Tracking Pitcher Performance with Instantaneous Component ERA and Moving Averages

John L. Salmon, Brigham Young University
Willie K. Harrison, University of Colorado Colorado Springs

1 Introduction

For more than 100 years the earned-run averaged, or ERA, statistic has been used to evaluate performance on the mound [6]. Despite its familiarity and general acceptance on baseball statistic tables, sabermetricians are aware of the significant limitations of using ERA in isolation to evaluate, compare, and compensate pitchers [5]. Managers should be likewise hesitant to use ERA excessively or even as the dominant criterion for pitching decisions, especially when such decisions are made independently of other readily available statistics. In addition, ERA statistics vary significantly from year to year, with a low correlation to the rankings of the final team’s performance suggesting reduced validity [3].

Two simple examples quickly illustrate these limitations. If a pitcher is pulled in the middle of an inning after leaving batters on base, and the reliever allows a base hit resulting in a run, the first pitcher’s ERA is adjusted higher while the reliever’s ERA does not change. In other situations, the pitcher is penalized when a fielder is not charged with an error, despite the latter’s evident mistake or missed play that results in a run, which is ultimately counted against the pitcher. Both of these examples should incite some uncertainty that the ERA is a reasonable metric for accurately assessing pitching performance.

In response to these limitations, a number of metrics have been proposed, sometimes referred to as defensive independent pitching statistics, or DIPS, to better isolate the performance of a pitcher from other teammates (DICE), the particulars of the ball-park (xFIP), or other factors (tRA) [7]. Many of these statistics seek to breakdown or decompose the various events and outcomes within the pitcher’s control, and appropriately remove or adjust those actions which should not be justifiably ascribed to the pitcher.

Although each DIPS will have associated assumptions with fewer weaknesses compared to the ERA, exploration into how the statistical values change over time has been lacking. This research seeks to explore the behavior of these metrics and how the trends and dynamic nature of these statistics over the course of a season could inform better pitching decisions. Questions potentially answerable with this methodology include: Is a given pitcher improving? Is their improvement accelerating or decelerating? Are they trending up or down? On what time frame? Are they experiencing a temporary slump? Are they expected to regress to the mean in the short-term?

While this research investigates the Component ERA (CERA) statistic in detail, other DIPS could potentially be applied with the various resultant visualizations and conclusions analyzed with an alternative statistic (e.g. DICE, LIPS, FIP, QERA, etc.) [3]. Therefore, one can easily replace CERA with a different statistic and apply the same analysis without loss of generality. Ultimately, the goal is to
answer how one can enhance these statistics in providing additional evidence to allow management to better play, rotate, and pull pitchers appropriately for overall increased performance.

2 Methodology

The Component ERA statistic, or CERA, enhances the traditional ERA by providing a more detailed metric to analyze the performance of pitchers [4].

While the formulation for the ERA is:

\[
ERA = \frac{9 \text{Earned Runs Allowed}}{\text{Innings Pitched}}
\] (1)

the formulation for CERA is more sophisticated, with the statistic decomposed into baserunners allowed (BA) multiplied by the Pitcher’s Total Base Estimate (PTB), with appropriate scaling:

\[
CERA = 9 \left( \frac{\text{BA} \times \text{PTB}}{\text{BFP} \times \text{IP}} \right) - 0.56
\] (2)

and BA is defined as:

\[
\text{BA} = \text{H} + \text{BB} + \text{HBP}
\] (3)

with,

\[
\text{PTB} = 0.89 \left( 1.255(\text{H} - \text{HR}) + 4\text{HR} \right) + 0.56 \left( \text{BB} + \text{HBP} - \text{IBB} \right)
\] (4)

where, BFP is the number of batter’s faced by the pitcher, IP is the number of innings pitched, BB is the number of bases on balls (walks), IBB is the number of intentional bases on balls (intentional walks), HBP is the number of times the batter was hit by a pitch, and H and HR are hits and homeruns, respectively.

Since the CERA uses hits and walks in place of runs, it provides a more accurate view of a pitcher’s performance, although not entirely removing the effect of team performance. Furthermore, by appropriately aggregating the results an “Instantaneous” CERA value can be calculated after the conclusion of each individual batter a pitcher faces throughout the season. This instantaneous CERA can be updated and tracked from batter to batter to gain insight into the overall behavior or trend of a pitcher’s performance.

When these values are plotted over time, or with respect to the number of batters faced, the individual pitcher performance and the comparison between other pitchers is readily available. Examples of these comparisons are presented and discussed in the next section.

Furthermore, with computed time series data of the instantaneous CERA throughout the season, rolling or moving averages at different time scales can be applied. The crossover points from a moving average approach, similar to some investing strategies, can suggest when a pitcher is trending up or down, providing additional evidence in making more informed pitching decisions.

The pitching data used in this research are publicly available and were obtained from Retrosheet [1]. At the time of data processing, the 2015 MLB Season was still underway and therefore only the 2014 data was accessible. The individual at bat results for all 30 teams and all 2430 games are processed and combined into one data table for analysis. The table is then sorted by time (i.e. by game day and inning) to enable instantaneous CERA calculations for every pitcher during the season and enable moving average evaluations to identify trends and other features.
3 Results and Discussion

3.1 Instantaneous CERA

The Instantaneous CERA profile for five different pitchers is shown in Figure 1. For every additional batter faced, the CERA for each pitcher can be recalculated and is plotted per number of batters faced (a surrogate for time). The log scale on the y-axis illustrates the high variance of the CERA statistic earlier in the season when fewer batters have been faced by each pitcher.

As shown in Figure 1, the individual CERA profiles become more stable over time since the CERA statistic is dependent on the number of batters. In other words, an event later in the season will have a smaller impact in calculating CERA than an earlier event. This is advantageous since analyzing overall pitcher performance is desirable as the CERA profile naturally dampens out random “streaks” of unusually favorable (or unfavorable) outcomes over longer periods of time (i.e. more batters). Furthermore, this suggests that significant movement later in the season could be indicative of something wrong or a changing trend that warrants investigation (e.g. potential discomfort or injury).

Therefore, the upward movement of Johnny Cueto’s CERA profile from Figure 1 with a slow but steady increase in his CERA (or decrease in overall performance), may suggest a pitching rotation with too many games too close to each other resulting in a slight issue with fatigue or recovery throughout the season. Clayton Kershaw’s profile decreases dramatically for the first 400 batters, flattening out for most of the second half of the season and then increasing slightly during the last few games. Scott Kazmir and Corey Kluber’s CERA profiles cross around 630 batters, suggesting Kluber and Kazmir are potentially improving and struggling, respectively. A.J. Burnett appears to reach steady state performance at an CERA around 4 early and continues around that level throughout the season.
The mean value of the CERA statistic over time (i.e. over batters faced) for these five pitchers hovers near 2.4. However, to overcome the issues with such a small sample size the analysis is repeated with the 100 pitchers who faced the largest number of batters during the 2014 season. The instantaneous CERA profiles of these 100 pitchers is shown in Figure 2 with the mean CERA value shown in black and the profile for Kershaw indicated in red, to facilitate comparisons within the group of pitchers. These CERA profiles illustrate the expected high variance from the few data points at the beginning of the season followed by CERA stabilization later in the season as discussed previously.

What is clearly evident is the outlier performance of Kershaw during the 2014 season and thus his apparent deserving of the Cy Young Award, NL MVP, and other accolades [8]. Almost all of the other pitchers’ performances stabilize above Kershaw’s CERA profile during the second half of the season. Also, the group mean value for the CERA, starting around 100 batters and continuing to 800 batters, hovers near 3.4. The higher variance at the end of the season is a result of the lower number of pitchers that faced more than 850 batters in the season. Kershaw's performance during 2014 was at worst average for only a few batters in a row when his profile is above the mean line. However, this point occurs after he had faced only 100 batters and near the boundary or region of high variance.

The individual pitching performance of the five pitchers previously introduced, with seven other pitchers, compared against the mean CERA line is shown in Figure 3. Within this trellis chart or small multiple, a wider range of performance is observed from the aggregated data presented above in Figure 2 [9]. With identical axes for each panel, this chart enables better comparisons between pitchers over time (i.e. over batters) and highlights how they performed with respect to the mean performance.
Figure 3: Comparison of 12 individual MLB pitchers to the mean CERA profile.

Eric Stults (middle row) starts the season above the CERA average and remains there for the entire season. Edwin Jackson (top right) and Vidal Nuno (bottom row) also start and remain above the mean line, but do have portions of the season which approach the mean pitcher performance. However, while Nuno is improving for most of the season Jackson is struggling, especially after 200 batters. Burnett, Weaver and Greinke generally follow the mean line throughout the season with a few deviations at the beginning or end. Tim Hudson’s and Tim Lincecum’s 2014 performances are much more interesting. Hudson starts the season better than Kershaw but unfortunately slowly climbs back to the average CERA by the end of the season. After a rough start, Lincecum works his way down to a CERA level much lower than the mean line, but can only maintain this for a few games before a sharp turn and eventual ascent of his CERA profile.

These small multiples can be useful to identify trends or trigger points of significance in changing pitching performance. In particular, movement during the later half of the season is more significant because the larger number of batters essentially lessens those effects. Therefore, with the higher general slope of Lincecum’s profile after 550 batters, a manager would be wise to, at minimum, investigate the apparent change in performance potentially due to fatigue, injury, or pitching rotation and make the necessary adjustments. Even if nothing problematic is identified, and the profile is a simple result of the random nature of the pitcher’s performance, investigating a few false positives is much better than missing a true one.
3.2 Moving Averages

Since a CERA profile resembles that of a time series chart of a stock price, albeit in the opposite direction since one seeks lower CERA values, various strategies of technical analysis of the price of an investment vehicle could be potentially applied to these profiles. Although multiple trading tools and analyses are available [2], the oldest and most popular, the simple moving average, can provide additional insights into pitcher performance. (The exploration of other technical analysis tools applied to a pitcher’s CERA profile defines a direction for future research efforts.)

When a moving average (MA) is applied to the CERA profile, crossovers between MA's with “time periods” or different numbers of batters, can be used as trigger points for the identification of potential trends. Using financial MAs in the context of CERA can be useful to explore the establishment of upward or downward trends, momentum within that trend, support, and “stop losses” (i.e. time to pull or rest a pitcher).

For example, simple MA's with 75, 150 and 300 batters are applied to Tim Lincecum's CERA profile. The result is shown on the left hand side of Figure 4. Up until Lincecum has faced 480 batters, the 300-batter-MA is above the 150-batter-MA, which is, in turn, above the 75-batter-MA. The right hand side of Figure 4 shows this same data with an applied smoothing function. These features confirm a downward trend of Lincecum’s CERA and improvement in pitching performance during the first part of the season (see Figure 3). However, near batter 480, the 75-batter-MA crosses the 150-batter-MA signaling a potential trend reversal. Later on, both the 75 and the 150-batter-MA cross the 300-batter-MA, confirming this trend with increasing momentum, which continues for the rest of Lincecum’s season. With the 300-batter-MA continuing to rise, Lincecum finishes the 2014 season with a disappointing “Instantaneous” CERA well above the average.

On the other hand, Tim Hudson’s 75, 150, and 300-batter-MA are shown in Figure 5. Here the trend is reversed, where the MA’s are all generally moving higher during the first half of the season and Hudson’s CERA is climbing. Although this may simply be a case of an unusually successful start with the remainder of the season regressing to the mean, one should, at minimum, investigate potential reasons for this established trend and the signs of upward strength and momentum. Furthermore, his second half of the season does not significantly improve. Despite the 75 and 150-batter-MA crossing multiple times, there seem to be indications of actual support for this trend because these “lower batter MAs”
stay close to, or above, the 300-batter-MA. This is most pronounced on the smoothed data (right hand side of Figure 5) where the 75 and 150-batter-MA struggle to cut below the 300-batter-MA within the range of 650 to 850 batters. A pitching coach could potentially benefit from this analysis and explore ideas as to why these trends are so well supported and change is resisted.

Finally, in Figure 6, Kershaw’s CERA profile, and the same set of MA’s, start low and stay low throughout the season. A downward trend, similar to Lincecum’s, is established early in the season up to 400 batters followed by a trend reversal for a short period of time, over a small number of batters. However, Kershaw, or the managers, were evidently able to figure out what was wrong and make the adjustment to overcome the momentum in the wrong direction. Kershaw ends the season quite strong despite some signs of another upward trend during his last few appearances on the mound. Still, his 300-batter-MA is quite flat in comparison to Lincecum’s and Hudson’s, resulting in a much better overall season with a much lower final CERA statistic.

Figure 5: 2014 CERA profile for Tim Hudson with MA’s of 75, 150 and 300 batters. (Left - CERA MA data, right - smoothed CERA MA data.)

Figure 6: 2014 CERA profile for Clayton Kershaw with MA’s of 75, 150 and 300 batters. (Left - CERA MA data, right - smoothed CERA MA data.)
Similar MA analyses to that described above, using different “time periods,” could add additional insight on a shorter or longer time scale suggesting particular issues such as pitcher fatigue or discomfort. When certain indicators are compelling, some managers may decide to apply a “stop loss” and pull pitchers in the middle of games (e.g. when the 150-batter-MA crosses a 50-batter-MA). Since other pitchers may have different characteristic CERA profiles with associated MA’s and crossover points, using this richer data set and associated analyses could suggest other ways to properly manage schedules and pitching styles.

4 Conclusion

With significant investment into the player on the mound, baseball teams should be continually looking for advanced methods and analyses to track and manage their pitchers. Instantaneous CERA profiles and analyzing moving averages shows a promising metric to better identify the performance trends of pitchers. Once a trend is identified with these metrics, an investigation into the cause can inform decision makers or players about additional rest, adjustments, or other actions that should be implemented as necessary. As such, these metrics could be added to the arsenal of tools available to a manager making decisions regarding pitcher performance, pulling pitchers in the middle of a game, and the overall team’s pitching rotation and schedule.

Acknowledgements

The authors would like to acknowledge the efforts of Calen McLean for some of the data acquisition and processing stages.

References


