

# PITCHER DYNAMICS XPLAINED

**CREATING A MODEL THAT DETERMINES THE RIGHT TIME TO PULL A  
STARTING PITCHER**

Atul Venkatesh, Ishan Kinikar, Levon Sarian  
2024 SMT Data Challenge

# SETTING THE SCENE

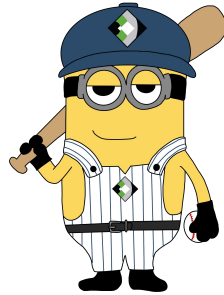


# A RECURRING ISSUE

**\*100 PITCH RULE\***



# DATA



# THE DATA WE USED

- Farm system data provided by SMT
- Focused on pitch-by-pitch data:
  - A unique game id
  - Home team
  - Visiting team
  - Day of game
  - Play and at-bat number in the game
  - Inning number and whether it was the top or bottom half
  - Unique player ids of everyone on the field



# PREPARING THE MODEL



# MODEL FACTORS

## Factors that were given:

- Batter id
- Level of farm system
- Baserunner situation  
(occupied/unoccupied)

## Factors that we created:

- Pitch count
- Pitches in inning
- Outs in inning
- Batting order
- **\*\*Difficulty\*\***  
(outs in inning + baserunner situation)

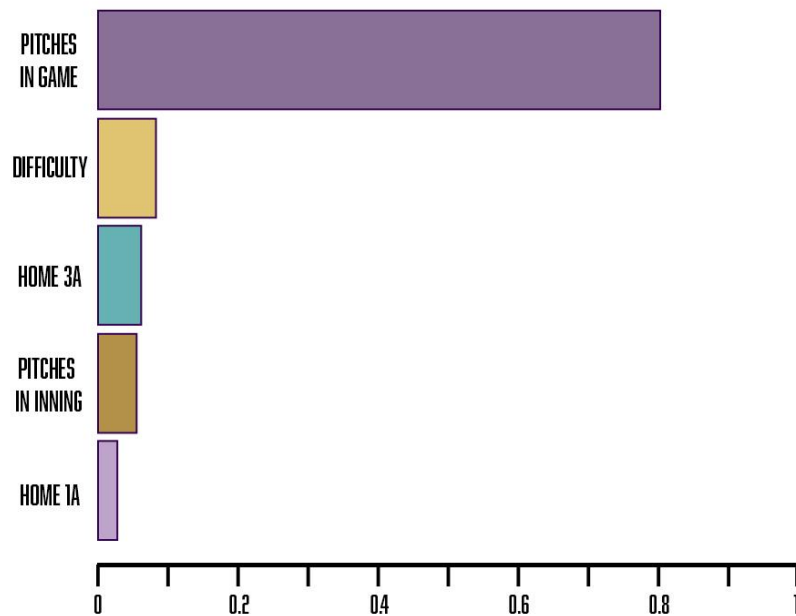
# MODEL PROPER





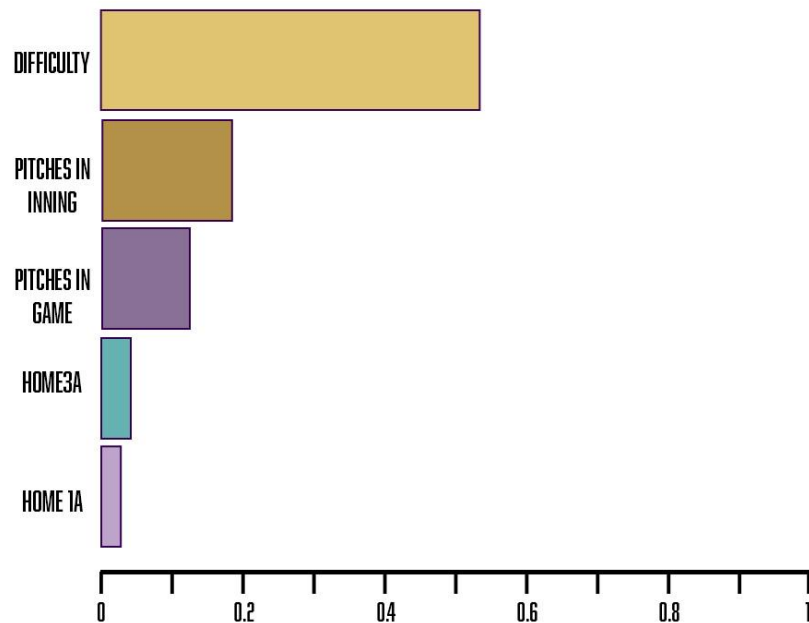
# EXPECTED PITCHES LEFT IN GAME (XPLG)

XBoost model 1: Based on factors like pitches thrown in the game, pitches thrown in the inning, baserunner situation, outs in inning, and batting order, how many pitches do we expect the pitcher to throw?



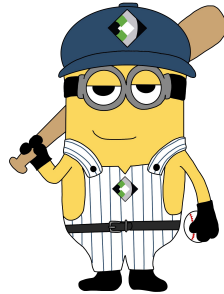
# EXPECTED PITCHES LEFT IN INNING (XPLI)

XBoost Model 2: Using the same factors as the xPLG model, how many pitches can we expect a pitcher to throw in the inning?



**When the xPLG is less than the xPLI, we know to pull the pitcher.**

# RESULTS AND LIMITATIONS



# PREFACE TO RESULTS

We can split the pulling data into three categories:

Pulling type	Description
Overworked	The coaching staff decides to leave the pitcher in too long, pulling him <b>after</b> we recommended them to.
Underworked	The coaching staff decides pull the pitcher too early, choosing to pull the pitcher <b>before</b> we recommended them to.
Appropriately used	The coaching staff chose to pull the pitcher at the <b>same time</b> as our recommendation.

# MODEL SUCCESS RATE

## How often does the model make the correct decision?

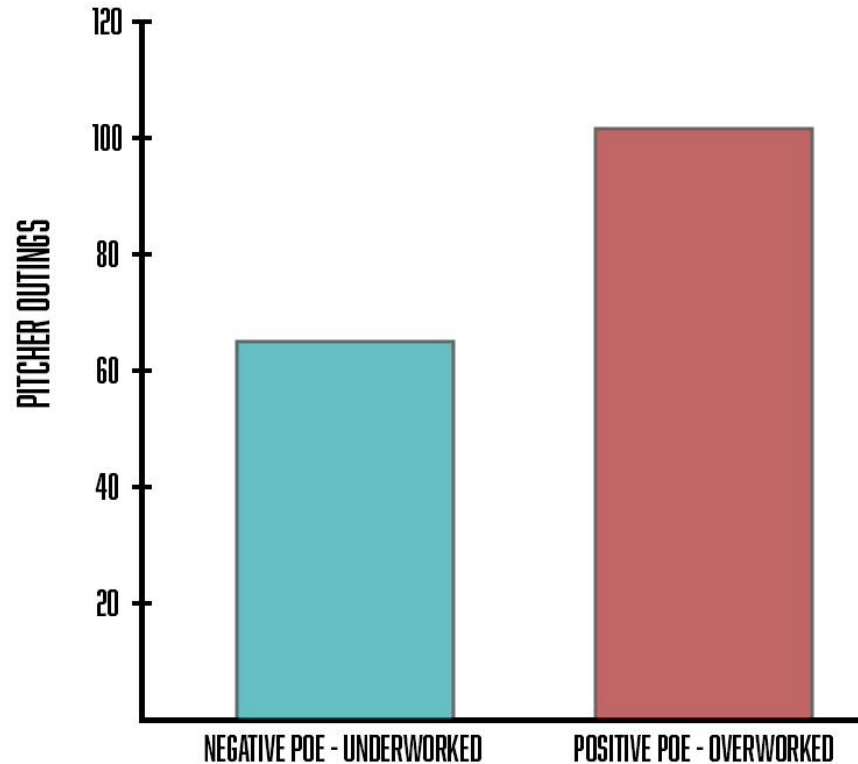
*Comparing model recommendation to manager's decision. Based on arm conservation and limiting pitch count.*

Situation	Count	Success rate
Overworked <sup>1</sup>	828	0.87
Appropriately used <sup>2</sup>	46	1.00
<b>Total</b>	<b>874</b>	<b>0.88</b>

<sup>1</sup> Pitcher gets pulled **after** our recommendation.

<sup>2</sup> Pitcher is pulled by the manager **at the same time** as the model's recommendation.

# PITCHES OVER EXPECTED (POE)



# ARE PITCHERS OVERWORKED IN THE MINORS?

## Which teams overwork their pitchers?

*Comparing each level of a farm system. A positive number indicates pitchers tend to be overworked.*

Team Name	Pitches Over Expected
Home1A	3.621
Home2A	3.825
Home3A	2.043
Home4A	0.144




# APPLICATIONS



# THE SNELL TOOL

- Built using React.js
- User inputs game state variables:
  - Pitch count
  - Pitches in inning
  - Outs in inning
  - Base runner situation
  - Hitter number
- Also inputs two qualitative variables:
  - Pitcher Confidence
  - Pitcher Fragility
  - These variables are based on the user's confidence in the pitcher and how much they want to push the pitcher physically and mentally.

# SNELL TOOL USER INTERFACE

Snell Tool 

Providing recommendations for when to pull a starting pitcher based on situational and qualitative factors.

Pitches Thrown

Pitches in Inning

Enter pitches thrown

Enter pitches in inning

Outs in Inning

Enter number of outs in inning

Base 1

Base 2

Base 3

Unoccupied

Unoccupied

Unoccupied

Hitter Number

Enter current position in the batting order (1-9)

Pitcher Confidence

How confident the staff is in the pitcher - based on recent form, how the pitcher is currently pitching, and the situation of the team's relief pitching


3 - Neutral

Pitcher Fragility

How worried the staff is about the pitcher's health - based on expected durability, past injuries suffered, and how concerned the staff is with preserving the pitcher's health

3 - Neutral

Submit



2024 SMT Data Challenge

\*After submitting, scroll to view results\*



Submitted Data:

Pitches Thrown: 60

Pitches in Inning: 2

Outs in Inning: 1

Baserunner Situation: 000

Hitter Number: 3

Pitcher Confidence: 3

Pitcher Fragility: 3

Recommendation: Keep the pitcher in

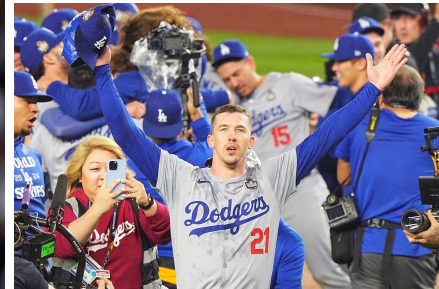
[snelltool.com](https://snelltool.com)

# FUTURE OF THE SNELL TOOL



- Managerial decisions - Provides context and statistical backing for when to pull the pitcher
- Coaching evaluation tool
  - If Snell Tool decisions consistently underperform due to qualitative inputs, it might be a sign of a below average understanding of player
- More data, new features, and advanced player tracking would significantly improve the model

# FINAL WORDS



# LOOKING BACK AT THE WORLD SERIES



## Gerrit Cole - Game 5

After the T. Hernandez double in the 5th inning

### Submitted Data:

Pitches Thrown: 79  
Pitches in Inning: 29  
Outs in Inning: 2  
Baserunner Situation: 010  
Hitter Number: 5  
Pitcher Confidence: 3  
Pitcher Fragility: 3

Recommendation: Keep the pitcher in

After pitching 6 full innings, going into the 7th inning

### Submitted Data:

Pitches Thrown: 98  
Pitches in Inning: 0  
Outs in Inning: 0  
Baserunner Situation: 000  
Hitter Number: 1  
Pitcher Confidence: 3  
Pitcher Fragility: 3

Recommendation: Pull the pitcher

After pitching 6 full innings, **high confidence + extremely healthy**

### Submitted Data:

Pitches Thrown: 98  
Pitches in Inning: 0  
Outs in Inning: 0  
Baserunner Situation: 000  
Hitter Number: 1  
Pitcher Confidence: 5  
Pitcher Fragility: 1

Recommendation: Keep the pitcher in



# LOOKING BACK AT THE WORLD SERIES



## Walker Buehler - Game 3

### Submitted Data:

Pitches Thrown: 76  
Pitches in Inning: 0  
Outs in Inning: 0  
Baserunner Situation: 000  
Hitter Number: 1  
Pitcher Confidence: 3  
Pitcher Fragility: 3

Recommendation: Keep the pitcher in

### Submitted Data:

Pitches Thrown: 76  
Pitches in Inning: 0  
Outs in Inning: 0  
Baserunner Situation: 000  
Hitter Number: 1  
Pitcher Confidence: 1  
Pitcher Fragility: 3

Recommendation: Pull the pitcher



## Jack Flaherty - Game 5

### Submitted Data:

Pitches Thrown: 35  
Pitches in Inning: 9  
Outs in Inning: 1  
Baserunner Situation: 100  
Hitter Number: 4  
Pitcher Confidence: 3  
Pitcher Fragility: 3

Recommendation: Keep the pitcher in



# ACKNOWLEDGEMENTS



- Huge thanks to Meredith Wills, Kenny Bores, and Andrew Steenkamer and the rest of the SMT staff for all of their help throughout this entire project
- Thank you to SMT for giving us the opportunity to come to North Carolina and talk about our research
- Thank you to the judges for taking time out of your days to watch our presentation
- Thank you to Atul's little brother Akhil for the cool drawings



# THANKS FOR WATCHING



**EXTRA  
SLIDES**

# **FAMOUS OUTINGS**

# BLAKE SNELL

## Submitted Data:

Pitches Thrown: 73

Pitches in Inning: 4

Outs in Inning: 1

Baserunner Situation: 100

Hitter Number: 1

Pitcher Confidence: 3

Pitcher Fragility: 3

Recommendation: Keep the pitcher in

## Submitted Data:

Pitches Thrown: 73

Pitches in Inning: 4

Outs in Inning: 1

Baserunner Situation: 100

Hitter Number: 1

Pitcher Confidence: 1

Pitcher Fragility: 3

Recommendation: Pull the pitcher

# MATT HARVEY

## Submitted Data:

Pitches Thrown: 99

Pitches in Inning: 0

Outs in Inning: 0

Baserunner Situation: 000

Hitter Number: 3

Pitcher Confidence: 5

Pitcher Fragility: 1

Recommendation: Keep the pitcher in

## Submitted Data:

Pitches Thrown: 99

Pitches in Inning: 0

Outs in Inning: 0

Baserunner Situation: 000

Hitter Number: 3

Pitcher Confidence: 3

Pitcher Fragility: 3

Recommendation: Pull the pitcher

**OUT COUNTER**

# OUT COUNTER - GENERAL

- If the game state remains unchanged from one at-bat to another, i.e. the next baserunners are the same as the current baserunners, then we are confident an out occurred. The only exception is a solo home run, but we will correct for that later.

13

- If the batter does not show up on base in the next at-bat, unless he hit a home run, then we know he got out.

# OUT COUNTER - DOUBLE PLAYS

## Coding for simple double play:

If there is a man on first and no one on second and third and there are no baserunners in the next at-bat, then we know a double play occurred.

- Coding for simple fielder's choice: If there is only a man on first and only a man on first in the next at-bat, but the ID of the next first baserunner matches the ID of the hitter, then an out occurred.

## Coding for complex double plays:

- Runner on first and second, the second baserunner advances to third base, but first and second base are empty in the next at-bat. We know a double play occurred here.

- Runner on first and third and the next at-bat, the bases are empty. Here we would assume that the runner on third scored, and the first baserunner and hitter were thrown out. We can be reasonably confident that this is the case.

- Bases are loaded and the next at-bat, the runner on second is now at third, and the bases are empty. Here we can also assume that the runner on third scored, and the first baserunner and hitter were thrown out.



# OUT COUNTER - FINAL METRICS

- When the home team pitches, the last play of the inning is always an out. If there is one out before the last play of the inning, we know an inning-ending double play occurred. If there are no outs before the last play of the inning, we know a triple play occurred. This provides some more accuracy to the model.
- Adding in home runs: SMT provides data that tells which at-bats were home runs. We merged this information with our existing data. In situations where a home run occurred, we made sure our model could recognize that there was no out on the play.

# **XGBOOST MODEL**

# XGBOOST ALGORITHMS

XGBoost algorithms serve as a method to understand the relationships within data and use those relationships to inform future predictions

- First step: Wrangling the data
  - We isolate all of the variables that we think are important: level in the farm system, current pitch count, pitches in inning, situation difficulty, and position in batting order
- Next step: Make sure all non-numeric variables would be accepted by the algorithm. This means converting character variables like level of farm system and hitter number to "factor" variables. This ensures that values are split into categories rather than numbers.
- Next step: split the data into training and testing. We decided to do an 80/20 split respectively.
- Conducting the algorithm: We used the xgboost package in R and specified the dimensions. With a combination of five-fold cross validation and trial and error, we determined the best possible dimensions to reduce root mean-squared error.
  - xPLG model: the number of rounds was 134, the number of early stopping rounds was three, the max depth was six, and the eta was 0.18.
  - xPLI model: the number of rounds was 230, and the number of stopping rounds, max

# **MICRO-LEVEL ANALYSIS**

# MICRO-LEVEL ANALYSIS - P1

- Home 1A vs. Vis1AL, Day 18 Pitcher pitched three full innings. We recommended to pull the pitcher out before the beginning of the 4th inning. What happened? The pitcher got one out and then proceeded to load the bases. The manager eventually pulled the pitcher at 1 out with the bases loaded.
- Home 1A vs. Vis1AU, Day 77 The pitcher is near the beginning of the fourth inning. Man on first, no outs. Already gave up one run in the innings. We say pull the pitcher. What happens? Next batter goes on base. Allows another base hit, but the runner on first gets out stretching to third. Ends up getting out of the inning. But some damage has been done. Interestingly enough, two pitches after our recommendation, we no longer recommended the pitcher would be pulled. However, five or so pitches after our first recommendation, we once again recommended the pitcher to be pulled.
- Home1A vs. Vis1AY, Day 109 - The pitcher has pitched four full innings. We recommend the pitcher to be pulled now. The pitcher proceeds to two base runners on base then gives up a double, allowing for a run to score. While the pitcher eventually gets two outs, he also allows another base hit and then another run. Eventually gets pulled out with two outs and men on first and third.

# MICRO-LEVEL ANALYSIS - P2

- Home 2A vs. Vis2AF, Day 11 - Pitcher goes 6 full innings. We recommend to pull him before the first pitch in the 7th inning. The pitcher is facing the cleanup hitter. The pitcher gets an out before letting the next two hitters get on base. Now the manager decides to pull him out.
- Home2A vs. Vis2AG, Day 15 - Pitcher pitched five full innings. We recommended the pitcher be pulled before the beginning of the 6th inning. The pitcher proceeded to record one out, give up a double, and get pulled mid-inning.
- Home2A vs. Vis2AI, Day 28 - Pitcher goes 5 full innings. We recommend pulling him before the 6th inning. Proceeds to allow the first three batters to reach base without getting an out. Gets pulled with the bases loaded and zero outs.
- Home2A vs. Vis2AT, Day 118 - The pitcher is midway through the fourth inning, man on first and second with no outs. We recommend the pitcher be pulled. The pitcher proceeds to give up a double and a single, letting up two runs before finally getting 1 out and getting pulled with men on second and third. After our recommendation, the pitcher let in multiple runs, left with baserunners on, and only got one out.

# MICRO-LEVEL ANALYSIS - P3

- Home 3A vs. Vis3AK, Day 38 - The pitcher pitched six full innings. We said to pull the pitcher 8 pitches into the 7th inning. The pitcher proceeds to give up two doubles and a single, letting up one run and only getting one out before he is pulled out of the game.
- Home 3A vs. Vis3AO, Day 78 - The pitcher is seven pitches into the fifth inning with a man on first and one out. We recommend pulling the pitcher. The pitcher proceeds to get one more out before allowing the next two hitters to reach base. With the bases loaded, the pitcher allows a bases clearing triple before getting pulled out.
- Home3A vs. Vis3AV, Day 140 - The pitcher is midway through the second inning and in a bit of a jam. Man on first and second with one out in the inning. We say pull the pitcher. What happens next? The next baserunner gets on base so the bases become loaded. The last batter the pitcher faces also gets on base, driving in a run. After our recommendation, the pitcher let two batters on base, giving up a run, and not getting an out.
- Home4A vs. Vis4AO), Day 27 - The pitcher pitched five full innings before giving up a leadoff single. We then recommended the pitcher be pulled. The pitcher proceeds to let up a two run home run.

# MICRO-LEVEL ANALYSIS - P4

- Home 4A vs. Vis4AV, Day 78 The pitcher is 11 pitches into the inning. No outs, man on second and third. We say pull the pitcher, but they wait one at bat. The batter ends up getting on base. With the bases loaded and nobody out, the starter gets pulled.
- Home 4A vs. Vis4AY, Day 104 - The pitcher has pitched six full innings. we recommend pulling the pitcher out before the 7th inning. What happened? Pitcher got two outs, let the third hitter get on base, and then gave up a hit to the fourth hitter. Gets pulled out, crossing 100 pitches. Leaves a man on first and third for the reliever.
- Home 4A vs. Vis4BD, Day 136 Pitcher pitched five full innings. We recommended pulling the pitcher before the start of the 6th inning. The pitcher had a rough 6th inning, facing 5 batters. While the pitcher did not give up any runs, the inning took 30 pitches, putting his final pitch count over 100



# PREFACE TO RESULTS

How do we evaluate how our model performed?

- For the pitchers who were overworked, if they threw more pitches than our model predicted (xPLG), then we considered this a success.
- For the pitchers who were appropriately used, we counted this as a success for ourselves.
- Evaluating the underworked pitchers is a little more complicated and we didn't have the data to do an accurate analysis.

# COMPARING OUR MODEL TO MANAGERS

## DECISIONS?

Recommendations were split into three categories:

- Overworked: the managers leave the pitcher in too late and choose to pull the pitcher **after** we recommended them to.
- Underworked: the managers pull the pitcher too early, choosing to take the pitcher out **before** we recommended them to.
- Appropriately used: the manager chose to pull the pitcher at the **same** time as our recommendation.



# EVALUATING COACHING USING THE SNELL TOOL

- Qualitative factors: **confidence in pitcher**, usage/injury constraints, command issues
- Example:
  - Coaching staff uses the Snell Tool to decide to keep a starter in relatively late
  - If the pitcher lets up several runs, the coaching staff would be partly to blame
- If Snell Tool decisions consistently underperform due to qualitative inputs, it might be a sign of a below average understanding of player

