

Quantifying Manager Pull Decisions



Owen St.Onge, Seth Missal, Max Kaplan

Introduction



Manager's biggest impact on an individual game – calls to the bullpen

Bullpen decisions create lots of controversy for individual managers: should they have made the move?

Yes?



No?



How can we optimize manager bullpen decisions?



Goal: Determine how many **runs** a pitching decision creates or loses

Model:
Probability of
runs scored
in inning



Determine:
Runs lost or
gained in
current game



Project:
Marginal effect
on next game

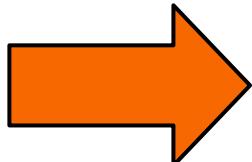


Propose:
Rule
Change



Starters get pulled when:

1. In trouble
2. Fatigued
3. Work is done: trust in the bullpen to close out the game



XGBoost Model:

Probability of runs scored in inning

Game State:

Base state, outs, pitcher and batter skill, platoon matchup, ground ball matchups

Pitcher Fatigue:

Pitch count, inning of work, velocity decline, command decline, times through order, game xwOBA

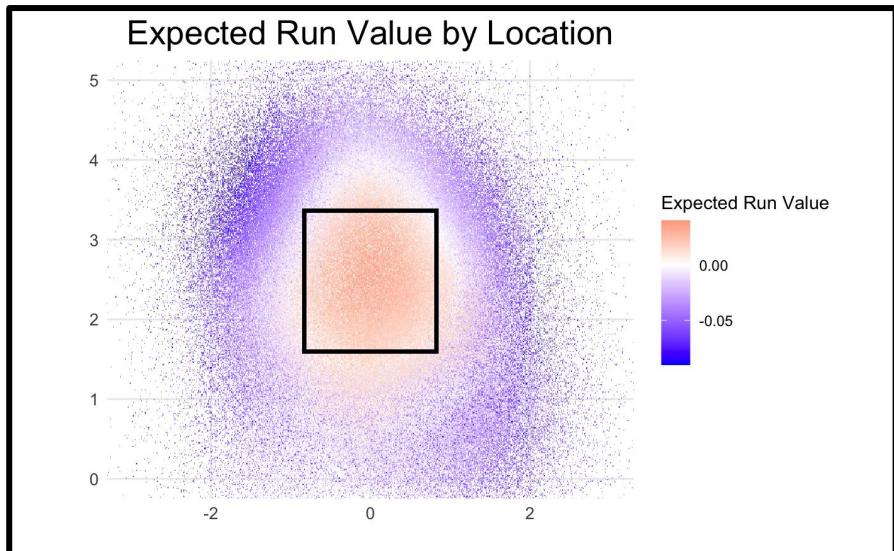
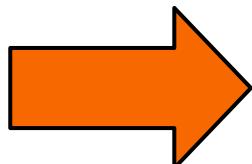
Command Decline



A manager may take a pitcher out because they have **lost command**. How do we account for this factor in our model?

GAM predicting ΔRE

- Pitch location
- Pitch Type
- Strike Zone Dimensions
- Batter Side



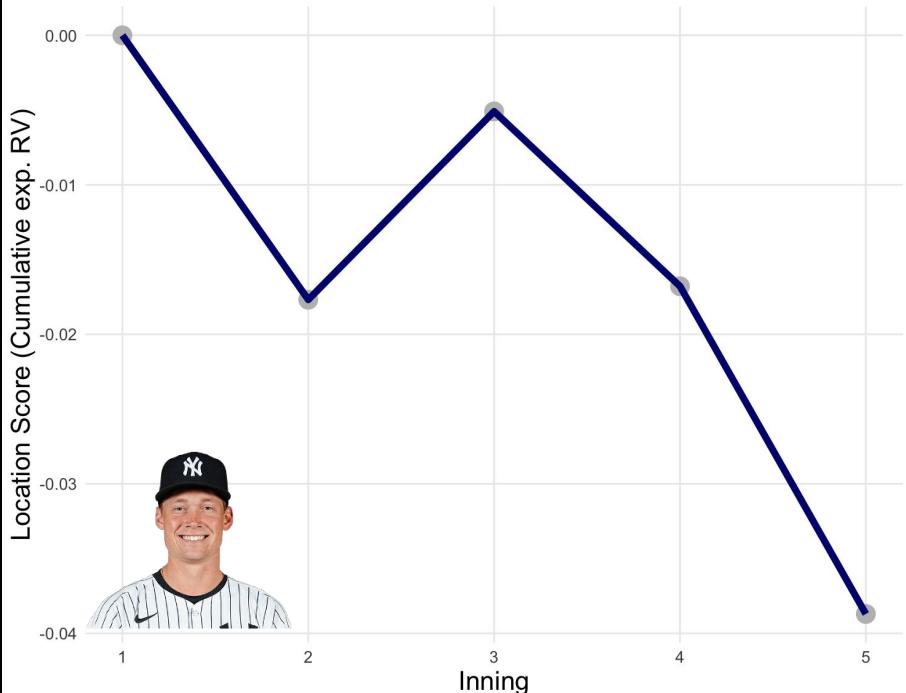
Location Scores were calculated for each inning and then **compared to the first inning to measure decline**. The same process was performed for **primary fastball velocity**



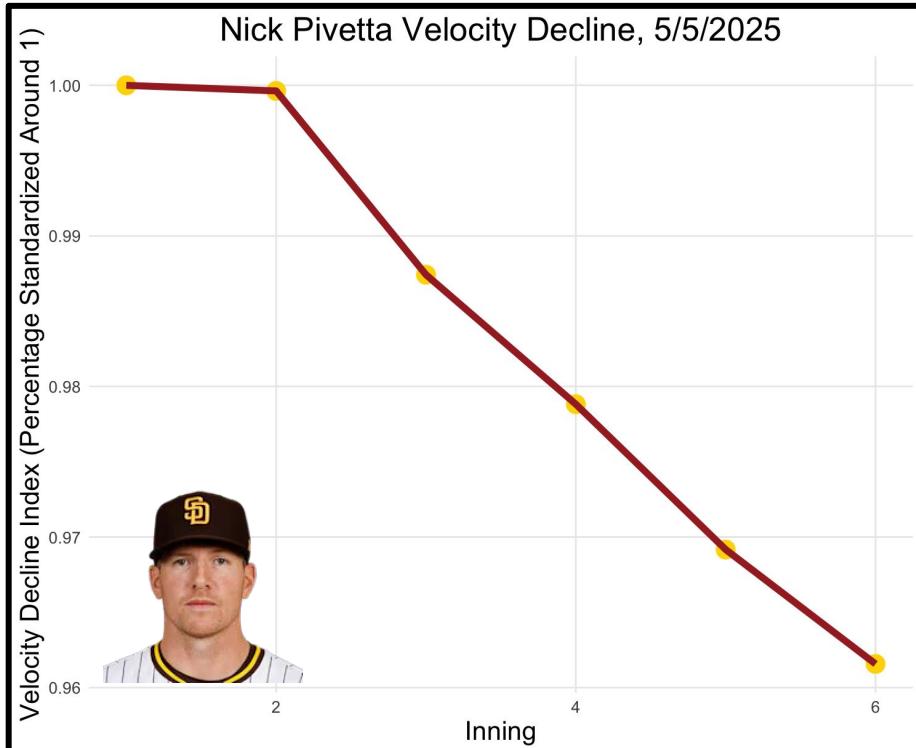
Command/Velocity Decline



Will Warren Location Decline, 8/23/2025



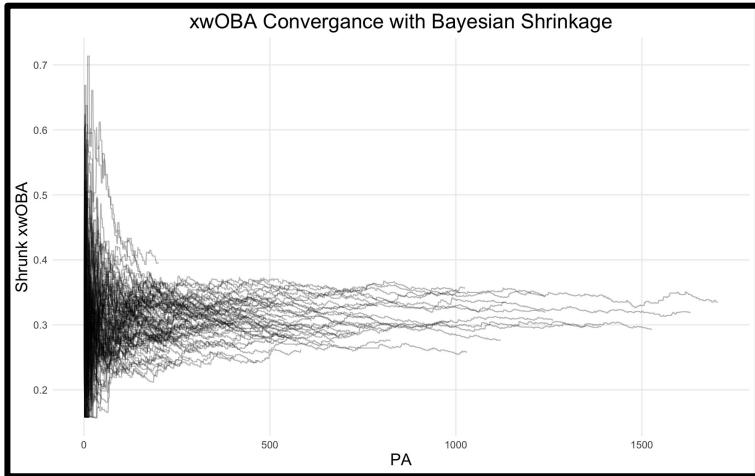
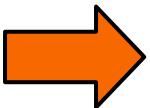
Nick Pivetta Velocity Decline, 5/5/2025



These scores for each inning were used as features in the model

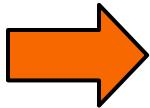
xwOBA Stabilizes after 350 PA (Carleton)

- Rolling mean of last 350 PA
- Players who did not meet this were given imputed values
 - Half league mean, half player mean
 - A stat each for righties and lefties



Shrunk values used to create a single xwOBA score for the next hitters based on the historical probability that they will be faced in the inning

- Single values for pitcher and hitter skill
- Same process done for ground ball rates



Weighted Calculation					
Position	xwOBA	×	Probability	=	Contribution
Current	0.320	×	1.00	=	0.320
+1	0.345	×	0.90	=	0.310
+2	0.298	×	0.83	=	0.247
+3	0.367	×	0.40	=	0.147
+4	0.312	×	0.15	=	0.047
+5+	0.289	×	0.03	=	0.009
				Sum of contributions:	= 1.080
				Sum of probabilities:	+ 3.31
					Weighted xwOBA: = 0.326



XGBoost Model:

Probability of runs scored in inning

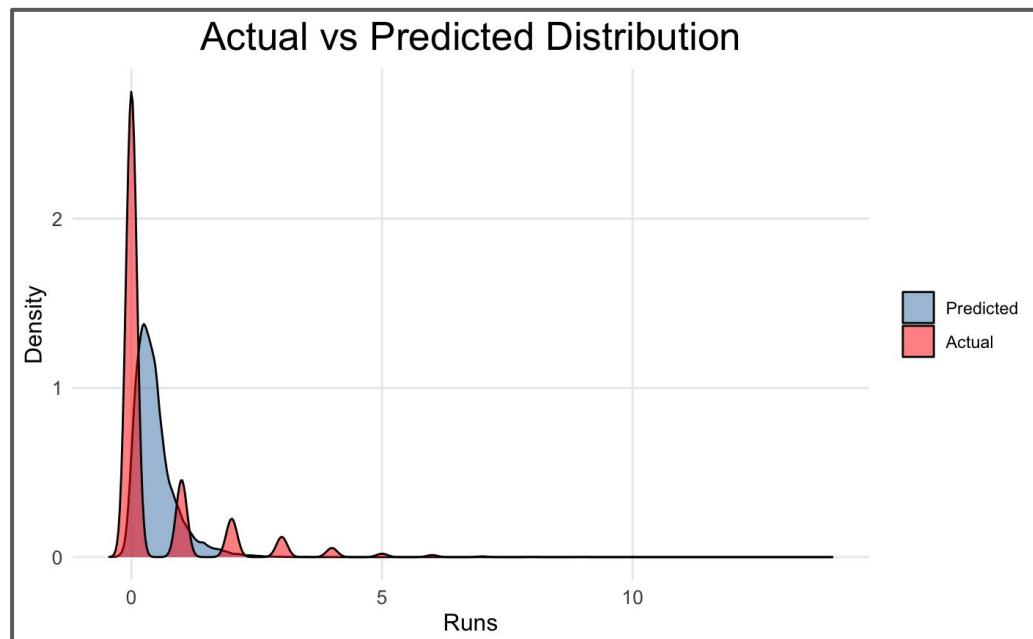
Game State:

Base state, outs, pitcher and batter skill, platoon matchup, ground ball matchups

Pitcher Fatigue:

Pitch count, inning of work, velocity decline, command decline, times through order, game xwOBA

Model prediction for runs scored in the remainder of inning:



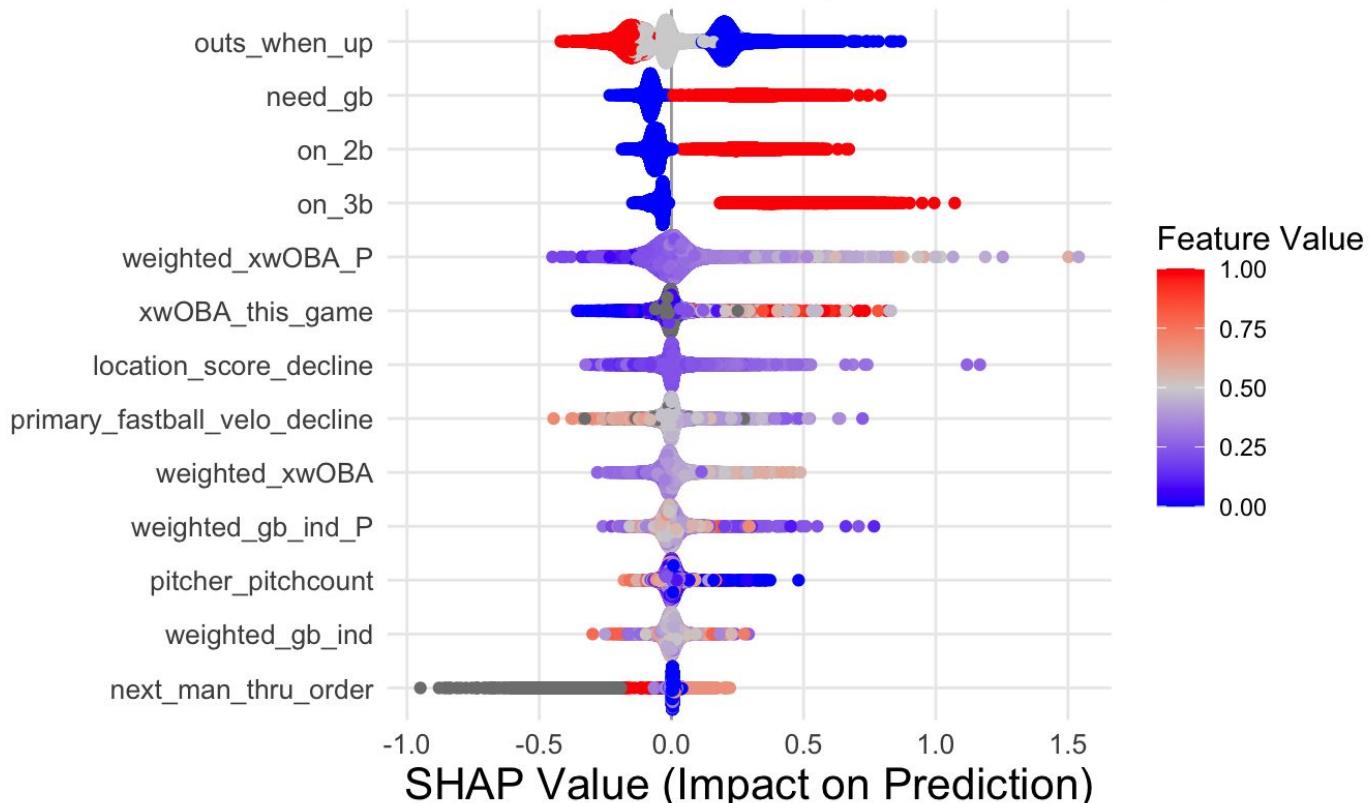
RMSE: .918 – overpredicting 0 runs, underpredicting positive runs, corrected with further tuning



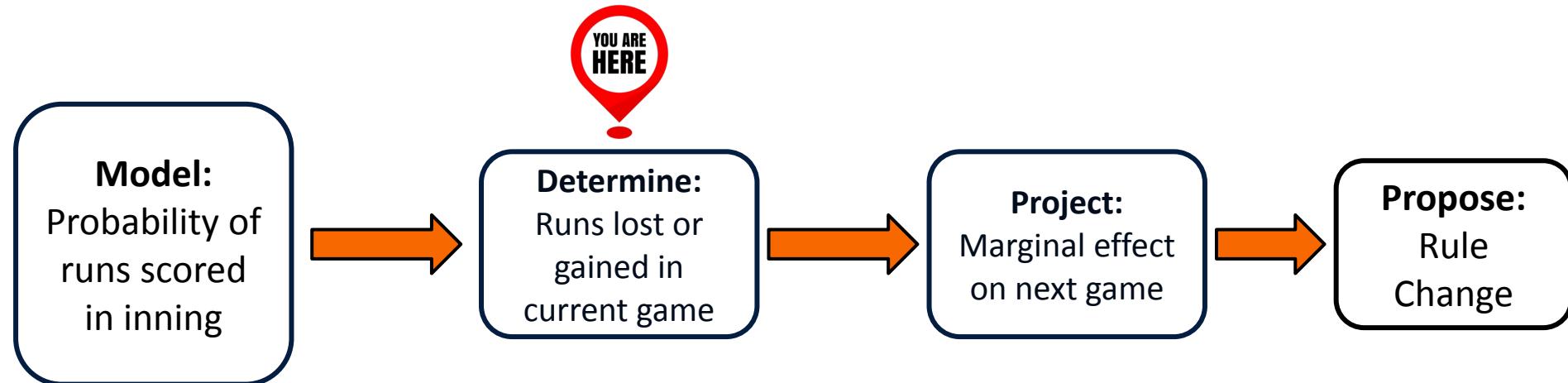
Regression XGBoost



Feature Importance (SHAP Values)



Next: Determine how many runs the manager's decision cost or saved



Example 1



Lorenzen is taken out mid-inning...

Michael Lorenzen



5th Inning

1 Outs



48% weighted ground ball rate
.317 opp. weighted xwOBA (R)
85 pitches, 3rd time through

Angel Zerpa



5th Inning

1 Outs



55% weighted ground ball rate
.286 opp. weighted xwOBA (L)
Fresh arm





Example 1



Michael Lorenzen



5th Inning
1 Outs



2.275 Predicted Runs

Angel Zerpa



5th Inning
1 Outs



0.477 Predicted Runs



$$2.274 - 0.477 = \text{Net gain } 1.80 \text{ Runs}$$

Example 2



Vásquez is taken out at the end of the 6th inning... Simulate next inning with Peralta and Vásquez

Randy Vásquez



Maintained location and velo
Batters have .107 xwOBA this game
Good matchup comparably (.337 xwOBA)

Wandy Peralta



Comes in to relieve Vásquez (0 RA)
Matchup not in his favor (.371 xwOBA)

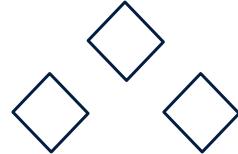
Example 2



Randy Vásquez



7th Inning
0 Outs

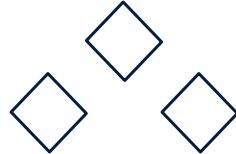


0.127 Predicted Runs

Wandy Peralta



7th Inning
0 Outs



1.054 Predicted Runs



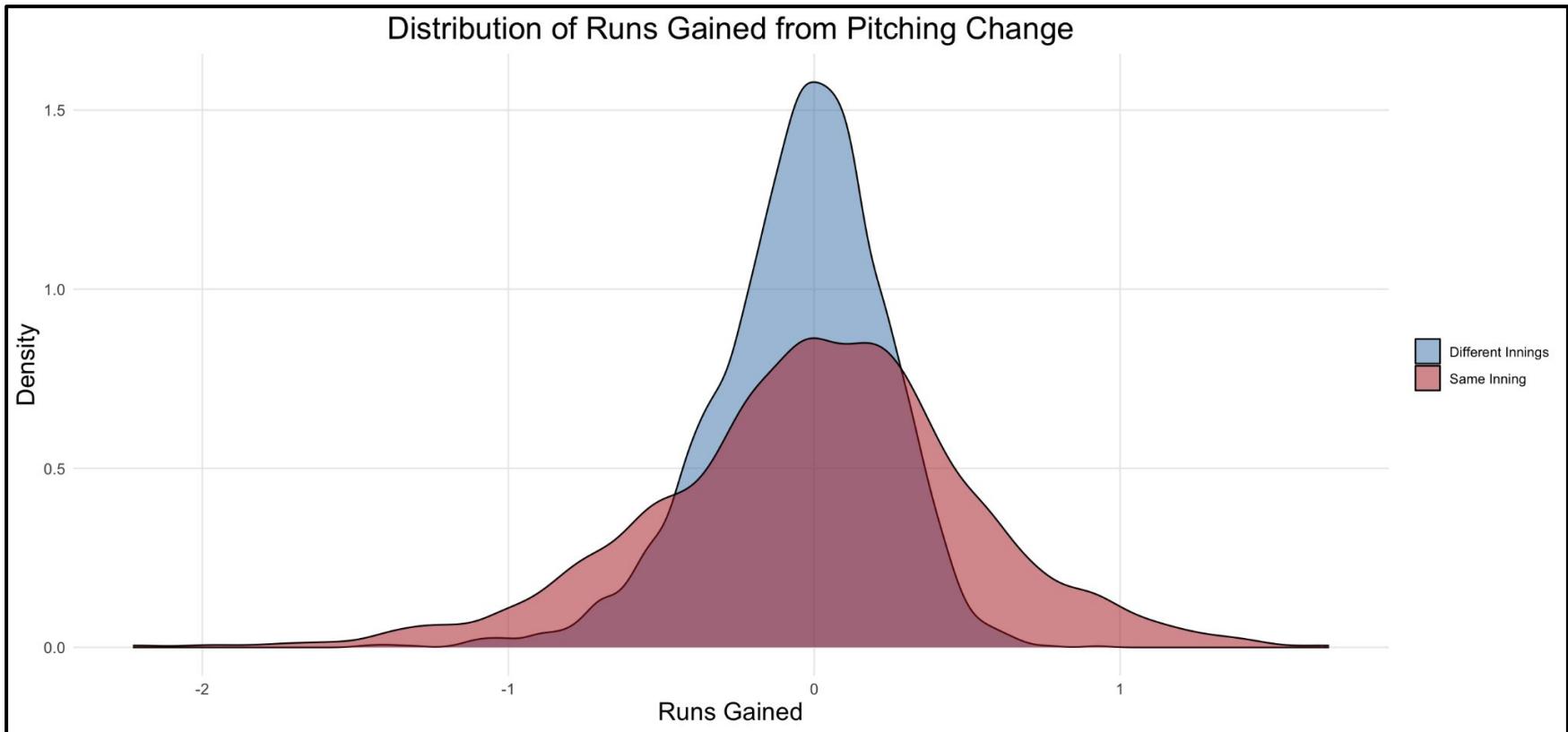
$$0.127 - 1.054 = \text{Net loss } 0.926 \text{ Runs}$$



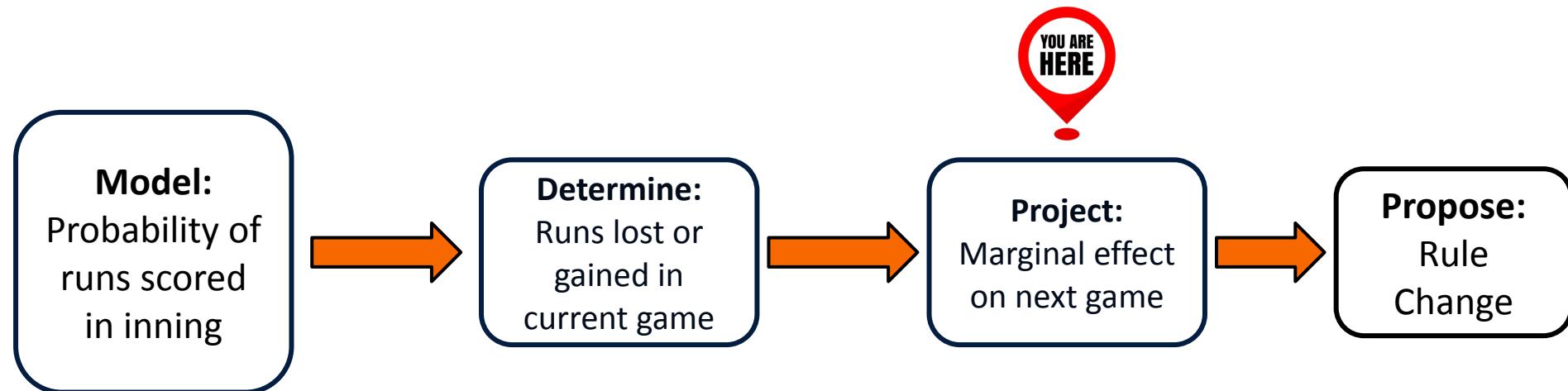
Bullpen Move Runs



Process is done to every starting pitcher pull in the 2025 season



Next: Determine how the bullpen was affected





Starter Pulls and Bullpen Workload



Assumption: The starter getting pulled impacts the **next** pitcher's workload only

Michael Lorenzen



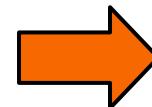
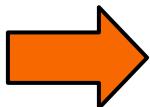
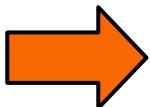
Angel Zerpa



John Schreiber

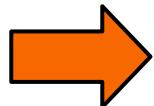


Lucas Erceg

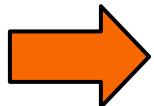


Assumption: The starter getting pulled impacts the **next** pitcher's workload only

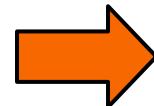
Michael Lorenzen



Angel Zerpa



John Schreiber



Lucas Erceis



We are going to focus only on the reliever that follows the starter:
how does Zerpa having to pitch in this game impact tomorrow's game?



Assumption: The starter getting pulled impacts the **next** pitcher's workload only

Pros:

- No “butterfly effect”
- Time convenient
- Don’t have to guess at managerial decisions

Cons:

- Only a piece of the puzzle
- If starter remains in game, different pitcher could come in

Although a limitation, we decided the assumption would improve our analysis

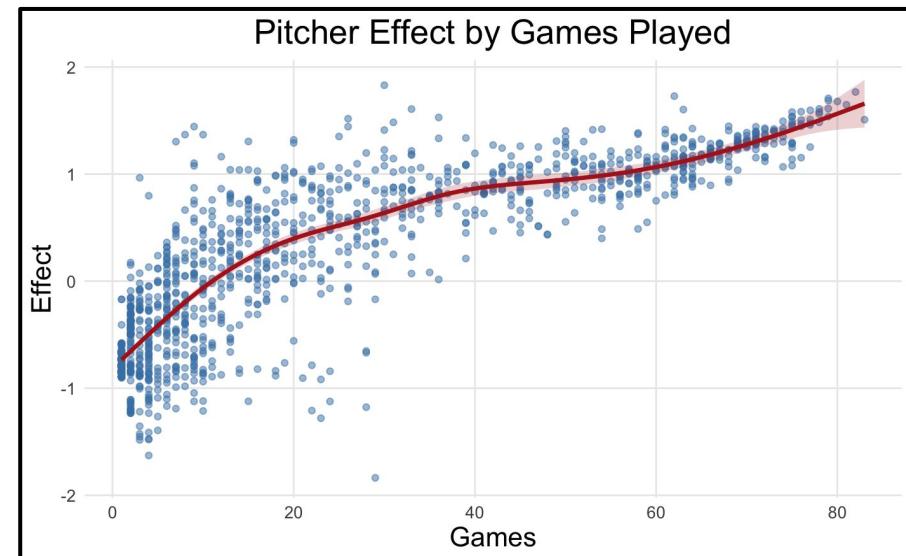
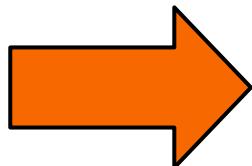
“Active” Pitchers



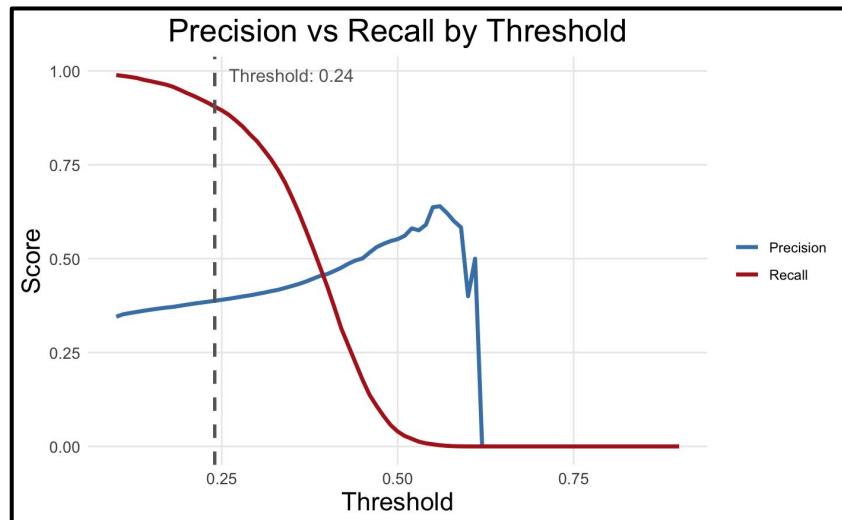
We want to know how having a reliever be “**available**” impacts a game.
First — how do we know if someone is **available**?

Mixed Effects Logit predicting reliever appearance

- Back to back indicator
- Days rest
- Random effect for pitcher
- Games pitched last 6 days
- Pitches in last 6 days
 - Diminishing marginal correlations after 6 days



Model wasn't very accurate, but that's okay – we don't want to know who played, we want to know who is available, so we minimize false negatives



Maximize precision while recall is at least 90%:
classification threshold of .24

Confusion Matrix for Pitcher Appearances

		Actual	
		Appeared	Did Not Appear
Predicted	Appeared	39767	25207
	Did Not Appear	21082	2647

Extremely poor at predicting appearances correctly,
but very good at predicting non-appearances
correctly, meaning we can **predict unavailability**

Converting Availability to Runs

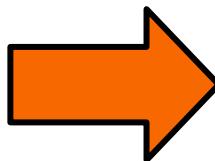


Now we will know if Zerpa will be available for his next outing based on the pitch counts from his appearance in relief of Lorenzen

Michael Lorenzen



Angel Zerpa



He pitched 20 pitches in relief, whereas if Lorenzen was left in the game, we are assuming Zerpa pitches 0 pitches. How can we convert this to runs so we can compare to the previous model?

Bullpen Pitcher xwOBAs were averaged to create a Strength of Bullpen metric where only the **available pitchers**, per our predictions, were used

Negative Binomial Regression for Opposing Runs Scored

- Strength of Bullpen
- Number of inactive pen arms
- Starting Pitcher and opp. lineup xwOBA
- Runs per game

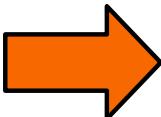
Characteristic	IRR	95% CI	p-value
strength_of_bullpen	1.03	1.01, 1.04	<0.001
number_inactive	1.01	0.99, 1.02	0.3
xwOBA_R_P	15.3	9.29, 25.3	<0.001
xwOBA_L_P	5.56	3.53, 8.78	<0.001
strength_of_lineup_R_other	143	43.5, 470	<0.001
strength_of_lineup_L_other	0.54	0.18, 1.65	0.3
rpg_allowed	1.02	1.0, 1.04	0.13
rpg_scored_other	1.03	1.01, 1.06	0.011

Strength of Bullpen is significant, with every **.1 increase in bullpen xwOBA** yielding a marginal effect of **.1092 runs**

Usage Future Runs Calculation



Our availability model tells us that Zerpa will not be available tomorrow if he pitches, but will be if he does not pitch



Bullpen With Zerpa:
.301 xwOBA

Bullpen Without Zerpa:
.316 xwOBA

.15 xwOBA lost,
.1 xwOBA = .1092 runs
.1092 x 1.5 = .1638 runs

Based on the availability model and the marginal effect of the strength of a bullpen, using Zerpa today cost the team .1638 runs tomorrow



Final Calculation



Finally, to calculate our final manager stat, we add the runs gained from making the bullpen move with the runs lost from the next day



Net gain 1.80 Runs
Relief Appearance



Net loss .1638 Runs
Unavailable Next Day

Matt Quatraro



1.6362 Runs Gained

This Calculation was performed for every starting pitching change in 2025



Leaderboards



Manager	Team	Runs Saved/Game
Torey Lovullo	Diamondbacks	0.119
AJ Hinch	Tigers	0.101
Dave Roberts	Dodgers	0.094
Warren Schaeffer Bud Black	Rockies	0.085
Clayton McCullough	Marlins	0.077
Matt Quatraro	Royals	0.065
Kevin Cash	Rays	0.055
Alex Cora	Red Sox	0.050
Brian Snitker	Braves	0.047
Craig Albernaz	Orioles	0.044

Manager	Team	Runs Saved/Game
Joe Espada	Astros	-0.023
Mike Shildt	Padres	-0.021
Stephen Vogt	Guardians	0.006
Ron Washington Ray Montgomery	Angels	0.004
Dan Wilson	Mariners	0.016
Dave Martinez Miguel Cairo	Nationals	0.016
Terry Francona	Reds	0.017
John Schneider	Blue Jays	0.019
Pat Murphy	Brewers	0.019
Bruce Bochy	Rangers	0.026

Suppose for a moment...

- MLB wanted to implement the Golden Bat Rule – once per game, a team can send any hitter they want up to the plate
- Managers put in their best bat in bases loaded situations against the first reliever to enter the game – high leverage situation against who is likely not to be the closer
- Would the distribution of estimated runs change significantly?





In Reality:

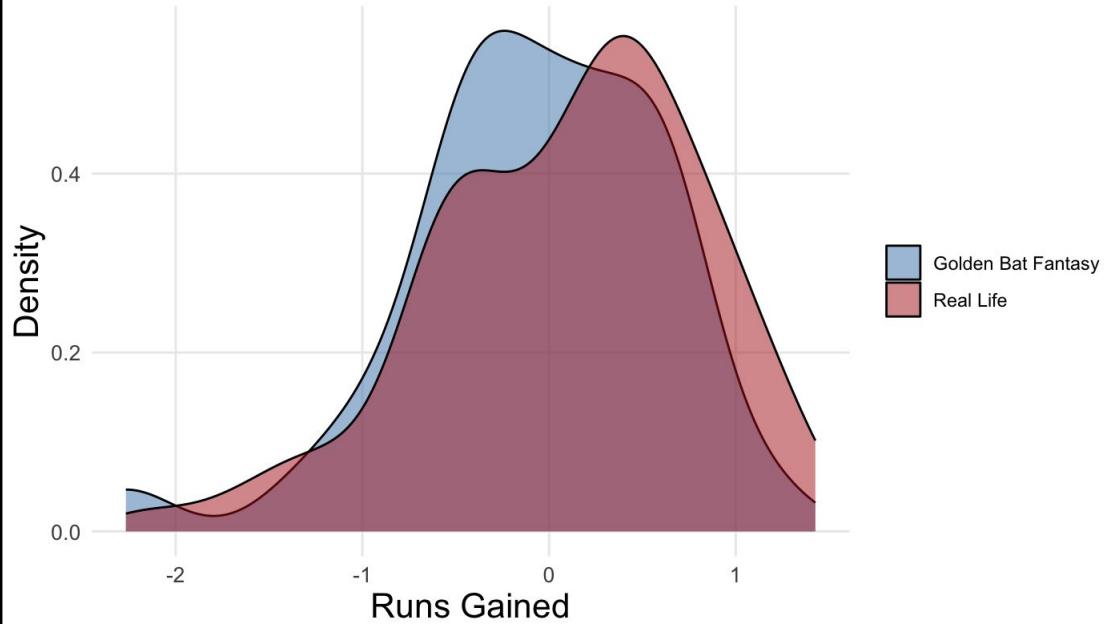
- Managers will use their higher leverage relievers earlier in the game
- Managers will wait for the Golden Bat to come out before going to the bullpen

Model Change:

- First hitter of relief pitcher has highest team xwOBA for bases loaded situations
- Impacts both the starting pitcher and relief pitcher in the model (weighted xwOBA)

Model change only performed on bases loaded opportunities

Distribution of Runs Gained



As expected: mean shifted to the left by .16 runs
negative runs on average gained



Biggest risers and fallers

Conclusions:

- Over the course of a season, managers have small positive impacts on the game through calls to the bullpen at the end of a start
- Bullpen moves directly impact future games in a statistically significant and quantifiable way

Limitations/Future Work:

- Strict assumptions made: time and complexity
- We rated existing manager pull decisions, we did not find an optimal pull time

