Interesting Patterns of Home Run Hitting

Jim Albert
Emeritus Professor
Bowling Green State University

February 2023

BGSU

Outline

- Introduction – my background and interests in baseball
- Baseball data (Statcast)
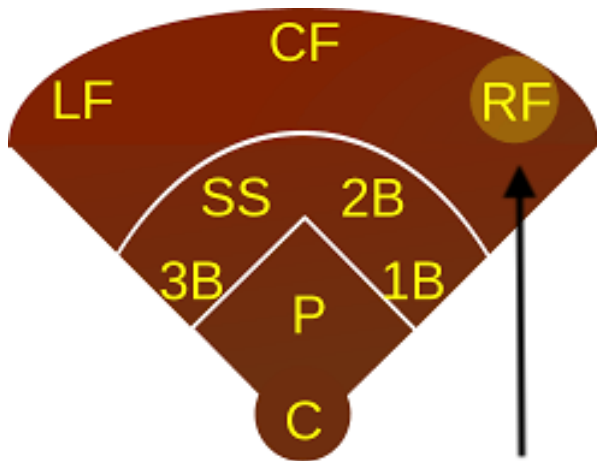- Home run hitting – changes over the Statcast era

Introduction

Background

Grew up around Philly ...

- Played Little League Baseball
- Followed the Phillies
- Collected baseball cards (liked the stats on the back of the cards)
- Played baseball simulation games (Strat-O-Matic)

What Position Did I Play?

My Academic Career

- Math major
- Doctorate in Statistics
- Professor at BGSU for 41 years, retired in 2020
- Interests in Bayesian modeling, stats education and applications of statistics in sports

My Baseball Books

- *Curve Ball* (with Jay Bennett)
- *Teaching Statistics Using Baseball*
- *Visualizing Baseball*
- *Analyzing Baseball Data with R* (with Max Marchi and Ben Baumer)

Retirement

- Play a lot of tennis
- Keep active in statistics
- Interested in baseball and tennis analytics
- Enjoy writing Shiny applications in R

Shiny - Building Web Interfaces with R

- Have a R package *ShinyBaseball* at
  `https://github.com/bayesball/ShinyBaseball`
- Book `https://bayesball.github.io/shinybaseball/`
  describing my Shiny baseball applications
- Great way to communicate findings

A Shiny App - Pitch Locations and Pitch Values
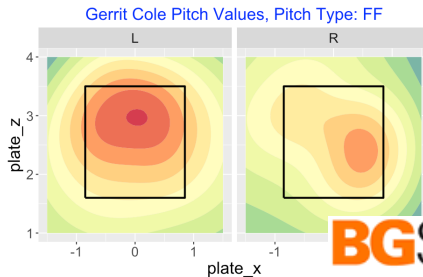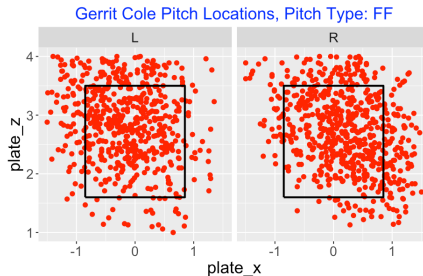
Exploring Baseball Data with R

Demise of the Sacrifice Hit

By Jim Albert on March 28, 2022 | Leave a comment | Edit

Introduction

There are two significant rule developments in Major League Baseball for the 2022 season. First, games will continue to use the so-called zombie runner rule where each half-inning in an extra inning will start with a runner on second base. Second, MLB will implement an universal

BGSU.

Baseball Data

Baseball Data

- Remarkable data is available in baseball
- Lahman database (season to season)
- Retrosheet (game data, play-by-play data)
- Statcast

BGSU.

Best Baseball Stats Sites

- Baseball Reference
- FanGraphs
- Baseball Savant

Support them!

Statcast

- Every movement of players and the ball is recorded
- Data on each pitch (velocity, movement, spln, location)
- Data on the result of bat (launch angle, exit velocity, spray angle)
- Some data is available through Baseball Savant
- Using this data to study home run hitting
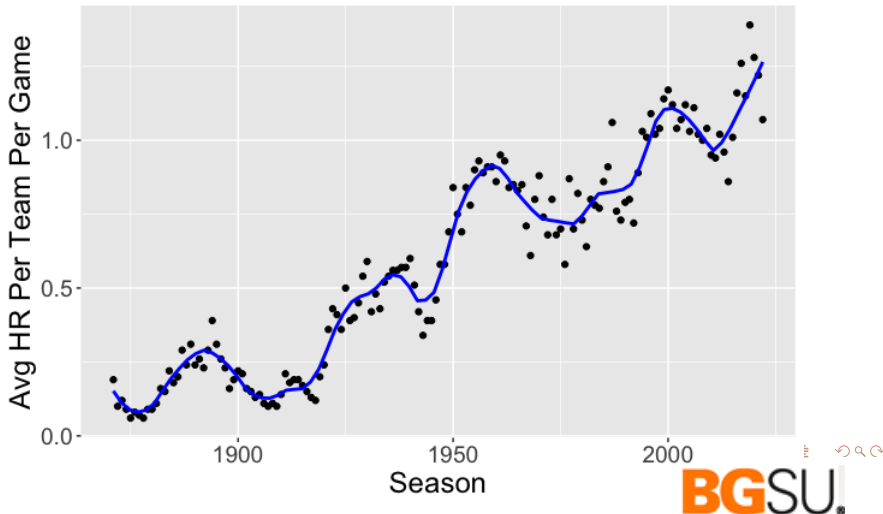
New Measures of Performance

- Catcher framing
- Catcher pop time
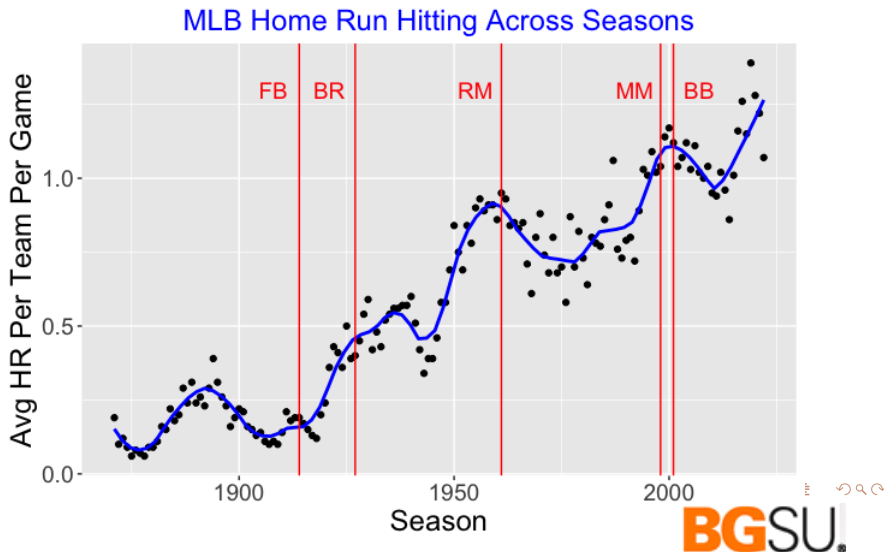- Runner speed
- Fielding catch probability
- Barrels

Home Run Hitting

History of Avg HR Hit Per Team Per Game
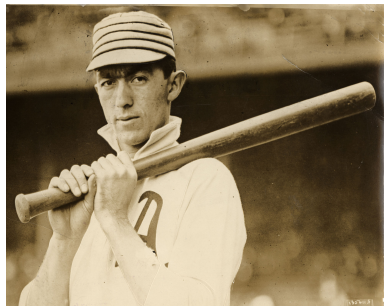
Seasons of Famous HR Hitters

MLB Home Run Hitting Across Seasons

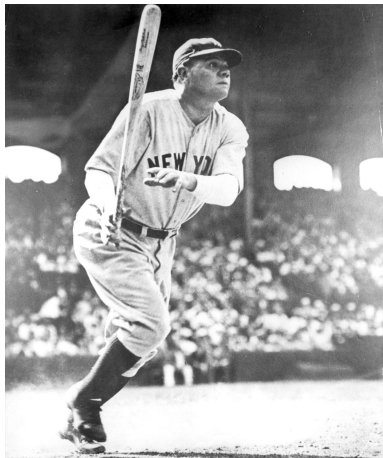# Some Famous Home Run Hitters and Seasons

- Frank (Home Run) Baker (1914)
- Babe Ruth (1927)
- Roger Maris (1961)
- Mark McGwire (1998)
- Barry Bonds (2001)

Home Run Baker

- Played during Deadball Era
- Was home run leader in 1914 with 9 HR
- Home runs were not a big part of the game

Babe Ruth

- With Babe, home runs became a big part of baseball
- Was home run leader in 1927 with 60 HR
- Maybe the greatest player of all time

Roger Maris

- Broke Ruth's record with 61 HR in 1961
- Played with Mickey Mantle
- Some controversy with record (famous asterisk)

Mark McGwire

- Played during "Steroids Era"
- Competed with Sammy Sosa in 1998 for the HR crown
- Hit 70 to break the record

Barry Bonds

- Played during "Steroids Era"
- Broke the HR record with 73 in 2001
- Career HR leader with 762

Increasing (and Decreasing) Home Run Totals in the Statcast Era

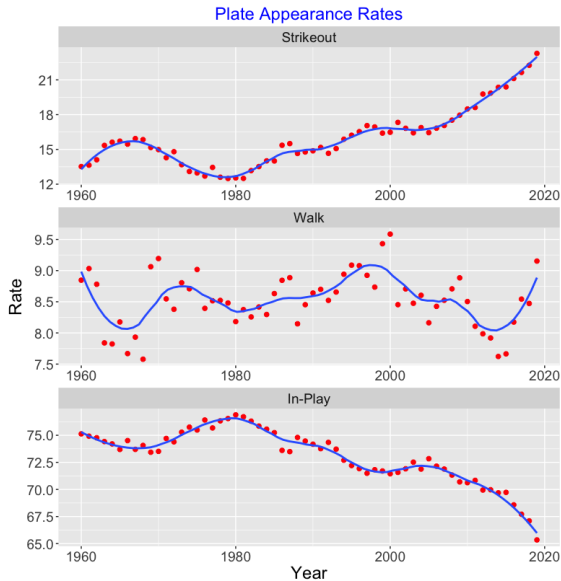| Season	HR Total
2015	4909
2016	5610
2017	6105
2018	5585
2019	6776
2021	5944
2022	5215

BGSU

A Plate Appearance

Three Basic Outcomes

- Strikeout
- Walk
- Ball is put in-play

How have the rates of the three basic outcomes changed over the last 50 years of baseball?

Historical Change in SO, BB, In-Play Rates
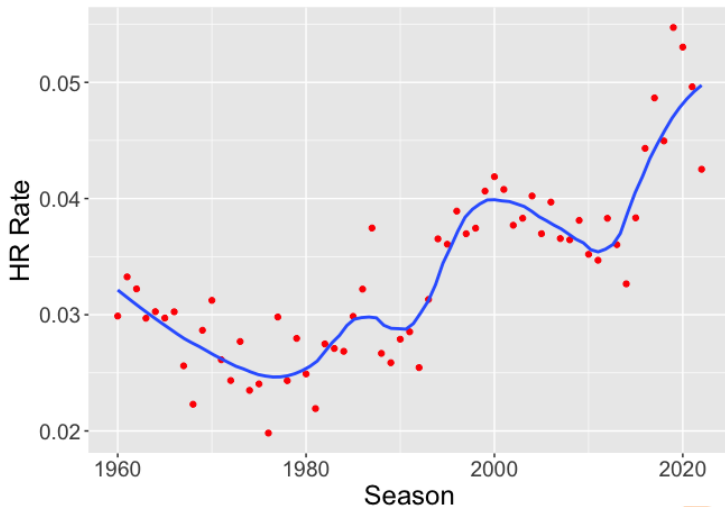
In-Play Home Run Rate

- Define the home run rate as the fraction of HR among all batted balls ($AB - SO$).

$$HR\ Rate = \frac{HR}{AB - SO}.$$

- Look at history of home run rates in last 60 years.

History - Home Runs per Batted Ball

MLB Home Run Report

- In Fall of 2017 a committee was charged by Major League Baseball to identify the potential causes of the increase in the rate at which home runs were hit in 2015, 2016, and 2017.
- Report was released in May 2018. (Available online.)

Reasons for the increase in home run hitting

The batters?

- Changes in characteristics of batted balls
- Launch angle, exit velocity, and spray angle

The pitchers?

- Changes in types of pitches
- Pitch location

Reasons for the increase in home run hitting

The ball?

- Changes in how the ball is made?
- Seam height, core?
- Drag coefficient?

Reasons for the increase in home run hitting

Game conditions?

- Ballpark effect
- Weather
- Cold vs. hot temperatures

Process of Hitting a Home Run

- **IN-PLAY**: Have to put the ball in play
- **HIT IT RIGHT**: The batted ball needs to have the "right" launch angle and exit velocity
- **REACH THE SEATS:** Given the exit velocity and launch angle, needs to have sufficient distance and height to clear the fence (the carry of ball)

Committee's Findings from 2015 - 2017 Data

- We found modest changes in launch angle and exit velocity among batters
- Focused on RED zone – launch angle in (15, 40) degrees, launch speed between 90 and 115 mpg
- The RED zone balls are showing more carry – they travel further

Committee's Findings from 2015 - 2017 Data

- Increase in home runs is due to better carry (less drag) for given launch conditions
- Likely due to the aerodynamic properties of the baseball
- Didn't appear to be a property of the manufactured baseballs
- Recommend that MLB monitor the climate environment of the baseballs

End of 2022 Season

- Eight seasons of Statcast data (2015 - 2022) are available
- Have launch speed and launch angle measurements for all seasons
- Take a broader perspective on home run hitting

Empirical Approach

- Look at region of launch angle and exit velocity where most of home runs are hit
- Look at **rate of batted balls** in this region – how does it vary by season?
- Look at **rate of home runs** for balls hit in this region – how does it vary season?

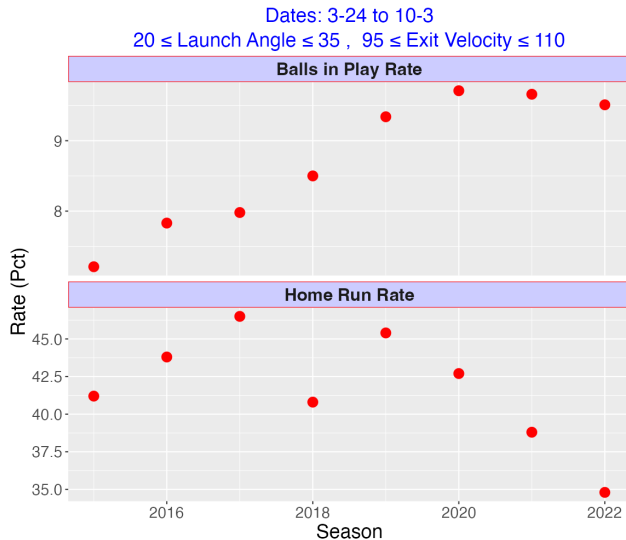# Focus on Region of Launch Angle and Launch Speed where Home Runs are Hit
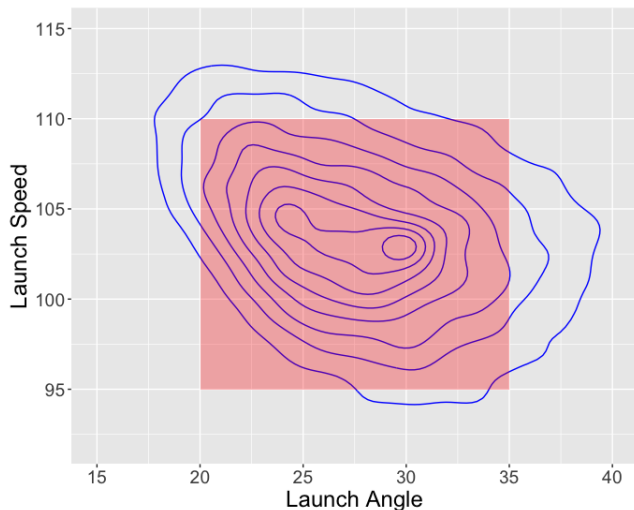
Balls in Play Rate

- Interested in rate of "home run likely" batted balls

$$BIP\ Rate = \frac{HR\ Likely}{BIP}$$

- Are batters changing their approach?
- Players getting stronger?

Balls in Play Rate

Balls in Play Rate (top panel)

- See a general increase in "home run likely" rates over Statcast period
- Players appear to be changing their hitting approach or they are getting stronger

Focus on Changes for Individual Batters

- Look at all hitters who had at least 200 batted balls in 2015 and 100 batted balls in 2019
- Collect fraction of hard-hit/good launch angle balls for each hitter
- How did hitters change in this period?

Scatterplot of Fraction of Hard Hit Balls for 2 Seasons – 75% Had Higher Fraction in 2019
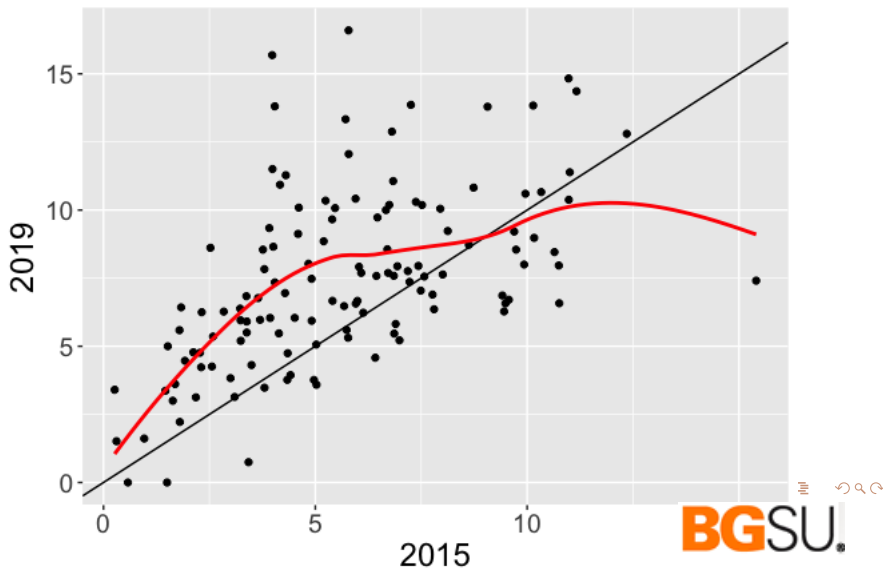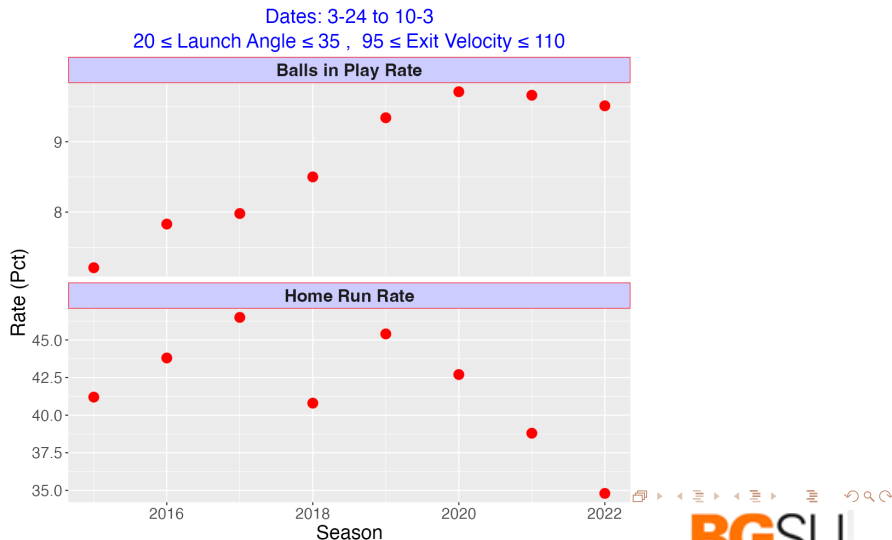
Home Run Rate

- What is the chance of a home run given good values of launch angle and exit velocity?

$$HR\ Rate = \frac{HR}{HR\ Likely}$$

- Characteristic of the baseball
- Changes in drag coefficient over seasons?

BGSU.

Home Run Rate (bottom panel)

Dates: 3-24 to 10-3
20 ≤ Launch Angle ≤ 35 , 95 ≤ Exit Velocity ≤ 110

Home Run Rate

- General increase from 2015 to 2017
- Big dip in 2018, followed by big increase in 2019
- Steady decrease from 2019 to 2022

Interpretation?

Shiny App LogitHomeRates

- Divide region of launch angle and exit velocity into subregions
- Choose two seasons
- Look at changes in batted ball rate and home run rate in each region
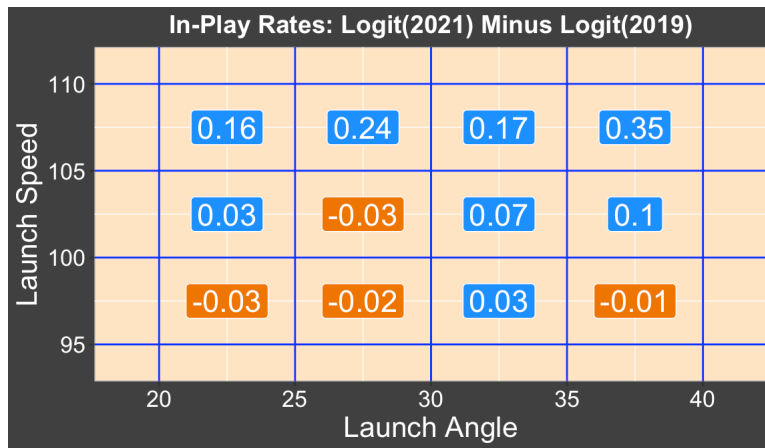- Using logit scale

What is a Logit?

- If P is a rate, then the logit of P is

$$logit\ P = \log\left(\frac{P}{1-P}\right)$$
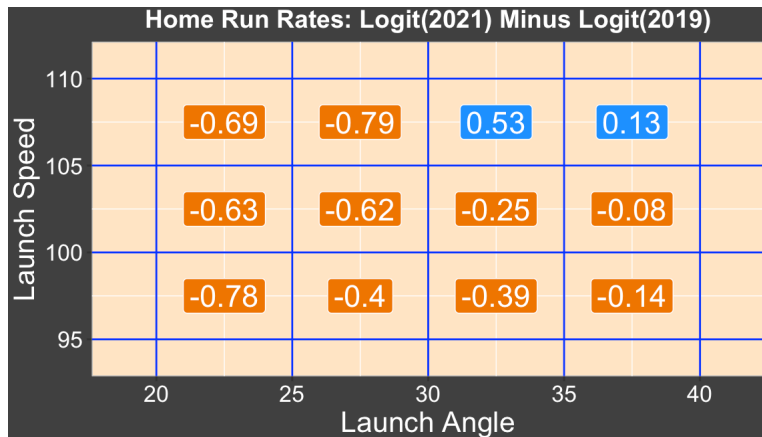
- When comparing two rates, say P_1 and P_2, better to compare on logit scale:

$$\log\left(\frac{P_2}{1-P_2}\right) - \log\left(\frac{P_1}{1-P_1}\right)$$

BGSU.

Batted Ball Rates - Comparing 2019 and 2021 Seasons

In-Play Rates: Logit(2021) Minus Logit(2019)

Home Run Rates - Comparing 2019 and 2021 Seasons

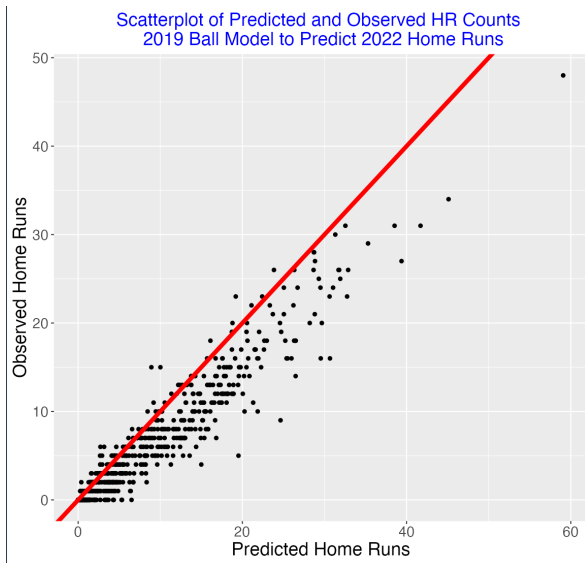Home Run Rates: Logit(2021) Minus Logit(2019)

Comparing 2019 and 2021

- Some increase in batted ball rates for high speeds and good launch angles
- Big decrease in home run rates
- Dead ball was used in 2021
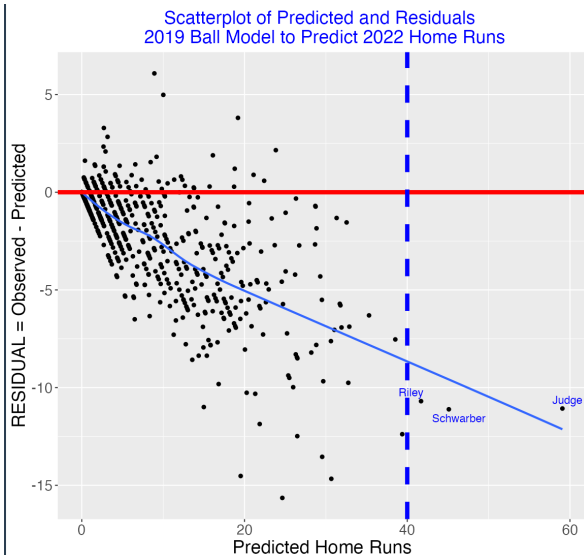- What happened in 2022?

Comparing 2019 and 2022 at Individual Batter Level

- Use model on 2019 data to predict home runs based on launch angle and exit velocity.
- For 2022 data (through August 23) predict probability of HR based on launch variables.
- Compare predicted (based on 2019 ball model) and observed.

BGSU.

Compare Predicted and Observed (Through August 23)

Scatterplot of Predicted and Observed HR Counts
2019 Ball Model to Predict 2022 Home Runs

Residual = Observed minus Predicted Plot (Through August 23)

Scatterplot of Predicted and Residuals
2019 Ball Model to Predict 2022 Home Runs
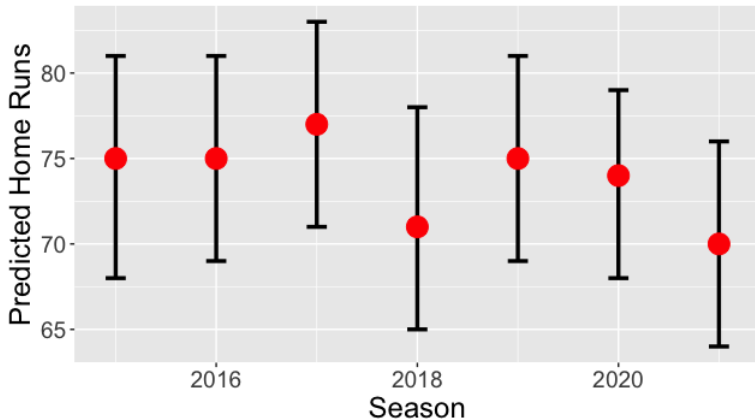
Riley
Schwarber
Judge

Aaron Judge

- Judge had a record 62 home runs hit in 2022.
- If he had been playing with a ball from previous seasons, he would hit in the 70's.

Home Run Predictions for Judge

90% HR Prediction Intervals for Judge
Using Different Season Balls

Next to Last Slide

- Great opportunities for baseball research
- Data is readily available
- Using R, easy to do one own's study
- Internships and jobs are available in MLB
- New opportunities in other sports

Paper and contact info

My baseball research:
https://bayesball.github.io/BLOG/

Blog: baseballwithr.wordpress.com

email: albert@bgsu.edu

Twitter: @albertbayes

BGSU.

Questions?