Chapman University

Chapman University Digital Commons

Mathematics, Physics, and Computer Science Faculty Articles and Research

Science and Technology Faculty Articles and Research

5-19-2021

Pitcher Effectiveness: A Step Forward for In Game Analytics and Pitcher Evaluation

Christopher Watkins

Vincent Berardi

Cyril Rakovski

Follow this and additional works at: https://digitalcommons.chapman.edu/scs_articles

Part of the Databases and Information Systems Commons, Numerical Analysis and Scientific Computing Commons, and the Other Computer Sciences Commons

Pitcher Effectiveness: A Step Forward for In Game Analytics and Pitcher Evaluation

Comments

This article was originally published in *Mathematics and Sports*, volume 2, issue 1, in 2021.

Mathematics and Sports

Pitcher Effectiveness: A Step Forward for In Game Analytics and Pitcher Evaluation

Christopher Watkins[1,*], Vincent Berardi[2], Cyril Rakovski[3]

[1]Computational and Data Sciences Program, Chapman University, Orange, CA, [2]Psychology Department, Chapman University, Orange, CA, [3]Mathematics Department, Chapman University, Orange, CA

[*]Corresponding Author E-mail:
watki115@mail.chapman.edu

Abstract

With the introduction of Statcast in 2015, baseball analytics have become more precise. Statcast allows every play to be accurately tracked and the data it generates is easily accessible through Baseball Savant, which opens the opportunity for improved performance statistics to be developed. In this paper we propose a new tool, Pitcher Effectiveness, that uses Statcast data to evaluate starting pitchers dynamically, based on the results of in-game outcomes after each pitch. Pitcher Effectiveness successfully predicts instances where starting pitchers give up several runs, which we believe make it a new and important tool for the in-game and post-game evaluation of starting pitchers.

Keywords: Baseball, MLB, Statcast, Analytics, Prediction

1 Introduction

1.1 Evolution of Statistics in Baseball

Baseball has quickly become one of the most analyzed sports with significant growth in the last 20 years (Koseler and Stephan 2017) with an enormous amount of data collected every game that requires professional teams to have a state-of-the-art analytics team in order to compete in today's game. As an example, the Houston Astros lost over 100 games each season from 2011-2013. In 2014, sports writer Ben Reiter predicted the Astros to win a World Series sooner rather than later, in 2017, because of the advanced analytics team they built (Reiter 2014). The Houston Astros ended up winning the 2017 World Series, as predicted by Mr. Reiter. In 2019 the Tampa Bay Rays brought Jonathan Erlichman, who had not played baseball past T-ball, into the dugout as "the first full-time analytics coach ever to join a major-league staff. In his new role, he will use his knowledge of data to assist manager Kevin Cash with in-game decisions and provide real-time information to players" (Diamond 2019). With an analytics minded approach, the Tampa Bay Rays made it to the 2020 World Series (Juan Toribio 2019). These examples highlight the role of data analytics within this sport.

The analysis of baseball has evolved over the years, with three major categories being sequentially developed. First, there was the traditional statistics such as home runs, batting average, and earned run average. Next, baseball statisticians developed Sabermetrics which further improved the analyses of player performance. The Sabermetrics movement led to the development of new statistics. This concept was pioneered by Bill James in the 1980's and defined Sabermetrics as "the search for objective knowledge about baseball" (Society for American Baseball Research 2019). Some examples of Sabermetric statistics include on-base percentage (OBP), on-base plus slugging (OPS), and wins above replacement (WAR). This started the Moneyball movement, where teams analyzed players differently than in the past with general manager Billy Beane of the Oakland Athletics leading the charge (Lewis 2003).

The final category of statistics was made possible by the introduction of Statcast in 2015, which revolutionized Major League Baseball (MLB Advanced Media 2019b). Statcast tracks every single play on the MLB field and "allows for the collection and analysis of a massive amount of baseball data, in ways that were never possible in the past" (MLB Advanced Media 2019b). The technology that makes this possible "is a combination of two different tracking systems – Trackman Doppler radar and high definition Chyron Hego cameras. The radar, installed in each ballpark in an elevated position behind home plate, is responsible for tracking everything related to the baseball at 20,000 frames per second. This radar captures pitch speed, spin rate, pitch movement, exit velocity, launch angle, batted ball distance, arm strength, and more. Separately, each ballpark also has a Chyron Hego camera system, where six stereoscopic cameras are installed in two banks of three cameras apiece down the foul line. The camera system tracks the movement of the people on the field, which allows for the measurement of player speed, distance, direction, and more on every play" (MLB Advanced Media 2019b). This new data from Statcast is easily accessible through Baseball Savant (Willman 2019). Proper analyses of these high precision, multidimensional data should be able to provide in-game analytics to coaches that would enable them to better evaluate player performance in real-time.

1.2 Decision to Remove a Starting Pitcher

One of the most difficult decisions a MLB manager has to make is when to remove a starting pitcher. Remove a starting pitcher too early, you do not maximize his use and risk overworking relief pitchers. Remove a starting pitcher too late when he is fatigued, and he will likely give up many runs and/or place your relief pitchers in difficult situations. A real-time predictive model could help the manager make the optimal decision during the game. One method that was proposed in 2017 (Soto-Valero, González-Castellanos, and Pérez-Morales 2017) considered pitches as time series data and used dynamic time warping and 1-nearest neighbor to classify the outing on an ongoing basis. Using the linear weights for all possible plays and count, the metric Linear Run Pitcher's Performance (LRPP) was built as a rolling sum to classify the performance as High Performance (HP) or Low performance (LP). When the result of 10 pitches were unknown, precision, recall and F1 values (prevalent binary classification measure ranging from 0 [inaccurate] to 1 [accurate])(Wood) had means 0.9, 0.8, and 0.89 and with 30 pitches unknown the F1 value was 0.78 (Soto-Valero, González-Castellanos, and Pérez-Morales 2017).

Another approach that had been taken was building a regression model that used past inning at bats, game situation, and historical data to predict Pitcher's Total Bases (PTB) for the following inning; a cut-off value was then used to determine if the pitcher should be taken out (Gartheeban and Guttag 2013). After the PTB model made a prediction, it was compared to a manager model, which was built from actual manager decisions (Gartheeban and Guttag 2013), which predicted the manager's decision correctly for 95% of the innings. Results suggested that the PTB model performed well; when the manager decided to leave the pitcher in and the PTB model agreed, 17.7% of the innings the pitcher gave up at least one run. In contrast, when the manager left a starting pitcher in and the PTB model disagreed, 31.5% of the innings the pitcher gave up at least one run (Gartheeban and Guttag 2013).

In 2017 Harrison and Salmon also addressed the question of when to remove a pitcher from the game by using a system that used pitch counts and strike-to-ball ratio (STB) (Harrison and Salmon 2017). A linear model, that regressed balls and strikes, was built to represent the expected (or average) strike-to-ball ratio for a pitcher (Harrison and Salmon 2017). The authors found that at high pitch counts there was a drop in the mean STB line with a unique number of pitches where this occurred for each pitcher, which they called a trigger point (Harrison and Salmon 2017). They proposed that a pitcher should be removed at this trigger point and managers using knowledge of the mental and physiological state of the pitcher could improve the decision making (Harrison and Salmon 2017).

The above-discussed models used to aid with the decision to remove a starting pitcher did not use the rich Statcast data that is now available. In this work we designed a a novel tool, Pitcher Effectiveness, that can be used to evaluate a starting pitcher on both an in-game and overall outing basis with its unique pitch by pitch structure. Pitcher Effectiveness is unique in comparison to other metrics because it does not take runs scored into account but is designed as a predictor of runs. In addition, unlike other evaluation tools, Pitcher Effectiveness used both pitch level event data and at-bat level event data. The goal of Pitcher Effectiveness was to measure how effective a pitcher is by only taking into account the variables that the pitcher can control. For example, a pitcher cannot control the defense so they should not be evaluated on runs caused by errors, but they do control working ahead of the count. Also, a pitcher who made a great pitch, with soft contact, but resulting in a hit should not be penalized because of a defensive shift or the hit falling between two fielders. Of course, baseball is a game where events are dependent on more than just the starting pitcher, but the starting pitcher has the biggest influence on the game. Using Statcast data of variables the pitcher can control, Pitcher Effectiveness was continuously calculated after each pitch to generate a rolling sum throughout the game. This paper discusses the data that was used, methods, results, and future work involved with Pitcher Effectiveness.

2 Data

The analyzed data was obtained from Baseball Savant and included pitchers that threw 2000 or more pitches in 2018 MLB season. This data set included mostly traditional starting pitchers, but there were also swing pitchers who made multiple spot starts and the opener was used by several teams in 2018. We considered every pitch that these pitchers threw and examined variables such as pitch speed, post-pitch score, fielding alignment, launch angle and exit velocity among others. This resulted in 115 pitchers, and 305,633 pitches from the query. We removed 1,150 pitches because of missing values produced by an error with Statcast, which left 304,483 total pitches. The analysis was restricted to 10 relevant variables that the pitcher had control over, plus 2 new variables were created using transformations of the original ones. These new variables that were created included Pitcher Effectiveness and slope. A detailed description of these variables is shown in Table 1.

Table 1: Description of variables (* New Variables)

Variable	Description
Player Name	Pitcher's Name
Events	Ball in play event
Description	Result of pitch
Balls	Number of balls in at bat
Strikes	Number of strikes in at bat
Launch Speed	Exit velocity of batted ball
Post Bat Score	Opposing team score post-pitch
Runs	Number of runs given up after the pitch
Pitches	Number of pitches thrown by the pitcher
Run Prediction	Binary variable for x of more runs given up in the next 5 or 10 pitches $x \in \{0, 1, 2, 3, 4\}$
Pitcher Effectiveness*	New tool created
Slope*	Fitted slope of Pitcher Effectiveness over the previous 5 or 10 pitches

3 Methods

Using 2018 Statcast data, the tool Pitcher Effectiveness was calculated for each pitch which resulted in a time series. A pitcher started with a Pitcher Effectiveness of zero and the value was continuously updated with each pitch throughout the game, until the pitcher was taken out of the game. This was based on three classes of variables, as seen in Table 2. The three variables used to calculate the Pitcher Effectiveness were event, ball and strike count and exit velocity. Using count, swinging strike, and exit velocity as events was important because they are indicators of a pitcher's performance. It has been shown that when a pitcher falls behind in a count the hitter has an advantage and performed better in those situations (Gray 2002). Swinging strikes have recently become tracked and are an indicator of a pitcher's ability to miss bats which influences performance (Willman 2019). Finally, we used exit velocity to combat defensive factors out of the pitcher's control, such as a bloop single due to defensive alignment, while penalizing a pitcher for getting lucky with an event like a hard hit line drive to an infielder. We decided on using 95+ mph exit velocity and 80 mph or less exit velocity as events based on hard-hit classification by Statcast(MLB Advanced Media 2020b) and Baseball Savant (Willman 2019). Each outcome for these three variables had a weight, and the sum of these weights for each pitch determined that pitch's contribution to Pitcher Effectiveness. The weights used for a single, double, triple, home run, walk, and hit by pitch were linear weights for calculating weighted on base average (wOBA) for the 2018 season (Fangraphs 2019). The linear weights for wOBA were found by calculating the run expectancy for each event using the data from the 2018 season and were taken from Fangraphs (Fangraphs 2019). All other weights (count, out, swinging strike, and exit velocity) were carefully chosen and can be found in Table 2. An extensive grid search of values between 0.1 to 1.5 by 0.2 (8 values, 4096 combinations) was implemented to find the optimal weights. The combination with the best area under the ROC curve (AUC) was chosen. Although finding a minimum is generally a continuous problem, a grid search is commonly used for optimizing weights. The computations took over 72 hours to complete on a 32 core computer. As an example of how Pitcher Effectiveness was calculated, let's say a pitcher had a Pitcher Effectiveness of 1 at a certain point of the game. On the next pitch the pitcher gives up a single with an exit velocity of 100 MPH while being ahead of the count. Then the updated Pitcher Effectiveness score would be: 1 – 0.88 (Single) - 0.5 (Exit Velocity 95+) + 0.5 (Ahead of the count) = 0.12. The higher the Pitcher Effectiveness, the better the pitcher was doing overall. Negative values for Pitcher Effectiveness indicate that a pitcher was ineffective. Summary statistics of final Pitcher Effectiveness scores for pitchers' entire games can be found in Table 3.

We hypothesized that trends in Pitcher Effectiveness would be a more useful predictive tool than the value associated with a single pitch. Therefore, for each pitch, a linear regression model was used to estimate the trend (i.e. slope) over the previous 5 pitches; this was repeated for the previous 10 pitches as well. As a result of this procedure, the first 4 or 9 pitches from each game were not considered. The data set used was sufficiently large so that the predictive power of Pitcher Effectiveness was not compromised.

A model was designed to predict a particular number of runs or more given up by a pitcher using the Pitcher Effectiveness score and the change in this score (slope) over a certain number of pitches as covariates.

For the run prediction variable, we examined several run-based outcomes including any number of runs, more than 1 run, more than 2 runs, and more than 3 runs given up in the next 5 or 10 pitches. This procedure meant that the last 5 or 10 pitches were ignored for each game respectively because there was nothing to predict when the pitcher was taken out of the game. Again, due to the large data, we didn't expect these omissions to materially affect predictive ability.

We used the presence and absence of runs as the outcome variable and Pitcher Effectiveness and slope of the recent performance trend as predictor variables in a logistic regression model. We used 5-fold cross validation to assess the performance of the model.

Table 2: Pitcher Effectiveness Weights

Event	Weight
Single	-0.88
Double	-1.25
Triple	-1.58
Home Run	-2.03
Walk	-0.69
Hit by pitch	-0.72
Swinging Strike	+0.5
Out (except sacrifice fly)	+0.5
Other Events (i.e. Foul ball, error, etc.)	+0
Count	
Ahead of count	+0.5
Behind count	-0.5
Exit Velocity	
95+ mph exit velocity	-0.5
80 mph or less exit velocity	+0.5

Table 3: Summary Statistics for Pitcher Effectiveness per Game

Statistic	Value
Minimum	-16.38
Quantile 1	6.93
Mean	15.01
Quantile 3	22.77
Maximum	55.18

We identified the best predictive model by comparing a range of potential models built using different combinations of the number of pitches used to calculate slope, the number of pitches used for run prediction, and the number of runs to predict. In addition, Pitcher Effectiveness by itself was compared to Game Score, to evaluate a pitcher's outing. The grid search tested for the highest cross validated area under the ROC curve for the particular pitch and run prediction combination of 4+ runs, 10 pitches for slope calculation, and 5 pitches for run prediction. The logistic model used can be found below.

$$log-odds(Run\ Prediction) \sim Pitcher\ Effectiveness + Slope \qquad \text{(Logistic Regression Model)}$$

4 Results

The models did very well to predict a big inning, 3+ or 4+ runs scored, within the next 5 or 10 pitches. The best combinations can be found in Table 4 with the 5-fold cross validated area under the curve (CV AUC).In addition, we reported the odds ratio for the variables in the logistic regression model. A 1-unit increase in one of these variables changes the odds of the event happening. For example, in line one of Table 4, a 1-unit increase in slope decreases the odds of 3+ runs given up in the next 5 pitches by about 78% (1-0.218). Both the Pitcher Effectiveness and slope variables were statistically significant (p-values of less than 0.001) in the models.

The accuracy of the models decreased in predicting 1+ or 2+ runs scored in the next 5 or 10 pitches. All CV AUC values were less than 0.7 for these combinations. There are a few reasons why this was the case. First, it takes much less to score 1-2 runs even if a pitcher was effective. For example, a pitcher could be doing well all game but miss a location once and give up a home run. Another example could be a pitcher giving off a leadoff double and a run scoring without another hit (i.e. Combination of moving the running via groundout or flyout and scoring on a groundout or sacrifice fly, etc.). Next, errors could cause runs to be given up by a pitcher, although unearned. Both Pitcher Effectiveness and the slope did not take errors into account, which meant a pitcher could still be effective, but the defense caused a run to be scored. Thus, predicting a larger number of runs was more successful. Since this model was predictive of giving up a large amount of runs, this could be a useful tool for indicating when a pitcher should be removed, but requires further analysis to determine criteria for removal and testing against a manager's decision.

To evaluate a pitcher's performance for a game, many look at the traditional pitching line to see the innings pitched, strikeouts, number of hits, runs, walks, and home runs given up by a pitcher. In a search for one number to describe a pitcher's performance, Bill James developed the metric Game Score in the 1980's (MLB Advanced Media 2019a). Each pitcher began with a score of 50, then their score would change depending on the play and associated weight from Table 5. In 2014 Tom Tango updated these weights to correlate

Table 4: Logistic Regression Model Results

CV AUC	Slope Odds Ratio	Pitcher Effectiveness Odds Ratio	Number of pitches for slope	Number of pitches for run prediction	Number of runs
0.716	0.218	0.964	5	5	3+
0.751	0.149	0.97	5	5	4+
0.723	0.206	0.97	5	10	4+
0.778	0.051	0.962	10	5	4+
0.743	0.09	0.975	10	5	3+
0.711	0.126	0.979	10	10	3+
0.744	0.078	0.982	10	10	4+

more with a pitcher's talent level (Tom Tango; MLB Advanced Media 2019a). The major difference from Bill James' formulation is starting from a score of 40, instead of 50, and taking home runs into account. Associated weights for Tom Tango's formula are in Table 5. A Game Score of 50 is an average performance for a pitcher and a Game Score of 40 indicates a replacement level outing (MLB Advanced Media 2019a). Both variations of Game Score fail to take pitch level events or Statcast metrics into account. Pitcher Effectiveness improved these deficiencies by utilizing pitch level events of ball-strike count and swinging strike in account. In addition, Pitcher Effectiveness took exit velocity into account as a measure of quality contact.

Table 5: Game Score Weights

Game Score Weights (Bill James)	
Event	**Weight**
Start of game	+50
Out	+1
Inning completed after 4th	+2
Strikeout	+1
Hit	-2
Earned Run	-4
Unearned Run	-2
Walk	-1
Game Score Weights (Tom Tango)	
Start of game	+40
Out	+2
Strikeout	+1
Walk	-2
Hit	-2
Any Run	-3
Home Run	-6

To evaluate Pitcher Effectiveness as a tool to evaluate a starting pitcher's game performance, it was compared to both the traditional pitching line and Game Score. For both the traditional line score and Game Score, data was used from Baseball Reference where they use Bill James' formula for Game Score (Baseball Reference 2019a). The worst and best pitched games according to Pitcher Effectiveness were more closely examined. For example, Dylan Bundy, on May 8th, 2018, had the lowest Pitcher Effectiveness for a game at -16.38 and the trend during the game can be seen in Figure 1. In this game Bundy did not record an out, while giving up 7 runs and a low Game Score of 10 with the pitching line in Table 6. Bundy's Pitcher Effectiveness in this outing agreed with Game Score, being the lowest of the season for him.

The best Pitcher Effectiveness for the season was an outing by multiple Cy Young award winner Max Scherzer on May 30th. He

Table 6: Dylan Bundy 5-8-18 (Baseball Reference 2019b)

Innings Pitched	Hits	Earned Runs	Walks	Strikeouts	home runs	Game Score
0+	5	7	2	0	4	10

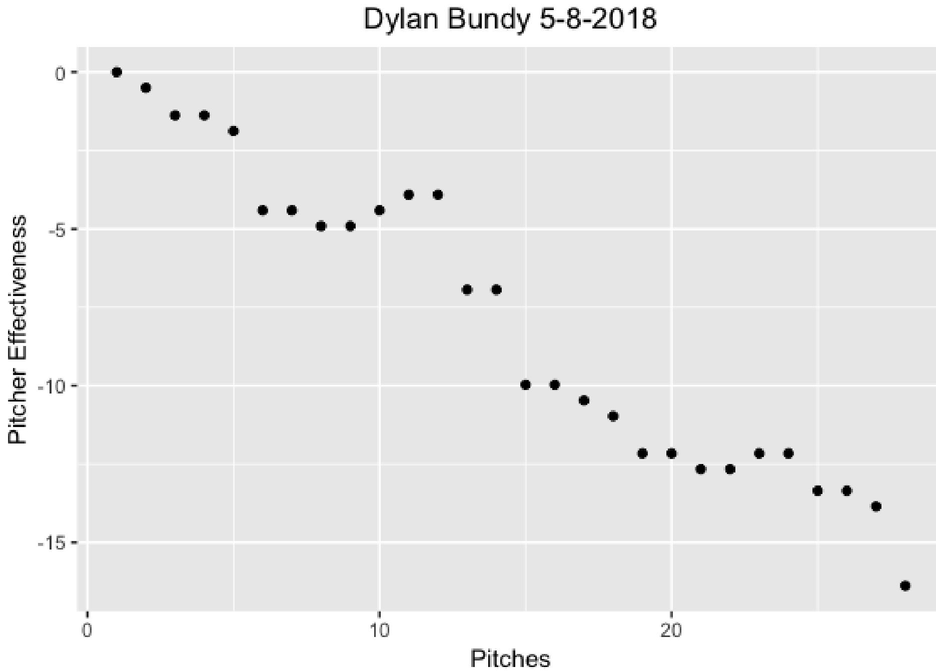

Figure 1: Dylan Bundy Pitcher Effectiveness (5-8-18)

pitched 8 innings giving up no runs, striking out 12, and a Game Score of 89. A small dip in Pitcher Effectiveness, in Figure 2, for Scherzer at around 94 pitches was due to a double by Manny Machado and walk to Mark Trumbo in the 7th inning of the game. According to Game Score, this was not quite the best pitched game by Scherzer but rather the April 9th game with a slightly better Game Score of 93. The pitching lines were similar, as seen in Table 7, but Scherzer pitched a shutout and did not walk anyone. However, the difference comes from Game Score valuing innings pitched after the 4th and Pitcher Effectiveness valuing pitchers missing bats with swing a miss strikes. In the May 30th game Scherzer had 11 of 12 strikeouts swinging while in the April 9th game he had 8 out of 10 strikeouts swinging.

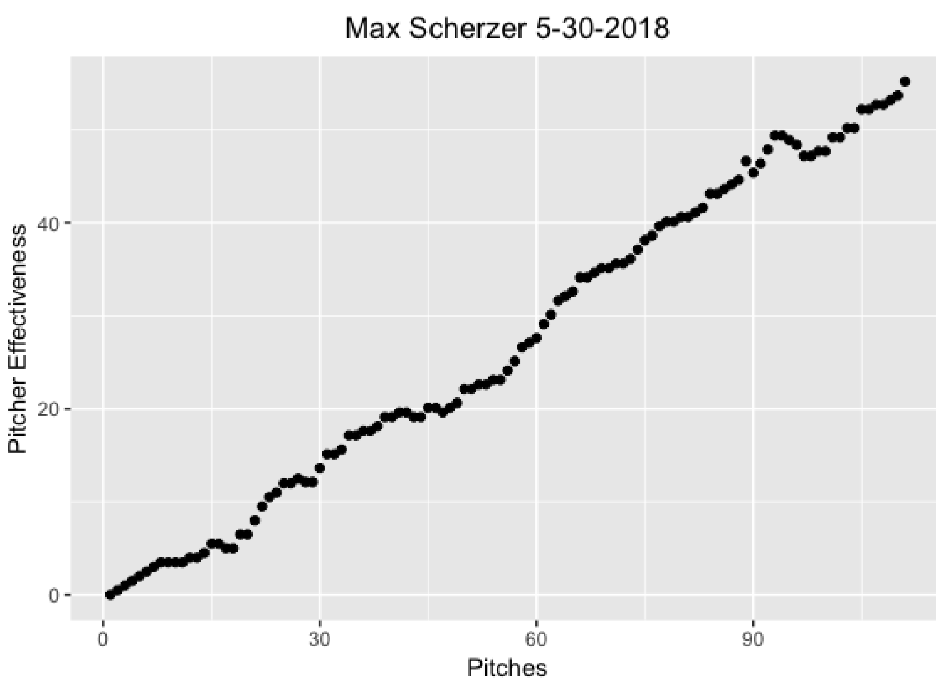

Figure 2: Max Scherzer Pitcher Effectiveness (5-30-18)

Table 7: Max Scherzer 5-30-18 (Baseball Reference 2019c)

Date	Innings Pitched	Hits	Earned Runs	Walks	Strikeouts	home runs	Game Score
5-30-18	8	2	0	1	12	0	89
4-9-18	9	2	0	0	10	0	93

5 Conclusion

In this paper Pitcher Effectiveness has been shown to be able to predict a pitcher giving up many runs, and a viable tool for evaluating a starting pitcher's performance for a game. Utilizing Statcast data, Pitcher Effectiveness gave a different evaluation of pitchers using data we never had before. Coaches could look back at the game pitch by pitch to find times the Pitcher Effectiveness dropped. This is a step forward for in game analytics, that already involves scouting reports available in the dugout, and the evaluation of how effective a pitcher is.

6 Future Work

There are many areas of future work involving Pitcher Effectiveness. First, the weights and variables can be better optimized for better predictive power and evaluation of pitchers. A wider grid search could be performed to find more optimal values. In addition, choosing the weights as whole numbers rather than some as fractional may be better for simplifying it. Statcast data is very new and the weights need further exploration. Also, both slope and number of pitchers for run prediction could be better optimized. There may be a better number of pitches to examine for both variables that would make the prediction better. From a front office perspective, Pitcher Effectiveness could also be used to find pitchers that may be undervalued. This would help teams with a limited budget to build a more competitive team.

The most exciting possibility for future work is improving the utilization of Pitcher Effectiveness for use by managers. The model could be compared to a manager's decision by looking at what the model is predicting versus what the manager did, similar to the Gartheeban and Guttag analysis (Gartheeban and Guttag 2013). For example, the model could predict many runs given up in the next 10 pitches, the manager leaves the pitcher in, and the pitcher gives up runs. Pitcher Effectiveness could extend the work of Harrison and Salmon (Harrison and Salmon 2017) by building a trend line of the average Pitcher Effectiveness per pitch for particular pitchers. Pitcher Effectiveness above the trend line could indicate good performance while under the trend line would indicate bad performance.

Next, exploring Pitcher Effectiveness for use with relief pitchers is an area of future research. The data set that was taken for this paper involved mainly traditional starting pitchers, with a few exceptions. However, getting ahead of the count, missing bats, and getting outs are important for a relief pitcher as well. One may argue that it is more important, especially in tight ballgames. A relief pitcher will pitch one to two innings maximum, in general, in a game. For in game analytics for relief pitchers the goal is to avoid giving up any runs, or few runs depending on the score of the game. With this in mind and less innings pitched by relief pitchers, a coach would want to know very quickly if they are at risk of giving up runs. This means the slope variable would have to use less pitches since a relief pitcher may pitch 20-25 pitches in a game. Also, evaluating relief pitchers on their outing using Pitcher Effectiveness would be different than starting pitchers. Their ceiling for Pitcher Effectiveness is much lower than a starting pitcher who are in the game longer. Either an adjustment of giving each relief pitcher an initial Pitcher Effectiveness or separating the evaluation of a starting and relief pitcher could be solutions to this problem. Looking at how a relief pitcher's Pitcher Effectiveness is affected by the number of days off would also be interesting to explore.

Finally, Pitcher Effectiveness can be adapted to incorporate additional efficiency metrics such as quality of contact. In the last couple years pitchers have been told to elevate the fastball as hitters are adapting their launch angle for the best contact. Pitchers that have been traditionally effective lower in the strikezone must adapt to this new trend. This could be captured using Statcast's barrel metric (MLB Advanced Media 2020a). Rewarding pitchers who induce non-quality contact could be added to Pitcher Effectiveness.

Acknowledgement

The authors would like to thank Alex Barrett for assisting with the parallelization of the grid search code.

References

Baseball Reference. 2019a. *Baseball Reference.* https://www.baseball-reference.com. Accessed: 2019-3-12.

Baseball Reference. 2019b. *Dylan Bundy.* https://www.baseball-reference.com/players/gl.fcgi?id=bundydy01&t=p&year=2018. Accessed: 2019-3-12.

———. 2019c. *Max Scherzer.* https://www.baseball-reference.com/players/gl.fcgi?id=scherma01&t=p&year=2018. Accessed: 2019-3-12.

Diamond, Jared. 2019. "The MLB Coach Who Played Only T-Ball." *The Wall Street Journal,* March 11, 2019. https://www.wsj.com/articles/the-mlb-coach-who-played-only-t-ball-11552312800.

Fangraphs. 2019. *Guts!* https://www.fangraphs.com/guts.aspx?type=cn. Accessed: 2019-3-12.

Gartheeban, Ganeshapillai, and John Guttag. 2013. "A data-driven method for in-game decision making in MLB," 973. https://doi.org/10.1145/2487575.2487660.

Gray, Rob. 2002. ""Markov at the Bat": A Model of Cognitive Processing in Baseball Batters." *Psychological Science* 13 (6): 542–547. ISSN: 09567976, 14679280. http://www.jstor.org/stable/40063736.

Harrison, Willie K, and John L Salmon. 2017. "Bullpen Strategies for Major League Baseball," 1–16.

Juan Toribio. 2019. *Rays headed to WS.* Accessed: 2020-10-28, October 18, 2019. https://www.mlb.com/news/rays-win-alcs-game-7-advance-to-world-series.

Koseler, Kaan, and Matthew Stephan. 2017. "Machine Learning Applications in Baseball: A Systematic Literature Review." *Applied Artificial Intelligence* 31 (9-10): 745–763. ISSN: 10876545. https://doi.org/10.1080/08839514.2018.1442991.

Lewis, Michael. 2003. *Moneyball: The art of winning an unfair game.* W.W. Norton.

MLB Advanced Media. 2019a. *Game Score.* http://m.mlb.com/glossary/advanced-stats/game-score. Accessed: 2019-3-12.

———. 2019b. *Statcast.* http://m.mlb.com/glossary/statcast. Accessed: 2019-3-12.

———. 2020a. *Barrel.* Accessed: 2020-8-28. https://www.mlb.com/glossary/statcast/barrel.

———. 2020b. *Hard-hit Rate.* http://m.mlb.com/glossary/statcast/hard-hit-rate. Accessed: 2020-8-28.

Reiter, Ben. 2014. "Your 2017 World Series Champs." *Sports Illustrated,* 30. https://doi.org/https://www.si.com/longform/astros/index.html.

Society for American Baseball Research. 2019. *A guide to Sabermetrics Research.* https://sabr.org/sabermetrics. Accessed: 2019-3-12.

Soto-Valero, César, Mabel González-Castellanos, and Irvin Pérez-Morales. 2017. "A predictive model for analysing the starting pitchers' performance using time series classification methods." *International Journal of Performance Analysis in Sport* 17 (4): 492–509. ISSN: 14748185. https://doi.org/10.1080/24748668.2017.1354544.

Tom Tango. *Game Score Version 2.0.* Accessed: 2020-8-28. https://blogs.fangraphs.com/instagraphs/game-score-v2-0/.

Willman, Daren. 2019. *Baseball Savant.* https://baseballsavant.mlb.com/statcast_leaderboard. Accessed: 2019-3-12.

Wood, Thomas. *F-Score Definition.* https://deepai.org/machine-learning-glossary-and-terms/f-score. Accessed: 2020-8-28.