

Construction de courbes de référence stratifiées Lichess pour Elo+Chess

Notes de recherche Elo+Chess

31 mai 2026

Abstract

Les rapports de référence Elo+Chess sont construits à partir de grands échantillons stratifiés de Jeux Lichess.org, avec l'échantillon de référence utilisé dans cette note contenant 483 974 jeux bullet 1+0, 479 974 jeux blitz 3+0, 925 636 jeux rapides 10+0, et 660 540 parties rapides plus longues. Pour chaque type de jeu majeur, le pipeline s'étend chaque jeu dans les états du plateau avant et après le coup, calcule des centaines de variables intermédiaires et de jeu du joueur, regroupe ces variables en métriques de coaching, puis fait la moyenne de ces métriques par compartiment Lichess Elo. Le l'échantillon est délibérément stratifié plutôt que proportionnel au joueur brut pool : l'objectif est d'estimer des courbes de comportement stables à chaque niveau de notation, y compris des seaux qui seraient autrement clairsemés, pour ne pas reproduire le naturel répartition des notes. Le résultat est une famille de courbes de référence empiriques : pour Pour chaque métrique, nous pouvons nous demander à quelle fréquence les joueurs de différents niveaux Elo affichent un niveau donné. habitude ou principe dans les jeux réels. Cette note décrit la méthode, l'échelle de le calcul, et plusieurs exemples de principes d'ouverture impliquant le roque, le développement de pièces mineures, mouvements répétés de la pièce d'ouverture et mouvement précoce de la reine. L'accent est mis méthodologique : Elo+Chess ne suppose pas que les principes des échecs soient universellement linéaire à tous les niveaux de compétence. Au lieu de cela, les courbes de référence sont inspectées par plage de notation, et le produit se concentre sur la région débutant à avancé où les relations sont les plus fortes et les plus interprétables.

1 But

L'objectif du système de référence est de convertir des conseils généraux sur les échecs en quantités mesurables. Une déclaration telle que « développez vos pièces mineures » est utile, mais cela devient plus exploitable lorsque nous pouvons dire combien de pièces mineures un le joueur se développe généralement au coup 10, comment cela se compare-t-il à ses pairs du même seau de notation et comment cette habitude change à travers le spectre de notation.

Elo+Chess construit donc une bibliothèque de métriques à partir de jeux réels. Chaque métrique est calculé pour chaque ligne de jeu de joueur, résumé par type de jeu et compartiment Elo, et utilisée comme courbe de référence dans le rapport. Le rapport peut ensuite placer un utilisateur sur la même courbe et montre si l'utilisateur est au-dessus, en dessous ou à proximité du comportement observé chez des acteurs comparables.

2 Données sources et stratification

La source de référence est la base de données publique de jeux Lichess.org. Elo+Chess utilise de grands échantillons stratifiés par type de jeu et par catégorie de notation plutôt que par un petit échantillon de commodité. Les artefacts d'exécution de production actuels couvrent quatre principaux catégories de jeux :

- 1+0 balle,
- 3+0 blitz,
- 10+0 rapide,
- jeux rapides de plus de 10 minutes.

L'objectif de stratification s'étend à tous les niveaux de notation. Un hasard proportionnel Un échantillon serait efficace pour décrire le pool global de joueurs, mais il être inefficace pour construire des courbes de référence. Les catégories de notation les plus courantes domineraient l'échantillon, tandis que les godets de queue à faible support auraient trop peu de jeux pour des moyennes métriques stables. Elo+Chess donc des échantillons par jeu moyen Seau Elo. Dans la construction actuelle d'échantillons propres, les jeux bruts sont les premiers filtré pour des jeux classés valides avec des joueurs connus, des notes connues, des résultats valides, et des contrôles horaires connus. Les jeux sont ensuite filtrés pour les joueurs ayant suffisamment de jeux dans cette vitesse et des notes suffisamment stables pour représenter un niveau de compétence cohérent : pour chaque joueur, Elo+Chess calcule l'intervalle de vitesse du joueur ($\max(\text{Elo}) - \min(\text{Elo})$) dans le pool source et conserve jeux uniquement lorsque les deux joueurs ont au moins le nombre minimum de jeux configuré et pas plus de 300 points Elo de dérive en vitesse. Le filtre à dérive n'est pas destiné pour supprimer une amélioration ou un écart normal ; cela réduit les cas où un joueur les jeux mélangerait des niveaux de compétence sensiblement différents, tels que des comptes provisoires, les joueurs qui reviennent ou les comptes qui ont évolué rapidement à travers plusieurs compartiments au cours la période échantillonnée. Enfin, les jeux éligibles sont classés au sein de chaque moyenne-Elo bucket par un hachage déterministe de l'identifiant et de la graine du jeu, et un nombre plafonné de Les jeux sont conservés dans chaque compartiment.

Cette conception rend les courbes de référence beaucoup plus utiles. Chaque compartiment Elo apporte suffisamment d'observations pour estimer la valeur moyenne de chaque métrique, et les rares seaux de qualité supérieure ou inférieure ne sont pas noyés dans la forme naturelle. de la répartition des joueurs actifs. Les courbes de référence utilisées par Elo+Chess sont puis restreint à la gamme la plus pertinente pour le produit, grossièrement débutant grâce à un jeu avancé précocce. Dans les rapports à l'échelle Lichess, cette plage principale est de 600 à 1600. Lorsqu'un utilisateur consulte le rapport sur l'échelle Chess.com, le les tranches de notation correspondantes sont mappées sur le même Lichess sous-jacent catégories de référence en utilisant la méthode distincte de cartographie des notations multiplateformes décrit dans le document d'accompagnement : [Estimating Cross-Platform Chess Rating Mappings with Régression modale](#).

3 Distributions des notes des joueurs actifs

Si l'on regarde la répartition des notes Elo des joueurs Lichess.org à la fin de mars 2026, par type de jeu majeur, nous voyons les décomptes et percentiles suivants. Pour cette vue de répartition,

chaque joueur est compté une fois par type de jeu exact, en utilisant la dernière note observée de ce joueur dans le brut correspondant lignes de contrôle du temps, après avoir nécessité au moins cinq jeux dans ce jeu exact tapez.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: Dernière répartition des notes Lichess des joueurs actifs observée par exact type de jeu après avoir requis au moins cinq parties dans ce type de jeu exact.

Cette distribution ne doit pas être confondue avec l'échantillon de référence stratifié. La distribution décrit le pool de joueurs Lichess actif observé dans le brut numériser. L'échantillon de référence est ensuite délibérément stratifié afin que les catégories Elo avoir suffisamment de support pour des courbes métriques stables.

Pour mapper une note Elo Lichess.org pour un type de jeu particulier vers un Note Chess.com équivalente, Elo+Chess utilise le modal spécifique au type de jeu équations de régression estimées dans le compagnon [Elo+Chess rating-mapping paper](#):

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

Ces équations cartographient les valeurs de notation, et non les rangs centiles. Les rangs centiles sont défini à l'intérieur d'un pool de joueurs particulier, et les joueurs Chess.com et Lichess les piscines ont des formes différentes. Dans le tableau 2, pour le Type de jeu rapide de 10 minutes, la première colonne donne un petit échantillon des possibilités Valeurs de notation Elo Chess.com. La deuxième colonne donne le correspondant Rang centile Chess.com affiché sur Chess.com au 31 mai 2026. A 71,7% Le classement sur Lichess.org pour un rapide de 10 minutes correspond à une note Elo Lichess de 1,644² et cette valeur est affichée dans le troisième colonne. Si nous appliquons l'équation de cartographie Chess.com–Lichess, 1 644 sur Lichess équivaut à environ 1 271 sur Chess.com. En supposant que l'estimation L'équation de cartographie est raisonnable, l'écart +522 entre la note Chess.com dans la colonne 1 et la note cartographiée dans la colonne 4 peuvent s'expliquer par un grand nombre de joueurs occasionnels dans le pool de joueurs Chess.com gonflant le Chess.com rang centile. En d'autres termes, être dans le 72e centile sur Lichess.org est un accomplissement plus important que d'être dans le 72e centile sur Chess.com, car les rangs centiles ne correspondent pas clairement d'un système à l'autre. Élo réel les valeurs peuvent être cartographiées à l'aide d'équations estimées à partir des scores du même joueur à travers les deux systèmes.

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 afficher les distributions avec et sans ce filtre.

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.

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: Comparaison illustrative entre le percentile rapide de 10 minutes Chess.com points et les mêmes points centiles dans le Lichess 10+0 rapide actif distribution après avoir nécessité au moins cinq jeux 10+0. “Chess.com mappé equiv.” applique le Elo+Chess actuel Conversion rapide de l’échelle de notation 10+0 vers la note Lichess du même centile. Le les rangées de queue la plus haute sont incluses pour montrer la distinction d’échelle, mais les rangées ajustées La ligne de notation est la plus fiable dans la gamme principale de coaching.

La raison des lacunes dans le tableau 2 est que le mappage estime les notes de force de jeu équivalentes, tandis qu’un centile décrit position au sein de la propre population d’une plateforme. Un joueur du 71,7e centile sur Chess.com n’est donc pas automatiquement un acteur du 71,7ème percentile sur Lichess, et vice versa. Les chiffres observés sont cohérents avec les données publiques de Chess.com. pool rapide ayant une masse de joueurs moins bien notés ou occasionnels beaucoup plus importante, tandis que le Le pool Lichess à contrôle exact actif est plus concentré parmi les répétitions expérimentées joueurs. Le point clé n’est pas qu’une échelle centile soit « juste » et que la autre est « faux » ; c’est que le rang centile n’est pas portable d’une plateforme à l’autre à moins que les distributions de joueurs sous-jacentes aient la même forme.

Note on the 1500 spike. The visible jump near 1500 est présent dans les données plutôt que créé par lissage. Dans le dernier distribution des notes observée, 1 500 notes exactes représentent 13,1% de 1+0 joueurs bullet, 5,6% des joueurs blitz 3+0, 9,4% des joueurs rapides 10+0, et 17,6% de joueurs rapides plus longs. Ceci est cohérent avec le fait que 1 500 agissent comme un ancrage de départ ou provisoire dans le système de notation.

Le pic exact de 1 500 est en grande partie dû à des comptes peu actifs. À vérifiez ceci, nous avons recalculé les mêmes distributions de notation après avoir demandé chaque joueur doit jouer au moins cinq parties dans le type de jeu exact. Le pic presque disparaît sous ce filtre : 1 500 notes exactes passent de 5,6 à 17,6 % de joueurs dans la distribution tous joueurs à 0,2–1,2% dans les cinq parties ou plus distribution indiquée dans le tableau 1.

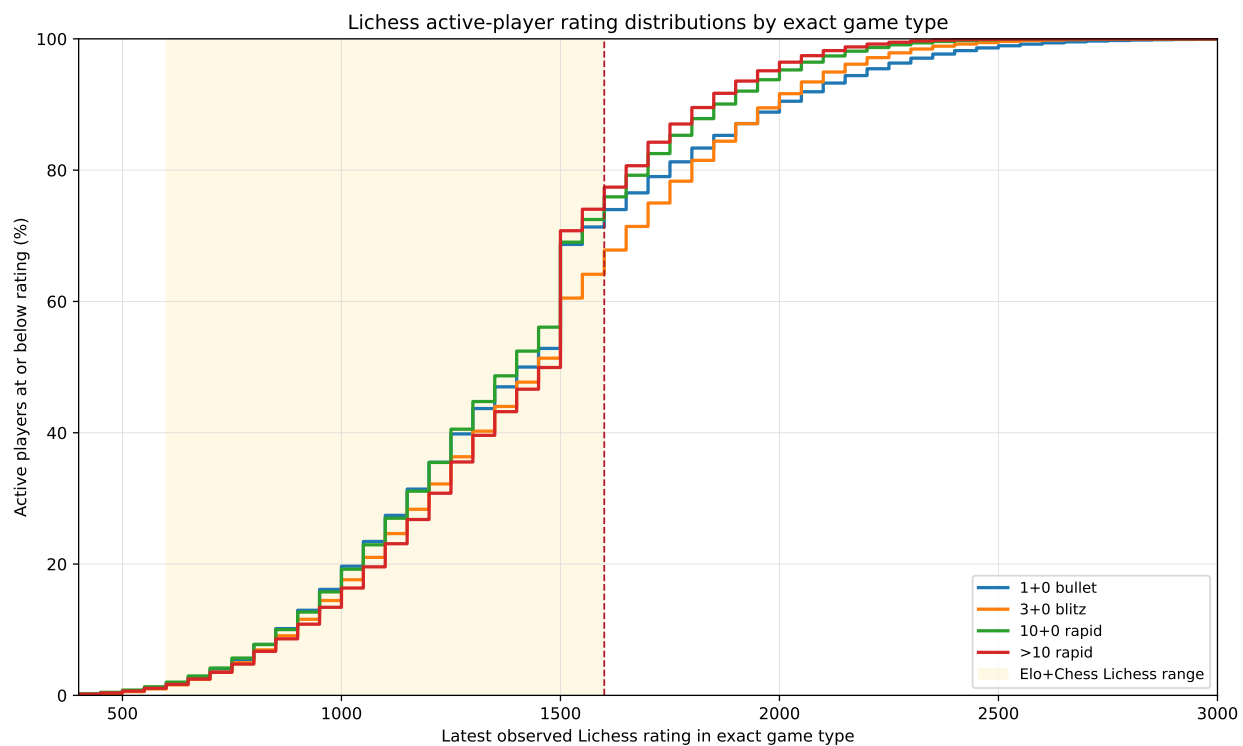


Figure 1: Répartitions cumulatives des notes des joueurs actifs par type de jeu exact. La bande jaune marque la gamme 600–1600 Lichess utilisée par le Elo+Chess principal repères de coaching.

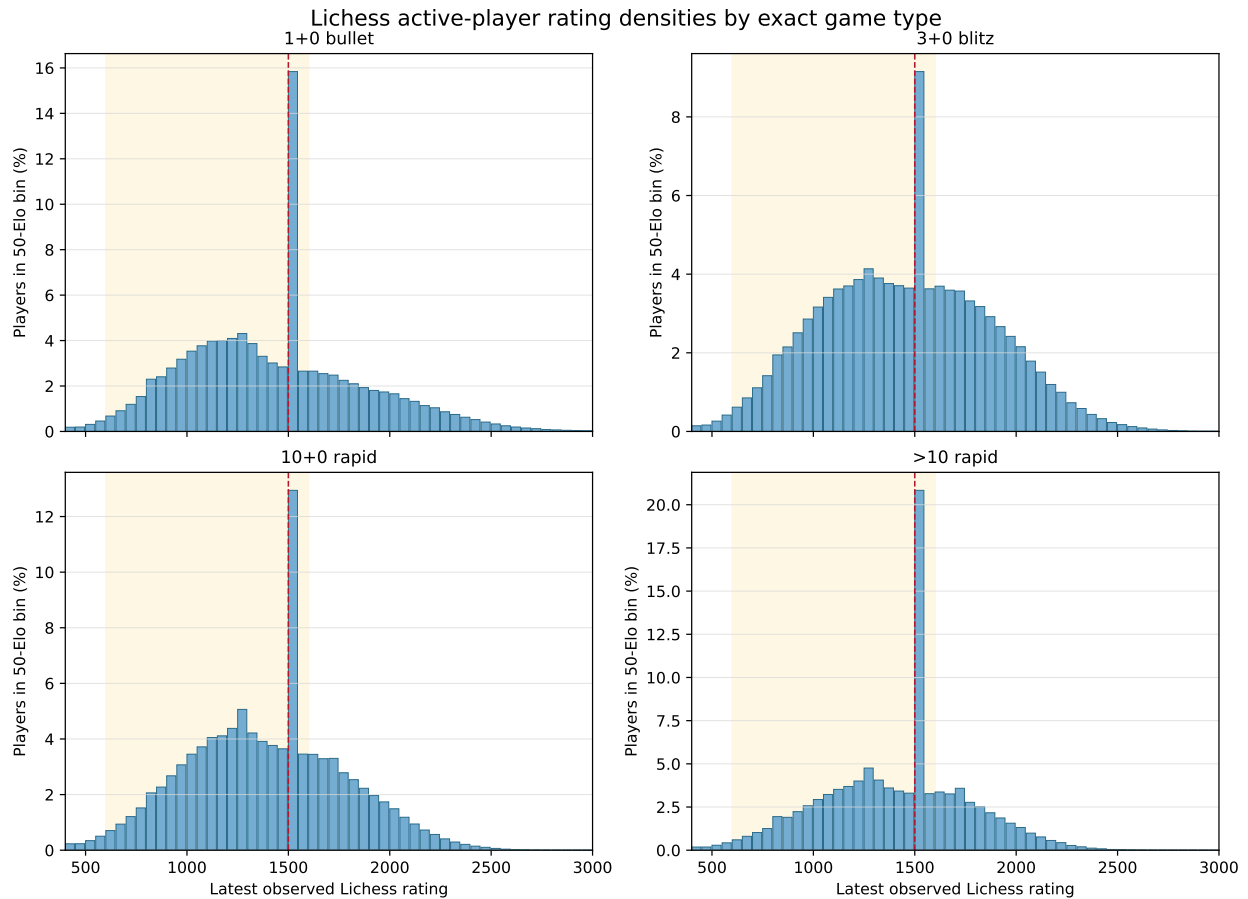


Figure 2: Vue en densité des mêmes distributions de notes des joueurs actifs. La ligne pointillée marque 1 500 Lichess, où les distributions brutes des dernières notes ont un grand point de masse précis. La bande jaune marque le 600–1600 Lichess plage utilisée par les principaux rapports de métriques Elo+Chess. Le pic est le plus visible dans la densité de tous les joueurs car de nombreux comptes peu actifs restent exactement à 1500 ; La figure 4 répète le tracé de densité après avoir exigé au moins cinq jeux dans le type de jeu exact, ce qui supprime largement cet artefact.

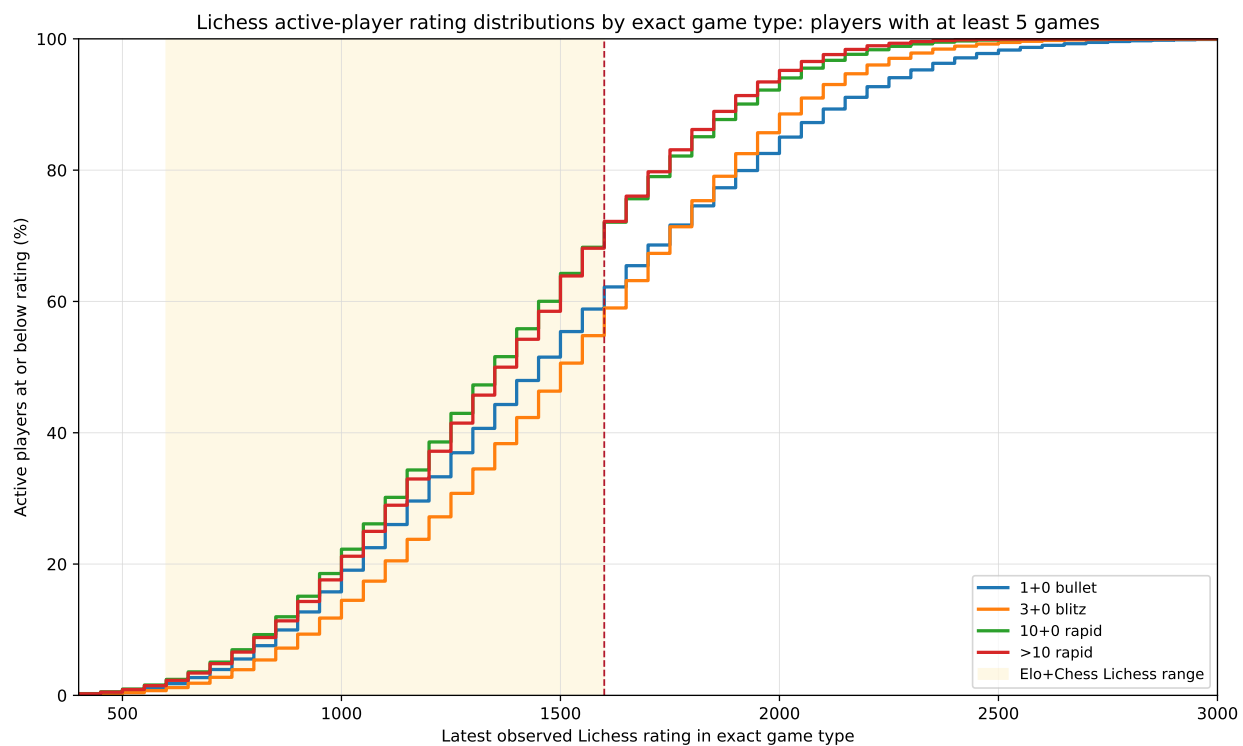


Figure 3: Distributions cumulatives des notes des joueurs actifs après avoir demandé au au moins cinq jeux dans le type de jeu exact. Le saut de 1500 est considérablement réduit.

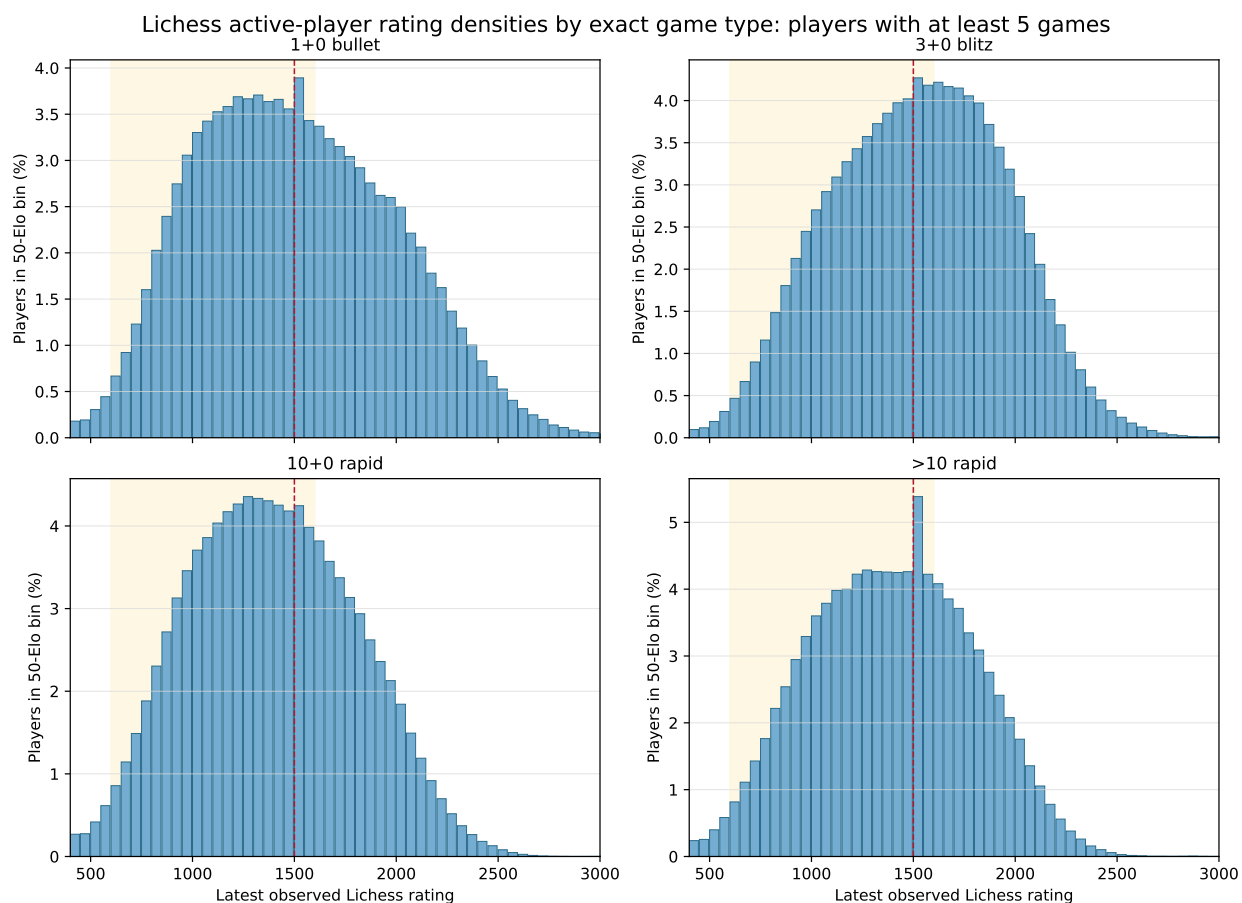


Figure 4: Vue de densité après avoir requis au moins cinq jeux dans le jeu exact tapez. Cette vue confirme que le grand point de masse de 1 500 dans le jeu pour tous les joueurs la distribution est principalement pilotée par des comptes peu actifs. La bande jaune marque la plage 600–1600 Lichess utilisée par les principaux rapports métriques Elo+Chess.

4 Échelle des artefacts de référence actuels

Le calcul est intentionnellement volumineux. Ce n'est pas une éraflure de profil, une petite un ensemble de PGN téléchargés ou une collection d'exemples instructifs sélectionnés à la main. Le pass Lichess initial couvrait les dossiers mensuels complets à partir de janvier 2025 jusqu'en mars 2026. La table d'analyse non stratifiée contient 1 382 178 087 jeux et 2 764 356 174 lignes d'évaluation des jeux des joueurs. Sur ces lignes, 6 623 438 uniques les noms d'utilisateur des joueurs apparaissent au moins une fois.

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: Échelle de numérisation Lichess non stratifiée avant échantillonnage de référence. L'analyse couvre les fichiers mensuels complets de l'historique du jeu de janvier 2025 à mars 2026.

Les artefacts de référence de production sont beaucoup plus petits que l'analyse brute car ils sont sélectionnés pour la construction métrique. Ils sont encore grands : la référence Cette étape est un pipeline d'expansion et d'agrégation d'états de déplacement sur des centaines de des milliers de jeux par type de jeu.

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: Échelle des artefacts de référence stratifiés après alignement sur le schéma commun de fonctionnalités d'exécution à 277 colonnes. Un « rangée joueur-jeu » signifie une partie du point de vue d'un joueur, donc la plupart des jeux produire deux rangées de jeu de joueur.

Pour le seul blitz 3+0, le benchmark de production actuel utilise 479 974 jeux. impliquant 129 398 joueurs uniques et 28 138 904 lignes d'état de mouvement distinctes. Le Le tableau des fonctionnalités du jeu blitz aligné contient 277 variables dérivées. Le les artefacts de création rapide de fonctionnalités sont encore plus larges au stade de l'état de déplacement : l'ensemble la table d'état de déplacement rapide utilisée lors de la construction des fonctionnalités contient 168 colonnes d'état de déplacement avant agrégation dans une fonctionnalité de jeu de joueur de 277 colonnes tableau. Les artefacts de référence du runtime frontend pour les quatre types de jeux sont puis aligné sur le même schéma de fonctionnalités de jeu de joueur de 277 colonnes, y compris le variables de répartition tactique-taxonomie étendues. À partir de ces tableaux de fonctionnalités, le Le rapport sélectionne et affiche les mesures de coaching les plus interprétables.

Ces décomptes sont importants pour l'interprétation des résultats. Les courbes de référence ne sont pas des estimations tirées de quelques centaines de jeux examinés manuellement. Au lieu de cela, ils proviennent de dizaines de millions d'États du conseil évalués et de plus de quatre

millions de lignes de jeu de joueur dans les quatre catégories de jeux de production actuelles.

5 Aperçu du pipeline

Le pipeline de construction de référence comporte six étapes principales :

1. Sélectionnez les jeux Lichess classés dans le type de jeu cible et les strates de notation.
2. Analysez chaque coup de jeu par coup et stockez l'avant-coup et l'après-coup état du conseil d'administration.
3. Calculez les variables intermédiaires d'état de mouvement à partir du tableau, du mouvement, le point de vue du joueur et le point de vue de l'adversaire.
4. Regroupez les variables d'état de mouvement en variables de fonctionnalités de jeu de joueur.
5. Convertissez les variables du jeu du joueur en métriques de rapport avec un sens d'interprétation : plus haut est généralement mieux, plus bas est généralement mieux, ou descriptif.
6. Faites la moyenne de chaque métrique par compartiment Elo et inspectez manuellement le résultat relation.

Cette structure est importante car de nombreuses mesures de rapport ne sont pas directement visible dans un en-tête PGN. Ils nécessitent de reconstituer la situation, de déterminer quelles pièces ont été déplacées à partir de quelles cases de départ, en comparant avant le mouvement et après le déplacement du matériau et de la structure, puis en agrégeant le résultat du passer du niveau au niveau du jeu du joueur.

6 Variables intermédiaires

Les variables intermédiaires sont conçues pour décrire ce qui a changé sur le tableau et pourquoi c'est important du point de vue du rapporteur. Les exemples incluent :

- numéro de coup, numéro de coup du joueur, côté à déplacer, couleur du joueur et joueur couleur;
- du carré, du carré, du type de pièce mobile, du type de pièce capturée et si le mouvement était une capture, un échec, un château, une promotion ou un compagnon ;
- bilan matière avant et après le déménagement ;
- si une pièce mineure a quitté sa case de départ selon un numéro de coup donné ;
- si la même pièce déplacée sur des joueurs consécutifs se déplace dans le ouverture;
- si la reine a bougé avant le coup 10 ;
- côté du roque, numéro du mouvement du roque et si les tours sont devenues connectées après le roque ;

- pièces détachées, pièces suspendues, pions doublés, pions isolés, passés les pions et les pions arriérés ;
- statut du fichier tour, fichiers ouverts, fichiers semi-ouverts et sécurité du roi caractéristiques;
- opportunités tactiques, tactiques exécutées, tactiques de pression matérielle, gains de suivi garantis et gains réalisés.

L'ensemble intermédiaire exact varie selon l'étape de construction. Les bases de données de rapports d'exécution conservent uniquement les colonnes nécessaires pour fournir rapidement des rapports, tandis que la création complète de fonctionnalités les artefacts incluent des tables d'état de déplacement plus larges et des correctifs de fonctionnalités utilisés pour construire les dernières lignes de fonctionnalités du jeu joueur.

7 Exemples de métriques de principe d'ouverture

Les principes d'ouverture sont des exemples utiles car ils sont faciles à comprendre pour les utilisateurs. comprendre et facile à connecter aux calculs d'état de mouvement.

7.1 Minor Pieces Developed by Move 10

Cette métrique compte combien des quatre chevaliers et évêques d'un joueur il reste leurs cases d'origine au 10ème coup du joueur. Le calcul intermédiaire utilise la couleur du moteur, le type de pièce en mouvement, la case de départ et le mouvement du joueur numéro. Pour les Blancs, les cases d'origine sont b1, g1, c1 et f1. Pour les Noirs, ce sont b8, g8, c8 et f8. Pour chaque jeu-joueur, Elo+Chess compte la maison distincte les carrés parmi les chevaliers et les évêques déplacés par le joueur se déplacent 10.

Il s'agit d'une mesure plus élevée est généralement meilleure. Il ne prétend pas que chaque mineur la pièce doit toujours se déplacer au coup 10, mais parmi les joueurs en développement, la relation est claire : les joueurs qui atteignent des niveaux Elo plus élevés ont tendance à se développer. des pièces plus mineures plus tôt.

7.2 Repeated Piece Moves in the Opening

Cette métrique compte les premiers mouvements supplémentaires passés à déplacer à nouveau la même pièce à la place. de développer quelque chose de nouveau. Le calcul intermédiaire utilise le précédent la case de destination, la case d'origine actuelle et le numéro du coup du joueur. Si la même pièce déplacée lors du coup précédent du joueur et est déplacée à nouveau dans le Lors des 10 premiers mouvements du joueur, le nombre de mouvements répétés augmente.

Il s'agit d'une mesure plus faible est généralement meilleure. Il mesure les tempi d'ouverture gaspillés dans un manière simple. Certains mouvements répétés sont tactiquement justifiés, mais à grande échelle, La courbe en seau reflète la tendance générale : des acteurs plus forts vers l'objectif les niveaux dépensent moins de premiers mouvements pour déplacer la même pièce à plusieurs reprises.

7.3 Queen Touched Before Move 10

Cette métrique compte les coups de la reine avant le coup 10 du joueur. Le calcul utilise le type de pièce en mouvement et le numéro de mouvement du joueur. Si la pièce mobile est la reine et que le numéro de coup du joueur est au maximum 10, le décompte augmente.

Ce chiffre est également inférieur, c'est généralement meilleur pour la plage cible Elo+Chess. Tôt les mouvements de reine peuvent gagner du matériel ou forcer des concessions, mais de nombreux joueurs en développement faire sortir la reine trop tôt, perdre le tempo face aux attaques contre la reine et retarder développement des pièces mineures et sécurité du roi.

7.4 Castling and King Safety

Les métriques de roque utilisent les drapeaux de roque, le côté du roque, le numéro de mouvement du roque et l'état du plateau après le roque. Les exemples incluent si le joueur a roqué, si le joueur a roqué au coup 10, si le joueur a roqué côté roi ou côté reine, si les tours étaient connectées après le roque et si les tours étaient ouverts ou entrouverts après le roque.

Ces métriques ne sont pas toutes interprétées de la même manière. Achèvement du roque et le roque au coup 10 est généralement plus élevé, c'est généralement meilleur dans la plage cible. Le roque Queenside ou le roque dans des fichiers ouverts peut être plus dépendant du contexte. C'est pourquoi Elo+Chess ne se contente pas de coder en dur une règle morale ; il construit un seau courbe et inspecte quelles relations sont suffisamment stables pour le coaching.

8 Courbes de godet et inspection manuelle

Une fois les métriques de jeu des joueurs calculées, Elo+Chess les regroupe par Lichess Elo bucket et calcule les moyennes du bucket. La moyenne de la tranche est la valeur empirique brute point de référence pour cette métrique. Une ligne de tendance est ensuite utilisée dans le rapport pour attribuer des notes de style Elo aux valeurs utilisateur.

Les courbes sont inspectées manuellement avant d'être utilisées. Une métrique est plus utile lorsque les moyennes du compartiment montrent une relation cohérente sur toute la plage de notation servie par le produit. Une métrique est moins utile lorsqu'elle est bruitée, plate, très non monotones ou motivés par des cas particuliers difficiles à expliquer à un utilisateur.

Cette étape d'inspection manuelle est particulièrement importante au-dessus de la partie supérieure du Elo+Chess. coupure. Les joueurs les plus forts savent quand enfreindre les principes d'ouverture et le font efficacement. Au-dessus de la fourchette débutant à avancé, la relation entre les habitudes simples et la notation s'aplatit ou change souvent de forme. Le produit souligne donc la gamme dans laquelle les relations empiriques sont fortes et pédagogiquement utile.

9 Exemples de courbes d'ouverture sur toute la plage

La figure 5 montre quatre courbes de principe d'ouverture blitz 3+0. Les points gris montrent la plage nominale Lichess plus large. Blue points and the blue trend la ligne montre la gamme 600–1600 Lichess utilisée par le coaching principal Elo+Chess des repères. The dashed red line marks the upper cutoff.

Le même schéma est visible dans la gamme complète 10+0 rapide et plus rapide. artefacts. La décision centrale de modélisation reste inchangée : utiliser la méthode empirique plage stable

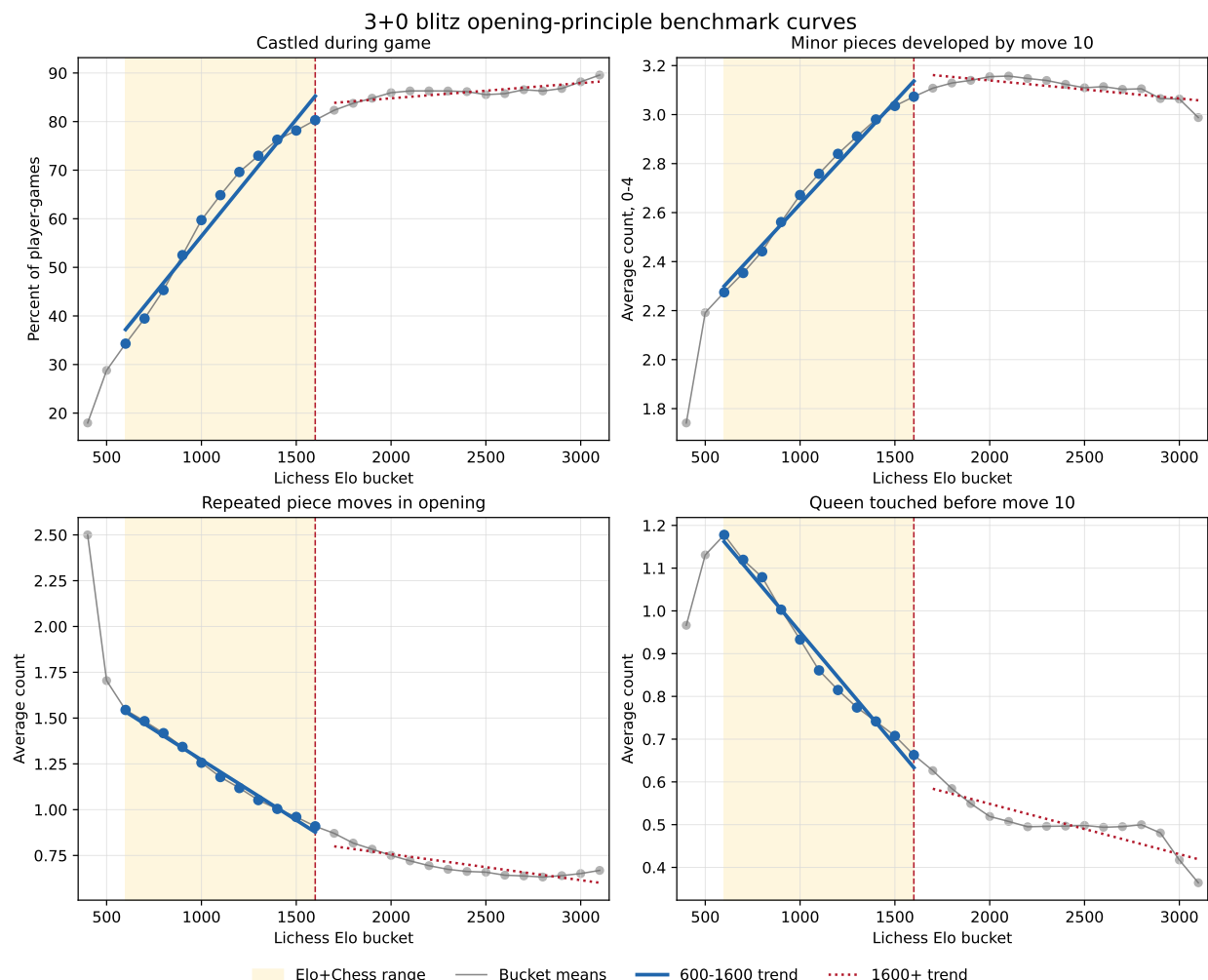


Figure 5: Courbes de référence sélectionnées selon le principe d'ouverture pour le blitz 3+0. Les relations sont les plus fortes et les plus utiles dans la gamme de coaching Elo+Chess. Au-dessus du seuil supérieur, plusieurs relations s'aplatissent considérablement.

pour les allégations de coaching et évitez de prétendre qu'une simple courbe d'habitude reste tout aussi informatif pour toujours.

Le même modèle apparaît numériquement dans le tableau 5. Les pentes sont signalé comme changement de valeur métrique pour 100 points Elo. Roque, pièce mineure le développement, le mouvement répété des pièces et les premiers mouvements de la reine montrent tous clairement mouvement sur 600–1600. Au-dessus de 1600, les pentes sont beaucoup plus faibles dans ces régions. exemples, ce qui est cohérent avec l'idée selon laquelle les joueurs les mieux notés cassent principes de base de manière plus sélective et avec plus de succès. Un deuxième mécanisme est saturation. Pour certaines habitudes, la métrique atteint un plafond pratique ou doux spot : les joueurs forts ont déjà pris l'habitude, il n'y a donc aucune raison de la courbe continue de progresser fortement vers le haut. Plus n'est pas toujours mieux après ce point. Les données s'aplatissent donc à la fois car des acteurs plus forts peuvent briser règles dans des positions concrètes et parce que de nombreuses habitudes de base ont déjà atteint leur plage de fonctionnement utile.

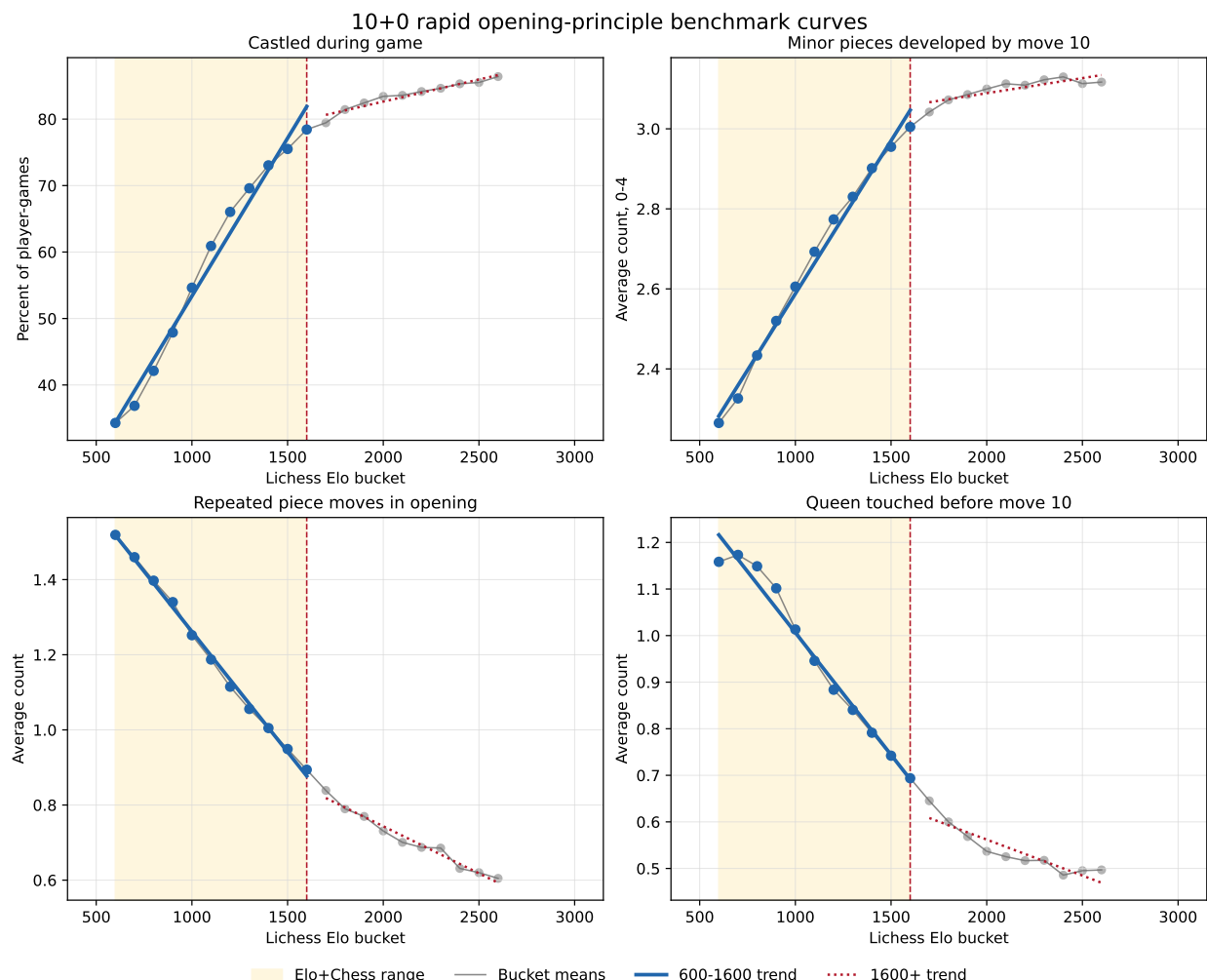


Figure 6: Courbes de référence du principe d'ouverture sélectionnées pour les jeux rapides 10+0. Les relations sont les plus fortes et les plus utiles dans la gamme de coaching Elo+Chess. Au-dessus du seuil supérieur, plusieurs relations s'aplatissent considérablement.

10 Pourquoi le seuil supérieur est important

L'utilisateur cible de Elo+Chess est un joueur débutant à avancé. Pour ces joueurs, les habitudes de base sont fortement associées au classement : développer des pièces, éviter les mouvements répétés des premières pièces, éviter les premiers mouvements inutiles de la reine, le château à temps, connectez les tours, gérez le matériel et évitez de laisser des pièces lâches ou suspendu.

À des notes plus élevées, ces mêmes habitudes de surface deviennent moins diagnostiques. Un fort le joueur peut retarder le roque parce que le centre est fermé, déplacer la reine plus tôt parce que la position contient une tactique concrète, ou déplacer deux fois la même pièce parce qu'il gagne un tempo ou force une concession structurelle. L'habitude brute toujours a un sens, mais la relation entre l'habitude et la note devient plus conditionnel. C'est pourquoi Elo+Chess utilise un seuil supérieur pour son coaching principal allégations et pourquoi les utilisateurs au-dessus de cette plage sont avertis que la relation étroite entre les habitudes de principe et les pannes de notation.

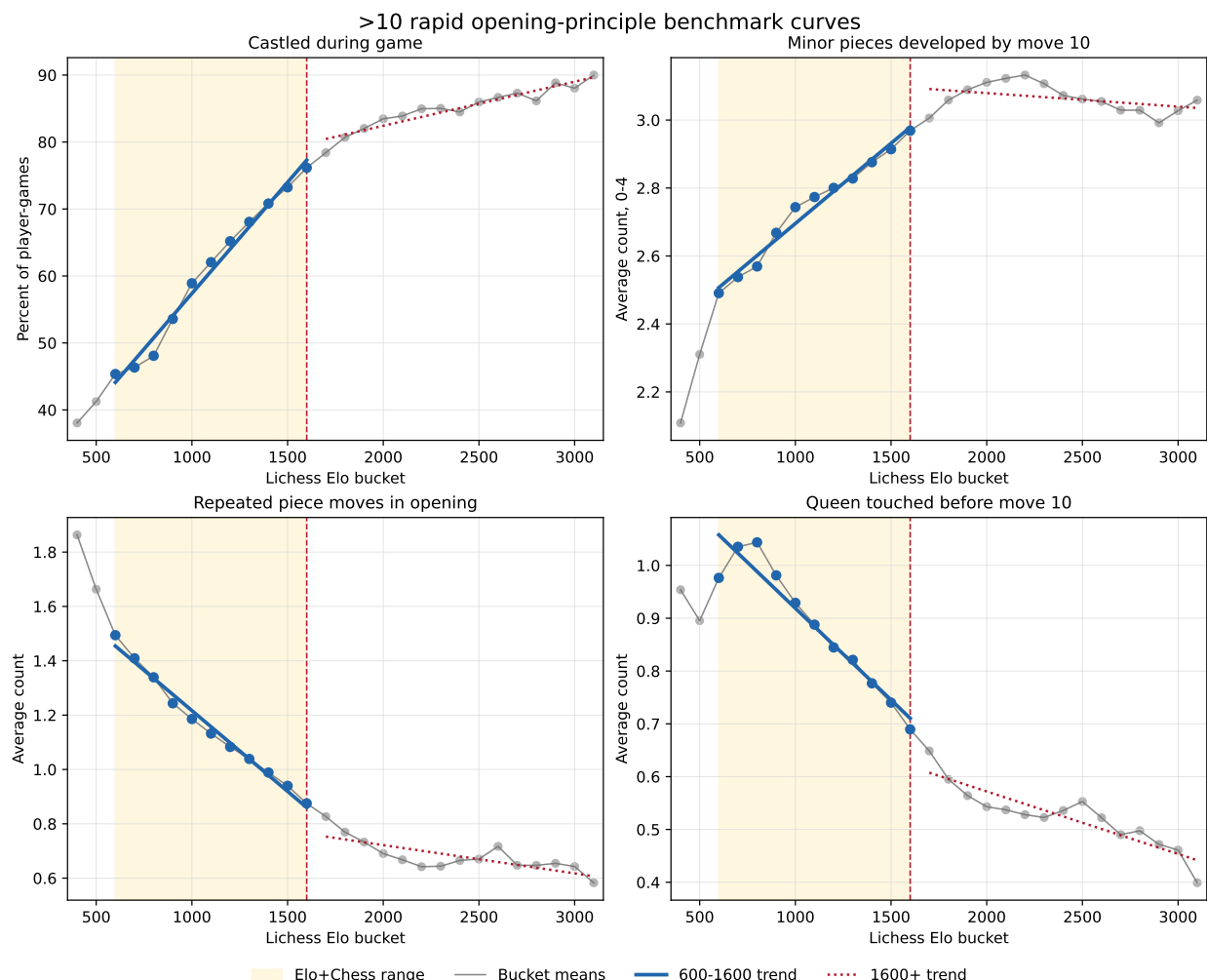


Figure 7: Courbes de référence sélectionnées selon le principe d'ouverture pour les jeux rapides plus longs.

11 Mesures tactiques bilatérales et objectifs centiles

Certaines métriques ont la forme opposée aux exemples de principe d'ouverture. Fourchettes sont un exemple utile car l'événement est intrinsèquement biface. Un joueur qui trouve et convertit des forks plus utiles que ses pairs fait quelque chose de bien, mais le taux de population de fourchettes converties diminue avec la notation car plus fort les adversaires accordent moins d'opportunités de fork et se défendent contre les menaces tactiques plus activement.

La figure 8 montre cet effet. Les deux fourchettes converties par le joueur de rapport et les fourchettes converties par l'adversaire deviennent moins fréquents à mesure les notes augmentent. Cela ne veut pas dire que l'exécution de forks est mauvaise. Cela signifie l'événement brut le taux est en partie contrôlé par le taux d'erreur de l'adversaire et la texture du jeu. Pour les métriques comme ceci, Elo+Chess utilise donc les percentiles de l'historique des pairs à l'intérieur du le propre bucket Elo du joueur plutôt que de s'appuyer uniquement sur une tendance globale moyenne du bucket.

Pour construire ces objectifs centiles, Elo+Chess échantillonne les joueurs par type de jeu et Le

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: Pentes approximatives de la courbe de godet pour les quatre exemples de principe d'ouverture métriques pour tous les types de jeux de production.

compartiment Elo utilise ensuite les jeux de qualification récents pour chaque joueur échantillonné. Le L'ensemble d'objectifs actuel utilise 11 compartiments de 600 à 1 600, 200 joueurs par seuil, et jusqu'à 50 jeux de qualification récents par joueur pour chaque production type de jeu. Cela donne 2 200 échantillons de joueurs par type de jeu et environ 110 000 lignes de jeu de joueur par type de jeu pour le percentile intra-bucket distributions.

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: Ensembles d'objectifs centiles actuels de l'historique des pairs. Ce sont des échantillons d'historique des joueurs utilisés pour estimer les distributions au sein du compartiment pour métriques de type percentile.

La figure 9 illustre la cible résultante distribution pour le taux de fork converti des joueurs échantillonnés en blitz 3+0. Chaque boîte est une distribution au sein du bucket entre les joueurs échantillonnés, et non une moyenne du bucket jeux individuels. Dans l'interface utilisateur du rapport, la valeur d'un utilisateur est comparée à la valeur pertinente. boîte pour le compartiment Elo mappé de l'utilisateur. C'est la base de déclarations telles que si le taux de fork converti d'un joueur est élevé ou faible par rapport à ses pairs au moment même niveau, même lorsque le taux d'événements dans la population globale diminue avec la notation.

12 Coût de réplcation

La méthode est reproductible en principe, mais elle n'est pas légère. Reconstruction les courbes de référence nécessitent :

Two-sided converted-fork rates by Elo bucket

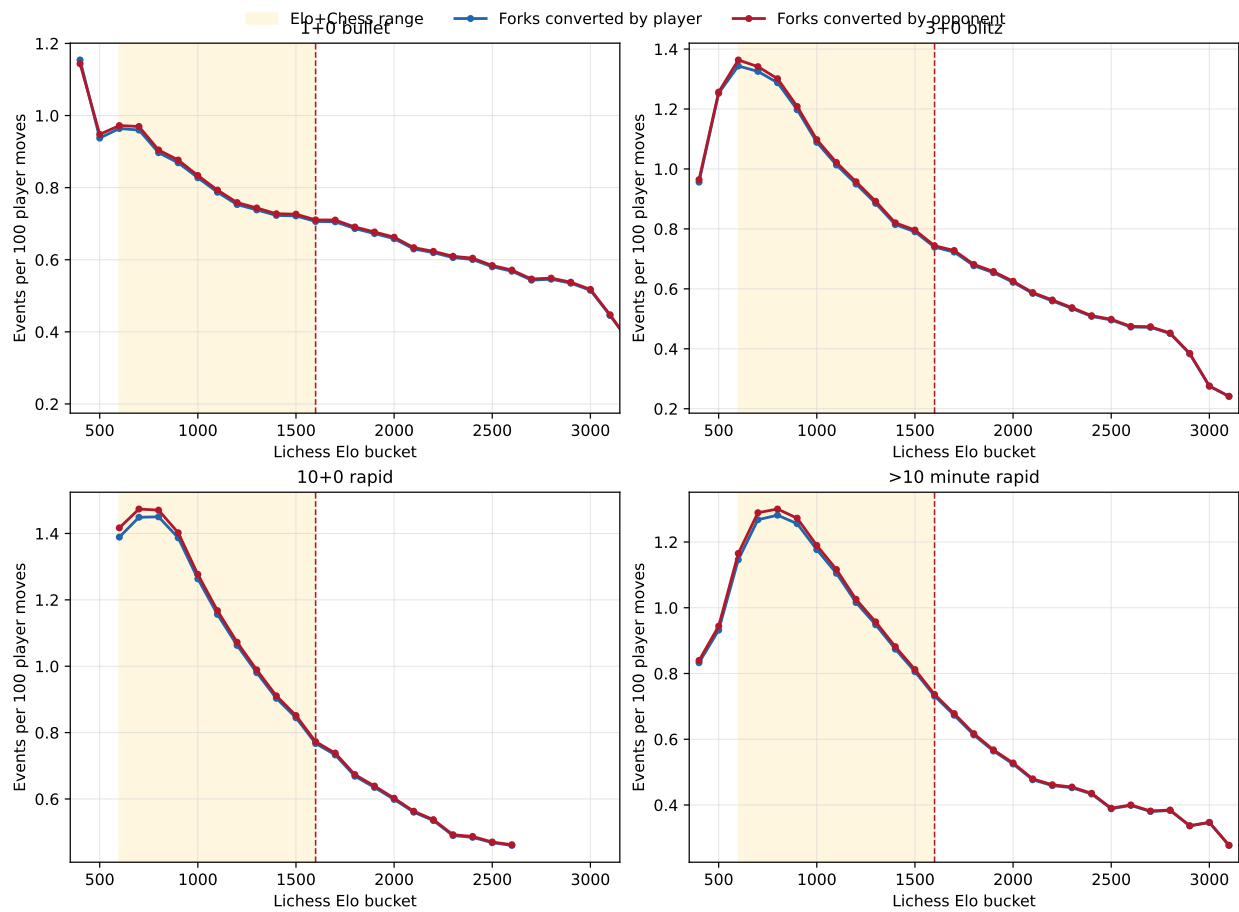


Figure 8: Tarifs de fourche convertis recto-verso par seau Elo. Événements de fork convertis chuter à des notes plus élevées parce que les deux camps défendent mieux et moins de fourchettes propres des conversions sont disponibles.

- accès aux gros fichiers mensuels d'historique de jeu Lichess ;
- filtrage stratifié par type de jeu, catégorie de notation et éligibilité règles;
- un analyseur de mouvements légaux et un évaluateur de conseil d'administration capable d'étendre des dizaines de millions d'états de déplacement ;
- logique des fonctionnalités pour l'ouverture, le matériel, la structure des pions, la sécurité du roi, la tour activité, tactiques, gestion du temps et résultats du jeu ;
- agrégation de l'état de déplacement aux fonctionnalités du jeu du joueur ;
- des résumés au niveau du compartiment pour chaque métrique du rapport ;
- examen manuel de chaque courbe de référence candidate avant son utilisation dans un rapport de coaching.

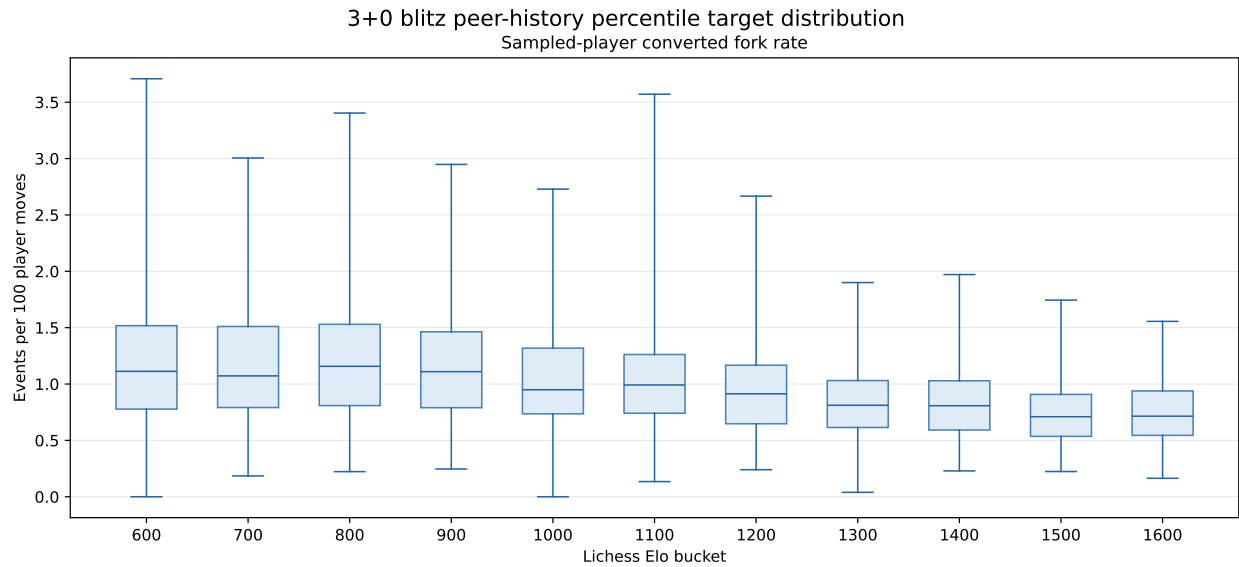


Figure 9: Exemple de distribution cible de l'historique des pairs blitz 3+0 pour un joueur échantillonné taux de fourchette converti. Les centiles sont calculés dans le compartiment Elo, de sorte que le rapport compare un joueur avec des pairs confrontés à des opportunités tactiques globalement similaires tarifs.

Pour un blitz 3+0, cela signifie près d'un demi-million de parties, près d'un million lignes de jeu de joueur et plus de 28 millions de lignes d'état de mouvement. À travers les quatre types de jeux de référence, les artefacts de référence actuels contiennent plus de 2,55 millions de jeux, plus de 5,10 millions de lignes de jeu de joueur et plus de 147 millions de lignes d'état de déplacement. L'échelle est au cœur de la valeur du benchmark : il permet au rapport de montrer aux utilisateurs ce qui se passe réellement dans les jeux joués par des joueurs proches de leur niveau plutôt que de s'appuyer uniquement sur des slogans d'échecs généraux.

13 Limites

Les courbes de référence sont des descriptions empiriques et non des preuves causales. Un joueur n'obtient pas de note simplement en correspondant à une moyenne de tranche. Les mesures sont également résumés imparfaits de positions complexes : un coup de dame précoce peut être excellent dans une position et imprudent dans une autre. Le but des courbes est de identifier les habitudes récurrentes et les comparer au comportement des pairs, pour ne pas remplacer le calcul ou le jugement du coaching.

Les courbes sont également des courbes à l'échelle Lichess. Lorsque les utilisateurs apportent le jeu Chess.com historiques, Elo+Chess mappe la tranche de notation pertinente sur le benchmark Lichess à l'échelle en utilisant le modèle de cartographie multiplateforme distinct. Cette couche cartographique est décrit dans la note de recherche complémentaire Elo+Chess sur Lichess-to-Chess.com conversion de note : [Estimating Cross-Platform Chess Rating Mappings with Régression modale](#). Cette cartographie ne doit pas être interprétée comme préservant le percentile rang. Il préserve une équivalence de notation estimée entre les deux notations systèmes; le rang centile reste spécifique à la plateforme car le joueur actif les répartitions sont différentes.

14 Graphiques du rapport de production

Le rapport de production transforme les artefacts de référence en panneaux de niveau métrique qui peut être lu sans connaître le pipeline de construction complet. Chiffres 10 and 11 montrent trois exemples de la conception du rapport Elo+Chess en direct.

Les panneaux du principe d'ouverture dans la figure 10 utilisent la présentation ordinaire de la courbe de référence. Chaque carte commence par le famille de métriques et nom de la métrique, donne ensuite une brève règle d'interprétation. Le la grande note dans le coin supérieur droit est intentionnellement séparée du brut valeur métrique : elle répond à la question posée à l'utilisateur : « comment cette habitude est-elle évaluée ? par rapport à la référence ? » tandis que les lignes ci-dessous préservent les valeurs mesurées. Les quatre lignes récapitulatives affichent la durée de vie, les 10 derniers, les victoires et les pertes. Cela rend le diagnostic du panneau plutôt que simplement décoratif. Si la durée de vie et les 10 derniers correspondent, l'habitude est probablement stable. Si les victoires et les défaites divergent, la mesure peut être dépendant du résultat ou lié à l'état du jeu plutôt qu'à une simple habitude.

La zone du graphique utilise un arrière-plan sombre et discret, une ligne de tendance dorée et des graphiques colorés. marqueurs pour les quatre tranches. L'annotation de tendance indique explicitement si la relation de référence augmente ou diminue avec Elo. C'est important parce que certaines mesures sont plus élevées, c'est mieux, comme le développement de pièces mineures, tandis que d'autres sont plus bas, c'est mieux, comme le mouvement précoce de la reine. La carte donc ne demande pas à l'utilisateur de déduire la direction à partir de la seule pente. La finale Les lignes « Ce que cela signifie » et « Conseil de coaching » traduisent le résultat du benchmark en une instruction d'échecs concrète tout en préservant le contexte empirique.

La figure 11 montre le style centile présentation utilisée pour les métriques dont le taux brut n'est pas bien représenté par un tendance mondiale unique. Les fourchettes gagnantes sont un exemple tactique : le taux de population des fourchettes converties changent à la fois avec les compétences du joueur et le taux d'erreur de l'adversaire. Le La carte compare donc la valeur de l'utilisateur avec les historiques des autres joueurs dans le compartiment Elo mappé. Le boxplot affiche la distribution des pairs, y compris les intervalle interquartile, médiane et moustaches. Le marqueur de l'utilisateur est tracé sur le même échelle et étiquetés avec à la fois la valeur brute et le rang centile. Adjacent les compartiments de comparaison restent visibles sous forme muette afin que l'utilisateur puisse voir que le Le bucket de pairs est local et non universel.

Cette conception rend les mesures centiles vérifiables. L'utilisateur voit la valeur brute, le percentile, le compartiment de pairs utilisé pour la comparaison et le traduit Mappage de compartiment Chess.com à Lichess. La même durée de vie, les 10 derniers, les victoires et Les lignes de pertes sont conservées afin que les scores centiles tactiques puissent toujours être vérifiés pour les effets de récence et la dépendance aux résultats.

15 Conclusion

Les rapports de référence Elo+Chess sont construits à partir d'un grand calcul d'état de mouvement sur échantillons stratifiés Lichess. Chaque métrique commence par un ensemble de données concrètes sur l'état du conseil d'administration. et les variables d'état de mouvement, deviennent une fonctionnalité de jeu du joueur, et sont ensuite moyennées par compartiment Elo pour former une courbe de référence empirique. The opening-principle des exemples montrent pourquoi cette approche est utile : les habitudes simples aux échecs ont des effets forts, relations visibles avec la notation dans la fourchette cible, tandis que celles les relations s'aplatissent souvent ou devi-

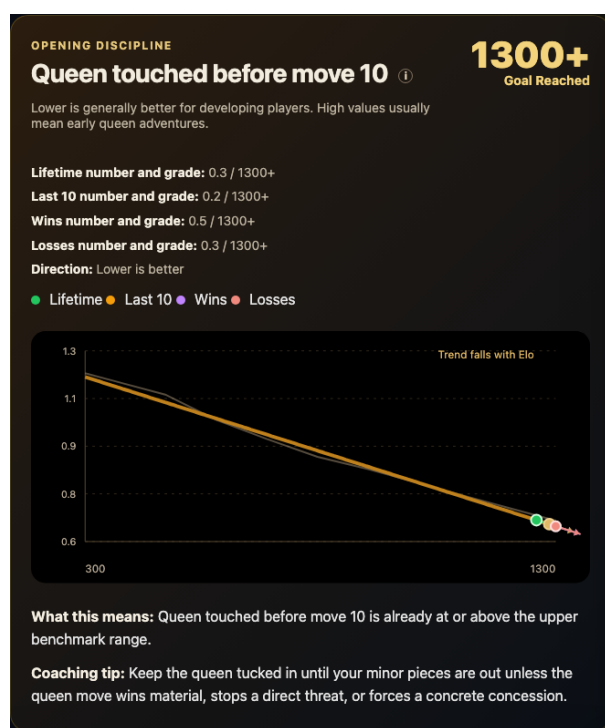
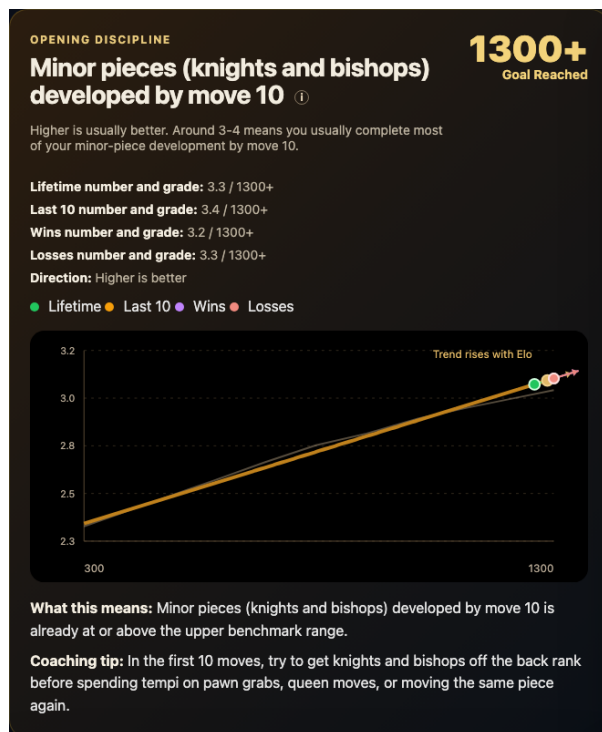


Figure 10: Panneaux selon le principe d'ouverture de la production. La même grammaire visuelle gère un la métrique la plus élevée est la meilleure, le développement de pièces mineures et la métrique la plus faible est la meilleure, premier mouvement de la reine. Les deux panneaux exposent la durée de vie, les 10 derniers, les victoires et les pertes afin que le rapport puisse séparer les habitudes stables des habitudes récentes ou dépendantes des résultats modèles.

ennent plus conditionnelles au-dessus du seuil supérieur. C'est la principale raison pour laquelle Elo+Chess concentre ses affirmations en matière de coaching sur le classement. plage où les données les soutiennent le plus clairement.

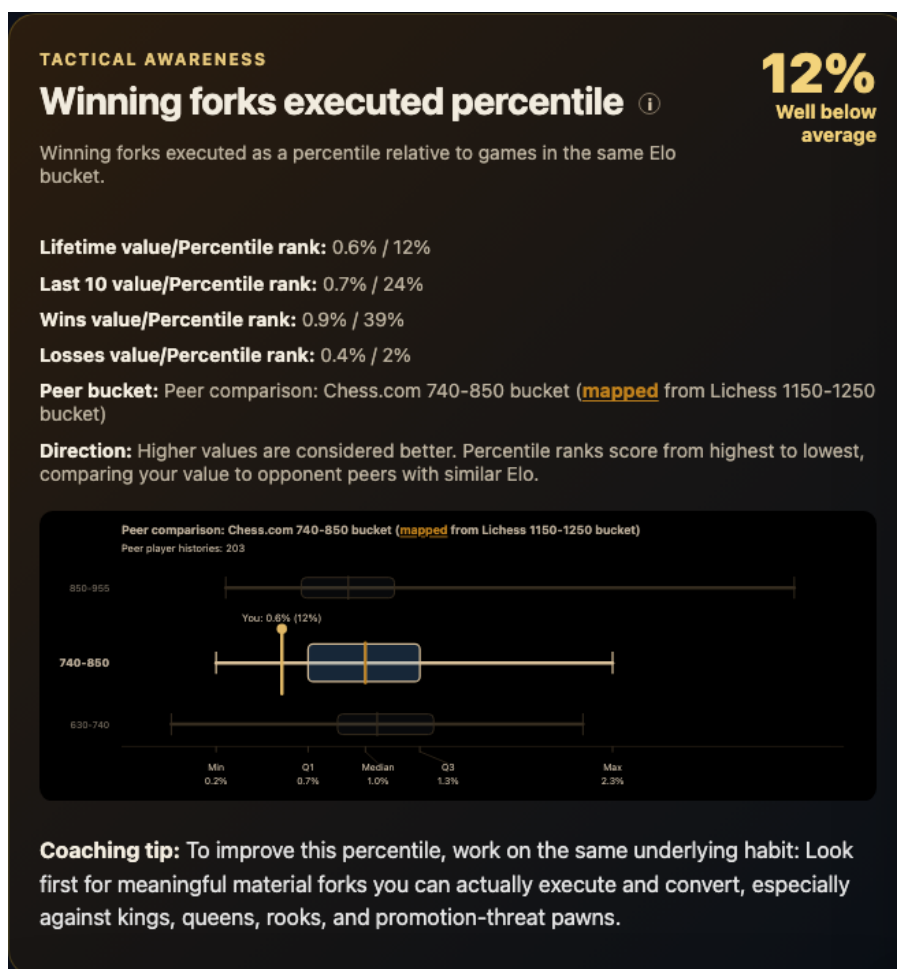


Figure 11: Panneau centile de production pour les fourches gagnantes exécuté. La boîte à moustaches montre la distribution des pairs dans le même compartiment utilisée par le rapport, tandis que l'or Le marqueur affiche la valeur brute et le rang centile de l'utilisateur.