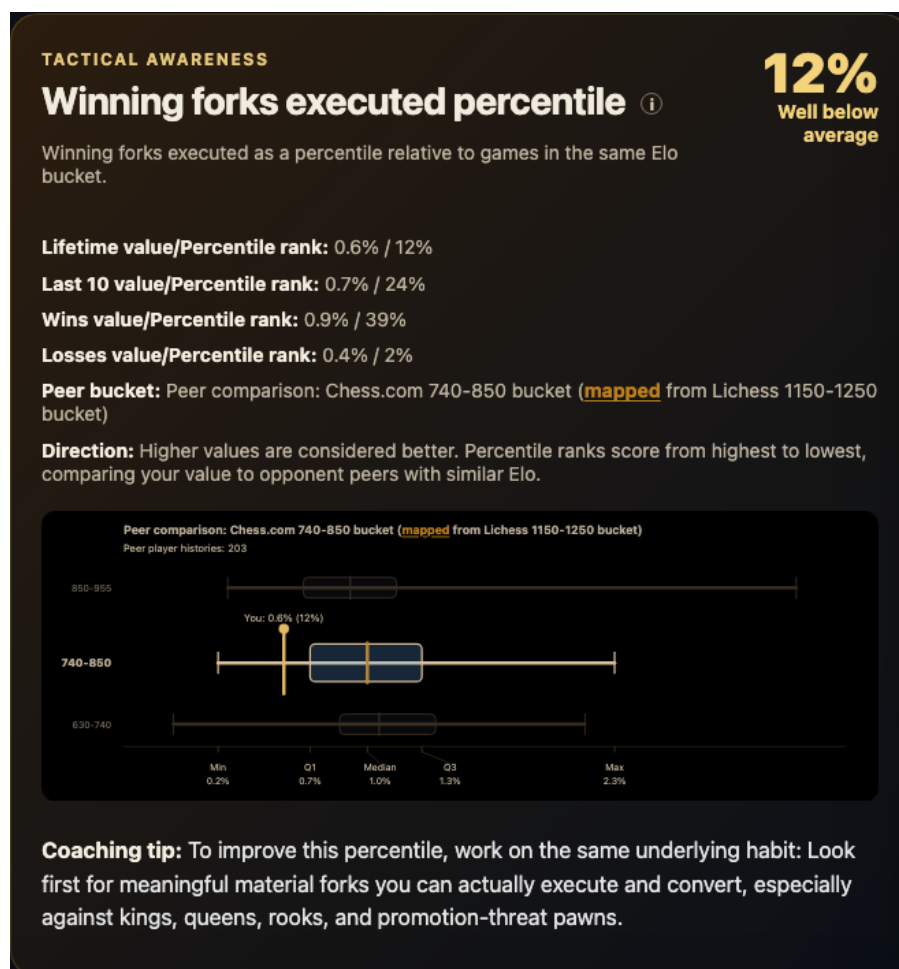


Figure 11: Byl proveden produkční percentilový panel pro vítězné vidličky. Krabicový plot ukazuje distribuci stejných segmentů, kterou používá sestava, zatímco zlato značka ukazuje hrubou hodnotu uživatele a procentuální hodnocení.