

為 Elo+Chess 建立分層 Lichess 基準曲線

Elo+Chess 研究筆記

2026 年 5 月 31 日

Abstract

Elo+Chess 基準報告是根據大量分層樣本建立的 Lichess.org 遊戲，本說明中使用的基準範例包含 483,974 1+0 子彈遊戲、479,974 3+0 閃電戰、925,636 10+0 快速遊戲、以及 660,540 個較長的快速遊戲。對於每種主要遊戲類型，管道都會擴展 每個遊戲進入移動前和移動後的棋盤狀態，計算數百個 中間變數和玩家遊戲變量，將這些變數聚合成 指導指標，然後按 Lichess Elo 儲存桶對這些指標進行平均。的 樣本是故意分層的，而不是與原始玩家成比例 池：目的是估計每個評級等級的穩定行為曲線，包括否則會稀疏的桶，而不是再現自然 評級分佈。結果是一系列經驗基準曲線：對於每個指標，我們可以詢問不同 Elo 等級的玩家展示給定的頻率 真實遊戲中的習慣或原則。本說明描述了方法、規模 計算，以及一些 涉及王位鑄造、小件開發的開放原則範例，重複的開局棋子動作和早期的皇后動作。重點是 方法論：Elo+Chess 並不假設西洋棋原理是普遍適用的 所有技能等級呈線性。相反，基準曲線透過以下方式檢查 評級範圍，產品專注於初級到早期高級區域 關係最強、最容易解釋的地方。

1 目的

基準系統的目標是將廣泛的國際象棋建議轉化為 可測量的數量。諸如“發展你的小作品”之類的陳述是 很有用，但是當我們可以說出一個有多少個小部分時，它就變得更具可操作性 玩家通常透過第 10 步進步，與同等程度的同齡人相比如何 評級範圍，以及該習慣如何在評級範圍內變化。

因此，Elo+Chess 根據實際遊戲建立了一個指標庫。每個指標是 計算每個玩家遊戲行，按遊戲類型和 Elo 桶進行匯總，以及 用作報告中的基準曲線。然後，該報告可以將用戶置於 相同的曲線並顯示使用者是否高於、低於或接近該行為 在可比較的玩家可以觀察到。

2 來源資料和分層

基準來源是公共 Lichess.org 遊戲資料庫。Elo+Chess 用途 按遊戲類型和評級桶劃分的大分層樣本，而不是小樣本 方便樣品。目前的生產運行時工件涵蓋四個主要 遊戲類別：

- 1+0 子彈，
- 3+0 閃電戰，
- 10+0 快速，
- 超過 10 分鐘的快速遊戲。

分層目標是評級等級的廣度。比例隨機 樣本對於描述整個玩家池是有效的，但它會 建構基準曲線效率低。最常見的評級桶 將主導樣本，而低支撐尾桶的數量太少 穩定指標平均值的遊戲。Elo+Chess 因此按平均遊戲進行採樣 埃洛桶。在當前的乾淨樣本構建中，原始遊戲是第一位

的篩選出已知玩家、已知評級、有效結果的有效評級遊戲，和已知的時間控制。然後將遊戲過濾給擁有足夠遊戲的玩家以足夠的速度和穩定的評級來代表連貫的技能水平：對於每位玩家，Elo+Chess 計算玩家的速度內評級跨度 ($\max(\text{Elo}) - \min(\text{Elo})$) 跨源池並且維持 僅當雙方玩家都至少具有配置的最小遊戲數時才進行遊戲，並且 速度內漂移不超過 300 Elo 點。漂移濾波器並不是指 消除正常的改進或差異；它減少了玩家的情況 遊戲會混合不同的技能水平，例如臨時帳戶，返回的玩家，或在多個儲存桶中快速移動的帳戶 採樣週期。最後，符合條件的遊戲在每個平均 Elo 內排名 儲存桶由遊戲 ID 和種子的確定性哈希以及上限數量組成 每個儲存桶中都會保留遊戲。

這種設計使基準曲線更加有用。每個 Elo 桶 提供足夠的觀察值來估計每個指標的平均值，罕見的高評級或低評級桶不會被自然形狀淹沒 活躍玩家分佈。Elo+Chess 使用的基準曲線是 然後限制在與產品最相關的範圍內，大致是初學者 透過早期的高級遊戲。在 Lichess 規模報告中，此主要範圍 是 600–1600。當使用者查看 Chess.com 規模的報告時，對應的評級桶映射到相同的底層 Lichess 使用單獨的跨平台評級映射方法進行基準儲存桶 配套論文中描述：[Estimating Cross-Platform Chess Rating Mappings with Modal Regression](#)。

3 活躍玩家評分分佈

如果我們看看截至最後 Lichess.org 玩家 Elo 評分的分佈 截至 2026 年 3 月，按主要遊戲類型，我們看到以下計數和百分位數。對於此分佈視圖，每個玩家按確切的遊戲類型計數一次，使用該玩家在相應原始數據中最新觀察到的評分 時間控制行，在該遊戲中至少需要五場比賽之後 型號。¹

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: 最新觀察到的活躍玩家 Lichess 評分分佈（以精確分佈）需要至少五場該確切遊戲類型的遊戲。

此分佈不應與分層基準樣本混淆。此分佈描述了在原始資料中觀察到的活躍 Lichess 玩家池掃描。然後故意對基準樣本進行分層，以便 Elo 桶 對穩定的度量曲線有足夠的支撐。

將特定遊戲類型的 Lichess.org Elo 評級對應到 等效 Chess.com 評級，Elo+Chess 使用遊戲類型特定模式 同伴中估計的迴歸方程 [Elo+Chess rating-mapping paper](#):

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

這些方程式映射的是評級值，而不是百分位數排名。百分位數排名是在特定玩家池中定義，以及 Chess.com 和 Lichess 玩家 水池有不同的形狀。在表 2 中，對於 10 分鐘快速遊戲類型，第一欄給出了可能的小樣本 Chess.com Elo 額定值。第二列給出了相應的 截至 2026 年 5 月 31 日，Chess.com 上顯示的 Chess.com 百分位排名。A 71.7

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 顯示有和沒有此過濾器的分佈。

Lichess.org 10 分鐘快速排名對應於 Lichess Elo 評級 1,644²，該值顯示在 第三欄。如果我們應用 Chess.com–Lichess 映射方程，則 1,644 Lichess 相當於 Chess.com 上的約 1,271。假設估計 映射方程式合理，+522與Chess.com評級之間的差距 第 1 列和第 4 列的映射評級可以用大量數字來解釋 Chess.com 玩家池中的休閒玩家讓 Chess.com 膨脹 百分位數排名。換句話說，在 Lichess.org 上處於第 72 個百分位數是 比 Chess.com 上的第 72 個百分位數更大的成就，因為 百分位等級不能清楚地從一個系統對應到另一個系統。實際Elo 可以使用根據同一玩家得分估計的方程式來映射值 兩個系統。

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: Chess.com 10分鐘快速百分位之間的說明性比較 積分與主動Lichess 10+0快速積分相同要求至少五场10+0比赛后分配。“映射的 Chess.com 等價。” 应用当前的 Elo+Chess 10+0 快速評級量表轉換為相同百分位數的 Lichess 評級。的 包括最高尾行以顯示尺度差異，但擬合的 評級線在主要教練範圍內是最可靠的。

表 2 中存在間隙的原因是映射 估計同等的演奏強度等級，而百分位數描述 在一個平台自身群體中的地位。71.7% 的玩家 因此，Chess.com 不會自動成為 Lichess 上第 71.7% 的玩家，反之亦然。觀察到的數字與Chess.com公開的一致 快速池擁有更大的低評級或休閒玩家群體，而 主動精確控制Lichess池更集中在經驗豐富的重複中 玩家。關鍵不在於一個百分位數是“正確的”，而且其他的是「錯誤」；這是百分位數排名不可跨平台移植 除非底層玩家分佈具有相同的形狀。

Note on the 1500 spike. The visible jump near 1500 存在於資料中，而不是透過平滑創建的。在最新的 觀察分數分佈，精確的 1500 評分佔 1+0 的 13.1

子彈型玩家、3+0 閃電戰型玩家中 5.6

17.6

評級系統中的起始或臨時錨點。

確切的 1500 個高峰很大程度上是由輕度活躍帳戶造成的。至 檢查這一點，我們在要求後重新計算了相同的最新評級分佈 每位玩家至少有五場相同遊戲類型的遊戲。尖峰幾乎 在此過濾器下消失：確切的 1500 個評級從 5.6–17.6

在五場或以上比賽中，所有球員分佈中的球員達到 0.2–1.2

在表 1 中報告了分佈。

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.

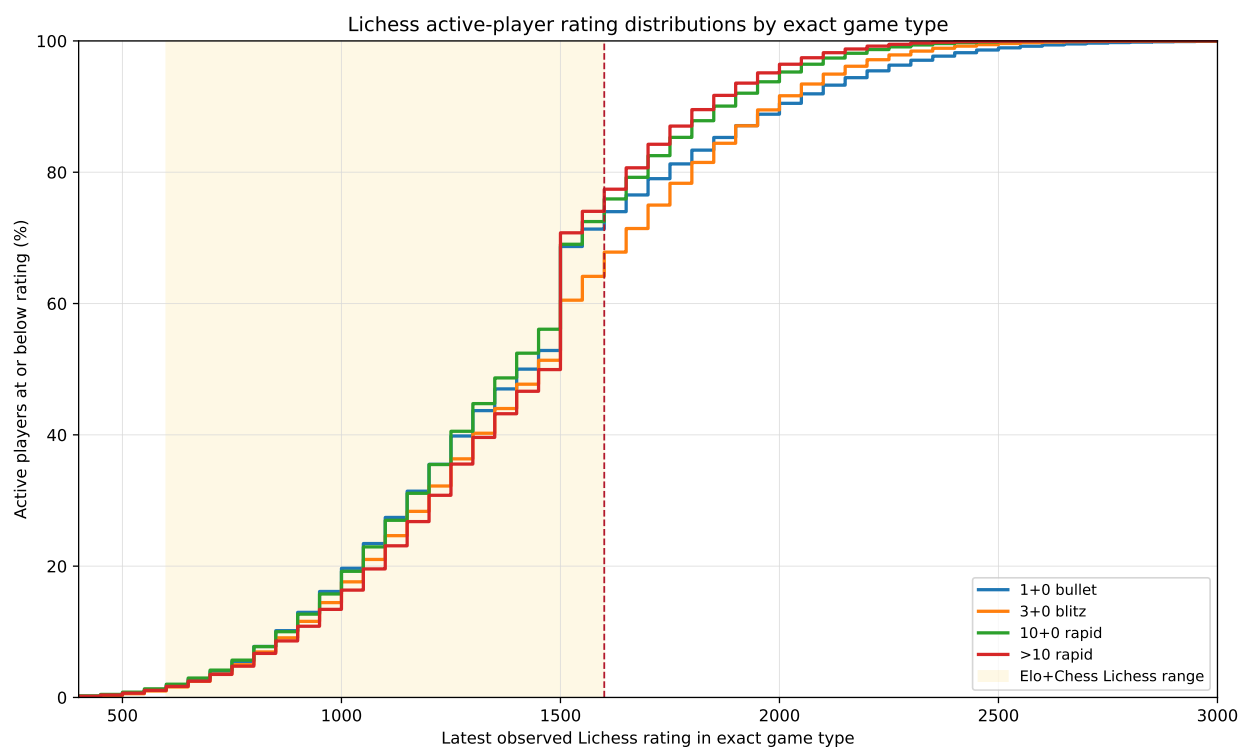


Figure 1: 以確切遊戲類型劃分的累積活躍玩家分數分佈。的 黃色帶標記了主要Elo+Chess使用的 600–1600 Lichess範圍 教練基準。

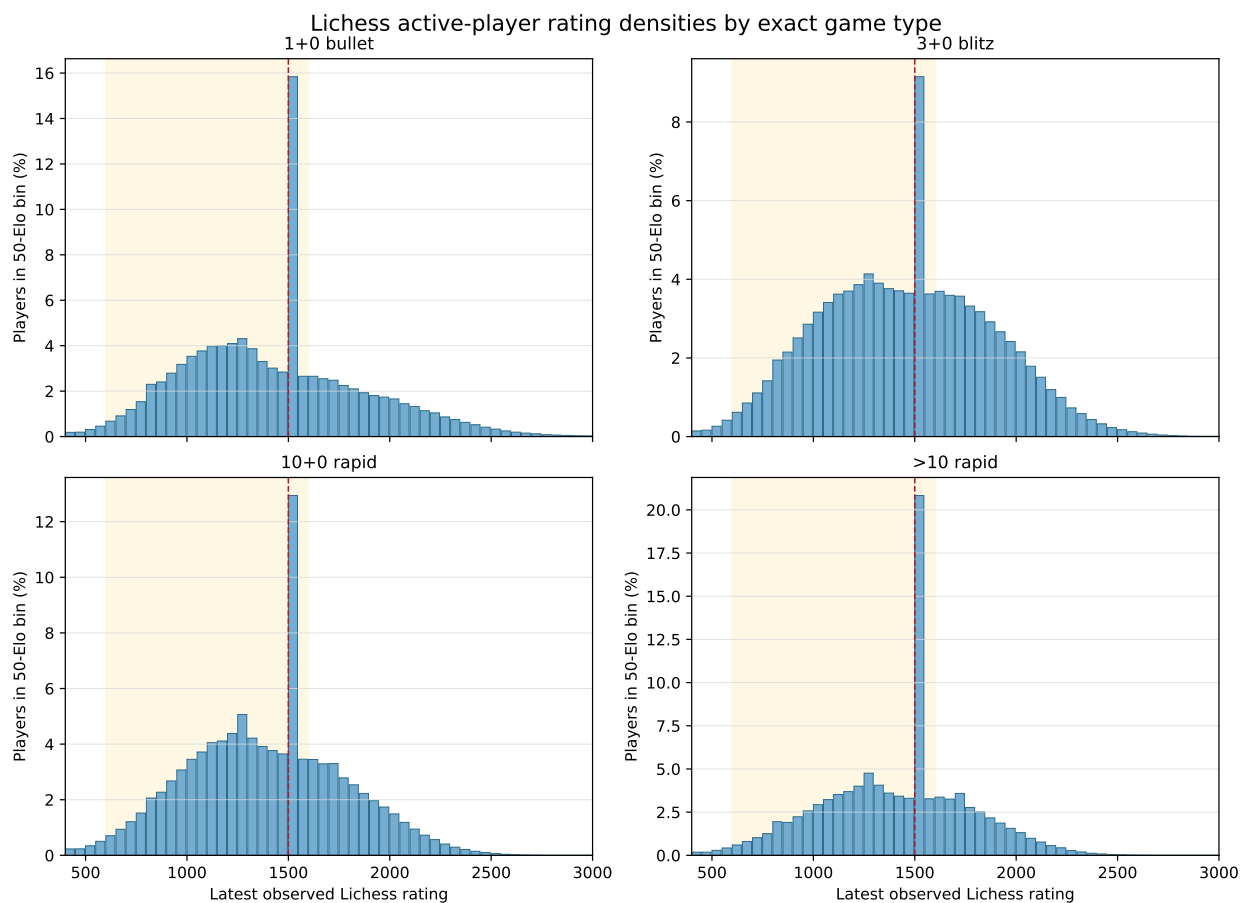


Figure 2: 相同活躍玩家評分分佈的密度視圖。的 虛線標記 1500 Lichess，其中原始最新評級分佈一個大的精確額定質量點。黃色帶標記600–1600 Lichess 主要 Elo+Chess 指標報告使用的範圍。峰值最明顯的是 所有玩家的密度，因為許多輕度活躍帳戶仍然保持在 1500；圖 4重複密度圖 在要求至少五場特定遊戲類型的遊戲之後，這在很大程度上消除了 這個神器。

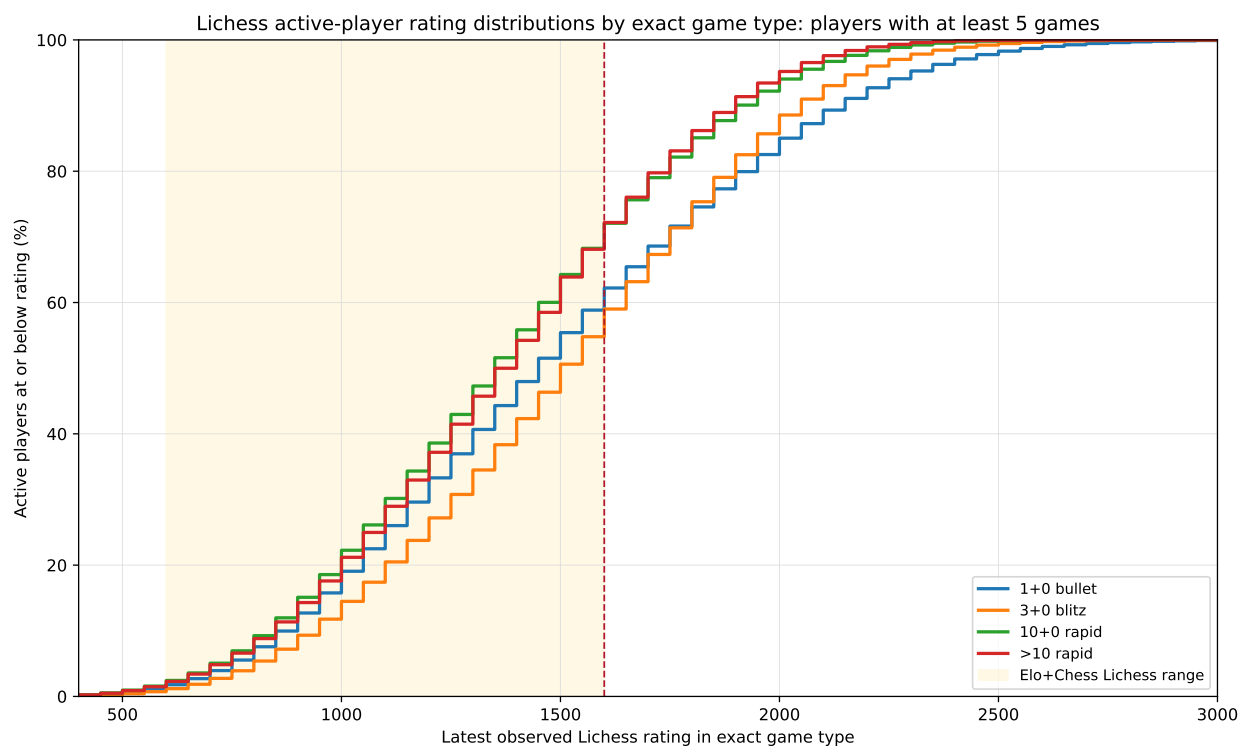


Figure 3: 要求後的累積活躍玩家分數分佈 確切遊戲類型中至少有五場比賽。 1500跳大大減少。

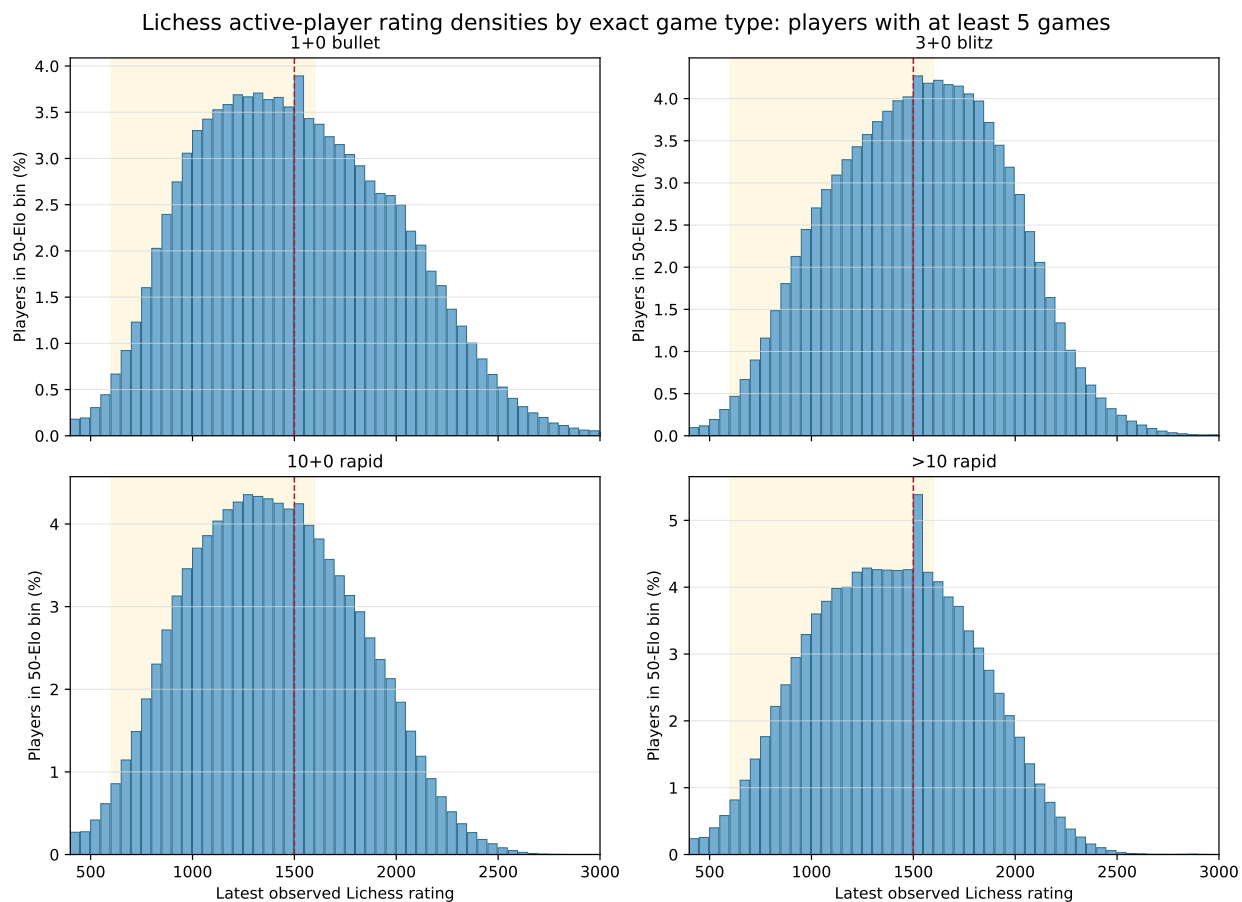


Figure 4: 要求在同一場比賽中至少進行五場比賽後的密度視圖 類型。該視圖證實了全玩家中的大 1500 個質點 分配主要是由活躍度較低的帳戶驅動的。黃色帶標記 主要 Elo+Chess 指標報告使用的 600–1600 Lichess 範圍。

4 目前基準工件的規模

計算量故意很大。這不是輪廓刮傷，而是一個小一組下載的 PGN，或手工策劃的指導性範例集合。初始 Lichess 通行證涵蓋 2025 年 1 月起的完整月度文件 截至 2026 年 3 月。非分層掃描表包含 1,382,178,087 場比賽 以及 2,764,356,174 個玩家遊戲評分行。在這些行中，有 6,623,438 個唯一 玩家用戶名至少出現一次。

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: 基準採樣之前未分層的 Lichess 掃描比例。掃描 涵蓋 2025 年 1 月至 3 月完整的每月遊戲歷史文件 2026 年。

生產基準測試工件比原始掃描小得多，因為 它們被選用於公制結構。它們仍然很大：基準 stage 是一個超過數百個的移動狀態擴展和聚合管道 每種遊戲類型有數千個遊戲。

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: 對齊後分層基準工件的規模 常見的 277 列運行時特徵模式。一個「玩家-遊戲行」是指從一個玩家的角度來看一場遊戲，因此大多數遊戲 產生兩個玩家遊戲行。

僅就 3+0 閃電戰而言，目前的生產基準就使用了 479,974 場比賽 涉及 129,398 個獨特的玩家和 28,138,904 個不同的移動狀態行。的 對齊的閃電戰玩家遊戲特徵表包含 277 個衍生變數。的快速功能建構工件在移動狀態階段甚至更廣泛：完整的 特徵建構期間使用的快速移動狀態表包含 168 聚合成 277 列玩家遊戲功能之前的移動狀態列表。所有四種遊戲類型的前端運行時基準工件是 然後與相同的 277 列玩家遊戲特徵模式對齊，包括 擴展的戰術分類突破變數。從這些特徵表中，報告選擇並顯示最容易解釋的輔導指標。

這些計數對於解釋結果很重要。基準曲線 不是根據幾百個人工審核的遊戲來估計的。相反，他們 來自數以千萬計的評估委員會國家和超過四個 目前四個製作遊戲類別中的百萬玩家遊戲行。

5 管道概覽

標竿建設流程有六個步驟：

1. 在目標遊戲類型和評級層中選擇評級為Lichess的遊戲。
2. 逐步解析遊戲的每一步棋並儲存棋步前和棋步後的資訊 董事會狀態。
3. Calculate intermediate move-state variables from the board, the move, 玩家視角和對手視角。

4. 將移動狀態變數聚合為玩家遊戲特徵變數。
5. 將玩家遊戲變數轉換為清晰的報告指標 解釋方向：通常越高越好，通常越低 更好，或描述性的。
6. 按 Elo 儲存桶平均每個指標並手動檢查結果 關係。

這種結構很重要，因為許多報告指標並不直接 在 PGN 標頭中可見。他們需要重建位置，確定哪些棋子從哪些起始方格移動，比較移動前和 移動後的材料和結構，然後總結結果 將等級移至玩家遊戲等級。

6 中間變數

中間變數旨在描述板上發生的變化 以及為什麼從報道者的角度來看它很重要。範例包括：

- 移動編號、玩家移動編號、移動方向、移動顏色和玩家 顏色；
- 從方格、到方格、移動棋子類型、捕獲棋子類型以及 此舉是否為佔領、過牌、城堡、晉升或交配；
- before-move and after-move material balance;
- whether a minor piece left its starting square by a given move number;
- whether the same piece moved on consecutive player moves in the 開幕；
- 皇后是否在第 10 步之前移動；
- 易車側、易車移動次數、車輛是否已連接 易位後；
- 鬆散棋子、懸掛棋子、雙棋子、孤立棋子、通過 兵和後兵；
- rook 檔案狀態、開啟檔案、半開啟檔案和 king-safety 特徵；
- 戰術機會、執行戰術、物質壓力戰術、 guaranteed follow-up gains, and realized gains.

確切的中間集因建置階段而異。運行時報告資料庫保存 只有快速提供報告所需的列，而完整的功能構建 工件包括更廣泛的移動狀態表和用於建置的功能補丁 最終的玩家遊戲特徵行。

7 開放原則測量範例

開放原則是有用的例子，因為它們很容易讓使用者理解 理解並易於連接到移動狀態計算。

7.1 Minor Pieces Developed by Move 10

此指標計算玩家的四個騎士和主教中還剩下多少個 當玩家第 10 步移動時，他們的主場方格。中間計算 使用移動者顏色、移動棋子類型、起始方格和玩家移動 數量。對於白方，本方格為 b1、g1、c1 和 f1。對黑方來說，他們是 b8、g8、c8 和 f8。對於每個玩家遊戲，Elo+Chess 計算不同的主場 由玩家移動 10 移動的騎士和主教之間的方格。

這是一個通常越高越好的指標。它並不聲稱每個未成年人 棋子必須始終以第 10 步移動，但在發展中的玩家中，經驗值 關係很明確：達到更高 Elo 等級的玩家往往會發展 更早的小作品。

7.2 Repeated Piece Moves in the Opening

該指標計算了再次移動同一棋子所花費的額外早期移動 开发新东西。中間計算使用之前的 目的地方格、目前起始方格和玩家移動編號。如果 前一個玩家移動時移動的同一棋子，在該棋子內再次移動 玩家在前 10 步移動後，重複移動次數會增加。

這是一個通常越低越好的指標。它測量浪費的打開速度 簡單的方法。一些重複的舉動在戰術上是合理的，但從規模上看 桶曲線捕捉了總體趨勢：目標上更強的玩家 關卡早期移動重複移動同一塊的時間較少。

7.3 Queen Touched Before Move 10

此指標計算玩家移動 10 之前的皇后移動。計算使用移動棋子類型和玩家移動數量。如果移動件是皇后且玩家的移動次數最多為 10 次，則計數增加。

對於 Elo+Chess 目標範圍來說，這通常也是越低越好。早期 皇后動作可以贏得物質或力量讓步，但許多發展中的玩家 太早帶出女王，失去攻擊女王的節奏，並拖延 小件開發，安全為王。

7.4 Castling and King Safety

易位指標使用易位標志、易位方位、易位步數和 易位後的棋盤狀態。例子包括玩家是否城堡，玩家是否通過第 10 步攻勢，無論玩家在王翼還是攻勢 後側，車易位後車是否連接，以及車是否文件易位後呈開放式或半開放式。

這些指標的解釋方式並不相同。易位完成及 步驟 10 的易位通常在目標範圍內越高，通常越好。後側易位或易位至開放文件可能更依賴上下文。這就是為什麼 Elo+Chess 不是簡單地硬編碼道德規則；而是簡單地硬編碼道德規則。它建造了水桶 曲線並檢查哪些關係對於指導來說足夠穩定。

8 鏟鬥曲線和手動檢查

計算玩家遊戲指標後，Elo+Chess 以 Lichess Elo 將它們分組 桶併計算桶平均值。桶平均值是原始經驗值 該指標的基準點。然後在報告中使用趨勢線來 將 Elo 風格的等級指派給使用者值。

曲線在使用前經過手動檢查。指標更有用 當桶均值在整個評級範圍內顯示出一致的關係時 由產品服務。當一個指標有噪音、平坦、高度集中時，它的用處就較小。非單調的，或由難以向使用者解釋的特殊情況所驅動的。

這個手動檢查步驟在 Elo+Chess 鞋面上方尤其重要 截止。更強的玩家知道何時打破開局原則並這樣做 有效地。在初級到早期高級範圍之上，關係 簡單的習慣和評級之間經常會變平或改變形狀。產品 因此強調經驗關係較強的範圍 並且在教學上有用。

9 全範圍開度曲線範例

Figure \reffig:opening-curves-blitz3 shows four 3+0 blitz opening-principle curves. Gray points show the wider Lichess rating range. Blue points and the blue trend line show the 600–1600 Lichess range used by the main Elo+Chess coaching benchmarks. The dashed red line marks the upper cutoff.

同樣的模式在全範圍 10+0 快速和更長的快速中可見 文物。中心建模決策不變：使用經驗 教練主張的穩定範圍，避免假裝簡單的習慣曲線 永遠保持同樣的訊息。

表 5 中的數字也出現了相同的模式。坡度是 報告為每 100 個 Elo 點的指標值變化。易位，小件發展、重複棋子運動和早期皇后運動都顯示清晰 運動跨越 600–1600。高於 1600 時，這些區域的

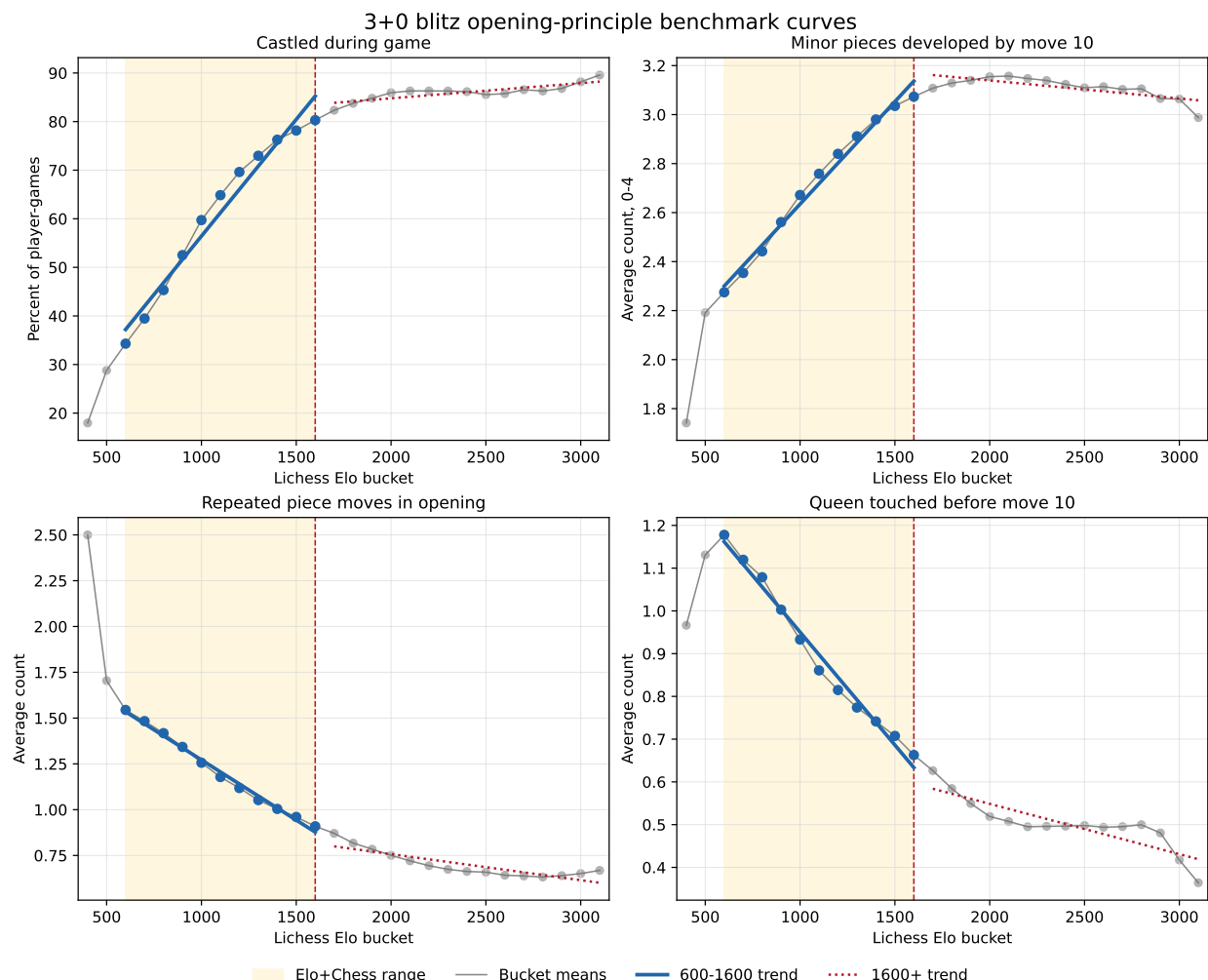


Figure 5: 3+0閃電戰精選開放原理基準曲線。的 在 Elo+Chess 教練範圍內，關係是最強、最有用的。高於上限時，有些關係會大幅趨於平緩。

坡度要小得多 例子，這與評價較高的玩家打破的想法是一致的 更有選擇性、更成功的基本原則。第二個機制是 飽和度。對於某些習慣，該指標達到了實際上限或甜蜜 現貨：強者已經養成習慣了，沒理由 曲線繼續急劇向上移動。更多並不總是更好 那一點。因此，數據會趨於平緩，因為更強的玩家可以打破 規則在具體的位置上，因為很多基本的習慣已經達到了 他們的有用操作範圍。

10 為什麼上限很重要

Elo+Chess的目標使用者是初級到初級高級玩家。對於這些 玩家的基本習慣與評級密切相關：發展棋子，避免重複的早期棋子移動，避免不必要的早期皇后移動，城堡 及時連接車輛、管理材料並避免留下鬆散的零件或 懸掛。

在較高的收視率下，那些相同的表面習慣變得不那麼診斷。一個強大的 由於中心關閉，玩家可能會延遲易位，請儘早移動王后 因為該位置包含具體的戰術，或將同一棋子移動兩次 因為它贏得了節奏或迫使結構性讓步。原始習慣依然存在 有意義，但是習慣和評級之間的關係變得更加重要 有條件的。這就是為什麼 Elo+Chess 使用上限作為主要訓練的原因 聲明以及為什麼高於該範圍的

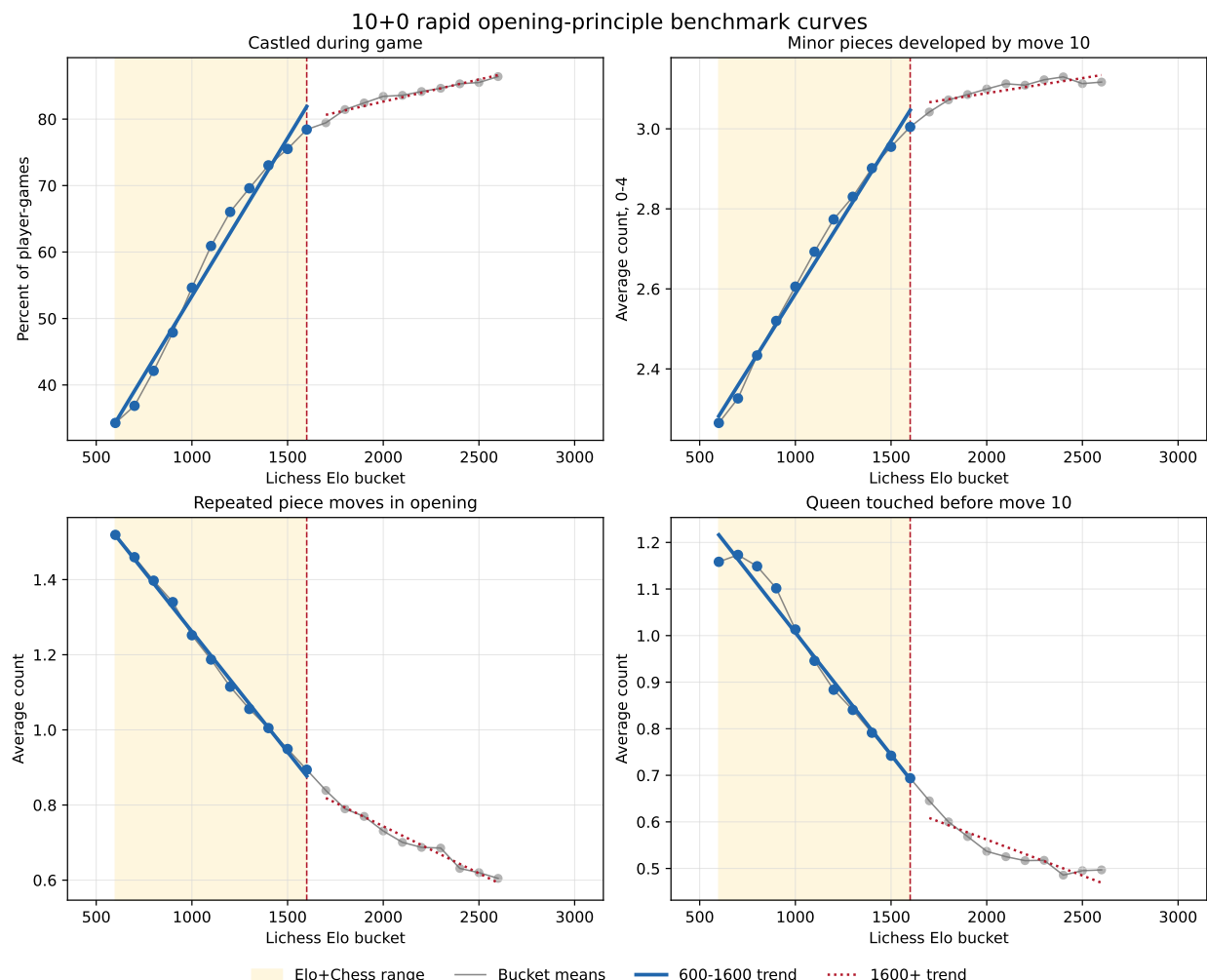


Figure 6: 10+0快速遊戲的開局原理基準曲線精選。的 在 Elo+Chess 教練範圍內，關係是最強、最有用的。高於上限時，有些關係會大幅趨於平緩。

用戶會被警告緊密關係 原則習慣和評級之間的關係崩潰了。

11 雙邊戰術指標與百分位目標

某些指標的形狀與開放原則範例相反。叉子 是一個有用的例子，因為事件本質上是雙向的。一位球員 發現並轉換比同行更有用的分叉是在做好事，但是 轉換後的叉子的數量隨著評級的增加而下降，因為更強 對手允許更少的分叉機會並防禦戰術威脅 更积极地。

圖 8展示了這個效果。兩個改裝貨叉均由 報告玩家和對手轉換的叉子變得不那麼頻繁，因為 收視率上升。這並不意味著執行分叉是不好的。這意味著原始事件 比率部分由對手錯誤率和遊戲紋理控制。對於指標 像這樣，Elo+Chess 因此在 玩家自己的 Elo 桶，而不是僅依賴全球桶平均趨勢。

為了建立這些百分位數目標，Elo+Chess 按遊戲類型對玩家進行採樣，Elo 桶，然後使用每個樣本球員最近的資格賽。的 目前的目標集使用 11 個桶，從 600 到 1600，每個桶 200 名玩家 桶，以及每個球員最近最多 50 場資格賽 遊戲類型。這為每種遊戲類型提供了 2,200 個玩家樣本，大約

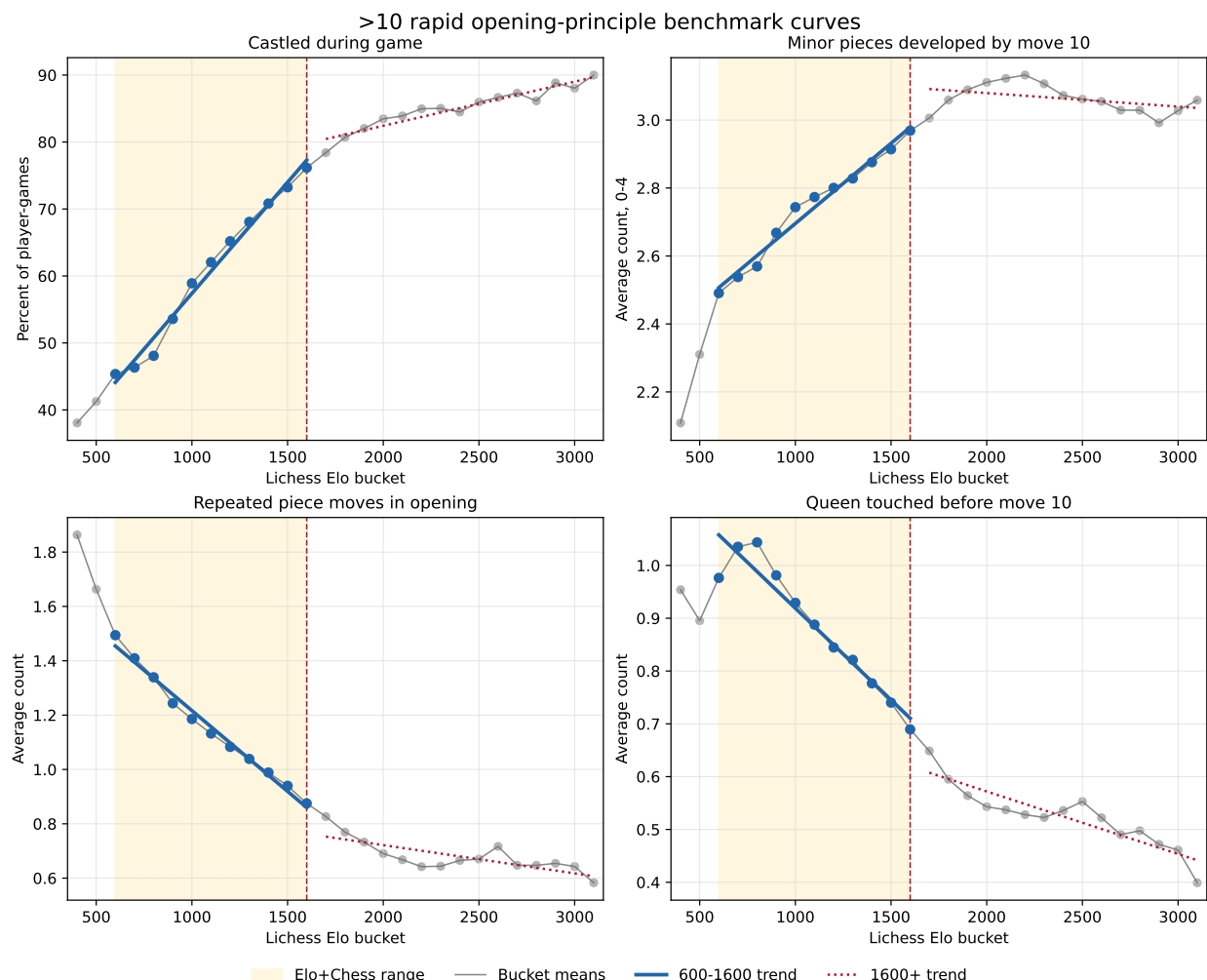


Figure 7: 為較長的快速遊戲選擇的開局原理基準曲線。

每個遊戲類型的桶內百分位有 110,000 個玩家遊戲行 分佈。

圖 9 說明了最終的目標 3+0 閃電戰中取樣玩家轉換分叉率的分佈。每個盒子都是 樣本玩家的桶內分佈，而不是桶內的平均值 個人比賽。在報告 UI 中，將使用者的價值與相關值進行比較 使用者映射的 Elo 儲存桶的框。這是諸如此類的陳述的基礎 相對於同儕來說，玩家的分叉轉換率是高還是低 相同的水平，即使總體人口事件發生率隨評級而下降。

12 複製成本

该方法原则上是可重复的，但不是轻量级的。重建 基准曲线要求：

- 存取大型 Lichess 每月遊戲歷史檔案；
- 按遊戲類型、評級範圍和資格進行分層過濾 規則；
- 能夠擴展的合法棋步解析器和局態評估器 數以千萬計的移動狀態；
- 開局、材料、當鋪結構、國王安全、車子的特徵邏輯 活動、戰術、時間管理和比賽結果；

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: 四個範例開口原理的近似桶曲線斜率 跨製作遊戲類型的指標。

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: 目前同業歷史百分位數目標集。這些是 用於估計桶內分佈的玩家歷史樣本 百分位式指標。

- 從移動狀態到玩家遊戲特徵的聚合；
- 每個報告指標的桶級摘要；
- manual review of every candidate benchmark curve before it is used in 教練報告。

For 3+0 blitz, this means nearly half a million games, almost one million player-game rows, and more than 28 million move-state rows. Across the four benchmark game types, the current benchmark artifacts contain more than 2.55 million games, more than 5.10 million player-game rows, and more than 147 million move-state rows. The scale is central to the value of the benchmark: it lets the report show users what actually happens in games played by players near their level rather than relying only on general chess slogans.

13 限制

基準曲線是經驗描述，而不是因果證明。一名球員 僅僅透過匹配桶平均值並不能獲得評級。指標也是 複雜局面的不完美總結：早期的皇后棋步可能會非常出色 在一種立場上而在另一種立場上魯莽。曲線的目的是 識別反覆出現的習慣並將其與同伴行為進行比較，而不是為了 取代計算或指導判斷。

這些曲線也是 Lichess 比例曲線。當用戶攜帶Chess.com遊戲時 歷史記錄，Elo+Chess 將相關評級桶對應到 Lichess 基準 使用單獨的跨平台映射模型進行縮放。此映射層是 在 Lichess-

Two-sided converted-fork rates by Elo bucket

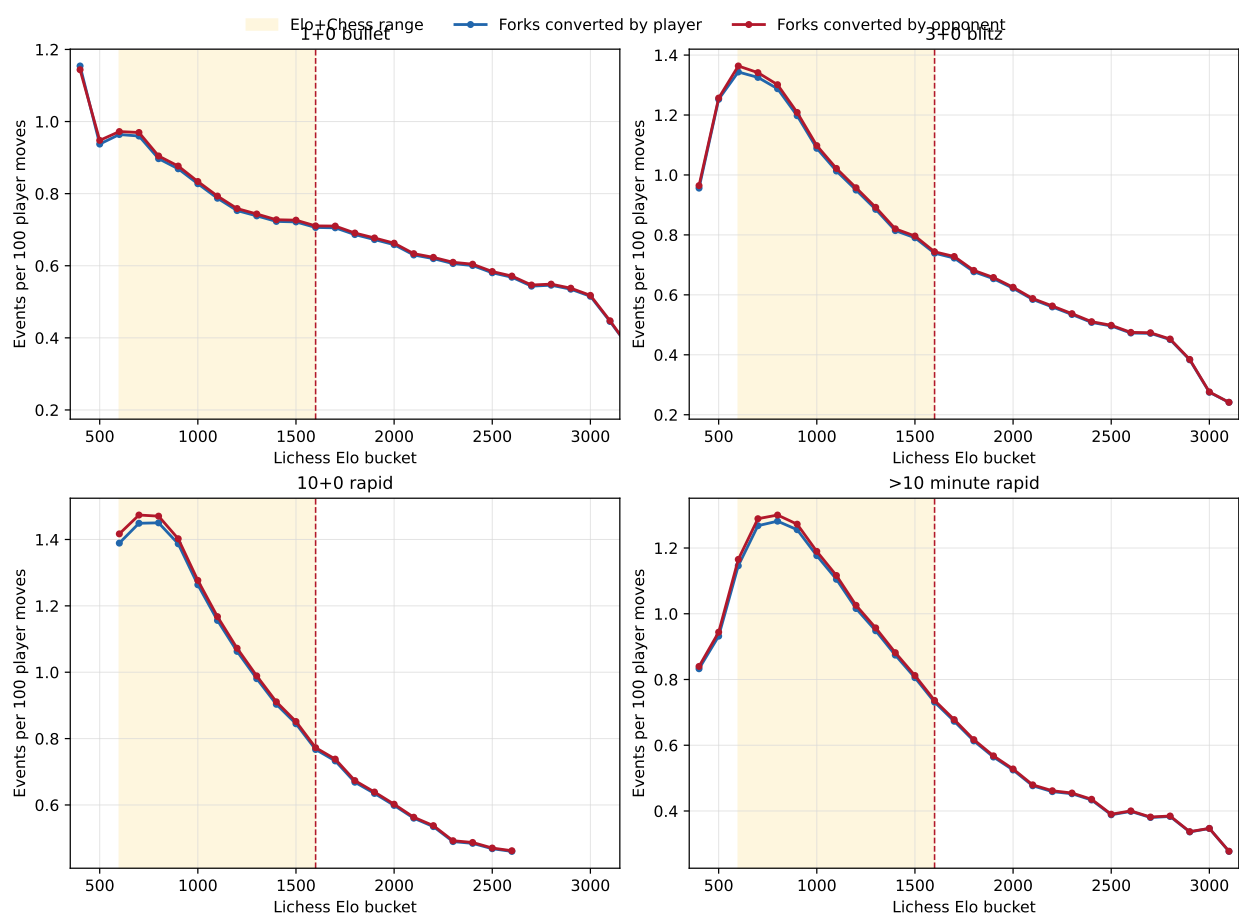


Figure 8: Elo 桶的兩側轉換叉率。轉換後的分岔事件 由於雙方防守更好且乾淨叉子更少，收視率下降 可以進行轉換。

to-Chess.com 的配套 Elo+Chess 研究說明中進行了描述 評級轉換：[Estimating Cross-Platform Chess Rating Mappings with 模態回歸](#)。此映射不应被理解为保留百分位 排名。它保留了兩個評級之間的估計評級等效性 系統；百分位數排名仍然是特定於平台的，因為活躍玩家 分佈不同。

14 生產報告圖表

生產報告將基準工件轉變為指標層級的面板 無需了解完整的建置管道即可閱讀。人物 ~ 10 and 11 顯示來自即時 Elo+Chess 報告設計的三個範例。

圖 10 開啟原理面板 使用普通的基準曲線演示。每張卡都以 度量標準系列和度量標準名稱，然後給出簡短的解釋規則。的 右上角的大等級是故意與原始的分開的 指標值：它回答了用戶面臨的問題“這個習慣如何評分”與基準相比？”，而下面的行保留測量值。四個摘要行顯示「一生」、「過去 10 場」、「勝利」和「失敗」。這使得 面板診斷而不僅僅是裝飾。如果 Lifetime 和 Last 10 一致，這個習慣可能是穩定的。如果輸贏不同，衡量標準可能是 取決於結果或與遊戲狀態相關，而不是簡單的習慣。

圖表區域使用內斂的深色背景、金色趨勢線和彩色 四個切片的標記。趨勢註釋明確指出是否 基準關係隨 Elo 上升或下降。這很重要，因為 有些指標越高越好，例如小件開發，而 其他則是越

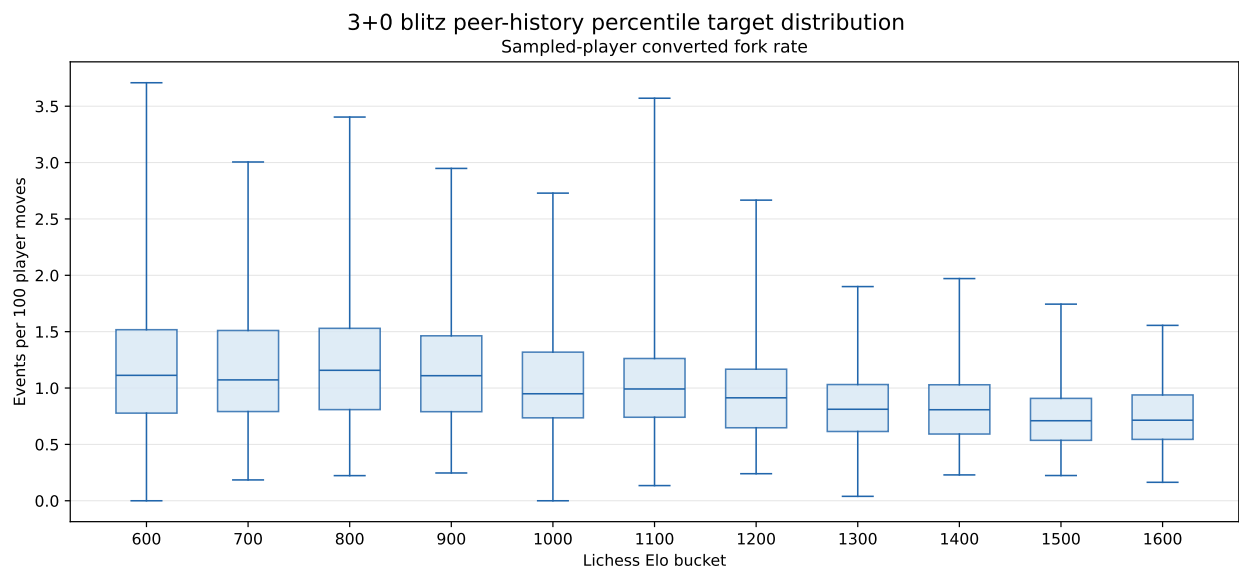


Figure 9: 樣本玩家的 3+0 閃電戰同行歷史目標分佈範例 轉換後的分叉率。百分位數是在 Elo 桶內計算的，因此報告 將球員與面臨大致相似戰術機會的同儕進行比較 費率。

低越好，例如早期的皇后運動。因此該卡 不要求使用者單獨從坡度推斷方向。決賽「這意味著什麼」和「指導提示」行翻譯了基準結果 轉化為具體的國際象棋教學，同時保留經驗背景。

圖 11 顯示百分位式 用於原始率不能很好地表示的指標的表示 單一的全球趨勢。獲勝分叉是一個戰術範例：人口比率 轉換後的叉子數量會隨著玩家技能和對手錯誤率的變化而變化。的 因此，卡片將用戶的價值與同行玩家的歷史記錄進行比較 映射的 Elo 桶。箱線圖顯示了對等分佈，包括四分位距、中位數和須線。使用者的標記繪製在 相同的尺度並標有原始值和百分位等級。相鄰 比較桶以靜音形式保持可見，因此使用者可以看到 對等儲存桶是本地的，而不是通用規模。

這種設計使百分位數指標可審計。用戶看到原始值， 百分位數、用於比較的對等儲存桶以及翻譯後的值 Chess.com 到 Lichess 儲存桶映射。相同的一生、最後 10 場勝利、以及 保留損失行，因此仍然可以檢查戰術百分位數分數 近因效應和結果依賴性。

15 結論

Elo+Chess 基準報告是根據大量移動狀態計算建立的 分層 Lichess 樣本。每個指標都以一組具體的板狀態開始 和移動狀態變量，成為玩家遊戲特徵，然後進行平均 透過 Elo 桶形成經驗基準曲線。開放原則 例子說明了為什麼這種方法有用：簡單的西洋棋習慣具有很強的， 與目標範圍內的評級有明顯的關係，而那些 高於上限的關係通常會趨於平緩或變得更有條件。這就是 Elo+Chess 將其教練主張集中在評級上的核心原因 數據最清楚地支持它們的範圍。

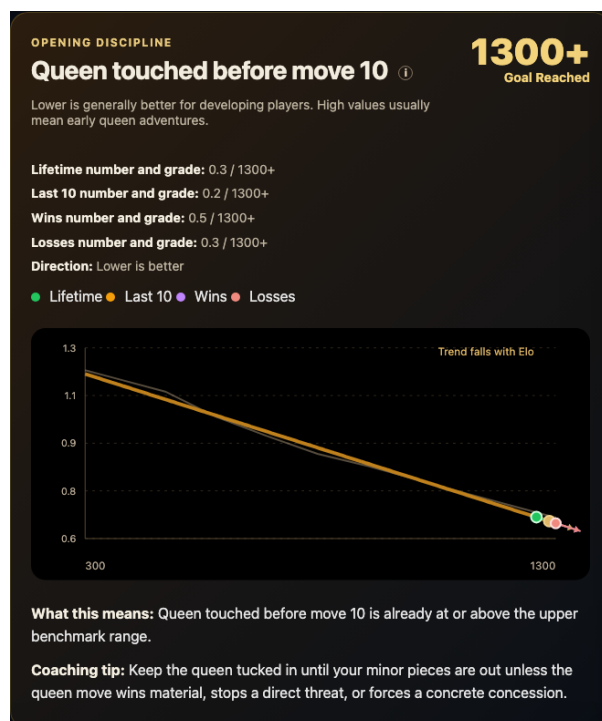
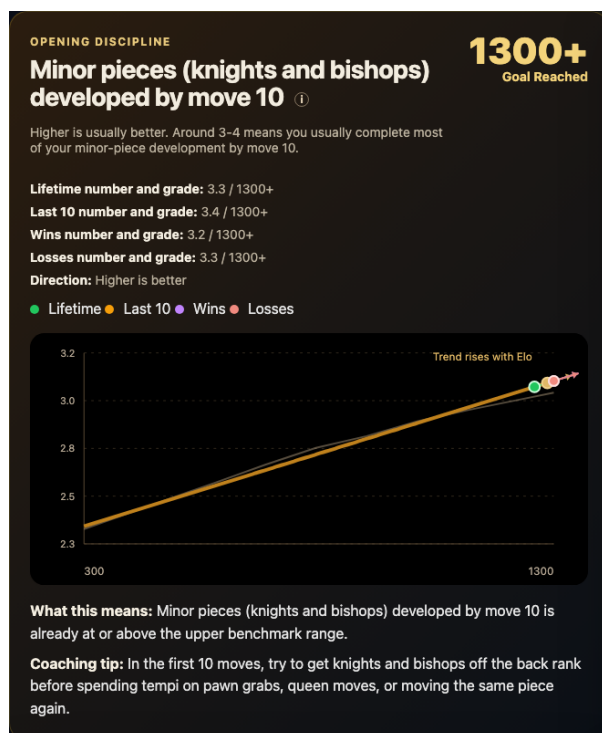


Figure 10: 生產開口原理面板。相同的視覺語法處理 越高越好的指標、小件開發和越低越好的指標，早期的女王運動。兩個面板都展示了一生、過去 10 場、勝利和失敗 因此報告可以將穩定的習慣與最近的或依賴結果的習慣區分開來 模式。

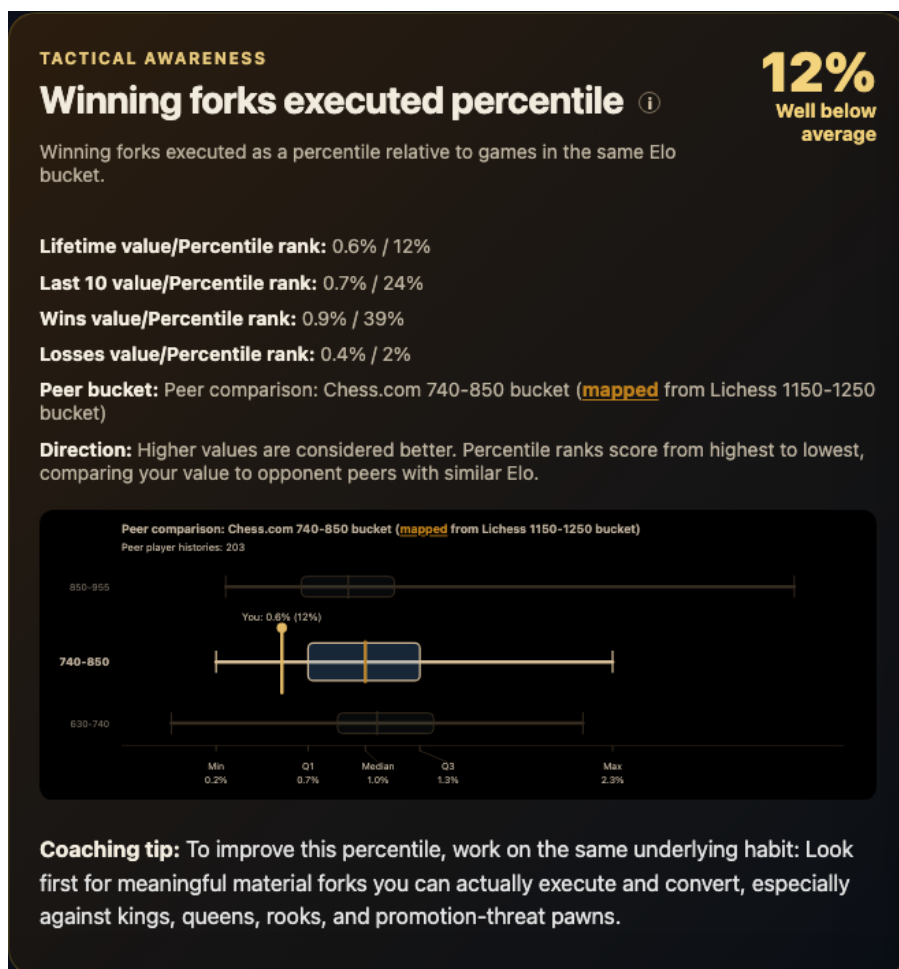


Figure 11: 執行獲勝叉子的生產百分位面板。箱線圖 顯示了報告使用的同桶同行分佈，而黃金標記顯示使用者的原始值和百分位數排名。