What New Technologies Are Teaching Us About the Game of Baseball

Alan M. Nathan
University of Illinois at Urbana-Champaign
a-nathan@illinois.edu

October 15, 2012

Abstract

The trajectory of a baseball moving through the air is very different from the one we teach in our introductory classes, in which the only force is that due to gravity. In reality, the aerodynamic drag force (which retards the motion) and the Magnus force on a spinning baseball (which causes the ball to curve) play very important roles that are crucial to many of the subtleties of the game. Despite their importance, our knowledge of how these forces affect the flight of the baseball has been qualitative at best. Recently, however, new tools have been developed for measuring accurate baseball trajectories during an actual game. These tools include both video tracking (the PITCHf/x and related systems) and the TrackMan Doppler radar system. In this article, I will discuss these new tools and give some examples of what they are teaching us about the game of baseball.

1 Introduction

The flight of the ball plays an integral role in the game of baseball. Whether the ball is pitched by the pitcher, struck by the batter, or thrown by a fielder, the physics that governs the flight involves the same forces. If these are known and the initial conditions are given, the full trajectory of the ball can be predicted by solving the equations of motion. So, what are the forces on a baseball as it travels through the air? Of course there is the downward pull of gravity, which is typically the only force treated in introductory physics courses. In real life, however, gravity must be supplemented with the two aerodynamic forces of air drag $\vec{F_D}$ and the Magnus force $\vec{F_M}$, the latter arising when the ball is spinning. The conventional way to parametrize the aerodynamic forces is as follows:

$$\begin{aligned} \vec{F_D} &= -\frac{\pi}{2}\rho R^2 C_D v^2 \hat{v} \\ \vec{F_M} &= \frac{\pi}{2}\rho R^2 C_L v^2 (\hat{\omega}\times\hat{v})\,, \end{aligned} \tag{1}$$

where C_D and C_L are the drag and lift coefficients, respectively, $\vec{v}$ is the velocity vector, $\vec{\omega}$ is the spin vector, R is the radius of the ball, and ρ is the air density. The direction of the drag is opposite to the instantaneous velocity vector, so it slows the ball without changing its direction. The direction of the Magnus force is perpendicular to both the velocity and spin vectors, so it only changes the direction of the ball. A convenient mnemonic is that the Magnus force points in the direction in which the leading edge of the rotating ball is turning. So, for example, for a ball moving horizontally with backspin, the Magnus force is directly up. As we will see in the examples that will be presented below, the aerodynamic forces play important and often subtle roles in the game of baseball.

Although not the primary focus of this paper, we start by reviewing briefly the current state of our knowledge of C_D and C_L. New measurements of C_D for a variety of sports balls are given by Kensrud and Smith.[1] For a baseball, C_D is approximately 0.50 for speeds below about 40 mph. At some higher speed, it undergoes a "drag crisis" as the air flow in the boundary layer makes a transition from laminar to turbulant, resulting in a drop in C_D to about 0.30 at a speed of around 90 mph. The range of speed over which the drag crisis occurs is not known with great precision, as different experiments seem to not agree. At 90 mph and typical air densities, the drag force is about equal to the weight of the ball. Many interesting questions abound, such as the dependence of C_D on the seam orientation or the spin. One interesting feature is that most trajectories in baseball fall right in the region of the drag crisis.

The lift coefficient is expected to be a function of the spin parameter $S = R\omega/v$. Some recent measurements and a review of earlier measurements are given by Nathan.[2] The existing data are reasonably well described by the relation

$$C_L = \frac{2.5S}{1+5.8S}\,. \tag{2}$$

For low values of spin, C_L has a linear dependence on S, vanishing when the spin vanishes. For a baseball moving at v=90 mph and spinning at ω=1800 rpm–values quite typi-

cal of both pitched and batted balls–S=0.22 and C_L=0.24; if the spin is perpendicular to the velocity, the magnitude of the Magnus force is approximately 0.8 of the weight of the ball. For high values of S, C_L asymptotically approaches 0.43.

The fact that the drag and Magnus forces can be appreciable fractions of the weight suggests that neglecting the aerodynamic forces in trajectory calculations is not a very good approximation, as demonstrated in Fig. 1. This

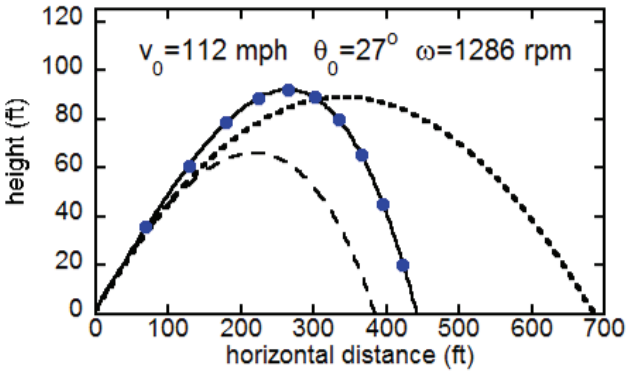

Figure 1: An example of a fly ball trajectory (solid curve), with the initial conditions shown. The dots show the location of the ball in 0.5-sec intervals. The other curves are the trajectories with the same initial conditions with neither drag nor Magnus (short dashed) and with drag but no Magnus (long dashed).

fly ball was hit with initial parameters taken from an actual home run: launch velocity of 112 mph, vertical launch angle of 27°, and backspin of 1286 rpm. Such a ball would carry nearly 700 ft in a vacuum but in reality only carries about 430 ft. Whereas the vacuum trajectory is symmetric about the apex, the effect of drag is to reduce the horizontal speed so that the ball travels farther prior to than after the apex. This particular feature is well known to baseball players, even though they are probably not aware of the physics that causes it. Also shown is a trajectory with drag but no Magnus force, which in this case is mainly in the upward direction, opposing gravity. Relative to the no-Magnus trajectory, the ball rises higher, stays in the air longer, and travels farther, although the latter effect is only about 50 ft.

2 The New Technologies

2.1 The f/x Video-Based Systems

The f/x systems are based on standard video technology. Most prevalent is the PITCHf/x system,[3] which is installed in all Major League Baseball (MLB) ballparks and has been used since the start of the 2007 season to track every pitch in each MLB game. The system consists of two 60-Hz cameras mounted high above the playing field with fields of view that cover most of the region between the pitching rubber and home plate and with roughly orthogonal principal axes. Proprietary software is used to identify the camera coordinates of the baseball in each image, which are then converted to a location in the field coordinate system. The conversion utilizes the camera transformation matrix, which is determined separately using markers placed at precisely known locations on the field. Depending on the details of each installation, the pitch is typically tracked in the approximate range y=5-50 ft, resulting in about 20 images per camera for each pitch, each with measurement precision of order 0.5 inch. Each trajectory is fitted using a constant-acceleration model, so that nine parameters (9P) determine the full trajectory: an initial position, an initial velocity, and an acceleration for each of three coordinates. Simulations have shown that such a parametrization is an excellent description of trajectories for most pitches.[4] The essential physics is that the aerodynamic forces on the ball, while not constant, change slowly enough and the flight path is short enough that the 9P model provides an excellent description of the actual trajectory. Using the 9P fit, various quantities of interest to both physicists (e.g, drag and lift coefficients) and baseball analysts (e.g., release speed, home plate location, movement) can be determined. The HITf/x system utilizes the same cameras to track the initial trajectory of batted balls, including the batted ball speed, vertical launch angle, and horizontal spray angle. The FIELDf/x system is an ambitious extension that tracks essentially everything on the playing field: the fielders, the base runners, the umpires, and the full trajectory of the batted baseball.

2.2 TrackMan Doppler Radar

TrackMan is a phased-array Doppler radar system that is installed in a number of MLB ballparks. The system is usually mounted high above home plate and has a field of view that covers most of the playing field. A single antenna transmits at approximately 10.5 GHz. The signal reflected by the baseball is detected in a 3-antenna array, allowing the determination of the Doppler velocity and angular location of the baseball in the radar coordinate system. Together with an initial location, that is sufficient to track the full trajectory of both pitched and batted baseballs. For pitched baseballs, the measurement precision is comparable to that of PITCHf/x. The precision for batted balls is not known. Unlike PITCHf/x, the TrackMan data are not publicly available, although some MLB teams are willing to share data with researchers.

3 Studies of Pitched Baseballs

The great lefthanded pitcher Warren Spahn once said, "Hitting is timing. Pitching is upsetting timing." In this section, I will use the publicly available PITCHf/x data from actual MLB games[5] to explore some of the ways that pitchers can upset the timing of batters.

Given the variety of possible pitches that can be thrown (e.g., fastballs, curveballs, sliders, etc.), one way to upset the timing of a batter is for the trajectory of each pitch to be nearly identical for as long as possible, giving the batter little opportunity to recognize and react to the different pitches. There is a perception among practitioners of the game that a pitched ball can deviate sharply from a straight line just before reaching the batter, giving rise to the concept of "late break." From a purely physics point of view, late break is a fiction, since it is not possible to change rapidly the trajectory of a pitched baseball without enormous forces acting on it. So, it is a fair question to ask why perception is different from reality, a topic I examine in Fig. 2. Shown is a bird's eye view of an actual pitch (the solid curve) thrown by the great New York Yankee pitcher Mariano Rivera. The dotted curve is a straight line extrapolated from the initial velocity vector. The two curves are virtually indistinguishable until about 20 ft from home plate, at which point the batter has already decided where he thinks the ball will end up and whether or not to swing. The two trajectories deviate by about 0.5 ft at home plate, which is more than enough to confound the batter, who swears the solid curve broke sharply at the last moment. Clearly it did not.

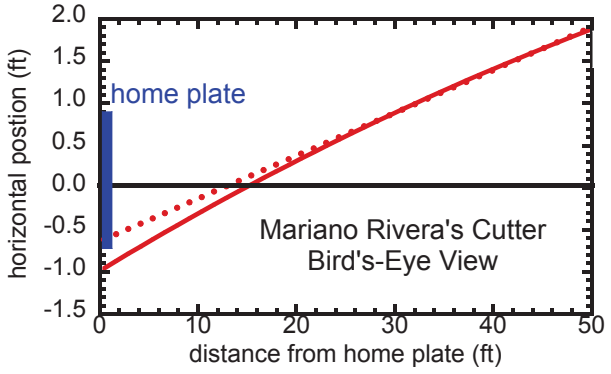

Figure 2: An example of "late break."

Another way a pitcher can upset the timing of the batter is by varying both the speed and the movement of the pitch. The movement is defined as the deviation of the pitch from a straight line trajectory, with the effect of gravity removed. For most pitches, it is determined by the magnitude and orientation of the spin axis. An example is shown in Fig. 3, in which the direction of the movement and release speed

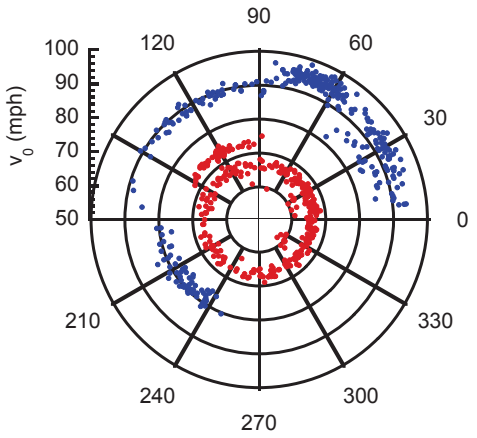

Figure 3: Polar plot showing the pitch repertoire of Jon Lester (blue) and knuckleball pitcher Tim Wakefield (red). The radial coordinate is the release speed and the angular coordinate is the direction of the movement, as viewed by the catcher.

for a few hundred pitches thrown by Boston Red Sox lefthander Jon Lester are shown as the blue points in a scatter plot, as viewed by the catcher. The spin axis is the movement angle plus 90° (see Eq. 1). For Lester, as with most MLB pitchers, the plot shows distinct clusters that serve as signatures for each of the five pitches in his repertoire. For example, the cluster at 95 mph and 65° (spin angle=155°, or primarily backspin) is his primary pitch, the so-called four-seam fastball, breaking primarily upward and slightly away from a right-handed batter. The cluster near 77 mph and 225° (spin angle=315°) is his curveball, breaking down (due to topspin) and toward a right-handed hitter. Now compare with the red points, the pitches thrown by recently retired Red Sox righthander Tim Wakefield, who occasionally throws fastballs (73 mph and 120°) and curveballs (60 mph and 315°). However, Wakefield's primary pitch is the knuckleball, shown as the amorphous ring at about 66 mph. Whereas the movement of "ordinary" pitches is predictable, the movement of a kunckleball is not.

Given the unusual nature of the knuckleball, it is interesting to examine its behavior in more detail. It is usually thrown at a speed significantly lower than that of ordinary pitches and with very little spin. The lack of spin means that the knuckleball does not experience the Magnus force that is responsible for the movement on ordinary pitches. But it does experience movement, as seen in Fig. 3. Wind tunnel experiments have shown that the movement is due to the disruption of the air flow over the seams of the ball.[6] The seemingly erratic movement has led to the perception that the knuckleball trajectory is not smooth but undergoes abrupt changes of direction. This is an issue that can be addressed with the tracking data. To do so requires having

access to the raw tracking data rather than the 9P fit, since we do not know how well the constant-acceleration model fits the actual data.

Therefore raw data were obtained for several games from the 2011 MLB season. Here I will give a detailed analysis for one game in which 278 pitches were thrown, of which 77 were knuckleballs. Each trajectory was fit to a smooth function and the root-mean-square (rms) deviation of the data from that function was determined. The essential idea is that the smaller the rms, the better the smooth function describes the actual data. Rather than use the constant acceleration function, a more exact model was used in which the aerodynamic forces are proportional to the square of the velocity. A nonlinear least-squares fitting program was used to adjust the parameters to minimize the rms value. The result of applying this procedure is presented in Fig. 4.

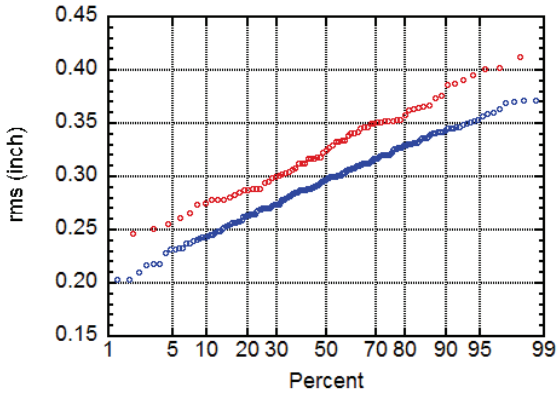

Figure 4: Distribution of rms values for normal (blue) and knuckleball (red) pitches. The nearly parallel linear contours of these samples indicate that they each have a Gaussian distribution with the same standard deviation but with the mean value slightly larger for knuckleballs.

The rms value for each of the 201 ordinary pitches (blue) and 77 knuckleball pitches (red) is plotted as a function of the percentage of pitches in each sample having a smaller rms value. The horizontal axis is such that samples following a Gaussian distribution appear as straight lines, with the central value at 50% and standard deviation proportional to the slope. This plot shows that the distribution of rms values is very similar for the ordinary and knuckleball pitches, each being approximately Gaussian with about the same standard deviation, but with the mean value for knuckleballs (0.33 inch) only slightly larger than that for ordinary pitches (0.30 inch). If we take the mean value for ordinary pitches as a measure of the statistical precision of the tracking data (approximately 0.3 inch), then the slight increase in the knuckleball values suggests that the latter pitches deviate from "smoothness" by at most 0.15 inch. The conclusion is that within the precision of the tracking data, knuckleball trajectories are just as smooth as those of ordinary pitches, thereby disproving the common perception.

4 Studies of Batted Baseballs

I now want to turn to a study of batted baseballs. Unfortunately, full trajectories of batted baseballs from MLB games are not yet available for analysis. However, two extremely useful pieces of information are more readily available: the initial velocity vector $\vec{v}_0$ from either the video or radar tracking systems; and the landing point $\vec{R}_f$ and flight time T, as determined either from the radar tracking or from direct observation, the latter only for batted balls that clear the fence for a home run.[7] Interestingly, this partial information is sufficient to place very tight constraints on the full trajectory, as demonstrated in Fig. 5. The data for the full trajectory shown in the figure come from a dedicated experiment in which baseballs were projected using a pitching machine and tracked using a portable version of the TrackMan system. The first 0.5 sec of the trajectory were used to determine $\vec{v}_0$, while $\vec{R}_f$ and T were directly measured by the tracking system. The trajectory was modeled using Eqs. 1-2, with three free parameters adjusted to fit $\vec{R}_f$, evaluated at the measured T: a *constant* drag coefficient C_D and the backspin ω_b and sidespin ω_s components of the spin vector. The spin components are both perpendicular to $\vec{v}_0$, with ω_b and ω_s lying in the horizontal and vertical planes, respectively. The curve in Fig. 5 was arrived at by this procedure. The closeness of the curve to the actual data over the full range of the trajectory demonstrates the point made above. It also shows the utility of the simplified model used in the analysis. Simulations of this procedure show that T determines ω_b; the horizontal deflection determines ω_s; and the total distance determines C_D. This procedure will be used in the some of the discussion that follows.

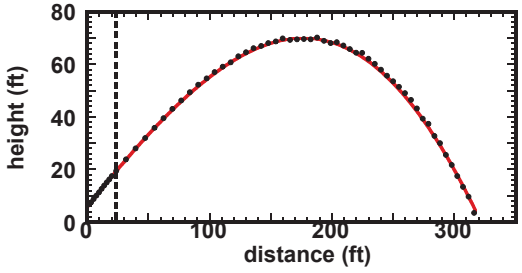

Figure 5: A fly ball trajectory measured with TrackMan (points). The curve is a fit to the data in which only the data to the left of the dashed line plus the landing point and flight time were used to constrain the parameters.

The first question I address is the dependence of fly ball distance on the initial speed v_0 and vertical launch angle θ. A total of 1257 batted balls were analyzed, using TrackMan from MLB games in St. Louis from 2009. The results are shown in Fig. 6, in which the fly balls are separated into 10 mph and 5° bins and mean values are taken for the range in each bin. Some interesting features emerge from the figure. First, the optimum θ is about 30°, which is considerably smaller than the vacuum value of 45°. Second, the range depends on v_0 approximately linearly, with a slope of about 5 ft/mph; in vacuum the dependence is quadratic. Finally, a ball hit with v_0=102.5 mph at and θ=30° will carry about 405 ft. These results are in accord with expectations based on calculations of trajectories using reasonable models for C_D and C_L, and it is nice to see that they are confirmed by actual data.

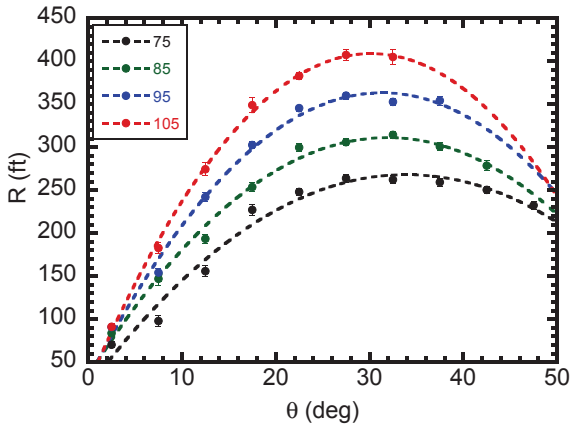

Figure 6: Dependence of fly ball distance on the vertical launch angle for different 10-mph ranges of v_0, with the central value indicated in the legend.

Next, I investigate the effect of temperature and altitude on fly ball distances. For this study, I analyze 8800 home runs from the 2009-10 MLB seasons, using $\vec{v}_0$ from HITf/x and direct observation of $\vec{R}_f$ and T. The fitting procedure described at the beginning of this section is applied to each home run to determine the average C_D and the two spin components, using an air density appropriate for the known temperature and elevation. Then for each home run, two new trajectories are calculated using the same $\vec{v}_0$, C_D, and spin: one in which the temperature is set to 72.7F, the average value for the full data set, and another in which the elevation is set to sea level. For each of these a new value of the range is determined. We denote by ΔR the difference between the actual and new value of range. In such a manner, we can determine ΔR as a function of temperature or elevation, and these results are shown in Fig. 7.

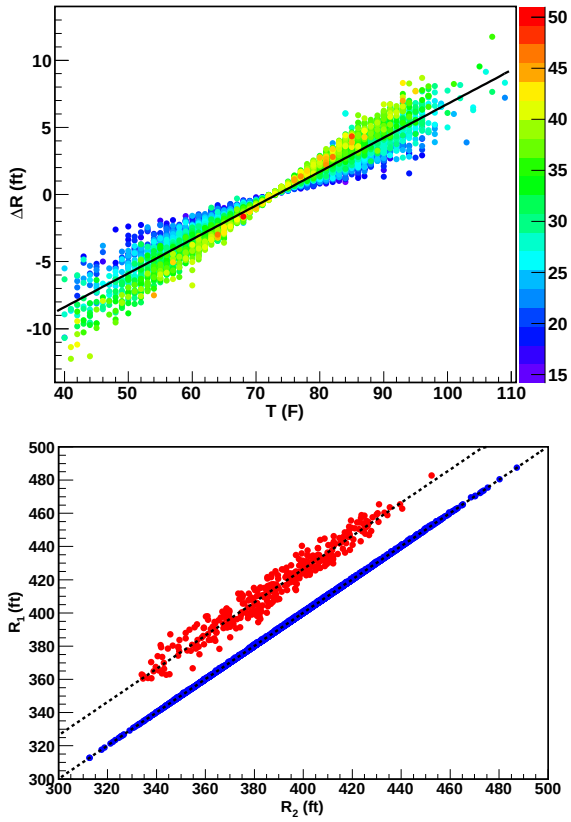

Figure 7: Dependence of home run distance on temperature (top) and elevation (bottom). The top plot is color-coded by vertical launch angle. The bottom plot displays the actual distance vs. the sea level distance for sea-level ball parks (blue) and Coors Field in Denver (red).

In the top part, the linear dependence of ΔR on temperature for a given θ is evident, with the slope increasing with larger θ. The average slope, corresponding to the optimum launch angle of 30°, is 2.5 ft/10F, so that a long fly ball will carry about 15 ft farther on a hot summer day at 100F than on a cold spring day at 40F. In the bottom figure, the actual distance R_1 is plotted versus the sea-level distance R_2. The blue points are those for ballparks within 200 ft of sea level, so the data is essentially a line of unit slope passing through the origin. The red points are those for Coors Field in mile-high Denver, for which there is also a line of unit slope but displaced higher by about 26 ft., from which we conclude that fly ball distances increase by about 5 ft for each 1000 ft of elevation.

Finally I address the question of how well the initial velocity vector $\vec{v}_0$ determines the landing point $\vec{R}_f$. In vacuum, of course, $\vec{v}_0$ is completely sufficient to determine $\vec{R}_f$. The actual situation is more complicated due to the aerodynamic forces. The ideal tool for making such a study

is TrackMan. Unfortunately there are not sufficient data freely available from that system for such a study. Instead, I will use the data from the 8800 home runs from 2009 and 2010, the same data used to investigate the dependence of a fly ball distance on elevation and temperature. The basic idea is to place a narrow cut on v_0 and θ and plot the resulting distribution of R_f^*, which is the distance expected with the same initial conditions at the standard air density at sea level and 72.7F. An example is given in Fig. 8 for which the restrictions were $99 \leq v_0 \leq 101$ mph and $26.7^\circ \leq \theta \leq 30.7^\circ$. Shown is a scatter plot of the inferred C_D values versus the total distance R_f^*, along with a histogram of R_f^* showing that the spread of distances is amazingly large, approximately 40 ft FWHM. It is natural

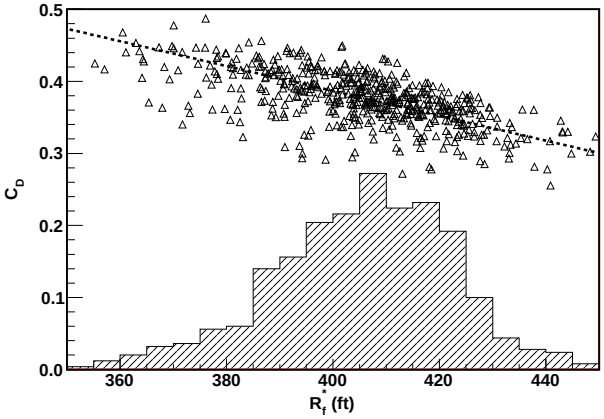

Figure 8: Plot of C_D vs. R_f^* with $99 \leq v_0 \leq 101$ mph and $26.7^\circ \leq \theta \leq 30.7^\circ$. The dotted line is a linear fit and shows clearly the correlation between C_D and R_f^*. The histogram is the projection onto the R_f^* axis.

to ask why this is so, keeping in mind that the distances have already been normalized to standard temperature and elevation. One clue comes from the inferred C_D values for which the FWHM is also large, about 20% of the mean value. Moreover, there is a clear correlation between and C_D and R_f^*, with larger distances associated with reduced drag. On the other hand, no such correlation is found between ω_b and R_f^*, suggesting that variation in backspin does not account for the variation in distance. So, why does the baseball "carry" better in some instances than in others? I offer two possible reasons. First, the effect of wind, while not explicitly included in the fitting procedure, is implicitly taken into account with an elevated or reduced C_D for an in-blowing or out-blowing wind, respectively. Second, it is possible that there are variations in C_D from one baseball to another due to subtle differences in the surface roughness. Neither possibility can be ruled out in the present analysis. Interestingly, a similar variation in C_D is observed from pitched baseballs in ballparks with both PITCHf/x

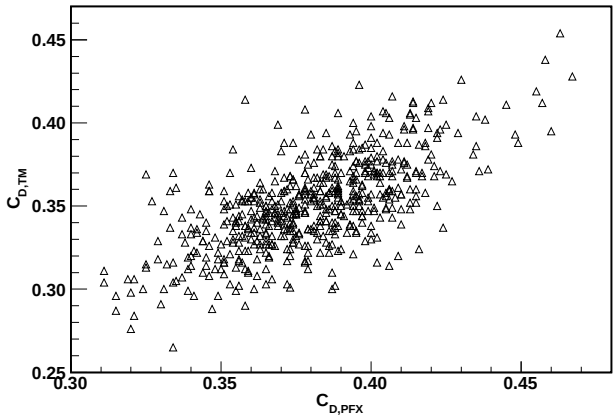

Figure 9: Plots of C_D for pitched baseballs determined from TrackMan vs. those determined from PITCHf/x.

and TrackMan installed, as shown in Fig. 9. Over a narrow range of release speeds, the inferred C_D for each system is approximately normally distributed with rms=0.033, due in part to random measurement noise. However, the values measured by the two independent systems are highly correlated, suggesting that a large part of the spread (0.028) is due to a common feature of the two systems rather than measurement noise. This analysis lends support to the notion that there is a real variation in C_D values over and above measurement noise. A controlled experiment is in the planning stage to study further this variation, with the goal of distinguishing wind from surface effects.

5 Summary

The new tracking systems that are used in MLB ballparks have provided a wealth of information about baseball aerodynamics that is teaching us much about how the game of baseball is played. A few examples have been presented in this article.

References

[1] J. Kensrud and L. Smith, "In situ drag measurements of sports balls", The Engineering of Sport **8**, Vol. 2, Issue 2, ISEA, Vienna, Austria, pp. 2437-2442.

[2] Alan M. Nathan, "Effect of spin on the flight of a baseball", Am. J. Phys. **76**, 119-124 (2008).

[3] Mike Fast, "What the heck is PITCHf/x?", The Hardball Times Annual, 2010. A copy may be downloaded at `http://webusers.npl.uiuc.edu/~a-nathan/pob/FastPFXGuide.pdf`.

[4] A web article about this analysis may be downloaded at http://webusers.npl.uiuc.edu/~a-nathan/pob/PitchFX_9P_Model-4.pdf.

[5] PITCHf/x data may be downloaded at www.brooksbaseball.net.

[6] R. Watts R and E. Sawyer, "Aerodynamics of a knuckleball", Am. J. Phys. **43**, 960-963 (1975).

[7] The web site for the ESPN Home Run Tracker is at http://www.hittrackeronline.com/.