A Potential Solution to the Pitch Command Modeling Problem

Project Outline

One of the most difficult skills to quantify with any real precision is pitcher command. At its core, command depends on two pieces of information: where the pitcher intended to locate the ball, and where the pitch actually ended up. Leaguewide current models that are informed with public data must attempt to infer the first component indirectly, relying on secondary statistics and mechanical proxies that are informative but imperfect reflections of true intent. The central goal of this project is to build a hybrid system that captures intended location directly in controlled settings, learns how to map those patterns onto in-game pitch characteristics, and ultimately produces a reliable, organization-wide command model that improves both player development and player evaluation.

The first layer of the approach would focus on collecting high-quality, high-intensity practice data. During bullpens and simulated game environments, pitchers would wear lightweight eye-tracking glasses (ex: Tobii Pro Glasses 3) that record gaze direction, fixation patterns, and head-eye alignment throughout the pre-release window. When synchronized with pitch-tracking data, this setup creates an exact pairing of intended target and final pitch location. Over time, these sessions would build a comprehensive internal database reflecting how each pitcher in the organization commands different pitch types, how stable their visual routines are, and how intent deviates under different conditions.

The second layer leverages in-game data, where exact intended locations cannot be directly observed. As other organizations have experimented with, coaching staffs can call pitches during games, providing a structured and repeatable set of target labels. While these labels do not offer the same precision as eye-tracking data, they capture meaningful information about intended location, pitch type, and sequencing decisions. By aggregating these calls alongside pitch location outcomes, the organization can build a broader dataset that incorporates the in-game context and indirect signals that form the basis of existing league-wide command models.

Alternatively, teams could maintain their existing leaguewide, publicly informed command models, but use the intent-labeled data as a precise calibration target rather than an entirely separate system. Once the organization trains a command model directly on the eye-tracking sessions, it establishes a ground-truth representation of each pitcher’s actual intent-to-location relationship. That internal model becomes the benchmark. The next step is to build the parallel model that operates on leaguewide pitch-tracking data and iteratively refine its inputs, feature construction, and functional form until its outputs reliably converge toward the ground-truth model produced from the intent-labeled data. When the public-data model begins reproducing the same command signals the club knows to be accurate, the organization gains a scalable, leaguewide tool that is both robust and empirically validated against its proprietary standard.

The central technical challenge, regardless of the calibration pathway the team adopts, is to formally bridge these two data sources. Using the eye-tracking sessions as a labeled foundation, a supervised learning model can be trained to predict intended location from a combination of pitch outcomes, contextual variables, and pitcher-specific tendencies. A gradient-boosted framework such as XGBoost in R would be appropriate for the first layer of prediction, while a Bayesian hierarchical model or mixed-effects structure can account for within-pitcher and between-pitcher variation. R would support the entire data processing and modeling workflow, with packages such as tidyverse for data manipulation and brms or rstanarm for hierarchical modeling. This structure allows the club to infer intent from in-game pitch data with calibrated uncertainty, creating a unified command score that is consistent across practice and competition.

Once established, this model opens several avenues of value to the front office. Player-development staff gain a clearer understanding of which pitchers possess repeatable visual routines and which exhibit instability prior to release. Analysts can identify organizational traits that correlate with strong command, providing a template for scouting and acquisition. The model also makes it possible to evaluate pitchers outside the organization with greater accuracy, since inferred intent provides a more stable indicator of underlying command quality than raw outcomes. Finally, by isolating the perceptual and mechanical components of command, the organization can experiment with targeted improvement plans and track whether adjustments produce measurable improvements in intent-to-location execution.

Ultimately, this layered framework provides a clearer and more empirically grounded way to measure command, and it offers a sustainable mechanism for generating competitive advantages in development, scouting, and player evaluation.