Optimizing Defensive Softball Shifting Models

Chloe Kelly, Matthew Lemoine, Gowri Priya Sunkara,
and Finley Westmoreland

May 5, 2026

Abstract

This project aims to discuss the viability of shifting within the sport of softball. To achieve this, we outline a method for finding the optimal placement of fielders to maximize the chances of catching a fly ball or getting a member of the opposing team out. Utilizing min-max optimization, we are able to find the best possible positions for fielders under certain limitations and scenarios.

1 Introduction

Softball is a game in which two teams take turns hitting a thrown ball. Each team receives an opportunity to hit/catch the ball in cycles called *innings*. The defending team has nine players on the field: the *pitcher* who throws the ball towards the offending player, the *catcher* who stands behind the defending player to catch the ball if the player misses it, and seven *fielders* who stand ready to catch the ball once it is in play. There are two distinct types of fielders: *in-fielders* and *out-fielders*. In-fielders are defensive team members that are playing the specific positions of pitcher, catcher, short stop, or first, second and third base. The out-fielders play further away from home plate than the in-fielders; those positions are referred to as right-fielder, left-fielder, and center-fielder. The offending team sets forth a *batter* who stands ready to hit a ball thrown by the defending team's pitcher, and when that ball is hit, then the ball is in play.

Shifting is a strategy in which the defending players move prior to the ball being set into play in order to predict where the batter will hit the ball. If successful, this will enable the defending team to prevent more runs than the standard positioning. However, if the defending team shifts inaccurately, they risk giving up far more runs to the offending team. As such, successful predictive models for shifting ensure that the fielding team maximizes their defensive capabilities and minimizes their risk of giving up more runs.

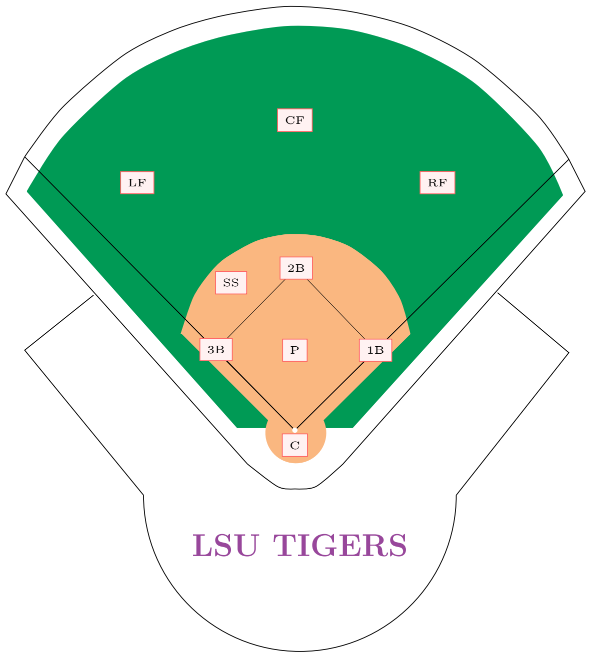

Figure 1: Standard Positioning for Defensive Fielders

The purpose of this project is to design an accurate shifting model using data from Louisiana State University's Trackman system for the LSU Softball Team. The Trackman system is a radar system that analyzes where the ball is at any given moment and records all information about the ball as it is in play. Things included in the record are: what the play was called (single, double, triple, strike, ball, etc.), where the ball landed, who was at bat, who was the pitcher, the handedness of the pitcher and batter, etc. The handedness of the batter is determined by which hand is higher when they hold the bat, not the direction they are facing relative to someone sitting in the stands.

2 Methodology

In order to determine where to move the defenders, we must first determine where the *buckets* are. Buckets designate the areas in which the balls are both aggregating and producing the highest run value for the offenders. Through the data provided by the Trackman system, we analyzed the variables (listed in the glossary at the end of this report) and created a chart to designate where the balls in play landed; this is referred to as a *Spray chart*.

We decided to apply seven buckets (one for each player we are considering in the specific shift). To make these, we took the variables listed in the glossary, designated the standard starting positions of all the defending players (except the pitcher and catcher), and used k-means clustering (k = 7) to find the centroid of 7 places where we could place our fielders. With these centroids we ran multi tree regression using the max-min equation to determine the best possible shifted position.

More specifically, after generating the cluster centroids, these locations, the buckets, are treated as initial candidate positions for the seven fielders. To refine these positions, we model the field on a two dimensional coordinate system with the x-axis directly behind home plate and the y-axis moving perpendicularly through the pitcher's mound and to the back fence.

Each batted ball is represented by it's distance and bearing that has been converted into Cartesian coordinates after each hit,

$$p_i = (x_i, y_i), \quad x_i = d_i \sin(\theta_i), \quad y_i = d_i \cos(\theta_i) \tag{1}$$

where p_i, d_i and θ_i denote the location, distance, and bearing (with respect to the positive y-axis) of the i^{th}hit. Each hit location is then assigned a weight that combines its expected run value with the locations of the runners on base, given by

$$w_i = w_i^{xRV} \times w_i^{zone} \tag{2}$$

This emphasizes the more strategically important regions of the field. Using these weighted hit locations, we apply an iterative min-max optimization procedure that seeks to minimize the worst-case weighted distance between hits and defenders by

$$\max_i w_i \times \min_j \|p_i - d_j\|$$

where d_j denotes the position of the j^{th} defender. We approximate the solution through an iterative procedure, where at each step, we identify the index

$$i^* = \arg\max_i w_i \times \min_j \|p_i - d_j\|\,, \tag{3}$$

corresponding to the worst covered hit, and determine the nearest defender.

$$j^* = \arg\min_j \|p_i - d_j\|$$

The selected defender is then updated toward that hit location via $d_{j^*} \leftarrow d_j^* + \eta\left(p_i^* - d_j^*\right)$ where $\eta > 0$ is a step size parameter. This process is repeated for a fixed number of iterations, producing defender positions that progressively reduce the worst-case weighted coverage gap in the field. The resulting positions are then constrained to lie within the field boundaries and are assigned to defensive roles based on depth and horizontal ordering.

The pitcher and catcher remain locked in every scenario but some of the mutable positions must also have limitations. The defender on first base is not allowed to move a certain distance from first base regardless of how many people were on any of the bases. This position still shifts, just minimally; the player will be moved toward their bucket via a vector whose magnitude cannot exceed its maximum scalar. Additional adjustments are applied to the optimized defender positions depending on the base runner configuration in order to reflect the game play constraints.

3 Results

Using the Trackman dataset and considering non-LSU balls put in play, the model computes data-driven defensive shifts by combining clustering of hit locations with a weighted max–min optimization framework. To account for differences in hitting tendencies, the analysis is divided into two groups based on batter handedness: left-handed and right-handed batters. This separation allows the model to generate more accurate and context-specific defensive shifts, as spray patterns vary significantly between the two groups.

For left-handed batters with a runner on first base, the optimized value decreases marginally from 88.19 to 87.18. This small improvement indicates that the defensive shift is highly constrained in this scenario. In particular, the first baseman must remain close to the base to hold the runner; this limits the flexibility of the infield and restricts how effectively the defense can redistribute coverage.

In contrast, the optimized value for left-handed batters with a runner on second base decreases significantly from 83.21 to 63.06. This reduction is substantial and reflects a more effective shift compared to the first scenario. The model is able to reposition both infielders and outfielders to better align with the observed hit distribution, thereby reducing large uncovered regions.

For the scenario with runner-on-third and left-handed batters, the optimized value further decreases from 83.21 to 57.24. This indicates an even greater improvement in defensive coverage. In this case, the shift prioritizes minimizing high-impact gaps while maintaining sufficient depth to prevent extra-base hits, resulting in a more balanced and efficient defensive configuration.

For right-handed batters with a runner on first base, the optimized value drops from 132.10 to 75.26. This represents a substantial improvement and shows that the shift effectively compensates for initial defensive gaps even under moderate positional constraints. The redistribution of defenders significantly reduces the largest uncovered region.

When a runner is on second base for right-handed batters, the optimized value decreases further from 132.10 to 58.64; this is one of the most effective scenarios. The absence of restrictive positioning allows the model to fully adapt the defensive shift to the tendencies of the hitter, resulting in a well-balanced coverage across the field.

Finally, for right-handed batters with a runner on third base, the optimized value reduces from 132.10 to 61.83. Although slightly less effective than the runner-on-second case, the shift still produces a strong improvement. The defensive positioning balances the need to prevent runs from scoring with maintaining adequate field coverage.

These results demonstrate that the effectiveness of the defensive shift is highly dependent on both batter handedness and runner situation. The analysis is based on 4,256 non-LSU balls put in play and consists of 1,902 left-handed and 2,354 right-handed batter observations. By separating left- and right-handed batters, the model is able to capture distinct hit distributions and produce more targeted and effective defensive shifts. On average, optimized shifts reduce the weighted min-max objective by approximately 35%, indicating a substantial improvement.

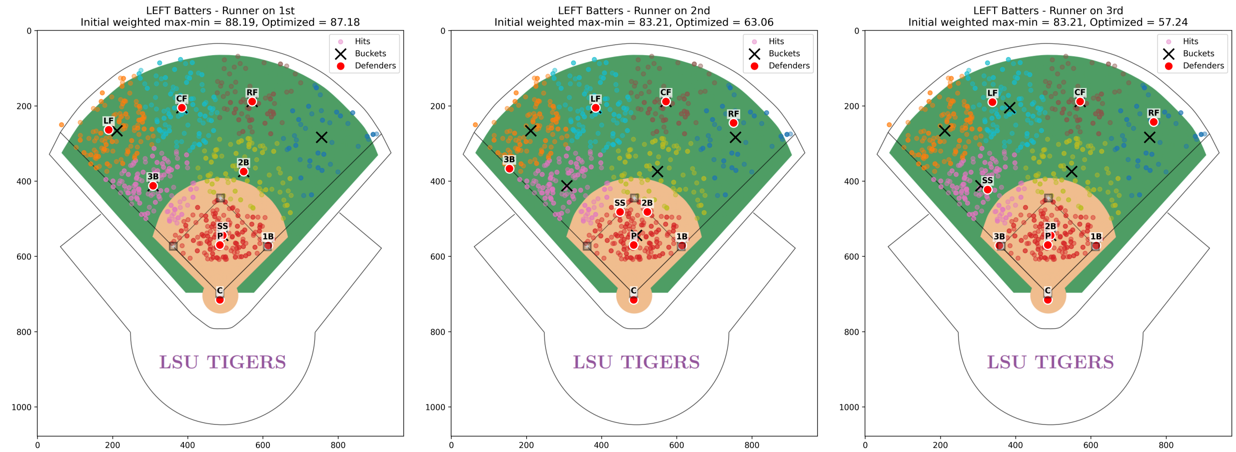

Figure 2: Defense Strategies for left-hand Runners shown when they are on different bases

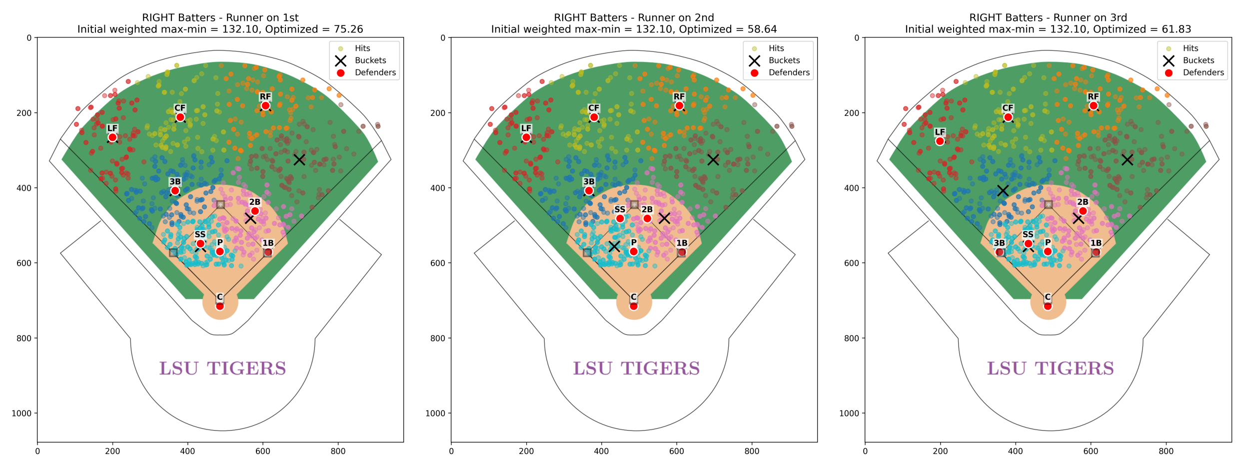

Figure 3: Defense Strategies for Right-hand Runners shown when they are on different bases

4 Conclusion

The techniques used here can be applied to help improve the defensive capabilities of the softball team while playing games; shifting based on the handedness of the batters can result in more outs and fewer runs for the opposing team. In future work, there are several directions that this project can take. An increased emphasis on specific teams that are batting could provide tailored plays and would be useful in minimizing runs given up. Another direction may be to focus on how different shifting models affect the game from a game theoretic stand-point. An important question to consider here may be "Does shifting make other areas of the field vulnerable to the offensive team?"

An assumption we made was that if a batter is left/right handed, then they will hit the ball more to the right/left side of the field, respectively. We also planned to categorize non-LSU batters into distinct hitter profiles based on their batted ball tendencies and apply the same min-max optimization framework to each profile to develop more precise defensive strategies that reduce expected runs.

We suggest exploring HDBSCAN and Gaussian Mixture Models (as alternatives to K-means) to better capture nonlinear patterns in sports analytics. Future iterations of this project may also wish to explore more complex simulations in which there are runners on multiple bases at once.

A Glossary of Utilized Trackman Variables

- ExitSpeed: Speed of the ball the moment it comes off the bat, reported in miles per hour or meters per second.

- Angle: The vertical angle formed by the intersection of the y-axis and the ball's path (in the z-direction) as it leaves the bat.

- Direction: The horizontal angle formed by the intersection of the y-axis and the ball's path (in the x-direction) as it leaves the bat.

- HangTime: Time elapsed from the moment the ball hits the bat until the ball lands or would have landed if not obstructed, reported in seconds.

- HitSpinrate: Speed at which the ball is spinning as it leaves the bat, reported in revolutions per minute (RPM).

- SpinRate: Speed at which the ball is spinning, reported in revolutions per minute (RPM)

- RelSpeed: Speed of the ball as it leaves the pitcher's hand, reported in miles per hour or meters per second.

- VertRelAngle: The vertical angle formed by the intersection of the y-axis and the ball's path (in the z-direction) as it leaves the pitcher's hand. A negative number means the ball is sloping downward, while a positive number means it's sloping upward.

- HorzRelAngle: The horizontal angle formed by the intersection of the ball's path (in the z-direction) as it leaves the pitcher's hand. A negative number means the ball is travelling towards the LHB batter's box, while a positive number means it's travelling towards the RHB batter's box.

- VertBreak: Distance between the height at which the ball crosses home plate and the height at which it would have crossed if it travelled in a straight line from release and were completely unaffected by gravity, reported in inches or centimeters.

- HorzBreak: Horizontal distance between where the ball crosses home plate and where it would have crossed if it had travelled in a straight line from release and were unaffected by gravity, reported in inches or centimeters.

- InducedVertBreak: Distance between the height at which the ball crosses home plate and the height at which it would have crossed if it travelled in a straight line from release and were affected by gravity, reported in inches or centimeters.

- ZoneSpeed: Speed of the ball as it crosses the front of home plate, reported in miles per hour or meters per second.

- ZoneTime: Time elapsed from when the pitcher releases the ball to when it crosses the front of home plate, reported in seconds.

References

[1] W. Melville, S. Wise, G. Nielson, T. Mott, C. Archibald, and D. Grimsman, "Optimizing baseball fielder positioning with consideration for adaptable hitters," 42 Analytics, 2024.

[2] J. Harris, E. Bouzarth, B. Grannan, A. Hartley, K. Hutson, and E. Morton, "Swing shift: A mathematical approach to defensive positioning in baseball," 42 Analytics, 2020.

[3] N. Paine, "Why baseball revived a 60-year-old strategy designed to stop ted williams," ABC News, 2016.