

FldrOp - Fielder Position Optimization

Jaret Prothro David Ryu

Johns Hopkins University | Whiting School of Engineering | Baltimore, MD
Design Day 2021

Introduction

The FldrOp project began when the Baltimore Orioles approached the Sports Analytics group with an idea of creating a program that would determine the optimal positioning for each position player on the field based on the data of a specific hitter. This program would allow the Orioles to make data-driven, game-time decisions about moving their defense around to gain an advantage over their opponents. It will also shed some light on situations in which very unconventional defensive arrangements could be viable or even beneficial. More recently, we were approached by another team asking for help in optimizing their outfielder's positions. For the past few weeks, we have been running tests and sending them reports for their upcoming games.

Approach

By breaking down the baseball field into a grid of potential locations for a player to be positioned, the program can play a "game" in which it moves the position players around in order to optimize the defensive score metric that we are measuring. The defensive score is based on the amount of time it would take for the nearest player to reach the ball in play and to make a play. The longer that it takes a player to get to the ball in play the worse the defensive score. The program will then be fed a set of data for a specific hitter facing left-handed or right-handed pitchers so that it can optimize the defensive score. In doing so, the program will determine the optimal position for each player on the field.

An alternative approach being worked on is to use a formula that relates expected WOBAs to the distance of a player from the ball. All a player's balls-in-plays will then be plotted onto the baseball field grid and players will be moved to locations that minimize the amount of expected WOBAs gained for the data set.

Procedures

The project can be broken down into 4 major steps.

- 1) Create an environment that allows for the mapping of a fielder's positioning and all of a hitter's balls in play.
- 2) Create formulas that utilize Statcast data to calculate the hangtime of balls in the air and the time a fielder would have to field groundballs.
- 3) Determine each fielder's unique range based on the amount of opportunity time of each ball in play and the fielder's speed and reaction time.
- 4) Have the program place fielders in the field environment and score each defensive set up based on the defensive score metric to find fully optimized defenses for each specific hitter.

Progress and Results

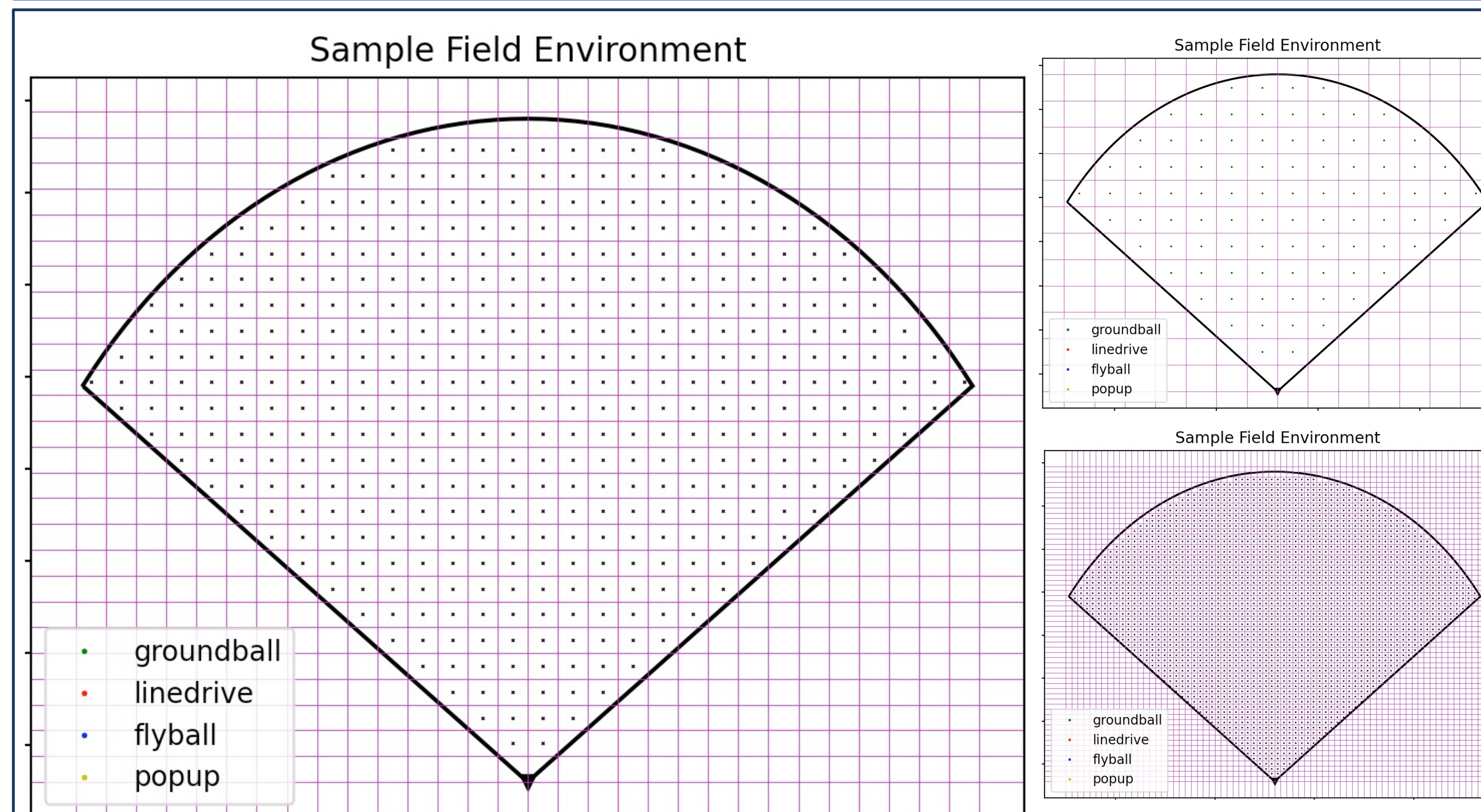


Figure 1 - Field Environment

This is a sample of the Field Environment where balls in play and fielder positions can be plotted. Each of the squares in the left image represents a 7 foot by 7 foot square on the baseball field and each of the black dots represents a potential location that the program could place a player. The precision of the program can be changed by altering the size of each of the squares, allowing for more potential locations a player could be placed. The top right image shows squares that are 15 feet by 15 feet and the bottom right image shows squares that are 3 feet by 3 feet. Additional constraints can be added to the potential positions a player can be placed by the program. For example, we can say that we want to make sure the left fielder stays within this general area and the program will make sure that all of the outfield setups that it generates has the left fielder within those constraints. This is especially useful for keeping the first baseman within a reasonable distance to first base.

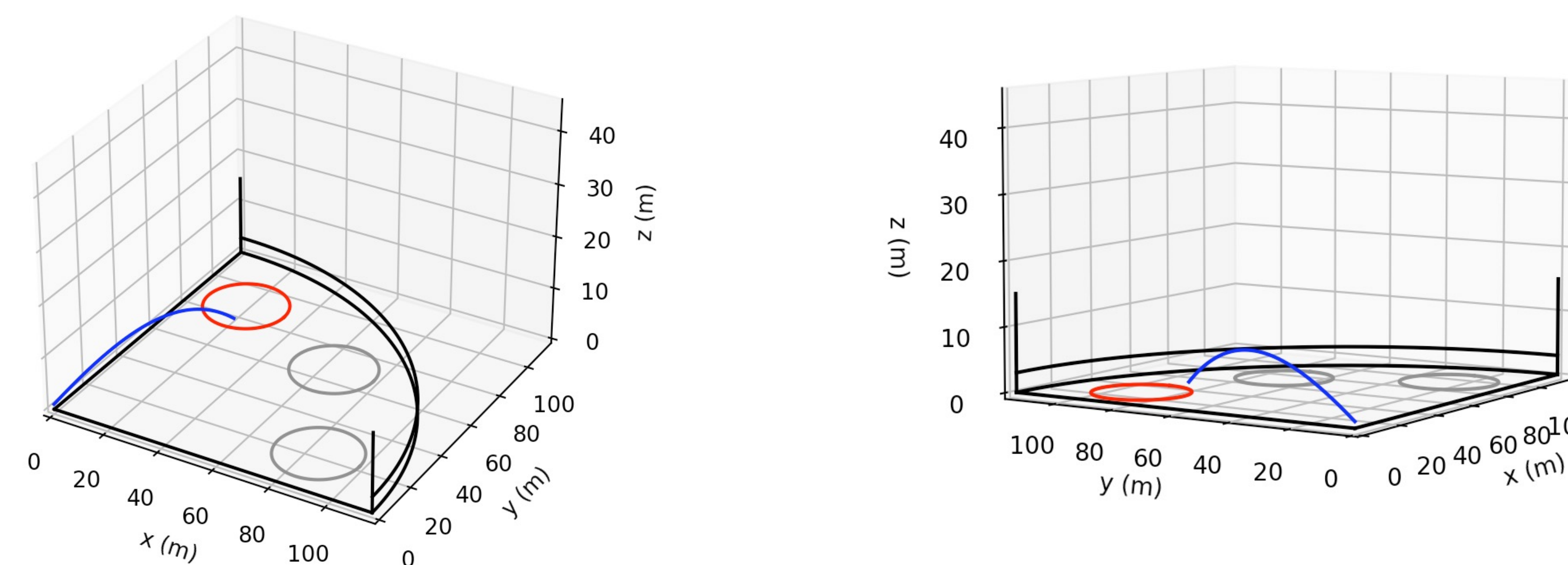


Figure 2 - Balls-in-Play

We have separated all of the hitter's balls-in-play based on the handedness of the pitcher. To simplify the program, we are only optimizing the positions of the outfielders. Therefore, we are only looking at line drives, popups, and flyouts. Each ball-in-play is plotted in this 3D model as the hangtime is calculated because Statcast does not provide this data. After the hangtime is calculated, the range of the outfielders can be calculated, as shown by the circles in the above images. The red circle represents the player that made the play on that particular ball.

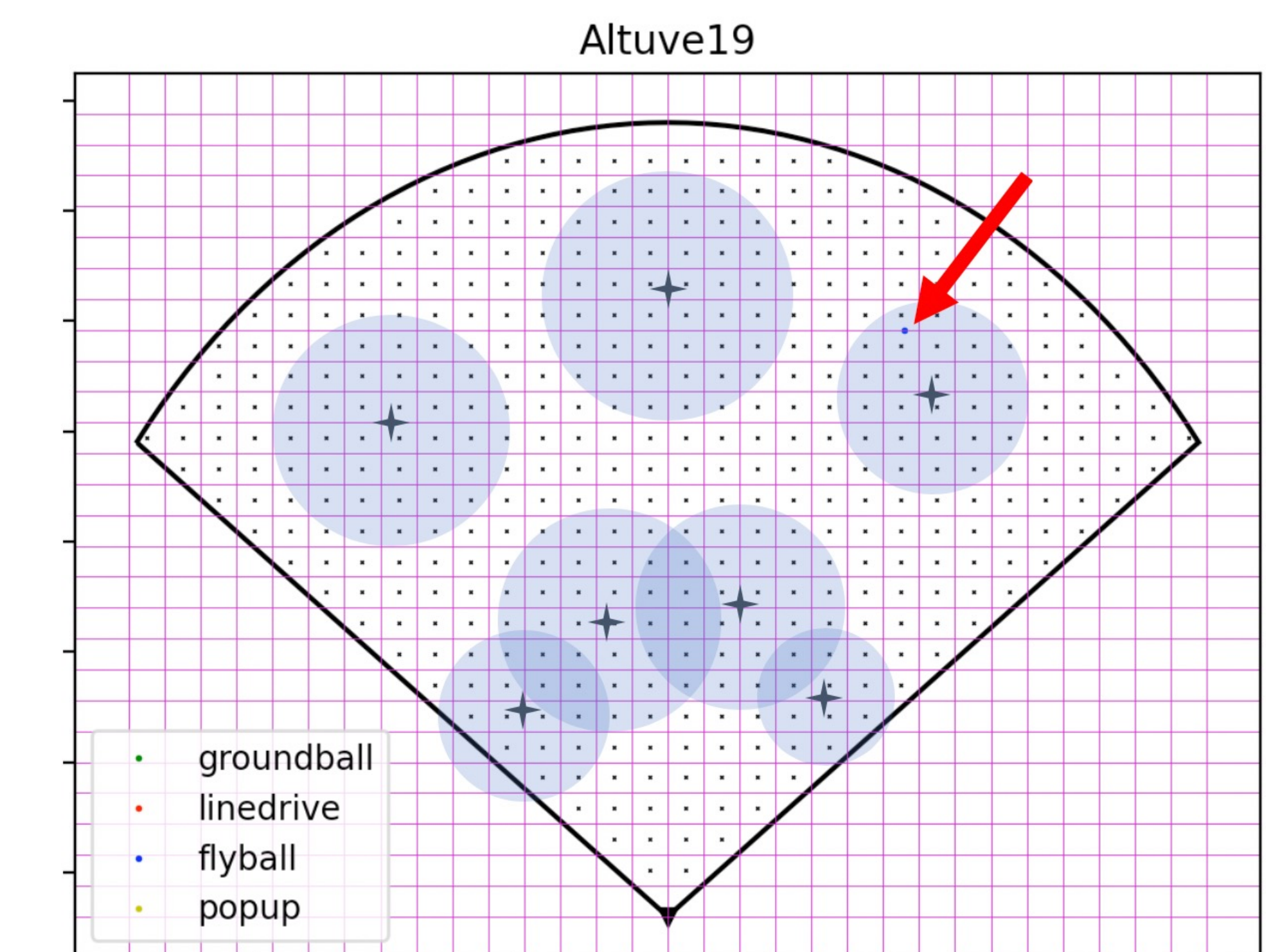


Figure 3 - Defensive Scoring

The defensive score generated for each different position setup is based off of the amount of time it takes a player to get to a ball in play. If the player can catch the ball, the program counts that ball-in-play as an out. Time thresholds are set so that if it takes longer than x seconds, the ball in play is considered a single. If it takes a player longer than y seconds, where $y > x$, the ball might be classified as a double, so on and so forth. The score of each of these balls in play is based on the wOBA weighting.

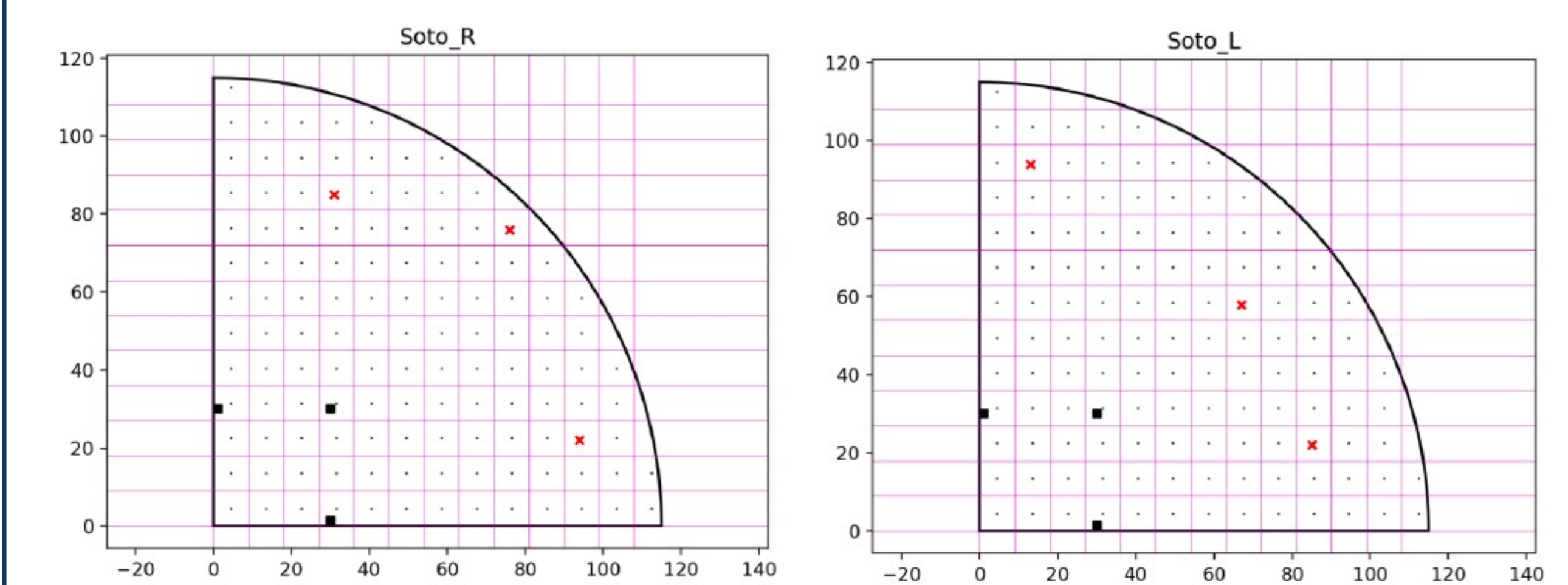


Figure 4 - Optimal Fielder Positions

After the program has computed defensive scores for at least 150 different outfield setups, the program can compare all of the defensive scores to a baseline score where the outfielders are playing in standard positions. The two plots above shown optimized positioning for Juan Soto against right-handed pitching and left-handed pitching. Sometimes there are multiple outfield setups that score better than the standard set up, and we have elected to include those in the player reports as well.

Conclusion and Future Work

We are very pleased with the results that the program has been able to generate in its current state. We will continue to optimize the process and better adjust the defensive scoring metrics as we run more tests. Future tests will also include four infielders for the program to move around as well in order to try to find an optimized setup for the entire defense.