Groundballs, Flyballs and Line Drives

by Dave Studeman

May 9, 2004

Our friends at **Baseball Info Solutions** track every single batted ball in every game. If a batted ball results in an official plate appearance (that is, a hit or an out), they place the batted ball in one of three buckets: groundballs, flyballs and line drives. Actually, they also track bunted balls, but we're not too concerned with those — the first three types make up about 99% of all batted balls. So let's talk about those.

Here's a little table, based on games played through May 4th, of the number of each type of batted ball, and what percent of each type was turned into an out:

	NUMBER	OUTS	OUT%
Groundballs	9952	7208	72%
Flyballs	7845	6179	79%
Line Drives	4082	1050	26%
Total	21879	14437	66%

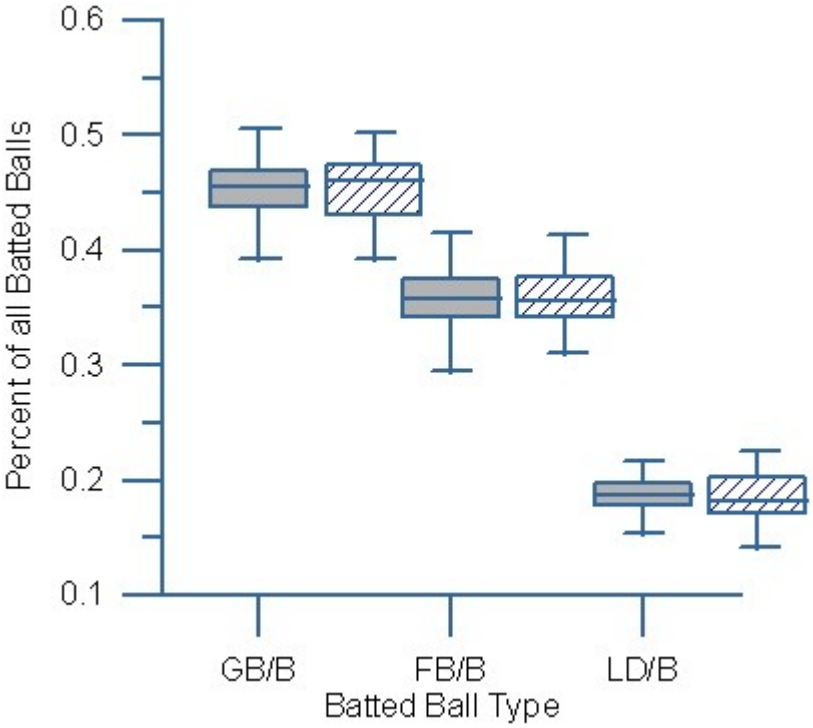
45% of batted balls were groundballs, 36% were flyballs and 19% were line drives. But line drives were much more likely to become hits than groundballs and flyballs, while flyballs were most likely to be turned into an out. Overall, 66% of all batted balls were turned into outs by the fielders.

Actually, this data is a little bit misleading because it includes home runs (which are mostly flyballs). There were about 880 home runs hit during this period. When you take them out of the data, you find that 69% of all fieldable balls were converted into outs by fielders — which is consistent with the DER (Defense Efficiency Ratio) that is listed in our **stats section**. Data consistency always makes me feel good.

There's gold in them thar data — so much that I hardly know where to start. For

instance, your first inclination might be to say that flyball pitchers are better than groundball pitchers (more outs per batted ball) but then you'd realize that flyballs are most likely to turn into home runs. We need to dive a little deeper into the data.

For this article, I'm going to take a closer look at team data only and leave player data for later. To start, here is a graph of batted ball type distribution by each team's pitching staffs (the solid gray box) and each team's batters (the lined box). All 30 major league teams are included.



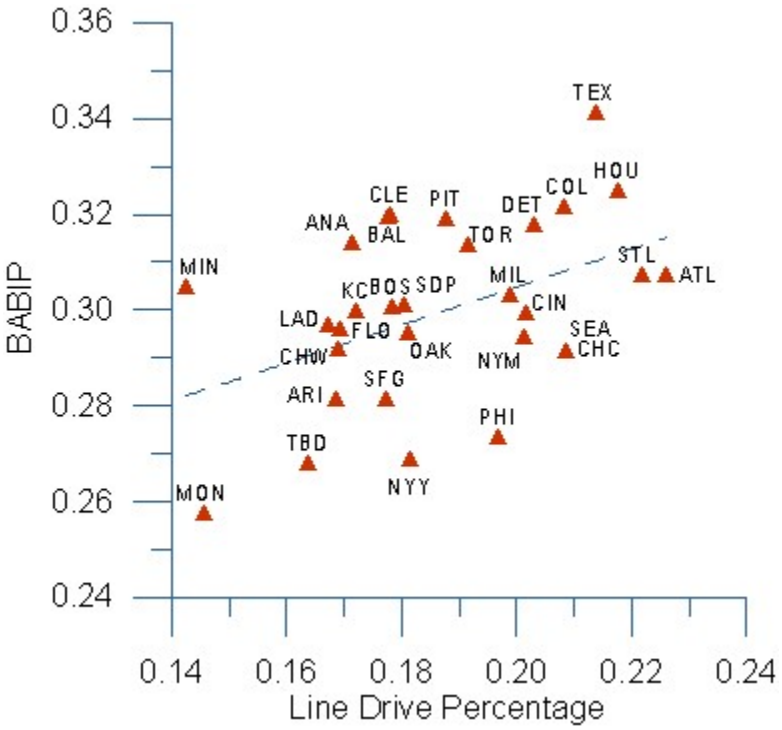
This is called a "box whisker graph." The median of each batted ball type is represented by the horizontal line in the middle of each box. The boxes represent the 25th and 75th percent quartiles, and the extended vertical lines include the rest of the teams.

This graph will be a lot more revealing with individual player data, but even at the team level, you can see one important point: batters, more than pitchers, determine if a batted ball is a line drive. Look at the pitching and batting boxes for line drives, and you'll see that the variance is greater for batters than pitchers. The difference from the top to the bottom is greater, for both the boxes and the extended lines.

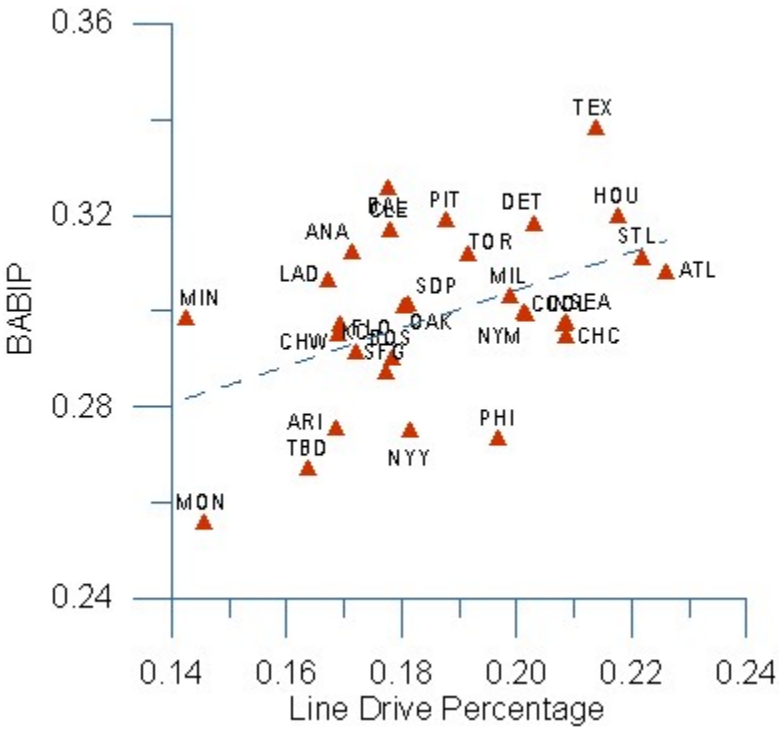
By the way, it may look like the variance for flyballs is wider for pitchers, but that is primarily driven by one outlier (the **Braves**, who have a very low FB%), so we can't call it conclusive. But the implication of the line drive data is pretty clear: major

league pitchers don't vary a lot in their ability to cause or prevent line drives, but major league batters do vary in their ability to hit them.

As you can imagine, batters who hit line drives are more likely to see them fall in for hits. Here's a graph of each team's LD% (percent of batted balls that are line drives) and its BABIP (Batting Average on Balls in Play).



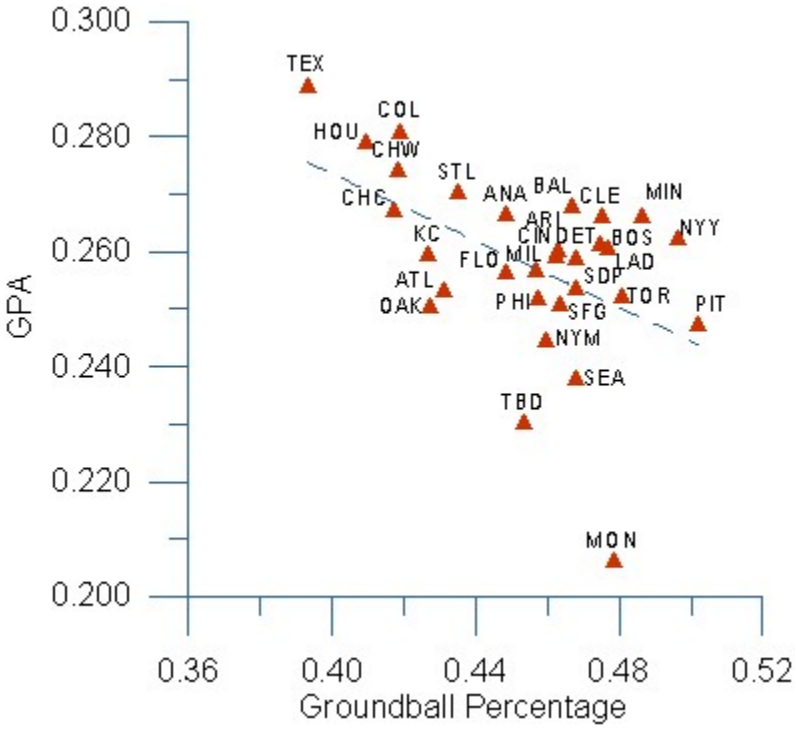
You can see that BABIP generally increases as line drives increase. Of course the correlation isn't perfect (R squared of .20), but the trend is unmistakable. Certain ballparks can also have a big impact on BABIP, so let's adjust this data by ballpark, **using some figures provided by the generous Tangotiger**.



It may be hard to see, but this adjustment helped the "fit" a little (R squared of .22). In fact, it's probably safe to say that the **Rangers** and **Orioles** have been lucky so far (BABIP is higher than LD% would predict), and the **Yankees**, **Expos**, **Devil Rays** and **Phillies** have been unlucky (vice versa). Time will tell; we'll keep an eye on this graph as the season progresses.

By the way, the ability to hit line drives is also correlated with overall batting effectiveness (GPA) and Isolated Power (ISO, which is SLG minus BA) but not home runs. Home runs are more closely tied to hitting flyballs.

What about groundballs, you ask? Don't hit them, unless you're **Juan Pierre** or **Scott Podsednik**. Here is a graph of groundball frequency against GPA:



In general, the more that teams hit the ball on the ground, the less runs they score. The R squared for this data is .25.

There are a lot of variables that need to be considered when analyzing offense. These graphs have only looked at batted balls, but obviously the ability to hit a ball in the first place helps, too. So does the ability to draw a walk. So to try and pull things together, I ran a multiple regression model based on a number of factors.

More from The Hardball Times

A Hardball Times Update
by RJ McDaniel
Goodbye for now.

Here's what I did: I regressed five different rates per plate appearance — strikeouts, walks, line drives, groundballs and flyballs — against the number of runs that each team scored (per plate appearance), to see which factor had the most impact.

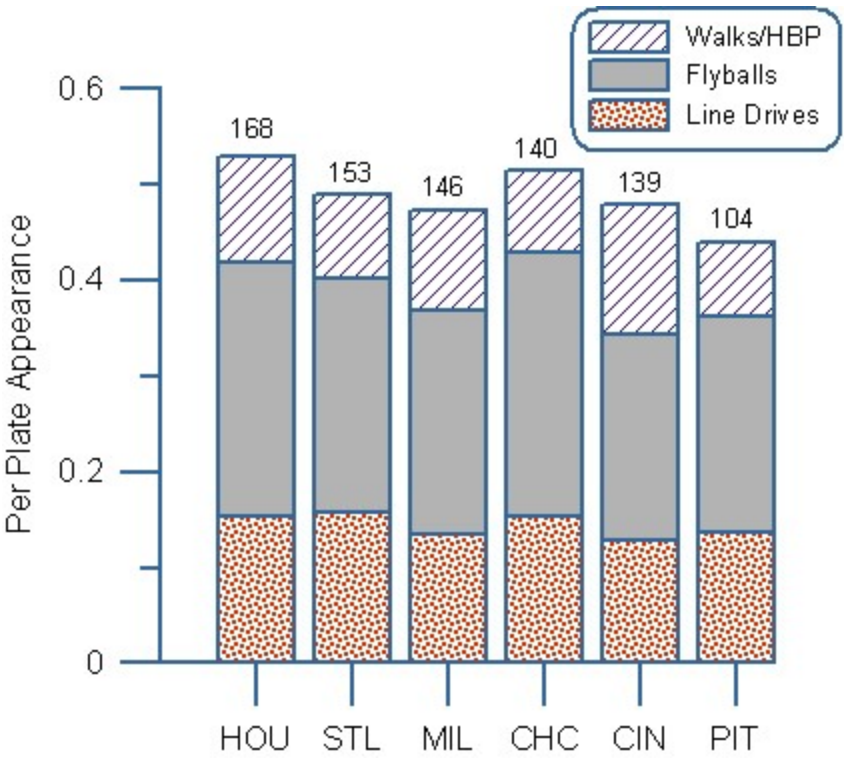
In a multiple regression, the best measure of impact is something called the "t-stat" (if I remember my statistics class correctly). Here's a list of each of the five factors, along with its t-stat:

– Line Drives: 3.29
– Flyballs: 2.89
– Walks/HBP: 1.79
– Groundballs: -0.99
– Strikeouts: -1.76

The R Squared of this model is .39. I won't go into the technical interpretation of these results, other than to say that line drives and flyballs are the most important factors in the ability to generate runs. Taking a walk is good, too. Strikeouts and groundballs are not good for scoring runs, though strikeouts are worse. Maybe you knew that already…

Okay, this is a very general model — a rough cut at the data. If you have suggestions for improving the model, please let me know.

But before you go, let me try one more graph — let's look at the offense of the six teams in the very competitive National League Central. I'll call this the "Three Good Outcomes" graph, because it includes the rate at which each team hit line drives, flyballs or drew a walk (or HBP). The number at the top is the total number of runs each team scored:



Not a perfect match, but not bad either. Remember, there are a lot of things that impact run scoring. For instance, the **Astros** play in a batter's park, and they've

batted .322 with runners in scoring position. The **Cubs**, on the other hand, have batted .232 with RISP. The **Brewers** don't look very impressive, but their offense features one of the few guys who can turn groundballs into a good thing (the aforementioned Mr. Podsednik). The **Reds** don't hit enough line drives (and don't connect often enough) and the **Pirates** could use a few more walks and a lot less groundballs.

This is not a definitive graph, but it's a fun graph, providing a different look at things. Hopefully, you've enjoyed it. I plan to turn my "graphical eye for the line drive and fly" to pitching next, as well as individual players, eventually. I've already started wondering what the "three good outcomes" looks like for **Barry Bonds**.

References & Resources
Update:

I just received a great e-mail from JC, who told me that the coefficient of each regressed factor, rather than its t-stat, is more indicative of the factor's impact. Both are important, but I should have listed the coefficient of each factor for you. So here they are:

Line drives: .464
Walks/HBP: .356
Flyballs: .212
Groundballs: -.069
Strikeouts: -.264

What this means is that walks actually have a bigger impact on scoring than flyballs do. Line drives and flyballs are not as similar in impact as their t-stats would indicate. And that's as technical as I'm going to get.

By the way, I want to express my thanks to a number of baseball writers and commentators who have stimulated my thinking in this area. They include Mike Emeigh, Chris Dial, Vinay Kumar, Bryan Smith, Tangotiger and the boys at Baseball Prospectus.

Dave Studeman was called a "national treasure" by Rob Neyer. Seriously. Follow his sporadic tweets **@dastudes**.

Never Swat an Infield Fly

by Dave Studeman

June 20, 2004

Hopefully, you've noticed that we collect a lot of batted ball information at The Hardball Times. I have to admit that I'm fascinated by the information, and I constantly look up the number of line drives so-and-so has hit, or what's-his-name has given up, just to see what it implies about who's-his-face. Now if only I could remember names ...

We're building on the work of many other people in this area, including the **original, provocative formation of DIPS** by Voros McCracken, and **his toned-down second look**, as well as **Tom Tippett's study**. I think the best synopsis of how to think like a DIP comes from Voros' second article:

> 1. The amount that MLB pitchers differ with regards to allowing hits on balls in the field of play, is much less than had been previously assumed. Good pitchers are good pitchers due to their ability to prevent walks and homers and get strikeouts in some sort of combination of those three.
>
> 2. The differences that do exist between pitchers in this regard are small enough so that if you completely ignore them, you still get a very good picture of the pitcher's overall abilities to prevent runs and contribute to winning baseball games.
>
> 3. That said, the small differences do appear to be statistically significant if generally not very relevant.

How we separate blame and credit between pitching and fielding is one of the most interesting aspects of baseball research these days. The issue is critical to total player evaluation systems, such as **Win Shares**, that strive to assign total credit and blame for a team's performance. Even if pitchers don't vary a lot in their ability to influence hits on balls in play, significant differences do exist between pitchers within any given year. And it is important to understand the cause of these differences if you

want to evaluate the contribution of pitchers and fielders.

Let me give an example. The Expos have been very good at preventing runs this year. Their **FIP is slightly below average, but their DER is the best in the league**. This would normally lead you to believe that their fielding is most responsible for that good run prevention record.

However, Expo pitchers have only given up lines drives at a .157 rate, which is way low — more than twenty points below average. So their pitchers do deserve a lot of credit for preventing those runs. Now, Expo pitchers probably won't keep yielding line drives at such a low rate for the rest of the year, but that's sort of irrelevant for understanding what's happened so far this year. If you were to evaluate the relative contribution of the Expos' pitchers and fielders, you'd want to know this sort of thing.

Baseball statistics companies, such as Stats Inc. and **Baseball Info Solutions**, have been collecting play-by-play data for many years now. Among the bits of data they collect are specific information on each batted ball, including the type of ball that was hit and where it landed. This has enabled baseball analysts, such as **Mike Emeigh** and **Mitchel Lichtman** (also known as MGL) to dive further into the pitching/fielding contribution issue.

It was actually an article from MGL called **DIPS Revisited**, as well as Mike Emeigh's comments on **this thread at Baseball Primer**, that got us thinking about collecting batted ball data for The Hardball Times. So we've been collecting groundball, flyball and line drive data this season, and we took an **in-depth look at the data a month and a half ago**. Some of our findings were:

- 79% of flyballs are fielded for outs, 72% of groundballs are fielded for outs, but only 26% of line drives are fielded for outs.
- Batters differ in their ability to hit line drives — at least more than pitchers differ in their ability to prevent them.
- 90% of home runs are flyballs, and the remaining 10% are line drives.
- Line drives, flyballs and walks help the offense the most, while groundballs and strikeouts are the worst things for an offense.

However, if you read MGL's article, you'll see that there is one other type of batted ball we need to track: infield flies. This was also a key insight of Michael Humphries' cutting-edge fielding study, Defense Regression Analysis (which was published at Baseball Primer during the offseason, though it doesn't seem to be available now). There are two reasons infield flies are important:

1. They are caught for an out 97% of the time.
2. Pitchers actually seem to have different levels of skill of inducing them.

So I'm happy to announce that our friends at Baseball Info Solutions are now sending us the number of infield flies allowed by each pitcher, and we are now reporting that info in our **American League** and **National League** pitching stats.

More from The Hardball Times

A Hardball Times Update
by RJ McDaniel
Goodbye for now.

What does this new data tell us? Well, let me start with a simple table, which lays out the percent of time each batted ball occurs, how often it is converted into an out, and how often it is hit for a home run. This data is based on 2004 data, through games of June 17.

Type	Percent	Out%	HR%
Groundballs	45%	72%	0%
OF Flyballs	30%	75%	12%
Line Drives	19%	26%	2%
IF Flyballs	6%	97%	0%

When you separate infield flies from outfield flies, you get some interesting results. For instance, the out percentage of outfield flies is only somewhat higher than groundballs — it is the infield fly that makes a flyball more desirable from a pitching perspective. Also, a non-caught outfield fly has about an even chance of being a home run or in-park hit.

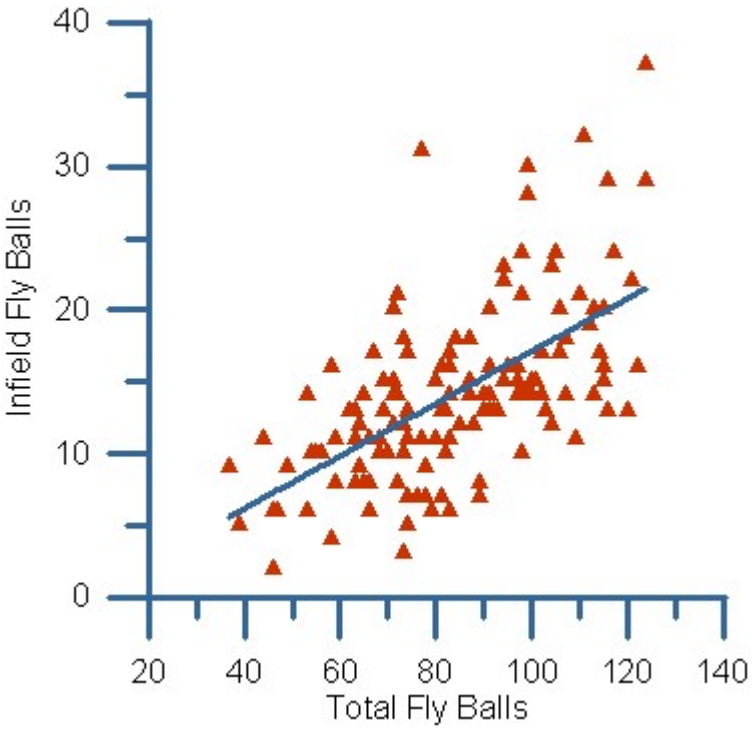
To get a better handle on what these stats mean, I calculated a run impact for each

type of batted ball by assigning a value for each out (-.27), home run (1.4) and hit allowed (educated guesses that vary by type of ball). The source of these values is the second part of Tangotiger's seminal analysis, **How Runs Are Really Created**.

Batted Ball	Run Impact
IF Flyballs	-0.25
Groundballs	-0.06
OF Flyballs	0.06
Line Drives	0.36

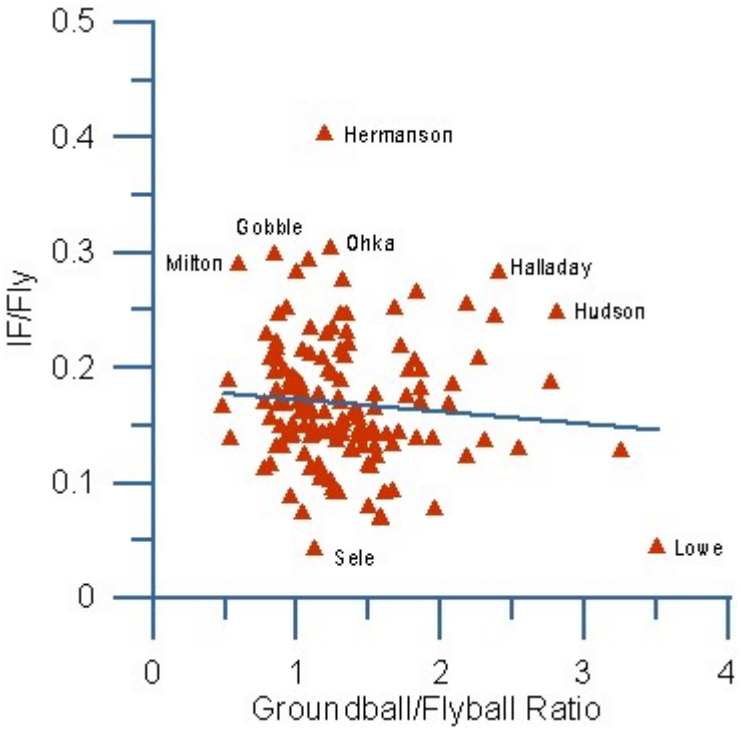
As you can see, hitters want to hit line drives, and pitchers want to induce infield flies. The difference is big.

We've already discussed the control that hitters have over line drives. But which pitchers are most likely to induce infield flies? Here's a graph of all pitchers who have faced at least 200 batters this year (sample size of 141 pitchers), with the total number of flyballs allowed and infield flyballs allowed:



There's no surprise here. Flyball pitchers allow more infield flies. For this reason, MGL reasoned that the most important measure here is the **percent of flyballs that are infield flies**, instead of just total infield flyballs. We label this measure IF/Fly in our data.

So let's change the graph a little bit. I'll put each pitcher's groundball/flyball ratio on the "X" axis (so that groundball pitchers are on the right and flyball pitchers are on the left) and IF/Fly on the "Y" axis. I'll also label a few of the outliers:



Now, the line slopes down slightly, which means that flyball pitchers are slightly more likely to induce infield flies than groundball pitchers. But there are some wild outliers on this graph — especially Dustin Hermanson. Let's take a closer look at their specific stats.

First, let's look at the overall league averages for the stats. Specifically, I'm going to list Runs Allowed Per Game (RA), the percent of batted balls that are line drives (LD%), the ratio of Groundballs/Flyballs (G/F), the proportion of total flyballs that are infield flies (IF/Fly), the percent of batted balls that are converted into outs by fielders (DER), and the three basic FIP rate stats.

	RA	LD%	G/F	IF/Fly	DER	K/9	BB/9	HR/9
American	4.92	.178	1.21	.176	.689	6.3	3.5	1.1
National	4.66	.182	1.30	.176	.698	6.6	3.4	1.1

You probably knew that the American League is the higher scoring league. But the two leagues also differ a bit in their batted balls. The National is more of a groundball league, with a slightly higher LD%. As a result, the league DER is higher (line drives drive DER down, but groundballs drive it up; park factors are also

important). Interestingly, the basic rate stats are virtually even, particularly IF/Fly!

Okay, let's look at the stats of some of our outliers, to see what we can see:

Player	Team	RA	LD%	G/F	IF/Fly	DER	K/9	BB/9
Hermanson D.	SFG	4.62	.123	1.19	.403	.702	6.4	2.6
Ohka T.	MON	3.77	.115	1.23	.303	.709	4.3	2.0
Gobble J.	KC	5.17	.157	.85	.298	.752	2.2	2.3
Milton E.	PHI	4.73	.225	.59	.288	.699	7.2	4.0
Halladay R.	TOR	4.25	.117	2.41	.282	.703	6.6	2.7
Hudson T.	OAK	3.04	.162	2.81	.247	.707	4.3	1.9

Wow. **Dustin Hermanson** is allowing line drives at a very low rate, and his IF/Fly rate is a staggering 40%. Yet he's allowed more than the average number of home runs per inning pitched, and about the average number of runs per game. One problem he's having is that he's not pitching well with runners in scoring position. But maybe his fielders are letting him down a bit, too.

Tomo Ohka is really having a fine season all around, or at least he was before his injury. **Eric Milton** is an extreme flyball pitcher, and he's also giving up line drives at an alarming rate. That won/loss record is very misleading. And look at **Roy Halladay**: all his batted ball stats are great. He's a groundball pitcher who has given up line drives at a very low rate and infield flies at a very high rate. Yet his DER is only .703. Fielders, you think?

Here are some of the outliers from the bottom of the graph.

Player	Team	RA	LD%	G/F	IF/Fly	DER	K/9	BB/9
Sele A.	ANA	3.72	.152	1.12	.041	.717	4.5	3.4
Lowe D.	BOS	6.65	.173	3.50	.043	.677	4.1	4.4
Thomson J.	ATL	5.14	.194	1.49	.079	.661	6.3	2.3
Pavano C.	FLO	2.81	.173	1.23	.101	.766	5.3	2.0
Loaiza E.	CHW	4.18	.165	1.18	.108	.739	5.2	2.7

It's hard to see why **Pavano** has such a high DER. His line drive and GB/FB ratios are about league average and his IF/Fly ratio is low. Either he's got some great fielders, or he's been lucky. Meanwhile, **John Thomson** is at the other end of the spectrum; he seems to deserve that low DER despite being a groundball pitcher.

Aaron Sele is a flyball pitcher who has not been inducing a lot of infield flies. That line drive percentage helps his performance. And it looks like you can put **Esteban Loaiza** in the lucky/good fielders bucket.

Okay, enough with the individual pitchers. Let me try to summarize what I think we know.

- There are definitely groundball and flyball pitchers, and there seem to be some pitchers who are somewhat better at inducing infield flies. They tend to be flyball pitchers.

- Line drives have a big impact on the game, and pitchers may differ quite a bit in the number of line drives they've allowed in a given year. However, most major league pitchers don't seem to have different levels of innate ability to stop line drives from being hit.

- Hitters, on the other hand, do have different line drive hitting abilities.

- If you want to establish credit and blame for allowing or preventing runs, batted ball types can tell you a lot.

Our next step will be to craft a formula for predicting DER and ERA based on batted ball types, park factors and FIP. Y'all come back now.

References & Resources

As always, our thanks to **Baseball Info Solutions** for their data. As a reminder, the definition of a batted ball type depends on the trajectory of the batted ball. Groundballs hit the ground before leaving the infield, line drives come off the bat on a relatively flat trajectory and flyballs follow an upward path off the bat. Infield flies are defined as flyballs that land in the infield (or would have landed in the infield if not caught). If an infielder runs into the outfield to catch a fly, it is still considered an outfield fly, because it would have landed there.

And don't forget that line drives are not necessarily hit harder than flyballs or groundballs. Batted ball type is not a function of how hard the ball is hit.

Dave Studeman was called a "national treasure" by Rob Neyer. Seriously. Follow his sporadic tweets **@dastudes**.

Comments are closed.

Pictures of Batted Balls

by Dave Studeman

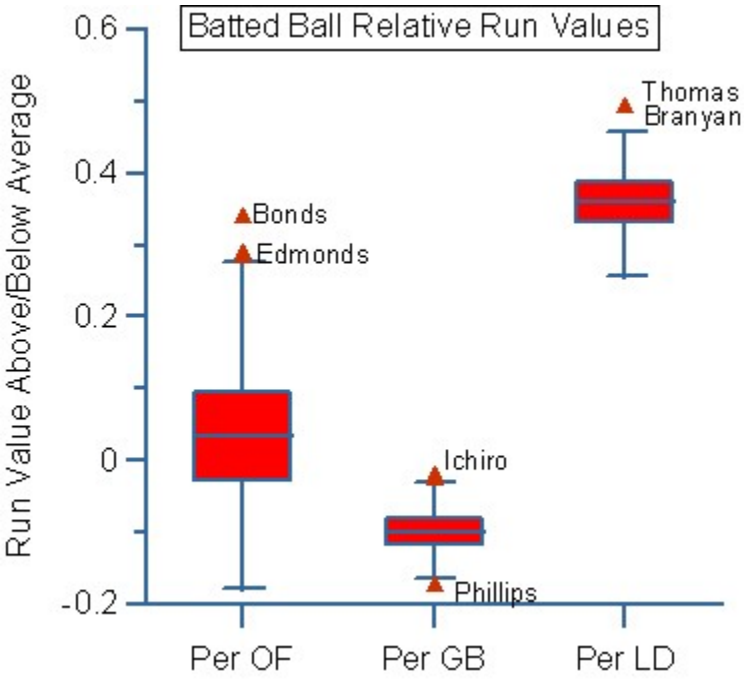
January 5, 2006

I'd like to draw a few pictures of batted balls. If you've read the ***Hardball Times Annual 2006***, you know what I'm talking about. If not, you can read the **Batted Ball Leaderboards article** posted a couple of weeks ago to get the gist.

Here's a nutshell: we applied the relative run values found in **Tom Ruane's excellent Retrosheet article** to the outcome of every batted ball over the last four years. That allowed us to assess the value of every batted ball type, explore team defense and analyze park factors, among other things, in the *Annual*. However, to perform some more definitive player analyses, I had to develop fully regressed batted-ball park factors. I finished that task last week.

So now I can tell you that **Barry Bonds** averaged .34 runs more than the average plate appearance for every outfield fly he hit, adjusted for ballpark. I can also tell you that every outfield fly **Einar Diaz** hit averaged .18 runs less than the average plate appearance, adjusted for ballpark. I can tell you this sort of thing for every major league player.

And now, because I am an insufferable graphing fool, I am going to show you some pictures of the player-specific, park-adjusted, batted-ball data. Ready? First up is something called a "box-whisker" graph. It shows the spread of the average run values for outfield flies, ground balls and line drives, by batter. I included all batters with at least 1,000 plate appearances over the last four years (a total of 300 batters).



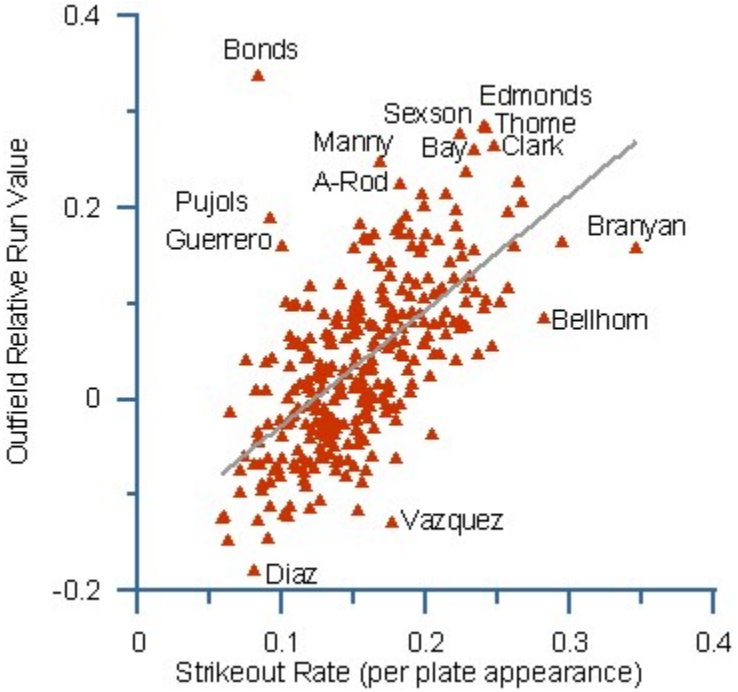
The blue line in the middle of each red box represents the median relative run value of each type of batted ball (for this sample). The red box encompasses the quartiles above and below the median for each type of batted ball (so 50% of the players are contained in the red box), while the lines extending out from the red box encompass the top and bottom quartiles. For each set of data, there are outliers identified by red triangles.

As you can see, the run values of outfield flies have a much wider variance than ground balls or line drives. Much, much wider. There's a little truism in this graph: For line drives and ground balls, the key is how often you hit them. For outfield flies, the key is how you hit them.

Like all truisms, this one shouldn't be applied indiscriminately. **Ichiro** gets a lot of good value out of his ground balls, and **Jason Phillips** should avoid them like the plague. **Frank Thomas** and **Russell Branyan** certainly get a lot out of their line drives.

But the chasm between the best and worst flyball hitters is wide. An outfield fly by Bonds is as good as a line drive by the average major leaguer. On the other hand, an outfield fly by the aforementioned Diaz is just as bad as a Phillips grounder.

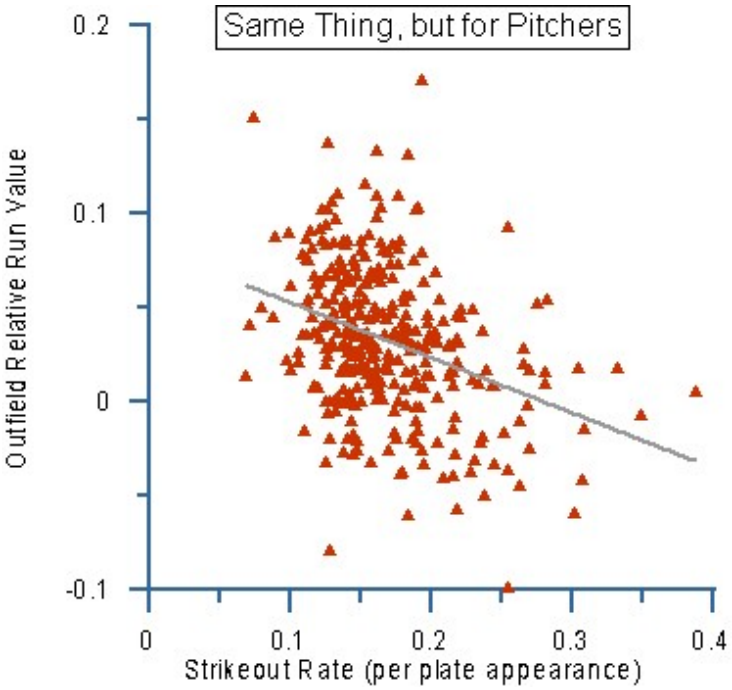
What makes a good flyball hitter? Swinging hard, at the risk of missing. To illustrate the point, here's a graph of strikeout rates and outfield fly values for these batters. I've added a "fitted line" to the graph to highlight the relationship between the two.



The line tells the story. The cost of hitting outfield flies well is striking out more often (for most everyone not named Bonds, **Pujols** or **Vladimir**). I've labeled some of the players who deviate the most from the fitted line. Those farthest above the line are among the very best hitters in baseball. The ones farthest below it aren't, including Diaz (who obviously just needs to strike out more!) and shortstop **Ramon Vazquez**. Branyan is a special case, by the way. His power shows up relatively more often in his line drives.

In case you're wondering (and who isn't?), the R-squared of the relationship between striking out and hitting valuable outfield flies is .40. If you remove Bonds from the sample, it's .43. Maybe you weren't wondering about that.

You know how some people say that strikeouts aren't that bad for batters, but they're great for pitchers? And you know how some other people say that makes no sense? Let me show you the basic difference between the two with another graph, a duplicate of the strikeout rate/outfield fly run value graph, but for pitchers instead of batters (same sample size).



Now scroll up to the batter graph. See the difference? For batters, the line goes up; for pitchers, the line goes down. Although the relationship isn't as strong (an R-squared of .13), it's there. Strikeout pitchers tend to give up lower-impact fly balls. Strikeout batters tend to hit higher-impact fly balls.

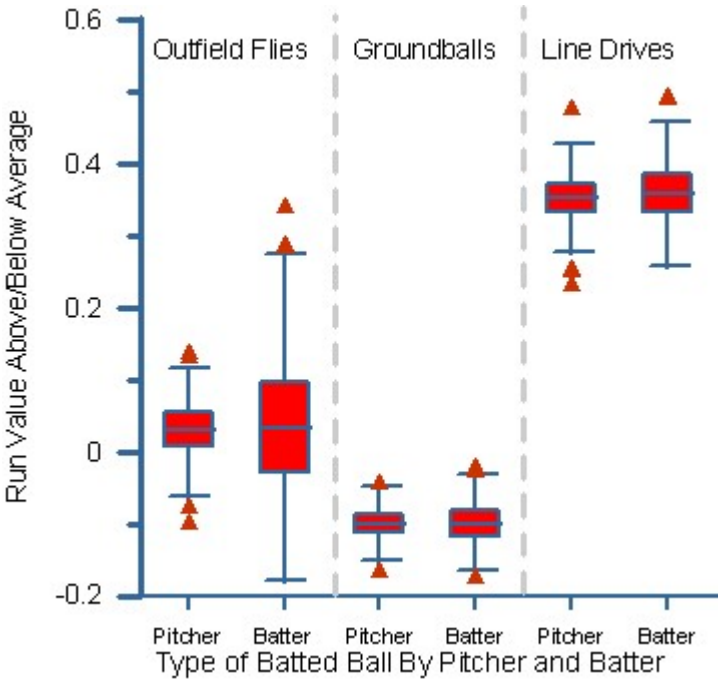
This is essentially the same conclusion J.C. Bradbury reached when he took **"Another Look at DIPS"** earlier this year. He found that strikeout pitchers tend to have favorable fielding stats, and he concluded that DIPS "works" because the formula includes strikeouts.

John Burnson, in the ***2006 Graphical Pitcher*** concluded that strikeout pitchers also have a lower home run-per-fly ball rate. I didn't reach the same conclusion, but I think that's because I calculate home run rates per outfield flies. I did find that strikeout pitchers have a higher rate of infield flies per total fly balls (R-squared of .10), which would yield the same result as Burnson's because infield flies don't go over the outfield fence. Strikeout rates don't have a measurable impact on the run values of line drives or ground balls, as far as I can tell.

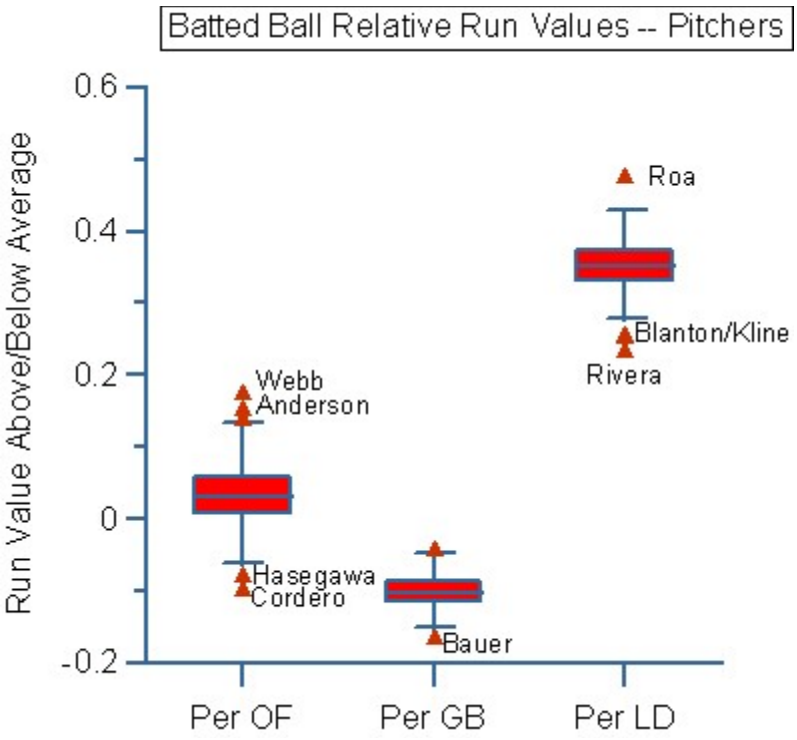
In general, however, batters have a larger impact on the outcome of a batted ball than pitchers do. Following is a combined box-whisker graph, which includes the outcomes of each type of batted ball for both batters and pitchers. As you can see, the width of each box/whisker is wider for batters than pitchers (much wider for outfield flies), which means that pitchers don't differ as much as batters in this regard.

More from The Hardball Times

A Hardball Times Update
by RJ McDaniel
Goodbye for now.



For all three types of batted balls, the outcome has more to do with who's batting than who's pitching. This is the "DIPS insight," if you will. But you also may notice that there are a few more outlier triangles for the pitchers. The outliers tell a story of their own, so let's take a closer look at the pitching box-whisker graph, labeling the outliers.



The two outliers on the top of the outfield fly data are **Brandon Webb** and **Jimmy Anderson**, who also happen to be two of the most extreme groundball pitchers in the sample. Another truism is that when groundball pitchers yield outfield flies, they tend to have more run value (R-squared of .11 between groundball rates and outfield fly values).

The bottom two outliers are **Francisco Cordero** and **Shigetoshi Hasegawa**, two of the better relievers of the past four years. You may have also noticed that **Mariano Rivera** is the low outlier in line drive run values. Which leads me to my last, surprising truism: great relievers do have some demonstrable control over their batted balls. They are the "anti-DIPS."

Consider this: **Billy Wagner's** career **Defense Efficiency Ratio** is .750 by my calculations. Rivera's is .745 if you exclude his first year. These are much higher figures than their respective team DERs, which implies that Wagner and Rivera's batted balls were more catchable than average.

One could argue that we're dealing with small sample sizes, but these two guys have pitched 630 and 740 innings in their careers, respectively. That's equal to three-to-four full seasons of a major league starter.

Or consider this table of the pitchers with the lowest outfield fly relative run values in our sample. If you can look past **Felix Rodriquez**, you'll see that it includes some

of the best relievers of the last four years.

Last	First	OF Run Value
Cordero	Francisco	-0.099
Hasegawa	Shigetoshi	-0.079
Rodriguez	Felix	-0.060
Nathan	Joe	-0.060
Foulke	Keith	-0.058
Rincon	Juan	-0.051
Marte	Damaso	-0.045
Wagner	Billy	-0.042
Hawkins	LaTroy	-0.041
Politte	Cliff	-0.040
Fultz	Aaron	-0.039
Walker	Jamie	-0.038
Smoltz	John	-0.037
Donnelly	Brendan	-0.037
Speier	Justin	-0.034
Benitez	Armando	-0.033
Stanton	Mike	-0.033
Quantrill	Paul	-0.032
Burnett	A.J.	-0.031

Why do top relievers tend to be anti-DIPS? Well, there is probably some selection bias in the data. Also, some relievers are saved for specific batter matchups, which undoubtedly helps. But many of the players on this list are closers or primary setup men; they typically enter a game regardless of who is batting.

My guess is that we're also seeing something fundamental to the relief role. Pitchers who only pitch an inning or two throw harder and only use one or two pitches. The best relief pitchers, like Wagner and Rivera, leverage this situation to overpower batters, even when they hit the ball. Whatever it is that makes them best suited for relief also gives them an upper hand with outfield flies.

This suggests that, when you're looking for top relievers, you might investigate not only their strikeout and walk rates, but their outfield fly run values too. Although there are some wild swings from year to year, the truly elite relievers will tend to be above average over the long term.

Here's a review of some of the game's other top relievers:

- Francisco Cordero is just dominant in this category. Over the past four years, his figures have been -0.138, -0.094, -0.097 and -0.082.
- If he had pitched enough to qualify, **Chad Cordero** would be third-best on the list of outfield fly run values, at -0.065 runs below the average plate appearance. It's true that RFK kills flyballs, but these figures are adjusted for ballparks. In other words, both Corderos are keepers.
- **Eric Gagne**: 0.006 outfield fly run value. Good, but not Corderoesque. When healthy, Gagne is so dominant in the strike zone that it doesn't matter.
- **Jason Isringhausen**: -0.022 for the last four years, but he's been up significantly the last two years. He appears to be a risky proposition going forward.
- **B.J. Ryan**: Four-year average of 0.017, albeit -0.019 last year. Not exactly confidence-building.
- **Brad Lidge**: Outfield fly run values of -0.080 in 2003, 0.046 in 2004 and 0.035 in 2005. That's not a good trend, as the Astros learned to their regret this past offseason.
- **Huston Street** registered -0.109 last year, which is Cordero territory. **Derrick Turnbow**, on the other hand, came in at 0.042, which isn't encouraging. Turnbow shows signs of following the Rivera pattern, however. His line drive relative run value was only 0.206 last year.

References & Resources

After posting this article, one reader asked if I adjusted relievers' outfield fly values by their strikeout rate. The answer is no, but every pitcher listed in the final table had outfield fly values lower than that predicted by their strikeout rate.

Dave Studeman was called a "national treasure" by Rob Neyer. Seriously. Follow his sporadic tweets **@dastudes**.

Comments are closed.

What's a Batted Ball Worth?

by Dave Studenmund

This past summer, Tom Ruane posted a study on Retrosheet (http://www.retrosheet.org) called *The Value-Added Approach to Evaluating Performance*. Don't be intimidated by the title; the idea behind the article was to determine how much each event on a baseball field was worth, and then add up the number of times each batter or pitcher did one of those things. Saying that the article does this well is like saying that Albert Pujols had a pretty good year.

For example, here is a table of how much each of the following events was worth from 2002 through 2004:

Single	.465	runs
Double	.772	runs
Triple	1.055	runs
Home Run	1.394	runs
Non-Intentional Walk	.315	runs
Intentional Walk	.176	runs
Hit by Pitch	.342	runs
Sacrifice Hit	-.127	runs
Sacrifice Fly	-.052	runs
Double Play	-.839	runs
Strikeout	-.287	runs
Other kinds of outs	-.250	runs

Tom developed this list by evaluating the impact of every play in every game on the number of runs each team eventually scored. As you can see, there is a whole lot of information packed into these little bitty numbers, to wit…

- A home run is worth about three times as much as a single. This is why slugging percentage is not a great stat, though it's still a lot better than batting average (in which a home run is the same as a single).
- A walk is worth about 2/3 of a single. But intentional walks, dictated by game strategy, usually yield about half as many runs as a regular walk.
- In general, sacrifice bunts and sacrifice flies add more outs than runs.
- A strikeout really is worse than a regular out (or better, if you're pitching).

At *The Hardball Times*, we also publish a lot of baseball information collected by our friends at *Baseball Info Solutions (BIS)*. One of the most unique items they collect is something called "batted-ball type." For each plate appearance, BIS notes whether the batter hit a ground ball, fly ball or line drive. On our website, for instance, you can see how many line drives each batter has smashed, or how many each pitcher has allowed.

Recently, we started playing around with the outcome (otherwise known as the baseball event) of every type of batted ball from 2002 through 2005. Using this data, we asked a bunch of questions, such as "How many times does a fly ball become a hit? Or a home run? What about line drives? Why do you park your car on a driveway but drive your car on a parkway?"

To answer some of those questions, we looked at the results of each kind of batted ball from 2002 through 2005:

	Outfield Fly	Groundball	Line Drive	Infield Fly	Bunt
Fair Out	74.4%	60.6%	25.4%	53.3%	67.1%
Foul Out	3.0%	0.0%	0.0%	45.3%	2.1%
Double Play	0.2%	6.4%	0.8%	0.1%	1.3%
Error	0.3%	2.5%	0.2%	0.3%	2.8%
Fielders' Choice	0.1%	7.8%	0.2%	0.1%	7.8%
Single	4.1%	20.7%	51.4%	0.3%	18.6%
Double	6.1%	1.9%	18.3%	0.0%	0.0%
Triple	0.9%	0.1%	1.6%	0.0%	0.0%
Home Run	11.0%	0.0%	2.0%	0.1%	0.0%

Take some time with this table, because there is a lot here. And while you're looking, allow me to make a few points:

- Flies to the outfield are either really good or really bad. They're caught for outs three quarters of the time, but they make it over the fence 11% of the time. Outfield flies are the dramatic flourishes of the baseball bat. Only when the ball lands do you know if you've witnessed a tragedy or comedy.
- If you include double plays and fielder's choices, ground balls are turned into outs about as often as outfield flies. Most other times, ground balls are singles. In fact, a batter is more likely to reach on a ground ball error than a ground ball double.
- Line drives are pure baseball. They're either caught for outs (a quarter of the time) or batted for singles and doubles.
- The infield fly is a pitcher's secret weapon, an automatic out. They are caught for outs nearly 99% of the time, almost half of those in the foul area.
- Bunts are almost as productive as regular ground balls. Of course, this depends on who's doing the bunting.

So we have runs per event and events per batted ball. Kind of like peanut butter and chocolate, horses and carriages, these are two things that really go together. By simply multiplying the two previous tables, we have a new kind of table: the value of each type of batted ball.

Batted Ball	Run Value
Line Drive	.356
HBP	.342
Non-Intentional Walk	.315
Intentional Walk	.176
Outfield Fly	.035
Groundball	-.101
Bunts	-.103
Infield Fly	-.243
Strikeout	-.287

I threw in batting events that don't involve batted balls, such as strikeouts and walks. Otherwise, this table is simply the product of multiplying the value of each type of event times the number of times it occurred for each batted ball.

As you can see, on average the best thing for a batter to do is to hit a line drive. The best thing for a pitcher is a strikeout. There is a lot of nuance in between, however. For instance, this table shows the power of the walk (and hit by pitch)—it is second to only the line drive in its value.

On the other hand, an infield fly is almost as good as a strikeout. To the extent pitchers can induce infield flies from batters, they are almost as effective as power strikeout pitchers.

Bunts aren't really much worse than ground balls. And the true difference between the outfield fly (a somewhat positive value) and the ground ball (negative value) is the home run.

Batted ball information like this allows you to look at the baseball diamond a little differently. It adds another dimension to what's happening on the field. Keep this chart in the back of your mind next time you watch a game. It will give you some brand new insights.

Batted-ball information also permits you to investigate a few things you might be curious about, such as…

- Why are some parks pitcher's parks and other parks batter's parks?
- Do pitchers and batters have a consistent ability to hit line drives, induce infield flies and other cool things?
- What does the batted-ball data tell us about fielders?

Read on.

Retrosheet is a miraculous site, containing detailed box scores, stats and research for the entire history of baseball. It is a nonprofit site, meaning that you won't run into a single ad or popup window. Donations are tax-deductible.

You can read Tom Ruane's article at http://www.retrosheet.org/Research/RuaneT/valueadd_art.htm. The article includes detailed lists of the best batters and pitchers from 1960 through 2004, as determined by the value added by the events on the field.

Batted Ball Fielding Stats

by Dave Studenmund

I've noticed a disturbing trend in baseball columns these days. More and more writers are referring to something called Defense Efficiency Ratio (or DER) to describe a team's fielding prowess. This is disturbing to me, because DER has some serious flaws.

The calculation for DER, which was introduced by Bill James over 20 years ago, is relatively simple. Take all the balls in play given up by a pitching staff (batters faced minus strikeouts, walks, batters hit by pitches and home runs) and then figure out how often the team's fielders recorded an out off those balls in play. The second part is a little tricky, because you don't want to include outs that occurred in other ways, such as a runner caught stealing or the first out of a double play. But you can usually find a way to get the right stats.

It's true that good fielders get to more balls than bad fielders. But it's also true that many other things can affect DER, such as:

- The ballpark. Try catching an easy flyball 15 feet up the Green Monster in Fenway.
- The type of batted ball. Line drives are hard to catch; infield flies are easy. Outfield flies, groundballs and bunts are in-between.
- Where the ball goes. Balls in the shortstop hole are harder to field than balls hit directly at the shortstop. I won't name names.
- How hard the ball is hit. Even Rafael Furcal will have trouble with a ball hit hard in the shortstop hole.

So when you see a columnist use DER to announce that one team's fielders are better than another's, you should be skeptical. It's not that DER is wrong; it's just not always right. Let me give you an example.

Last year, the Yankees' DER was .691, according to our stats from *Baseball Info Solutions*; 69% of qualified batted balls were fielded for outs. That would place the Yankee fielders slightly below the major league average DER of .695. But the Yankee fielders were actually much worse than that.

You see, the much-maligned Yankee staff actually yielded the most fieldable batted balls in the majors last year. If you add up all of their batted-ball types and assume that each type was turned into an out at the average major league rate, their DER would be .721. So when judging the Yankee fielders, you should compare them to .721, not .695.

Luckily, *The Hardball Times* can help. We used our batted-ball data to develop better fielding stats for 2005, and we found that the Yankees were actually the third-worst fielding team in the majors last year.

Here's what we did.

- We added up the number of batted balls allowed by each team's pitching staffs. The Yankees allowed 2,161 groundballs, 1,280 outfield flies (not including home runs), 798 line drives (also not including home runs), 188 infield flies and 57 bunts for a total of 4,484 batted-balls in play.
- We then applied the major league average out percentage for each type of batted ball (such as 99% for infield flies and 25% for line drives) to each total to generate the number of expected outs from the batted balls. We also adjusted the out percentage for each team based on the ballpark factors discussed in the previous article. (The fielding impact of Yankee Stadium is pretty small). This produced a total of 3,235 expected outs off those batted balls.
- Next, we compared the expected total to the actual number of batted balls turned into outs: 3,146, or 89 fewer than expected.
- Finally, we converted each unfielded ball into a run value, based on how often each type of batted ball hit is a single, double or triple, on average. As you can imagine, an unfielded outfield fly does more damage than an unfielded bunt.

We're still missing two important elements, where the ball was hit and how hard it was hit. But with just the stats we have, we find that Yankee fielders allowed 51 runs below average. When you consider that every nine-to-ten runs equal a win, this means that their fielders cost them at least five wins compared to the average major league team.

I don't mean to pick on the Yankees; they're just the example I chose. Actually, the Reds (57 runs below average) and Royals (67 runs below average) were worse. The best fielding teams were the Indians (49 runs above average), Athletics (46) and Phillies (40). The difference between the best (Cleveland) and worst (Kansas City) fielding teams was 116 runs.

To put that in perspective, the difference between the best and worst defensive teams (pitching and fielding) last year was 302 runs. (Tampa Bay allowed 936

runs and the Cardinals allowed 634.) Pitching is still the most important aspect of total defense, but fielding matters too.

The following table has more information than you can shake a stick at. By team, it shows the number of runs allowed above/below average for each type of batted ball and in total, as well as each team's rank in DER and their fielding runs above/below average in 2004. Let me lay it out for you and add comments afterward:

DER			Fielding Runs Above/Below Average							
Rank	Rank	Team	IF	OF	LD	GB	Bunt	**Total**	2004	Diff
1	3	Indians	0.10	22.52	21.52	4.19	0.76	**49.09**	-11.1	60.2
2	1	Athletics	0.58	9.41	13.48	22.11	0.40	**45.97**	4.7	41.3
3	5	Phillies	0.05	13.50	5.76	20.85	-0.11	**40.04**	0.1	39.9
4	2	White Sox	-0.38	21.23	4.53	13.78	0.64	**39.79**	5.0	34.8
5	4	Astros	-0.06	5.35	9.21	21.00	1.33	**36.82**	-4.6	41.4
6	19	Braves	-1.06	24.66	4.81	-3.43	0.87	**25.84**	17.1	8.7
7	8	Mets	0.47	19.29	-11.55	12.58	4.03	**24.81**	29.7	-4.8
8	11	Cubs	-0.09	6.97	11.60	2.08	3.22	**23.79**	19.8	3.9
9	6	Cardinals	-0.61	-17.26	2.86	27.61	3.48	**16.08**	39.8	-23.7
10	10	Blue Jays	-0.45	7.64	-6.88	15.51	-0.77	**15.05**	8.7	6.4
11	20	Orioles	-0.48	12.09	6.06	-7.07	-3.67	**6.94**	-14.3	21.3
12	15	Brewers	0.56	-3.75	11.61	-2.05	0.29	**6.65**	1.9	4.7
13	18	Pirates	0.55	-16.35	2.75	20.26	-1.70	**5.51**	-13.5	19.0
14	16	Nationals	0.15	-5.00	9.42	-3.17	0.82	**2.22**	15.9	-13.7
15	7	Twins	0.12	-3.73	-6.63	9.83	2.16	**1.75**	-23.5	25.3
16	12	Dodgers	-0.96	3.19	-5.44	-2.19	4.61	**-0.80**	28.6	-29.4
17	14	Giants	-0.29	-8.38	7.33	-3.15	3.17	**-1.33**	0.5	-1.8
18	9	Mariners	0.58	15.72	-5.17	-17.88	-0.74	**-7.49**	9.0	-16.5
19	23	Diamondbacks	0.00	-12.70	-1.71	4.00	2.09	**-8.33**	3.9	-12.2
20	22	Padres	0.04	-10.05	-3.43	4.93	-2.13	**-10.64**	5.0	-15.6
21	17	Tigers	0.55	-2.15	-7.25	0.04	-2.24	**-11.04**	-21.7	10.6
22	13	Angels	0.08	-10.73	-0.18	4.93	-6.71	**-12.61**	-40.5	27.9
23	24	Red Sox	-0.34	2.31	-17.71	-0.44	-0.86	**-17.04**	18.5	-35.5
24	26	Rangers	0.51	4.52	5.31	-32.48	-2.51	**-24.65**	-3.6	-21.0
25	25	Devil Rays	0.13	7.83	-3.99	-26.99	-2.71	**-25.73**	10.5	-36.2
26	29	Rockies	-0.03	-11.61	-10.19	-3.07	-2.31	**-27.21**	-33.3	6.1
27	27	Marlins	0.07	-13.54	-2.88	-14.74	3.82	**-27.28**	12.5	-39.8
28	21	Yankees	0.56	-25.62	2.05	-24.00	-3.68	**-50.69**	-32.8	-17.9
29	28	Reds	-0.37	-27.16	-18.86	-9.82	-0.56	**-56.75**	-4.6	-52.1
30	30	Royals	0.02	-14.89	-16.41	-35.24	-0.97	**-67.49**	-32.2	-35.3

It's really not that bad; let me point out a few things.

The first two columns show each team's rank in Fielding Runs Above/Below Average— let's just call it FRAA for this article—and its rank in DER. As you can see, DER is close, but it misses badly on a few teams, such as the Braves, Orioles, Mariners, Angels and Twins.

The five middle columns show each team's FRAA by batted-ball type. This allows you to say a few things about each team's outfield and infield defense.

Outfields

- The Braves, with Andruw Jones in center and Jeff Francoeur in right, had the best outfield in the majors last year. The Indians and White Sox also had great outfields.
- On the other hand, the worst outfields were the Reds' and Yankees'. Bernie Williams' outfield limitations are fairly well known, but it appears that Mr. Griffey Jr. has lost his outfield panache as well.

Infields

- Even though the Cardinals' infield almost had a complete turnover this year at second, shortstop and third base, they still had the best infield defense in the majors. Findings like this make you wonder if the ballpark is having some sort of impact, but no such impact is apparent in the data.
- There were a number of other fine infields last year, including the A's, Astros, Phillies and Pirates.
- The Royals' and Rangers' infields were truly terrible in 2005. A number of teams had big differences between their infields and outfields, but the Royals were really, really bad in both.

By the way, research shows that when line drives are caught for outs, the outfield accounts for a little more than half of those outs. This makes it tough to say whether a good record at turning line drives into outs is the result of good plays by the outfield or infield, or just plain good luck.

The last two columns list each team's 2004 Fielding Runs, as well as the difference between this year and last. The team with the greatest improvement from 2004 to 2005 was the Cleveland Indians, the best story in the American League the second half of the season. During the Indians' mad run for the Wild Card slot, a lot of attention was paid to their improved pitching and second-half hitting. Not many people mentioned their improved fielding.

In retrospect, it should have been obvious. The Indians allowed 857 runs in 2004 and only 642 in 2005—a difference of 215 runs. It's hard to make that much of an improvement in pitching alone. Indeed, the Indians' fielders contributed 60 of that 215 run difference.

So here's a salute to the Indians' reconfigured outfield of left fielder Coco Crisp, center fielder Grady Sizemore and right fielder Casey Blake. I nominate them for the unsung fielding heroes of 2005.

They Play in Parks

by Dave Studenmund

Ballparks have long been the "invisible hand" of the ol' ballgame. From the crazy dimensions of the Polo Grounds inflating Mel Ott's home run totals to Sandy Koufax's lifetime 1.37 ERA at Dodger Stadium, ballparks have had a huge impact on baseball games and the players that have played in them. And I haven't even mentioned the "House that Ruth Built."

In today's baseball world, ballparks range from the rarefied atmosphere of Denver's Coors Field to the low-scoring environment of PETCO Park in San Diego. Various ballparks have turf, ivy, short porches, deep alleys, roofs, retractable roofs and Green Monsters. Ballplayers spend a lot of time getting to know each angle off every wall, and general managers spend a lot of energy figuring out which type of player will feel most at home in their parks.

As a result, baseball fans sometimes like to track something called a "park factor," which is a ratio that represents the level of offense at each park compared to other parks. According to last year's *Bill James Handbook*, park factors ranged from 82 at PETCO to 136 at Coors. Since 100 represents average, this means that runs scored 36% more often at Coors and 18% less often at PETCO. It also means that runs scored 65% more often at Coors than PETCO (136 divided by 82).

But sometimes park factors are impish, inconstant measures. For instance, the park factors at Cincinnati's Great American Ballpark (GAB) each of the three years since it opened have been 99, 85 and 114 (according to my simple calculations), a swing of almost 30 points. Someone once e-mailed me to say (and I'm paraphrasing) "Well, the GAB was a pitcher's ballpark last year, but it's a hitter's park this year." But how can a ballpark fundamentally change like that?

Typical culprits include the weather (wind and humidity, for instance, are two meteorological conditions that influence what happens to a ball), or maybe something subtle (the slope of the mound, for instance). Maybe the players themselves started playing differently at home for whatever reason.

But if park factors can change so much, can they be helpful at all?

I'd like to partially answer this question by discussing a complex mathematical calculation called regression to the mean. I promise I'll be quick.

Statisticians, baseball and otherwise, engage in sample sizes. When statisticians want to predict which candidate is likely to win the next election, they don't ask everyone. They only ask some of the voters, and they're careful to make sure the sample size is large enough and that the folks they poll are representative of the greater voting population. In baseball, we don't have that luxury.

In baseball, we get 162 games a year, like it or not. We get unbalanced schedules, which means that external comparisons between teams aren't the same. And yes, we get weather and changing mound slopes.

So think of a baseball season as an imperfect sample of a ballpark's tendencies. It takes more than 162 games to really know the impact of a park, weather or not. For instance, Arizona's Chase Field, a domed stadium in the desert, has ranged from 107 to 121 in each of the past four years. Weather had nothing to do with it.

To get a better handle on park factors, baseball statisticians can do three things (at least!):

- Use multiple years for park factors. The more years, the better—unless something has changed in the park (such as moving the fences in). The *Bill James Handbook* calculates three-year park factors, for instance.
- Adjust the sample. The *Handbook* excludes interleague games, because the designated hitter skews AL ballparks to higher offense. Great point but, unfortunately, steps like this decrease the sample size.
- Regress the sample to the mean.

Here's what you do to regress to the mean: Using regression analysis, you find out how much one-year park factors are correlated. You can also find out how much two-year average factors correlate with a third year, and how much three-year factors correlate with a fourth year, etc. etc. Correlation is measured using something called R (see the other article (must be changed to refer to JC and David's article) for a definition of R), where one (1) means that you can exactly predict next year's park factor from the previous year's.

Here's the mathematical paragraph: One-year park factors typically have an R of about .6, two-year factors are around .7 and three-year factors are around .8. So to regress a one-year park factor to the mean, you multiply it by the relevant R (.6 in this case) and you multiply the average by 1-R. Since average equals 100 for park factors, the math is PF*.6 + 100*.4. If your park factor

is 120, your one-year regressed factor is 112, your two-year factor is 114 and your three-year factor is 116. So you can see how larger sample sizes increase your confidence in the factors.

Going from 120 to 112 might not sound like a big change, but go back and calculate Cincinnati's regressed park factor at the end of its first two years. You'll see that a two-year regressed park factor will cut down the error by a third, as opposed to just using a one-year factor.

I learned all of this from an Internet baseball wonk who calls himself U.S. Patriot. Patriot maintains his own website with a host of useful (albeit highly mathematical) essays and spreadsheets (http://gosu02.tripod.com/id7.html). With his permission, we are reprinting his Run and Home Run Park Factors for 2006. They include up to five years' worth of data, if appropriate, and they're regressed to the mean. Please note that Patriot has taken the added step of essentially cutting his park factors in half, because teams play only half their games at home. In other words, you can multiply or divide these factors directly onto a player's stats since a player plays half his games at home.

Team	Run PF	HR PF	Team	Run PF	HR PF
ARI	105	106	BAL	97	101
ATL	100	99	BOS	102	97
CHN	100	104	CHA	102	113
CIN	99	106	CLE	98	96
COL	115	114	DET	97	93
FLA	96	93	KC	98	91
HOU	102	104	LAA	98	97
LA	94	101	MIN	101	95
MIL	100	104	NYA	99	103
NYN	97	95	OAK	99	101
PHI	103	107	SEA	95	96
PIT	100	95	TB	99	96
SD	94	91	TEX	106	107
SF	97	89	TOR	103	105
STL	98	96			
WAS	96	94			

I'm sometimes asked what THT's park factors are. There's your answer.

Park factors like these are extremely important but, in some ways, they're only the tip of the iceberg. Ballparks have quirks, like big walls in left field, short fences in right, low air pressure, bad infields, poor sightlines, artificial turf, etc. etc. Park factors pick up the sum impact of these quirks, but they miss the details. And one of the reasons park factors change is that different batters and pitchers have different styles of play themselves. Sometimes they match the ballpark, sometimes they don't.

So we've done something at *The Hardball Times* that I haven't seen before. We took the batted-ball information from our buddies at *Baseball Info Solutions* and looked at what happened in every park over the last four years. Essentially, we developed annual park factors for every type of batted ball in each ballpark. We looked at how often batters struck out and walked, or how often they hit flies, line drives and grounders at each park. We also looked at how often each batted ball was an out or a hit (single, double, triple or home run). We even looked at how often there were errors on each type of batted ball.

We found some things that we expected, but we found some other things that blew us away. Here are some of those things …

Strikeouts per Plate Appearance

In our data, the most persistent trait of a ballpark was its strikeout ratio. We correlated three years of data against the fourth year, to see how predictable trends were, and strikeout factors were highest at .79. Average strikeout park factors ranged from 88 at Coors to 112 at Florida's Dolphins Stadium. In fact, the most important aspect of Dolphins Stadium is its strikeout ratio.

Why does this happen? I'm guessing factors like sightlines and heavy atmosphere causing balls to break more. Whatever the reason, the data is fairly clear.

Outs and Home Runs Per Outfield Fly

These two factors were the second and third most persistent ones, which is not really a surprise. The outcome of an outfield fly has the biggest impact on park factors in general. As you can imagine, outs and home run factors are related, because the more outfield flies go over the fence, the fewer are caught for outs.

Some of the exceptions to this rule are…

- Fenway, where the monster wall in left field turns both would-be home runs and outs into singles and doubles,
- Yankee Stadium, a relatively average home run park where outfield flies are caught for outs more often than any other park, and
- Dolphins Stadium, again. Dolphins Stadium is an average park for outs per outfield fly but the second lowest in home runs per outfield fly. Home runs

don't turn into outs at Dolphins Stadium; they turn into singles, doubles and triples.

Foul Outs Per Outfield Fly

It's fairly well known that parks have different foul areas, and our data showed this. But I was a little surprised to find that foul outs per outfield flies have a much stronger pattern than foul outs per infield flies. The three-year correlation for foul out per outfield fly was .637 but was only .316 for infield flies. By the way, BIS considers anything beyond the base running paths to be an outfield fly, regardless of who catches it.

In addition, some of the differences were extreme. The factor at Oakland's Network Associates Coliseum is 192, and it's 160 at Tampa Bay's Tropicana Field. On the low end, it's 40 at Yankee Stadium and 46 at Fenway.

Extra Base Hits per Ground ball

This was initially a surprise to me. We looked at the factor of doubles and triples per ground ball at each park and found a three-year correlation of .64, a higher factor than I expected. Upon closer inspection, however, we found that the ballparks with the highest factors in this were mostly turf parks, where grounders can skip through infields and into corners very quickly. These parks were Rogers Centre in Toronto (137), the Humphrey Dome in Minnesota (126), Tampa Bay's Tropicana (123) and Chase Field in Arizona (123). That last park doesn't have turf, by the way.

On the other end of the spectrum, here are the parks where it is least likely that a ground ball will be a double or triple: Yankee Stadium (69), Comerica Park in Detroit (80) and Baltimore's Camden Yards (80).

There was also a good correlation in doubles and triples per line drive (.54) and most of the same parks made the best and worst lists. One exception is Milwaukee's Miller Park, which has the second-highest factor (117) for line drives. I have no idea why.

Ground balls and Line Drives Per Batted Ball

Speaking of being clueless, this finding truly shocked me. Essentially, we found that ballparks have a persistent trend in the types of batted balls that are hit. In other words, a batted ball is more likely to be a groundball in some parks, or a line drive in others. The impact isn't huge in most cases, but it's persistent.

I'd never heard of this before, so I double-checked the findings many times. I put out feelers to various folks asking if they had heard of such a thing and received a negative reply in each case. In fact, when I mentioned the results on *The Hardball Times* website, one reader said that I "had gone horribly wrong." Maybe, but let me share the results with you.

Line drives per batted ball have a three-year correlation of .59 and ground balls per batted ball have a three-year correlation of .54. We didn't leave home runs out of the equation, so different home run factors are not to blame. The foul area does have some impact, because parks with large foul areas will wind up with more fly balls as more of them are caught for outs. But that impact is relatively small.

The highest line drive factors are at Coors (116) and Texas's Ameriquest Field (105), while the lowest line drive factors are at Dodger Stadium (93) and Dolphins Stadium (95).

The highest ground ball factors are at Cleveland's Jacobs Field (107) and Dodger Stadium (104), and the lowest ground ball factors belong to Yankee Stadium (96), US Cellular Field in Chicago (97) and Safeco in Seattle (97).

In each of the four years we examined, for both batters and pitchers, Jacobs Field was a ground ball park. This was the single most surprising finding to me. The ground ball factor is Jacobs's most important ballpark influence.

I don't know why this is, but I do know that it could have implications for building a team in Cleveland. With more ground balls, the Indians can put a relatively higher value on infield defense and less of an emphasis on outfield defense. In other words, if you're going to sacrifice defense for offense, the Indians should do it in the outfield. This is a good example of how strategy can evolve from a careful analysis of park factors.

Please note that we only considered ballparks that have been around, unaltered, for four straight years. We left a number of new parks out of the analysis, including PETCO, Citizen's Bank and the Great American Ballpark, as well as Kansas City's Kauffman Stadium, which had its fences moved out after the 2003 season. In retrospect, we shouldn't have included Detroit in our sample, because they lowered their fences after the 2003 season. As proof, here is their HR/OF park factors from 2002 to 2005: 70, 77, 85 and 92. Better, but it's still not the Polo Grounds.

So, that's what we found. I haven't mentioned every ballpark or finding, just the highlights. I know you probably have a favorite team you're wondering about. So following is a table of commentary and statistics for each major league ballpark. There are two stats, Run Impact and Ball Factor.

Run Impact shows the value of each batted ball type relative to the major league average. In other

words, an outfield fly in Fenway has a run impact of 0.043, because of the net effect of lots more singles and doubles but fewer triples and home runs off the Green Monster yields a total run impact of .078 per fly ball, or .043 runs more than the MLB average of .035. You can find average run values of all batted ball in the article "What's a Batted Ball Worth?"

Ball Factor is kind of like Park Factors, only for batted balls. This is a number that represents the relative impact of both changes in frequency (such as more ground balls in Jacobs) and run impact for a type of batted ball. I've expressed ball factor in the same format as park factor.

By the way, these figures aren't regressed to the mean. They're raw four-year averages (or less, where noted). We thought you might like to see the raw data in this case.

In the following shorthand comments, "rate" is the word I use for a true rate, such as strikeouts per plate appearance, while "factor" is the word I use for each type of park/ball factor, where 100 is average. K stands for strikeouts, OF for outfield flies, GB for ground balls and LD for line drives. Notations such as "2B/OF" stand for doubles per outfield flies, for instance. And where I say "OF factor" or "GB factor," I'm referring to the factor of how often a certain type of batted ball occurs relative to other types of batted balls.

AMERICAN LEAGUE BALLPARKS

Baltimore Orioles		**K**	**OF**	**GB**	**LD**
Camden Yards	Run Impact:		-0.008	-0.005	-0.020
	Ball Factor:	103	98	99	96

Comments: K factor of 93%, which helps offset the higher out factors of batted balls. Avg HR/OF park, but fewer other types of hits from OF.

Boston Red Sox		**K**	**OF**	**GB**	**LD**
Fenway	Run Impact:		0.043	0.001	-0.010
	Ball Factor:	101	109	100	100

Comments: It's all about the Green Monster: 2B/OF factor is 176

Chicago White Sox		**K**	**OF**	**GB**	**LD**
U.S. Cellular	Run Impact:		0.040	-0.014	0.014
	Ball Factor:	100	110	97	101

Comments: HR/OF factor is 127, second highest behind Coors. Also, OF factor is 105. Hence the fireworks.

Cleveland Indians		**K**	**OF**	**GB**	**LD**
Jacobs	Run Impact:		-0.001	-0.009	-0.021
	Ball Factor:	97	99	95	96

Comments: Ground balls. Higher frequency, plus they're turned into outs more often. Also, Out/LD factor is 110.

Detroit Tigers		**K**	**OF**	**GB**	**LD**
Comerica	Run Impact:		-0.037	-0.010	-0.001
	Ball Factor:	103	92	97	102

Comments: Out/OF factor is 103 fair, 112 foul. In addition to changing HR/OF factors, 2B factor is 80 or less for all types of batted balls.

Kansas City Royals		**K**	**OF**	**GB**	**LD**
Kauffman	Run Impact:		-0.049	-0.010	0.001
	Ball Factor:	104	89	96	100

Comments: Two years of data, since they moved fences back. HR/OF factor is 73. 3B/OF factor is 143.

Anaheim Angels		**K**	**OF**	**GB**	**LD**
Edison Park:	Run Impact:		-0.014	-0.002	-0.011
	Ball Factor:	100	97	99	97

Comments: HR/OF is 90, 3B/OF is 83, but 3B/GB is 221. Small sample size flukes, but interesting. 1B/OF factor is 120.

Minnesota Twins		K	OF	GB	LD
Humphrey Dome	Run Impact:		-0.013	0.013	-0.001
	Ball Factor:	97	97	104	99

Comments: Strikeout rate is 107%, tendency to be ground ball park (factor of 102). 3B/GB factor is 428.

New York Yankees		K	OF	GB	LD
Yankee Stadium	Run Impact:		-0.007	-0.002	0.001
	Ball Factor:	98	99	101	100

Comments: Overall average, but a lot going on underneath. Induces outfield flies more than ground balls. HR factors OK, but 2B and 3B factors are low.

Oakland Athletics		K	OF	GB	LD
Network Associates	Run Impact:		-0.015	-0.013	0.024
	Ball Factor:	101	97	97	105

Comments: Huge foul area, even for line drives. Foul out per LD factor is 342. But Fair out/LD factor is only 93.

Seattle Mariners		K	OF	GB	LD
Safeco	Run Impact:		-0.033	0.012	-0.014
	Ball Factor:	96	93	105	95

Comments: Fly ball park, OF factor is 106, Out per OF is 103, foul out per OF is 114. All base hits off OF are 90 or lower.

Tampa Bay Devil Rays		K	OF	GB	LD
Tropicana	Run Impact:		-0.042	0.022	0.007
	Ball Factor:	99	91	107	99

Comments: Foul outs per OF factor is 160 and fair outs factor is 103. Favors ground ball hitters: 106 for 1B/GB, 121 for 2B/GB, 183 for 3B/GB.

Texas Rangers		K	OF	GB	LD
Ameriquest	Run Impact:		0.057	-0.004	0.017
	Ball Factor:	102	113	98	106

Comments: Masher's park: only 62 factor foul out per OF, 97 for fair outs. LD factor is 105 and are outs/LD factor is 95.

Toronto Blue Jays		K	OF	GB	LD
Rogers Centre	Run Impact:		0.026	0.021	0.011
	Ball Factor:	100	105	106	101

Comments: OF factors are a mixed bag: OF factor is 97, foul out is 118, 1B factor is 87, HR factor is 118. GB hit more often (103) and have much more run value.

NATIONAL LEAGUE BALLPARKS

Arizona Diamondbacks		K	OF	GB	LD
Chase Field	Run Impact:		0.015	0.004	0.025
	Ball Factor:	102	104	101	105

Comments: Hitter's park across all batted balls. Out/LD factor only 91. 2B/GB is 121. 3B factors for both OF and LD above 150.

Atlanta Braves		K	OF	GB	LD
Turner Field	Run Impact:		-0.017	0.007	-0.012
	Ball Factor:	101	96	103	100

Comments: OF trap. Rate is 102, fair outs 101, foul outs 116. 3B come from GB (235) not OF (73). May be due to Andruw effect.

Chicago Cubs		**K**	**OF**	**GB**	**LD**
Wrigley	Run Impact:		0.045	-0.006	-0.017
	Ball Factor:	98	110	99	95

Comments: OF factor is 101, but fair out per OF factor is 96 and foul out per OF is 81. HR/OF factor is 117. Not a great place for a fly ball pitcher.

Cincinnati Reds		**K**	**OF**	**GB**	**LD**
Great American	Run Impact:		0.016	-0.021	0.005
	Ball Factor:	100	104	95	102

Comments: Three years of data says: Death to ground balls. GB factor is 96, and Out per GB is 105. HR/OF factor is 111. Find GB pitchers!

Colorado Rockies		**K**	**OF**	**GB**	**LD**
Coors Field	Run Impact:		0.077	0.003	0.016
	Ball Factor:	106	117	101	113

Comments: It's Coors. More line drives, fewer outs.

Florida Marlins		**K**	**OF**	**GB**	**LD**
Dolphins Stadium	Run Impact:		-0.012	-0.001	-0.012
	Ball Factor:	94	97	102	94

Comments: K rate is biggest issue. High OF factor, with fewer HR per OF (83) but more 1B, 2B and 3B. LD factor only 95.

Houston Astros		**K**	**OF**	**GB**	**LD**
Minute Maid Park	Run Impact:		0.019	0.001	0.014
	Ball Factor:	100	105	99	101

Comments: Low LD out factor (92) and high HR/OF factor (108) are key.

Los Angeles Dodgers		**K**	**OF**	**GB**	**LD**
Dodger Stadium	Run Impact:		0.008	-0.004	-0.008
	Ball Factor:	97	101	98	95

Comments: Favors GB (104), not LD (93), which is why there are fewer 2B and 3B in general. Still, 2B/OF factor is 90 and 2B/ LD is 89. HR/OF is actually 110.

Milwaukee Brewers		**K**	**OF**	**GB**	**LD**
Miller Park	Run Impact:		-0.008	0.001	0.018
	Ball Factor:	98	98	101	101

Comments: K rate is 105, OF out factor is 102, but foul out factor/OF is 77. All OF hit factors below 90 except HR/OF (104). 2B/LD is 116, 3B per LD is 121 and HR/LD is 155. All small samples, but line drives appear to pay off a bit more here.

New York Mets		**K**	**OF**	**GB**	**LD**
Shea Stadium	Run Impact:		-0.014	-0.001	0.003
	Ball Factor:	100	97	99	99

Comments: OF factors: 1B: 124, 2B: 108, 3B: 67, HR: 88. LD factor is 97, though Out/LD factor is 94.

Philadelphia Phillies		**K**	**OF**	**GB**	**LD**
Citizens Bank	Run Impact:		-0.002	0.011	0.024
	Ball Factor:	102	100	103	107

Comments: Only open for two years. LD factor is 105. K rate is 95%. Foul outs on OF factor is 135.

Pittsburgh Pirates		**K**	**OF**	**GB**	**LD**
PNC Park	Run Impact:		-0.021	0.009	-0.011
	Ball Factor:	103	95	103	101

Comments: OF out factor is 103, HR factor is 89. More LDs are hit (104) but more are outs (105).

St. Louis Cardinals		K	OF	GB	LD
Busch Stadium	Run Impact:		-0.020	0.007	0.000
	Ball Factor:	100	96	102	98

Comments: OF factor is 103, HR/OF is 90. LD factor is low at 97. Will next year's park be different?

San Diego Padres		K	OF	GB	LD
PETCO Park	Run Impact:		-0.043	-0.009	-0.025
	Ball Factor:	96	90	98	94

Comments: Favors all kinds of pitchers. Higher OF factor (102) and lower LD (97) factor; Very high out factors: OF (104), GB (106), LD (108). Two years of data.

San Francisco Giants		K	OF	GB	LD
Pac Bell Park	Run Impact:		-0.029	-0.006	0.027
	Ball Factor:	101	93	96	105

Comments: Groundball (104) much stronger than fly ball (95) factors. HR/OF only 82. Hitter's opening: Out/LD is 89.

Washington Nationals		K	OF	GB	LD
RFK	Run Impact:		-0.056	0.013	-0.033
	Ball Factor:	96	87	107	92

Comments: Only open one year, extreme park. K rate is 109, OF factor is 110 vs. GB factor of 93. Out per OF factor was 105, per LD was 117. Best hitter strategy: stay out of town.

Batted Balls Redux

by Dave Studenmund

A New Run Table

In last year's *Annual,* I published a table that listed the number of runs each type of batted ball and batting event produces, on average. Actually, it didn't quite show that. To get technical, it showed the average number of runs each type of batted ball produces, relative to the overall average.

If you think that's confusing, so did some readers. For instance, the table showed that the relative run value of a ground ball is -0.16. In other words, the average ground ball is worth .16 runs less than the average batting event. I noticed during the season, however, that some bloggers cited that number as proof that ground balls are a negative event. Not true! Ground balls do produce runs; they just produce fewer runs than line drives and outfield flies.

Still confused? Well, allow me to post the following table for you, which shows the *absolute* number of runs each type of batted ball produced in 2006:

Event	Run Impact
Line Drive	0.391
HBP	0.355
Walk	0.355
Outfield Fly	0.192
IBB	0.075
Ground ball	0.045
Bunts	0.021
Infield Fly	-0.088
Strikeouts	-0.113

In other words, if you multiply these numbers by the average number of times each event occurred during a game last year, you'll get 4.8 runs, which was the average number of runs scored per game. If you were to follow the same math with last year's table, you'd get zero, because the numbers were calibrated around the average.

Some people like the average approach, while some like the absolute. Now you can choose between the two. And, as you can see, ground balls do produce some runs.

To remind you, I developed these tables by calculating the "run impact" of each type of hit (such as a single, double, etc.) as well as the impact of strikeouts and other kinds of outs. I then applied those "run impact" figures to the specific number of times they occur for each type of batted ball.

For 2006, I used linear weights that were derived specifically from 2006 stats (thanks to David Gassko). Technically, actual run impact values will change slightly from year to year, but this table can serve as a good overall guide to the relative value of each batting event in just about any recent year.

The 2006 Batted Ball Champs

We've got something new in our stats section this year. For the first time, we've included comprehensive batted-ball statistics for every batter and pitcher with at least 100 plate appearances or batters faced. This is information you can't get anywhere else, and we think it's pretty interesting stuff. I could spend hours just staring at them. Let me run you through some of the details.

As you now know (you did read the previous section, right?), line drives are worth more than outfield flies, which are worth more than ground balls. So, in general, batters want to hit more line drives and outfield flies, while pitchers want to induce ground balls.

Batters

For an average major league hitter, 20% of batted balls are line drives, but there were some big differences among individuals last year. Following is a list of the top 10 in most line drives hit and fewest line drives hit (as a percent of all batted balls; minimum of 502 plate appearances; figures not adjusted for ballpark).

Most Line Drives			Fewest Line Drives		
Player	**Tm**	**LD%**	**Player**	**Tm**	**LD%**
Sanchez F.	PIT	28	Inge B.	DET	14
Loretta M.	BOS	27	Bay J.	PIT	15
Kennedy A.	LAA	27	Iguchi T.	CHA	16
Young M.	TEX	25	Willingham J.	FLA	16
Mauer J.	MIN	25	Giambi J.	NYA	16
Konerko P.	CHA	25	Cedeno R.	CHN	16
Youkilis K.	BOS	24	Feliz P.	SF	16
Cabrera M.	FLA	24	Glaus T.	TOR	17
Kendall J.	OAK	24	Durham R.	SF	17
Helton T.	COL	24	Ortiz D.	BOS	17

The line drive leader was the National League batting champ, Pittsburgh's Freddy Sanchez. Also, AL batting champ Joe Mauer isn't far down the list. If you want to win a batting championship, it helps to hit line drives.

But there are some pretty good hitters on the "fewest line drives batted" list, too, such as Pittsburgh's Jason Bay, Yankee Jason Giambi, Troy Glaus of Toronto and Boston's David Ortiz. If these guys aren't hitting line drives, what are they hitting?

Well, since fly balls are the second-most valuable type of batted ball, let's draw a list of batters who hit the most and fewest fly balls, as a percentage of all batted balls. As you can see, a few players, such as Giambi and Glaus compensated for their lack of line drives by hitting lots of fly balls.

Most Fly Balls			Fewest Fly Balls		
Player	**Tm**	**FB%**	**Player**	**Tm**	**FB%**
Thomas F.	OAK	57	Jeter D.	NYA	18
Giambi J.	NYA	53	Castillo L.	MIN	21
Soriano A.	WAS	51	Pierre J.	CHN	24
Crede J.	CHA	51	Murton M.	CHN	24
Glaus T.	TOR	49	Grudzielanek M.	KC	25
Dunn A.	CIN	49	Roberts D.	SD	25
Hall B.	MIL	48	Jones J.	CHN	26
Swisher N.	OAK	48	Mauer J.	MIN	26
Barmes C.	COL	48	Kendall J.	OAK	26
Burrell P.	PHI	48	Young M.	TEX	27

Actually, the list of batters who hit the fewest fly balls is fascinating. For instance, players like Joe Mauer and the Rangers' Michael Young didn't hit many flies, but they did hit a lot of line drives so their low fly ball rate didn't hurt their productivity. But the MVP of the American League, Derek Jeter, hit the fewest fly balls of all. Admittedly, at a 22% line drive rate, he fell only a little behind the line drive leaders, but what's going on with that guy?

Let's fill out the picture by listing the players who hit the fewest and most ground balls.

Fewest Ground Balls			Most Ground Balls		
Player	**Tm**	**GB%**	**Player**	**Tm**	**GB%**
Thomas F.	OAK	24	Castillo L.	MIN	61
Dunn A.	CIN	28	Jeter D.	NYA	59
Soriano A.	WAS	29	Murton M.	CHN	58
Giambi J.	NYA	30	Jones J.	CHN	56
Youkilis K.	BOS	31	Roberts D.	SD	56
Burrell P.	PHI	31	Taveras W.	HOU	56
Crede J.	CHA	31	Pierre J.	CHN	55
Rolen S.	STL	33	Ausmus B.	HOU	53
Konerko P.	CHA	33	Berroa A.	KC	53
Swisher N.	OAK	33	Grudzielanek M.	KC	52

The fly ball and ground ball lists are almost mirror images of each other. Ground ball hitters hit fewer fly balls, and vice versa. Really, these two tables are completely redundant, but I thought you'd like to see the stats anyway.

If line drives are usually good, ground balls usually only a little good and fly balls in between, why do we see a mix of good and bad batters on all lists? The answer is that not every batter gets the same result from the same type of batted ball. In fact, there can be some big differences between them.

On the next page is a list of the number of runs generated by each batter's ground ball, based on the number of outs, strikeouts, singles, doubles, etc. that he compiled on all his ground balls. There's not really a big difference between the most extreme hitters—only one-tenth of a run overall—but some of the differences are telling.

Most Runs per Ground Ball			Fewest Runs per Grounder		
Player	**Tm**	**R**	**Player**	**Tm**	**R**
Cameron M.	SD	.11	Dunn A.	CIN	-.02
Ramirez H.	FLA	.11	Kennedy A.	LAA	-.02
Crawford C.	TB	.11	Giles B.	SD	.00
Freel R.	CIN	.10	Millar K.	BAL	.00
Suzuki I.	SEA	.10	LaRoche A.	ATL	.00
Uggla D.	FLA	.09	Giambi J.	NYA	.00
Byrnes E.	ARI	.09	Gonzalez A.	SD	.00
Betancourt Y.	SEA	.09	Chavez E.	OAK	.00
Granderson C.	DET	.09	Ausmus B.	HOU	.00
Matthews Jr. G.	TEX	.08	Jacobs M.	FLA	.01

See how important speed can be? Speedsters like Mike Cameron of the Padres and Tampa Bay's Carl Crawford get the max out of their ground balls, but fly ball hitters like Adam Dunn of the Reds and Giambi hit fly balls for a reason. They're not fast enough to produce with their ground balls.

By the way, double plays are included in ground ball run values, too. For instance, Adrian Gonzalez of San Diego tied for the National League lead with 24 GIDPs, which decreased his ground ball run value by, well, a lot.

Let's look at a couple of other personalized run value lists. For instance, the average line drive is worth .39 runs. Which players got the most out of the line drives, and which got the fewest?

Most Runs per Line Drive			Fewest Runs per Liner		
Player	**Tm**	**R**	**Player**	**Tm**	**R**
Drew J.D.	LAN	.52	Figgins C.	LAA	.31
Hall B.	MIL	.49	Hatteberg S.	CIN	.32
Matthews Jr. G.	TEX	.49	Kendall J.	OAK	.32
Sizemore G.	CLE	.49	Phillips B.	CIN	.32
Cabrera M.	FLA	.49	Cedeno R.	CHN	.32
Dye J.	CHA	.48	Taveras W.	HOU	.32
Hawpe B.	COL	.48	Ausmus B.	HOU	.32
Ramirez H.	FLA	.47	Castillo J.	PIT	.33
Holliday M.	COL	.47	Berroa A.	KC	.33
Brown E.	KC	.47	Crawford C.	TB	.33

The Dodgers' J.D. Drew garnered over half a run for every line drive he hit because only 14% of his line drives were fielded for outs (the major league average was 31%). Unfortunately (for him) his line drive frequency was 19%, slightly below the major league average.

Conversely, 7% of Jermaine Dye's line drives were home runs, tied with Travis Hafner for the highest percentage in the majors. Among other notables, 8% of Hanley Ramirez's line drives were triples and 32% of Scott Rolen's line drives were doubles. Both figures led the majors.

You may have noticed that the difference between the best and worst line drive hitters is about 0.20 runs, while the difference between the best and worst ground ball hitters is about half that. The difference between fly balls hitters is even more dramatic, so let's talk fly balls.

When you hit a fly ball, the very first important thing to do is to get it out of the infield, because 99% of infield flies are caught for outs. Here are the batters who had the fewest and most infield flies as a percent of all flies (a fly is considered an infield fly if it falls inside the basepaths):

Fewest Infield Flies per Fly Ball			Most Infield Flies per Fly ball		
Player	**Tm**	**IF/F**	**Player**	**Tm**	**IF/F**
Mauer J.	MIN	.02	Byrnes E.	ARI	.26
Jones J.	CHN	.02	Francoeur J.	ATL	.21
Gonzalez A.	SD	.02	Lopez J.	SEA	.18
Jeter D.	NYA	.02	Encarnacion J.	STL	.18
Roberts B.	BAL	.03	Betancourt Y.	SEA	.17
Giles M.	ATL	.03	Everett A.	HOU	.17
Howard R.	PHI	.03	Vizquel O.	SF	.16
Hafner T.	CLE	.04	Thomas F.	OAK	.16
LaRoche A.	ATL	.04	Hunter T.	MIN	.16
Kennedy A.	LAA	.04	Chavez E.	OAK	.16

Already, you can pick up something that differentiates some of the league's best batters such as Joe Mauer and Derek Jeter: they avoid infield flies. The list of players with the highest rate of infield flies is a mixed one, including some poor hitters (such as Everett), great hitters (Thomas) and enigmas (Francoeur).

If a player manages to get a fly ball out of the infield, it's really nice (for the batter) if it clears the outfield fence altogether. In fact, there is probably no batted-

ball stat that separates batters as much as the percentage of home runs per outfield fly. Check out the leaders and laggards:

Most Home Runs per Outfield Fly			Fewest Home Runs per Outfield Fly		
Player	**Tm**	**HR/OF**	**Player**	**Tm**	**HR/OF**
Howard R.	PHI	.39	Kendall J.	OAK	.01
Thome J.	CHA	.29	Punto N.	MIN	.01
Hafner T.	CLE	.28	Taveras W.	HOU	.01
Ortiz D.	BOS	.27	Eckstein D.	STL	.02
Berkman L.	HOU	.27	Ausmus B.	HOU	.02
Jones A.	ATL	.26	Roberts D.	SD	.02
Ramirez M.	BOS	.25	Pierre J.	CHN	.02
Pujols A.	STL	.24	Lofton K.	LAN	.02
Dye J.	CHA	.23	Vizquel O.	SF	.02
Dunn A.	CIN	.23	Loretta M.	BOS	.03

On average, 11% of outfield flies left the playing field, but a staggering 39% of Ryan Howard's outfield flies were home runs. You're probably not surprised by the other batters on these lists—they include some of the top sluggers in baseball, and some of the worst.

So when you compile the total runs produced by outfield flies, home run rate has the biggest impact. Really, the lists of most and fewest runs generated per outfield fly contain just about the same cast of characters as the homer lists…

Most Runs per Outfield Fly			Fewest Runs per Outfield Fly		
Player	**Tm**	**R**	**Player**	**Tm**	**R**
Howard R.	PHI	.58	Eckstein D.	STL	.00
Hafner T.	CLE	.47	Kendall J.	OAK	.01
Thome J.	CHA	.44	Punto N.	MIN	.04
Berkman L.	HOU	.42	Loretta M.	BOS	.05
Ortiz D.	BOS	.39	Pierre J.	CHN	.05
LaRoche A.	ATL	.39	Barmes C.	COL	.05
Ramirez M.	BOS	.38	Lofton K.	LAN	.06
Jones J.	CHN	.36	Wilson J.	PIT	.06
Beltran C.	NYN	.36	Vizquel O.	SF	.06
Dye J.	CHA	.36	Betancourt Y.	SEA	.06

Think about it: the top fly ball hitters generate half a run more than the worst for every single outfield fly they hit. Yes, there are differences between the best and worst ground ball and line drive hitters, but it is the fly ball that truly separates the best from the worst.

Well, okay, there is one other thing. There are times batters don't hit a pitch at all. For instance, Adam Dunn didn't hit the ball in 46% of his plate appearances. He struck out, he walked, he was hit by a pitch. No batted balls at all. So I ought to include those outcomes in the analysis too, don't you think?

Basically, it takes one walk to offset the damage of three strikeouts (as I showed in the run values table). Some batters use that ratio to their advantage, while others don't. Here's a table in a slightly different format: it ranks all batters according to how many runs they produced in total by striking out or walking, compared to the average batter. I've done it this way because this format captures both each batter's strikeout/walk ratio and how often he strikes out or walks in total:

Most Runs on Balls Not In Play			Fewest Runs on Balls Not in Play		
Player	**Tm**	**R**	**Player**	**Tm**	**R**
Giambi J.	NYA	24	Cedeno R.	CHN	-13
Johnson N.	WAS	23	Francoeur J.	ATL	-13
Giles B.	SD	21	Berroa A.	KC	-11
Ortiz D.	BOS	20	Feliz P.	SF	-10
Pujols A.	STL	19	Monroe C.	DET	-9
Helton T.	COL	18	Rodriguez I.	DET	-8
Hafner T.	CLE	17	Betancourt Y.	SEA	-8
Ramirez M.	BOS	16	Jones J.	CHN	-7
Hatteberg S.	CIN	15	Cano R.	NYA	-6
Beltran C.	NYN	15	Peralta J.	CLE	-6

There was almost a difference of 40 runs between the best players at controlling the plate, like Jason Giambi and Nick Johnson, and the worst, like Ronny Cedeno and Jeff Francoeur. I'm going to assume you're not surprised by the players on this list.

Now let's put it all together. Here is a table of the best major league batters last year, ranked by runs created above average. The "runs created" part of it is based on all of the batted-ball metrics I just described (basically, frequency of each batted ball times run value) and compared to the major league average. The overall

results differ a bit from other approaches (such as Lee Sinins' Runs Created Above Average), but not by a lot.

Plus, with this approach you can see something you have never seen before: how much above or below average each batter ranks for each kind of batted ball. Like so…

Total Runs Above/Below Average

Player	Tm	NIP	GB	LD	Fly	Tot
Howard R.	PHI	11	0	-1	52	61
Pujols A.	STL	19	0	5	33	58
Ortiz D.	BOS	20	-7	-4	43	52
Hafner T.	CLE	17	-7	2	38	49
Berkman L.	HOU	15	2	-3	35	48
Cabrera M.	FLA	12	4	18	10	44
Ramirez M.	BOS	16	-3	2	29	44
Beltran C.	NYN	15	0	-6	33	42
Dye J.	CHA	1	2	8	31	42
Thome J.	CHA	14	-1	-5	33	41

Most of the top batters in the majors are pure sluggers, adept at controlling the plate and blasting fly balls. They tend to not get as much from ground balls and line drives. David Ortiz may be the most extreme example, because he's definitely below average on his ground balls and line drives. Among other types of hitters, Miguel Cabrera is much more of a line drive hitter. Jermaine Dye had the most unique profile of the top 10 batters; he displayed top-notch fly ball and line drive power, but didn't control the plate any better than the average batter.

Let's look at a few unique batting profiles. Alfonso Soriano had a great year in Washington but, as you can see, he is all about the fly ball…

Player	Tm	NIP	GB	LD	Fly	Tot
Soriano A.	WAS	-2	1	-3	35	32

…and even his fine speed doesn't get him more than the major league average on ground balls. On the other hand, take a look at this fine all-around hitter…

Player	Tm	NIP	GB	LD	Fly	Tot
Utley C.	PHI	2	7	7	15	31

Philly Chase Utley can pretty much do it all, fly balls, line drives and ground balls. As pitchers give him more respect over time, that NIP figure will probably increase, too.

Here are a couple of unique American League profiles:

Player	Tm	NIP	GB	LD	Fly	Tot
Mauer J.	MIN	13	3	12	-1	28
Jeter D.	NYA	7	12	10	-1	28

Joe Mauer and Derek Jeter are average fly ball hitters; Mauer gets more out of strikeouts and walks, while Derek Jeter is an extreme ground ball hitter. Few hitters are as successful as Jeter with such a heavy ground ball approach. For instance, compare Jeter to Ichiro Suzuki:

Player	Tm	NIP	GB	LD	Fly	Tot
Suzuki I.	SEA	1	21	5	-16	10

Suzuki is the best ground ball hitter in the majors, but he didn't fill in with line drives, fly ball power and plate discipline the way Jeter did last year.

Just for fun, here are two more unique batting profiles:

Player	Tm	NIP	GB	LD	Fly	Tot
Helton T.	COL	18	-6	17	-5	23
Matthews Jr. G.	TEX	1	13	10	-3	21

Todd Helton was all about plate discipline and line drives last year—not fly balls. And Gary Matthews' big year was the result of ground balls and line drives—a rare combination.

Finally, our change of scenery award goes to Bobby Abreu, who was a very different hitter with the Yankees compared to the first half of the year spent with the Phillies.

Player	Tm	NIP	GB	LD	Fly	Tot
Abreu B.	PHI	16	0	2	-5	13
Abreu B.	NYA	3	0	2	6	10

In 438 plate appearances with the Phillies, Abreu worked walks but was average, at best, at everything else. He was particularly below average with fly balls. Once he went to New York, however, he became more powerful and created six fly ball runs above average in only 248 plate appearances.

Pitchers

Want to do the same thing for pitchers? Of course, you do, but let's start at the most fundamental level for pitchers: strikeouts and walks. Here are the run value leaders and laggards for balls not in play last year (minimum of 502 batters faced):

Fewest Runs on Balls Not In Play			Most Runs on Balls Not In Play		
Player	**Tm**	**NIP**	**Player**	**Tm**	**NIP**
Santana J.	MIN	-22	Cabrera D.	BAL	11
Schilling C.	BOS	-21	Trachsel S.	NYN	9
Mussina M.	NYA	-16	Marquis J.	STL	9
Harang A.	CIN	-16	Zito B.	OAK	9
Oswalt R.	HOU	-16	Maholm P.	PIT	8
Smoltz J.	ATL	-15	Fossum C.	TB	8
Carpenter C.	STL	-14	Wright J.	SF	8
Webb B.	ARI	-14	Davis D.	MIL	7
Haren D.	OAK	-13	Zambrano C.	CHN	7
Halladay R.	TOR	-13	Marshall S.	CHN	7

The difference between the best strikeout/walk pitchers (Minnesota's Johan Santana) and the worst (Baltimore's Daniel Cabrera) is slightly less than that between the best and worst strikeout/walk batters, 33 vs. 37 runs. For all the attention spent on pitcher strikeout/walk ratios, the K/BB ratio of major league hitters can vary even more.

In another *Annual* article, John Burnson shows that the outcome of a plate appearance tends to vary more with the batter than the pitcher, and we've uncovered the same dynamic with our strikeout and walk stats. Wondering about batted balls? Well, here are the pitching leader and laggards in percent of batted balls that are ground balls:

Most Ground Balls			Fewest Ground Balls		
Player	**Tm**	**GB%**	**Player**	**Tm**	**GB%**
Lowe D.	LAN	67	Young C.	SD	25
Webb B.	ARI	66	James C.	ATL	28
Wang C.	NYA	63	Milton E.	CIN	31
Westbrook J.	CLE	61	Lee C.	CLE	33

Most Ground Balls			Fewest Ground Balls		
Player	**Tm**	**GB%**	**Player**	**Tm**	**GB%**
Wright J.	SF	58	Williams W.	SD	36
Cook A.	COL	58	Cain M.	SF	36
Hernandez F.	SEA	58	Martinez P.	NYN	36
Hudson T.	ATL	58	Lowry N.	SF	36
Halladay R.	TOR	57	Schmidt J.	SF	37
Saarloos K.	OAK	54	Moyer J.	SEA	37

Thanks to Derek Lowe and Brandon Webb, there is a slightly larger difference between the most and least extreme ground ball pitchers. However, just like the batters, there are some pretty good pitchers among those who gave up the fewest ground balls. The two guys at the top, Chris Young and Chuck James, had pretty good years for San Diego and Atlanta, respectively. Let's see what else we can find out about them.

Here are the pitchers who gave up the most and fewest line drives, as a percent of their total batted balls:

Fewest Line Drives			Most Line Drives		
Player	**Tm**	**LD%**	**Player**	**Tm**	**LD%**
Lowe D.	LAN	16	Byrd P.	CLE	24
Johnson R.	NYA	16	Wright J.	NYA	24
Cain M.	SF	16	Pineiro J.	SEA	23
Contreras J.	CHA	16	Glavine T.	NYN	23
Wakefield T.	BOS	16	Shields J.	TB	23
Zito B.	OAK	17	Maddux G.	CHN	23
Hensley C.	SD	17	Verlander J.	DET	23
Wang C.	NYA	17	Hernandez R.	KC	23
Marquis J.	STL	17	Kim B.	COL	23
Webb B.	ARI	17	Suppan J.	STL	23

See, the good thing about being a ground ball pitcher is that you don't give up as many line drives and fly balls. The top ground ballers, Lowe, Webb and Chien-Ming Wang, are on the "least line drives" list…

Fewest Fly balls			Most Fly balls		
Player	**Tm**	**FB%**	**Player**	**Tm**	**FB%**
Webb B.	ARI	16	Young C.	SD	56
Lowe D.	LAN	17	James C.	ATL	53
Wang C.	NYA	20	Milton E.	CIN	50
Westbrook J.	CLE	22	Lee C.	CLE	48
Halladay R.	TOR	22	Cain M.	SF	48
Wright J.	SF	23	Lowry N.	SF	45
Cook A.	COL	24	Zito B.		45
Hudson T.	ATL	24	Martinez P.		44
Hernandez F.	SEA	25	Santana E.	LAA	44
Saarloos K.	OAK	25	Wakefield T.		44

…and they are also the leaders on the "fewest fly balls" list. Conversely, take a look at the pitchers who have given up the most fly balls. Yeah, it's guys like Young and James, who gave up the fewest ground balls. As I said before, when it comes to frequency, you really only need to know one batted-ball stat: ground ball percentage.

But do pitchers give up the same number of runs on each type of batted ball? Well, here's a look at runs given up per ground ball:

Fewest Runs per Ground ball			Most Runs per Ground ball		
Player	**Tm**	**R**	**Player**	**Tm**	**R**
Lidle C.	PHI	.00	Lee C.	CLE	.11
Glavine T.	NYN	.00	Madson R.	PHI	.11
Rogers K.	DET	.00	Kazmir S.	TB	.10
Halladay R.	TOR	.00	Byrd P.	CLE	.09
Robertson N.	DET	.00	Hernandez L.	WAS	.08
Suppan J.	STL	.01	Vargas C.	ARI	.08
Bush D.	MIL	.01	Milton E.	CIN	.08
Peavy J.	SD	.01	Snell I.	PIT	.07
Hensley C.	SD	.02	Harang A.	CIN	.06
Cook A.	COL	.02	Lopez R.	BAL	.06

Yes, pitchers do differ about as much (though slightly less) than batters in ground ball run values. This is actually a little surprising to me, because pitchers have better and worse infields behind them, while batters hit to all sorts of different infields. In fact, you can pick out some of the best infields on this list: there are two Detroit and San Diego pitchers on the "fewest runs" list, and two Cincinnati and Cleveland pitchers on the "most runs" list. That's partially a reflection of the quality of their respective infields.

Next up is the line drive list:

Fewest Runs per Line Drive			Most Runs per Line Drive		
Player	**Tm**	**R**	**Player**	**Tm**	**R**
Santana E.	LAA	.30	Lidle C.	PHI	.46
Fossum C.	TB	.31	Francis J.	COL	.46
Lee C.	CLE	.31	Nolasco C.	FLA	.46
Burnett A.	TOR	.32	Moehler B.	FLA	.46
Benson K.	BAL	.33	Schilling C.	BOS	.45
Mussina M.	NYA	.33	Bush D.	MIL	.45
Beckett J.	BOS	.33	Myers B.	PHI	.44
Arroyo B.	CIN	.34	Capuano C.	MIL	.44
Lowry N.	SF	.34	Davis D.	MIL	.44
James C.	ATL	.34	Meche G.	SEA	.44

Once again, the difference between the fewest and most runs per line drive is greater for batters than pitchers (.21 runs vs. .16). Simply put, line drive hitting (both in terms of frequency and getting hits out of line drives) changes more with batters than pitchers.

At this stage, I'm going to do you a favor and skip the infield flies and home run per outfield fly table. Let's go straight to those pitchers who gave up the fewest runs per outfield fly and those who gave up the most (which is, again, primarily driven by home run rates):

Least Runs per Outfield Fly			Most Runs per Outfield Fly		
Player	**Tm**	**R**	**Player**	**Tm**	**R**
Wright J.	NYA	.09	Saarloos K.	OAK	.34
Cabrera D.	BAL	.12	Lidle C.	PHI	.31
Schmidt J.	SF	.12	Pettitte A.	HOU	.29
Francis J.	COL	.12	Hernandez F.	SEA	.28
Lackey J.	LAA	.13	Silva C.	MIN	.28
Loewen A.	BAL	.13	Burnett A.	TOR	.28
Escobar K.	LAA	.13	Webb B.	ARI	.27
Bedard E.	BAL	.14	Snell I.	PIT	.27
Blanton J.	OAK	.15	Beckett J.	BOS	.26
Cain M.	SF	.15	Santos V.	PIT	.25

The difference in home run rates between Jaret Wright and Kirk Saarloos isn't nearly as great as that between Ryan Howard and David Eckstein. For pitchers, the best way to keep your home run rate down is to induce ground balls, because pitchers' rates of home runs per outfield fly tend to be much closer to the overall average.

And now we'll put it all together again. Here is the same "runs created vs. average" table, but this time the leaders are those pitchers who have given up fewer runs than average.

Total Runs Above/Below Average

Player	Tm	NIP	GB	LD	Fly	Tot
Santana J.	MIN	-22	-9	0	-8	-39
Carpenter C.	STL	-14	-7	-8	-5	-35
Webb B.	ARI	-14	-1	-5	-13	-32
Halladay R.	TOR	-13	-9	1	-8	-30
Mussina M.	NYA	-16	-1	-11	4	-25
Zambrano C.	CHN	7	-5	-16	-11	-24
Sabathia C.	CLE	-13	3	-5	-10	-24
Lackey J.	LAA	-6	5	-9	-14	-24
Lowe D.	LAN	-5	5	-8	-14	-23
Smoltz J.	ATL	-15	0	-6	-2	-22
Oswalt R.	HOU	-16	-4	2	-5	-22

There is a less consistent pattern among pitching leaders than batting leaders. This list includes strikeout pitchers (Santana), pitchers who "controlled" line drives (Carlos Zambrano and Mike Mussina) and those who "controlled" fly balls (John Lackey and Derek Lowe). The key to Lackey's performance was his home run rate (6% of outfield flies) while Lowe's key was his ground ball rate, as we have seen.

And what about those two youngsters with the low ground ball rates, Chris Young and Chuck James?

Player	Tm	NIP	GB	LD	Fly	Tot
Young C.	SD	-3	-8	-10	1	-20
James C.	ATL	1	-4	-5	2	-5

Young and James followed similar patterns, though James was less so. Neither one was terribly hurt by his high fly ball rates, primarily because 79% of their outfield flies were caught for outs vs. the major league average of 74%. Outfield defense?

Speaking of defense, look at the ground ball performance of these two Tigers hurlers:

Player	Tm	NIP	GB	LD	Fly	Tot
Rogers K.	DET	2	-12	3	-5	-12
Robertson N.	DET	-1	-12	1	7	-5

Both Kenny Rogers and Nate Robertson had relatively high ground ball rates last year. Usually, a higher rate will produce more runs, but these two Tigers hurlers actually yielded fewer overall runs on ground balls thanks to their fine infielders.

And our change of scenery award goes to Greg Maddux, who was traded from Chicago to Los Angeles at the trading deadline, with these results…

Player	Tm	NIP	GB	LD	Fly	Tot
Maddux G.	LAN	-3	-4	0	-2	-10
Maddux G.	CHN	-9	-5	12	-1	-3

In 572 plate appearances with the Cubs, Maddux was slightly better than average. But in just 290 plate appearances with the Dodgers, Maddux was outstanding. The difference was almost entirely in the line drive column. Fielders, skill, ballpark or just plain luck? Hmm.

As I said, these stats are available in our stats section for all players with at least 100 plate appearances/batters faced. I hope you enjoy them as much as we have.