

Abandon Rates: Modeling How Often an Outfielder Gives Up on a Play

At Statcast, our goal is to make every metric and model as intuitive as possible - we want the models we create to reflect what we're seeing on the field. At times, however, the metrics don't pass the eye-test, and this causes us to dive deeper into our models and investigate the disconnect. One of these times happened recently with our [catch probability](#) model when a team member watched what looked like an easy play, but had a catch probability of only 40%. The rest of the team agreed it looked far easier than 40%, so they found a sample of 50 plays with similar parameters (distance needed, opportunity time) and calculated the catch rates. Sure enough, 20 of the 50 plays were outs – a catch rate of 40%. How did it look so easy? What's causing the difference between the optics and the numbers?

After watching the 50 plays, they noticed that the outfielder effectively gives up, or “abandons”, half of them. The outfielder decides to let the ball drop and play it off the bounce rather than go for the out. Trout demonstrates an abandoned play below:

http://tangotiger.net/files/abandon_trout.mp4

So, in the 50-play sample, the outfielder has 25 “abandon plays”, which means he has 25 “try plays”. 20 of these plays were outs, so the catch rate *given he's trying* is 20/25, or 80%. A catch probability of 80%, rather than 40%, passes the eye-test for that play. We've found the disconnect between the optics of a play and its probability – abandon rates introduce a significant bias to catch rates.

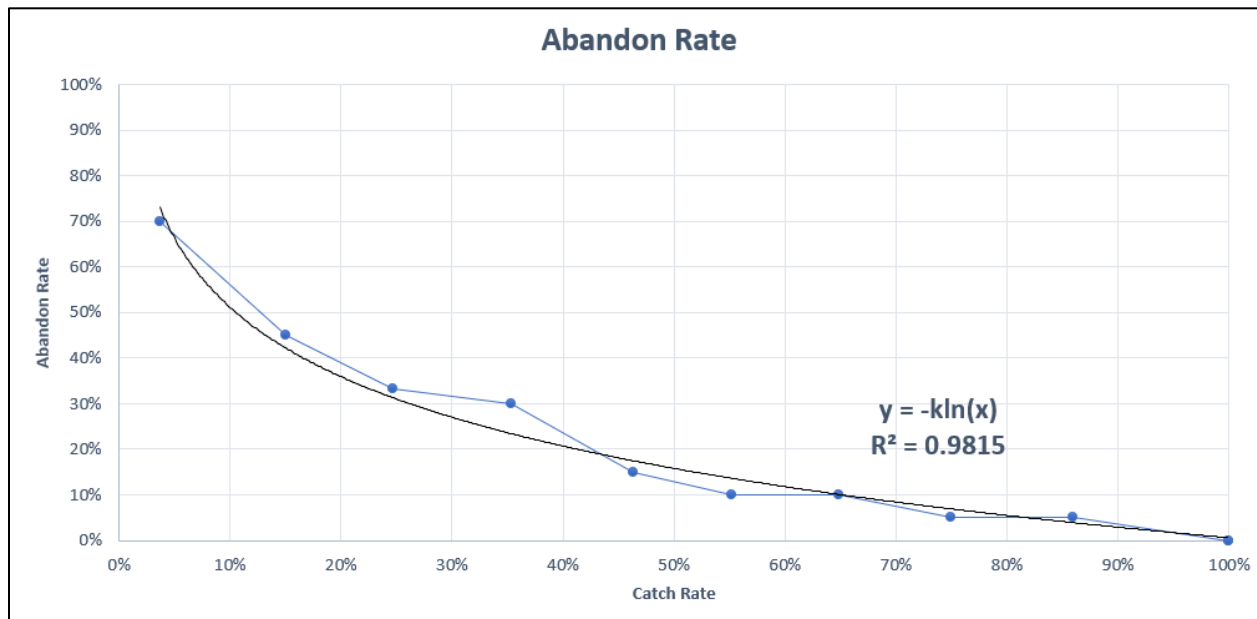
To better understand when outfielders give up on plays and how much of an impact it has on overall catch rates, I built an Abandon Rate model to predict how often a certain type of play will be abandoned. Because we don't have a data field indicating whether a play was abandoned, the only way to collect this information was to roll-up my sleeves and watch the plays myself. As a proof of concept, I watched a sample of 220 plays, about 20 per 10% catch-rate band. I split up no out/out plays by what we would expect based on the catch probability buckets. For example, in the 50%-60% bucket, we would expect at most 60% of the 20 plays to be outs, so I looked at 12 out plays and 8 no out plays.

Bucket	No Out	Out	Total
0%-10%	18	2	20
10%-20%	16	4	20
20%-30%	21	9	30
30%-40%	18	12	30
40%-50%	10	10	20
50%-60%	8	12	20
60%-70%	6	14	20
70%-80%	4	16	20
80%-90%	2	18	20
90%-100%	0	20	20
Grand Total	103	117	220

Once I had my sample set of plays, I watched each one and indicated if the play was abandoned using a binary code (Abandon = 1, Try = 0). I then calculated the abandon rate in each bucket. As expected, abandon rates fall as catches become easier.

Catch Probability Bucket	Abandoned Plays	Total Plays	Abandon Rate
0%-10%	14	20	70%
10%-20%	9	20	45%
20%-30%	10	30	33%
30%-40%	9	30	30%
40%-50%	3	20	15%
50%-60%	2	20	10%
60%-70%	2	20	10%
70%-80%	1	20	5%
80%-90%	1	20	5%
90%-100%	0	20	0%

I plotted the average catch rates in each bucket against the abandon rate for that bucket (so that I had a numerical catch rate input to the model rather than a categorical bucket). The data fit very well to a logarithmic model, proving the relationship between the catch probability of a play and the outfielder's decision to abandon.

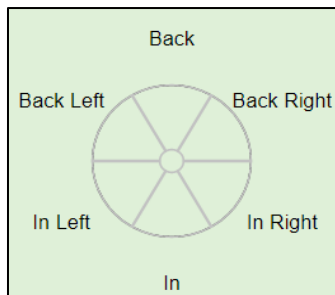


Using this model, I could predict the catch probability *given the outfielder is trying* by estimating an Abandon Rate which leads to the Try Rate (1-Abandon Rate) and calculating the conditional catch probability on try plays as $\frac{\text{Catch Probability}}{\text{Try Rate}}$. The Try Rates have a larger effect on the harder plays as

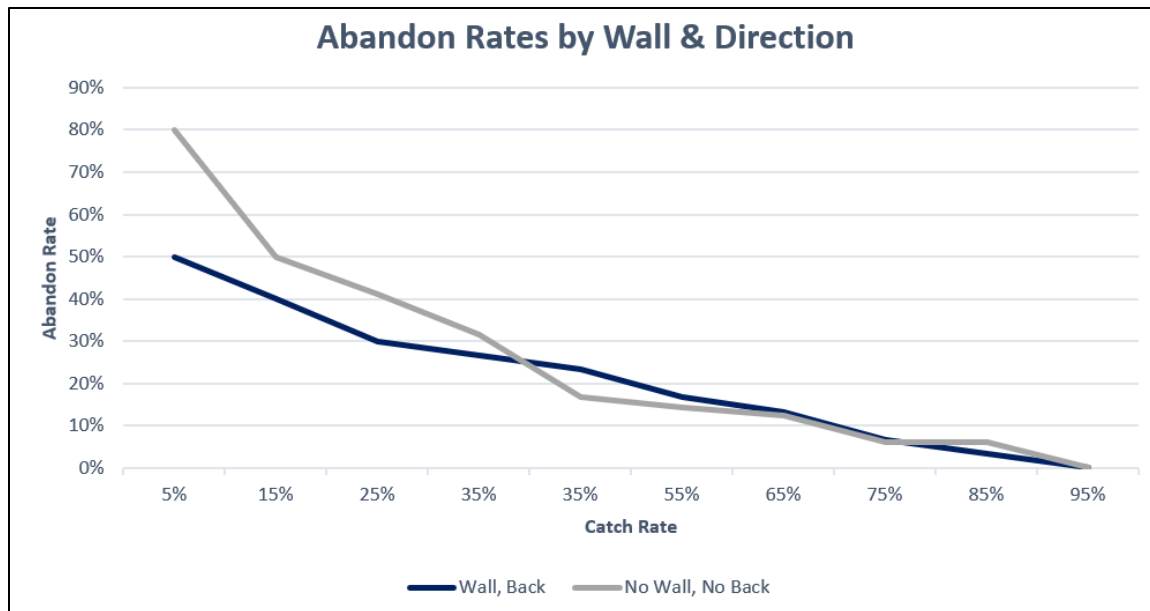
expected, increasing their catch probability on try plays about 10%. The table below answers the question *given the outfielder is going for it, what is the probability he makes the catch?*

Catch Probability Bucket	Abandon Rate	Try Rate	Catch Probability (Given Try)
0%-10%	80%	20%	0%-20%
10%-20%	50%	50%	20%-31%
20%-30%	41%	59%	31%-41%
30%-40%	32%	68%	41%-50%
40%-50%	17%	83%	50%-59%
50%-60%	14%	86%	59%-68%
60%-70%	13%	88%	68%-76%
70%-80%	6%	94%	76%-84%
80%-90%	6%	94%	84%-92%
90%-100%	0%	100%	92%-100%

Now that I had a general model, the next steps were to subset the data into the two factors which have the highest effect on outfield plays: [the impediment of the wall and direction](#). Specifically, the direction of “back” as shown in the diagram below:



Our catch probability models are adjusted for these factors, so the abandon model should be as well. I wanted to see the impact the wall and direction had on abandon rates, so I compared No Wall/No Back plays to Wall/Back plays. I had about 140 No Wall, No Back plays already from the 220-play sample I started with, and I watched over 300 Wall, Back plays, following the same methodology as my proof of concept model. Both subsets fit well to the logarithmic model $Abandon\ Rate = -k \ln(catch\ rate)$. The comparison of No Wall/No Back to Wall/Back plays led to an interesting insight on the behavior of outfielders.



On Wall/Back plays, outfielders actually abandon less than No Wall/No Back plays up until about 40% catch probability. After that point, the outfielder's approach is about the same. This can be explained with the risk-reward dynamic of abandoned plays. For Wall/Back plays, because the ball is already well-struck to a deep part of the outfield, it's likely going to produce at least a double, even if the outfielder plays it well off the bounce. Thus, the outfielder may as well try for the out since they don't have much to lose. On the other hand, with No Wall/No Back plays, it's more likely to result in a single, and the throw to the infield is shorter than from the wall. Therefore, going for the out and missing would turn a single into a double or worse, so the outfielder has more to risk. Thus, they take the safer approach of abandoning more often in these harder plays.

Using the Abandon Rate model, we can better understand the decision-making process of outfielders and reconcile the visual ease of a play and the estimated catch probability. For example, we can explain why this play looks easy, despite having a 35% catch probability:

http://tangotiger.net/files/rockies_CF.mp4

For No Wall/No Back plays with this catch rate, the outfielder will abandon about 30% of the time. Thus, when the outfielders are trying, they have closer to a 50% catch probability, aligning with the visuals of the play.

This model has power to answer many questions about outfield plays. It already predicts how often a play will be abandoned and is used to calculate catch probabilities given the outfielder is trying. The next question I'm interested in using this model to answer is *should* the outfielder abandon a play to gain further insight into outfield defense.