AN EXPECTED OUTCOME FRAMEWORK FOR EVALUATING BATTING AND PITCHING PERFORMANCE IN MAJOR LEAGUE BASEBALL WITH APPLICATIONS TO THE "JUICED BALL" AND THE "FLY BALL REVOLUTION"

by

TAYLOR SMITH

(Under the Direction of L. Jason Anastasopoulos)

ABSTRACT

I utilize Major League Baseball Statcast data from 2015-2017 to build batted ball classifiers using state-of-the-art gradient boosting trees in conjunction with hyperparameter optimization techniques. Visual and numeric summaries of the model results are used to glean insights into batted balls in MLB. Further, the model framework is used to create new batting and pitching metrics with demonstrated advantages over previously used metrics. Using the batted ball classifiers and the introduced metrics, I investigate the "Juiced Ball" and "Fly Ball Revolution" phenomena in MLB, quantify the respective impacts of both phenomena, and present a manner for evaluating batter and pitcher performance across different ball environments.

INDEX WORDS: sabermetrics, juiced ball, Major League Baseball, hyperparameter optimization, gradient boosting trees, baseball statistics

# AN EXPECTED OUTCOME FRAMEWORK FOR EVALUATING BATTING AND PITCHING PERFORMANCE IN MAJOR LEAGUE BASEBALL WITH APPLICATIONS TO THE "JUICED BALL" AND THE "FLY BALL REVOLUTION"

by

TAYLOR SMITH

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2018

AN EXPECTED OUTCOME FRAMEWORK FOR EVALUATING BATTING AND PITCHING PERFORMANCE IN MAJOR LEAGUE BASEBALL WITH APPLICATIONS TO THE "JUICED BALL" AND THE "FLY BALL REVOLUTION"

by

TAYLOR SMITH

Major Professor: L. Jason Anastasopoulos
Committee: Nicole Lazar
Jaxk Reeves

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2018

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

The first two decades of the twenty-first century have seen an explosion of the usage of data across many industries; baseball has been no exception. While baseball has a rich and storied relationship with statistics, there has been a proliferation of statistical analysis devoted to the game of baseball both among the teams and in the public sector. It is now common for most Major League Baseball teams to employ at least one or two personnel devoted to analytics and several teams have devoted Research & Development departments [1]. Since the publication of *Moneyball* [2] and the movie of the same name, baseball analysis has taken off in the public world with websites such as Fangraphs, The Hardball Times, and Baseball Prospectus devoted to analytics in baseball. Even the most casual of fans are becoming familiar with the world of sabermetrics, the term used to define the statistical analysis of baseball, as sabermetric concepts are becoming increasingly used by broadcasters and writers.

The pace at which baseball digests information is only increasing. While baseball has long had radar guns and the pitch tracking system PITCHf/x since 2006 [3,4], in the 2015 season Major League Baseball began using a Doppler radar and high-definition video system called Statcast to track nearly every aspect of the game. Examples of the kinds of measurements that Statcast records are the exit velocity and launch angle of batted balls; velocity, break, and spin rate of pitches; and sprint speed and first step quickness for both fielders and baserunners [5–7].

One goal of this work is that it will be accessible to both statistical and baseball audiences. For this reason, in the second chapter, we provide a primer on key concepts in the statistical analysis of baseball. Various metrics for both batting and pitching performances are introduced along with their strengths and weaknesses as assessments of player performance. We also present the concept of linear weights which is essential to the remainder of the thesis. Lastly, we present two new metrics: Expected Batting Runs Above Average (xbRAA) and Expected Pitching Runs Above Average (xpRAA) that will be used heavily in the remainder of our work. These two new metrics are compared with alternatives already existing in the public sphere.

In the third chapter, we more fully describe Statcast data. We outline some of the ways that it has been used in the public sabermetric community and also look at some strengths and shortcomings of the system. In particular, we outline the data collection and cleaning process used to obtain the subset of the Statcast data that will be employed here.

In the fourth chapter, we provide a high-level overview of the statistical techniques used to perform the analysis. The first sub-chapter overviews LightGBM, a gradient boosting tree method that is heavily utilized in the ensuing analysis. The second and third sub-chapters overview Bayesian optimization and tree-structured Parzen estimators, two hyperparameter optimization techniques used here to select an optimal set of hyperparameters for the LightGBM models.

The fifth chapter details the two versions of the LightGBM batted ball classifiers, one that utilizes exit velocity, vertical launch angle, batter handedness, year, and half of season as predictors and a second that includes the aforementioned variables along with a proxy to horizontal angle. We begin by presenting the methodology and results of the hyperparameter optimization routines. From there, we present the model results in both a numeric and visual fashion with a particular emphasis on using the results to visualize and better understand the dynamics of batted balls in Major League Baseball.

Using the two batted ball classifiers, we present two versions of xbRAA and xpRAA. We compare and contrast the two versions of the metrics against each other according to properties that have been deemed desirable in the sabermetric community. We demonstrate that the introduced metrics have more desirable empirical properties than the traditional metrics, in addition to the philosophical superiority discussed in the first chapter. We also use the introduced metrics to examine some of the best and worst performers and those with the largest gaps between their expected performance and realized performance.

In the seventh chapter, we present a new framework for examining the "Juiced Ball" hypothesis and the so-called "Fly Ball Revolution". Over the last three years, offensive performance in Major League Baseball has increased precipitously. Currently, the sabermetric community believes that this is the product of two forces: first, the balls used by Major League Baseball have changed in some capacity, leading to balls that come off the bat faster and/or fly further; second, that batters across the MLB have been intentionally seeking to hit balls in the air more in hopes of hitting more extra-base hits. We begin the section by providing an overview of the research done on the juiced ball hypothesis up to this point. We then present a framework where through counterfactual analysis, we are able to quantify how much of the change in the offensive environment can be attributed to the change in the ball used by Major League Baseball.

# Chapter 2

# Baseball Analytics Background

## 2.1 Basic Definitions

For completeness, we overview a few key baseball terms that are necessary for understanding the remainder of this section and the rest of the thesis as a whole. A plate appearance refers to a completed turn batting. An at-bat is any plate appearance that does not end with a walk, a hit-by-pitch, a sacrifice bunt, a sacrifice fly, interference or obstruction, or if a hitter is removed and replaced with another batter (except in the case of being replaced with two strikes in which the replacement subsequently strikes out). While the official definition of a hit is rather extended, a hit occurs when a batter makes contact with the ball such that it lands in fair territory and the batter reaches base without the aid of a defensive player making an error. There are four kinds of hits: singles, doubles, triples, and home runs. The name refers to how far the batter advances on a hit: single means the batter got to first base, a double means to second, a triple to third base, and a home run all the way back around to home. Home runs usually occur by hitting the ball with such force that the ball clears the outfield wall in fair territory. One other definition that will be needed below is that of an earned run. An earned run simply refers to when an offensive player is able to get to home plate without the aid of an error by the defensive team. All of these definitions and more can be found in the official rulebook of Major League Baseball [8].

## 2.2 Batter Statistics

### 2.2.1 Batting Average

Batting average is probably the most recognizable of all baseball statistics, and most readers are likely familiar. It is simply defined as the number of hits over the number at-bats.

$$\text{BA} = \frac{H}{AB}$$

While it is a very popular statistic, it is a deeply flawed measure of a player's offensive talent. First, it completely ignores both walks and HBPs. This makes very little sense as walks are valuable as they avoid an out and get a runner on base. A second major flaw with batting average is that it weights all hits the same. Even the most casual observer of the game can see that this is a poor metric for that reason. Obviously, a home run is worth more than a single, but determining how much more remains to be seen.

### 2.2.2 Slugging Percentage

$$\text{SLG} = \frac{\text{1B} + 2 \times \text{2B} + 3 \times \text{3B} + 4 \times \text{HR}}{\text{AB}}$$

One attempt at improving on batting average is slugging percentage. Slugging percentage attempts to address the flaw in batting average of weighting all types of hits the same. As opposed to assigning a weight of one to each hit type, slugging percentage assigns a weight corresponding to the number of total bases associated with that kind of hit. While this is certainly an improvement over batting average. It is still flawed in that it also ignores walks and HBPs. Additionally, slugging percentage assumes that the value of hits relative to each other can be seen in the number of total bases they provide. We will see that this is, in fact, a poor assumption.

### 2.2.3 On-Base Percentage

$$\text{OBP} = \frac{H + BB + HBP}{AB + BB + HBP + SF}$$

Whereas slugging percentage addresses the problem of equally weighting all hits, on-base percentage corrects the problem of ignoring walks and HBPs. On-base percentage effectively measures how good players are at avoiding outs and equivalently getting on-base. The proper evaluation of on-base percentage has been one of the hallmarks of the sabermetric mindset. Baumer and Zimbalist [9] showed that the degree to which teams have valued on-base skills has steadily climbed from the 1980s and 1990s through the early 2010s with a spike occurring after the publication of *Moneyball* [2].

### 2.2.4 OPS

$$\text{OPS} = \text{OBP} + \text{SLG}$$

On-base percentage more properly evaluates the ability to avoid making outs and getting on-base, while slugging percentage acknowledges that different kinds of hits should be valued differently. A very popular way to combine the strengths of these two measurements separately is to simply add them together. While a crude technique that lacks a natural interpretation on account of the difference in denominators, OPS has been shown to be a very effective measurement of offensive performance in MLB. As is seen in the Table 2.2.1 taken from Baumer and Zimbalist, we can see that of the measurements considered thus far that team

Table 2.2.1: Correlation with Runs Scored

| Statistic | Correlation |
|---|---|
| Batting Average | 0.822 |
| Slugging Percentage | 0.910 |
| On-Base Percentage | 0.885 |
| OPS | 0.946 |

OPS has the strongest correlation to the total number of team runs scored from 1954-2011 [9].

While it is a crude measure, it overall does a very good job of capturing offensive performance. However, it lacks a natural interpretation, and it still fails to properly assign values relative to one another.

### 2.2.5 Linear Weights

The motivating idea behind linear weights is to assign the proper value to each event that occurs on the baseball field. The idea can be summarized as weighting events according to the average change in expected runs that an event provides. Early versions of linear weights were suggested by Lindsey [10] and were brought to great popularity in the sabermetric community by Palmer [11]. Tango has continued to popularize and advance linear weights usage in the public sabermetric community in [12–15].

In order to understand linear weights, one must first understand the run expectancy matrix. At any moment in any half of an inning, the game is in one of twenty-four base out states: there can be zero, one, or two outs and there are three bases that may or not be occupied. The run expectancy for that base-out state can be calculated by considering all such times that a team was in that base-out state and then taking the average number of runs scored over the remainder of the inning. This gives rise to a matrix like the one below taken from Fangraphs [16]:

Table 2.2.2: Run Expectancy Matrix

| Runners | No Outs | One Outs | Two Outs |
|---|---|---|---|
| Empty | 0.461 | 0.243 | 0.095 |
| 1 _ _ | 0.831 | 0.489 | 0.214 |
| _ 2 _ | 1.068 | 0.644 | 0.305 |
| 1 2 _ | 1.373 | 0.908 | 0.343 |
| _ _ 3 | 1.426 | 0.865 | 0.413 |
| 1 _ 3 | 1.798 | 1.140 | 0.471 |
| _ 2 3 | 1.920 | 1.352 | 0.570 |
| 1 2 3 | 2.282 | 1.520 | 0.736 |

For a particular event, we can calculate its run expectancy based on the twenty-four base-out states (RE24). This is defined as the sum of the difference between the run expectancy after the event has ended and the starting run expectancy and the number of runs scored on the play.

$$\text{RE24} = \text{RE of Ending State} - \text{RE of Starting State} + \text{Runs Scored}$$

To calculate a linear weight for a particular event one simply averages the RE24 values for all such events. For example, to find the linear weight for a single, we take the average of the RE24 values for all singles.

While the concept of linear weights can be extended to a pitch-by-pitch basis, the RE24 version will suffice here. We utilize the linear weights that Fangraphs reports for the 2015 season [16].

### 2.2.6 RAA

The advantage of linear weights is that it allows us to assign proper weights to different events based on their impact on the expected run values. This allows sabermetricians to design metrics that are both strongly interpretable and properly capture the value of different events which previous attempts such as slugging percentage and on-base percentage failed to do.

The first such metric that we present is batting runs above average (bRAA). For all plate appearances that end in an official at-bat, an unintentional walk, a sacrifice fly, or a hit by pitch, bRAA is defined as the the linear weight of that event. The name comes from the fact that the average plate appearance would add zero runs to the run expectancy by construction, thus considering the linear weight represents how much above or below average that outcome was than an average plate appearance. Thus the sum of all such plate appearances represents the total number of runs that a batter was above or below average over the course of the season. While bRAA itself is a counting statistic, it is easily turned into a rate statistic by dividing by the number of relevant plate appearances.

$$\mathrm{bRAA} = -.26 \times \mathrm{Out} + .29 \times \mathrm{uBB} + .31 \times \mathrm{HBP} + .44 \times \mathrm{1B} + .74 \times \mathrm{2B} + 1.01 \times \mathrm{3B} + 1.39 \times \mathrm{HR}$$

$$\mathrm{AbRAA} = \frac{-.26 \times \mathrm{Out} + .29 \times \mathrm{uBB} + .31 \times \mathrm{HBP} + .44 \times \mathrm{1B} + .74 \times \mathrm{2B} + 1.01 \times \mathrm{3B} + 1.39 \times \mathrm{HR}}{\mathrm{AB} + \mathrm{BB} - \mathrm{IBB} + \mathrm{SFF} + \mathrm{HBP}}$$

The choice of what plate appearances are considered was inspired by wOBA [12]. Plate appearances not included consist of situations in which the batter was not given a chance to show his offensive value. For example, if a batter is intentionally walked, while that may demonstrate he is viewed as a dangerous batter (or that the batter behind him in the order is much worse), it does not serve as a demonstration of his offensive skill in its own right. Similarly, if a batter attempts a sacrifice bunt, he was intentionally trying to get out in order to advance a runner, and it does not accurately reflect his offensive ability.

### 2.2.7 wOBA

As mentioned above, weighted on-base average (wOBA) [12] was introduced by Tango and is one of the most popular linear weight based statistics in the sabermetric community. It relays the same information as AbRAA does, but it has been put on a different scale in order to be more interpretable to a larger audience. The reasoning goes that common fans have very little intuition for run values, so wOBA is put onto a scale with which fans will be more familiar. The weights are shifted to be relative to an out and then multiplied by a constant in order to get wOBA to have the same scale as OBP.

$$\text{wOBA} = \frac{.69 \times \text{uBB} + .72 \times \text{HBP} + .89 \times 1\text{B} + 1.27 \times 2\text{B} + 1.62 \times 3\text{B} + 2.10 \times \text{HR}}{\text{AB} + \text{BB} - \text{IBB} + \text{SFF} + \text{HBP}}$$

It is our opinion that this shifting and scaling is very detrimental. First, as a result, wOBA has no natural interpretation. While the values of events relative to each other have been preserved, they are no longer anchored to an actual baseball meaning. Similarly, the sum of wOBA no longer has an interpretation as the sum value of a batter's contribution above average.

## 2.3 Pitching Statistics

### 2.3.1 ERA

$$\text{ERA} = 9 \times \frac{\text{ER}}{\text{IP}}$$

Earned run average (ERA) is a very simple metric that has traditionally been used to evaluate pitcher performance. It is defined as the number of earned runs a pitcher has given up divided by the number of innings he has pitched. It is then multiplied by nine since there are nine innings in a regulation baseball game, thus it can be interpreted as the number of runs we expect the pitcher would give up in a whole nine-inning game.

### 2.3.2 FIP

In the early 2000s, the sabermetric community became aware of the fact that outcomes on balls in play were largely luck driven and were not something that most pitchers could control well [17]. However, the things that were largely in the pitchers' control were the number of home runs they give up, the number of walks and hit by pitches they allow, and the number of batters they strikeout. This has led to the popularity of statistics such as Fielding Independent Pitching (FIP) introduced by Tango [18].

$$\begin{aligned} \text{FIP} &= \frac{13 \times \text{HR} + 3 \times (\text{BB} + \text{HBP}) - 2 \times \text{K}}{\text{IP}} + C \\ C &= \text{League ERA} - \frac{13 \times \text{League HR} + 3 \times (\text{League BB} + \text{League HBP}) - 2 \times \text{League K}}{\text{League IP}} \end{aligned}$$

In a similar manner to wOBA, FIP is intentionally put back onto the scale of a more familiar metric, ERA. While it only considers home runs, walks, hit by pitches, and strikeouts, it has been shown to be a better predictor of future ERA than past ERA [19,20].

### 2.3.3 xFIP

$$\begin{aligned} \text{xFIP} &= \frac{13 \times (\text{Fly Balls} \times \frac{\text{LgHR}}{\text{Lg Fly Balls}}) + 3 \times (\text{BB} + \text{HBP}) - 2 \times \text{K}}{\text{IP}} + C \\ C &= \text{League ERA} - \frac{13 \times \text{League HR} + 3 \times (\text{League BB} + \text{League HBP}) - 2 \times \text{League K}}{\text{League IP}} \end{aligned}$$

Expected Fielding Independent Pitching (xFIP) assumes that not only is preventing damage on balls in play not a skill for pitchers but also that luck plays a large role in the number of home runs that pitchers give up. As opposed to including the number of home runs, xFIP estimates the number of home runs that should have been given up by multiplying the number of flyballs the pitcher allowed by the league average home run per flyball ratio. While this may seem like an oversimplifying assumption, it has been shown to be a better predictor of future ERA than both ERA and FIP [19,20].

### 2.3.4 SIERA

The last of the ERA estimators outlined here is Skill-Interactive Earned Run Average. There are two common versions of SIERA: one produced by Baseball Prospectus and one produced by Fangraphs [21,22]. Here we outline the Fangraphs version developed by Swartz. For a more thorough treatment see Swartz's series of articles where he fully explains the metric [20,23–26].

From a statistical standpoint, SIERA is the fitted value of a linear model with ERA as the response variable and a selection of expertly chosen pitching statistics used as explanatory variables. In addition to being a better predictor of future ERA than the other ERA estimators discussed [19,20], the model provides strong insight into the dynamics of effective pitching. While particular interpretations of the coefficients are not discussed here, we include the model in Table 2.3.1 to give readers intuition on how the model works.

### 2.3.5 Slash Line Against

$$\text{BAA/OBPA/SLGA}$$

While ERA and its estimators are a very popular set of pitching statistics, they are not always the most appropriate. For one, all the estimators are on the scale of the number of earned runs we could expect them to give up over a nine-inning game. This is somewhat more

Table 2.3.1: SIERA Table

| | Coefficient |
|---|---|
| Constant | 5.534 |
| $K/PA$ | -15.518 |
| $(K/PA)^2$ | 9.146 |
| BB/PA | 8.648 |
| $(BB/PA)^2$ | 27.252 |
| netGB/PA | -2.298 |
| $\pm(netGB/PA)^2$ | -4.92 |
| $K/PA \cdot BB/PA$ | -4.036 |
| $K/PA \cdot netGB/PA$ | 5.155 |
| $BB/PA \cdot netGB/PA$ | 4.546 |
| Year Coeff. | -.02 - .289 |
| Percent Innings as SP | 0.367 |

[1] netGB is the difference between the number groundballs and flyballs allowed

appropriate for starting pitchers who routinely pitch multiple innings; however, one could argue that it might be more appropriate to report the number of runs to be given up over six or seven innings as opposed to nine as it is now very rare for a starting pitcher to throw a complete game. For relievers, who routinely pitch just an inning or possibly only to a couple of batters, the ERA interpretation is not particularly natural. In many situations, it would be better to know how we expect a pitcher to perform against a single batter. Thus we have a collection of "against" statistics: batting average against, on-base percentage against, etc. Essentially any rate offensive statistic can be turned into a measurement of pitcher quality by examining the batter values of that offensive statistic against that pitcher. A commonly used version of this principle is to report the "slash line" of batting average, on-base percentage, and slugging percentage against a pitcher. The strengths and weaknesses of these metrics as an evaluation of pitcher performance are the same as they are for measures of offensive performance but further complicated by the pitcher's dependence on his defense.

### 2.3.6 pRAA, ApRAA, and wOBAA

In the same way that batting average, on-base percentage, and slugging percentage can be adapted to be measurements of pitching performance, we can adapt the linear weight-based metrics to be measures of pitcher performance. We define pitching runs above average to be the batting runs above average against and analogously for average batting runs above average and weighted on-base average.

## 2.4 Proposed Metrics

The goal of baseball statistics is very often to quantify how good we think a particular player is at a specific skill. Batting average measures how good a batter is at getting hits, on-base percentage measures how good a batter is at getting on-base, and bRAA provides a catch-all measure of how good a player is with a bat in his hand. The common shortcoming with all these metrics is that they are all limited to what actually happened and do not consider what could have or should have happened. Said a different way, these statistics hope to capture what a player's true underlying skill level is, but traditional statistics capture two sources of randomness: randomness coming from the random variation in player performance and randomness coming from environmental factors outside of the player's control. For example, consider a hypothetical MLB game between the Tampa Bay Rays and the Seattle Mariners at Seattle's Safeco Field. Suppose that Seattle's slugging designated hitter, Nelson Cruz, hits a ball hard to straight-away centerfield against Tampa's Chris Archer. Due to the climate and the dimensions, Safeco Field is known for suppressing fly balls. Further, suppose the Rays' centerfielder, Kevin Kiermaier who is one of the premier defensive baseball players in all of baseball, races back to the wall, leaps, and makes a great catch to rob Cruz of a home run. According to traditional statistics, Cruz gets no more credit for that batted ball than had he hit a shallow flyball. In the same way, Archer is not punished by traditional statistics for giving up what should have been a home run had it been hit at an average park and had he not had one of the best defensive players in baseball playing behind him. These problems give rise to a class of statistics called expected statistics which are generally denoted by appending a lower case "x" to the front of the name. We have already discussed a fairly naive one in xFIP. Here we propose the addition of several such metrics.

### 2.4.1 xbRAA and AxbRAA

We propose a metric that will here be called expected batting runs above average (xbRAA). For non-batted balls, the xbRAA associated with an event is the linear weight associated with the event. However, for batted balls, the xbRAA is the dot product of the linear weights for batted ball events (out, single, double, etc) with the probability vector that a batted ball is an out, single, double, etc based on the way that it was hit. As with the non-expected version, we only consider plate appearances that are an official at-bat, unintentional walks, sacrifice flies, and hit by pitches.

$$\text{xbRAA} = \begin{cases} .29 & \text{uBB} \\ .31 & \text{HBP} \\ -.26 & \text{K} \\ -.26 \times \text{P(Out)} + .44 \times \text{P(1B)} + .74 \times \text{P(2B)} + 1.01 \times \text{P(3B)} + 1.39 \times \text{P(HR)} & \text{Batted Ball} \end{cases}$$

Similarly, to evaluate performance on a plate appearance by plate appearance basis we define the average denoted here as AxbRAA.

$$\text{AxbRAA} = \frac{\sum_N \text{xRAA}}{N}$$

$$N = \text{AB} + \text{uBB} + \text{SFF} + \text{HBP}$$

The advantage of a statistic such as this can clearly be seen in the above example illustrated above. In that example, even though the ball was caught, Nelson Cruz's batted ball would have had a very high xbRAA associated with it as it would have had a very high probability of being a home run in a neutral setting. xbRAA removes the randomness that can be attributed to factors such as weather, park dimensions, and the quality of defensive players in the field. It instead gives a better measurement of the skill of interest which is the ability to hit the baseball well and in such a manner that it produces runs. However, this metric now has the concern of how to estimate the probability of each of the possible events, a problem that is addressed in the fourth chapter. It should be noted that similar metrics have been proposed in the past. An expected weighted on-base average, expected batting average, and expected slugging percentage are provided in the Statcast data [5]. However, we believe that the formulation we present offers several advantages over the xwOBA as provided by Statcast. First, we apply state-of-the-art classification techniques to estimate the probabilities of the batted ball types and use cross-validation to ensure that overfitting does not occur; whereas it is unclear how the xwOBA version is calculated. Secondly, we provide the probability estimates for each batted ball outcome in addition to just presenting the xbRAA. These probabilities are useful for investigating a number of baseball phenomenon as will be shown here. Further providing the probability estimates allows one to create a variety of expected statistics such as a true expected FIP or an expected on-base percentage. Lastly, we believe that bRAA is philosophically superior metric to wOBA for the reasons presented above, thus xbRAA is superior to xwOBA in our opinion.

### 2.4.2 xpRAA and AxpRAA

We also introduce expected pitching runs above average and a per batter faced version of it. The advantages of these metrics over traditional metrics are the same as for the batting versions.

# Chapter 3

# Statcast Data

The primary source of data used here consists of publicly available Statcast data provided by MLB Advanced Media (MLBAM). Statcast is a hybrid radar-camera system implemented in all Major League Baseball stadiums beginning in 2015. The system uses a Doppler radar-based system to track the velocity, spin rate, release location, etc of pitches and the exit velocity, launch angle, and hang time. The camera part of the system tracks the location of every player on the field allowing measurements such as their route efficiency, sprint speed, and first step quickness [5–7]. The system presents a potential gold mine of information allowing essentially all aspects of baseball to be measured and then analyzed. However, with any project as bold and large-scale as this, there are difficulties to be expected. The Statcast system has been no exception. While it has revolutionized baseball, there have been problems with large measurement errors [27,28]. Nevertheless, Statcast has equipped both the teams and the public to better understand the sport of baseball than ever before.

While the player location and the more granular data is only available to MLBAM and the front offices of Major League Baseball teams, the pitch level data has been made available to the public through baseballsavant [29]. Utilizing Bill Petti's baseballr package [30] we scraped every pitch from the regular MLB regular season from 2015 to 2017. From there we filtered to only pitches on which an outcome from a plate appearance was reached. Plate appearances not pertinent to the analysis (interference, intentional walks, and sacrifice bunts) were also filtered out. This left us with a dataset consisting of nearly 550,000 plate appearances of which more than 380,000 were batted balls. Each observation consists of an event outcome, the batter, the pitcher, as well other identifying variables. For batted ball observations the additional fields of exit velocity/launch speed (the speed with which the ball comes off the bat), the vertical launch angle, and the coordinates on the field in which the ball was fielded or landed in the case of home runs. From the fielding coordinates, we derived a proxy to the horizontal launch angle off the bat by measuring the horizontal angle from home plate to that coordinate on the field.

Secondary sources of data were taken from Fangraphs Leaderboards [31]. This consisted of traditional yearly baseball data where each observation corresponds to a particular player in a particular year along with their statistics for the year.

# Chapter 4

# Methodological Background

## 4.1 LightGBM

Gradient boosting trees [32] are a very popular set of machine learning algorithms with widely-used implementations including Xgboost [33] and LightGBM [34]. While XGBoost has long been one of the most popular implementations due to its excellent performance and its prominence in machine learning competitions, LightGBM has been gaining in popularity due to its comparable performance and considerable increase in training speed.

LightGBM employs a histogram-based algorithm [35] for finding split points in the construction of the tree. This is done by binning continuous variables and then making splits based on the bins as opposed to the individual values. While an approximation, it is much more efficient in memory consumption and training speed as training now consists of a one-time histogram building step of $\mathcal{O}(n \times p)$ and finding splits from the histogram which is $\mathcal{O}(k \times p)$ as opposed to finding exact split points which is $\mathcal{O}(n \times p)$ at each step where $n$ is the size of the data, $p$ is the number of features, and $k$ is the number of bins used in the histogram. Additionally, LightGBM introduced two new algorithms for speeding up training on datasets with a large number of observations and a large number of features respectively.

Gradient-based One-Side Sampling (GOSS) is introduced as a manner of reducing the number of data instances used in training. Observations with large gradients (instances for which the model is well trained) are kept and random sampling is performed on instances with small gradients (instances for which the model is not well trained) in such a way that the emphasis is placed on under-trained instances without changing the original distribution by much, an assertion defended by mathematical and empirical results. A second method, Exclusive Feature Bundling (EFB), is introduced for reducing the number of features through exploiting the sparsity of features. LightGBM is demonstrated to achieve comparable results to exact methods while offering major speedups.

While the Statcast data here is neither large in the number of observations or in the number of features, we elect to use LightGBM because of its excellent performance and its vast superiority in training times over other methods. This advantage in speed is particularly advantageous because it allows us to perform the hyperparameter optimization much quicker than it would with other methods such as XGBoost.

## 4.2 Bayesian Optimization with Gaussian Process Priors

In Bayesian optimization [38] of machine learning algorithms, the relationship between a model's hyperparameters and its cost function is modeled by a Gaussian process. As additional trials are run, the Gaussian process is updated reflecting the current beliefs about the relationship between the choice of hyperparameters and the model's performance. The posterior distribution after one iteration becomes the prior for the next iteration. An acquisition function is defined to dictate which set of hyperparameter values would be best to use in the next iteration. Common acquisition functions include expected improvement (EI) [38] which chooses the set of hyperparameter values which maximize the expected improvement over the current best trial, probability of improvement (PI) [39] where the set of hyperparameter values that maximizes the probability of improving over the current best iteration are chosen, and confidence bounds methods where the next set of hyperparameter values to be evaluated are chosen based on where the confidence bounds are at their minimum or maximum depending on the goal. In practice, a Bayesian optimization routine usually begins by several rounds of evaluating the machine learning algorithm at random hyperparameter values before proceeding to use the acquisition function for choosing the next set of hyperparameters to be evaluated. For a more rigorous and complete introduction to Bayesian Optimization see [40] on which this overview was based.

Based on the recommendations put forth in [40], we utilize automatic relevance determination (ARD) Matern 5/2 for the kernel of the Gaussian Process and expected improvement as the acquisition function. We utilize GPyOpt [41] for performing the Bayesian optimization of the LightGBM models.

## 4.3 Tree-Structured Parzen Estimators

The tree-structured Parzen estimator [42] assumes a hyperparameter space $\chi$ that follows a tree-structure (for example, choosing how many hidden layers in a neural network, and then choosing the parameters for a particular choice of layers). The algorithm seeks to model the probability distribution of the cost function we are seeking to optimize given the choice of hyperparameters ($p(y|x)$) by modeling $p(x|y)$ and $p(y)$. This contrasts with the Gaussian Process Bayesian Optimization overviewed above which directly models $p(y|x)$. The TPE models $p(x|y)$ by replacing with priors over the hyperparameter space with non-parametric densities. In particular, TPE defines $p(x|y)$ as:

$$p(x|y) = \begin{cases} \ell(x) & \text{if } y < y^* \\ g(x) & \text{if } y \geq y^* \end{cases}$$

where $\ell(x)$ is the density formed by all observed hyperparameter sets with a corresponding loss of less than $y*$ and $g(x)$ formed by the remaining observations. The value $y^*$ is chosen to be some quantile $\gamma$ of the observed $y$ values so that $p(y < y^*) = \gamma$, but a specific model for $p(y)$ is not necessary.

TPE is based on optimizing the expected improvement over $y^*$. It is shown in that:

$$EI_{y^*}(x) = \int_{-\infty}^{y^*} (y - y^*) p(y|x) \propto (\gamma + \frac{g(x)}{\ell(x)}(1 - \gamma))^{-1}$$

Thus to maximize improvement, we choose the next point to evaluate according to $g(x)/\ell(x)$. TPE uses an Adaptive Parzen estimator to estimate $\ell(x)$ and $g(x)$. For each continuous hyperparameter a uniform prior over some interval, a Gaussian prior, or a log-uniform prior is specified. TPE uses an equally-weighted mixture of the specified prior and a Gaussian centered at each evaluated point where the cost is less than $y^*$ for $\ell(x)$ and greater than $y*$ for $g(x)$. The standard deviation of each Gaussian is set to the greater of the distances between a point and its left or right neighbor, but where the distribution is truncated to keep it in a reasonable range. For categorical variables where the prior is a vector of length N with probabilities $p_i$, the posterior is proportional to $Np_i + C_i$ where $C_i$ is the number of times that value was chosen for observations belonging to the set corresponding to $\ell(x)$ or $g(x)$.

We utilize the implementation of TPE provided by hyperopt [43].

# Chapter 5

# Batted Ball Classifier

In order to better understand the behavior of batted balls in Major League Baseball and to obtain outcome probabilities to fuel a class of "expected" metrics namely the xbRAA and xpRAA introduced above, we present two variations of a batted ball classifier using the LightGBM algorithm. The first model seeks to predict whether a batted ball is an out, single, double, triple, or home run based on the launch speed, the vertical launch angle, the handedness of the batter as well as the year and the half of the season which are used to model suspected changes in batted ball dynamics across and within seasons. The second model includes the same variables as well as a proxy to the horizontal launch angle of the batted ball. Similar investigations of Statcast batted ball data has been performed by Arthur and Petti in [44] and [45].

## 5.1 Hyperparameter Optimization

In order to find an optimal set of hyperparameters for the Lightgbm batted ball classifier, we employ the hyperparameter optimization techniques discussed in the previous chapter. We perform a five-fold cross validation where the function that we seek to minimize is the

Table 5.1.1: Hyperparameter Space

| | Space | Bayes Opt Partial Model | TPE Partial Model | Bayes Opt Full Model | TPE Full Model |
|---|---|---|---|---|---|
| Learning Rate | (.01, .3) | 0.0226 | 0.0445 | 0.01 | 0.0225 |
| Number of Leaves | 127, 255, 511, 1023, 2047, 4095 | 127 | 255 | 127 | 127 |
| Min Data in Leaf | 1,...,50 | 1 | 3 | 50 | 46 |
| Min Gain to Split | (0,5) | 0.581 | 1.029 | 1e-06 | 7.157 |
| Min Sum Hessian in Leaf | (0, 10) | 5.879 | 0.435 | 0 | 0.1229 |
| L2 Regularization | (0, 100) | 11 | 1.66 | 15.26 | 99.51 |
| Bagging Fraction | (.5, 1.00, .025) | 1 | 0.825 | 1 | 0.875 |

Table 5.1.2: Hyperparameter Optimization Results

| | Log Loss | Time Per Iteration (sec) |
|---|---|---|
| Bayes Opt Full Model | 0.4247 | 279.72 |
| TPE Full Model | 0.4236 | 114.42 |
| Bayes Opt Partial Model | 0.5392 | 190.00 |
| TPE Partial Model | 0.5391 | 60.42 |

out-of-sample log-loss. For both the full model including the proxy to horizontal launch angle and the partial model, we run three hundred iterations of both the Tree-Structured Parzen Estimator and the Bayesian Optimization algorithm. In Table 5.1.1, we display the hyperparameters that we chose to optimize, the space of searchable values, as well as the optimal values chosen for each of the optimization routines.

In Table 5.1.2, we display the best log-loss found for each optimization routine as well as the average time in seconds per iteration. We have that the tree-structured Parzen estimator out-performed the Bayesian Optimization routine both in terms of log-loss for both model formulations. Additionally, the tree-structured Parzen estimator technique was considerably faster for both model formulations offering speed-ups of 244% and 314% for the full and partial models respectively. For these combined reasons, we elect to use hyperparameter set from the tree-structured Parzen estimator.

Furthermore in Figure 5.1.1, we display the results of the optimization routines with the red line tracing the current minimum log-loss. For the full model specification, we can observe that for the Bayesian optimization there is more variation in the log-losses than there is in

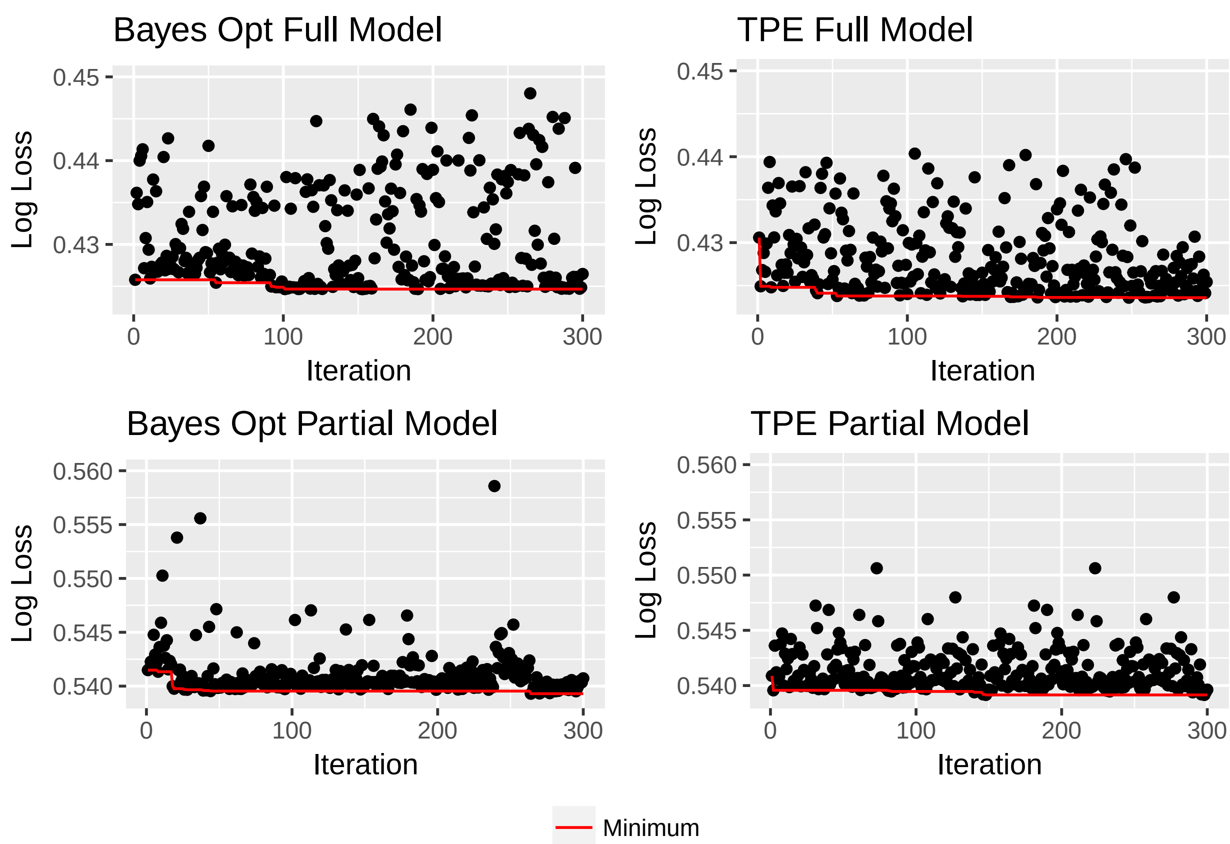


Figure 5.1.1: Hyperparameter Optimization Trials

the TPE log-losses. The opposite seems to be true of the partial model, where the log-losses for the Bayesian optimization model are very tightly grouped with a few outliers.

## 5.2 Model Results

### 5.2.1 Model Performance

In Table 5.2.1, we report the overall model performance on the out-of-sample cross-validation folds for both the full model which includes the horizontal component and the partial model. The full model achieves an overall accuracy of about 83% whereas the partial model achieves an accuracy of about 78%. These both offer very substantial improvements over the null model (which would be to predict an out for all outcomes) which has an accuracy of 67%.

Table 5.2.1: Overall Model Performances

| | Accuracy | Kappa | AccuracyLower | AccuracyUpper | AccuracyNull |
|---|---|---|---|---|---|
| Full Model | 0.829 | 0.644 | 0.828 | 0.830 | 0.67 |
| Partial Model | 0.781 | 0.521 | 0.779 | 0.782 | 0.67 |

Proper calibration of probability predictions is much more important to us than actually predicting the correct label, but nevertheless, the models perform very well in terms of accuracy.

In Tables 5.2.2 and Table 5.2.3, we report various performance metrics by batted ball outcome for both the full model and the partial model respectively. For both models, we see from the F1 score along with other metrics that both models perform the best on recognizing outs and home runs. This is to be expected as home runs are very differentiated from most batted ball types and outs are the most common class. Furthermore, we see that both models have very poor performance on triples. This is very reasonable as triples are very rare events and typically look very similar to doubles off the bat. The biggest improvement from the partial model to the full model is that the full model achieves much better performance on doubles. This makes sense as doubles are very influenced by the horizontal component. Almost all doubles occur either right down the foul lines, the gap between rightfield and centerfield, or between leftfield and centerfield.

## 5.3 Model Visualizations

### 5.3.1 Launch Angle/Exit Velocity Visualizations

In both plots in Figure 5.3.1, we plot the exit velocity of the batted ball along the horizontal axis and the vertical launch angle along the vertical axis. For the left plot, the color is determined by the predicted class for the partial model and in the right plot the color is

Table 5.2.2: Full Model Classification Results by Class

| | Outs | Singles | Doubles | Triples | Home Runs |
|---|---|---|---|---|---|
| Sensitivity | 0.920 | 0.663 | 0.536 | 0.005 | 0.819 |
| Specificity | 0.728 | 0.937 | 0.977 | 1.000 | 0.989 |
| Pos Pred Value | 0.873 | 0.742 | 0.620 | 0.560 | 0.777 |
| Neg Pred Value | 0.817 | 0.910 | 0.968 | 0.993 | 0.992 |
| Precision | 0.873 | 0.742 | 0.620 | 0.560 | 0.777 |
| Recall | 0.920 | 0.663 | 0.536 | 0.005 | 0.819 |
| F1 | 0.896 | 0.700 | 0.575 | 0.011 | 0.797 |
| Prevalence | 0.670 | 0.215 | 0.065 | 0.007 | 0.043 |
| Detection Rate | 0.616 | 0.143 | 0.035 | 0.000 | 0.035 |
| Detection Prevalence | 0.706 | 0.192 | 0.056 | 0.000 | 0.046 |
| Balanced Accuracy | 0.824 | 0.800 | 0.756 | 0.503 | 0.904 |

Table 5.2.3: Partial Model Classification Results by Class

| | Outs | Singles | Doubles | Triples | Home Runs |
|---|---|---|---|---|---|
| Sensitivity | 0.914 | 0.579 | 0.202 | 0.008 | 0.703 |
| Specificity | 0.606 | 0.919 | 0.985 | 1.000 | 0.987 |
| Pos Pred Value | 0.825 | 0.662 | 0.489 | 0.808 | 0.716 |
| Neg Pred Value | 0.776 | 0.889 | 0.947 | 0.993 | 0.987 |
| Precision | 0.825 | 0.662 | 0.489 | 0.808 | 0.716 |
| Recall | 0.914 | 0.579 | 0.202 | 0.008 | 0.703 |
| F1 | 0.867 | 0.618 | 0.285 | 0.016 | 0.709 |
| Prevalence | 0.670 | 0.215 | 0.065 | 0.007 | 0.043 |
| Detection Rate | 0.613 | 0.125 | 0.013 | 0.000 | 0.030 |
| Detection Prevalence | 0.743 | 0.188 | 0.027 | 0.000 | 0.043 |
| Balanced Accuracy | 0.760 | 0.749 | 0.593 | 0.504 | 0.845 |

determined by the xRAA from the partial model. These models offer very interesting insight into the outcomes for batted balls in MLB. We observe that there is a strip of batted balls that are most likely to be singles that range from softly hit balls with a moderately high launch angles to balls hit very hard on a very flat trajectory. Singles, with the exceptions of groundballs hit through holes in the infield, typically land in front of the outfielders but behind the infielders. Thus a ball hit softly with enough loft to get over the infielder but not enough power to get to the outfielders is likely to fall for a single in the same way a ball that is hit very hard but flat will not give the infielder much time to move towards and make a play on the ball but will not carry all the way to the outfielders. We see the same corresponding strip in the xRAA version of the plot.

Predicted doubles occupy a very narrow and short strip situated above the right end of the strip of singles but beneath the clump of home runs. This makes intuitive sense based on our baseball knowledge. As said above doubles usually occur down the foul lines or in the right- or left-centerfield gap. Excluding doubles on groundballs hit hard down the foul lines, a double must carry further than a single in order for it to get down in the gaps between where the outfielders play. Thus doubles will typically have more loft than a single, but not too much that they allow the outfielders sufficient time to get underneath the ball or in the case of hard-hit batted balls go over the wall for a home run.

Home runs are seen in the dark purple clumps in both plots. The home run clump is bordered beneath and slightly to the left by the strip of the doubles which reflects the intuition that home runs are very similar to doubles but they carry further which is either a result of increased exit velocity or launch angle.

Another interesting aspect of batted balls is that as exit velocity increases there is a fanning in of the launch angles. In order to hit a ball at an extreme angle, one must have contacted either near the top or the bottom of the ball. Thus less force will be imparted on the ball and it will leave at lower exit velocity than if the ball had been hit nearer to its center. While

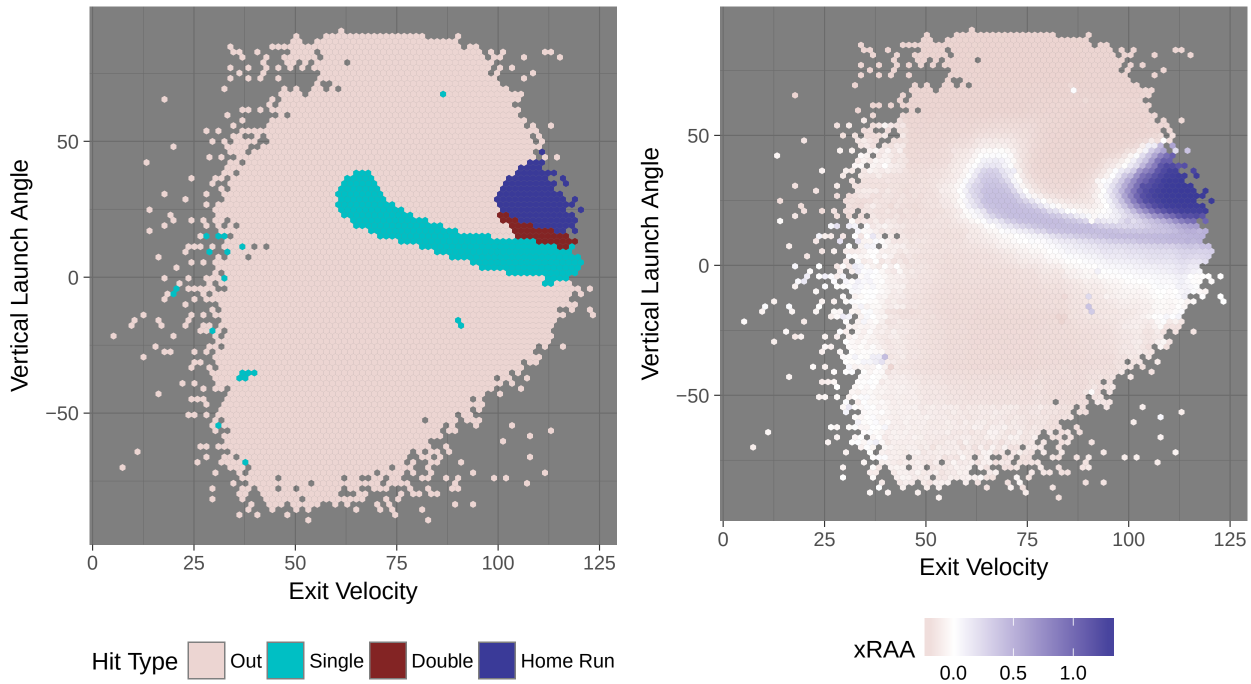


Figure 5.3.1: Batted Ball Outcomes by Exit Velocity and Launch Angle

none of these conclusions are new, visualizing the results in this way provides us with a sanity check to make sure that the predictions are lining up with what we know about baseball and providing a way to quantify what we previously knew qualitatively about batted balls.

### 5.3.2 On-Field Visualizations

Another way of validating that our model corresponds with our prior baseball knowledge is to plot the probability of a batted ball outcome based on the location it was fielded (or in the case of home runs where it landed out of play). These results are displayed in the six-paneled figure 5.3.2 where the full model predicted probability of a hit, out, single, double, triple, and home run are plotted based on the position on the field where they were fielded (or landed out of play). In the first row, we plot the probability of a hit and the probability of an out which are complements of each other. As is expected, we see very high probabilities of outs where we know defensive players are positioned and high probabilities of hits in areas unoccupied by defenders.

In the next row, we plot the probability of a single and the probability of a double based on the batted ball locations. As is expected the probability of single is highest for those balls fielded in front of the outfielders and behind the infielders. We can also see where balls that were hit up the middle, in the hole in between where the third baseman plays and where the shortstop plays, and in the hole between where the first baseman plays and where the second baseman plays. We can also observe that there are balls fielded on the left half of the infield in front of the third baseman with a high probability of a single. These correspond to bunts and so-called swinging bunts that are hit softly and in such a location that the batter is able to beat a throw to first. We see that those kind of results are strongest on the left half of the infield and away from the pitcher because those locations will require a longer and more difficult throw to first base.

As stated above, we see that the probability of a double is strongest down the foul lines and in the two outfield gaps. It is also interesting to note that some of the balls with a higher probability of a double were in fact fielded on the infield indicating groundballs or line drives where the third or first baseman was either positioned more down the line or quickly adjusted to a ball hit down the line. The model also does a very good job of picking up the characteristics of a triple. The high triple probability areas occur in the right-center gap and down the rightfield line, and to a lesser extent in the left-center gap. There is a higher probability of a triple on balls hit to the right half of the field because that requires a longer throw to third base. It is encouraging to see that the model is accurately capturing that dynamic. In the last plot of figure 5.3.2, we are able to validate that the model is properly predicting home runs. It accurately captures how the fences are shorter in left and right field than in center.

Combining these sources of information and weighting them according to their run values yields figure 5.3.3 where plot the full model xRAA of a batted ball based on the location it

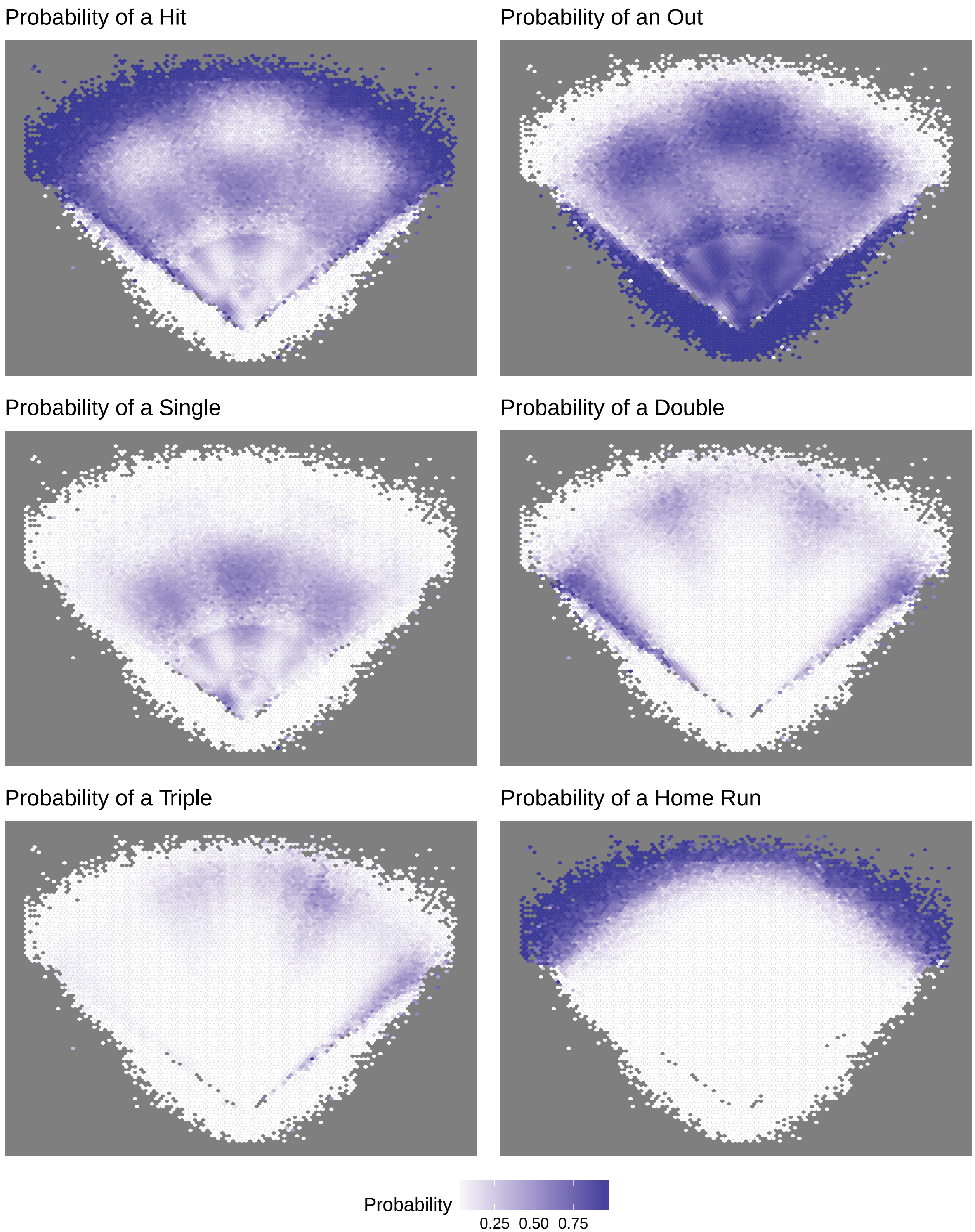


Figure 5.3.2: On-Field Batted Ball Probabilities

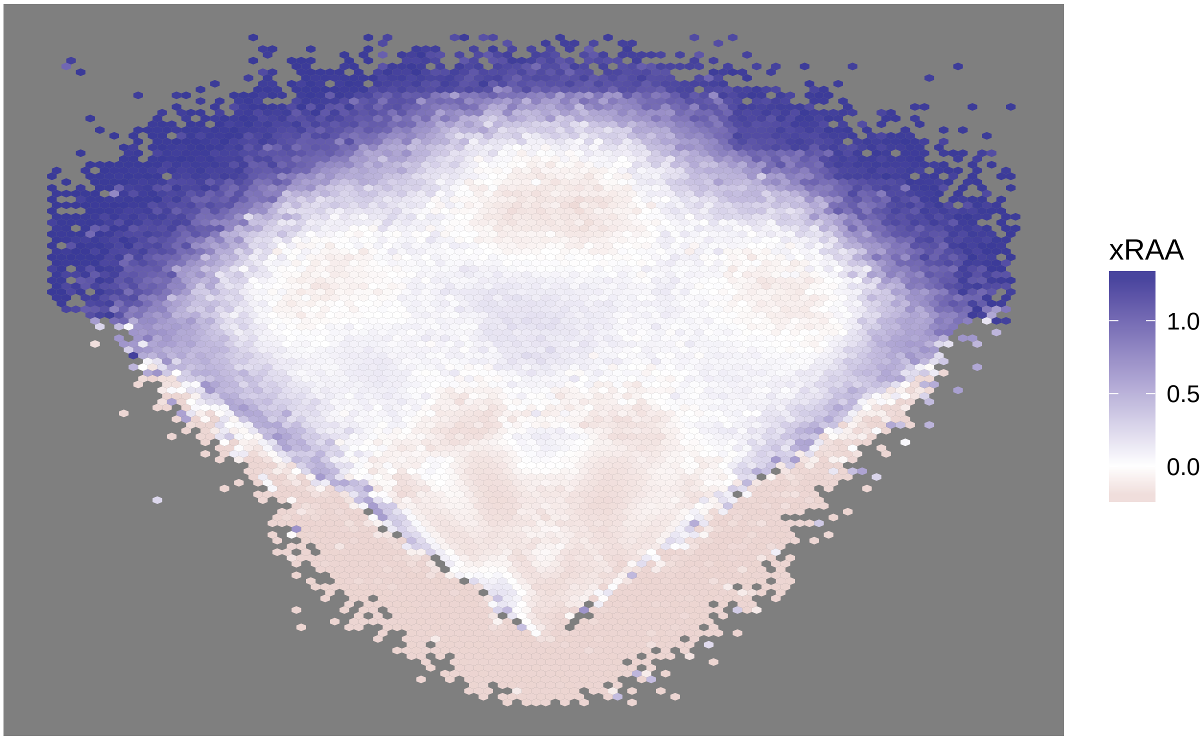


Figure 5.3.3: On-Field xRAA

was fielded or landed in the case of home runs.

# Chapter 6

# Expected Outcome Framework for Evaluating Player Performance

Using the two batted ball classifiers introduced in the previous chapter, we now present two corresponding versions of expected runs above average calculated using the two different models. We denote these by xRAA and xRAA-Partial throughout this section with xRAA corresponding to the full model which includes the proxy to the horizontal angle and xRAA-Partial corresponding to the model that only uses the velocity and launch angle components of a batted ball. We will examine both of these metrics in the context of batting and pitching performance beginning with an inspection of their respective distributions. We then proceed to examine how they compare to traditional measures of performance through year-to-year correlations, predictive power for future performance, and reliability. Further, we explore, for both batters and pitchers, some of the best performances and worst performances from the 2015-2017 seasons. We also address players who overperform and underperform these expected statistics and demonstrate that it is possible to consistently overperform and or underperform based on attributes beyond those expressed in these expected metrics. Only players with at least 250 batters faced or relevant plate appearances will be considered here.

Table 6.1.1: AbRAA, AxbRAA, AxbRAA-Partial Quantiles

| | 5% | 10% | 25% | 50% | 75% | 90% | 95% |
|---|---|---|---|---|---|---|---|
| AbRAA | -0.046 | -0.037 | -0.018 | 0.000 | 0.019 | 0.038 | 0.051 |
| AxbRAA | -0.046 | -0.037 | -0.020 | -0.001 | 0.019 | 0.039 | 0.051 |
| AxbRAA-Partial | -0.046 | -0.037 | -0.022 | -0.001 | 0.019 | 0.039 | 0.053 |

## 6.1 Distribution

### 6.1.1 Batters

In order to develop some intuition for these metrics, we begin with a simple examination of their distributions. In Table 6.1.1, we report several quantiles for AbRAA, AxbRAA, and AxbRAA-Partial. As can be seen, the quantiles are very similar across each of the metrics. As is to be expected, the median batter seasons from 2015-2017 results is about zero for each of the three metrics. Similarly, an above average batter (75%) added approximately .02 runs above an average plate appearance per time to the plate. Likewise, a below average batter (25%) costs his team approximately .02 runs per plate appearance when compared to an average batter. Great and very bad batters gain and cost their teams .04 runs per plate appearance respectively when compared to the average batter. At the extremes, we have that the best batters in the league can contribute more than .05 runs more per plate appearance than an average batter, whereas the worst batters with at least 250 plate appearances cost their teams about .045 runs per plate appearance when compared to the performance we could expect from an average batter. We suspect that the true underlying skill is actually symmetric between the best and worst batters but there is a selection bias at work here in that the worst batters will likely not get enough plate appearances to qualify as they are an offensive liability.

In Figure 6.1.1, we plot the distributions of AbRAA, AxbRAA, and AxbRAA-Partial. We can see the same general results discussed above. Almost all the of the mass is concentrated

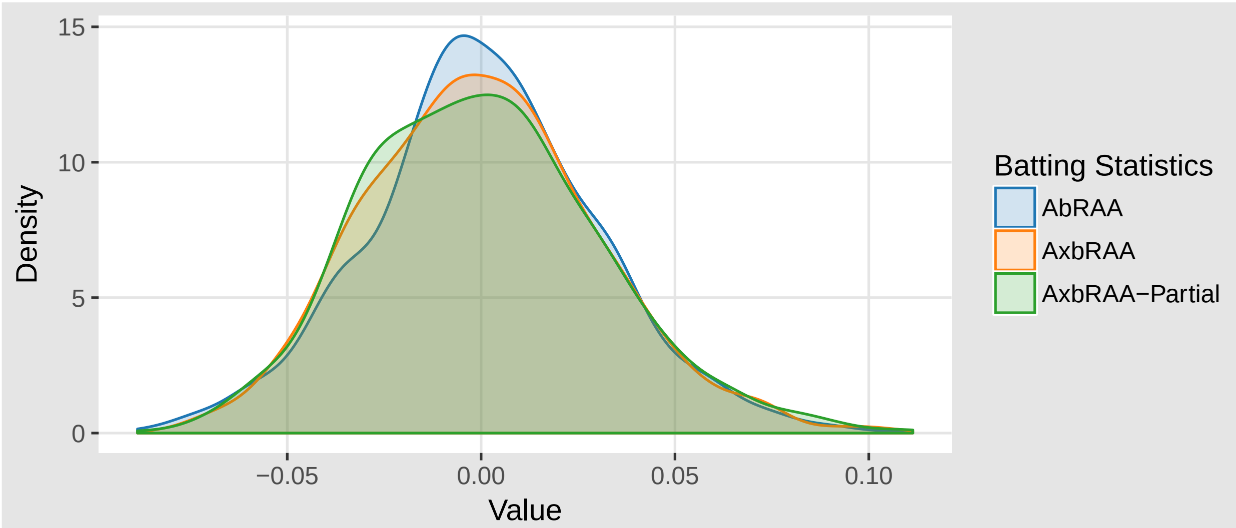


Figure 6.1.1: AbRAA, AxbRAA, AxbRAA-Partial Distributions

between -.05 and .05; however, we see the distributions are skewed with considerably more mass in the .05 to.10 range than in -.10 to -.05, a result of the fact that the worst batters simply are not given many plate appearances.

### 6.1.2 Pitchers

We now examine the distribution of ApRAA, AxpRAA, and AxpRAA-Partial. It is important to note that good pitchers will have a negative value for these metrics. That is because batters are expected to produce fewer runs per plate appearance against them when compared to an average pitcher. As in seen in Table 6.1.2 the skew for pitchers is more obvious than it was for batters. We can see the median of pitchers who faced at least 250 batters is approximately -.008 across each of the three metrics. An above average pitcher (25%) saves his team between .024 and .028 runs per batter faced when compared to an average pitcher. A below average pitcher costs his team somewhere around .01 runs per batter faced when compared to an average pitcher. Great pitchers (10%) save their teams about .045 runs per plate appearance according to the expected statistics and about .048 runs according to the

Table 6.1.2: ApRAA, AxpRAA, AxpRAA-Partial Quantiles

| | 5% | 10% | 25% | 50% | 75% | 90% | 95% |
|---|---|---|---|---|---|---|---|
| ApRAA | -0.060 | -0.048 | -0.028 | -0.008 | 0.012 | 0.029 | 0.043 |
| AxpRAA | -0.054 | -0.045 | -0.026 | -0.008 | 0.011 | 0.027 | 0.037 |
| AxpRAA-Partial | -0.053 | -0.045 | -0.024 | -0.007 | 0.010 | 0.026 | 0.033 |

traditional version of the metric. This gap between the expected versions of the metric and the traditional versions continues to widen at the extremes. According to the traditional version of the metric, the best pitchers save their team about .06 runs per batter faced when compared to a typical pitcher, whereas the best 5% for the expected versions save more than .053 runs per batter faced. This gap is seen on the opposite side of the spectrum as well: the traditional version has a larger magnitude than the expected versions. This is a product of the fact that by traditional metrics their skills are conflated with the skills of the defense that plays behind them. If a pitcher has an elite defense that plays behind him that will make him appear considerably better than he is. The expected metrics are able to isolate only pitcher performance in a manner similar to what FIP and xFIP did but without ignoring the information communicated by batted balls.

The increased skew for pitchers is very likely a result of the fact that pitchers really only contribute through their pitching skills whereas position players can provide value through batting, baserunning, and fielding. Many below average batters are able to continue to play and get plate appearances based on their strengths in these other areas. Pitchers, by and large, are not able to make up for their deficiency in any meaningful way.

In Figure 6.1.2, we plot the density of each of the three metrics to give a further sense of how ApRAA, AxpRAA, and AxpRAA-Partial are distributed.

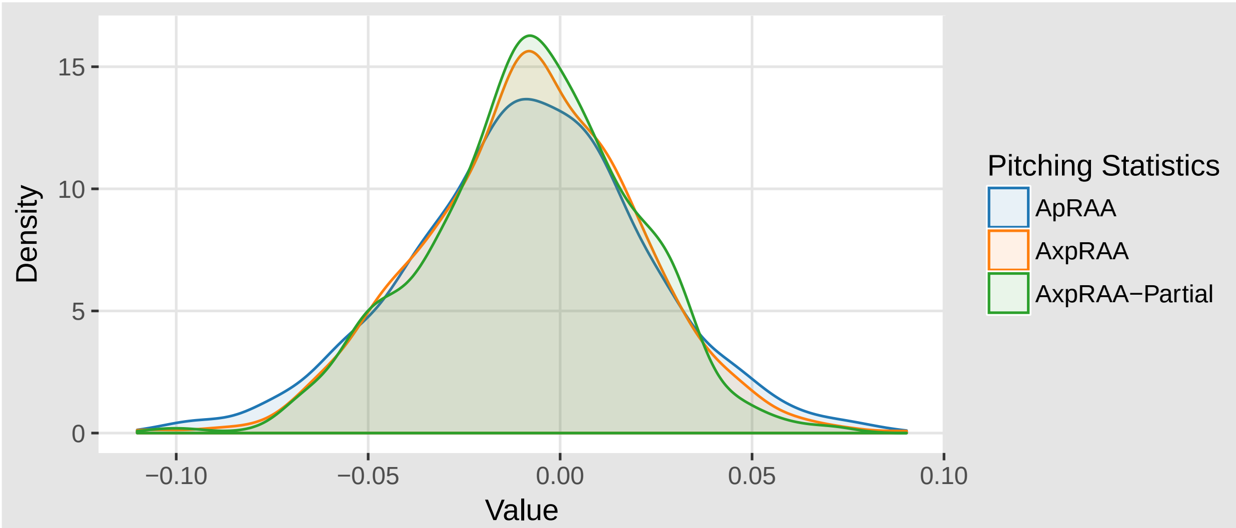


Figure 6.1.2: ApRAA, AxpRAA, AxpRAA-Partial Distributions

## 6.2 Year-to-Year Correlations

A very common way of comparing statistics within the sabermetric community is to look at the year-to-year correlation. While it does not show that one measurement of performance is superior to another, a higher year-to-year correlation would suggest that the metric is more robust to noise and better captures an underlying skill (or is possibly over-regressed).

### 6.2.1 Batters

In Table 6.2.1, we report the year-to-year correlation for the introduced metrics along with the traditional metrics discussed above. As expected, we find that AxbRAA and AxbRAA-Partial have a much higher year-to-year correlation than the traditional metrics suggesting that they better measure the underlying skill and are not as influenced by other random factors. Of the two, AxbRAA-Partial's correlation is higher suggesting that there is more year-to-year variation in the horizontal launch angle than when considering just the vertical launch angle and exit velocity in the model.

Table 6.2.1: Year-to-Year Correlations for Various Batting Metrics

| | Correlation |
|---|---|
| AxbRAA | 0.622 |
| AxbRAA-Partial | 0.680 |
| AbRAA | 0.433 |
| BA | 0.407 |
| OBP | 0.475 |
| SLG | 0.468 |
| OPS | 0.449 |

Table 6.2.2: Year-to-Year Correlations for Various Pitching Metrics

| | Correlation |
|---|---|
| AxpRAA | 0.499 |
| AxpRAA-Partial | 0.514 |
| ApRAA | 0.429 |
| ERA | 0.336 |
| FIP | 0.504 |
| xFIP | 0.613 |
| SIERA | 0.654 |

### 6.2.2 Pitchers

In Table 6.2.2, we present the year-to-year correlation for ApRAA, AxpRAA, AxpRAA-Partial, and ERA and its aforementioned estimators. One point to mention is that ERA and its estimators are on an innings pitched scale as opposed to a batters faced scale and would thus be expected to have higher correlations. Nevertheless, we see that AxpRAA and AxpRAA-Partial have a higher year-to-year correlation than pRAA and ERA and very similar correlations to FIP. SIERA and xFIP both have higher correlations, but that is expected as they are on an innings pitched scale and are heavily regressed (xFIP assumes that every pitcher gives up home runs at the same rate per flyball which is a faulty assumption).

## 6.3 Prediction of Future Performance

While a strong year-to-year correlation is a desirable property in a baseball metric, an even more desirable property is its ability to predict the skill of interest moving forward. For example, you are not usually interested in a pitcher's FIP on its own but rather in the context that you believe it to be a better indicator of what a pitcher's ERA will look like moving forward than his ERA currently. In a similar manner, we are interested in these expected metrics not necessarily in their own right but in their ability to better capture the talent level of the player and serve to give us a better idea of how the batter will perform moving forward. In this section, we will compare the predictive ability of ARAA, AxRAA, and AxRAA-Partial to predict future ARAA. More specifically, we will use the statistics from the first half of the season to predict performance in the second half of the season. Note that the halfway mark in the baseball community typically refers to the All-Star break which confusingly does not actually occur halfway through the regular season. While unclear, this language will be used here. In order to predict second-half performance, we fit a simple linear regression model with the chosen first half metric as the explanatory variable. Only players with at least 125 relevant plates appearances or batters faced in both halves are considered.

### 6.3.1 Batters

In Table 6.3.1, we report the $R^2$ and root of the mean squared error of the regression using metric from the first half to predict the second half AbRAA. We see that both expected versions are better predictors of performance in the second half of the season with the partial version of the metric slightly outperforming the full model version. This is a very interesting result especially since we know that full model performs better in the classification task. It would appear that horizontal angle (or at least the proxy used here) is subject to more random variation than the launch angle and exit velocity.

Table 6.3.1: Predictions of Second Half AbRAA

| | $R^2$ | RMSE |
|---|---|---|
| First Half AbRAA | 0.1133 | 0.0355 |
| First Half AxbRAA | 0.1247 | 0.0353 |
| First Half AxbRAA-Partial | 0.1262 | 0.0353 |

Table 6.3.2: Year-to-Year Correlation In Difference between AbRAA and AxbRAA

| | Diff Correlation |
|---|---|
| AxbRAA - AbRAA | 0.354 |

As mentioned above, some batters provide additional value through defensive performance but others through their skill in baserunning which is often closely tied to their speed. The classification models that AxbRAA is based only consider the characteristics of the batted ball when making their predictions. So in that way, it provides an excellent assessment of the quality of the batted balls that a player hits; however, it ignores the fact that the outcome probabilities could be different for two batters on identical batted balls. A very fast batter might be able to beat the throw to first base and thus get a single where a slower batter would be more likely to be thrown out. Similarly, a faster batter can turn would be singles into doubles by trying to take the extra base. Additionally, there are some batters who typically hit balls towards one side or the other with much greater regularity and as a result opposing teams will shift their defensive players to better defend those specific areas. Thus, it is possible for batters to consistently under or overperform their AxbRAA. In Table 6.3.2, we report the year to year correlation between the difference in AxbRAA and AbRAA strongly supporting the notion that there are batters who as a result of additional offensive skills or lack thereof over or underperform their expectation based on the way they contact the ball.

Even with the exclusion of additional information, we found that AxbRAA and AxbRAA-Partial were both better predictors of future AbRAA performance than was past AbRAA. However, by including the difference between their AxbRAA (or AxbRAA-Partial) and their

Table 6.3.3: Predictions of Second Half AbRAA with First Half Difference

| | $R^2$ | RMSE |
|---|---|---|
| First Half AbRAA | 0.1133 | 0.0355 |
| First Half AxbRAA | 0.1247 | 0.0353 |
| First Half AxbRAA and Diff | 0.1471 | 0.0349 |
| First Half AxbRAA-Partial | 0.1262 | 0.0353 |
| First Half AxbRAA-Partial and Diff | 0.1632 | 0.0345 |

Table 6.3.4: Predictions of Second Half ApRAA

| | $R^2$ | MSE |
|---|---|---|
| First Half ApRAA | 0.0991 | 0.0345 |
| First Half AxpRAA | 0.0950 | 0.0346 |
| First Half AxpRAA-Partial | 0.0766 | 0.0349 |

AbRAA in the first half along with their first half AxbRAA (or AxbRAA-Partial) we are able to better predict performance in the second half by a meaningful margin as is seen in Table 6.3.3.

### 6.3.2 Pitchers

In Table 6.3.4, we perform the same analysis as above. Here we have that traditional metric has a slightly higher $R^2$ value and lower RMSE. Contrasting with the batters, the partial model version of AxpRAA performs considerably worse than either the full model version or the traditional statistic. While the traditional statistic did moderately outperform the full model expected version, as we noted above defensive performance is not factored into expected statistics whereas it will be implicitly included in the traditional statistic.

In the same way we did for batters, we examine the year-to-year correlation in the difference between AxpRAA and ApRAA with the result displayed in 6.3.5. The year-to-year correlation in the difference is much lower for batters than for pitchers. This is because for batters the primary reasons for maintained under or overperformance were player-specific (speed,

Table 6.3.5: Year-to-Year Correlation In Difference between ApRAA and AxpRAA

| | Diff Correlation |
|---|---|
| AxpRAA - ApRAA | 0.136 |

Table 6.3.6: Predictions of Second Half ApRAA with First Half Difference

| | $R^2$ | MSE |
|---|---|---|
| First Half ApRAA | 0.0991 | 0.0345 |
| First Half AxpRAA | 0.0950 | 0.0346 |
| First Half AxpRAA and Diff | 0.1183 | 0.0342 |
| First Half AxpRAA-Partial | 0.0766 | 0.0349 |
| First Half AxpRAA-Partial with Diff | 0.1165 | 0.0342 |

baserunning ability) whereas for pitchers it is likely a product of their defense which will change from year-to-year.

Nevertheless, by including the gap between AxpRAA (or AxpRAA-Partial) and ApRAA alongside AxpRAA (or AxpRAA-Partial), we find that we can substantially increase our ability to predict future performance as is seen in Table 6.3.6

## 6.4 Reliability

Another area of interest in sabermetrics is determining at what sample size performance statistics become meaningful. There has been a well-established line of research measuring the reliability of various batter and pitching statistics including the work of Russell Carleton [46–48], Tango & Lichtman [12], Carty [49,50], and Sean Dolinar and Pemstein [51–53]. Here we apply a methodology most similar to Dolinar and Pemstein to compare the reliability of the expected statistics with their traditional versions.

Our methodology is as follows for both batters and pitchers. We separate the data into player-season combinations from the 2015-2017 seasons. In increments of 25 beginning with 25 and going to 600, we sample $k$ pertinent plate appearances or batters faced from each player-season combination with at least $k$ appearances. Six hundred such plate appearances/batters faced is chosen as the upper limit because the sample size of batters and pitchers with more than 600 appearances in a particular season is rather small and susceptible to noise. We could get larger samples by grouping by player as opposed to grouping by player and season; however, we believe that the differences from season to season as a result of aging and other factors would neutralize whatever advantage we get from being able to consider larger samples. After taking a sample of k appearances, we then compute Cronbach's alpha for average runs above average, the full model average expected runs above average, and the partial model average expected runs above average.

### 6.4.1 Batters

As can be seen in figure 6.4.1, the reliability of AxbRAA and AxbRAA-Partial is much higher across all sample sizes. For the traditional statistic, we see that the coefficient of reliability is only approaching .6 after 600 plate appearances. However, with only 200 plate appearances the same level of reliability has been achieved by AxbRAA-Partial and after about 300 plate appearances for the full model AxbRAA. Furthermore, by 600 plate appearances, the partial model version of AxbRAA has a reliability coefficient of about .8 and for the full model version a coefficient of about .75. These are extremely powerful properties for properly evaluating baseball players. The expected metrics which are driven based on the way the batter actually hits the ball as opposed to only what happens become more reliable much quicker meaning that sabermetricians both in the public and private sphere would be better able to better know a player's true talent level in much fewer plate appearances.

Another interesting feature to note in figure 6.4.1 is how the reliability coefficient appears to be starting to approach an asymptote below 1. We believe that this is the product of the fact that acquiring 600+ plate appearances takes a considerable amount of time during

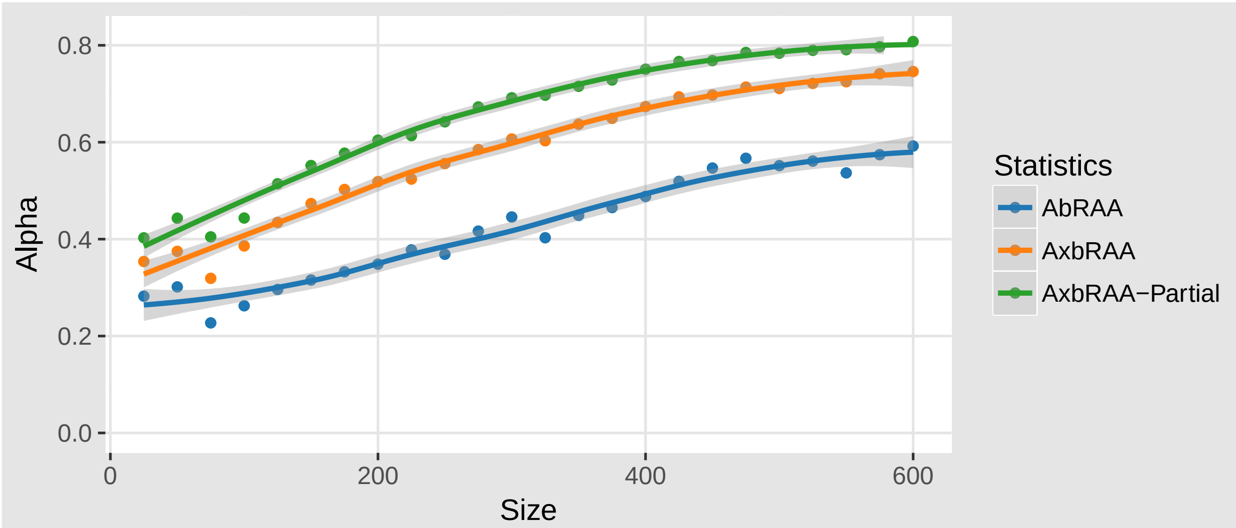


Figure 6.4.1: Reliability of bRAA, xbRAA, and xbRAA-Partial

which the true talent level of the batter will fluctuate due to aging, injury, rest, etc. So even if we were able to collect samples of thousands of plate appearances per batter per season we believe that the curve would not eventually approach a reliability coefficient of 1. It is also interesting to observe that the partial model AxbRAA is always a more reliable metric than the full model version. As touched on above, this seems to be a result of the fact that there is more variance in the horizontal launch angle (or at least the proxy used here) than in vertical launch angle or the exit velocity.

### 6.4.2 Pitchers

The pitcher reliability curves shown in figure 6.4.2 display considerably different behavior than those for batters. First, across all three metrics and all plate appearances, the reliability is much lower than it is for batters. Additionally, whereas there were the potential beginnings of an asymptotic behavior for the expected statistics by 600 plate appearances for the batters, this does not seem to be the case for pitchers and rather appears that with more plate appearances the reliability would continue to increase by large amounts. Furthermore, the

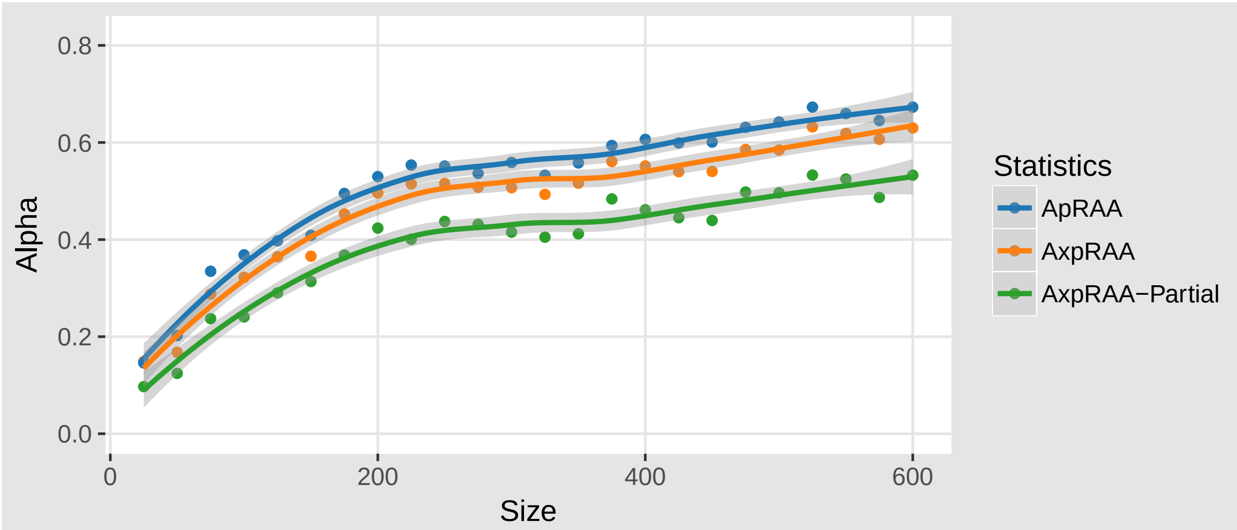


Figure 6.4.2: Reliability of pRAA, xpRAA, and xpRAA-Partial

partial model AxpRAA has considerably lower reliability than either the full model version or the observed results. An additional surprise is that the metric based on the observed result, in fact, has a higher reliability than the full model version. We hypothesize that this is once again a result of the fact that the traditional metric incorporates defensive information implicitly.

## 6.5 League Leaders

In this section, we utilize the introduced metrics in a method similar to how they would typically be used in public or private sabermetric analysis to evaluate the performance of individual players. We take a look at batters who had the best and worst seasons according to the full model AxbRAA as well as those who most under or overperformed their expected metrics. We then do the same for pitchers.

Table 6.5.1: Best Batting Seasons by AxbRAA 2015-2017

| Player | Season | PA | Team | Median Launch Angle | Median Exit Velo | K% | BB% | AbRAA | AxbRAA | bRAA | xbRAA |
|---|---|---|---|---|---|---|---|---|---|---|---|
| David Ortiz | 2016 | 610 | BOS | 16.9 | 93.8 | 13.8 | 10.70 | 0.078 | 0.107 | 47.7 | 65.5 |
| Miguel Cabrera | 2016 | 664 | DET | 12.5 | 95.3 | 17.3 | 9.04 | 0.061 | 0.101 | 40.2 | 67.0 |
| Aaron Judge | 2017 | 667 | NYY | 15.7 | 97.7 | 31.0 | 17.40 | 0.089 | 0.095 | 59.3 | 63.5 |
| Bryce Harper | 2015 | 635 | WSH | 14.7 | 94.2 | 20.3 | 17.20 | 0.108 | 0.077 | 68.3 | 48.8 |
| Mike Trout | 2016 | 669 | LAA | 12.8 | 91.8 | 20.3 | 15.50 | 0.076 | 0.076 | 50.6 | 51.1 |
| Joey Votto | 2017 | 687 | CIN | 16.0 | 90.3 | 12.1 | 16.60 | 0.086 | 0.076 | 58.9 | 52.4 |
| Freddie Freeman | 2016 | 675 | ATL | 18.0 | 93.5 | 25.3 | 10.50 | 0.063 | 0.074 | 42.5 | 49.8 |
| David Ortiz | 2015 | 593 | BOS | 15.5 | 94.9 | 15.9 | 10.10 | 0.040 | 0.074 | 23.5 | 43.7 |
| Giancarlo Stanton | 2017 | 679 | MIA | 10.9 | 92.3 | 23.7 | 10.60 | 0.074 | 0.073 | 50.3 | 49.3 |
| Josh Donaldson | 2016 | 692 | TOR | 13.7 | 94.8 | 17.1 | 14.90 | 0.064 | 0.072 | 44.3 | 50.2 |

[1] Team indicates the team that the player played for at the end of the season
[2] League average strikeout rate is approximately 20 percent
[3] League average walk rate is approximately 8 percent
[4] Team indicates the team that the player played for at the end of the season

### 6.5.1 Best AxbRAA Performances

In Table 6.5.1, we report the ten best seasons by AxbRAA by players with at least 502 plate appearances (the number of plate appearances needed to qualify for being a league leader) that count towards the calculation of AxbRAA. Other statistics of interest that have been included are the median vertical launch angle, the median exit velocity, the player's strikeout percentage, the players walk and hit-by percentage, as well as the xbRAA and bRAA. The names that we find here a very unsurprising for even casual baseball fans as the list is entirely composed of slugging superstars. While they all have excellent reputations as elite batters, there are similarities and differences between them all. We find that for all of them that they tend to hit the ball in the air and at high exit velocities. Further, we see that almost all of them have double-digit walk rates well above the league average. However, there are substantial differences between these batters. Consider Joey Votto and Aaron Judge. Votto has the lowest strikeout rates of the group and one of the best walk rates in the group, but also the lowest median exit velocity. Judge on the other hand strikes-out an absurd amount but also has the highest walk rate of the group and hits the ball very hard when he does make contact with it.

Turning our attention to the last column of Table 6.5.1, we get a sense for how much value these best of the best batters add to their team over the course of the season. We see that these batters in expectation can create more than 50 runs when compared to an average MLB batter. A commonly used rough conversion between runs and wins is that for each additional 10 runs a team scores we would expect a team to win approximately one more game. Thus we see that these elite batters can add five or more wins to their team with their bat when compared to an average batter.

### 6.5.2 Worst AxbRAA Performances

On the other end of the spectrum in Table 6.5.2, we report the worst seasons by AxbRAA by players with at least 502 plate appearances that count towards AxbRAA. Whereas the best batters had median launch angles in the teens and accompanying median exit velocities over 90mph, no batter here has a median exit velocity above 90 with most in the mid-eighties and lower median launch angles. Additionally, the walk rates are all in the single digits. It can also be observed that these batters in expectation cost their teams between about 23 and 35 runs over the course of the season when compared to an average batter which would correspond to about 2.5-3.5 wins.

### 6.5.3 Overperforming Batters

In Table 6.5.3, we report the batters who most outperformed their expected results on average. As we noted above there are batters who can consistently outperform their AxbRAA as a result of speed and baserunning abilities; we observe several such players here. Dee Gordon, Jonathan Villar, Jose Altuve (who appears twice), Jose Reyes, and Ender Inciarte are all known as exceptionally fast baseball players and thus are able to outperform their expectation. However, we also observe several batters who based on our prior knowledge of their skill-sets,

Table 6.5.2: Worst Batting Seasons by AxbRAA 2015-2017

| Player | Season | PA | Team | Median Launch Angle | Median Exit Velo | K% | BB% | AbRAA | AxbRAA | bRAA | xbRAA |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Chris Owings | 2015 | 538 | ARI | 11.80 | 87.5 | 26.8 | 4.28 | -0.057 | -0.056 | -30.70 | -30.2 |
| Billy Hamilton | 2017 | 628 | CIN | 8.80 | 81.9 | 21.0 | 7.01 | -0.037 | -0.055 | -23.10 | -34.7 |
| Jose Reyes | 2015 | 506 | COL | 9.74 | 84.0 | 12.3 | 5.14 | -0.021 | -0.051 | -10.60 | -25.7 |
| Jean Segura | 2015 | 575 | MIL | 1.44 | 87.6 | 16.0 | 1.91 | -0.046 | -0.049 | -26.60 | -28.4 |
| Billy Burns | 2015 | 549 | OAK | 5.14 | 83.0 | 14.8 | 4.55 | -0.007 | -0.049 | -3.89 | -26.9 |
| Jimmy Rollins | 2015 | 558 | LAD | 13.80 | 86.6 | 15.4 | 7.89 | -0.037 | -0.044 | -20.80 | -24.7 |
| Anthony Gose | 2015 | 533 | DET | 3.70 | 87.9 | 27.2 | 8.44 | -0.018 | -0.044 | -9.51 | -23.4 |
| Josh Harrison | 2016 | 518 | PIT | 12.70 | 88.0 | 14.7 | 3.47 | -0.019 | -0.044 | -10.00 | -22.8 |
| Alcides Escobar | 2017 | 621 | KC | 11.50 | 86.3 | 16.3 | 2.25 | -0.044 | -0.042 | -27.10 | -26.3 |
| Alcides Escobar | 2016 | 670 | KC | 8.23 | 85.2 | 14.3 | 3.73 | -0.038 | -0.041 | -25.30 | -27.8 |

[1] Team indicates the team that the player played for at the end of the season
[2] League average strikeout percentage is approximately 20 percent
[3] League average walk rate is approximately 8 percent
[4] Team indicates the team that the player played for at the end of the season

we believe to have gotten extremely lucky. Billy Burns, for example, seems to have benefited tremendously from good fortune. He rated as an almost average batter despite the fact that in expectation we would have expected him to be one of the worst batters in the league. We already saw Bryce Harper's 2015 season in the best AxbRAA seasons above, but we have reason to believe that he also go extraordinarily lucky that season as his expected AxbRAA was an elite but .077 but his AbRAA was an otherworldly .108. While Harper is a solid athlete, we have no reason to believe that he should have been able to outperform his expected stat line like this.

### 6.5.4 Underperforming Batters

In 6.5.4, we report the players who most underperformed their expected statistics. This list is heavily dominated by first baseman and designated hitters who tend to be very slow. Furthermore, there are also batters who are very easy to shift against. For example, David Ortiz was both very slow and an extreme pull-hitter so teams would often shift their defense over to be in better position to take away would have been hits. Miguel Cabrera is another interesting one as he is both very slow and plays half of his games in the Detroit's Tiger

Table 6.5.3: Overperforming Batters by AxbRAA

| Player | Season | PA | Team | Median Launch Angle | Median Exit Velo | AbRAA | AxbRAA | Diff | Diff Next Year | Diff Previous Year |
|---|---|---|---|---|---|---|---|---|---|---|
| Dee Gordon | 2015 | 640 | MIA | -0.216 | 85.3 | 0.007 | -0.041 | -0.048 | -0.026 | NA |
| Billy Burns | 2015 | 549 | OAK | 5.140 | 83.0 | -0.007 | -0.049 | -0.042 | 0.005 | NA |
| Didi Gregorius | 2016 | 590 | NYY | 13.700 | 86.8 | -0.004 | -0.037 | -0.033 | -0.021 | -0.003 |
| Jonathan Villar | 2016 | 670 | MIL | 3.400 | 90.7 | 0.025 | -0.007 | -0.032 | -0.012 | -0.038 |
| Bryce Harper | 2015 | 635 | WSH | 14.700 | 94.2 | 0.108 | 0.077 | -0.031 | 0.000 | NA |
| Jose Altuve | 2015 | 678 | HOU | 9.600 | 88.2 | 0.016 | -0.014 | -0.030 | -0.018 | NA |
| Jose Altuve | 2017 | 658 | HOU | 10.100 | 87.8 | 0.067 | 0.036 | -0.030 | NA | -0.018 |
| Jose Reyes | 2015 | 506 | COL | 9.740 | 84.0 | -0.021 | -0.051 | -0.030 | -0.022 | NA |
| Zack Cozart | 2017 | 507 | CIN | 15.200 | 88.0 | 0.057 | 0.028 | -0.029 | NA | -0.010 |
| Ender Inciarte | 2017 | 712 | ATL | 9.760 | 85.0 | 0.003 | -0.026 | -0.029 | NA | -0.008 |

[1] Diff indicates the difference in AxbRAA and AbrAA.

[2] Team indicates the team that the player played for at the end of the season

Comerica Park which is very spacious. In addition to these factors, research has indicated that Detroit's tracking system reads values higher than what they are in reality which further contributes to Miguel Cabrera's gap between AxbRAA and AbRAA [54].

### 6.5.5 Best AxpRAA Performances

In Table 6.5.5, we report the best seasons by AxpRAA by starting pitchers with at least 502 batters faced. As with the batters, this list is populated by the superstars that we would expect. Perhaps the most obvious takeaway from this table is the dominance of the Los Angeles Dodgers' Clayton Kershaw who has all three of his seasons from 2015-2017 in the top ten as well as the two top spots overall. Part of the reason for Kershaw's exceptional success his high strikeout percentage in the low 30s and an extremely low BB%. This higher than

Table 6.5.4: Underperforming Batters by AbxRAA

| Player | Season | PA | Team | Median Launch Angle | Median Exit Velo | AbRAA | AxbRAA | Diff | Diff Next Year | Diff Previous Year |
|---|---|---|---|---|---|---|---|---|---|---|
| Miguel Cabrera | 2016 | 664 | DET | 12.5 | 95.3 | 0.061 | 0.101 | 0.040 | 0.039 | 0.030 |
| Miguel Cabrera | 2017 | 523 | DET | 13.2 | 94.6 | -0.008 | 0.031 | 0.039 | NA | 0.040 |
| Kendrys Morales | 2016 | 616 | KC | 11.1 | 94.8 | 0.012 | 0.048 | 0.035 | 0.025 | 0.022 |
| Albert Pujols | 2016 | 644 | LAA | 11.0 | 93.0 | 0.006 | 0.040 | 0.034 | 0.025 | 0.022 |
| David Ortiz | 2015 | 593 | BOS | 15.5 | 94.9 | 0.040 | 0.074 | 0.034 | 0.029 | NA |
| Brandon Moss | 2015 | 519 | STL | 19.6 | 91.0 | -0.013 | 0.018 | 0.031 | 0.004 | NA |
| Kyle Seager | 2015 | 676 | SEA | 15.5 | 90.6 | 0.007 | 0.037 | 0.030 | 0.016 | NA |
| Mitch Moreland | 2017 | 570 | BOS | 12.2 | 90.4 | 0.003 | 0.032 | 0.029 | NA | 0.017 |
| David Ortiz | 2016 | 610 | BOS | 16.9 | 93.8 | 0.078 | 0.107 | 0.029 | NA | 0.034 |
| Manny Machado | 2017 | 687 | BAL | 12.8 | 93.5 | 0.005 | 0.034 | 0.028 | NA | 0.015 |

[1] Diff indicates the difference in AxbRAA and AbrAA.
[2] Team indicates the team that the player played for at the end of the season

average strikeout rate and lower than average walk/HBP rate is a key part of the success of many of these pitchers as all posted above-average to elite strikeout rates and all but one posted walk rates of less than 8.

These elite starting pitchers in expectation saved between .05 and .08 runs per batter when compared to the performance of an average starting pitcher. Over the course of the season, this can translate to saving their team in excess of 40 runs and more than approximately four wins over an average pitcher.

In Table 6.5.6, we report the best seasons by AxpRAA by pitchers with at least 200 batters faced so as to allow relief pitchers to be included. Relief pitchers typically only pitch about an inning at a time whereas starters are responsible for pitching many innings typically between five and seven. As a result, relief pitchers are usually able to give a higher per pitch effort

Table 6.5.5: Best Starting Pitching Seasons by AxpRAA 2015-2017

| Player | Season | TBF | Team | Median Launch Angle | Median Exit Velo | K% | BB% | ApRAA | AxpRAA | pRAA | xpRAA |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Clayton Kershaw | 2016 | 539 | LAD | 7.58 | 87.3 | 31.9 | 2.23 | -0.0965 | -0.0819 | -52.0 | -44.2 |
| Clayton Kershaw | 2015 | 885 | LAD | 5.89 | 86.1 | 33.9 | 5.20 | -0.0762 | -0.0758 | -67.4 | -67.1 |
| Zack Greinke | 2015 | 836 | LAD | 7.77 | 87.8 | 23.8 | 5.26 | -0.0812 | -0.0697 | -67.9 | -58.2 |
| Jake Arrieta | 2015 | 833 | CHC | 2.34 | 85.0 | 27.4 | 6.24 | -0.0775 | -0.0639 | -64.6 | -53.3 |
| Corey Kluber | 2017 | 772 | CLE | 11.00 | 87.5 | 34.2 | 5.05 | -0.0663 | -0.0592 | -51.2 | -45.7 |
| Max Scherzer | 2017 | 773 | WSH | 18.20 | 87.6 | 34.5 | 8.15 | -0.0606 | -0.0582 | -46.9 | -45.0 |
| Jacob deGrom | 2015 | 739 | NYM | 10.10 | 88.4 | 27.6 | 5.14 | -0.0602 | -0.0561 | -44.5 | -41.4 |
| Clayton Kershaw | 2017 | 674 | LAD | 9.49 | 86.4 | 30.0 | 4.45 | -0.0513 | -0.0535 | -34.6 | -36.0 |
| Matt Harvey | 2015 | 718 | NYM | 10.40 | 89.0 | 25.5 | 5.43 | -0.0475 | -0.0513 | -34.1 | -36.9 |
| Justin Verlan-der | 2015 | 533 | DET | 19.20 | 87.8 | 21.2 | 6.38 | -0.0398 | -0.0499 | -21.2 | -26.6 |

[1] Team indicates the team that the player played for at the end of the season
[2] League average strikeout rate is approximately 20 percent
[3] League average walk rate is approximately 8 percent

and also have the advantage of facing batters only once. Because of this elite relief pitchers are often better on a rate basis than the best starting pitchers. We observe this in Table 6.5.6 as nine of the ten spots are occupied by relief pitcher seasons with the one exception being Clayton Kershaw's 2016 season.

With these elite reliever seasons, we see that many of them have extremely high strikeout rates and low walk rates as well. From, the median launch angle, we can see that there are a few separate groups. Zach Britton throws a sinking fastball that generates exorbitant amounts of groundballs, whereas Kenley Jansen and Josh Fields are known for throwing a cutter and fastball respectively that rise relative to what we would expect from gravity alone and generate a lot of flyballs and pop-ups.

Table 6.5.6: Best Pitching Seasons by AxpRAA 2015-2017

| Player | Season | TBF | Team | Median Launch Angle | Median Exit Velo | K% | BB% | ApRAA | AxpRAA | pRAA | xpRAA |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Kenley Jansen | 2016 | 246 | LAD | 23.30 | 86.4 | 42.3 | 4.47 | -0.106 | -0.115 | -26.0 | -28.4 |
| Zach Britton | 2016 | 250 | BAL | -15.80 | 83.2 | 29.6 | 6.00 | -0.107 | -0.110 | -26.8 | -27.5 |
| Andrew Miller | 2016 | 274 | CLE | 5.03 | 88.8 | 44.9 | 4.01 | -0.091 | -0.103 | -24.9 | -28.3 |
| Aroldis Chap-man | 2016 | 222 | CHC | 8.01 | 90.0 | 40.5 | 8.11 | -0.095 | -0.094 | -21.2 | -20.8 |
| Seung-Hwan Oh | 2016 | 308 | STL | 15.80 | 89.4 | 33.4 | 5.52 | -0.081 | -0.090 | -25.0 | -27.7 |
| Kenley Jansen | 2017 | 257 | LAD | 17.40 | 86.0 | 42.4 | 3.50 | -0.093 | -0.089 | -23.9 | -22.8 |
| Zach Britton | 2015 | 249 | BAL | -17.00 | 88.1 | 31.7 | 5.22 | -0.067 | -0.085 | -16.7 | -21.1 |
| Josh Fields | 2015 | 204 | HOU | 20.10 | 89.3 | 32.8 | 8.33 | -0.051 | -0.084 | -10.4 | -17.2 |
| Andrew Miller | 2017 | 242 | CLE | 14.40 | 83.8 | 39.3 | 10.70 | -0.094 | -0.083 | -22.8 | -20.1 |
| Clayton Kershaw | 2016 | 539 | LAD | 7.58 | 87.3 | 31.9 | 2.23 | -0.096 | -0.082 | -52.0 | -44.2 |

[1] Team indicates the team that the player played for at the end of the season
[2] League average strikeout rate is approximately 20 percent
[3] League average walk rate is approximately 8 percent

### 6.5.6 Worst AxpRAA Performances

In Table 6.5.7, we report the worst seasons by AxpRAA from 2015-2017 by starting pitchers with at least 500 batters in each season. When contrasted with the elite starting pitchers we see that these pitchers give up much harder contact and also produce strikeouts at a lower than average rate and/or walk than is average as well. This combination of giving up hard contact, failing to get high strikeout totals, and walking batters proves costly as these pitchers cost their team between .035 and .06 runs per batter faced when compared to an average pitcher which translates over a full to season to up to 40 runs worse than an average pitcher which roughly equates to about four wins lost over the course of the year.

Table 6.5.7: Worst Pitching Seasons by AxpRAA 2015-2017

| Player | Season | TBF | Team | Median Launch Angle | Median Exit Velo | K% | BB% | ApRAA | AxpRAA | pRAA | xpRAA |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Derek Holland | 2017 | 622 | CWS | 13.90 | 91.0 | 16.7 | 13.00 | 0.053 | 0.060 | 33.20 | 37.1 |
| Jeremy Guthrie | 2015 | 658 | KC | 16.30 | 90.0 | 12.8 | 7.90 | 0.040 | 0.055 | 26.20 | 36.4 |
| Ricky Nolasco | 2017 | 777 | LAA | 13.20 | 92.0 | 18.3 | 7.46 | 0.029 | 0.052 | 22.90 | 40.3 |
| Jordan Zimmer-man | 2017 | 708 | DET | 16.90 | 90.1 | 14.4 | 6.92 | 0.041 | 0.041 | 29.10 | 29.0 |
| Kyle Gibson | 2017 | 692 | MIN | 6.84 | 90.6 | 17.3 | 9.54 | 0.026 | 0.039 | 18.00 | 26.9 |
| Bartolo Colon | 2017 | 642 | MIN | 13.00 | 90.6 | 13.9 | 5.61 | 0.049 | 0.038 | 31.20 | 24.4 |
| Yovani Gallardo | 2017 | 571 | SEA | 11.00 | 89.8 | 16.5 | 10.30 | 0.022 | 0.038 | 12.40 | 21.5 |
| Mike Pelfrey | 2017 | 542 | CWS | 6.72 | 90.3 | 14.6 | 12.50 | 0.032 | 0.037 | 17.60 | 19.8 |
| Kyle Kendrick | 2015 | 622 | COL | 13.90 | 90.6 | 12.9 | 8.04 | 0.053 | 0.036 | 32.70 | 22.1 |
| Jeremy Hellick-son | 2017 | 688 | BAL | 17.60 | 89.0 | 14.0 | 7.70 | 0.014 | 0.035 | 9.69 | 24.3 |

[1] Team indicates the team that the player played for at the end of the season
[2] League average strikeout percentage is approximately 20 percent
[3] League average walk rate is approximately 8 percent

### 6.5.7 Overperforming Starting Pitchers

Unlike with batters, the starting pitchers with the largest positive gap between their expected and realized pitching runs above average (displayed in Table 6.5.8) are not quite as obvious as with the batters. Some are explainable as Jake Arrieta and Tim Hudson pitched in front of the best and second best defense in baseball according to Fangraphs' defensive metrics [31]. Other pitchers may have just gotten extremely lucky in a given year. Nevertheless, these pitchers benefited in the neighborhood of .025 runs per batter faced when compared to an average pitcher. Over the course of six hundred batters faced this would equate to 15 runs better than their expectation which would substantially influence how that pitcher is evaluated. Teams with access to this kind of information could better avoid signing players who have gotten lucky and would be overvalued. The presence of Verlander (who spent

Table 6.5.8: Overperforming Pitchers By AxpRAA

| Player | Season | TBF | Team | Median Launch Angle | Median Exit Velo | K% | BB% | ApRAA | AxpRAA | Diff | Diff Next Year | Diff Previous Year |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Jake Arrieta | 2016 | 792 | CHC | 6.42 | 87.8 | 23.9 | 10.20 | -0.051 | -0.024 | 0.027 | -0.001 | 0.014 |
| Justin Verlan-der | 2017 | 844 | HOU | 16.60 | 89.3 | 25.8 | 8.53 | -0.030 | -0.003 | 0.027 | NA | 0.005 |
| Chris Tillman | 2016 | 712 | BAL | 13.20 | 89.8 | 19.5 | 10.10 | -0.006 | 0.020 | 0.026 | 0.000 | 0.009 |
| Jose Urena | 2017 | 715 | MIA | 13.30 | 88.6 | 15.7 | 10.30 | -0.005 | 0.020 | 0.025 | NA | -0.008 |
| Andrew Cashner | 2017 | 702 | TEX | 10.30 | 89.0 | 12.1 | 10.40 | -0.015 | 0.009 | 0.024 | NA | -0.019 |
| Michael Fulmer | 2017 | 671 | DET | 8.70 | 87.4 | 17.0 | 6.86 | -0.035 | -0.011 | 0.024 | NA | 0.018 |
| Erasmo Ramirez | 2015 | 664 | TB | 7.48 | 88.8 | 18.8 | 7.38 | -0.030 | -0.006 | 0.024 | -0.010 | NA |
| Sonny Gray | 2015 | 830 | OAK | 4.71 | 89.7 | 20.4 | 7.35 | -0.051 | -0.028 | 0.023 | -0.007 | NA |
| Alex Cobb | 2017 | 738 | TB | 8.19 | 90.1 | 17.3 | 6.50 | -0.014 | 0.009 | 0.023 | NA | -0.046 |
| Tim Hudson | 2015 | 522 | SF | 3.35 | 90.6 | 12.3 | 8.24 | 0.000 | 0.023 | 0.023 | NA | NA |

[1] Diff indicates the difference in AxpRAA and AprAA.
[2] Team indicates the team that the player played for at the end of the season
[3] League average strikeout rate is approximately 20 percent
[4] League average walk rate is approximately 8 percent

most of 2017 with Detroit before being trade to Houston) and Fulmer on this list should be discounted somewhat as it is known that Detroit's exit velocity readings are higher than in reality [54].

### 6.5.8 Underperforming Starting Pitchers

On the reverse, in Table 6.5.9 we report the starting pitchers who had the largest negative gap between their expected and realized pitching runs above average. These pitchers lost between .020 and .029 runs per batter faced based on the difference between their expected and realized performances. The opposite conclusion holds true for these pitchers; they are very likely undervalued compared to their actual talent level and could likely be had for a lower salary.

Table 6.5.9: Underperforming Pitchers by AxpRAA

| Player | Season | TBF | Team | Median Launch Angle | Median Exit Velo | K% | BB% | ApRAA | AxpRAA | Diff | Diff Next Year | Diff Previous Year |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Juan Nicasio | 2016 | 505 | PIT | 11.600 | 89.7 | 27.3 | 9.70 | 0.005 | -0.024 | -0.029 | 0.004 | -0.006 |
| Jameson Taillon | 2017 | 576 | PIT | 8.950 | 88.8 | 21.7 | 8.16 | 0.012 | -0.016 | -0.027 | NA | -0.007 |
| CC Sabathia | 2015 | 718 | NYY | 8.130 | 89.1 | 19.1 | 7.38 | 0.014 | -0.011 | -0.025 | -0.019 | NA |
| Michael Pineda | 2015 | 664 | NYY | 7.420 | 89.2 | 23.3 | 3.61 | -0.003 | -0.028 | -0.024 | -0.022 | NA |
| Chris Rusin | 2015 | 564 | COL | 5.700 | 89.2 | 14.9 | 6.91 | 0.037 | 0.013 | -0.023 | -0.006 | NA |
| Kyle Freeland | 2017 | 670 | COL | 5.920 | 89.2 | 16.0 | 10.00 | 0.014 | -0.009 | -0.023 | NA | NA |
| Michael Pineda | 2016 | 754 | NYY | 10.800 | 90.0 | 27.2 | 7.69 | 0.009 | -0.013 | -0.022 | -0.012 | -0.024 |
| Luis Perdomo | 2016 | 654 | SD | 0.455 | 91.3 | 16.1 | 7.03 | 0.029 | 0.007 | -0.022 | -0.005 | NA |
| Tim Adelman | 2017 | 529 | CIN | 18.000 | 89.2 | 20.2 | 10.60 | 0.036 | 0.015 | -0.021 | NA | 0.006 |
| Jacob deGrom | 2017 | 819 | NYM | 10.400 | 88.0 | 29.2 | 6.84 | -0.025 | -0.045 | -0.020 | NA | 0.004 |

[1] Diff indicates the difference in AxpRAA and AprAA.
[2] Team indicates the team that the player played for at the end of the season
[3] League average strikeout rate is approximately 20 percent
[4] League average walk rate is approximately 8 percent

# Chapter 7

# Juiced Ball

## 7.1 Motivation

Over the past three seasons there has been a very dramatic increase in offensive output across the league. As can be seen in Table 7.1.1, the number of runs across the league has increased by more than 2,000 runs which corresponds to four-tenths of a run more per team per game. Even more dramatic has been the increase in the number of home runs going from 4,909 in 2015 to an all-time MLB high of 6,105 in 2017. This corresponds to nearly a quarter of a home run more per team per game. Associated with these events has been an increase in the average launch angle. More and more players are putting an emphasis on hitting the ball in the air, a so-called fly ball revolution. While launch angle has gone up steadily, average exit velocity went up in 2016 but was actually on average below its 2015 levels. Nevertheless, xRAA on batted balls has increased tremendously going from .027 in 2015, to .037 in 2016, all the way to .042 in 2017.

The shift in offensive output can be seen even more clearly when broken down by halves of the season. The most dramatic jump in offensive performance took place between the first and second half of the 2015 season where there was an uptick in average exit velocity,

Table 7.1.1: Increasing Offensive Production In MLB

| Year | Runs Scored | Runs Per Game | Total Home Runs | Home Runs Per Game | Avg. Exit Velo | Avg Launch Angle | xRAA-Batted Balls |
|---|---|---|---|---|---|---|---|
| 2015 | 20647 | 4.25 | 4909 | 1.01 | 87.7 | 10.5 | 0.027 |
| 2016 | 21744 | 4.48 | 5610 | 1.16 | 88.1 | 11.2 | 0.037 |
| 2017 | 22582 | 4.65 | 6105 | 1.26 | 87.0 | 11.5 | 0.042 |

Table 7.1.2: Batted Ball Summary 2015-2017

| Year | Half | Avg Exit Speed | Avg Launch Angle | Avg HR Prob | xRAA |
|---|---|---|---|---|---|
| 2015 | First | 87.5 | 10.3 | 0.035 | 0.022 |
| 2015 | Second | 88.1 | 10.9 | 0.043 | 0.037 |
| 2016 | First | 88.1 | 11.0 | 0.044 | 0.037 |
| 2016 | Second | 88.0 | 11.6 | 0.043 | 0.036 |
| 2017 | First | 87.1 | 11.3 | 0.048 | 0.043 |
| 2017 | Second | 86.8 | 11.8 | 0.048 | 0.041 |

average launch speed, and very sharp jumps in the average home run probability and the xRAA. The offensive performance stabilized at about this level through the end of the 2016 season before another substantial jump at the start of the 2017 season that was sustained through the second half of the year as well.

While the increased emphasis on hitting the ball in the air has been offered as an explanation for the increase in scoring and home runs most in the baseball community believe that this alone could not have caused the observed increase in offensive production. Past upticks in offensive performance have been the result of the introduction of the live ball in the 1920s, the lowering of the mound after the 1968 season, and the steroid era in the late 1990s and the early 2000s. The most popular theory currently is that there have been changes in the ball that have led to the increase in offensive performance. We will begin by recapping the previous research that has been done regarding the juiced ball hypothesis. From there we will introduce our methodology for analyzing the effects of the juiced ball. This framework allows us to predict how batted balls would have performed had they been contacted in a different ball environment which we use to separate out the effects of the juiced ball and the fly revolution, to visualize how batted ball behavior has changed as a result of the juiced ball, and to examine the effects of the juiced ball on a player-by-player basis examining which batters have benefited the most and which pitchers have been the most harmed by the effects of the juiced ball.

## 7.2 Previous Research Done on the Juiced Ball

The first major piece done concerning the scoring surge in the second half of the 2015 season was published by Arthur and Lindbergh at FiveThirtyEight before the start of the 2016 MLB season [55]. After summarizing the magnitude of the increase in scoring and contrasting it with the decade-long trend of decreasing offensive, they consider several hypotheses for the increase in scoring in the second half of 2015: a smaller strike zone, the weather, better rookie batters, diminished pitching quality, and a bouncier baseball. Citing the work of Roegele who showed that the strike zone continued to expand in 2015 [56], they dismiss the first hypothesis. While limited in their ability to analyze humidity and wind, they dismiss temperature as the cause showing that it could not have produced an effect of the magnitude that was observed. One theory that was very popular at the time was that a considerably better than average rookie class of batters fueled the increase in offense [57]. Another popular theory was that the quality of pitching was down in the second half of the season as a result of injury, trade, or resting for the playoffs. In order to test this hypothesis, Arthur and Lindbergh develop a model to predict exit velocity based on the pitcher, the batter, the temperature, the count, the pitch velocity, and the called strike probability of the pitch. While their model does predict somewhat higher values later in the season, it dramatically underestimates the exit velocities observed in September and October. Additionally, they attempt to predict the outcomes of second-half matchups based on the quality of the players involved and observe

7% decrease in predicted strikeout rate, an 8% increase in walk rate, and a 31% increase in home run rate. These results combine to discredit the idea that the massive increase in offense is largely attributable to a tremendously successful rookie class or diminished league average pitcher quality. Lastly, they suggest that an increased liveliness of the baseballs would explain the increase well. They demonstrate how the average daily deviation from expected exit velocity began increasing after the All-Star break when there would have been a new batch of balls brought into circulation. Using the Sports Science Laboratory at Washington State University, they measured the coefficient of restitution of balls from the 2014 season and those from the second half of the 2015 season but found inconclusive results.

Midway through the 2016 season in [58], Nathan compares home run behavior in the first half of the 2015 season with home run behavior in the first half of the 2016 season. He concludes that the primary reason for the increase in home runs in 2016 is the result of more hard-hit balls as opposed to more balls hit at ideal home run trajectory demonstrating that there was 27% increase in batted balls hit between twenty-five and thirty degrees with an exit velocity of greater than 95 mph while there had only 7.5% increase in batted balls simply hit in the twenty-five to thirty degrees range. He repeats this result using different bin dimensions with similar results. In an addendum, he shows that while exit velocity is up by large amounts for balls with vertical launch of greater than ten degrees, it is very similar for balls hit at line drive angles of between zero and ten degrees. He interprets this as potential evidence against a juiced ball, as based on his physics knowledge, he would expect that if the ball was juiced line drives would show the greatest increase in exit velocity as a result of a higher coefficient of restitution.

Shortly after Nathan's article was published in The Hardball Times, Lindbergh and Arthur published a second piece at FiveThirtyEight with evidence in favor of a juiced ball [59]. They look at instances where batters and pitchers squared off against one another in the MLB and also at the highest level of minor league baseball, AAA, which uses a baseball manufactured in a different location. They built a mixed logit model with random effects for batter, pitcher, and park and a fixed effect for league, and found the league effect to be statistically significant in predicting home run probability.

In May of 2017, Lindbergh in [60] relays the findings of a report of the testings of MLB baseballs done at the Baseball Research Center at the University of Massachusetts Lowell. The results showed that game balls taken from five different teams were within MLB's manufacturing specifications and that the weight, circumference, and coefficient of restitution were comparable to past quality checks. This has been the strongest evidence against a juiced ball hypothesis so far.

However, a month later Lindbergh and Lichtman return to the juiced ball hypothesis [61]. They begin by recapping the continued increase in home runs and scoring, the rejected theories that had been previously considered, and the history of changes in the ball producing dramatic changes in Japanese baseball, the NCAA, and the Mexican League. They also argue that attributing the increase in offensive performance entirely to changes in batter philosophy is unlikely to have taken place so quickly. They also explain how the coefficient of restitution (effectively the bounciness of the ball) can affect the velocity of the ball, while the circumference and seam height can have a considerable impact on the air resistance of the ball which impacts how far the ball can carry. The publication focuses on a new set of laboratory testing commissioned by Lichtman using 17 game balls from before the 2015 All-Star break and from the 2016 season. The testings revealed significant differences in the coefficient of restitution, the seam height, and the circumference between the balls used in before the 2015 All-Star break and the balls used in 2016. They estimate that the differences in the batted balls could have produced results similar to what has been observed in Major League Baseball over the past few seasons. They also outline how lax MLB's allowable ranges are for their baseballs, and show that it could be possible for the ball to increase offensive performance while remaining within the boundaries. Lindbergh and Lichtman also acknowledge that many hitters have been intentionally trying to hit more balls in the air, suggesting that part of the reason for the rapid adoption has been that fly balls are becoming more profitable. They also call attention to a spike in the AAA home run rate that took place after Lindbergh's and Arthur's above work suggesting that the adoption of a flyball driven approach may be spreading to other levels.

The next week Arthur published [62] at FiveThirtyEight with additional evidence supporting a juiced ball. By measuring the amount of speed lost on four-seam fastballs from when the ball is released to when it crosses the plate, Arthur was able to approximate the average drag coefficient per month for baseball data from 2013 to 2017. He demonstrates that reduced monthly drag coefficients account for nearly 25 percent of the variation in the ratio of home runs to fly balls. He also finds evidence that the yearly drag coefficient has decreased from 2014 through 2017. He estimates that of the 47% increase in home run rate between 2014 and that point in 2017 that about half of it can be attributed to higher exit velocities from a springier ball, and that the remaining half of the increases can be attributed to changes in drag forces on the ball as well as adjustments made by batters across the league. In a follow-up study [63], Arthur shows that baseball's drag coefficient has become considerably more consistent from 2008 to 2017.

It would appear that the MLB is beginning to take some measures surrounding the juiced ball. Following the end of the 2017 season, MLB Commissioner Rob Manfred mandated that all teams must begin storing their game balls in a climate-controlled, enclosed room in an effort to create a more uniform ball environment across the league. Major League Baseball will be installing climate sensors in each such room to determine if a humidor, a device used to reduce the liveliness of batted balls by storing them in a moist and cool environment that has been effective in reducing offensive performance at the home of the Colorado Rockies, will be necessary league wide beginning in the 2019 season [64]. Additionally, Manfred has commissioned a team of scientists to investigate whether the ball was in fact juiced during the 2017 season [65].

At the time of this writing, the most recent major work in the public sphere on the juiced ball was done by Arthur and Dix at FiveThirtyEight [66]. Arthur and Dix discuss results of studies commissioned by "ESPN Sport Science" (ESPN is FiveThirtyEight's parent company) and performed by the Keck School of Medicine at the University of Southern California and Kent State University's Department of Chemistry and Biochemistry. The two studies reveal differences in the density and chemical composition of the cores of two groups of baseballs: four balls used in games between August 2014 and May 2015 and three balls used in games from August 2016 and July 2017 along with a brand new MLB ball directly from Rawlings. The balls were first analyzed at the Keck School of Medicine using CT scans that revealed the outer layer of the baseball's core was 40 percent less dense in the new group of balls. After the tests at USC, the balls were sent to Kent State where through thermogravimetric analysis it was found that the region of the core that was less dense in the new group was on average composed of seven percent more polymer and ten percent less silicon which would create a less dense core validating the results of the CT scan study. While the changes to the weight of the ball as result of the differences in core density would have a negligible impact on the flight of the ball, these changes in core composition coincide with the increase in the bounciness of the ball that has been previously observed. At this juncture, Arthur and Dix believe that more than half of the approximately 46 percent increase in home run rate can be attributed to the increased bounciness of the ball corresponding to an increase in velocity and the decreased air resistance of balls that would allow balls to carry further. He believes that the remainder of the increase in offensive performance could be attributed to the fly ball revolution.

## 7.3 Methodology

The core idea behind our analysis of the juiced ball is that using the batted ball classifier introduced above we can make a prediction for what the outcome of a batted ball would have been in a different ball environment. While most of the analyses above have focused on quantifying and describing the changes in the ball itself, we seek to predict at a batted ball-by-batted ball level what would have happened if an average ball from a particular ball environment had been used.

Since we do not have ball specific measurements for each baseball, we must find a way to indirectly capture the effects of changes in ball composition. While there is substantial in-batch variations within baseballs, we decide to capture ball environment effects on a season and half of season basis. For this reason, both season and the half of the season were included as predictors in the batted ball classifiers introduced to supplement the ball specific characteristics (exit velocity, launch angle). By using the season and half of season as proxies for measuring changes in average ball environment, we are making the assumption that there are no other changes that would impact the outcomes of batted ball that are changing by year or half of season that would not be picked up by the batted ball specific measurements of exit velocity and launch trajectory. We can obtain a counterfactual prediction for a batted ball under a different ball environment by inputting the batted ball characteristics and the counterfactual season and half of season into the batted ball classifier model. However, this is not quite sufficient. As previous research has suggested there seem to be two ways that the balls have been juiced: carrying further through less air resistance and coming off the bat faster through a springier ball. By inputting a counterfactual year and half, we would only be accounting for the reduced air resistance of the ball. We will also need to produce a counterfactual exit velocity for a different ball environment.

In order to obtain a counterfactual velocity prediction, we employ another LightGBM model but this time in a regression context. Exit velocity is modeled as a function of vertical launch angle, the proxy to horizontal launch angle, batter handedness, the season, and the half of the season. There is an obvious relationship between launch angle: balls hit at a line drive launch angle tend to be hit harder, and balls hit at very high vertical launch angles or very low launch angles reflect poor contact and thus lower velocity. We know that balls hits to a player's pull side are on average hit at a higher exit velocity which motivates the inclusion of horizontal launch angle and the batter's handedness. Season and half of season are included for capturing the changing ball environment's impact on velocity and can be manipulated to get counterfactual predictions. While these are all relevant predictors for estimating a batted ball's exit velocity, it is very limited. Ideally, we would like to have measurements such as the bat speed, the contact angle, where on the bat the ball hit, and where on the ball the bat hit. With this set of information, we would be much better equipped to predict batted ball exit velocities than with the current model. One other potential shortcoming in the velocity predictions is measurement errors in the Statcast data itself. The time span in question represents the first three years of Statcast's use in MLB. We feel it is very likely that the way raw information is measured and converted into the data used here has undergone adjustments during this time. While this was checked against, it is possible that there could be changes that would impact the results of this analysis

In order to find a good set of hyperparameters for the velocity model, we use the tree-structured Parzen estimator routine as used to fit the classification models. The best out-of-sample root mean squared error is 11.66 mph.

Below, we describe the full process employed here for acquiring the counterfactual batted ball predictions. For a particular batted ball, we predict the exit velocity both in the original batted ball environment and in the counterfactual ball environment (lines (1) and (2)). We then take the difference between the predicted velocity in the counterfactual ball environment and the predicted velocity in the original batted ball environment (line (3)), this difference is then added to the observed batted ball velocity (line (4)) to obtain an adjusted exit velocity to be used in the counterfactual prediction of the batted ball outcome which is done in line (5).

$$(1)\text{Velo}_{\text{Pred}} = LGBM_1(\text{Launch Angle}, \text{Horizontal Angle}, \text{Batter Handedness}, \text{Year}, \text{Half})$$

$$(2)\text{Velo}_{CF} = LGBM_1(\text{Launch Angle}, \text{Horizontal Angle}, \text{Batter Handedness}, \text{CF Year}, \text{CF Half})$$

$$(3)\text{Velo Difference} = \text{Velo}_{CF} - \text{Velo}_{\text{Pred}}$$

$$(4)\text{Adjusted Velo} = \text{Observed Velocity} + \text{Velo Difference}$$

$$(5)\text{CF Prediction} = LGBM_2(\text{Adjusted Velo}, \text{Launch Angle}, \text{Batter Handedness}, \text{CF Year}, \text{CF Half})$$

## 7.4 League-Wide Results

Before looking at its impact on production, we examine the effect of changes in ball environment on the exit velocity shown in table 7.4.1. Here we plot the velocity adjustments for a given ball environment when compared to the ball environment from the first of the 2015 season based on the vertical launch angle and the observed exit velocity. Agreeing with the findings of Nathan [58], we see for the batted ball environments from the second half of 2015 and both halves of 2016 that balls hit in the air display velocity increase of between one and two miles per hour. Also agreeing with his results, we find that for balls hit on a nearly flat trajectory there is little to no evidence of an increase in the bounciness of the ball and in fact exhibits slight evidence of the opposite. Nevertheless, we find strong evidence that the ball was causing increases in exit velocity for balls hit in the air. In the batted ball environments from both halves of the 2017 season, we have that exit velocities on balls hit with a launch angle of above twenty degrees are still elevated when compared to the balls used in the first half of 2015; however, there is a very strong decrease in exit velocity on balls hit below this launch angle. The elevated exit velocity on balls hit with higher launch angles continues to fit with the belief that the ball is juiced, but this drop in velocity for balls hit on flatter trajectories is extremely puzzling. It could be a product of the nature of the juiced ball, a relic of the changes in exit velocity measurement, or it could potentially be a product of more swings designed to hit the ball in the air. Additional research will be needed to confirm the source of this result. For now, we will operate under the assumption that it is in some way connected to the juiced ball.

In Table 7.4.1, we report the predicted xRAA for each batted ball and ball environment pairs. By moving down the columns, one can examine how the same set of batted balls would have been projected to have faired in different ball environments. For example, the first column denotes the batted balls from the first half of 2015 and by moving down the column, we can see how we would have expected the performance of batters from the first half of 2015 to have faired in other ball environments. By moving across the rows, we can examine how batted ball performance changed over time while keeping the batted ball environment fixed. In effect, we can estimate the change in true batting performance by fixing a row (ball environment) and moving across it. The juiced ball effect can be estimated by fixing a column (set of batted balls) and then moving down it.

Perhaps the single most interesting juncture in Table 7.4.1 is the change from the first half of 2015 to the second half. As we know, in reality, there was a dramatic increase in offensive performance during that time, but how much of that was due to the batters and how much to a juiced ball? According to the predictions, xRAA in the first half of 2015 with the first half of 2015 batted ball environment resulted in an xRAA of .0215. However, if those same set of batted balls had been hit in the batted ball environment from the second half of 2015 we would have expected an xRAA of .0258. This means that we believe that the juiced ball contributed an increase of about .004 runs on average which becomes a very large impact when we consider the total number of batted balls. However, we actually estimate that the impact in overall offensive performance might have been more strongly influenced by a change in true underlying batter performance as we estimate that the set of batted balls

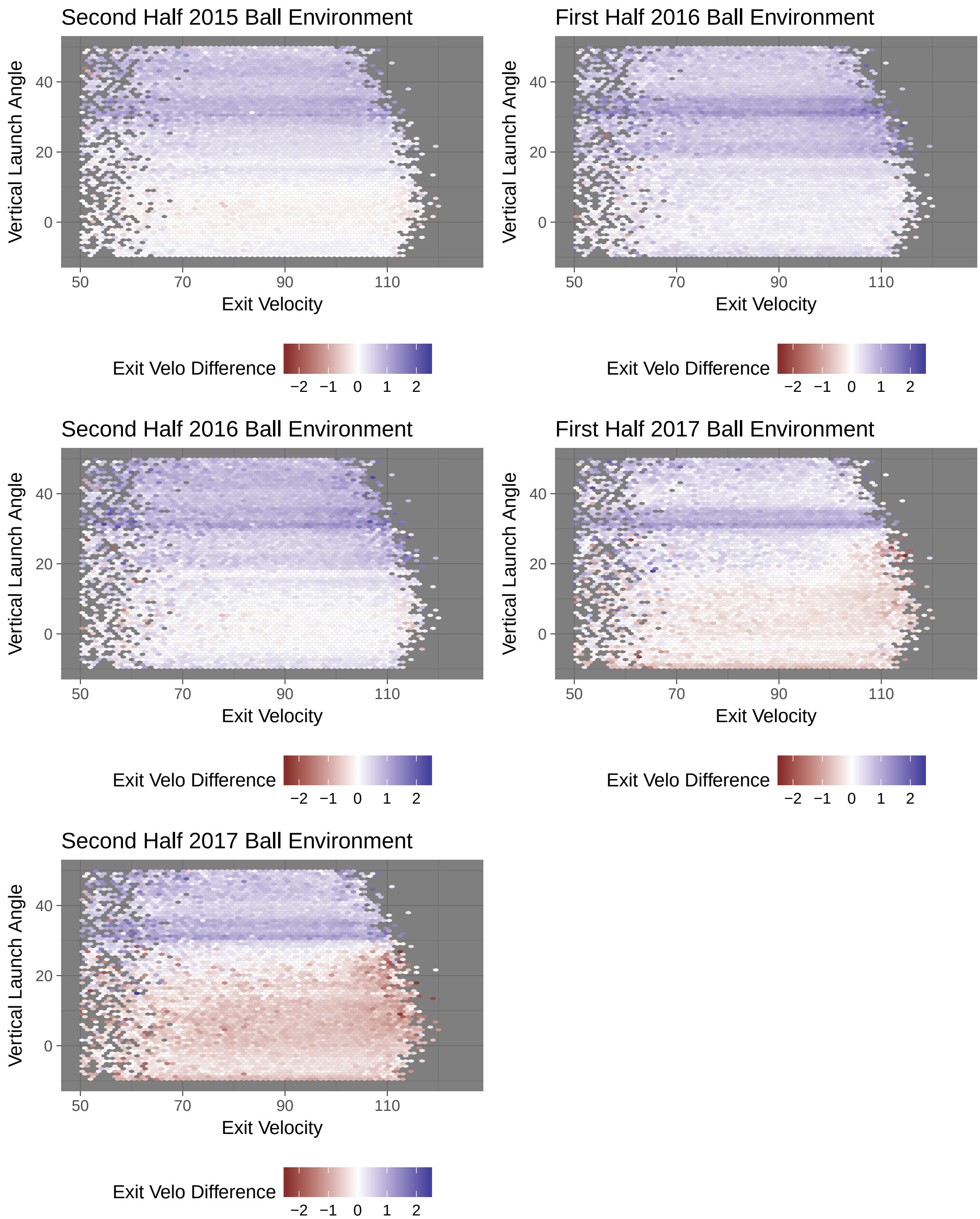


Figure 7.4.1: Velocity Adjustments for Batted Balls with Launch Angles between -10 degrees and 50 degrees

from the second half of the 2015 season would have produced an xRAA of .0296 in the batted ball environment from the first half of 2015. We observe that the sum of these two effects is less than the observed change in xRAA from the first half of 2015 to the second half of 2015, possibly a result of batters attempting to take advantage of a more profitable environment for balls hit in the air and adjusting accordingly.

Returning to the batted balls from the first half of the 2015 season, we can see that they would have performed best in the ball environments from the 2016 season and would have faired substantially better than the original batted ball environment in both halves of 2017 season, but not to the same capacity as we would have predicted them to do in 2016. This is very interesting as it is generally believed that the ball has been at its most "juiced" during the 2017 season. However, as we saw in figure 7.4.1, we estimate that exit velocity was down for balls hit at lower launch angles in both of the 2017 ball environments. These factors combined would serve to explain why it is that we see a substantial drop in xRAA by changing the batted ball environment from the second half of 2016 to the first half of 2017 for batted balls from the 2015 and 2016 seasons. However, this same drop off is not observed for batted balls from the 2017 season as we move from the ball environment from the second half of 2016 to the first half of 2017. We believe this is the result of shifting batter tendencies. By 2017, the adjustments made by batters attempting to drive the ball in the air and take advantage of the juiced ball would have let to more balls being hit at the trajectories where exit velocity was not adversely affected. It is also important to remember that there are multiple sources of the increased liveliness of the ball. While increased velocity may not have been what it was during the 2016 season, Arthur's results [58] suggest that the air resistance of the balls was at its lowest levels in 2017 which would have contributed to increased carry on balls hit in the air.

Another conclusion, that can be drawn from Table 7.4.1 is that the best actual offensive performance occurred during the second half of the 2015 season and during the 2017 season.

Table 7.4.1: xRAA by Batted Ball/Ball Environment Pairs

| | Batted Balls First Half 2015 | Batted Balls Second Half 2015 | Batted Balls First Half 2016 | Batted Balls Second Half 2016 | Batted Balls First Half 2017 | Batted Balls Second Half 2017 |
|---|---|---|---|---|---|---|
| Ball Environment First Half 2015 | 0.0215 | 0.0296 | 0.0265 | 0.0239 | 0.0277 | 0.0297 |
| Ball Environment Second Half 2015 | 0.0258 | 0.0371 | 0.0327 | 0.0302 | 0.0337 | 0.0360 |
| Ball Environment First Half 2016 | 0.0272 | 0.0381 | 0.0371 | 0.0335 | 0.0364 | 0.0385 |
| Ball Environment Second Half 2016 | 0.0292 | 0.0399 | 0.0376 | 0.0364 | 0.0383 | 0.0402 |
| Ball Environment First Half 2017 | 0.0252 | 0.0362 | 0.0340 | 0.0313 | 0.0427 | 0.0388 |
| Ball Environment Second Half 2017 | 0.0256 | 0.0364 | 0.0342 | 0.0313 | 0.0398 | 0.0413 |

Across all of the batted ball environments, these three halves have the best xRAA regardless of the batted ball environment. By moving across any of the rows, we can observe that there has been an upward shift in offensive performance across Major League Baseball even after we have accounted for the effect of the juiced ball. There is a very large spike as we move from the first half of 2015 into the second half of 2015, a decrease into 2016 albeit to levels still considerably higher than the first half of 2015, and then another substantial increase during the 2017 season. It would appear that either the juiced ball or the improvement in batter performance would have caused a considerable increase offensive performance, but it is has been the conjunction of the two together that has created the dramatically more offensive driven game that we observe today.

In Table 7.4.2, we display the results in the same manner as above but where batted balls have been grouped by year instead of by half in order to present a less overwhelming table. Once again, we can observe by moving down the columns that there has been a substantial impact from the juiced ball with effects as large as .0087 for 2015 batted balls, .0116 for 2016 batted balls, .013 for 2017 batted balls. Similarly, we can observe that there has been a substantial increase in overall batting performance from 2015 to 2016 to 2017. Moving along the rows, we see that the difference between the predicted performance on batted balls in 2017 and the batted balls from 2015 has been .0039, .0045, .0059, .0058, .0121, and .0107 respectively. It is interesting to note that how much better we estimate batted

Table 7.4.2: xRAA by Year and Ball Environment

| | Batted Balls 2015 | Batted Balls 2016 | Batted Balls 2017 |
|---|---|---|---|
| Ball Environment First Half 2015 | 0.0245 | 0.0256 | 0.0284 |
| Ball Environment Second Half 2015 | 0.0300 | 0.0318 | 0.0345 |
| Ball Environment First Half 2016 | 0.0312 | 0.0359 | 0.0371 |
| Ball Environment Second Half 2016 | 0.0332 | 0.0372 | 0.0390 |
| Ball Environment First Half 2017 | 0.0293 | 0.0331 | 0.0414 |
| Ball Environment Second Half 2017 | 0.0296 | 0.0332 | 0.0403 |

ball performance to have improved varies substantially across the ball environment. This is because the balls that benefit the most from the juiced ball environment are those that are driven in the air and emphasizing driving the ball in the air has been one of the hallmarks of the changes in batter approach during the 2015-2017 seasons. Even in the 2015 batted ball environment, batters would have benefited from trying to hit the ball in the air, but because of the fact that the balls are flying further as a result of the juiced ball, they are even more strongly incentivized to hit the ball in the air. It would be interesting to observe the counterfactual world in which the ball had never been juiced: would the change in approach in MLB been as strong if it not been further incentivized by the juiced ball?

In particular one of the biggest features of the increased offensive performance during the 2015-2017 seasons has been the dramatic increase in the number of home runs. In Tables 7.4.3, 7.4.4, and 7.4.5, we explore how average home run probabilities and home run totals in the same fashion as we did for xRAA in the above tables. While the overall conclusions are very much the same as with xRAA, it is worthwhile to consider the home runs in isolation as they have been key driver in the overall increase in offensive performance. Perhaps the most interesting additional contribution from the home run tables is that we have that in

Table 7.4.3: Avg. Predicted Home Run Probability by Batted Ball/Ball Environment Pairs

| | Batted Balls First Half 2015 | Batted Balls Second Half 2015 | Batted Balls First Half 2016 | Batted Balls Second Half 2016 | Batted Balls First Half 2017 | Batted Balls Second Half 2017 |
|---|---|---|---|---|---|---|
| Ball Environment First Half 2015 | 0.0347 | 0.0413 | 0.0395 | 0.0379 | 0.0434 | 0.0429 |
| Ball Environment Second Half 2015 | 0.0368 | 0.0435 | 0.0416 | 0.0401 | 0.0455 | 0.0452 |
| Ball Environment First Half 2016 | 0.0392 | 0.0462 | 0.0444 | 0.0428 | 0.0483 | 0.0479 |
| Ball Environment Second Half 2016 | 0.0398 | 0.0468 | 0.0449 | 0.0435 | 0.0489 | 0.0486 |
| Ball Environment First Half 2017 | 0.0392 | 0.0463 | 0.0443 | 0.0426 | 0.0483 | 0.0479 |
| Ball Environment Second Half 2017 | 0.0392 | 0.0461 | 0.0442 | 0.0423 | 0.0480 | 0.0475 |

terms of impact on home runs it would appear that both the ball environments from the 2016 contributed as strong and potentially a little stronger impact on the average probability of a home run than 2017. This is in line with what we saw above in Table 7.4.1; however, for batted balls from the 2015 and 2016 there was an observed substantial drop-off in xRAA when comparing the second half of the 2016 ball environment with the ball environment from the 2017 season. This effect is not observed in anywhere near the same magnitude in terms of the home run probabilities (though a very mild one is still present). This is in line with the fact that velocity on batted balls hit at lower launch angles was down in 2017 and the belief that air resistance of the balls was at a lower level in the 2017 season.

As was the case above with xRAA, we can see in 7.4.3 and 7.4.4 that once the effects of the juiced ball have been accounted for, the second half of the 2015 and both halves of the 2017 season produced the most prolific home run behavior once the juiced ball has been taken into effect.

Lastly, in Table 7.4.5, we aggregate batted balls based on year and display home run totals by ball environments in order to display the magnitude of the increase in home runs in a context for which we have intuition. We see that the ball environment from the second half of 2015 would have produced between 250-300 more home runs when we consider the batted balls from the 2015-2017 seasons. We see further evidence for a ball environment that is becoming progressively more juiced as we see another elevation in home run totals as we

Table 7.4.4: Total Predicted Home Runs by Batted Ball/Ball Environment Pairs

| | Batted Balls First Half 2015 | Batted Balls Second Half 2015 | Batted Balls First Half 2016 | Batted Balls Second Half 2016 | Batted Balls First Half 2017 | Batted Balls Second Half 2017 |
|---|---|---|---|---|---|---|
| Ball Environment First Half 2015 | 2818 | 1953 | 3261 | 1714 | 3573 | 1898 |
| Ball Environment Second Half 2015 | 2982 | 2056 | 3439 | 1815 | 3750 | 2001 |
| Ball Environment First Half 2016 | 3180 | 2187 | 3663 | 1936 | 3981 | 2121 |
| Ball Environment Second Half 2016 | 3227 | 2213 | 3706 | 1965 | 4028 | 2151 |
| Ball Environment First Half 2017 | 3177 | 2189 | 3660 | 1925 | 3973 | 2120 |
| Ball Environment Second Half 2017 | 3177 | 2180 | 3650 | 1915 | 3950 | 2105 |

[1] Recall that second half of a season refers to after the All Star Break.

Table 7.4.5: Total Predicted Home Runs by Year and Ball Environment

| | Batted Balls 2015 | Batted Balls 2016 | Batted Balls 2017 |
|---|---|---|---|
| Ball Environment First Half 2015 | 4771 | 4975 | 5471 |
| Ball Environment Second Half 2015 | 5038 | 5253 | 5752 |
| Ball Environment First Half 2016 | 5367 | 5599 | 6102 |
| Ball Environment Second Half 2016 | 5441 | 5671 | 6178 |
| Ball Environment First Half 2017 | 5366 | 5585 | 6093 |
| Ball Environment Second Half 2017 | 5357 | 5564 | 6055 |

examine the batted ball environments from 2016 and 2017. We also see the strong evidence that once ball environments have been accounted for that home run hitting performance has increased substantially across the league from 2015 to 2017 with estimates somewhere in the neighborhood of about 700 home runs. Furthermore, we observe that a similar number of home runs can be attributed the juiced balls for the ball environments from 2017.

The effects of the juiced ball can also be visualized in a very informative manner. In Figures 7.4.2 and 7.4.3, we plot the difference in predicted xRAA and home run probability respectively based on launch angle and exit velocity for each ball environment when compared to the first half of the 2015 season. In the first plot in both panels, we plot xRAA and home run probability respectively based on the launch angle and exit velocity to provide context for the differences displayed in the following plots.

In Figure 7.4.2, we see how the effects of the juiced have increased and changed over time. In the second plot, we observe the difference between the predicted xRAA for the second half of the 2015 ball environment and the original 2015 ball environment. We observe a light purple cluster on the edges of where the original home cluster shown in the first plot. This is the most prominent area of changes in xRAA as we move from one batted ball environment to the other. We see for the ball environments in 2016 that these areas expand and darken indicated a ball environment that displays stronger "juiced" effects. This cluster of increased xRAA seems to be at its largest in the second half of the 2016 ball environment. When we predict the outcomes for the 2015 balls under the 2017 ball environments, we have a differently shaped cluster that is somewhat thinner and focused around a dark center. We also have the development of prominent strip of decreased xRAA on the bottom edges of the home run cluster that was somewhat present before but becomes considerably strong in the 2017 ball environments as a result of the diminished exit velocity on balls hit at lower launch angles. Both of these same developments are seen in 7.4.3 confirming that they are a product of changes in home run behavior. The areas of increased home run probability and xRAA occur in areas that match with our intuition. The thickening of the strip in the 2017 ball environments fit with the hypothesis that the ball was less springy and resulted in less velocity. Low line drive home runs are dependent on a very high exit velocity to get out of the park and thus would be less common if the composition of the ball decreased the velocity of balls contacted in this manner.

In Figure 7.4.4, we present the differences in predicted home run probability for the batted balls from the first half of 2015 based on their original on field locations. In the first plot, we display the on-field home run probability under the first half of 2015 ball environment to

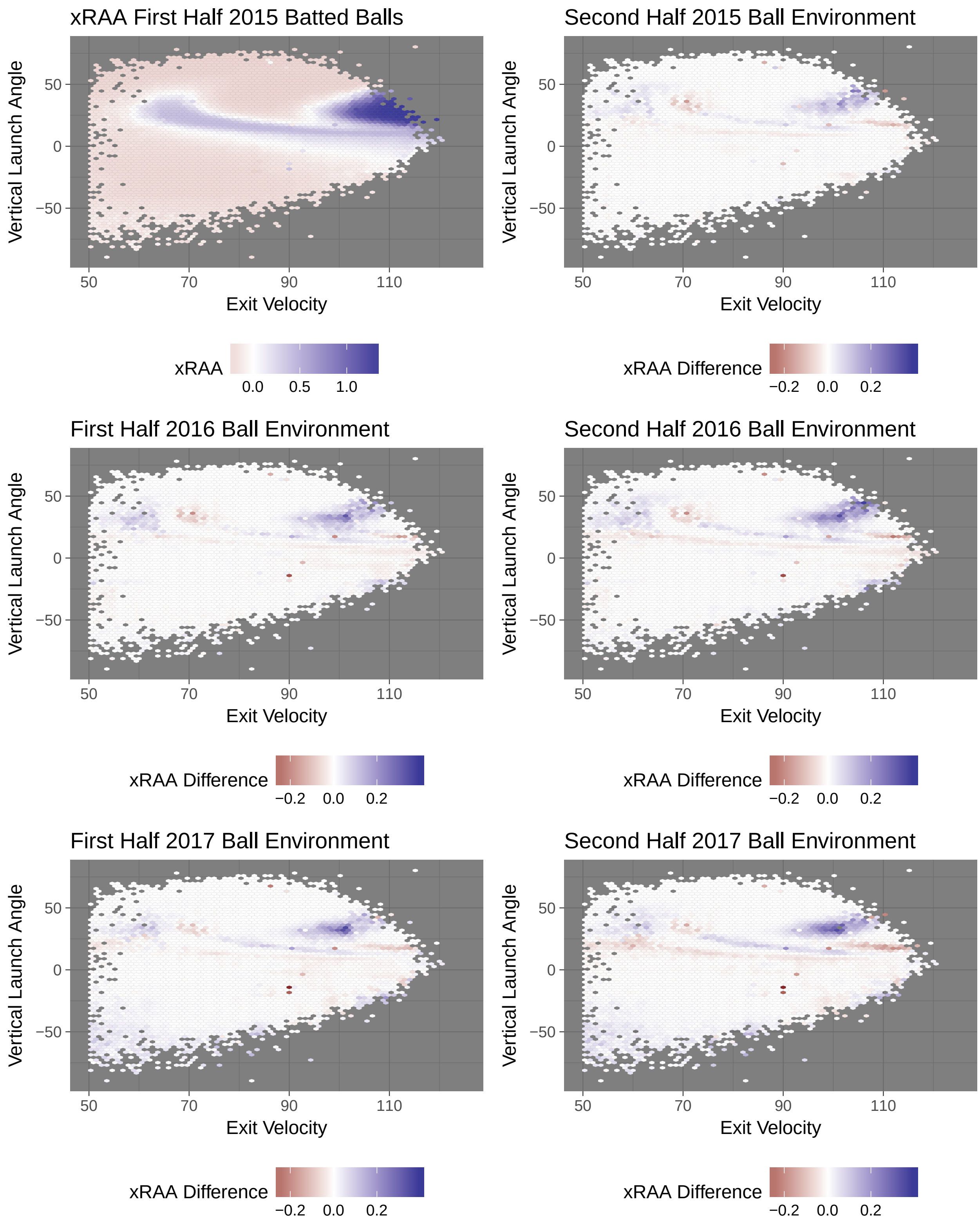


Figure 7.4.2: Counterfactual xRAA for Batted Balls from First Half of 2015

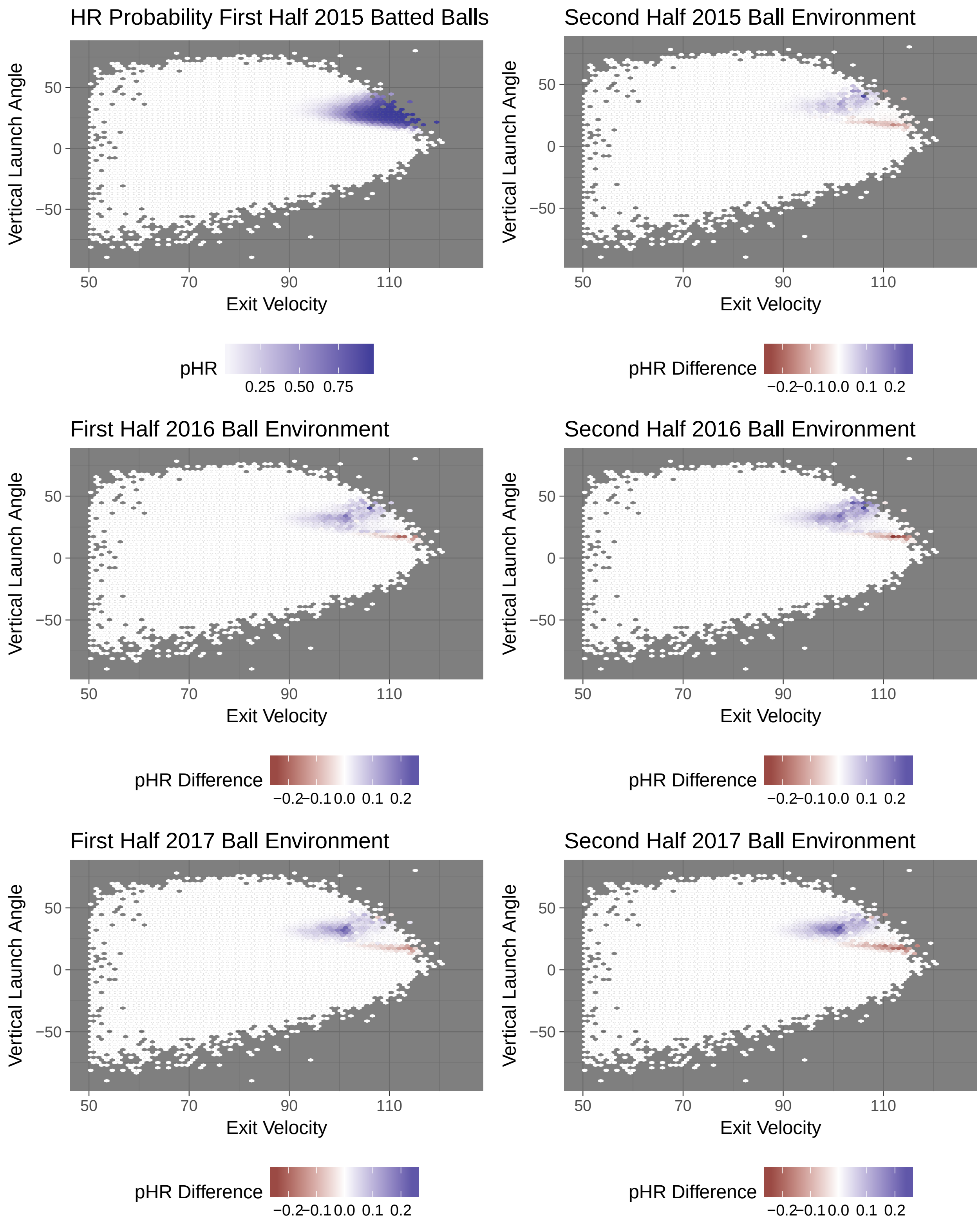


Figure 7.4.3: Counterfactual Home Run Probability for Batted Balls from First Half of 2015

provide context for the differences under the other ball environments. For the second half of 2015 ball environment, we can see that there is the start of a faint purple band in the deep outfield on the fringes of the high home run probability areas from the initial plot. We can see this area expand and darken in the both of the 2016 ball environments. There are balls with a moderate probability of home run that would have been zero under the original batted ball environment. Under the 2017 ball environment, we once again see areas with very sharp increases in home run probabilities that would have been fly balls in the original ball environment. One interesting development is the set of balls right along the left field line with a considerably decreased home run probability. This corresponds to the area of diminished home run probabilities at lower launch angles seen in Figure 7.4.3 as a result of diminished velocity on balls hit at lower launch angles.

## 7.5 Individual Batter Results

Perhaps even more interesting than examining league-wide effects of the batted ball is looking at how individual players have been affected by the juiced ball. There have been several players in recent years who have elevated their game from very good to great, and it is interesting to see if this is a result of individual improvement or if their batted ball profile benefited strongly from a juiced ball. Further, it may be possible to identify prospects who fit the profile of those who most benefit from the juiced ball. Additionally, if a team believes that a player has achieved success largely as a result of the juiced ball, they may wish to avoid paying those players at an elite rate if they believe their performance will return to more pedestrian levels if the ball environment was to revert back to a less juiced ball. On the reverse, teams may want to target flyball pitchers who have been disproportionately affected by the juiced ball if they believe that the ball environment will return towards the pre-spike levels.

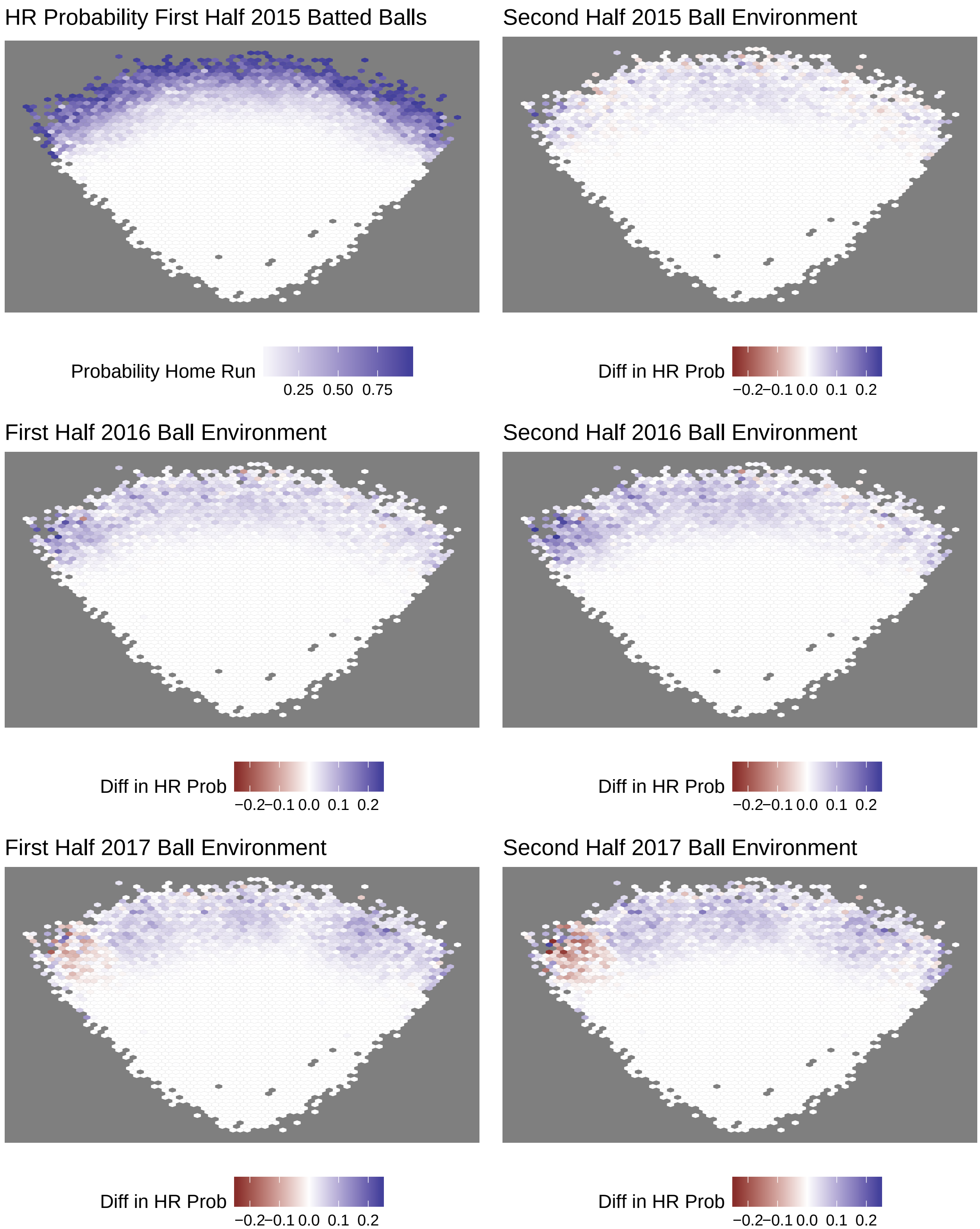


Figure 7.4.4: Counterfactual Home Run Probabilities on Field

In Table 7.5.1, we report those players from 2017 with at least 200 batted balls who had the largest increase in AxbRAA as a result of the ball environment from the 2017 season when compared to a batted ball environment from the first half of the 2015 season. We present similar results to above in Table 7.5.2 and Table 7.5.3 but for the case of benefit in total xbRAA and expected home run totals. In Table 7.5.1, we observe that the strongest beneficiaries gained between .027 and .036 runs per plate appearance as a result of the juiced ball in 2017 when compared to the first half of 2015. These players benefited more than twice as much from the juiced ball as the average individual. Perhaps the most interesting name on this list is Braves first baseman, Freddie Freeman. Freeman beginning in 2016 and extending to 2017 elevated himself from a very good player to one of the best hitters in all of baseball largely because of an increase in power, but these results suggest that some of this can be attributed to a change in the ball environment. One thing worth noticing that is reinforced in Tables 7.5.2 and 7.5.3 is that none of these players are generally regarded as having elite tier power (such as players like Aaron Judge, Giancarlo Stanton, and J.D. Martinez who had differences of .011, .0082, and .014 respectively). This makes sense as the batted balls that benefit the most from the juiced ball are those that would have come up just a bit short of a home run in a non-juiced environment. If one has elite power, when you make very solid contact with a ball, it does not change the outcome of that batted ball much if it goes 410 feet vs. 400 feet. Judge, Stanton and J.D. Martinez had the highest expected home run totals of 60, 53, and 48 respectively in the 2017 ball environment but were predicted for the same number, one more, and one less respectively under the first half of 2015 ball environment. Contrast that below with the results of Table 7.5.3 where players like Joey Votto and Freddie Freeman gained 6.17 and 5.69 expected home runs. Votto and Freeman are two of the best hitters in all of baseball in terms of making quality contact but are not prolific power hitters; however, they benefited from very substantial increases in home run totals based on the fact that they make a lot of quality contact and drive balls in the air without elite tier power. Others in the Table 7.5.3 fit a similar billing of having flyball tendencies with good but not

Table 7.5.1: Largest Average Benefit from Juiced Ball in 2017

| Player | Batted Balls | AxbRAA | AxbRAA CF | AxbRAA Diff | xbRAA | xbRAA CF | xbRAA Diff |
|---|---|---|---|---|---|---|---|
| Aledmys Diaz | 245 | -0.005 | -0.042 | 0.036 | -1.329 | -10.186 | 8.857 |
| Trea Turner | 333 | 0.043 | 0.010 | 0.033 | 14.477 | 3.492 | 10.986 |
| Luis Valbuena | 246 | 0.060 | 0.029 | 0.031 | 14.720 | 7.138 | 7.581 |
| Freddie Freeman | 347 | 0.132 | 0.102 | 0.030 | 45.803 | 35.235 | 10.568 |
| Keon Broxton | 240 | 0.121 | 0.092 | 0.029 | 29.081 | 22.128 | 6.953 |
| Willson Contreras | 281 | 0.094 | 0.065 | 0.029 | 26.396 | 18.341 | 8.054 |
| Yasmani Grandal | 311 | 0.048 | 0.020 | 0.028 | 15.008 | 6.297 | 8.712 |
| Matt Adams | 255 | 0.102 | 0.074 | 0.027 | 25.912 | 18.970 | 6.942 |
| Xander Bogaerts | 457 | 0.005 | -0.022 | 0.027 | 2.311 | -10.021 | 12.331 |
| Kevin Kiermaier | 282 | 0.030 | 0.003 | 0.027 | 8.475 | 0.947 | 7.527 |

[1] Mean increase in AxbRAA was .0145

[2] CF denotes the counterfactual prediction in ball environment from the first half of 2015

elite tier power and thus benefit the most from the added distance as the result of the juiced ball.

In Table 7.5.2 are the players who had the greatest total gain from the juiced ball over the course of the season. With the exception of Lindor who gained 9.9 expected runs above an average batter, all players gained 10 or more runs above an average batter over the course of the season and all posted above average benefits on a rate basis. Batters like Altuve, Galvis, and Merrifield highlight the profile of another set of batters who can strongly benefit from the juiced ball: those with very good bat to ball skills who put a lot of balls in play without having a high groundball rate.

## 7.6 Individual Pitcher Results

In Table 7.6.1 and Table 7.6.2, we look at the pitchers from the 2017 season who had the strongest adverse effects from the juiced ball. As is expected, both tables are dominated by pitchers with high average launch angles, with the exception of Trevor Cahill, a groundball pitcher who appears to have been incredibly unlucky as he posted a home run-to-fly ball ratio

Table 7.5.2: Largest Total Benefit from Juiced Ball in 2017

| Player | Batted Balls | AxbRAA | AxbRAA CF | AxbRAA Diff | xbRAA | xbRAA CF | xbRAA Diff |
|---|---|---|---|---|---|---|---|
| Charlie Blackmon | 512 | 0.090 | 0.065 | 0.026 | 46.33 | 33.16 | 13.2 |
| Joey Votto | 482 | 0.092 | 0.066 | 0.026 | 44.23 | 31.87 | 12.4 |
| Xander Bogaerts | 457 | 0.005 | -0.022 | 0.027 | 2.31 | -10.02 | 12.3 |
| Freddy Galvis | 501 | 0.007 | -0.016 | 0.023 | 3.52 | -7.83 | 11.4 |
| Jose Altuve | 510 | 0.053 | 0.031 | 0.022 | 27.15 | 15.90 | 11.2 |
| Trea Turner | 333 | 0.043 | 0.010 | 0.033 | 14.48 | 3.49 | 11.0 |
| Freddie Freeman | 347 | 0.132 | 0.102 | 0.030 | 45.80 | 35.23 | 10.6 |
| Whit Merrifield | 506 | 0.030 | 0.010 | 0.021 | 15.43 | 5.04 | 10.4 |
| Eric Hosmer | 501 | 0.056 | 0.036 | 0.020 | 28.27 | 18.19 | 10.1 |
| Francisco Lindor | 561 | 0.056 | 0.038 | 0.018 | 31.38 | 21.48 | 9.9 |

[1] Mean increase in AxbRAA was .0145

[2] CF denotes the counterfactual prediction in ball environment from the first half of 2015

of 25% this past season when league average is about 9.5%. The fact that fly ball pitchers have been the most harmed by the fly ball is not surprising. It is balls hit in the air that will have the most benefits from a ball that carries further, groundballs will be largely unaffected. The interesting thing to examine is how the breakdown of pitchers in the league will respond to the juiced ball. If the fly balls continue produce runs at an elevated rate, we will likely see ground ball pitchers become a more prized commodity both among already established Major League talent and also with regard to how pitchers are drafted and developed. Like predator and prey evolving in conjunction with one another, there is a constant co-evolution between batters and pitchers. In recent years, the hitters have been making adjustments by trying to hit the ball in the air more and have benefited tremendously from a juiced ball. The balance of power has shifted dramatically in the batters favor, but as recently the first half of 2015 the league was described as being in “the new age of mound excellence” [67] with scoring having dropped to levels not seen since 1981. For baseball fans and analysts everywhere watching this next step in the evolution of the league will be a fascinating one.

Table 7.5.3: Largest Total Benefit from Juiced Ball in 2017

| Player | Batted Balls | xbRAA | xbRAA CF | xbRAA Diff | xHR | CF xHR | xHR Diff |
|---|---|---|---|---|---|---|---|
| Joey Votto | 482 | 44.2 | 31.9 | 12.36 | 35.5 | 29.3 | 6.17 |
| Charlie Blackmon | 512 | 46.3 | 33.2 | 13.17 | 35.6 | 29.5 | 6.07 |
| Jay Bruce | 419 | 37.9 | 28.3 | 9.57 | 33.3 | 27.3 | 5.97 |
| Freddie Freeman | 347 | 45.8 | 35.2 | 10.57 | 35.9 | 30.2 | 5.69 |
| Kyle Seager | 474 | 27.5 | 18.0 | 9.53 | 30.4 | 25.3 | 5.12 |
| Eddie Rosario | 444 | 22.6 | 12.7 | 9.90 | 28.2 | 23.3 | 4.96 |
| Justin Smoak | 434 | 51.8 | 45.5 | 6.27 | 42.0 | 37.3 | 4.64 |
| Curtis Granderson | 329 | 19.5 | 10.8 | 8.62 | 25.7 | 21.3 | 4.39 |
| Cody Bellinger | 337 | 41.0 | 33.5 | 7.43 | 36.5 | 32.2 | 4.29 |
| Matt Carpenter | 377 | 29.3 | 22.7 | 6.58 | 27.4 | 23.3 | 4.11 |

[1] Mean increase in AxbRAA was .0145
[2] CF denotes the counterfactual prediction in ball environment from the first half of 2015

Table 7.6.1: Pitchers Most Hurt by the Juiced Ball on Average

| Player | Batted Balls | Launch Angle | AxpRAA | AxpRAA CF | AxpRAA Diff | xpRAA | xpRAA CF | xpRAA Diff |
|---|---|---|---|---|---|---|---|---|
| Trevor Cahill | 244 | 3.57 | 0.047 | 0.021 | 0.026 | 11.5 | 5.22 | 6.30 |
| Dinelson Lamet | 285 | 14.28 | 0.049 | 0.024 | 0.025 | 13.9 | 6.88 | 7.04 |
| Hector Neris | 201 | 16.47 | 0.053 | 0.028 | 0.025 | 10.6 | 5.63 | 4.95 |
| Trevor Williams | 455 | 9.41 | 0.033 | 0.009 | 0.024 | 15.2 | 4.27 | 10.94 |
| Wade LeBlanc | 210 | 11.06 | 0.065 | 0.042 | 0.023 | 13.6 | 8.88 | 4.76 |
| Anibal Sanchez | 343 | 16.67 | 0.078 | 0.055 | 0.022 | 26.6 | 18.90 | 7.72 |
| Chris Tillman | 325 | 12.83 | 0.098 | 0.076 | 0.022 | 31.8 | 24.64 | 7.19 |
| Jason Hammel | 602 | 15.68 | 0.069 | 0.048 | 0.021 | 41.5 | 28.61 | 12.86 |
| Tom Koehler | 227 | 12.86 | 0.068 | 0.046 | 0.021 | 15.3 | 10.50 | 4.85 |
| Daniel Norris | 324 | 13.92 | 0.074 | 0.053 | 0.021 | 24.0 | 17.23 | 6.79 |

[1] Mean increase in AxpRAA was .0134
[2] CF denotes the counterfactual prediction in ball environment from the first half of 2015

Table 7.6.2: Pitchers Most Hurt by Juiced Ball Total

| Player | Batted Balls | Avg Launch Angle | AxpRAA | AxpRAA CF | AxpRAA Diff | xpRAA | xpRAA CF | xpRAA Diff |
|---|---|---|---|---|---|---|---|---|
| Jason Hammel | 602 | 15.68 | 0.069 | 0.048 | 0.021 | 41.47 | 28.61 | 12.86 |
| Julio Teheran | 575 | 13.55 | 0.043 | 0.023 | 0.020 | 24.62 | 13.35 | 11.27 |
| Jordan Zimmerman | 556 | 16.91 | 0.071 | 0.051 | 0.020 | 39.56 | 28.48 | 11.08 |
| Martin Perez | 623 | 9.69 | 0.057 | 0.039 | 0.018 | 35.56 | 24.59 | 10.96 |
| Trevor Williams | 455 | 9.41 | 0.033 | 0.009 | 0.024 | 15.21 | 4.27 | 10.94 |
| Justin Verlander | 553 | 16.43 | 0.059 | 0.040 | 0.019 | 32.45 | 22.17 | 10.29 |
| Rick Porcello | 649 | 14.15 | 0.066 | 0.051 | 0.015 | 42.99 | 33.02 | 9.97 |
| Ervin Santana | 623 | 15.48 | 0.013 | -0.002 | 0.016 | 8.16 | -1.52 | 9.69 |
| Jason Vargas | 557 | 14.80 | 0.038 | 0.021 | 0.017 | 21.02 | 11.79 | 9.23 |
| Kevin Gausman | 560 | 11.65 | 0.070 | 0.053 | 0.016 | 38.98 | 29.84 | 9.14 |

[1] Mean increase in AxpRAA was .0134

[2] CF denotes the counterfactual prediction in ball environment from the first half of 2015

# Chapter 8

# Conclusion

In conclusion, we began this work by recapping the sabermetric foundations that the rest of the work was based on. We compared various metrics for both batters and pitchers highlighting the strengths and weaknesses of each. We also discussed the advantages of expected statistics. We then summarized the Statcast system and the specific Statcast data that fueled the analysis done here.

A brief summary of the ideas and theory behind the techniques used here was presented before introducing the batted ball classifier which serves as the backbone of the analyses done in the remainder of the work. We evaluated the model performances for the two versions of the classifiers in both a numeric and visual manner. In the next chapter, we used the results of batted ball classifier to form expected statistics for evaluating batting and pitching performance. We demonstrated that these metrics had properties superior to those of many traditional metrics that make the expected metrics superior in some capacities for evaluating the performance of MLB players. We also examined some of the best and worst performances according to these metrics and paid particular attention to how players could potentially overperform or underperform their expected statistics.

Lastly, we addressed the juiced ball and the fly ball revolution. We first summarized past research on the topic before presenting our novel method for predicting batted ball outcomes in a counterfactual ball environment. We used this technique to evaluate the impact of the juiced ball and changes in batter philosophy at both the league level and for individual players.

# Works Cited


[1] B. Lindbergh, R. Arthur, Statheads are the best free agent bargains in baseball, FiveThirtyEight. (2016). https://fivethirtyeight.com/features/statheads-are-the-best-free-agent-bargains-in-baseball/.

[2] M. Lewis, Moneyball: The art of winning an unfair game, W.W. Norton & Company, 2004.

[3] N. Dimeo, Baseball's particle accelerator, Slate. (2007). http://www.slate.com/articles/sports/sports_nut/2007/08/baseballs_particle_accelerator.html.

[4] M. Fast, What the heck is pitchf/x, The Hardball Times Baseball Annual. (2010). http://baseball.physics.illinois.edu/FastPFXGuide.pdf.

[5] MLB.com, ed., Statcast, (n.d.). http://m.mlb.com/glossary/statcast.

[6] MLB.com, ed., Statcast: Glossary of terms, (2015). https://www.mlb.com/news/major-league-baseballs-statcast-glossary-c-118508858.

[7] P. Casella, Statcast primer: Baseball will never be the same, (2015). https://www.mlb.com/news/statcast-primer-baseball-will-never-be-the-same/c-119234412.

[8] Official baseball rules, (2018). http://mlb.mlb.com/documents/0/8/0/268272080/2018_Official_Baseball_Rules.pdf.

[9] B. Baumer, A. Zimbalist, The sabermetric revolution: Assessing the growth of analytics in baseball, University of Pennsylvania Press, 2014.

[10] G.R. Lindsey, An investigation of strategies in baseball, Operations Research. 11 (1963) 477–501.

[11] J. Thorn, P. Palmer, The hidden game of baseball: A revolutionary approach to baseball and its statistics, The University of Chicago Press, 1985.

[12] T.M. Tango, M.G. Lichtman, A.E. Dolphin, The book: Playing the percentages in baseball, TMA Press, 2006.

[13] T. Tango, SABR 201: Linear weights by the 24 base/out states, 1999-2002, Tangotiger. (n.d.). http://www.tangotiger.net/lwbymob.htm.

[14] T. Tango, Linear weights by men on base, Tangotiger. (n.d.). http://www.tangotiger.net/lwbymob.htm.

[15] T. Tango, Sabermetrics 301: Custom linear weights, Tangotiger. (n.d.). http://www.tangotiger.net/customlwts.html.

[16] Linear weights, Fangraphs. (n.d.). https://www.fangraphs.com/library/principles/linear-weights/.

[17] V. McCracken, Pitching and defense: How much control do hurlers have?, Baseball Prospectus. (2001).

https://www.baseballprospectus.com/news/article/878/pitching-and-defense-how-much-control-do-hurlers-have/.

[18] T. Tango, Defensive responsibility spectrum (drs), Tangotiger. (n.d.). http://www.tangotiger.net/drspectrum.html.

[19] M. Swartz, Are pitching projections better than era estimators, Fangraphs. (2012). https://www.fangraphs.com/blogs/are-pitching-projections-better-than-era-estimators/.

[20] M. Swartz, New siera, part four (of five): Testing, Fangraphs. (2011). https://www.fangraphs.com/blogs/new-siera-part-four-of-five-testing/.

[21] SIERA, Fangraphs. (n.d.). https://www.fangraphs.com/library/pitching/siera/.

[22] SIERA, Baseball Prospectus. (n.d.). https://legacy.baseballprospectus.com/glossary/index.php?search=SIERA.

[23] M. Swartz, New siera, part one (of five): Pitchers with high strikeouts have low babips, Fangraphs. (2011). https://www.fangraphs.com/blogs/new-siera-part-one-of-five-pitchers-with-high-strikeouts-have-low-babips/.

[24] M. Swartz, New siera, part two (of five): Unlocking underrated pitching skills, Fangraphs. (2011). https://www.fangraphs.com/blogs/new-siera-part-two-of-five-unlocking-underrated-pitching-skills/.

[25] M. Swartz, New siera, part three (of five): Differences between xFIPS and sieras, Fangraphs. (2011). https://www.fangraphs.com/blogs/new-siera-part-three-of-five-differences-between-xfips-and-sieras/.

[26] M. Swartz, New siera, part five (of five): What didn't work, Fangraphs. (2011). https://www.fangraphs.com/blogs/new-siera-part-five-of-five-what-didnt-work/.

[27] R. Arthur, Baseball's new pitch-tracking system is just a bit outside, FiveThirtyEight. (2017). https://fivethirtyeight.com/features/baseballs-new-pitch-tracking-system-is-just-a-bit-outside/.

[28] R. Arthur, MLB's hit-tracking tool misses a lot of hits, FiveThirtyEight. (2016). https://fivethirtyeight.com/features/mlbs-hit-tracking-tool-misses-a-lot-of-hits/.

[29] D. Willman, Baseballsavant, (n.d.). https://baseballsavant.mlb.com/.

[30] B. Petti, Baseballr: A package written for r focused on baseball analysis, (2016).

[31] Fangraphs, Leaderboards data, (2018).

[32] J.H. Friedman, Greedy function approximation: A gradient boosting machine, Annals of Statistics. 29 (2001) 1189–1232.

[33] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, CoRR. (2016). http://arxiv.org/abs/1603.02754.

[34] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, LightGBM: A highly efficient gradient boosting decision tree, in: Advances in Neural Information Processing Systems 30, 2017: pp. 3146–3154.

[35] S. Ranka, V. Singh, Clouds: A decision tree classifier for large datasets, in: In Proceedings of the 4th Knowledge Discovery and Data Mining Conference, 1998: pp. 2–8.

[36] R. Jin, G. Agrawal, Communication and memory efficient parallel decision tree construction, in: In Proceedings of the 2003 Siam International Conference on Data Mining, 2003: pp. 119–129.

[37] P. Li, Q. Wu, C.J. Burges, McRank: Learning to rank using multiple classification and gradient boosting, in: Advances in Neural Information Processing Systems 20, 2008: pp. 897–904.

[38] J. Mockus, V. Tiesis, A. Zilinskas, The application of bayesian methods for seeking the extrememum,

Towards Global Optimization. 2 (1978) 117–129.

[39] H. Kushner, A new method for locating the maximum point of an arbitrary multipeak curve in the presence of noise, Journal of Basic Engineering. (1964) 86.

[40] J. Snoek, H. Larochelle, R.P. Adams, Practical bayesian optimization of machine learning algorithms, in: Advances in Neural Information Processing Systems 25, 2012: pp. 2951–2959.

[41] T.G. authors, GPyOpt: A bayesian optimization framework in python, (2016).

[42] J.S. Bergstra, R. Bardenet, Y. Bengio, B. Kégl, Algorithms for hyper-parameter optimization, in: Advances in Neural Information Processing Systems 24, 2011: pp. 2546–2554.

[43] T.H. Authors, Hyperopt: Distributed asynchronous hyper-parameter optimization, (2016).

[44] R. Arthur, The new science of hitting, FiveThirthyEight. (2016). https://fivethirtyeight.com/features/the-new-science-of-hitting/.

[45] B. Petti, Using statcast data to predict hits, The Hardball Times. (2016). https://www.fangraphs.com/tht/using-statcast-data-to-predict-hits/.

[46] P. Cutter, 525,600 minutes: How do you measure a player in a year?, Fangraphs. (2007). https://www.fangraphs.com/blogs/525600-minutes-how-do-you-measure-a-player-in-a-year/.

[47] R.A. Carleton, Baseball therapy: It's a small sample size after all, Baseball Prospectus. (2012). https://www.baseballprospectus.com/news/article/17659/baseball-therapy-its-a-small-sample-size-after-all/.

[48] R.A. Carleton, Baseball therapy: Should i worry about my favorite pitcher?, Baseball Prospectus. (2013). https://www.baseballprospectus.com/news/article/20516/baseball-therapy-should-i-worry-about-my-favorite-pitcher/.

[49] D. Carty, Resident fantasy genius: When hitters' stats stabilize, Baseball Prospectus. (2011). https://www.baseballprospectus.com/fantasy/article/14215/resident-fantasy-genius-when-hitters-stats-stabilize/.

[50] D. Carty, Resident fantasy genius: When pitchers' stats stabilize, Baseball Prospectus. (2011). https://www.baseballprospectus.com/fantasy/article/14293/resident-fantasy-genius-when-pitchers-stats-stabilize/.

[51] J. Pemstein, S. Dolinar, A new way to look at sample size, Fangraphs. (2015). https://www.fangraphs.com/blogs/a-new-way-to-look-at-sample-size/.

[52] S. Dolinar, J. Pemstein, A new way to look at sample size: Math supplement, Fangraphs. (2015). https://www.fangraphs.com/blogs/a-new-way-to-look-at-sample-size-math-supplement/.

[53] J. Pemstein, A long-needed update on reliability, Fangraphs. (2016). https://www.fangraphs.com/blogs/a-long-needed-update-on-reliability/.

[54] J. C, Detroit's batted ball readings are hot, Fangraphs Community Research. (2017). https://www.fangraphs.com/community/detroits-batted-ball-readings-are-hot/.

[55] R. Arthur, B. Lindbergh, A baseball mystery: The home run is back, and no one knows why, FiveThirtyEight. (2016). https://fivethirtyeight.com/features/a-baseball-mystery-the-home-run-is-back-and-no-one-knows-why/.

[56] J. Roegele, The 2015 strike zone, The Hardball Times. (2015). https://www.fangraphs.com/tht/the-2015-strike-zone/.

[57] O. Watson, The 2015 rookie class was the best in 100 years, Fangraphs. (2015). https://www.fangraphs.com/blogs/the-2015-rookie-class-was-the-best-in-100-years.

[58] A. Nathan, Exit speed and home runs, The Hardball Times. (2016). https://www.fangraphs.com/tht/exit-speed-and-home-runs/.

[59] B. Lindberg, R. Arthur, Are juiced balls the new steroids: The ball may be responsible for mlb's latest home run surge., FiveThirtyEight. (2016). https://fivethirtyeight.com/features/

are-juiced-balls-the-new-steroids/.

[60] B. Lindbergh, As home run rates rise, mlb offers evidence that the ball isn't juiced, The Ringer. (2017). https://www.theringer.com/2017/5/9/16040456/2017-mlb-home-run-rate-is-the-ball-juiced-report-results-6e1dd0233203.

[61] B. Lindbergh, The juiced ball is back, The Ringer. (2017). https://www.theringer.com/2017/6/14/16044264/2017-mlb-home-run-spike-juiced-ball-testing-reveal-155cd21108bc.

[62] R. Arthur, In mlb's new home run era, it's the baseballs that are juicing, FiveThirtyEight. (2017). https://fivethirtyeight.com/features/in-mlbs-new-home-run-era-its-the-baseballs-that-are-juicing/.

[63] R. Arthur, Baseballs are more consistently juiced than ever, FiveThirtyEight. (2017). https://fivethirtyeight.com/features/baseballs-are-more-consistently-juiced-than-ever/.

[64] T. Verducci, MLB to mandate that baseballs are stored in air conditioned room for 2018, Sports Illustrated. (2018). https://www.si.com/mlb/2018/02/23/mlb-mandate-teams-store-baseballs.

[65] M. Townsend, MLB is determined to find out if baseballs were 'juiced' in 2017, Yahoo! Sports. (2017). https://sports.yahoo.com/mlb-determined-find-baseballs-juiced-2017-035422507.html.

[66] R. Arthur, T. Dix, We x-rayed some mlb baseballs. here's what we found, FiveThirtyEight. (2018). https://fivethirtyeight.com/features/juiced-baseballs/.

[67] D. Schoenfield, The pitcher strikes back, ESPN. (2015). http://www.espn.com/mlb/story/_/id/12842711/mlb-hurlers-rise-dominance.