

Comparing Baseball Player Performance Metrics: NCAA (D1) vs. Frontier League

Zac Ambrose, Robert Cowan, Nathaniel Ascher, Youssef Gehad

Table of contents

Introduction	2
The Data	2
Random Forest	3
Pitch Metric Definitions	3
Velocity	3
Spin Rate	3
Horizontal Break	3
Induced Vertical Break	4
Pitch Type Definitions	4
Fastballs, Sinkers, and Cutters	4
Offspeed	4
Breaking Balls	5
3D Plot: Random Forest Auto Pitch Type Visualization	5
Pitch Type by Inning	6
Release Speed by Inning	7
Spin Rate by Pitch Type	8
Dead Zone Fastballs and Sinkers	9
Situational Pitch Type Run Values and Exit Speeds: Killer Plot	10
Introduction	10
Specific Combinations	11
Summary of Results	13
Run Value and Release Speed Regression	13
Conclusion	14
Overall Takeaways	14
Future Analysis Opportunities	14
References	15

Introduction

Every year, MLB teams scour the globe to find the greatest talent to improve their teams. It culminates in a 20-round draft where hundreds of prospects from college are selected, and hundreds more are signed into ‘farm systems’, which are lower-level professional teams that send their talent, once developed, up to the MLB level. With millions of dollars and armies of scouts and analysts, teams should be able to accurately determine who are the best prospects. But, how does MLB determine who the best prospects are, which ones are worth drafting, and which ones aren’t? Can MLB’s best prospects always be selected in the draft, or do they exist elsewhere, such as in independent leagues?

Teams often evaluate player success in the pre-farm system leagues, from NCAA to independent leagues and so on. Due to the many contexts and backgrounds from which players can be selected, MLB teams must understand the nuances and differences of these leagues in order to properly assess, and select, talent from within them. MLB has a uniquely difficult problem of identifying talent as there is a large farm system comprising multiple collegiate leagues as well as independent leagues such as the Frontier League. This means there can be even more variance between prospects that makes evaluation tougher. In this paper, we highlight that variance and use the differences in success across leagues to illustrate the difficulty that comes with prospect evaluation. To simplify this already gargantuan process, we will focus on pitching, as well as a few hitting metrics, as a lens through which differences in leagues can be gauged.

The Data

In this project we will be analyzing every single Auto Pitch Type on Trackman for the NCAA D1 Baseball 2023 season, our main data set. These metrics include, but are not limited to, pitch velocity, horizontal break, induced vertical break, spin rate, spin axis, and exit velocity. A lot of our analysis will be on this data set, where we will analyze any trends across the season and the correlation between pitching metrics. Our second data set is every single Auto Pitch Type on Yakkertech for the 2023 Frontier League Baseball season.

While the Trackman and Yakkertech files are very similar, they are not the same. They use different hardware to track speeds and locations of game events. Additionally, the data analyzed is on two very different groups of players: college athletes and professional athletes. Despite these limitations, the hardware for each company is professional-grade. We have subsetting the data from Trackman (NCAA) and Yakkertech (Frontier League) already, getting rid of miscellaneous Auto Pitch Types, as well as Knuckleballs.

The Rice Baseball Analytics team has given us access to the Trackman data set, while Nathaniel Ascher has gotten access to the Yakkertech data set through his internship this past summer. Because of confidentiality issues, we did not do any analysis on individual players, teams, or stadiums.

Random Forest

While both the Trackman and Yakkertech datasets have values for Tagged Pitch Type, Yakkertech does not have a column for Auto Pitch Type. This is a problem because Tagged Pitch Type, the other pitch type column consisting of pitches individually tagged by a human, is filled with inconsistencies and NA values, which may bias our analysis. Since the Tagged Pitch Types on Yakkertech are often incorrect, we thought it would be best to somehow use Trackman's Auto Pitch Type formula for Yakkertech and thus compare the data sets through Auto Pitch Type.

To accomplish this, we ran a random forest model on Trackman with the Auto Pitch Type as our target variable with features release speed, spin rate, spin axis, induced vertical break, horizontal break, and the pitcher's handedness. The random forest algorithm creates a series of decision trees, slicing the features into subsets, learning the decision trees of various shapes and returning a predicted Auto Pitch Type. The decision is found at the bottom of the tree, also known as the leaf node. The model then takes the decision from each leaf node and finds the majority vote predicted Auto Pitch Type until it has found the model that maximizes the information gained. The model achieves a 96.6% accuracy on the test set. Then, using the decision boundaries generated by the random forest model, we used the same features on Yakkertech to create a new Auto Pitch Type column. This way the Auto Pitch Types are directly comparable to Trackman's Auto Pitch Type column because they are generated in the exact same way, and therefore we can compare pitch metrics across leagues.

Pitch Metric Definitions

Before getting into the pitch metrics across the Frontier League and the NCAA, we will give definitions for the pitch metrics themselves: velocity, spin rate, horizontal break, and induced vertical break.

Velocity

Pitch velocity, or RelSpeed in Trackman and Yakkertech, refers to the speed of the ball as it leaves a pitcher's hand, in mph (Trackman glossary). This is simply how hard a pitch is thrown, and has been tracked ever since the invention of radar guns.

Spin Rate

Spin rate is "the speed at which the ball is spinning, reported in revolutions per minute (RPM)" (Trackman glossary). The spin rate of each pitch type can vary a bit throughout a game or an entire season, but it is a good metric to see the quality of each pitch and to differentiate between pitches that are similar in other metrics, such as fastballs and changeups.

Horizontal Break

Horizontal break (HB) is what Trackman and Yakkertech call HorzBreak and is a measure of horizontal distance in inches. Precisely, it is the difference in the horizontal distance from where the ball crossed home plate and where it would have crossed the plate without gravity and traveling in a perfectly straight line from a pitcher's release point (Trackman glossary). It is important to note that when you go from a right-handed to a left-handed pitcher, horizontal break gets negated.

Induced Vertical Break

Induced vertical break (IVB) is what Trackman and Yakkertech call InducedVertBreak. It is the difference in the height that the ball crosses home plate and the height that the ball would cross home plate if it was thrown at the same velocity, but with no spin (Trackman glossary). While HB gets negated when you are switching the handedness of a pitcher, IVB does not.

Pitch Type Definitions

After defining the pitch metrics, it is important to discuss the metric differences within the pitch types. They can be classified into three main categories: fastballs/sinkers/cutters, offspeed pitches, and breaking balls.

Fastballs, Sinkers, and Cutters

Fastballs

In this analysis, fastballs refer to four-seam fastballs. This pitch will be the fastest (highest RelSpeed) in a pitcher's arsenal, with the highest amount of positive IVB (Rapsodo). Fastballs that have a higher spin rate are typically more effective than those with lower spin rates, although the two most important metrics for fastballs are velocity and the absolute value of the difference between the absolute value of HB and IVB. We will get into both of these topics, spin rate and HB and IVB difference, later in this analysis. For right-handed pitchers, fastballs have positive HB, while for left-handed pitchers fastballs have negative HB. Fastballs will almost always have more IVB than the absolute value of their HB. Fastballs, especially those with a lot of rise (IVB), are great at inducing strikeouts, but also lead to a large number of fly balls as well.

Sinkers

Sinkers are another velocity-oriented pitch, using high speed to create difficulty for hitters. Sinkers are almost identical to fastballs, except that the absolute value of their HB will be greater than their IVB. Because sinkers have a lot of horizontal movement, they are great at inducing ground balls.

Cutters

Cutters are typically thrown softer than fastballs and sinkers, and with slightly more spin. They always have positive IVB, while their HB hovers around zero. For right-handed pitchers, the HB is either around or slightly below zero, and for left-handed pitchers it is either around or slightly above zero. As we will go more in-depth soon, cutters are sliders, but with more IVB, less HB, and are thrown harder.

Offspeed

Changeups

Changeups are offspeed fastballs and sinkers. They are thrown much softer and almost always with less spin than fastballs and sinkers. However, the HB and IVB characteristics of fastballs/sinkers and changeups are very similar. This means that pitchers who throw fastballs/sinkers and changeups utilize this congruence to induce swings and misses on changeups and weak contact, as long as they are not thrown up in the zone.

Splitters

Splitters are identical to changeups, except that they spin significantly less. Splitters are quite rare.

Breaking Balls

Sliders

Sliders are pitches that can significantly vary in velocity. They will never be faster than a pitcher's fastball, sinker, or cutter, but they could be faster than their changeup or splitter. Sliders spin a lot, have IVB around zero, and have significant glove-side movement. For most right-handed pitchers, sliders move towards the first base side, away from right-handed hitters, and have negative HB. For most left-handed pitchers, sliders move towards the third base side, away from left-handed hitters, and have positive HB. Compared to cutters, sliders have more HB. For left-handed pitchers, cutters move less toward third base, and for right-handed pitchers less toward first base. Since sliders have a lot of HB, they can be hard to throw for strikes. Pitchers generally throw sliders to induce swings and misses and generate ground balls.

Curveballs

Curveballs are typically the slowest pitch a pitcher will throw. They are similar to sliders, except that their IVB is much more negative and they have more overall total movement (abs. value of HB + IVB). Similar to fastballs, the more spin a curveball has the higher the quality of a pitch it is.

3D Plot: Random Forest Auto Pitch Type Visualization

Now we will visualize how the above pitch types are distributed across HB and IVB.



Figure 1: Legend for 3D Plots

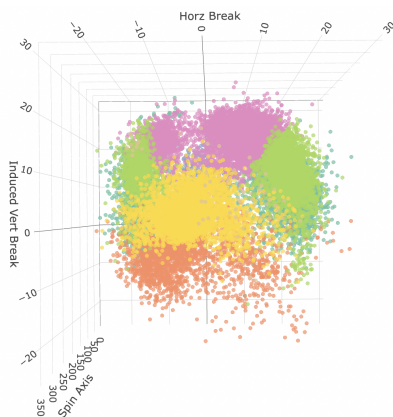


Figure 2: NCAA 3D Plot

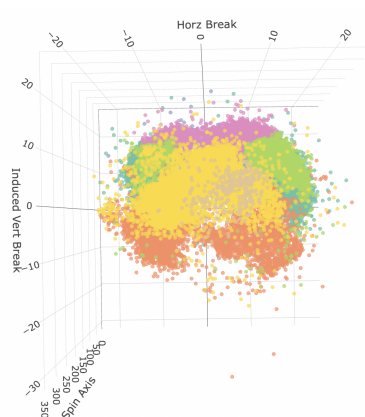


Figure 3: Frontier League 3D Plot

Looking at these 3D plots, which demonstrate the clusters of pitch types against each other across HB and IVB, we can see that the pitch types across leagues are similarly distributed. Because our random forest model has accurately clustered pitch types, we can now compare pitch metrics across leagues.

Pitch Type by Inning

One way to analyze the differences between pitching in the NCAA and the Frontier League is to compare the distribution of pitch types by inning across the two leagues. Through this, we can begin to understand the differences in how pitchers attack batters between the two leagues.

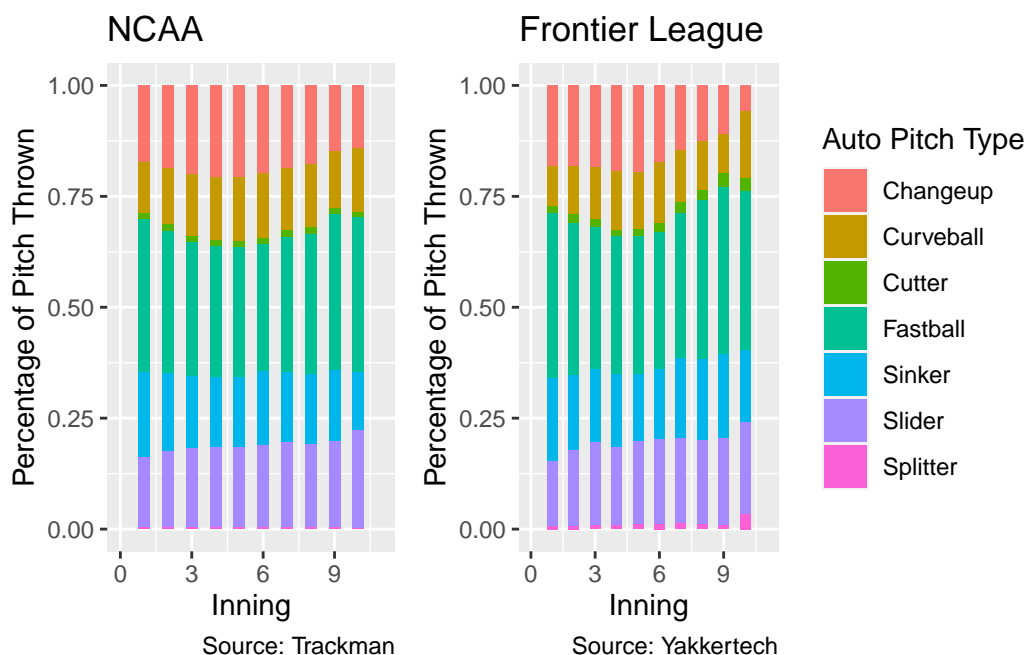


Figure 4: Pitch Type by Inning

Both leagues do seem to agree on fastball/sinker volume, with fastballs/sinkers constituting around 50% of the mix in each inning. In the NCAA, fastball usage decreases from the first to the sixth inning, reaching a low of 28.55%, before increasing again. Fastballs in the Frontier League follow a similar pattern; however, for each inning, the amount of fastballs in the Frontier League is around 2% higher. Furthermore, in the Frontier League, it appears that fastballs/sinkers are used more in the 8th and 9th innings than in the NCAA, especially sinkers. Looking at changeups, they are used a similar amount in both leagues for the first six innings. However, from the 7th to the 10th inning, NCAA pitchers throw significantly more changeups than Frontier League pitchers. For splitters, it seems as if they are used more in the Frontier League than in the NCAA, especially in the latter half of games.

Moving on to breaking balls, the usage of sliders and curveballs appears roughly the same, although it seems as if curveballs are thrown slightly more often in the NCAA than in the Frontier League, especially at the beginning and end of games. Overall, the pitch distribution across both leagues is pretty similar.

The one major difference between the NCAA and the Frontier League is that pitchers in the Frontier League use their fastballs and sinkers more in the 7th, 8th, and 9th innings of games. NCAA pitchers seem to trust their breaking balls and offspeed pitches late in games, while Frontier League pitchers go back to their most comfortable pitches, fastballs and sinkers, in the late innings. We think that this represents a slight skill difference between the two leagues. Fastballs and sinkers are the pitches that are most likely to get hit hard. They are also the most predictable pitches. Therefore, NCAA pitchers keep hitters more off balance at the end of games, most likely leading to better results.

Release Speed by Inning

In order to understand the relationship between pitch type, release speed, and inning, it is important to know the types of pitchers. In a baseball game, pitchers will typically work in a rotation. There are three kinds of pitchers in a rotation, known as ‘starters’, ‘relievers’, and ‘closers’. Starting pitchers aim to pitch through as many innings as possible before their pitch count becomes too high. This can range from 75-100 pitches, and they pitch from the start of the game until they are taken out (usually anywhere between the 4th and 7th innings). Starters are followed by relievers, who are intended to work after the starting pitcher through the end of the game. Relievers can last one to three innings and sometimes fail to get to the end of the game before being removed. In this case, managers will insert either another reliever or a closer. Closers pitch typically one or two innings, and look to end the game without giving up any runs or hits. Closers tend to pitch in the highest leverage innings, which is usually the 9th inning when they are trying to end the game.

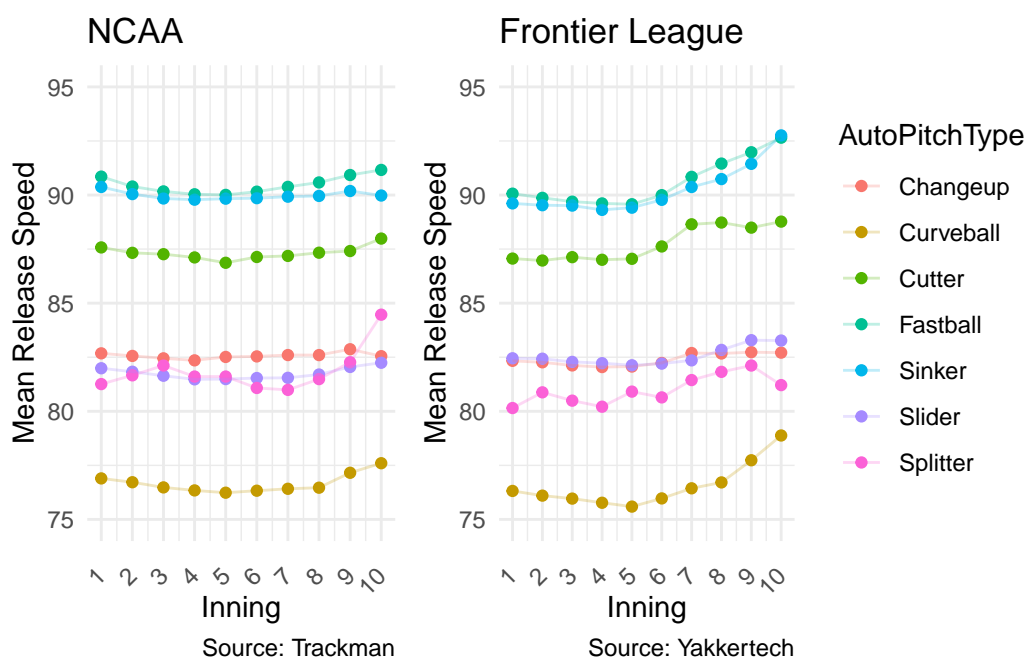


Figure 5: Average Pitch Velocity per Inning by Pitch Type

From our graphs, we can see that for the most part, pitching is similar across leagues. Pitch types start at very similar speeds, and across the first 5 innings, they are nearly identical. The Frontier League begins to break away around the 6th and 7th inning, illuminating the time at which fresh arms are inserted. Therefore it appears that Frontier League starters do not last as long in the game as NCAA starters do. In the Frontier League, curveballs, cutters, sinkers, and fastballs all jump in average release speed, indicating more difficult pitches for batters to hit. Calling back to our pitch types per inning chart, we know that these four pitch types constitute over 60% of pitches in each inning. This disparity in high-volume pitches is highly relevant to batters and shows that relievers and closers in the Frontier League have a larger talent gap over their starting counterparts when compared to NCAA pitching groups.

Spin Rate by Pitch Type

An example of a pitching metric that impacts performance is spin rate. While spin rate is just one metric of a pitcher's skill set and within their control, it can be used in different ways on different pitches. It can generate vertical or horizontal break, deceiving batters by moving the ball into or out of the strike zone. When pitchers look to throw a certain type of pitch, they are using spin, among other things, to reach the desired outcome. For each pitch type, you would expect a similar distribution of spin rates between the two leagues. As we see in the figures below, this is pretty much the case.

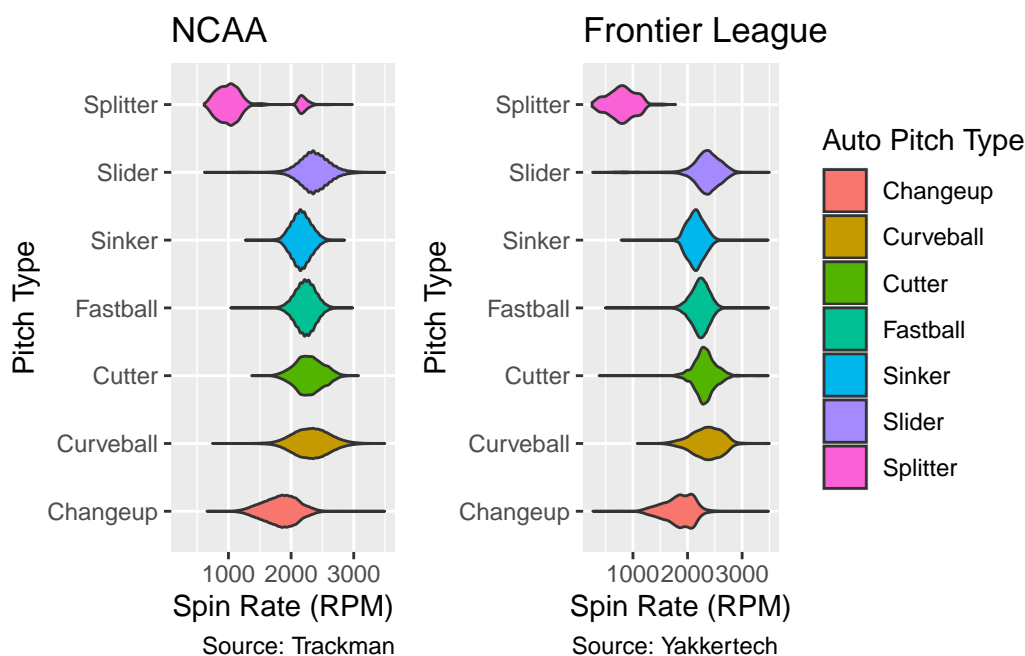


Figure 6: Violin plots of Spin Rates (RPM) by Pitch Type

Pitches such as the fastball, sinker, slider, cutter, curveball, and changeup have comparable distributions around the same spin rate, indicating that they are thrown similarly across both leagues.

While the distribution of spin rate for each Auto Pitch Type is pretty much the same, there are differences in the distributions of spin rates between the NCAA and the Frontier League. The largest difference is the range of spin rates. The Frontier League has a significantly wider range of spin rates for each Auto Pitch Type, with more pitches in the 0-500 rpm range and 3000-3500 rpm range. This is most likely due to errors in Yakkertech, as pitches will rarely be in this range. Trackman is a much more accurate system of pitch tracking, shown by the fact that the NCAA and MLB use it, compared to the Frontier League, where money is much harder to come by and expensive software more difficult to obtain.

Another large difference can be seen for splitters. Where the NCAA has the majority of its splitters at 1000 RPMs, the Frontier League sees a dip at that same value. The Frontier League's large range in splitters indicates a potential lack of control by pitchers, as a higher spin rate splitter will not be as effective.

Why is this the case? How do spin, and other factors, such as HB, IVB, and velocity affect pitch effectiveness and results? Next we are going to discuss the importance of HB and IVB for "dead zone" fastballs/sinkers and the tangible difference between the NCAA and the Frontier League within this pitching metrics.

Dead Zone Fastballs and Sinkers

A “dead zone” fastball/sinker is a pitch in which the difference between the absolute value of the HB and the IVB is less than or equal to 2 (we chose this mark, but essentially a “dead zone” pitch is one in which its IVB and the absolute value of the HB are close together). We take the absolute value of the HB because lefties have negative HB on their fastballs/sinkers, as explained earlier in this analysis. With dead zone pitches, the idea is that they should be easier to hit than fastballs and sinkers which have a larger difference between their HB and IVB. This is because the dead zone is in line with hitters’ most common swing planes, and thus the barrel of their bats (Medium). Separating each league into pitch velocity bins, we wanted to see if there were differences between the leagues in the average exit speed for these pitches.

The two graphs depict the average exit speed by velocity bin for “dead zone” fastballs and sinkers in the Frontier League and NCAA.

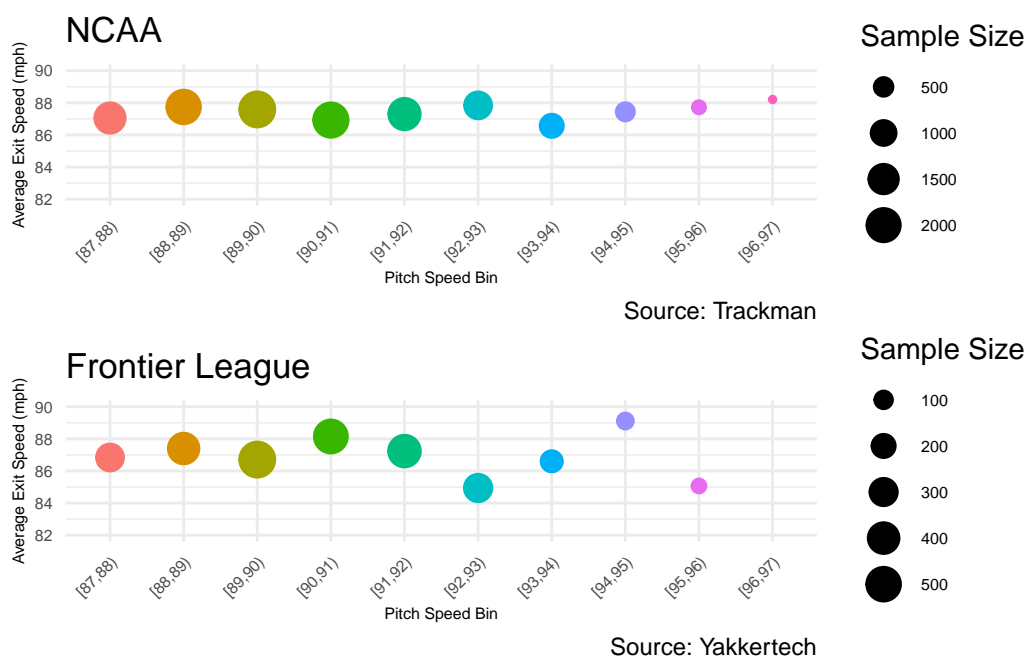


Figure 7: Avg Exit Speed by Pitch Speed Bin of Deadzone Fastballs and Sinkers

Looking at the two graphs, there is evidence for the claim that NCAA hitters hit dead zone fastballs and sinkers better. NCAA hitters have a higher average exit speed for each speed bin that starts with 87, 88, 89, 91, 92, 95, and 96 mph. The only speed bins in which Frontier League hitters had a higher average exit speed were dead zone fastballs and sinkers between 90-91 mph, 93-94 mph, and 94-95 mph. Overall, for 7/10 of these speed bins of all dead zone fastballs and sinkers, NCAA hitters are better at hitting them than Frontier League hitters. This does not take into consideration any other metrics or data about these pitches, including, but not limited to, spin rate, location, spin axis, and stadium altitude, which all certainly have an effect on the quality of contact a hitter can make with a pitch.

What further explains why NCAA hitters are better at hitting dead zone fastballs and sinkers than Frontier League hitters, as well as maybe being better hitters overall, is the difference in the average exit speed for the 95-96 and 96-97 speed bins. NCAA hitters hit the dead zone fastballs and sinkers in the former bin 2.6 mph harder on average, and 7.6 mph harder on average for the latter bin. These two differences are the largest between all of the bins, besides the 92-93 mph bin, and represent the dead zone fastball and sinker velocity bins that these hitters would face in the highest level of competition, MLB.

One interesting trend to note is that the sample sizes for both leagues follow the same pattern. The sample size increases until the 89-90 speed bin, then decreases after that. This is very telling, as in both leagues the same phenomenon occurs for dead zone fastballs and sinkers. This corroborates our findings from earlier about release speed by inning and spin rate. NCAA and the Frontier League are very similar in terms of pitch mixes and pitches overall. It now appears that they are similar as well in terms of batted ball data.

This analysis on dead zone fastballs and sinkers reveals a major difference between Frontier League hitters and NCAA hitters for the 2023 season. NCAA hitters made better contact on dead zone fastballs and sinkers. But does this mean that they are better hitters? Because dead zone pitches are, in theory, the easiest types of fastballs and sinkers to hit, this may not necessarily be the case. Next we dive much deeper into investigating this question, this time not just looking at dead zone fastballs and sinkers that were put in play, but all fastballs and sinkers for both leagues, as well as all other pitch types.

Situational Pitch Type Run Values and Exit Speeds: Killer Plot

Introduction

Our killer plot visualizes the differences in pitching outcomes across leagues by placing markers for the expected run value of each pitch across leagues at their average, standard deviation above average, and two standard deviations above average on a baseball diamond. By placing each along the base path, we create an effective visualization of what occurs in-game in each scenario across leagues because a higher number implies the base runner is expected to get farther along the base path. The average run value is calculated as the sum of the run values in each scenario, averaged over the number of times the scenario occurred. A single is given a value of 1, double 2, triple 3, and home run 4. Outs, errors, sacrifices, and fielder's choices are given values of 0.

However, expected run values do not tell the whole story—the exit speed coming off the bat of each scenario can also be an indicator of how well the pitcher has performed. Better pitchers can prevent high exit speeds by inducing weak contact. To visualize this, we plotted the exit speeds along the same axis stacked vertically with the average, one standard deviation above and below, as well as two standard deviations above and below to visualize the spread of expected exit speeds in each scenario.

To determine differences between the leagues, we want to compare the outcomes of pitches across various scenarios.

Specific Combinations

Killer plot

Choose Ball

Count:

Any

Choose Strike

Count:

Any

Choose

Pitcher

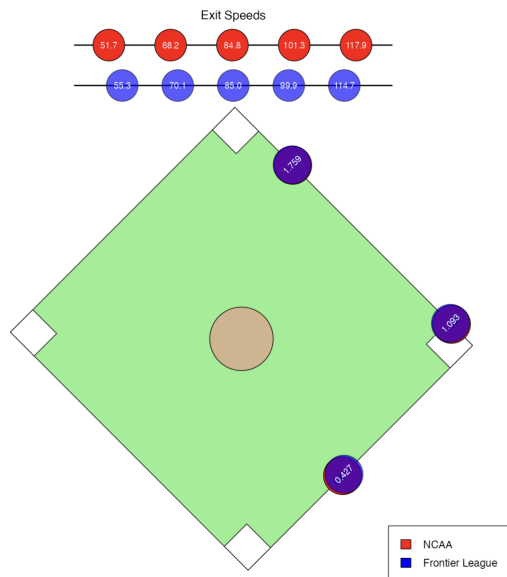
Handedness:

Any

Choose Pitch

Type:

Any



% of total
balls in
play
(NCAA):
100%

% of total
balls in
play
(Frontier
League):
100%

Figure 8: Killer Plot Combination #1

Above is our killer plot with run values at the default values. The markers for each league overlap because the run values, averaged over all at-bats, converge over time.

Killer plot

Choose Ball

Count:

3

Choose Strike

Count:

1

Choose

Pitcher

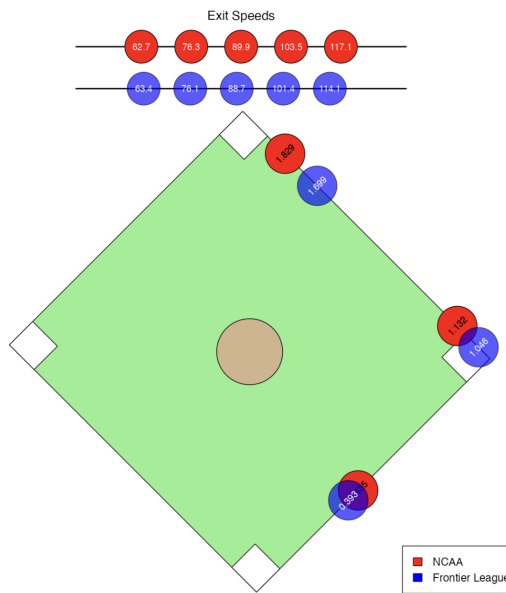
Handedness:

Any

Choose Pitch

Type:

Fastball



% of total
balls in
play
(NCAA):
1.5%

% of total
balls in
play
(Frontier
League):
1.3%

Figure 9: Killer Plot Combination #2

Looking specifically at the results for 3-1 fastballs in play, the advantage here definitely goes to the NCAA. Their average run value is higher, their average exit speed is more than one mph higher, and their standard deviation values for run value and exit speed are all higher, except two standard deviations below the mean for exit speed. Considering that the percent of total balls in play is above one percent in both leagues, we can determine that this is a significant difference between the NCAA and Frontier League. While some game states such as this lean towards the NCAA, note that the general trend may lie in the other direction.

Table 1: Count of Situations Where Average Run Value and Exit Speed are Higher

Pitch Type	NCAA better	Frontier League better
Any	13	28
Changeup	3	25
Curveball	7	16
Cutter	3	0
Fastball	25	9
Sinker	8	24
Slider	7	14

Summary of Results

When looking at all combinations of game states, adjusted for sample size (game state occurring over 0.5% of the time in each league), the Frontier League leads the NCAA in most game states with both higher average run value and higher average exit velocity. While we don't delve into the exact sample size of each game state in the table, we can determine that the Frontier League's hitters may be more effective than the NCAA's in many frequent game states. This further supports our takeaway that the Frontier League is more comparable to the NCAA than many scouts may think.

Run Value and Release Speed Regression

In order to determine if the metrics we've been looking at actually have predictive power on outcomes, we ran a regression analysis of run value on release speed to determine the impact increasing a pitcher's release speed has on the outcome of each individual pitch. We separated the regression across pitch types to get a further detailed idea of how release speed impacts run value. Looking at the results, across the board a single mph increase in release speed leads to about a 0.005 increase in run value. For curveballs, this value increases to 0.006 which indicates greater returns from increasing release speed on curveballs. Each is significant at a 99% significance level.

Table 2: NCAA Pitch Speed and Run value

	<i>Dependent variable:</i>				
	Fastball	Curveball	Slider	Changeup	Sinker
	(1)	(2)	(3)	(4)	(5)
RelSpeed	0.005*** (0.00003)	0.006*** (0.0001)	0.005*** (0.00005)	0.005*** (0.00004)	0.005*** (0.00004)
Observations	52,689	19,244	28,083	35,545	33,207
R ²	0.279	0.289	0.274	0.284	0.297
Adjusted R ²	0.279	0.289	0.274	0.284	0.297

Note:

*p<0.1; **p<0.05; ***p<0.01

Interestingly enough, the regressions in the Frontier league, shown below, produced the exact same results as the prior regressions. This indicates that the impact of release speed on run values is independent of the league one plays in, providing support for the claim that the two leagues are more similar than most baseball analysts think.

Table 3: Frontier League Pitch speed and Run value

	<i>Dependent variable:</i>				
	Fastball	Curveball	RunValue Slider	Changeup	Sinker
	(1)	(2)	(3)	(4)	(5)
RelSpeed	0.005*** (0.0001)	0.006*** (0.0002)	0.005*** (0.0001)	0.005*** (0.0001)	0.005*** (0.0001)
Observations	7,266	2,095	3,759	3,888	3,975
R ²	0.285	0.309	0.277	0.302	0.298
Adjusted R ²	0.285	0.309	0.276	0.302	0.297

Note:

*p<0.1; **p<0.05; ***p<0.01

Conclusion

Overall Takeaways

Fundamentally, finding statistical differences across leagues by only looking at the raw data of pitch-by-pitch data is no easy task. Typically, measures of pitcher value are based on runs allowed which depend on the skill of the hitter. This makes isolating the contribution of the pitcher more difficult. Despite this, we have some evidence that the leagues are very similar in terms of pitch metrics and exit speeds of hits. NCAA hitters are better at hitting fastballs, cutters, and MLB-level dead zone pitches than Frontier League hitters. Additionally, Frontier League relievers/closers seem to throw more fastballs/sinkers and throw harder than Frontier League starters.

In terms of pitching metrics, release speed, spin rate, HB, and IVB, the leagues are very similar. The most significant differences within pitching are the greater difference in quality of members of pitching groups in the Frontier League and the higher usage of fastballs/sinkers at the end of Frontier League games.

On the hitting side, the leagues are also quite similar, with much evidence that NCAA hitters are better against fastballs (both dead zone and not). But for every other pitch type, Frontier League hitters hold their own.

While it appears that the leagues are quite similar, many MLB scouts look for quality hitting (exit velocity and extra base hits) on fastballs. Thus, while the cream of the crop of hitters seem to be in the NCAA, scouts should look into Frontier League relievers and closers more than they currently do.

Future Analysis Opportunities

One avenue of possible future analysis is measuring the accuracy of each tracking system. As mentioned earlier in our discussion of the random forest model, one way we saw to get around this issue was by using a random forest model trained on the Auto Pitch Type column from the NCAA to predict the Pitch Type in the Frontier League (as the Frontier League lacked this feature). Both the NCAA and the Frontier League had a manual Tagged Pitch Type feature that we opted not to use as some of the pitches tagged by humans seemed inaccurate given the details of the pitch (eg. HB and IVB). Therefore, future iterations and steps for this research could focus on seeing the differences and similarities in results between using automatic pitch tagging in both leagues as compared to manual pitch tagging. Additionally, it would be useful to see if it is possible to remove the inaccurate human pitch tagging entries and compare if the human pitch tagging was more accurate, the same, or worse than the automatic pitch tagging.

Another possible avenue for future analysis exploration is regarding the datasets used. In our analysis, we compared data from the Frontier League and the NCAA. Future research can utilize the MLB, other independent leagues (eg. Atlantic League), or international baseball leagues (eg. Nippon Professional Baseball, which is the highest level of baseball league in Japan).

References

Data:

Every single tagged pitch on Trackman for the NCAA D1 2023 baseball season, courtesy of Rice Baseball Analytics internal data, 2023.

Every single tagged pitch on Yakkertech for the 2023 Frontier League baseball season, courtesy of Nathaniel Ascher's summer internship.

Other:

Trackman glossary: <https://trackmanbaseball.zendesk.com/hc/en-us/articles/5089413493787-V3-FAQs-Radar-Measurement-Glossary-Of-Terms>

Rapsodo: <https://rapsodo.com/blogs/baseball/understanding-rapsodo-pitching-data-break-profile-fastball/>

Medium: <https://medium.com/iowabaseballmanagers/dealing-with-dead-zone-how-to-optimize-fastball-performance-d186e2bb07ef>