History of the Fielding

History of the Fielding, Part 1 **2**
Basic Batting Performance 2
Basic Fielding Performance 3
Opportunity Space of a Batter 4
Opportunity Space of a Fielder 5
Enter Statcast 6
Intercept Model 6
Time After Time 7
Recap 7
History of the Fielding, Part 2 **8**
Catch Probability 8
Turn down the noise 9
Turn up the signal 10
Weighted opportunities 12
To Distance/Time, and beyond 13
History of the Fielding, Part 3 **14**
The Play 14
Probability Distributions 16
The Process 17
The Results 18
The Leaders 19
Fielder Roles 20
Future Iterations 21
References 21

History of the Fielding, Part 1

We have been very successful in measuring batter performance, and somewhat successful in measuring pitching performance. So why is it so difficult to measure fielding performance? Before I answer that question, or at least provide an answer to that question, let's take a few steps back and talk about batting.

Basic Batting Performance

Suppose we wanted to measure "Reaching Base Safely" performance. What would you do? Maybe we'd just start off with guys who Reached Base Safely the most. Seems obvious right? In that case, we'd get this top 5:

H+BB+HB	Player
292	Alex Bregman
276	Mookie Betts
276	Marcus Semien
272	Carlos Santana
269	Freddie Freeman

Not coincidentally, every one of those players played in at least 150 games. Mike Trout reached base 263 times while playing only 134 games, and misses out being in the top 5. Ok, so maybe the next step is to look at Reaching Base Safely Per Game. If we do that, we get this list:

Per Game	Player
1.96	Mike Trout
1.92	Christian Yelich
1.87	Alex Bregman
1.84	Mookie Betts
1.82	Anthony Rendon

That looks more promising. But you know, the higher up in the batting lineup the more opportunities you have to reach base, each game. **Opportunities** will be key, so remember that, as we'll circle back to it.

Anyway, so instead of per game, how about per plate appearance? If we do that however, we get a bunch of guys who came up to bat once and reached base. We need to have some sort of cutoff. I'll use 300 plate appearances, which gives us 273 batters, or just over 9 per team. When I do that, we get this list:

Per PA	Player
0.438	Mike Trout
0.429	Christian Yelich
0.423	Alex Bregman
0.412	Yordan Alvarez
0.412	Anthony Rendon

Alvarez joins the fun at the expense of Betts, who drops all the way to 14th. Betts came to bat 4.70 times per game, compared to Alvarez who came to bat only 4.24 times per game. So in terms of opportunity, the plate appearance makes more sense than the game.

What this shows is something akin to On Base Percentage (OBP). The official rule of OBP has the same numerator I noted (hits, walks, hit batters), but the denominator is slightly different (with sac hits not included, basically). In any case, the key point is that we put the successes in the numerator (times reaching base safely) and the opportunities in the denominator (plate appearances).

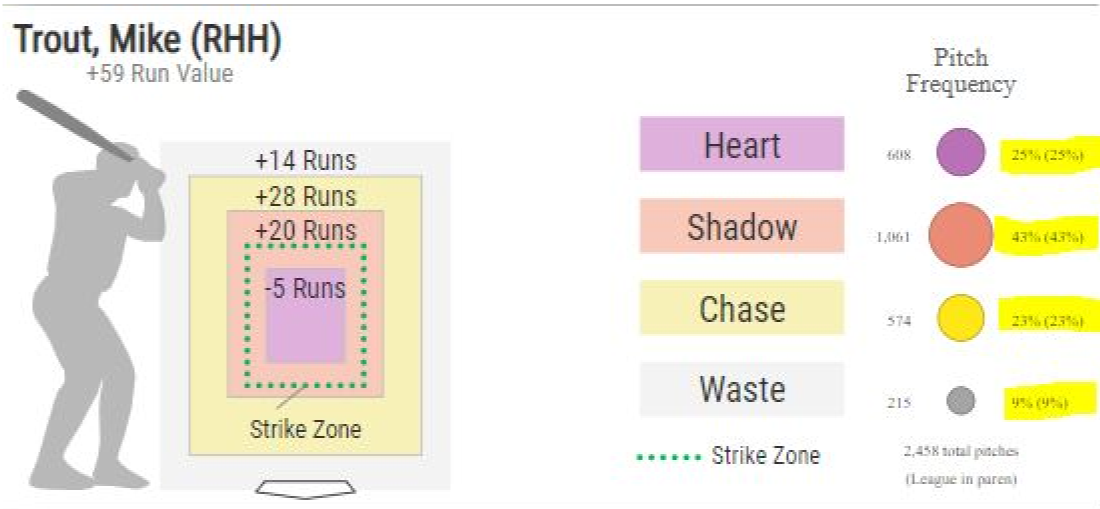
Basic Fielding Performance

Let's follow the similar logic for fielding. We want to measure fielding performance, and so, we need to count the successes. In the official stats, we have putouts and assists. So, we have our first problem: the league leaders in putouts are all first basemen and catchers. That's largely as a result of the other infielders making the initial play, while the first baseman catches the ball standing. Or, the pitcher getting a strikeout, and the catcher catches the ball squatting. In either case, the putout is largely influenced by the other infielders or pitcher.

So, maybe we look at assists only. When we do that, the list is all shortstop and secondbasemen, with the top third baseman, Arenado, appearing at 18th place. While for batting, most of the times reaching base is because of a strong batter influence, for fielding, it's quite the opposite: most assists are on fairly routine plays. And shortstops and second basemen especially face the most routine plays.

Opportunity Space of a Batter

The issue therefore is the *opportunity space*. This for example is how often Mike Trout sees a pitch in the various regions around the plate (what we call the Attack Regions). The league average is in parenthesis.



See the yellow highlights on the far right? Mike Trout sees 25% of his pitches in the Heart of the Plate. The league average batter sees 25% of his pitches in the Heart of the Plate. Mike Trout sees 43% of his pitches in The Shadow Zone, that region that straddles the edge of the strike zone. The league average batter also sees 43% of his pitches in The Shadow Zone. Mike Trout sees 23% in the Chase Region. So does the league average batter. Mike Trout sees 9% of his pitches in the Waste Area. As does the league average batter.

In other words, Mike Trout's opportunity space to reach base is similar to the opportunity space of the league average batter. And this is true, even though him being at the plate would have some influence on the pitches he sees, based just on his talent. To the extent that it does, it doesn't really matter. The pitches that Trout sees in and around the strike zone is pretty much the same pitches than any batter sees in and around the strike zone. The strike zone is a strong focal point.

Opportunity Space of a Fielder

For fielders, things are much different, as the playing field is very vast. First, we have the general areas on the field where you would stand, the shortstop, the center fielder, the right fielder, and so on. So, fielders have the extra complication because of their zones of responsibility.

In addition to that, the fielders are on the receiving end of the batter-pitcher matchup. Whereas a Verlander-Trout matchup will generally provide a similar opportunity set as say an average pitcher facing an average batter, the Astros shortstop (Carlos Correa principally) won't face a similar opportunity set as the Angels shortstop (Andrelton Simmons principally). Indeed, Carlos Correa with Verlander as his pitcher won't be the same opportunity set as Carlos Correa with each of the other Astros pitchers on his team. And whatever Correa happens to see on average won't be the same as whatever Simmons happens to see on average.

This is why Zone Ratings became the next step in fielding metrics. Every batted ball was either assigned to a specific fielder, or, that batted ball was assigned to NO fielder at all, a no-man's-land batted ball. So, this gave us the denominator that we were looking for, or at least, a denominator that was better than using games or innings. In addition, the denominator was set to the fielder who either fielded the ball (the first touch fielder), or the fielder who was assigned responsibility for the base hit that went through his zone of responsibility.

This was, in essence, the flip side of OBP: outs made per responsible ball. At the team level, Bill James called this the Defensive Efficiency Record (DER). At the individual level, the data providers called this Zone Rating (ZR). The next step was to improve that denominator by **weighting** the opportunities.

Rather than treat every batted ball in a fielder's zone of responsibility equally, each batted ball was instead given a certain weight based on where in the zone the ball travelled. Balls that were closer to where a fielder normally stands was given a high presumed out value (say 95% expected outs made for a presumed routine play), while balls that were much farther from a fielder's normal standing spot was given a low out value (say 20% expected outs made for a ball presumably almost out of reach). So, based on the location of the batted ball, each batted ball was given a value from 0% to 100% of expected outs made. Comparing these expected outs to the actual outs made gave us the Extra Outs Made. On Fangraphs, this is called Ultimate Zone Rating (UZR, a nod to ZR), and Defensive Runs Saved (DRS).

Enter Statcast

Notice that I said “where a fielder **normally** stands”. Prior to 2015, we never tracked the actual location of the fielder. With Statcast starting in 2015, we have. And so now, we can take the next step and instead of treating every batted ball that went to a particular zone equally, we can now compare that batted ball to where the responsible fielder is actually standing.

In other words, if we have a ball that was hit half-way between the normal SS spot and the normal 3B spot, ZR would count that as a no-mans-land ball (and remove it from the opportunity set of either fielder), UZR would give it some low expected out value. But, if we had a RHH at the plate and the shift was on, the SS would actually be standing right in line for that **particular** batted ball: what might in one system be considered a no-man’s-land batted ball (and given a presumed out rate of 0%), in the Statcast system, it would be considered routine (and given a presumed out rate of 95% or higher). Now we can separate two batted balls hit to the same spot, based on the proximity of the fielder of when the play started (and we’d define the start of the play at pitch release).

Intercept Model

Great, right? Yes, but still not enough. Why? Because there are two central things that affect the chance of a play being converted into an out: the proximity of the fielder’s starting location to the eventual path of the batted ball, **and** how much time the fielder has to reach that spot. After all, a batted ball that goes half way between the SS and 3B in 1 second is quite different (*) from getting there in 2 seconds (**). And if it gets there in 4 seconds (***), the batter will have almost reached first base (****).
(*) might skip past the infielders for little chance of an out
(**) enough time to get to the ball for a great chance of an out
(***) plenty of time to get there
(****) for almost no chance of an out

So, we have the Opportunity Distance, which is how much distance the fielder has to cover. And the Opportunity Time, which is how much time the fielder has to cover. And the two gives us the fielder’s Opportunity Space. And that point when fielder and ball cross is called the Intercept Point. In the event the ball and fielder don’t meet, we determine the Intercept Point as the point when the distance of the ball to the plate and the starting distance of the fielder to the plate are equal.

Finally, we have what we want. Now we can move forward. And this model is the Distance-Time Intercept Model, the point at which the fielder and ball cross or would have crossed paths.

Time After Time

Note that so far, all we've talked about is the fielder reaching a spot on the field. Once he has the ball, if it's a ground ball, he needs to make the force play, usually at first base. So, in addition to establishing the Intercept Point, we need to establish how far away the batter is to reaching first base. And since we know how fast each runner is, knowing how far away he is in *feet*, we can also estimate how far away he is in *seconds*. As for the fielder, we know the Intercept Point, and so, we also know how long the throw is in feet. And that means, we can estimate how much time in seconds the throw will take from the Intercept Point.

With an estimate of the time for the throw and with an estimate of the time for the run, it comes down to a simple comparison: if the time for the throw is longer than the time for the run, the runner is safe; if the time for the throw is shorter than the time for the run, the runner will be out. Sometimes, the two times are very close. And in those cases, rather than estimating a pure safe/out, we'd have a sliding scale from 1% to 99%: if the two times are identical, we'd estimate the Outs Made at about 50%, and the more there is a gap in the estimated time, the more our estimate will deviate from that 50% point.

Recap

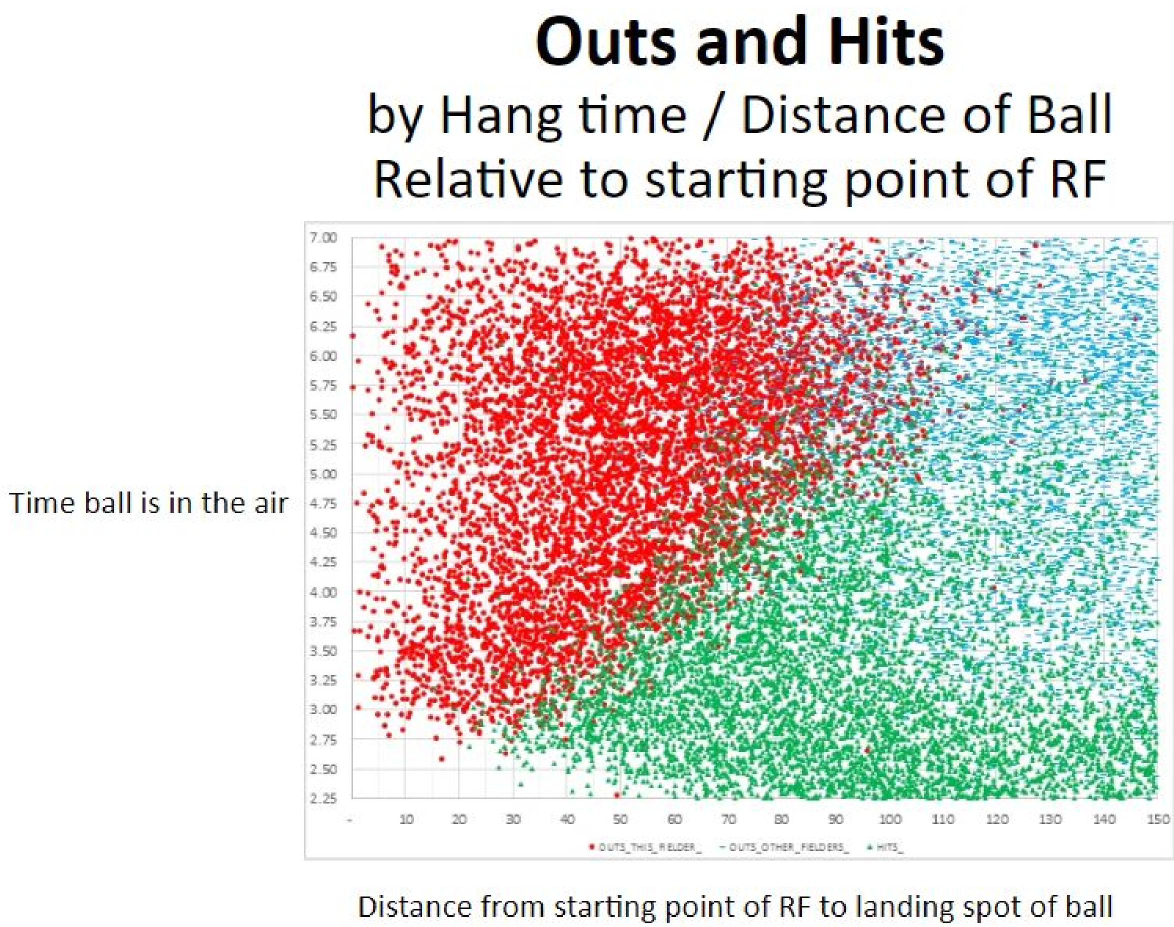
So there you go, this is the basis for Infield Defense Outs Above Average: we apply an Intercept Model based on distance and time. And the reason this model will resonate is because it is *intuitive*. When you watch a play develop, you are inherently making those calculations already. You are taking a snapshot of the play, and you are taking an educated guess as to whether the runner will beat the play or not. You may not be calculating the actual distance and actual time, but neither do you do that when you cross a busy street. Your experience lets you make an excellent guess. All we are doing with the Intercept Model is applying math and probability at the Intercept Point, and answering the questions that we are intuitively asking.

In Part 2, I'll describe how the math works for Catch Probability for outfielders. And in Part 3, I'll lay out how the math works out for the infield.

History of the Fielding, Part 2

There’s going to be charts. Before we talk about Infield Defense, let’s talk about Outfield Defense. Back in 2016, we rolled out Catch Probability. And this is how it was introduced.

Catch Probability



Along the horizontal, left to right, is the distance of the right fielder to the landing spot of the ball. Along the vertical, bottom to top, is the amount of time the ball is in the air.

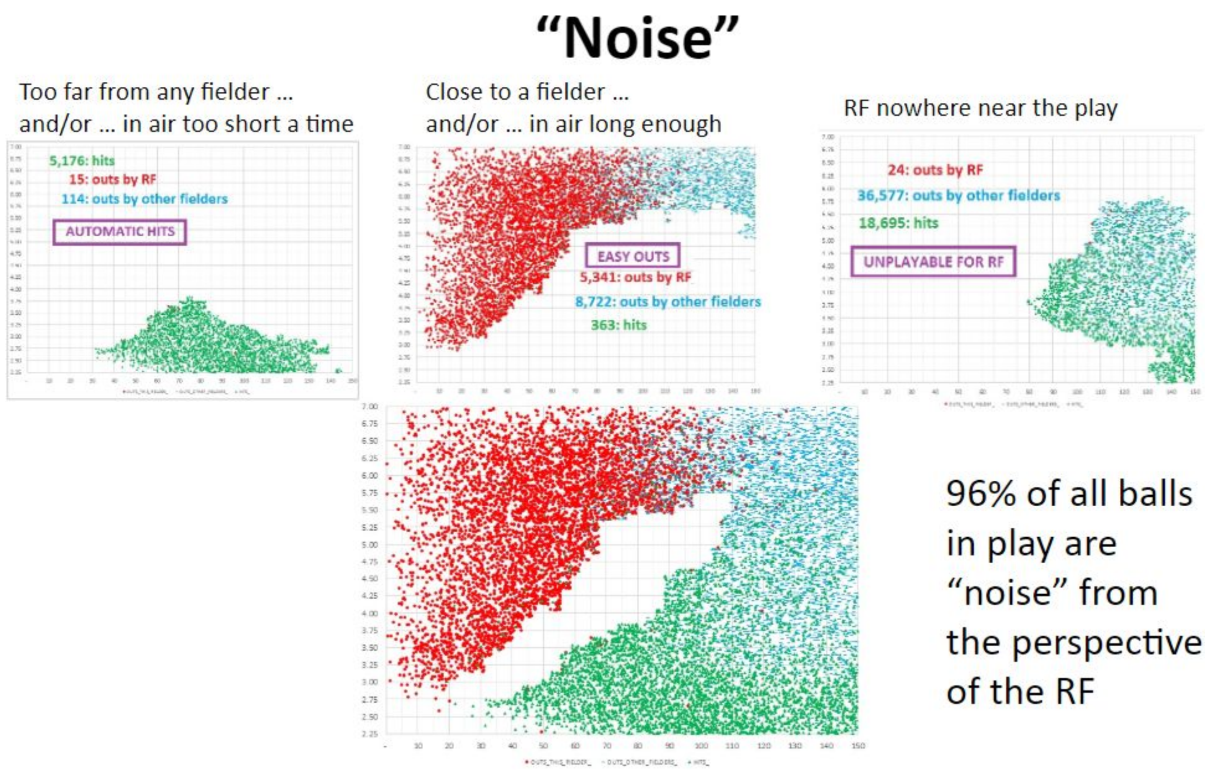
Those dots in red are all the outs made by this fielder, the RF. As you can imagine, when he has little distance to cover, he’ll make the out. Or when he has alot of time to make the play, he will get the out. But what happens when those two are a bit at odds?

Those green dots represents all the base hits. So, any combination of alot of distance to cover and a short amount of time to cover it will result in a base hit.

Those blue dashes represents outs made by other fielders, plays that are simply too far for the RF to make, and much closer to the CF to make.

Turn down the noise

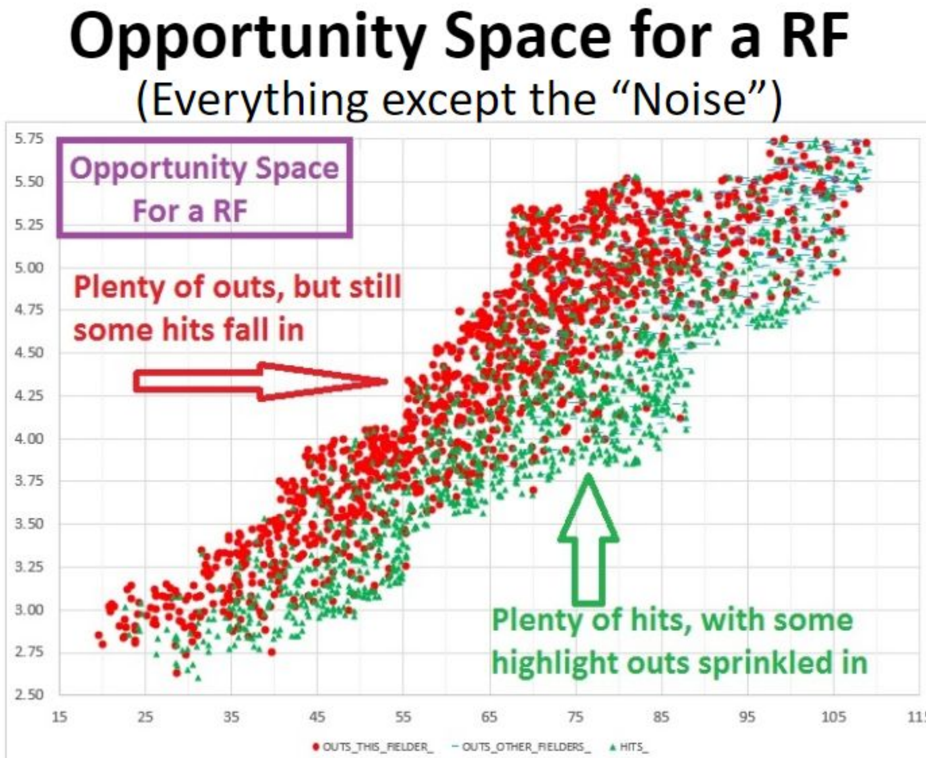
We can break this up more clearly this way. First, all the routine hits or routine outs, what is essentially "noise" in the data. This is plays that don't reflect the skill of the fielder involved.



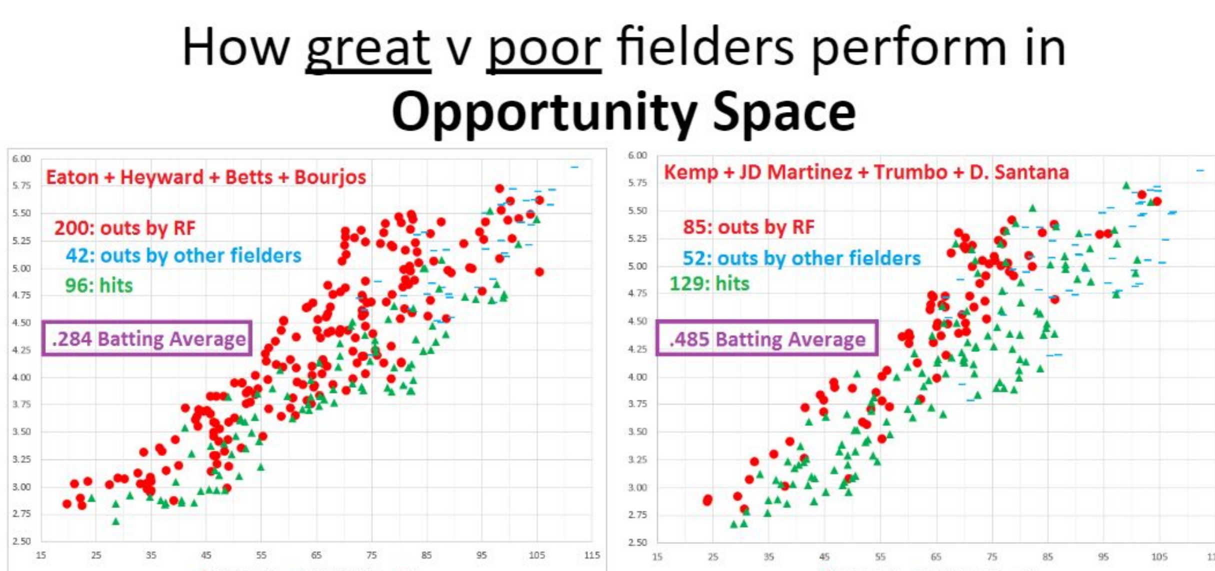
We can see that these plays are either automatic hits, easy outs, or simply unplayable for our fielder. And these plays represent 96% of all balls in play. What's left is the 4% of the plays that are close enough to the outfielder (in terms of distance and time) that puts his skill on display. It's that gap you see above, where there's an overlap in hits and outs. These plays:

Turn up the signal

Remaining 4% of balls in play is our area of interest.



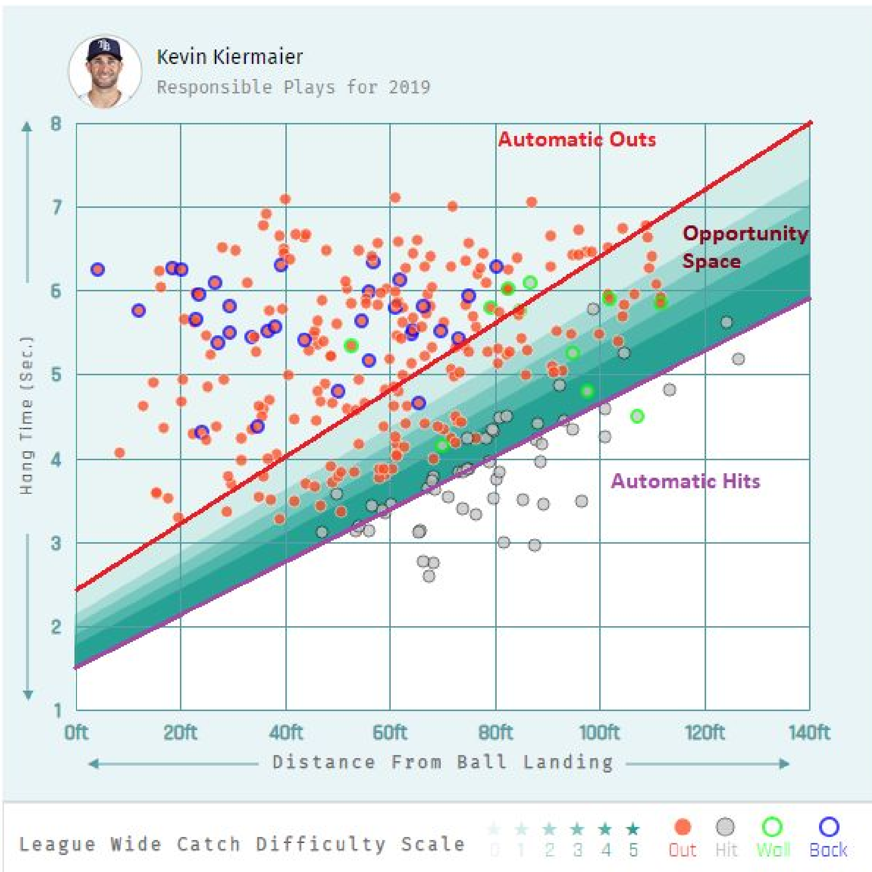
And here we see what is happening: at some combination of distance versus time, the fielder's skill comes into play. Indeed that line above is fairly straight. And a straight line of a Distance v Time chart, the slope, represents constant Speed.



To make it even clearer, we can see how great fielders and poor fielders perform in the **same** Opportunity Space. This is the key to Catch Probability: that we are able to remove all the data that represents noise, those plays that require almost no skill, that are automatic hits or automatic outs. These are plays that we want to remove from our numerator *and* denominator.

Once we do that, what we are left with is a set of plays that look similar to all the fielders: they all require some level of skill, and they are presented to the fielder at a somewhat similar level of difficulty: neither very hard, nor very easy to make. This is the opportunity space of our fielder. This is what counts in the denominator.

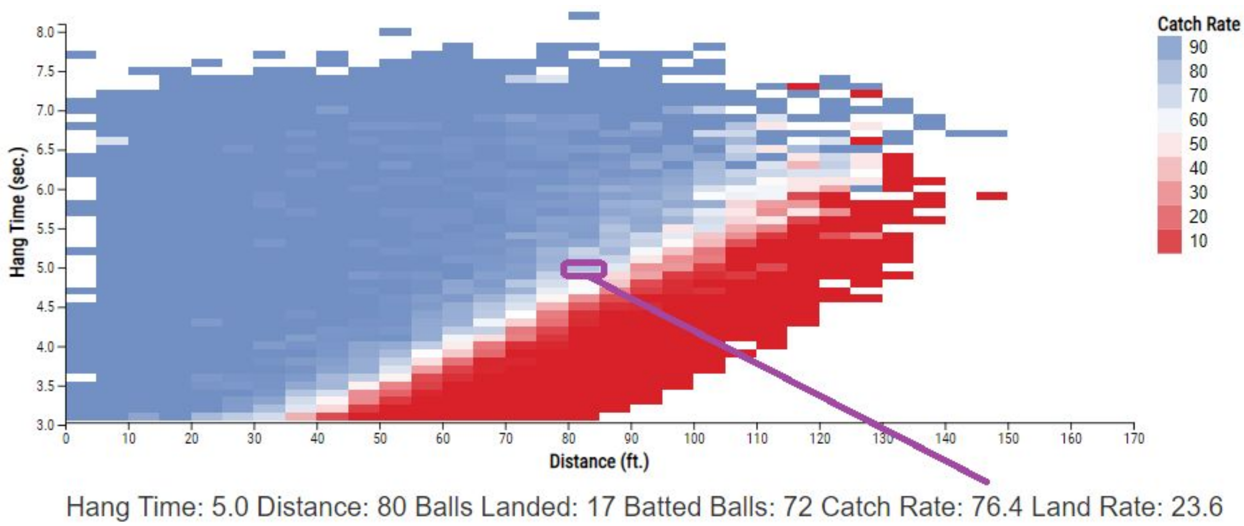
And here we see that our great fielders made 200 outs, with 96 hits dropping in. (And 42 outs made by a teammate.) As for our poor fielders, they made only 85 outs, with 129 hits dropping in. (And 52 outs made by a teammate.) Added all up, in the same opportunity space, the good fielders allowed a .284 batting average, while the poor fielders allowed a .485 batting average.



This breakdown also allows us to show how a particular fielder performs. This is Gold Glover Kevin Kiermaier. You can see that in the section marked *Automatic Outs*, every play was either an out (labelled in orange), or it hit the wall (circled in green, with a gray dot). And in the section marked *Automatic Hits*, every play was a hit (labelled in gray). All of these plays are noise. The remaining plays are his Opportunity Space, the plays that require some level of effort to turn into an out. And the vast majority of these plays were in fact turned into outs. Those that landed for hits are those that were very close to the Automatic Hit threshold. In other words, except for two base hits, every Hit that landed was either impossible, or close to impossible to catch.

Weighted opportunities

In Catch Probability, we weight those plays even further, which you can see here: https://baseballsavant.mlb.com/statcast_catch_probability



Here we see that balls where the outfielder had to cover 80 feet in 5 seconds were caught 76% of the time. What we do is create a smoothed out model that adheres to a simple principle: the more distance you have to cover, the lower the catch rate (so in the above chart, it would be more red). And the more time you have to cover, the higher the catch rate (in the above chart, it would appear more blue). And in those 50/50 plays, we'd see it in white in the above chart. And if you go back to earlier in this section, you will note that the above white, light blue and light red regions, essentially a straight line slope, is akin to the Opportunity Space.

To Distance/Time, and beyond

While the core for Catch Probability is a Distance/Time math model, there are two additional parameters that we use:

- Is the fielder running back?
- Is the wall an impediment to the play?

As you'd expect, if the fielder runs back, the catch probability goes down. And if the wall is an impediment, the catch probability also goes down. We can see the impact these parameters have, when we focus on a sample set of plays, which the chart below is 80-90 feet and 4.5-5.0 seconds:

dist	time	resp_fielder_is_back	resp_fielder_is_wall	catch_rate	actual_outs	n	diff
80	4.5	0	0	54%	55%	1101	1%
80	4.5	1	0	19%	17%	192	-2%
80	4.5	0	1	11%	8%	131	-3%
80	4.5	1	1	4%	3%	73	-1%

To read the first line: we have 1101 plays since 2016 where the fielder had to cover 80 to 90 feet in 4.5 to 5.0 seconds, where he did not have to run back, nor was the wall an impediment. The Estimated Catch Probability was 54%, while the actual catch rate under those conditions was 55%. The last line shows that the outfielder had to run back and that the wall was an impediment. Under those conditions, they caught the ball 3% of the time, compared to an estimated 4%.

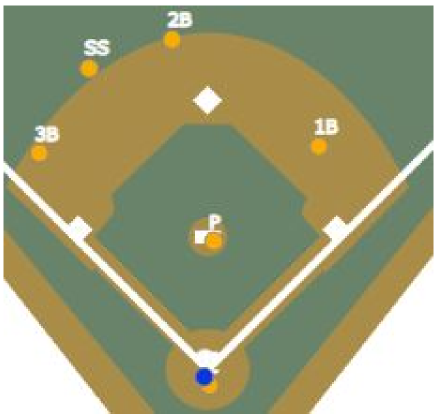
All these parameters taken together forms the Catch Probability Distance-Time model. Now that we know how it works in the outfield, we can describe how it works in the infield, using the same core idea, but with additional considerations.

History of the Fielding, Part 3

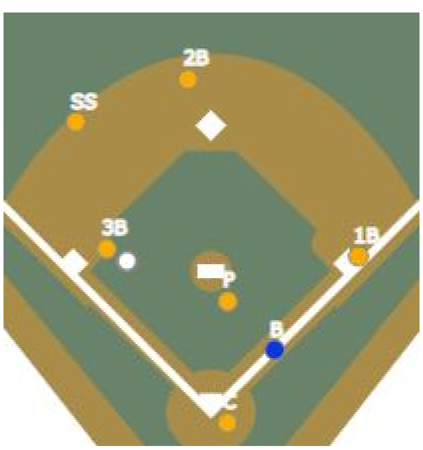
Infield Defense finally joins Outfield Catch Probability. Here's how it works, along with some math.

The Play

On July 28, 2019, Brian Anderson of the Marlins faced this Diamondbacks shifted fielding alignment, with all three infielders on the pull side playing somewhat deep.

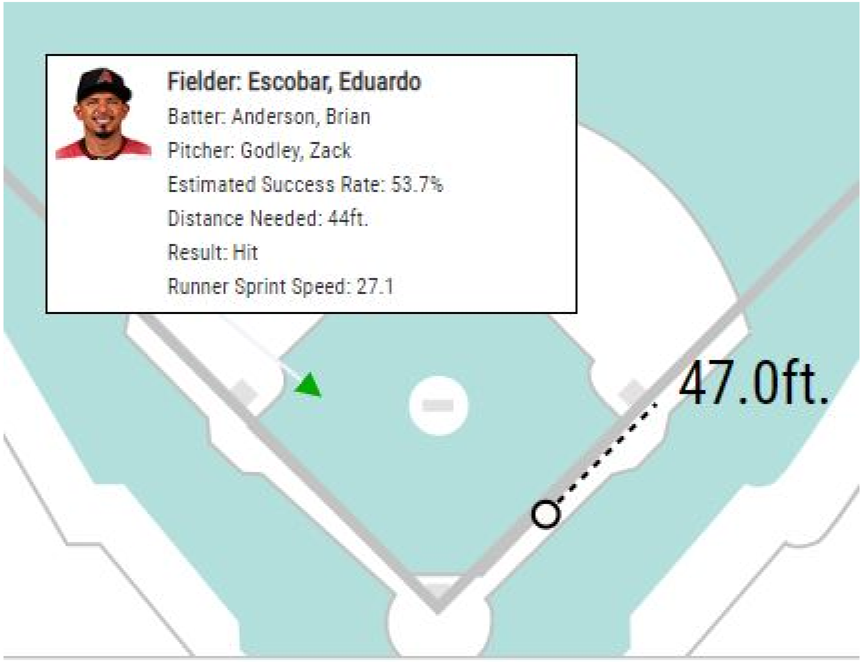


Anderson hit a slow roller toward the third baseman Escobar, who had to charge in. The Intercept Point, the point at which the fielder and the ball crossed paths, took 3.2 seconds from pitch release. Escobar had to cover 44 feet. At the Intercept Point, Escobar was 107 feet to first base. With a speed of a bit over 100 feet per second, the thrown ball will be in the air for roughly 1 second. It will also take about ¾ of a second for the fielder to release the ball. All told, Escobar would have 1.7 to 1.8 seconds from the Intercept Point to get the ball to first base (¾ of a second to get rid of it, and 1 second for the ball traveling through the air).



The batter at the Intercept Point was 47 feet from reaching first base. At his Sprint Speed of 27.1 feet per second, we'd estimate he had 1.7 to 1.8 seconds to reach first base (47 feet divided by 27.1 feet/second is 1.73 seconds; though he was still not on a dead run by that point, so it's an estimated 1.77 seconds).

Here's what the Intercept Point data looks like:



You will note that with the fielder having 1.7 to 1.8 seconds to make the play and the runner having 1.7 to 1.8 seconds to reach first base, both starting from the Intercept Point, then the chances of the fielder getting the out is around 50/50. In the above chart, our estimate is 53.7%.

Fast forward 1.7 to 1.8 seconds, and this is what happened.

The runner was initially ruled out, and the call was overturned to safe on a replay challenge. In other words, an expected 50/50 play based on the *model* was in fact a 50/50 play based on *reality*.

And that's all we are trying to do when we create a model: how can we represent reality? And we represent reality by modeling how we process the play. And we process the play by *implicitly* measuring everything in terms of distance and time. And so the model will *explicitly* measure or estimate all the touch points in distance and time. Welcome to the Distance/Time Intercept Model for Infield Defense.

Probability Distributions

So, how do we arrive at 53.7%? It's actually the product of three separate estimates. As I mentioned, we have a few touch points to consider, what you can consider Action Events:

- The first Action Event is the fielder arriving at the Intercept Point. In the above play, with the ball being hit in line with the fielder, there was essentially a 100% chance that the fielder would arrive at the Intercept Point. This isn't always the case, especially if you think of base hits that go through the SS/3B hole. The range is naturally from 0% to 100%. This play is an example of 100%.

- The second Action Event is the fielder retrieving the ball cleanly. In this play, it was fairly routine, and so, we estimate the ball being picked up at 97.5%. In tougher plays, such as a sharp liner directly at a fielder, he may have a near 100% chance of **arriving** at the Intercept Point, but he may have an 80% of **retrieving** the ball (think of deflections or simply muffing the ball)

- The third Action Event is the race: given that the fielder has the ball in his hands for a throw, how often will he beat the runner to the bag.
  - In the conditions represented by this play, the typical fielder will get the ball to the first baseman in 1.72 seconds. It won't be exactly 1.72 seconds. It'll be a distribution centered around 1.72 seconds, which you can consider to be +/- 0.25 seconds. He might get off a very quick release and very strong throw, or he might double-pump or lob it.
  - A runner with the given speed and distance to the bag of this play will get there in 1.77 seconds. It won't be exactly 1.77 seconds. It'll be a distribution centered around 1.77 seconds, which you can consider to be +/- 0.25 seconds. He might kick it up an extra gear, or he may have a misstep.
  - So we have two distributions of races denoted in time, and we come up with a probability that the fielder will beat the runner given those two distributions. And in this case, we estimate the fielder will beat the runner, *given* that he has ball in hand, 55.1% of the time.

- With all three Action Events probabilities estimated, we simply multiply them: the chance of arriving at the ball; then given that he has arrived at the ball, the chance of retrieving the ball; then given that he's retrieved the ball, the chance he'll win the race to the bag. In Probability Distribution parlance, that's respectively: 100% x 97.5% x 55.1%, which gives us 53.7%. And that 53.7% gives us our estimated success rate, or the Out Probability, on this play.

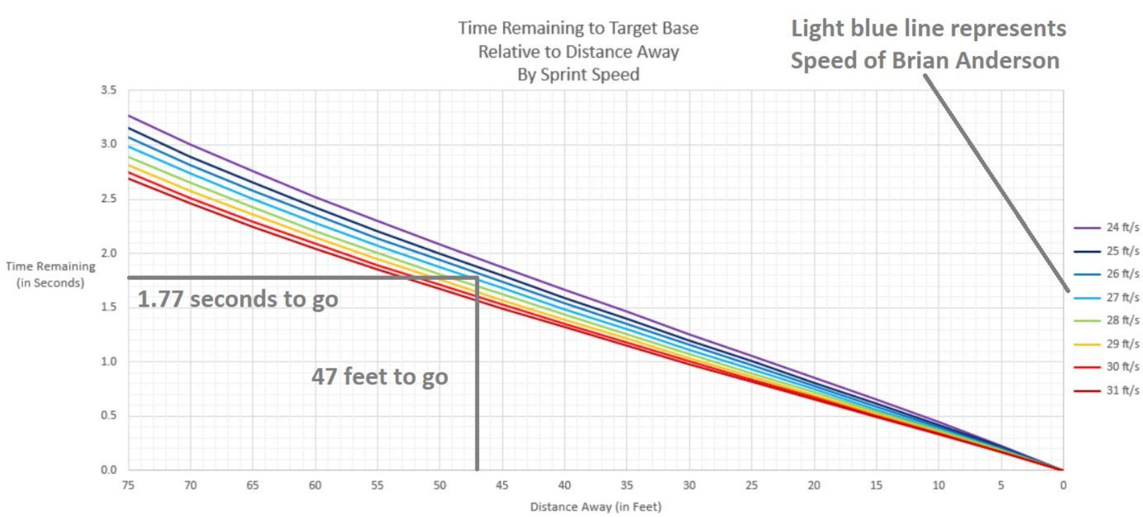
The Process

For every infield play, we establish the Intercept Point, as well as the responsible fielder's initial starting point. This allows us to determine how much distance and time the fielder has to cover. In the above play, there was a 100% chance the fielder would reach the ball, since it was a slow roller hit right toward him. But in plays where the ball is hit in the hole and roll toward the outfield, there is anywhere from a 0 to 100% chance of the fielder reaching the ball. Therefore, our first step is to determine the probability of the fielder arriving at the Intercept Point in time to pick up the ball.

We need to know how long it takes for a ball to reach the Intercept Point (on its potential path toward the outfield), as well as how far the fielder's starting point is to the Intercept Point. Knowing the distance of the infielder to the Intercept Point in *feet*, we now know his distance in *time*. And since we know how long the ball took to reach the Intercept Point, we simply compare the two times. When the two times are really close, we come up with a probability number. But when the ball skips through too fast (0%), or it's slow enough for the fielder to reach it, like the Escobar play (100%), then those are binary.

Having established the Intercept Point, we need to know if the fielder would reach it on the fly, or it was at least a one hopper. Then we need to freeze the play at the Intercept Point, and establish the remaining distances of the fielder and runner in feet (which we then convert into seconds). And we make the comparison of the two competing races (ball being thrown v runner running) in the common unit of seconds.

For the runner, we can see the *to go,* distance and time, graphically for all speeds of runners at every point along the baseline. Noted in grey is Brian Anderson on the play from the preceding section.



There are additional wrinkles, such as a first baseman fielding a ball and being able to run to the bag. As well as sharp liners that don't need to be caught for the out, since the fielder still has a chance for a force play after knocking the ball down.

The Results

So every play tracked is assigned an Estimated Success Rate, or an Out Probability. This is how it looks when we group all the tracked plays into one of four ranges of Out Probability:

Estimated Group	Estimated Success Rate	Actual Outs Per Play
0-19	3%	6%
20-79	61%	64%
80-89	87%	84%
90-99	95%	95%

The second line reads: for plays where the Out Probability was between 20% and 79%, the average of those plays was estimated to become an out 61% of the time. And in reality, those plays were turned into an out 64% of the time.

In order to get an out *value* on an individual play, it's straightforward:

- if you make the out, then your value-added is the difference between 100% and the out probability. So, if the out probability was 60%, and you made the out, you get 40%, or +0.40 outs.
- If you don't make the out, then you have a negative value, and this time, it's the difference between 0% and the out probability, or simply the negative of the out probability. Using the same above example of a 60% out probability, if you don't make the out, then you get negative 60%, or -0.60 outs.

We add up all these partial pluses and minuses. And the total of these becomes your Outs Above Average. (This is the same process followed for Outs Above Average for the Outfield.)

The Leaders

Naturally, what we care about is showing all this at the individual fielder level. Here's how the top seven infielders look. Javier Baez has a total Outs Above Average (OAA) of +19 outs. This means that adding up all his partial pluses and subtracting all his partial minuses left him with an overall total of +19. And that was good enough to lead the league.

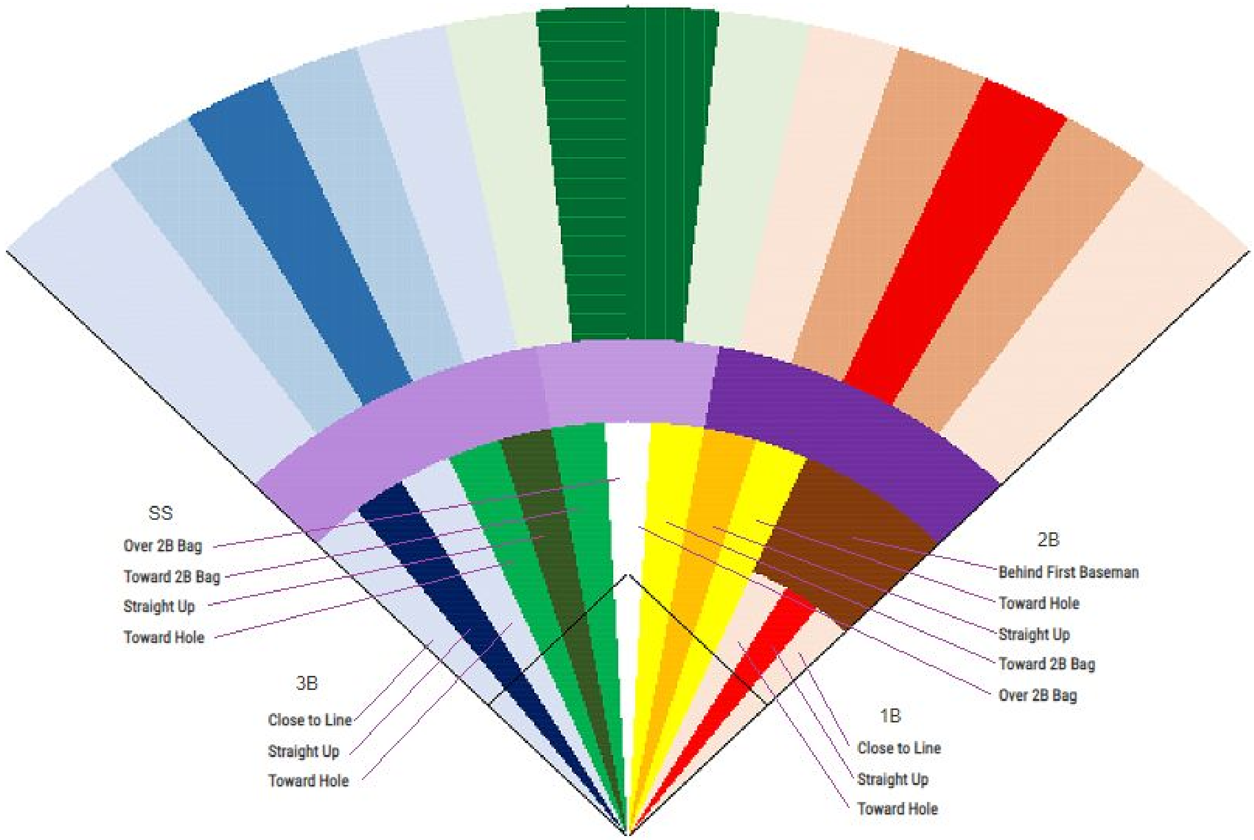
The chart below shows additional breakdowns. For Baez for example, the Estimated Success Rate of his average play was only 83%, indicating that he faced a bit tougher set of plays than his peers. You can see the other shortstops on the leaderboard were all at 87% or 88%. As a result of his great fielding, he was able to convert 88% of his responsible plays into outs. While that number is lower than his peers Story, Ahmed and DeJong (90% to 91% for each), when you compare their actual results to their individual Opportunity Space, Baez is ahead. Note that Simmons converted an astounding 93% of his responsible plays into an out, compared to his Opportunity Space value of 87%. However, because he missed alot of playing time, his overall OAA was "only" +16. Had he played as often as Baez, he might have led the league.

Rk.	Player	Team	Year	Pos	OAA	Outs Above Average: In	Lateral toward 3B	Lateral toward 1B	Back	RHB	LHB	Success Rate	Estimated Success Rate	Success Rate Added
1	Baez, Javier		2019	SS	19	7	4	7	1	13	6	88%	83%	5%
2	Arenado, Nolan		2019	3B	17	0	4	12	1	12	5	89%	85%	4%
3	Simmons, Andrelton		2019	SS	16	7	1	8	0	14	3	93%	87%	6%
4	Ahmed, Nick		2019	SS	16	10	0	5	2	8	8	90%	87%	3%
5	Story, Trevor		2019	SS	15	9	-1	7	0	14	1	90%	87%	3%
6	Chapman, Matt		2019	3B	14	8	1	4	1	13	1	89%	85%	4%
7	DeJong, Paul		2019	SS	13	6	-5	10	1	8	4	91%	88%	3%

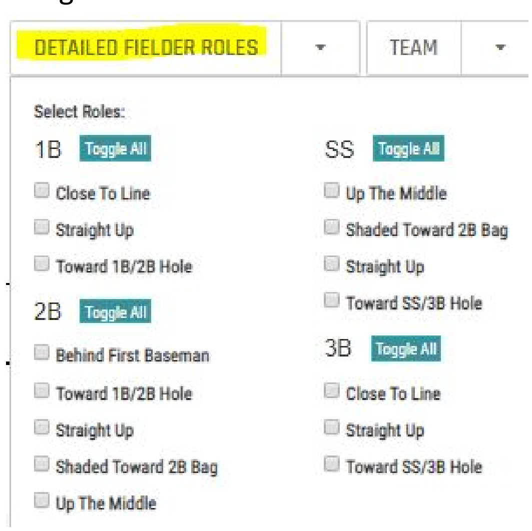
In addition, by having every play tagged, we can do additional breakdowns. Above you can see a breakdown based on direction: charging in, running back, or moving laterally toward the 3B or 1B sides.

Fielder Roles

What helps make the system work is the focus on Roles rather than Positions. In this new world order of Shifting, we don't want to compare SS Baez standing in short right field to other shortstops. But rather, we want to compare players based on where they are standing. If (Official) 3B Chapman stands in a (traditionally) SS slice, we compare Chapman to other fielders standing in the same slice. So as to not confuse *locations* from official *positions*, we've designated these locations or slices as Fielder Roles. The field is split into slices like so:



And when you go to Baseball Savant to slice and dice all the tracked plays, you will be able to do so based on where they are standing:



Future Iterations

As you might expect, Infielders pose a more complex challenge than Outfielders. So this is our first iteration, as we continue to improve and develop the model. This iteration only looks at the “first out” of a play. So, double plays are not considered (beyond the first out), but will be in the next iteration.

Another kind of play that was not considered for the initial iteration, but became obvious once we saw the results was: pitcher failing to cover first base. There were a few plays that we had tagged at 90%+ outs, but that the first baseman failed to convert into an out. And a good portion of these we watched visually was that the pitcher failed to cover the base. We do not currently have those plays tagged, but they will be handled in the next iteration.

In addition, our upgraded tracking system is going to offer more data, and so, we’ll be able to do even more. This first iteration is a leap forward for Infield Defense, and each future iteration will build from there. We’re excited to finally bring this to the public, and we can’t wait to bring even more.

If you think of assists, putout, and errors as the first inning of fielding metrics, Zone Rating as the second inning, and UZR/DRS as the third inning, we’re now in the fourth inning. We’ve come a long way, but we’ve got plenty of room to develop more.

References

https://www.mlb.com/news/statcast-introduces-outs-above-average-for-infield-defense

http://baseballsavant.mlb.com/leaderboard/infield_outs_above_average

http://baseballsavant.mlb.com/visuals/statcast-infield-defense?playerId=595879&year=2019&result=&direction=&normalize=undefined&roles=&esrGT=0&esrLT=1&evGT=0&evLT=125&distGT=0&distLT=200&batside=&viz=box-start