An Introduction to FIELDf/x

by Kate McSurley and Greg Rybarczyk

Sportvision first arrived on the baseball scene with the pitch tracking PITCHf/x system, which allows viewers to identify the speed and trajectory of each pitched ball during a game. The technology moved beyond the limitations of the radar gun and allows fans a more robust knowledge of the ball's overall activity before and at the plate. It also provides a digital record of the pitch sequences for future evaluation and analysis.

By the end of the 2007 season, the system was installed in nearly every major league ballpark, and it was clearly becoming an important element of the baseball experience. Baseball enthusiasts were finally able to access supportive data to discuss the full breadth of a pitcher's performance, as well as exciting new offensive data for every major leaguegame.

The PITCHf/x system continues to be well received by fans, but teams and coaches also recognize the value of objective data like accurate release points and movement trajectories when evaluating their players. Now minor leagues and baseball schools rely on Sportvision's system for informed roster decisions and strategies for player development. A coach's raw intuition can be supported with hard, comprehensive data from the PITCHf/x system.

In just a few seasons, the baseball community came to expect PITCHf/x data for every pitch and every game, but it also became clear that there was still a lot of untracked action left on the field. The natural next step in the tracking process was to record data from the other side of the pitch. HITf/x tracks and digitally records data for the contact point, speed off the bat, elevation angle and field direction.

Still, objective defensive data in baseball has always been somewhat of an enigma, or at least an unattainable dimension of the game. Sportvision had to respond with something just as telling, but, what's more, just as accurate as the PITCHf/x and HITf/x systems. This season Sportvision did respond with continued testing of the FIELDf/x tracking system in a few major league parks, and the baseball community is anxious to see how this new data will once again change the way we rank and evaluate player performance.

Soon people will be able to quantify new baseball statistics and unveil intriguing information about players. The FIELDf/x system identifies individual players and tracks player positions at a high enough resolution to understand granular movements like reaction times. This tracking technology reveals gripping new fielding data that may allow an explanation for the unexplainable. It is one thing to say that Derek Jeter is a great shortstop, but with FIELDf/x, his range can be directly compared with other players for an impartial assessment. It can also shed new light on fielding errors, which are traditionally subjective yet equally unyielding.

With FIELDf/x, we will be able to determine the percentage likelihood that a play was even possible to add validity to a called error. Sportvision's tracking technology can determine a third baseman's reaction time to a hit ball and even the path a base runner took to his base. The system also captures the who, when and where of ball events like pitches, hits, catches and throws for an all-inclusive record of the events on the field.

This new tracking system will certainly allow for fresh commentary and interesting player comparisons, but Greg Rybarczyk of Hit Tracker has identified one of the first pertinent applications for the FIELDf/x system. In the following section he explains the significance of being able to calculate defensive metrics with some interesting examples, and he alludes to ways the baseball community will be able to interact with this data.

A FIELDf/x Application: True Defensive Range

by Greg Rybarczyk

With the introduction of technology that permits the accurate tracking of the baseball and players throughout the entire play, across the entire field, Sportvision has created an opportunity to revolutionize defensive metrics by allowing the calculation of True Defensive Range, or TDR. This is a metric that will go beyond the current zone-based, event-counting methods, using continuous measurements of parameters such as fielder starting position, batted ball hang time and landing point to generate the most accurate, meaningful and transparent defensive metric ever.

Why we Should Move Beyond Zones

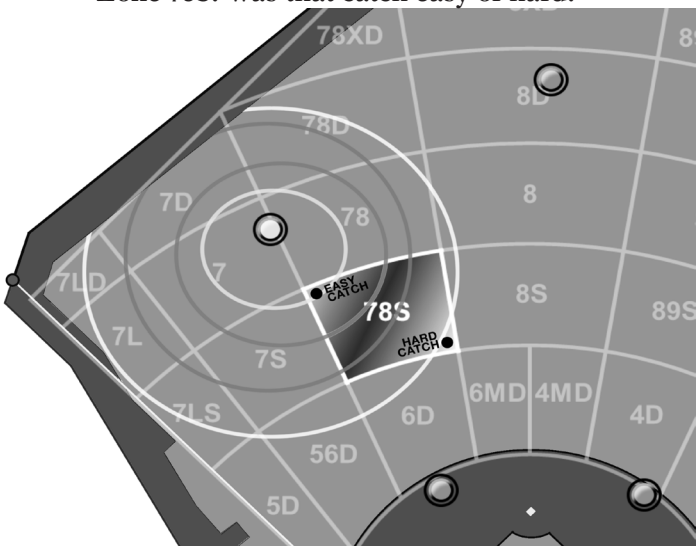
Many contemporary defensive metrics use "zones," discrete areas on the baseball field where all balls landing there are grouped and counted together. One prominent example of the use of zones is the Project Retrosheet Scoring System, which divides the fair territory on a standard baseball field

into 55 zones. The average zone in the Retrosheet system covers just under 2,000 square feet, equivalent to an area 40 by 50 feet, but there is much variation: the zones go from tiny (zone 1, the pitcher's mound) to enormous (zone 8XD in deepest center field, spanning more than 130 feet corner to corner).

The size of the typical zones used in defensive metrics is one disadvantage; any system that requires balls landing 40 or 50 feet apart (or 130!) to be considered equivalent is clearly oversimplifying things. However, there is an even more troubling characteristic of the Retrosheet zones: They are not the same in every ballpark. The standard zones are defined for a field with 330 foot foul lines and a 405 foot center field fence, but for ballparks that do not match this standard, it is unclear whether the existing zones are to be extended, or stretched, or compressed to conform to the actual home run fences.

Perhaps the most serious drawback to zone-based systems is that by their very nature, zones mix together balls that were easy for fielders to reach with balls that were difficult or impossible to reach. Retrosheet zone 78S is a good example: Nearly every ball hit to the back left corner of 78S, located a few feet from the left fielder's typical starting position, will be caught, while nearly every ball hit to the opposite side of 78S, a spot more than 100 feet from the shortstop, left fielder and center fielder, will fall safely for a hit. Because any individual ball hit to this zone (and to others like it) can be simple to field, impossible to field or anywhere in between, metrics based on this zone are particularly susceptible to bias, and without any way to distinguish the easy chances from the difficult, there is little an analyst can do about it.

Zone 78S: Was that catch easy or hard?



With so many important drawbacks, why are zones used in defensive analysis? The answer is that, for many years, no better alternatives came along. Precise tracking of the landing spots of batted balls by human eye is possible when the ball lands beyond the fence, among numerous easily located landmarks such as seats, aisles, access tunnels, etc., but it is much more difficult when the ball lands out in the middle of an unmarked expanse of green grass. The typical accuracy and precision in a "naked eye" estimate of a ball's landing point, when no landmarks are nearby, is poor. Zones came about as a way to group balls that landed near each other, upping the sample size and somewhat diminishing the importance of the poor measurement capability.

FIELDf/x is the Answer

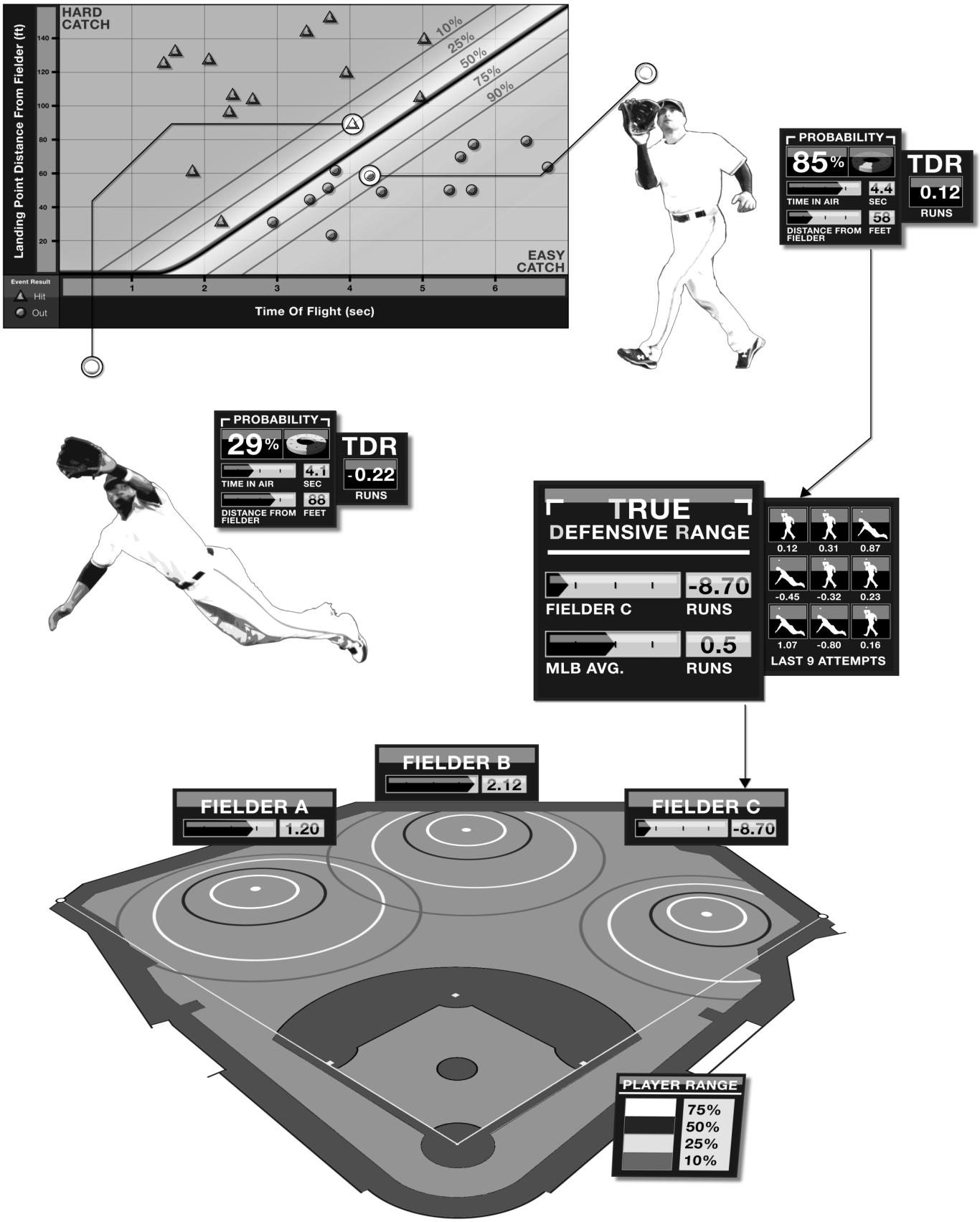
The answer to all these zone-based problems is to know where the fielders were standing when the ball was hit, and to know exactly where the ball was throughout the play. Only FIELDf/x can deliver this.

FIELDf/x uses cameras mounted high above the field to track the position of the baseball and every human being on the field for the duration of each play. Fifteen times per second, position data are captured and archived, allowing analysts and color commentators to precisely reconstruct the movement of the ball, the fielders, the runners, and even the umpires and base coaches. For every batted ball, FIELDf/x can tell where the fielder started, where the ball landed (or was caught, or where it passed by a fielder), and how long the ball took to get there. Together these three factors describe the difficulty of any particular play, and enable the calculation of what I have named True Defensive Range, or TDR.

True Defensive Range is the Question

With complete data for all batted balls, it becomes possible to determine the relative difficulty of a play by looking at the outcomes of similar plays, with that similarity being based on flight time and the distance the fielder had to cover to reach it. If only 10 percent of all balls hit a certain distance from a fielder in a certain amount of time are caught for outs, then the fielder who makes this play has done something quite unusual, and deserves to be credited highly for having done it.

On the other hand, if a fielder does not convert this play, he will not be penalized greatly, since it was a very difficult play to make. The opposite is of course true: Fielders who make a play that is successfully converted into an out 90 percent of the time receive only a very small amount of credit, while fielders who fail to convert this same play will face a substantial penalty.


HARD CATCH
EASY CATCH
10%
25%
50%
75%
90%
Landing Point Distance From Fielder (ft)
Time Of Flight (sec)
Event Result
Hit
Out
PROBABILITY
85%
TIME IN AIR
4.4
SEC
DISTANCE FROM FIELDER
58
FEET
TDR
0.12
RUNS
PROBABILITY
29%
TIME IN AIR
4.1
SEC
DISTANCE FROM FIELDER
88
FEET
TDR
-0.22
RUNS
TRUE DEFENSIVE RANGE
FIELDER C
-8.70
RUNS
MLB AVG.
0.5
RUNS
0.12
0.31
0.87
-0.45
-0.32
0.23
1.07
-0.80
0.16
LAST 9 ATTEMPTS
FIELDER A
1.20
FIELDER B
2.12
FIELDER C
-8.70
PLAYER RANGE
75%
50%
25%
10%

In the TDR system, the base amount of run credit awarded for plays is based on the area of the field where the ball is hit, and linear weights that assign a run value for each play category (single, double, triple, out, etc). Balls landing or being caught in front of an outfielder would be assigned the base value for the difference between an out and a single (-0.30 runs and 0.47 runs, respectively, in one model, for a difference of 0.77 runs), which would then be multiplied by either the league-wide out percentage for that distance/time or the quantity one minus that value. The result is the TDR for that particular play.

For example, suppose a ball is hit to shallow left field, and lands at a spot 55 feet from the left fielder's starting position after 2.75 seconds in the air. Suppose further that aggregated data from the entire league indicate that balls at that distance from the left fielder, for that time in the air, are caught 29 percent of the time (note: these values are not yet known, and will need to be calculated once FIELDf/x goes live across the major leagues, and a suitable volume of data have been captured).

If the left fielder did not catch the ball, the base value of 0.77 runs would be multiplied by 29 percent, and that value, 0.22 runs, would be subtracted from the left-fielder's TDR. If the left fielder did make the catch, the base value of 0.77 runs would be multiplied by (1-29 percent), and that value, 0.55 runs, would be added to the left fielder's TDR.

Because of the difficulty of the play, the fielder gets more credit for making the play than he would lose for missing it. This is not a new feature in defensive metrics, as zone based metrics have previously incorporated different values for different zones and types of hits, but what is new is the fact that the difficulty of the play will finally be *accurately* captured and built into the TDR metric.

We will no longer have to wonder if a player earned himself a sparkling zone-based metric by catching a lot of easy fly balls; nor will we have to rationalize away a weak metric by suggesting that a traditionally outstanding fielder must have had a lot of balls hit toward the outer edges of his zone while he was on the field. Every ball hit to a fielder will be individually measured, and individually scored. The TDR metric thus promises not only to be accurate, but also to be entirely transparent.

The key elements of the TDR metric, generated using the raw data that FIELDf/x provide, will obviously appeal to the sabermetrics crowd, but they will also resonate with the broader base of baseball fans, if presented in the right way. Imagine that during a baseball telecast, a hitter lines a double just beyond the outstretched glove of the left fielder, a slugger known more for his power than his defensive prowess. Within seconds, FIELDf/x can tell the audience that:

- 85 percent of all major league left fielders would have made that play
- The left fielder's failure to catch the ball cost his team, on average, 0.9 runs.
- The left fielder's TDR over the last week has been -3.2 runs.
- The left fielder's season total for TDR has been -8.7 runs over the first 95 games of the season, and projects his total defensive contribution over 150 games at -13.7 runs. This is the 26th best, or the fifth worst, total among major league left fielders.
- When the left fielder is replaced for a better defensive player late in the game, the screen could show the likelihood, in percentages, that a ball will be hit in the rest of that game that the first man would have missed, but the replacement would catch. The average run value could also be shown. For example, the FIELDf/x database of all batted balls might show a seven percent chance of a ball being hit in the eighth and ninth innings that would be missed by the original left fielder but caught by the substitute. This might work out to an expected run value of 0.1 runs (for example). This can provide a concrete way of showing how managers "play the percentages."

For a more visual presentation of FIELDf/x-generated information, a network televising the game could show a wide-angle shot of left field, with the left fielder's starting point marked, and then trace his path toward the ball, and show how he got there just too late. If a similar play has happened earlier in the game, it can be shown in an overlay, illustrating how the other team's left fielder, with his superior range, would have caught the ball.

There are many possible uses of FIELDf/x data, transformed into meaningful baseball information via the True Defensive Range metric, for all types of interested people: sabermetric analysts, television networks and viewers, and the baseball teams themselves.

Greg is right that people from all sides of the baseball community will be able to respond to FIELDf/x data in a unique way. Sportvision will also be able to apply FIELDf/x data to consumer applications so fans can see the defensive metrics illustrated on mobile devices as well as online applications and player analysis tools. Used in conjunction with Sportvision's PITCHf/x and HITf/x systems, FIELDf/x data will fill in the holes of what was once missing play action with meaningful information and revealing visuals.

Like the True Defensive Range metric, FIELDf/x data will give fresh fodder for analysts and bloggers to chew on, encouraging new baseball statistics and engaging conversations. Sportvision's systems set up new and better analysis, since the human error element is diffused and replaced with hard, objective data.

It is an exciting thing to unleash baseball facts and figures that have been untapped for so long, and Sportvision is once again at the forefront. Yet, for every uncovered answer, baseball has a way of coming up with several new questions. Sportvision will continue to respond to the expected queries of the baseball community with innovative technology, humbled to know that we may never know all there is about the game.