Dylan Guzman

Professor Agnellini

DIDA 325

May 9, 2024

What Elements of a Pitch Induce Soft Contact?

A new age of MLB data is readily available to MLB teams due to a new advanced tracking technology called Statcast. Statcast tracks the baseball at all times, and can offer how fast it's going, how much a pitch has moved, and how far it has traveled. The MLB is more competitive than ever because of it, as teams have a brand new field of statistics to compare players with and determine who would be the most valuable to their team. One of the most valued Statcast statistics by MLB teams is exit velocity, which measures how hard the baseball is hit. This paper will attempt to investigate which aspects of a pitch, such as velocity, spin rate, and vertical movement, have the greatest impact on its ability to result in low exit velocity. It will begin by explaining the choice to examine exit velocity, followed by visualizations and an explanation of the statistics, and will ultimately lead to the conclusion that these variables have little effect on exit velocity.

So why is exit velocity so valued among MLB teams? It's because it's thought to be strongly correlated with a batter's overall performance. Take hard-hit balls for example. MLB.com defines a hard hit as any baseball with an exit velocity of at least 95 miles per hour, claiming that it's the cutoff where "exit velocity begins to matter".[1] This is because there is a

[1] *Hard-hit Rate*

steep increase in positive results for the batter once the ball is hit at least 95 MPH. The source uses results from the 2018 season to explain why: balls hit under 95 MPH resulted in a hit only 22% of the time, where hard-hit balls were a hit 52% of the time. No matter how soft the ball was hit, whether it be 30 MPH or 80 MPH, the results hovered around that 22% mark. However, as the exit velocity increases past 95 MPH, the player's batting average rapidly increases. In 2018, balls hit at least 100 MPH resulted in a hit 62% of the time, and balls hit at least 110 MPH had a batting average of 76% of the time. MLB teams value hard-hit baseballs because they are much more likely to result in a hit. Each individual pitch can either be a hit or an out, meaning any test regarding pitches and results would have to be a categorical analysis. So, in order to carry out a quantitative analysis, MLB teams use exit velocity instead, due to its belief to lead to positive results.

The baseball community has provided mixed results in regards to what a pitcher can do to prevent high exit velocities. Some people claim that pitchers have little control over exit velocity. For instance, baseball writer Jeff Zimmerman did a study on how the location and horizontal movement of a pitch can lead to weak contact. He came to the conclusion that they held little effect on exit velocity, stating "I've found it's nearly impossible to constantly generate desirable outcomes".[2] Desirable outcomes in this case referring to low exit velocities. On the flip side, writer Matt Pettit posted results showing that the probability of soft contact generally increases with the pitcher's velocity.[3] Overall, there is no consensus about whether or not pitchers have an effect on exit velocity, and what factors of a pitch would have an effect on exit velocity. This study is an attempt to get closer to a clear consensus on the issue.

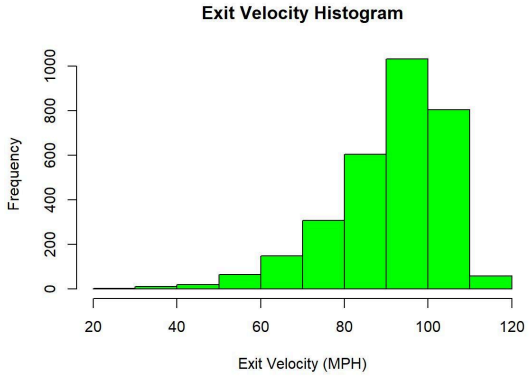
---

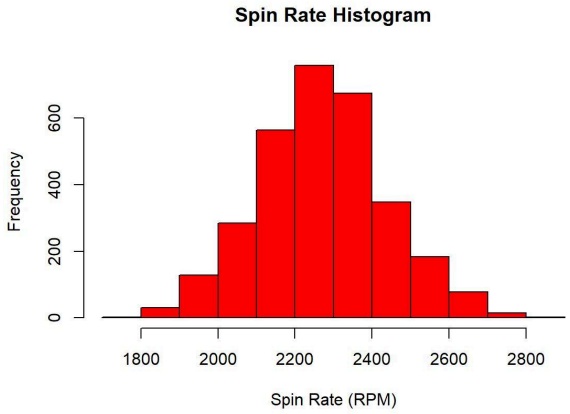
[2] Jeff Zimmerman, *Generating Weak Contact: Bringing it All Together*
[3] Matt Pettit, *Spin Rate and Soft Contact Probabilities*

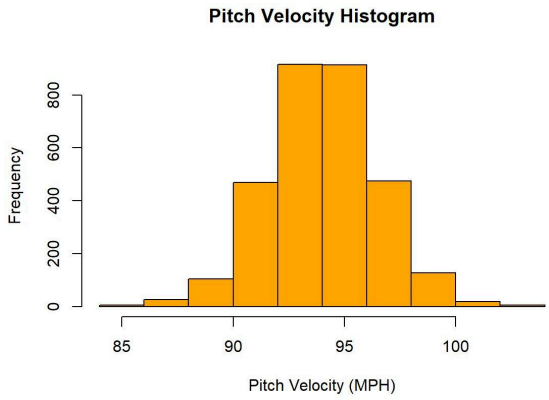
The dataset chosen for this test is a random two-week sample of all four-seam fastballs put in play from June 5, 2023 to June 19, 2023, which was chosen this way for a few reasons. First, a two-week sample was chosen to prevent overfitting in a machine learning sample. Additionally, a two week sample included over 3000 observations, and should be large enough for a successful study. Second, the sample includes only four-seam fastballs to reduce variability among different pitch types. A curveball behaves vastly differently from a fastball, and they would produce different results. By focusing on one type of pitch, the study can expect a stronger result. An additional thing to note about this dataset is that it excludes any balls that were bunted. Bunts are purposely supposed to be hit softly, and do not represent exit velocities when hitters are actually swinging. If bunts were included, they would skew the dataset as those results yielded vastly lower exit velocities than normal swings.

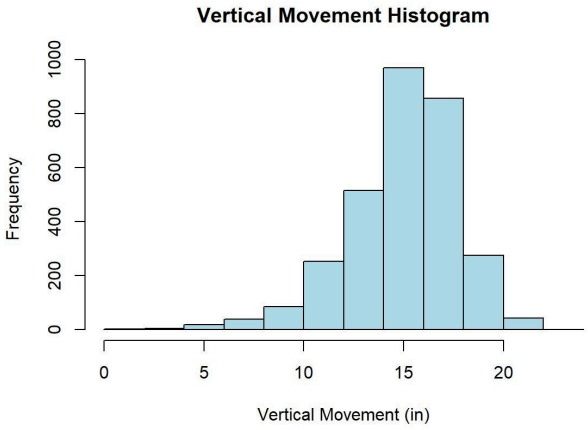
This study will utilize results from the search engine on baseballsavant.com, which is an official MLB website for all things statcast related. Factors that will be evaluated include pitch velocity, spin rate, and vertical movement. Pitch velocity measures how fast the pitch is thrown, and is measured in miles per hour. Spin rate measures how fast the pitch spins, and is measured in revolutions per minute. Vertical movement measures the difference between the height of the pitch when it was thrown and when it was caught by the catcher, or hit by the batter, and is measured in inches. Ultimately, this study will attempt to find what factors of a four-seam fastball lead to the weakest exit velocity. In order to do so, the exit velocity of each ball put in play from the chosen dataset will be compared to the pitch's velocity, spin rate, and vertical movement, in hopes of finding a correlation between the two.

Below are histograms for the four variables used in the study. It appears that spin rate and pitch velocity roughly have a normal distribution centered around 93 MPH, and 2250 RPM respectively. Exit velocity and vertical movement are both skewed to the left, with centers around 95 MPH, and 15 inches respectively.

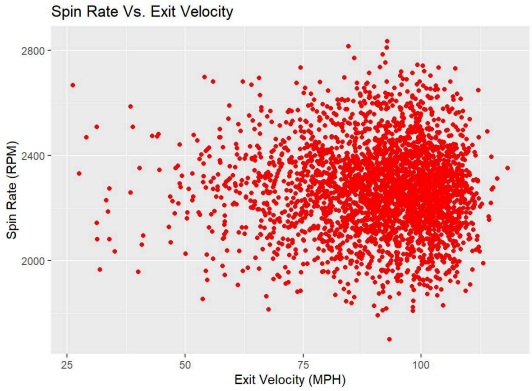


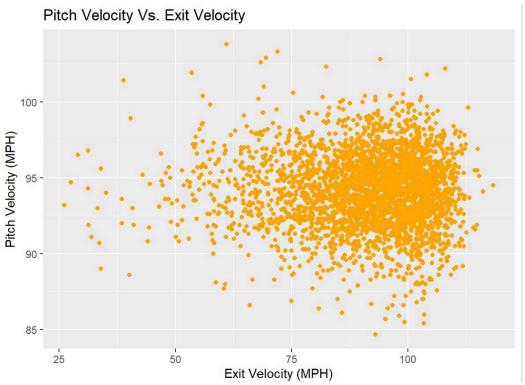


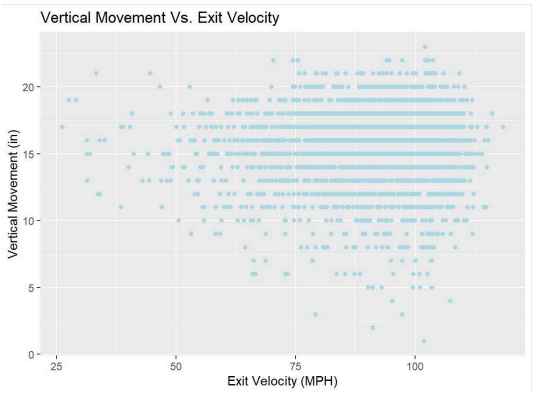




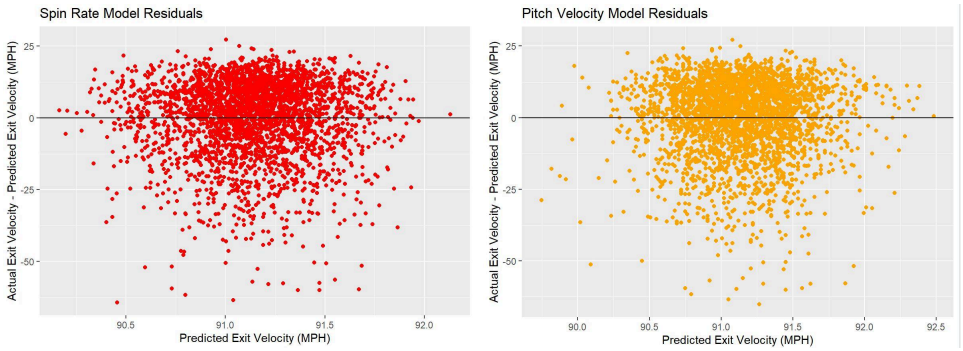
The scatter plots for these variables are quite telling. All three graphs are spread out toward the lower values, then bunch together as the variables increase. However, all three graphs very clearly show that there is not a linear regression between the variables:

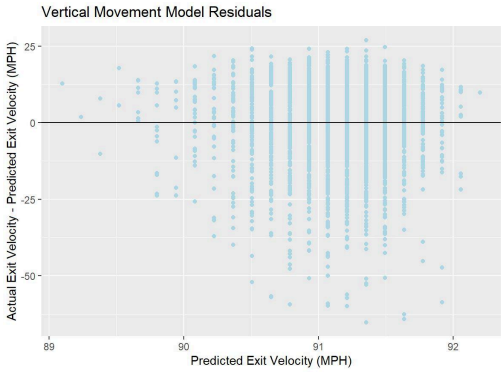






This claim is easily backed up by the r-values between the variables. Exit velocity has a -.0260, -.0212, and .0288 r-value with pitch velocity, spin rate, and vertical movement respectively. The r-values suggest that there is little to no correlation between exit velocity and pitch velocity, spin rate, or vertical movement. This can be further proven by plotting the residuals of their line of best fit models.





While the mean of the residuals for all three models are extremely close to zero, the distribution of residuals are not evenly distributed. The maximum residual for all three models are around 25, where the minimum of the models hover around -60.

The information from the study demonstrates that there is not a relationship between the variables of a pitch such as velocity, spin rate, and vertical movement and the exit velocity exerted on said pitch. When examining a pitcher's four-seam fastball, these statistics cannot be used on their own in order to predict how well the pitch can induce low exit velocity. In the future, it would be interesting to see how well only aspects of a pitch can have an effect on exit velocity, such as pitch location, and horizontal movement. Or, if given more time, a multi-linear regression analysis could pose more enticing results.

Sources:

"Hard-Hit Rate: Glossary." *MLB.Com*, www.mlb.com/glossary/statcast/hard-hit-rate. Accessed 9 May 2024.

Matt Petitt on January 23, 2017. "Spin Rate and Soft Contact Probabilities." *Exploring Baseball Data with R*, 23 Jan. 2017, baseballwithr.wordpress.com/2017/01/23/spin-rate-and-soft-contact/.

Zimmerman, Jeff. "Generating Weak Contact: Bringing It All Together." *RotoGraphs Fantasy Baseball*, 23 Sept. 2021, fantasy.fangraphs.com/generating-weak-contact-bringing-it-all-together/#:~:text=The%20batter%20chases%20a%20pitch,or%20handle%20of%20the%20bat.