Factoring Strength of Schedule into MLB Player Analysis

playerElo

Jacob Richey, Professor Abraham Wyner | University of Pennsylvania

Statistics updated to September 10th, 2019.

Shiny App: https://jrichey.shinyapps.io/playerelo/

FanGraphs Article: https://community.fangraphs.com/playerelo-factoring-strength-of-schedule-into-player-analysis/

Abstract

With the sabermetric revolution of the MLB, a plethora of new statistics have come into the mainstream, and a growing number of fantasy owners, ballclubs, and regular fans are turning to these new statistical methods for player analysis. However, I propose even advanced metrics such as wOBA, FIP, xwOBA, xFIP, and wRC+ are all missing a crucial element to accurately represent player performance thus-far. The playerElo system is able to reveal in aggregate the effects of previously unconsidered aspects of the game. Using an Elo ranking system determined by run-value calculations of all major league baseball players, the model incorporates context-dependent analysis and quality of competition to produce a proper evaluation of batters and pitchers. This enables playerElo to appropriately credit pitchers, especially relievers, for their true impact on the game, particularly when called upon in disadvantageous situations. Additionally, playerElo does not allow relative team strength, which confounds common counting statistics, to influence the evaluation of a player. The model is a holistic approach to the assessment of major league players and has incredible ramifications on player projections during free agency and player acquisition.

Introduction

Consider the following comparison between Freddie Freeman (29) and Carlos Santana (33). Both players were starters for the 2019 All-Star teams of their respective leagues and are enjoying breakout seasons beyond their usual high production level with nearly identical statistics across the board.

	PA	**wOBA**	**xwOBA**	**wRC+**
Freeman, 1B	643	0.398	0.396	144
Santana, 1B	624	0.389	0.371	141

However, I argue there is an underlying statistic that makes Santana's success less impressive and Freeman's MVP-consideration worthy. Recall the quality of competition of pitchers faced. The Atlanta Braves' division, the NL East, contains the respectable pitching competition of the Mets (11th league-wide in ERA), Nationals (12th), Phillies (17th), and Marlins (21st). Contrast this with the competition of the Cleveland Indians in the AL Central: The Twins (8th), White Sox (23rd), Royals (26th), and Tigers (28th). Over his first 500 plate appearances, Santana faced a top 15 pitcher (ranked by FIP) just 15 times, compared to 43 times by Freeman. wRC+ controls for park effects and the current run environment, while xwOBA takes into account quality of contact, but all modern sabermetrics fail to address the problem of Freeman and Santana's near-equal statistics, despite widely different qualities of competition. Thus, I present the modeling system of playerElo.

Methodology

Conceived out of inspiration from Arpad Elo's rating system for zero-sum games like chess, as well as FiveThirtyEight's use of an Elo modeling scheme for MLB team ratings and season-wide predictions, playerElo treats all at-bats as events and maintains a running power ranking of all MLB batters and pitchers. The system uses expected run values over the 24 possible base-out states. Additionally, run values are calculated for each at-bat event by subtracting the run expectancy of the beginning state from the ending state, and adding the runs scored.

$$Run\ Value\ of\ Play = RE\ End\ State - RE\ Beginning\ State + Runs\ Scored$$

The following run expectancy matrix presents the expected runs scored for the remainder of the inning, given the current run environment, baserunners, and number of outs. Data is sourced from all at-bats from 2016-2018, and expected run values are rounded to the second decimal place. For example, a grand slam hit with one out would shift the run expectancy from 1.54 to 0.27 and score four runs, so the run value of the play would be 2.73.

1B	2B	3B	0 outs	1 out	2 outs
-	-	-	0.51	0.27	0.11
1B	-	-	0.88	0.52	0.22
-	2B	-	1.15	0.69	0.32
-	-	3B	1.39	0.97	0.36
1B	2B	-	1.45	0.93	0.44
1B	-	3B	1.77	1.20	0.48
-	2B	3B	1.97	1.40	0.56
1B	2B	3B	2.21	1.54	0.75

The model begins with a calibration year of 2018, and for 2019, players begin with their previous seasons' ending playerElo, regressed to the mean slightly. If a player did not have a single plate

appearance or batter faced pitching in 2018, such as Vladimir Guerrero Jr. or Chris Paddack, then they are assigned a baseline playerElo of 1000 (calibration year of 2018 began every player at 1000). For every at-bat, given the current base-out state, an expected run value for both the batter and pitcher is calculated, based on quadratic formulas of historic performance of players of that caliber in the given situation. The dependency of the Elo formula on the base-out state ensures the model is context-dependent, meaning it incorporates the fact that a bases-loaded double is far more valuable than a double with the bases empty, however, it also takes into account that runs were more likely to be scored in the former situation compared to the latter.

It is important to note playerElo is a raw batting statistic and does not evaluate overall production, meaning stolen bases are not factored into the ranking system. Additionally, while the model does not take defense into account, it also does not count stolen bases or passed balls negatively against a pitcher, and likewise does not count changes in game states due to wild pitches positively for a batter.

Once an expected run value is synthesized from the current state and the playerElo of the batter and the pitcher, park factor and home field advantage adjustments (if applicable) are made, and the expected run value of the play is then compared to the true run value outcome. The playerElo of both the batter and pitcher are then updated accordingly, dependent on the difference between the true run value and the expected run value. For example, if an excellent pitcher strikes out a mediocre batter, the batter will not lose much Elo, and the pitcher will not gain much Elo. Likewise, if a below-average batter does extremely well against a top pitcher, there will be a far greater change in the Elo of both players. Errors are also taken into account and will prevent a positive run value from counting against a pitcher or positively for a batter.

Refer to the Technical Appendix at the end of the README.md on GitHub for further details regarding the playerElo methodology.

Player Analysis

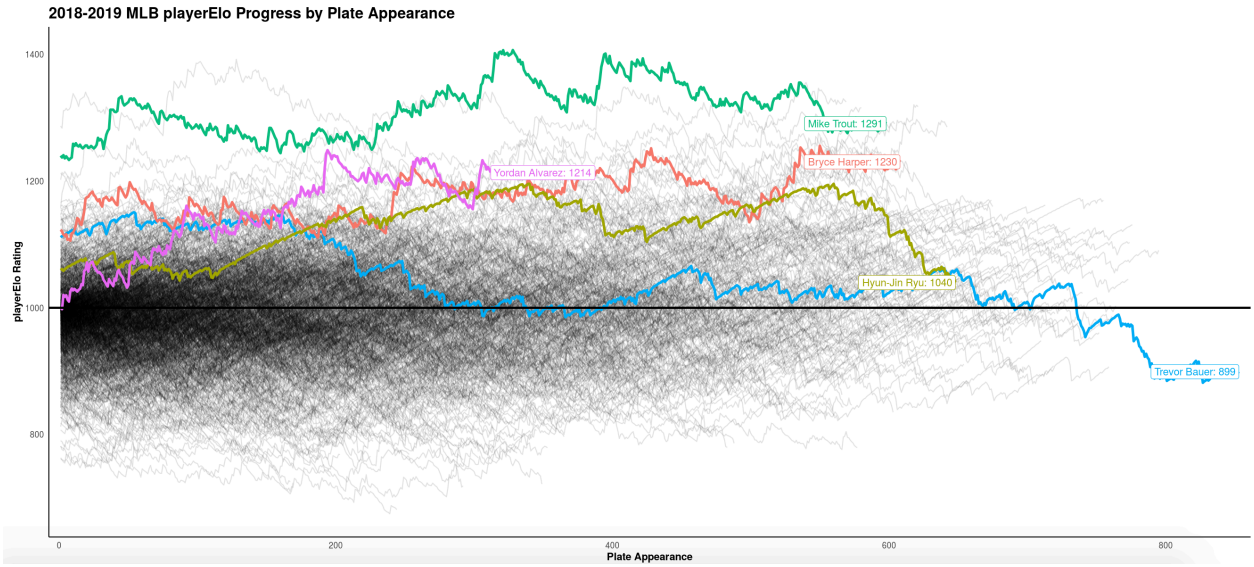
Rank	Name.x	Team.x	Position.x	playerElo.x	Name.y	Team.y	Position.y	playerElo.y
1	Anthony Rendon	Washington Nationals	3B	1370	Justin Verlander	Houston Astros	SP	1195
2	Christian Yelich	Milwaukee Brewers	OF	1356	Jack Flaherty	St. Louis Cardinals	SP	1194
3	Alex Bregman	Houston Astros	SS	1356	Gerrit Cole	Houston Astros	SP	1185
4	Mike Trout	Los Angeles Angels	OF	1333	Mike Clevinger	Cleveland Indians	SP	1183
5	Freddie Freeman	Atlanta Braves	1B	1329	Sonny Gray	Cincinnati Reds	SP	1179
6	Ketel Marte	Arizona Diamondbacks	2B	1324	Max Scherzer	Washington Nationals	SP	1175
7	Bryce Harper	Philadelphia Phillies	OF	1269	Zack Britton	New York Yankees	RP	1174
8	Cody Bellinger	Los Angeles Dodgers	1B	1264	Kirby Yates	San Diego Padres	RP	1170
9	Yordan Alvarez	Houston Astros	OF	1262	Yusmeiro Petit	Oakland Athletics	RP	1162
10	Anthony Rizzo	Chicago Cubs	1B	1257	Taylor Rogers	Minnesota Twins	RP	1158
11	Max Muncy	Los Angeles Dodgers	1B	1250	Felipe Vazquez	Pittsburgh Pirates	RP	1157
12	Pete Alonso	New York Mets	1B	1250	Giovanny Gallegos	St. Louis Cardinals	RP	1155
13	Nelson Cruz	Minnesota Twins	DH	1248	Shane Bieber	Cleveland Indians	SP	1154
14	Yuli Gurriel	Houston Astros	1B	1247	Aroldis Chapman	New York Yankees	RP	1153
15	Juan Soto	Washington Nationals	OF	1245	Josh Hader	Milwaukee Brewers	RP	1149
16	George Springer	Houston Astros	OF	1243	Jacob deGrom	New York Mets	SP	1149
17	David Freese	Los Angeles Dodgers	3B	1242	Brandon Workman	Boston Red Sox	RP	1146
18	Matt Olson	Oakland Athletics	1B	1231	Eduardo Rodriguez	Boston Red Sox	SP	1145
19	Josh Bell	Pittsburgh Pirates	1B	1225	Tyler Clippard	Cleveland Indians	RP	1145
20	Edwin Encarnacion	New York Yankees	1B	1211	Aaron Bummer	Chicago White Sox	RP	1142
21	Wilson Ramos	New York Mets	C	1207	Liam Hendriks	Oakland Athletics	RP	1138
22	Jeff McNeil	New York Mets	2B	1206	Will Harris	Houston Astros	RP	1137
23	Nolan Arenado	Colorado Rockies	3B	1206	Adam Ottavino	New York Yankees	RP	1133
24	Justin Turner	Los Angeles Dodgers	3B	1203	Scott Oberg	Colorado Rockies	RP	1129
25	Mookie Betts	Boston Red Sox	OF	1203	Mike Minor	Texas Rangers	SP	1129

It is interesting to note Nolan Arenado and Edwin Encarnacion do particularly well in the model, even with park factor adjustments. This is can be attributed to the difficulty of schedule of the Rockies and Yankees, facing the tough pitching competition in the NL West and AL East respectively. The average pitching Elo faced by Arenado and Encarnacion is 1010.5 and 1009.6 (20th and 25th highest overall). Quality of contact does leave room to be desired for Arenado, however playerElo does not incorporate statistics like exit velocity and launch angle in its calculations, and thus the model is a better reflection of on-field performance than underlying swing metrics. In contrast to Arenado and Encarnacion, Yordan Álvarez has played incredibly since

called up in June but has faced some of the easiest competition in the league, with an average Elo faced of only 986.8.

Pitching wise, the breakout performances of Gray, Gallegos, and Bieber are all captured and supported by the playerElo model. Verlander recently wrenched the pitching Elo throne away from Scherzer following his no-hitter and several dominant outings. Mike Minor ranks particularly high after seeing success against the high-powered offense of the Astros (3rd in runs scored), Athletics (7th), Angels (13th), and Mariners (17th), and could have been an impact pitcher if dealt at the deadline.

In contrast, playerElo has little faith in Trevor Bauer, one of the most discussed pitchers at the deadline and for whom the Reds paid a good price for. Bauer is ranked 166th among SPs with more than 100 batters faced, with a playerElo of 919 after beginning the season with a preseasonElo of 1111. He has struggled mightily against relatively easy divisions of the AL Central and NL Central, posting an xFIP of 4.40 against an average Elo faced of 979.5. We can visualize Bauer's season long downward playerElo trend within the context of the league with the following graphic. Displayed below is the playerElo trends of the entire MLB for 2019, with specific batters and pitchers highlighted. The bold line on both graphs denotes the average playerElo.



The graph illustrates Mike Trout has dominated since the start of the year and Harper has consistently been great and continues to rise in playerElo. Alvarez elevates from a rookie with a playerElo of 1000 to the top tier of players. On the pitcher side, Hyun-Jin Ryu was briefly the best pitcher in the game, as ranked by playerElo, for several starts, but has since come back to Earth hard with a sharp decline. Bauer's long decline is evidenced throughout the season, with his playerElo reaching all-time lows over his past few starts.

Team Analysis

Rank	Team	Batting Elo	Pitching Elo	Team Elo	Runs Scored	Runs Against
1	Houston Astros	1095	1047	1068	823	589
2	Los Angeles Dodgers	1066	1049	1056	792	554
3	Washington Nationals	1046	1053	1049	747	645
4	Oakland Athletics	1055	1026	1039	756	611
5	New York Yankees	1048	1027	1036	828	662
6	Atlanta Braves	1034	1037	1035	772	665
7	St. Louis Cardinals	1010	1054	1034	654	579
8	Cleveland Indians	987	1060	1027	677	570
9	Minnesota Twins	1038	1020	1027	828	645
10	New York Mets	1040	1013	1025	690	669
11	Boston Red Sox	1024	1011	1017	800	718
12	Chicago Cubs	1003	1020	1012	696	618
13	Tampa Bay Rays	983	1028	1008	670	567
14	Arizona Diamondbacks	1033	986	1007	726	652
15	Pittsburgh Pirates	1036	963	996	686	785
16	Milwaukee Brewers	997	993	995	669	695
17	Cincinnati Reds	963	1018	994	637	634
18	Philadelphia Phillies	1009	978	992	689	701
19	San Diego Padres	975	1003	990	618	682
20	San Francisco Giants	987	990	989	617	686
21	Los Angeles Angels	1000	969	984	707	762
22	Colorado Rockies	999	958	977	732	853
23	Chicago White Sox	952	984	970	607	739
24	Texas Rangers	916	1011	968	702	749
25	Kansas City Royals	967	965	966	606	757
26	Miami Marlins	980	953	966	536	705
27	Toronto Blue Jays	945	966	957	619	716
28	Seattle Mariners	964	944	954	684	810
29	Baltimore Orioles	941	943	943	614	879
30	Detroit Tigers	907	932	921	516	801

Team Batting and Pitching Elo power rankings can also be created from playerElo, with weights for each individual playerElo determined by the plate appearances of the batter or innings pitched of the pitcher. To account for differences in variance, z-scores are used to weight team Batting and Pitching Elo to create an aggregate Team Elo. Recall playerElo does not take into account defense or stolen bases, which will influence the ratings slightly. Speedster teams such as the Royals, Rangers, and the Mariners will rank slightly lower than their true abilities, and conversely the Padres and Cubs, who have had atrocious defensive performances thus far, will rank slightly higher

than true value. However, the Team Elo rankings still give a fairly accurate view of the current state of the MLB, and particularly highlight upstart teams such as the Nationals and Athletics that could be even better than they have seemed thus-far.

Sabermetric Context of playerElo

As playerElo only evaluates historical season performance to-date, and is not a prediction statistic, comparing playerElo with underlying swing metrics can enable accurate forecasts as to potential second-half decliners, and identify first-half breakouts that are for real. The below heat-map style table displays the top 40 players ranked by Expected Weighted On-Base Average (xwOBA), per Baseball Savant, with playerElo, Exit Velocity (EV), and Hard-Hit Percentage (HH%) also displayed. By evaluating the four statistics in collectively, it is easy to identify players such as J.D. Davis, Justin Smoak, C.J. Cron, and Gary Sanchez, whose quality of contact has been excellent, but have performed worse contextually or faced relatively easier pitching than other batters with similar quality of contact. Additionally, players such as Trout, Yelich, Rendon, Freeman, Bregman, and Marte have performed extremely well, and their contextual performance and level of competition backs up their success.

Name	Team	Position	EV	HH%	xwOBA	playerElo
Mike Trout	Los Angeles Angels	OF	90.7	44.3	0.454	1333
Cody Bellinger	Los Angeles Dodgers	1B	90.9	46.3	0.435	1264
Christian Yelich	Milwaukee Brewers	OF	93.1	49.1	0.423	1356
Anthony Rendon	Washington Nationals	3B	90.6	47.7	0.422	1370
Yordan Alvarez	Houston Astros	OF	92.6	50.3	0.421	1262
Nelson Cruz	Minnesota Twins	DH	93.5	51.4	0.417	1248
J.D. Martinez	Boston Red Sox	OF	91.5	47.1	0.410	1190
Juan Soto	Washington Nationals	OF	91.0	47.2	0.409	1245
Howie Kendrick	Washington Nationals	2B	92.1	50.0	0.407	1138
Matt Olson	Oakland Athletics	1B	91.9	52.0	0.398	1231
Mookie Betts	Boston Red Sox	OF	90.9	46.8	0.398	1203
George Springer	Houston Astros	OF	90.0	45.8	0.397	1243
Freddie Freeman	Atlanta Braves	1B	89.9	43.0	0.396	1329
Aaron Judge	New York Yankees	OF	95.7	56.3	0.396	1125
Joey Gallo	Texas Rangers	1B	93.0	52.3	0.396	1099
Josh Donaldson	Atlanta Braves	3B	92.8	51.1	0.392	1201
Anthony Rizzo	Chicago Cubs	1B	89.0	36.6	0.388	1257
Max Muncy	Los Angeles Dodgers	1B	89.8	44.0	0.388	1250
Carlos Correa	Houston Astros	SS	89.0	41.8	0.388	1066
Bryce Harper	Philadelphia Phillies	OF	91.4	46.3	0.386	1269
Pete Alonso	New York Mets	1B	90.7	41.9	0.384	1250
David Freese	Los Angeles Dodgers	3B	91.1	49.1	0.384	1242
Jorge Soler	Kansas City Royals	OF	92.4	49.2	0.384	1165
Marcell Ozuna	St. Louis Cardinals	OF	91.8	49.8	0.384	1058
Josh Bell	Pittsburgh Pirates	1B	92.4	47.0	0.383	1225
J.D. Davis	New York Mets	3B	91.4	47.6	0.383	1014
Ronald Acuna Jr.	Atlanta Braves	OF	90.4	46.4	0.381	1170
DJ LeMahieu	New York Yankees	2B	91.7	47.2	0.376	1151
Justin Turner	Los Angeles Dodgers	3B	90.2	43.1	0.373	1203
Mitch Garver	Minnesota Twins	C	91.1	48.8	0.372	1195
Justin Smoak	Toronto Blue Jays	1B	90.3	39.2	0.372	1012
Carlos Santana	Cleveland Indians	1B	92.0	45.7	0.371	1146
Alex Bregman	Houston Astros	SS	89.1	37.3	0.370	1356
C.J. Cron	Minnesota Twins	1B	91.0	44.8	0.370	999
Ketel Marte	Arizona Diamondbacks	2B	89.9	40.2	0.368	1324
Gary Sanchez	New York Yankees	C	91.2	41.6	0.368	992
Kyle Schwarber	Chicago Cubs	OF	92.4	49.9	0.367	1130
Rafael Devers	Boston Red Sox	3B	92.2	47.7	0.366	1157
Luke Voit	New York Yankees	1B	89.6	40.4	0.366	1136
Yoan Moncada	Chicago White Sox	2B	92.6	47.1	0.366	1124

Application

The playerElo system is able to reveal characteristics of the game and performance current metrics miss. For example, the context-dependent nature and run value calculations of playerElo appropriately credits a reliever who gets three outs without allowing a run after the previous pitcher loaded the bases, and similarly splits the credit of those runs being allowed between a pitcher who loaded the bases, and a reliever who failed to get the necessary outs to end the inning.

Additionally, playerElo does not assign extra value to players who simply come up to the plate in high-powered offenses who often come up to the plate with runners on, nor docks players who simply do not get the same opportunities as their peers to accumulate as many counting statistics such as RBIs by nature of playing for a bad team.

The difference between the playerElo rankings when run context-dependent, treating at-bats differently depending on the runs-out state, and context-free, meaning it treats all at-bats with equal weights of importance, highlight the important of including context in player evaluation. The greatest negative differences of rankings among the two models are all of hitters on the best offenses in the league: Yuli Gurriel, Justin Turner, and J.D. Martinez. The playerElo system accurately captures the favorable situations these hitters consistently bat in, and appropriately does not allow this to overinflate their value. Conversely, Alex Gordon of the atrocious Royals offense, rises significantly from, as does Renato Nunez from the Orioles.

But most importantly, the fundamental aspect of the playerElo system is that the quality of competition is factored into player analysis. Once again, consider the comparison between Freeman and Santana, who have nearly identical statistics but clearly face different levels of pitching competition.

	PA	**wOBA**	**xwOBA**	**wRC+**	**playerElo**
Freeman, 1B	643	0.398	0.396	144	1329
Santana, 1B	624	0.389	0.371	141	1146

There is now a distinguishing factor between the two players, reflected by Freeman's 5th best overall playerElo among batters versus Santana's 55th overall playerElo.