PROJECTING X 2.0

How to Forecast Baseball Player Performance

GB%, LD%, FB%   BABIP

HR/FB   K%

VERSION 2.0

BB%   WOBA

MIKE PODHORZER

Introduction

Welcome to *Projecting X 2.0*, the follow-up to the original *Projecting X*. Think of this edition as the pizza with even more of your favorite toppings, the ice cream sundae with an extra drizzle of delectable syrup, or the baseball game that treats you to a rare fireworks display after the final out is made (because, ya know, we all love fireworks, right?).

Though this version is not *completely* revised, it is most certainly updated with statistics through the 2015 season. Furthermore, and this is where your toppings, syrup, and fireworks make their highly anticipated appearances, *2.0* is chock full of new research, new metrics, new ideas, and new methods for projecting baseball player performance. And of course, the new metrics and methods are all explained in the included step-by-step guide to creating your own projections in Microsoft Excel.

Throughout these digital pages, you will find many links to websites, articles, and research studies. I encourage you to visit the content I have linked to in order to expand your knowledge base even further. The following sites play a vital role in my projection process and research:

FanGraphs – my primary source for statistics
Baseball-Reference – for several advanced pitching metrics
Pro Sports Transactions – injury and disabled list data
Baseball Heat Maps – hitter batted ball data
Brooks Baseball – PITCHf/x player cards
Roster Resource – depth charts and projected lineups/batting orders

Projection Systems

ZiPS. Steamer. Marcel. PECOTA. These are not the names of my pet hamsters or the latest fad diet trends. They are the names of some of the most well-known baseball player projection systems around. All of them derive their forecasts by mixing historical data in a blender and cooking up a projected performance line. These systems are pretty darn good, but for the majority of them, precious little is shared about their inner workings, making them essentially black boxes.

Accordingly, at some point, you might get the urge to project players yourself. And because you are reading this book, that urge may have already been felt. Why might you decide to project players on your own rather than rely on one of the aforementioned systems? Well for one, you may not agree with every projection a system generates. If you were to project a player yourself, you would hold full control over every individual statistical measure in the forecast (strikeout and walk percentages, BABIP, etc.) and be granted the opportunity to instantly discover how a change in one metric affects another.

Furthermore, the forecasting process requires a combination of both art and science. Since the projection systems are built using sophisticated computer models, they have the science part

Chapter 1: Projecting Hitter Performance

Plate Appearances (PA)

Projecting playing time is just as important as projecting performance. Unfortunately, there is not a whole lot of science behind this part of the forecasting process. It is the one offensive category that needs constant tweaking leading right up to the start of the season as position battles are won, injuries occur, and recovery timetables become clearer.

Aside from the player's historical totals, a plate appearance projection should take into account the hitter's spot in the batting order. John Dewan, co-founder of STATS, Inc., found that on average, the next slot in the batting order receives about 18 fewer plate appearances than the previous one. Also worth considering is how good the offense is that the batter hits in, as a better offense will turn the lineup over more frequently, resulting in additional plate appearances for all.

Fortunately, we now have more tools at our disposal to assist with factoring in a player's lineup spot and his team's projected runs scored total. *Baseball-Reference.com* features a page of myriad splits (this is for all of MLB, these are the AL and NL pages), including those by spot in the batting order. The table features various metrics broken down by lineup slot. Ever been curious how many runs scored the average leadoff hitter tallies over 650 plate appearances? Simple math could tell you (in 2015, it was approximately 85).

Of course, the table also allows you to calculate the average number of plate appearances recorded by players filling each spot. Remember that it's rare for a player to play in all 162 games and to hit in the same lineup spot in 100% of his appearances, so do not makes the mistake of automatically forecasting an American League leadoff hitter to record the 750 plate appearances the first spot in the batting order yielded on average in 2015.

Projecting a team's offense isn't as difficult as you may think. I go into more detail in Chapter 5, but the idea is to gather various standings projections, which typically include runs scored and runs allowed forecasts. Bump up your plate appearance forecast if the player's surrounding offensive cast has improved over the off-season and reduce it if it has weakened.

Tanner Bell of *Smart Fantasy Baseball* takes the lineup spot and team runs scored projections one step further and explains how you could develop a formula to calculate plate appearances by simply inputting the hitter's expected lineup spot. You could use what the formula spits out to help guide your own forecast, as it works well for full-time starters. I caution the use of the formula for reserve players, though (reserve players will enter the game at different times and in various roles, making use of the formula for these types problematic). Because the formula cannot be used for all hitters, it is not included in the Excel tutorial.

When projecting plate appearances, it is important to consider injury risk. Health is a skill and an injury-prone player usually rarely becomes an iron man candidate. In applying this concept, it is vital to distinguish between a freak injury, such as a player crashing into the outfield wall that damages his knee, and a standard everyday muscular injury, which has a higher probability of

Chapter 2: Projecting Pitcher Performance

Innings Pitched (IP)

Perhaps surprisingly, innings pitched does not correlate very well on a year-to-year basis, finishing in the bottom half of the metrics tested. Obviously, injuries are a primary culprit here, which more often rob pitchers of significant playing time than hitters. As usual, age plays a meaningful role. Sabermetrics extraordinaire Tangotiger found that a pitcher's seasonal innings total increases until its peak at age 27 and then declines consistently thereafter.

When projecting a pitcher's innings total, let the previous season(s) and his age guide you. Of the top 20 pitchers in innings pitched in 2015, only four of them were over the age of 30, and one of those was a knuckleballer, for whom the rules do not apply.

Minor league seasons are shorter than Major League seasons. As a result, minor league starting pitchers record fewer innings than their Major League counterparts. When projecting rookie pitchers, it would be unwise to forecast a significant jump in innings pitched or a total that exceeds 200. In fact, since 2000, only three rookie starting pitchers have eclipsed the 200 innings pitched barrier.

Similar to your common sense check after finishing your offensive projections for a particular team, it is essential that you do the same for your starting pitcher projections. Add up your games started projections for all the team's pitchers and ensure the total makes logical sense. Obviously, with 162 games in the season, you should not be forecasting a particular team's starters to start more than that many games. It is okay to be under that total as minor leaguers are recalled that you did not project, but do not go above it!

Benchmarks

The league leaders in innings pitched typically throw 230 to 240 frames. The top rookies and fifth starters who manage to stick in the rotation all season will normally throw 160 to 180 innings. Fifth starters don't start enough games to exceed the high end of that range and rookies are often saddled with innings limits to preserve their arms. For relief pitchers that remain in a Major League bullpen all year, between 55 and 80 innings is the sweet spot, with lower totals recorded by the situational guys.

Strikeout Percentage (K%)

A pitcher's strikeout rate is one of the most stable skills he possesses and it has one of the highest year-to-year correlations among all pitching metrics. That is a positive for forecasters, as it makes it that much easier to project.

As you might expect, pitchers lose their strikeout ability as they age. According to research by *FanGraphs* writers Bill Petti and Jeff Zimmerman, a starting pitcher maintains a K/9 rate within a 0.5 point range through age 29. After that age, the rate declines steadily, as he loses about a

Chapter 3: Creating an Excel Spreadsheet to Project Hitter Performance

Now that you have become an expert at analyzing the statistics and metrics used in forecasting baseball player performance, it is time to put those newly learned skills into action. Over the next several chapters, I will teach you how to create an Excel spreadsheet to generate your projections and walk you through the projection process for a hitter and a pitcher.

Begin by opening a new Excel workbook. Next, click on the "File" tab, open the "Options" dialog box, navigate to the "Formulas" section, and make sure that "Enable iterative calculation" is checked off. Some of the formulas you will be inputting into your spreadsheet will create a circular reference, which happens when a cell references itself in a calculation, either directly or indirectly. Checking this option will prevent an error from occurring and guarantee an accurate value is displayed. You may have to check off this option again each time you open Excel.
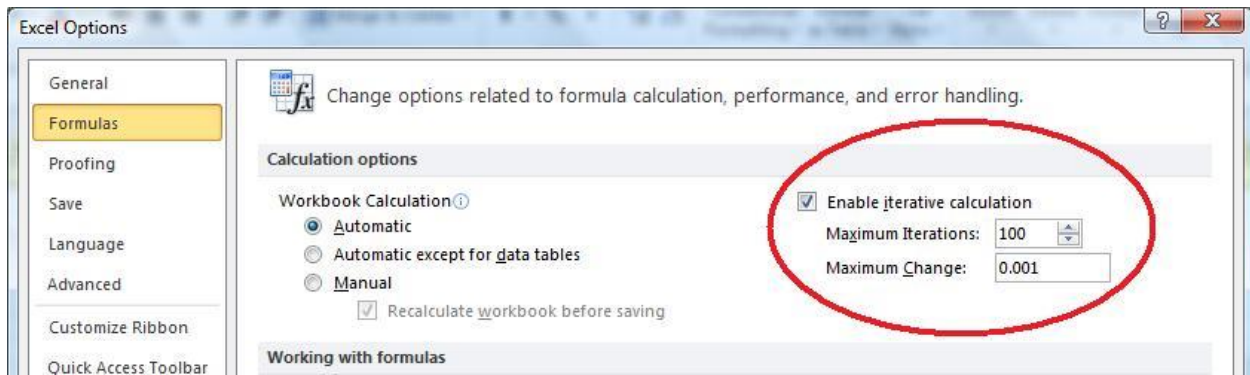
Now, let us start creating our hitter projection spreadsheet. For our hitter example, we will use Paul Goldschmidt, mashing first baseman for the Arizona Diamondbacks.

Fill in the cells in rows 1 and 2 to match the below three screen shots:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Player	Lg	Team	Pos	AB	PA	HITS	1B	2B	3B	HR	R	RBI	BB	IBB	SO
2	Paul Goldschmidt	NL	ARI	1B												

	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE
1	HBP	SF	SB	CS	AVG	OBP	SLG	OPS	ISO	wOBA	BIP	AB/2B	AB/3B	PA/HBP	PA/SF
2															

	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ
1	SBA/TOB	SB%	BB%	IBB%	K%	BABIP	GB%	LD%	FB%	HR/FB	R/TOB	RBI/BIP
2												

The highlighted cells will function as your inputs, while the cells without a fill color will include formulas. We will type in all the formulas first, as follows:

Chapter 4: Developing a Hitter Projection

Now comes the fun part, developing your projection! As discussed earlier, you will be filling in the yellow highlighted cells, and then the formulas in the non-highlighted cells will work their magic and reveal the calculated values automatically. Unless otherwise noted, all historical statistic screen shots are from *FanGraphs*.

We will start in cell F2 to project plate appearances. Below is how Goldschmidt has performed in the category over the last three seasons (depending on the hitter in question, you may choose to look at more than three season's worth of performance):

Season	Team	G	AB	PA
2013	Diamondbacks	160	602	710
2014	Diamondbacks	109	406	479
2015	Diamondbacks	159	567	695

We first notice that dip in plate appearances in 2014, which would require further research if we are unable to recall what led to the missed time. Turns out, Goldschmidt was hit by a pitch which fractured his left hand and ended his season prematurely. This is a perfect example of a freak injury that should not reduce future playing time projections.

Goldschmidt will be hitting in the same slot in the Diamondbacks batting order as he did in 2015 (third) and will be 28 years old for the majority of the 2016 season. Therefore, a 685 PA projection seems appropriate and we type that into cell F2. You will notice that suddenly the AB and BIP columns now show a projection as their formulas reference the plate appearance cell, while some of the rate categories have become 0.000, rather than an error.

We then meander over to the next pair of input cells, skipping a group of formulas, and find ourselves hovering over AB2 and AC2 in order to project doubles and triples rates. Below is Goldschmidt's performance in those two categories over the past three seasons:

Season	Team	AB	2B	3B	AB/2B	AB/3B
2013	Diamondbacks	602	36	3	16.7	201
2014	Diamondbacks	406	39	1	10.4	406
2015	Diamondbacks	567	38	2	14.9	284

Above, I added calculations for the rate of at-bats per double (AB/2B) and at-bats per triple (AB/3B) so you could more effectively forecast Goldschmidt's 2016 rates of extra-base hits. Admittedly, it could be cumbersome to calculate historical rates for every single player. You have two options to speed up the process:

1) In a new worksheet, copy and paste Goldschmidt's historical stats from a website like *FanGraphs* and then insert the appropriate formulas next to the data. Then you can