Maximizing Baseball Run Creation

Statistical Learnings of Batted Ball Data for Make Benefit Glorious Game of Baseball

The Question

Can we use a player's batted ball profile to determine their ability to create runs?

How do we determine a player's ability to create runs?

wRC+ (Weighted Runs Created Plus) determines how a player performs at the plate compared to his peers normalized at 100

What is batted ball profile?

Direction

- Pull 0.14
- Center -0.11
- Opposite -0.12

Strength

- Soft
- Medium
- Hard 0.5
  - Exit Velocity
  - Launch Angle

Type

- Groundball
- Line Drive 0.13
- Flyball
  - Home Runs 0.59

What is batted ball profile?

Non-Contact

- Walks
- Strikeouts

Most important features

Exit Velocity

Launch Angle

Home Runs per Flyball

Line Drive Percentage

Walk Percentage

Strikeout Percentage

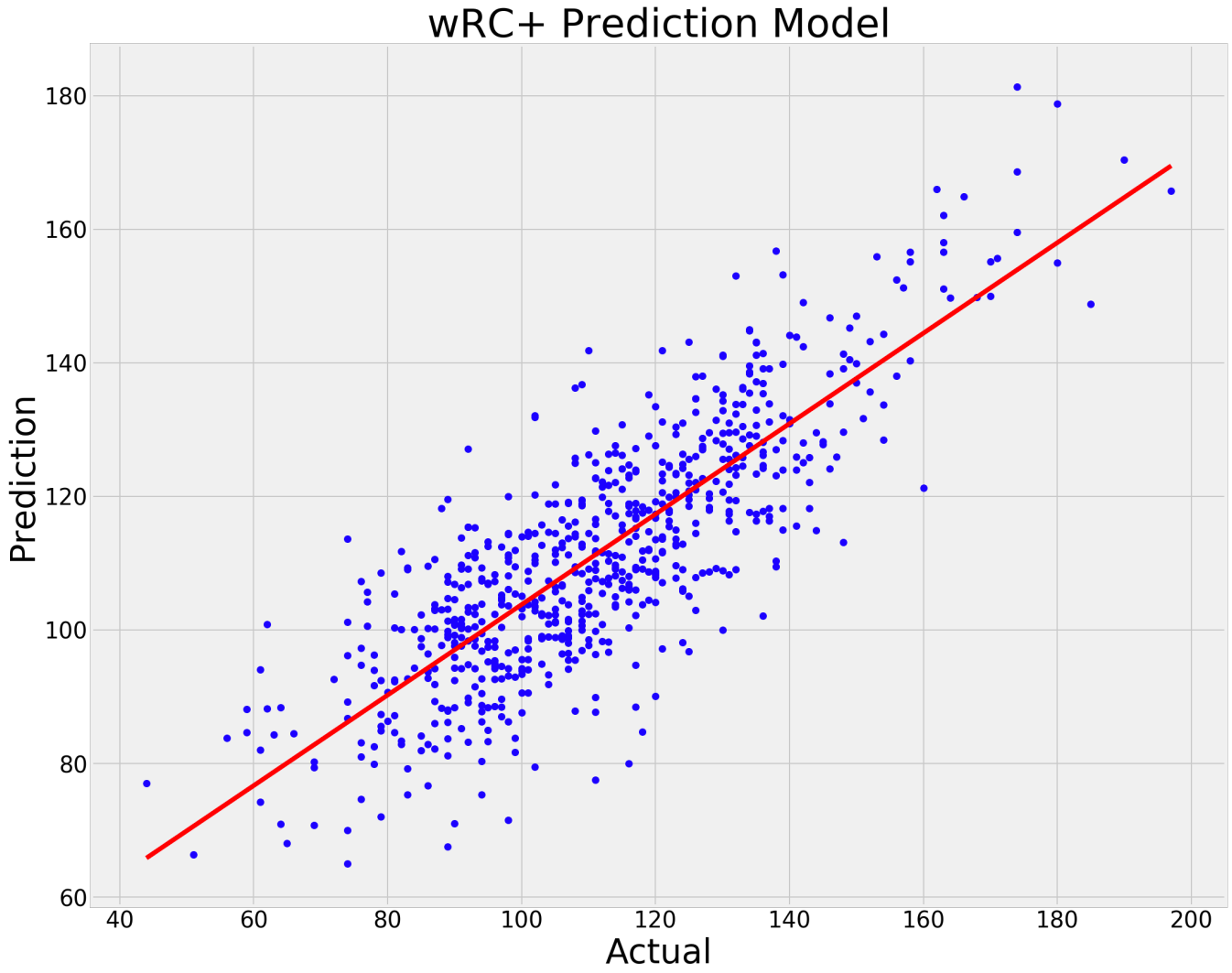
Model

$R^2 = 0.678$

$P < 0.05$

RMSE = 12.85

MAE = 10.14



What kind of players create the most runs?

1

Hit the ball hard

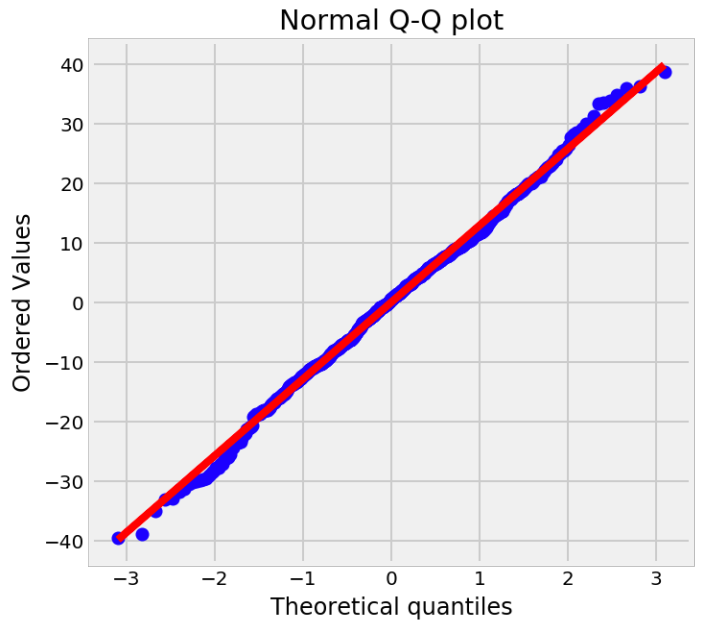
2

Draw more walks

3

Strike out less

Issues



Model has trouble at the extremes

High variance

Future

Remove outliers

More advanced regression

Q&A