



Predicting a Pitcher's Success with Stuff+

Michael Brennan Rosemary O. Ogbumba Lance Oubre
Zachary Jermain Austin C. Louque

Department of Mathematics, Louisiana State University, Baton Rouge,
Louisiana

Faculty Advisors: Dr. Nadejda Drenska and Dr. Peter Wolenski



Motivation

An effective softball pitcher mixes speed, movement, location, and pitch type to keep batters from making productive contact. Traditional results such as ERA describe what happened, but they do not isolate the intrinsic quality of an individual pitch.

Stuff+ grades a pitch using only its measured physical characteristics. Our goal was to build softball-specific Stuff+ models that can evaluate pitches, pitch types, and complete pitcher arsenals on a common scale.

By focusing on physical pitch characteristics, Stuff+ aims to separate a pitcher's underlying arsenal quality from defense, opponent quality, and short-term variation.

Data and Pitch Types

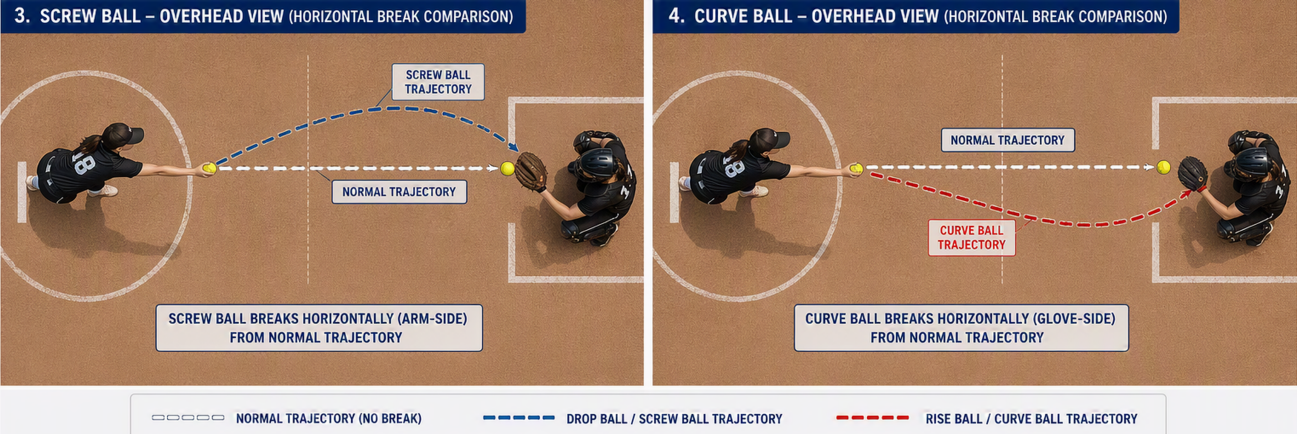
TrackMan combines radar and synchronized optical cameras to record more than 120 measurements for each pitch. Our 2026 dataset contains more than **250,000 pitches from over 1,000 games**, representing approximately 40 collegiate programs.

The analysis centers on the movement profiles of four common pitch types:

- **Rise ball:** upward trajectory; approximately 12-o'clock rotation.
- **Drop ball:** downward trajectory; approximately 6-o'clock rotation.
- **Curve ball:** glove-side movement for a right-handed pitcher.
- **Screw ball:** arm-side movement for a right-handed pitcher.



Vertical-break pitch types



Horizontal-break pitch types

Model Inputs and LightGBM Design

We trained two gradient-boosted tree models from the same nine TrackMan features, making their grades directly comparable:

Release	speed, height, side, extension
Movement	induced vertical break, horizontal break
Flight	vertical and horizontal approach angle
Rotation	spin rate

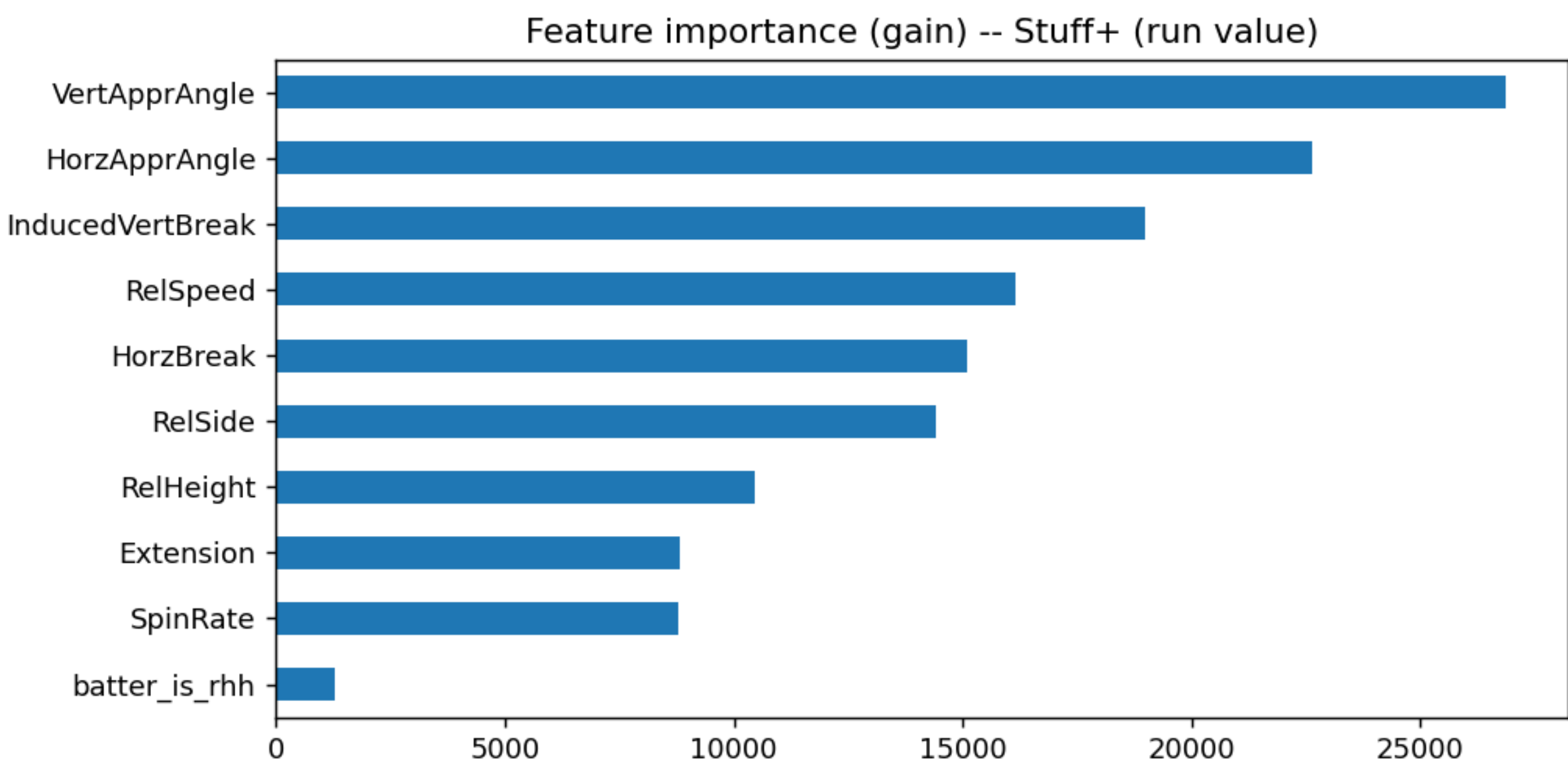
- **RV+:** a regression model fit to per-pitch run value, measuring run-prevention quality.
- **CSW+:** a binary classifier fit to called strikes plus whiffs, measuring strike-generating quality.

Together, the scores distinguish pitches that suppress runs from pitches that most reliably generate called strikes or swings and misses.

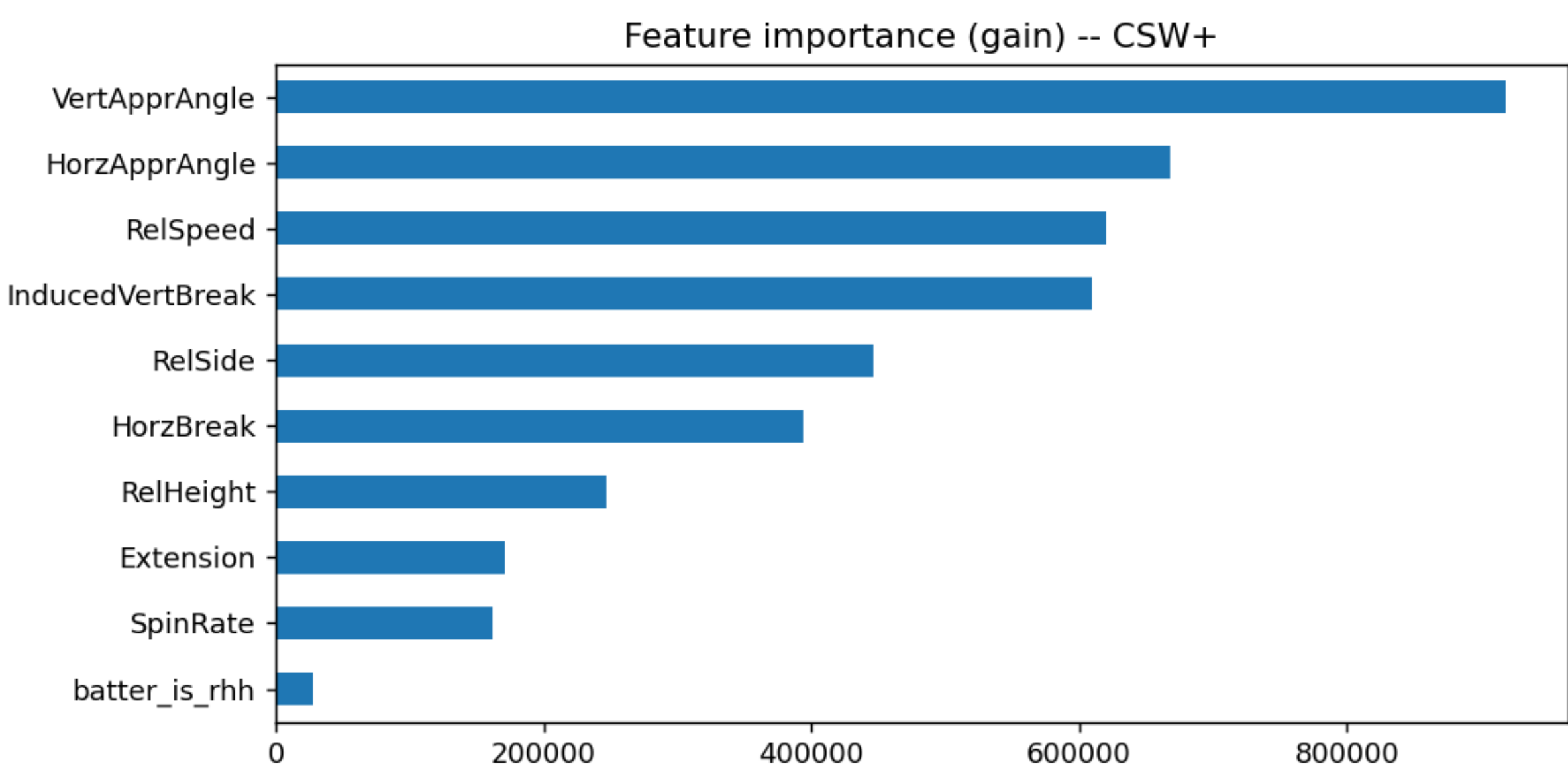
Validation and Aggregation

1. Apply five-fold **GroupKFold** cross-validation so a pitcher never appears in both training and validation data.
2. Use early stopping to select the number of boosting rounds in each fold: RMSE for RV+ and AUC for CSW+.
3. Select the median best iteration across folds and refit each model on the full dataset.
4. Report scores at the individual-pitch, pitcher–pitch-type, and overall pitcher levels.
5. Filter pitch-type summaries by minimum sample size and calculate overall pitcher grades using usage-weighted averages.

What Drives Each Grade?



Feature importance for the RV+ model



Feature importance for the CSW+ model

Results

- Higher RV+ and CSW+ grades were associated with lower ERA and FIP, demonstrating the expected relationship between Stuff+ and pitcher performance.
- Elite pitchers led both RV+ and CSW+ leaderboards, providing an initial face-validity check for the learned grades.
- Rise balls and fastballs ranked highest in RV+, reflecting their run-prevention value.
- Changeups and curve balls ranked highest in CSW+, reflecting their strike-generating ability.
- RV+ and CSW+ broadly agree while capturing complementary aspects of pitch effectiveness.

Top 5 Pitches by RV+ (Min. 100)				
Pitcher	School	Pitch Type	RV+	CSW+
Anneca Anderson	North Texas	Fastball	117.2	101.0
Erin Nuwer	Tennessee	Riseball	110.7	107.6
Sage Mardjetko	Tennessee	Riseball	110.6	107.5
Sidne Peters	Texas A&M	Riseball	110.0	101.8
Tatum Clopton	LSU	Riseball	109.9	99.8

RV+ leaderboard

Top 5 Pitches by CSW+ (Min. 100)				
Pitcher	School	Pitch Type	RV+	CSW+
Allyssa Parker	Oklahoma	Changeup	106.4	115.1
Alyssa Faircloth	Miss. State	Changeup	109.0	111.8
Randi Roelling	Georgia	Curveball	103.5	109.6
Alexis Jensen	Nebraska	Curveball	108.7	108.2
Miali Guachino	Oklahoma	Changeup	105.1	108.0

CSW+ leaderboard

Key Takeaway

RV+ and CSW+ provide two complementary views of pitch quality. RV+ emphasizes run prevention, while CSW+ emphasizes a pitcher's ability to generate called strikes and swings and misses. Together, they provide a more complete description of a pitcher's arsenal than either metric alone.

Limitations and Next Steps

- Expand data coverage beyond the approximately 40 collegiate programs currently represented.
- Transition to a higher-quality dataset containing only confidently labeled pitch types.
- Detect and correct TrackMan spin-rate errors, including occasional doubled measurements.
- Engineer interactions among velocity, movement, spin, and pitch type, and normalize approach angle for release height.
- Address uneven pitch-type sample sizes and develop a softball-specific reference for deception and tunneling.
- Use Optuna with pitcher-grouped cross-validation to tune both models and allow the CSW+ model additional boosting rounds.

References and Acknowledgements

References: Encyclopedia Britannica, "Softball." Nestico (2024), "Modelling tjStuff+ v1.0." TrackMan Baseball, "Radar Glossary of Terms."

Acknowledgements: We thank LSU Softball, the LSU Math Consultation Clinic, Dr. Nadejda Drenska, and Dr. Peter Wolenski for their support and guidance.