Revisiting the SAFE Framework in the Statcast Era: A Modernized Approach to Evalutating MLB Infield Defense

Tyler McCord[1]

[1]Amherst College,

Corresponding author: Tyler McCord, `tmccord26@amherst.edu`

Abstract

High resolution tracking data has transformed player evaluation in Major League Baseball (MLB), enabling high-level analysis of player performance. While public analyses on batting and pitching have advanced rapidly, defensive evaluation has been comparatively underdeveloped. The SAFE (Spatially Adjusted Fielding Evaluation) framework, introduced by Jensen et al. (2009), was the first effort in the public sphere to evaluate defense as a continuous space. We revisit the SAFE framework using modern Statcast data with an emphasis on infield defense, a notable struggle for prior defensive metrics.

1 Introduction

The evaluation of batting and pitching in baseball has been at the forefront of sports analytics for decades, mostly due to their discrete nature and the availability of relevant, quantifiable data. It is relatively simple to measure the outcome of a plate appearance or a pitch, making it easier to develop metrics that accurately reflect player performance in these areas. In contrast, defensive evaluation has lagged behind due to the continuous, spatio-temporal nature of fielding.

Still, Major League Baseball (MLB) organizations are faced with important decisions regarding defense, such as positioning players, making defensive substitutions, and evaluating trade-offs between offensive and defensive abilities. At the end of each season, MLB issues Gold Glove awards to the best defenders at each position, highlighting the importance of defense in the game.

Before the advent of high-resolution player tracking data, teams relied on simple defensive metrics such as fielding percentage, which calculates the percentage of plays a fielder successfully makes, and errors, which count the number of plays that the player does not make that the average fielder would. However, these metrics fail to capture the full scope of a player's defensive contributions, as they do not account for factors such as range, positioning, and the difficulty of plays made.

Statisticians have tried to find ways to quantify the nuances of defense. In 2003, Mitchel Lichtman introduced the Ultimate Zone Rating (UZR) metric, which attempted to evaluate defense by dividing the field into discrete zones and assigning run values to plays made or not made within those zones. This was the first run-based attempt at quantifying defense over a continuous space. In 2009, Jensen et al. (2009) introduced the SAFE (Spatially Adjusted Fielding Evaluation) framework, which built upon UZR by using a hierarchical Bayesian model to evaluate defense as a continuous surface. The SAFE framework uses estimates of player location, ball location, and ball velocity to model the probability of a fielder making a play on a batted ball, allowing for a more nuanced evaluation of defensive performance. The hierarchical Bayesian structure also allows for the sharing of information across players and positions, improving estimates for players with limited data. However, this model is limited by the accuracy and reliability of the underlying data used to estimate player and ball locations. These data, provided by Baseball Info Solutions, used hand-annotated video footage to estimate ball location and velocity. Even then, the starting location of the fielder at a given position was estimated by the authors by using the average location of balls caught by that position.

Notably, the results of Jensen et al. (2009) showed that the autocorrelation of defensive metrics from year to year was quite low, most notably for infielders. This shortcoming suggests that the original SAFE model performed poorly in evaluating infield defense, relative to outfield defense.

Since the publication of the SAFE framework, MLB has introduced Statcast, a high-resolution player tracking system that uses a combination of radar and camera technology to track the movement of players and the ball in real-time. Statcast

provides a wealth of data that was previously unavailable, including precise measurements of player and ball locations, velocities, and trajectories. This data has the potential to revolutionize defensive evaluation in baseball, allowing for more accurate and reliable estimates of defensive performance.

In this paper, we modernize the original SAFE framework for infielders using Statcast data. We also pose an improved model, with additional covariates that were not available in the original SAFE framework.

2 Data
References

Jensen, S. T., Shirley, K. E., & Wyner, A. J. (2009). Bayesball: A bayesian hierarchical model for evaluating fielding in major league baseball. *The Annals of Applied Statistics*, *3*(2), 491–520. https://doi.org/10.1214/08-AOAS228