Infielder Error Analysis Proposal

Long Nguyen

April 2021

1 Introduction

This proposal details a tentative analysis plan in response to the infield coach's request, quoted as follows:

> "One of our infielders, Player X, seems to be struggling in the field. He's got a great arm, but he's made a few errors this season and is failing to get to some balls. Could you look into this and identify any problem areas that we can target with drills?"

2 Potential Approaches

Since player X was reported to have good arm, our analysis presumes his problem is primarily concerned with fielding batted balls. As such, we identify three approaches to examine possible performance issues:

1. Assessing difficulty of fielding based on distance. A closer look at player X's theoretical feasible range compared with actual play will reveal where he might be lagging behind, e.g. slow reaction time or wrong judgment on batted ball location.

2. Assessing difficulty of fielding based on direction. Orientation of the player's movement is taken into consideration to inspect whether player X is more prone to errors during forward, backward, or lateral movement.

3. Assessing difficulty of fielding based on type of batted ball. Narrowing down on the trajectory type player X typically makes errors on would inform us whether he is having issue with fielding ground ball in-between hop or hard line drives.

3 Data

The data required for this analysis is accessible via StatCast, and the table below presents the schema of the dataset.

Data	Definition
hc_x	Batted ball's hit x coordinate
hc_y	Batted ball's hit y coordinate
hit_distance	Projected batted ball's hit distance
launch_speed	Exit velocity of the batted ball
launch_angle	How high, in degrees, a ball was hit by a batter
bb_type	Batted ball type
Player Location	Player's coordinate at time of hit
First Step	The time elapsed from time of bat-on-ball contact to the fielder's first movement toward the ball
Speed	Player X's average speed.

4 Methodology

With the available data above, it is possible to conduct a comprehensive analysis of the approaches detailed in the second section of the document.

With respect to difficulty based on location, we first visualize batted ball fielding success/miss on a scatter graph of distance from ball versus hang time. Patterns of high success and high failure regions could be gleaned from the plot, and fielding rate is then calculated for each combination of space and time range. Additionally, data from sampled catching rate from other players can be overlayed to make comparisons between player X and the average player. We can then posit hypothetical scenarios, such as where player X's time to first step is reduced, to see how these improvements might adjust his probability of successful fielding.

In regards to classifications based on direction and type of batted ball, we can separate data into categorical bins and conduct statistical tests such as chi square test for independence to determine whether these factors might influence player X's error rate. Drills recommendations can then be made to focus training on specific type of batted balls or movement exercises.

5 Analysis

6 Resources

- New Metric for Infield Defense
- StatCast Search CSV Documentation