Catcher Spotting: Measuring Pitcher Intent

Andrew C. Thomas

This draft: October 19, 2009

It's an element of the game of baseball that sportscasters mention all the time, but is missing from most numerical analysis. A pitcher is said to be "missing his spots" if the final location of a pitch is not where the catcher had signaled for the ball to be thrown. The data available through Major League Baseball cover almost every aspect of the ball's trajectory, both before and after contact with the bat, but the intended target is not systematically recorded.

When the precision and accuracy of a pitcher can be known with respect to the time of game, the type of pitch, the men on base and the number of pitches thrown, there are many questions we can investigate in each situation that may not have been as easy to disentangle without a measure of intent.

First I describe exactly what is meant by pitcher intent in terms of measurement, then suggest methods for collecting the data along with the potential cost for assembling a complete data set. I then speculate on a number of questions that can be asked with this type of data set.

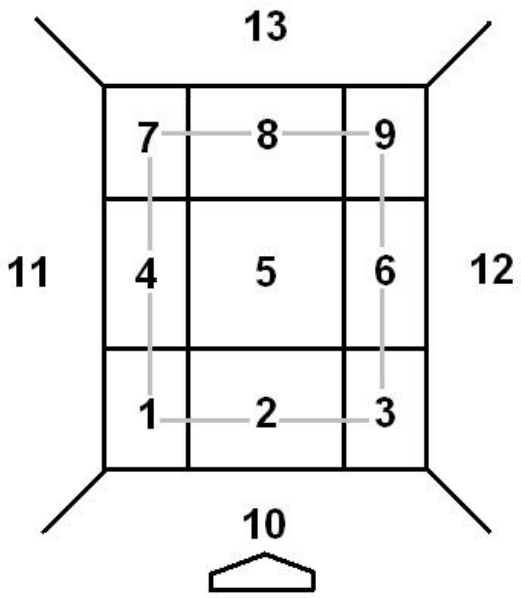

Figure 1: A map of potential zones for the target and actual destination of each pitch.

Collection of Spotting Data

There are two pieces of data to record in each instance: the intended target and the actual destination point. There are two basic ways of collecting these data: by considering the target area as a continuous coordinate space, or by dividing the target into zones as in Figure 1: a pitch can be aimed vertically (high, middle or low) or horizontally (inside, straight or outside) relative to a pitch right down the middle of the plate.

Not all batters are of the same height, however, and so "high" and "low" correspond to different absolute positions for different batters. More than that, pitchers do realize that their aim is not perfect, and the target is not necessarily set to an exact position, more to a particular zone. This also suggests a reasonably robust data collection method at this scale: simply estimate the target and destination zones by visual observation, as television commentators do – e.g. the pitch is aimed low and away (zone 1), but ends up inside at waist level (zone 6).

There is one particular difficulty with this method: the inter-rater agreement between what is considered a zone can vary, just as different umpires may call strikes with slightly different strike zone. Therefore, a system that can minimize this factor would prove to be extremely useful.

Continuous-space Observation

I have prepared an applet for collecting catcher spotting data at a crude level at my website, http://acthomas.ca/catcher-old/. This applet demonstrates both the concept of collecting the data – point and click the target and destinations respectively – and the difficulty in estimating the error in the observation. Figure 2 shows an example of pitches tracked separated by their target zone to demonstrate the general accuracy given the input.

We can assess inter-rater reliability for this model relatively easily: several people watch the same inning of baseball and make their own calls. Preliminary investigation suggests that this eyesight-based method is unreliable for large-scale data collection.

An alternate method with great potential is to use a telestrator approach: while viewing a video of the game in action, a rater can simply click on the screen to indicate the target location as well as the true destina-

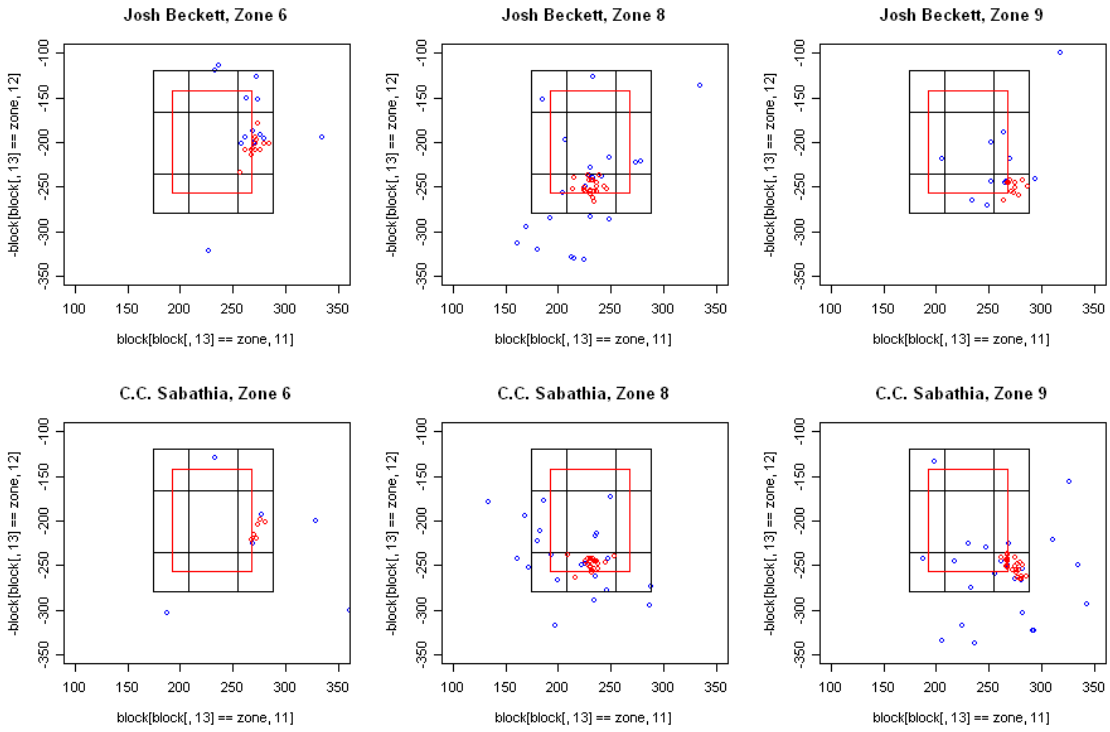

Figure 2: Pitching charts for two innings of a game in the 2007 ALCS for Josh Beckett and C.C. Sabathia, where each chart covers one intended target zone (points in red) and their corresponding true destinations (in blue).

tion of every pitched ball.[1] Once assembled, this process has two immediate benefits: the destination data can be matched and checked against MLB's pitch information data, and there is the possibility that the method can be completely automated to remove the human element and greatly speed up data collection; or, at the very least, to trim all non-pitch-related video from the raw game feed.

Applications

There are a number of concepts that can be investigated, and specific questions asked, once this class of data is available.

How much variation is "reasonable", per pitch type?

Baseball space is not isotropic: a miss in one direction may be more or less consequential than a miss in the opposite direction or a perpendicular dimension. A vertical curve ball, for example, has considerably more variation on the y-axis than on the x-axis, with different consequences for each type of miss; a "hanging" curve ball typically has very different results than a ball in the dirt. This also means that each type of pitch can be assessed a class of risk from a greater family.

Does injury follow inaccuracy?

Pitch precision could very well be an early indicator of injury in various situations. Since the physical act of pitching is an act of the whole human body, the smallest change in pitching behaviour may be attributable to a weakness or strain that may quickly blossom into a more severe tear or pull. The potential financial benefit for pre-empting these types of injury, given a new data source, is considerable.

How much deception do catchers use?

A catcher's signals are visible to a runner on second base, which has two pertinent consequences: an alternate set of signals are used to call for a particular type of pitch, and a catcher may not overtly signal the intended location, and if at all, the target usually only appears after the pitcher's action has begin. Rather than using the catcher's mitt as the indicator for location, the rater is then forced to guess the intended target from the catcher's body language. Rather than asking the catcher what their intention was (a difficult task for an academic), this is an opportunity to ask experts on the

[1] I am currently experimenting with Gromit, a free open-source telestrator program, to mark pitch locations in combination with the video player Totem on an Ubuntu Linux framework.

subject what the intended target was given their experience.

Should Grady Little have pulled Pedro in 2003?

My interest in intended pitch target grew from rewatching Pedro Martinez's performance in the seventh game of the 2003 ALCS. As the game goes on, catcher Jason Varitek appears to be laboring more to reach each pitch as the game progresses, even though Martinez did not appear to have a measurable drop in velocity. While many baseball analysts have argued that Martinez should have been pulled based on his opponent's batting average (indeed, based on a disobeyed directive from the front office that he should not remain past seven innings for that reason), there is a considerable potential for insight on the way pitchers should be managed based on their precision throughout a game.

Project Timeline and Cost Estimation

There is a considerable amount of work to be done before we can even attempt to answer these questions; at the very least, it must be proven that the data can be collected reliably. Roughly speaking, here is what needs to be done to investigate the phenomenon further:

1. Pilot data collection. To shake out the difficulties in collecting this kind of data, I propose the collection of 20 games, which at roughly 250 total pitches per game gives a prototype data base of 5000 pitches. At a fee of $50 per game scored, this is $1000 for this part of the project.

2. Refinement of tools. A baseball game takes roughly 3 hours to play, given commercial breaks: this is roughly 45 seconds per pitch overall. I propose a target maximum of 10 seconds per pitch for data collection once methods for trimming game data down to the essential elements, estimated at $5000 for one student to investigate and build.