Baseball's 'Hot Hand' Is Real

After analyzing MLB's fastballs, we found that every pitcher has both an ace and a bum lurking inside them.

By Rob Arthur and Greg Matthews
Filed under MLB
Published Aug. 11, 2017

ILLUSTRATION BY KELSEY DAKE

e've all seen a pitcher when he's zeroed in: His mechanics are clean, his curveball is dropping off the table and he's painting the edge of the zone. But just as often, a hurler

W can lose command, and we see pitches that normally look sharp getting hammered into the stands. These streaks are confounding; for no apparent reason, a journeyman can look like a Hall of Famer, or a Cy Young winner can look like an ordinary junkballer.

Sabermetricians usually insist that such streaky performances are really just an artifact of fans and journalists forcing narratives onto random patterns. No matter how much it seems like a pitcher is getting hot, the more likely explanation is that they just happen to have bunched a few good innings together. But that's not quite correct: Using a new method that focuses on fastball velocity, we found a way to detect whether a pitcher is actually throwing with a hot hand — and just how big of a difference it can make.

Arguments over streakiness in baseball are almost as old as the sabermetric movement itself, and for a long time, research seemed to disprove the notion of a "hot hand" effect or argue that it played only a severely limited role. But although the sabermetric community felt certain that these streaks were just coincidence, nearly every baseball player insists that hot streaks are real. And fresh evidence has muddied the debate. Recent papers have called into question some of the foundational research on which statheads built their skepticism. It turns out there might be room for hot streaks after all.

Of course, even if streakiness exists, predicting whether a player will go on a tear is even harder than proving tears are real. But it's also much more useful. Knowing that a pitcher is on a hot streak, for instance, could tell you whether a team is primed for victory or defeat in a short playoff series. So we set out to find a way to forecast the hot hand using a method called a Hidden Markov Model. And not only did we find that hot streaks are real, but we can also tell whether a pitcher is hot or cold at any given moment.

The statistical details of a Hidden Markov Model are complex, but here's a nontechnical description of the idea. Imagine that two pitchers were throwing from behind a curtain: flamethrowing Stephen Strasburg and soft-tossing Jered Weaver. The curtain completely conceals their uniforms and arm mechanics, and the two pitchers are each only allowed to toss a few dozen fastballs in a row before the other player takes a turn. The only way to tell who's throwing is by looking at their pitches.

Despite the curtain, any dedicated baseball fan would be able to tell who's who. Weaver's average fastball clocks in at a pedestrian 84 miles per hour, while Strasburg's zooms by at 96. With a glance at the radar gun, you'd easily be able to identify which pitcher was throwing, and you'd have a pretty good guess at how fast the next pitch would be.

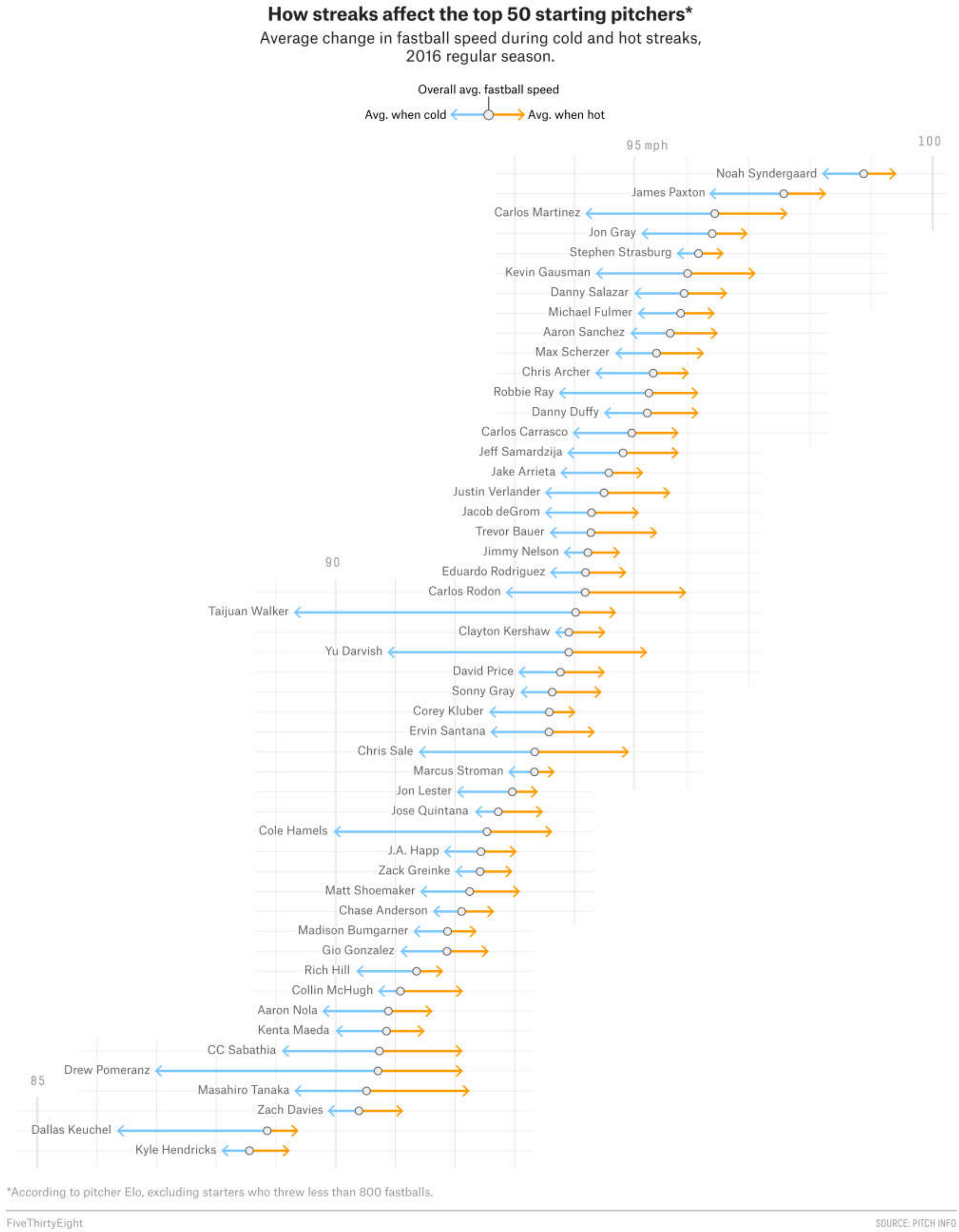
Our analysis assumes that deep within every hurler, there's a Strasburg and a Weaver lurking behind the curtain. When a pitcher is feeling healthy, or when he's had a good night's sleep, or when countless other factors fall into place to allow him to be in peak form, he can throw his hardest. On the other hand, if his mechanics are out of whack, or if he's having troubles in his personal life, then maybe he starts to toss more like Weaver. Using a sequence of fastball velocities recorded by MLB's pitch-tracking system[1] for each pitcher in the 2014-2016 seasons, our model guesses who's behind the curtain: the hot version or the cold version. We can then say whether a pitcher is likely to keep throwing hard or lose his flamethrowing touch.

According to conventional wisdom, Clayton Kershaw is a steady pitcher. But the data shows that even he runs hot and cold.

To do this, though, we need to eliminate some common factors that can influence fastball velocity but don't tell us anything about the pitcher's condition. First, we focused on each thrower's hardest types of fastball, which allowed us to avoid situations where pitchers were intentionally varying speeds, such as between two-seam and four-seam heaters.[2] Individual ballparks' speed measurements sometimes run hot or cold, which can make it look like a player is throwing harder or softer than usual, so we removed the effect of each stadium. We also accounted for the fatigue a pitcher experiences during the game, plus a host of other elements.[3]

Once we've done all that, we can go through a year's worth of fastball velocity readings and determine when a starting pitcher is hot or cold.[4] We found that the typical pitcher goes through 57 streaks in a season, jumping between hot and cold every 24 pitches. And not only did we find streaks, but we also found that the difference between being in the freezer and on fire is huge: The average pitcher moves up or down by an average of about two miles per hour when shifting between states.

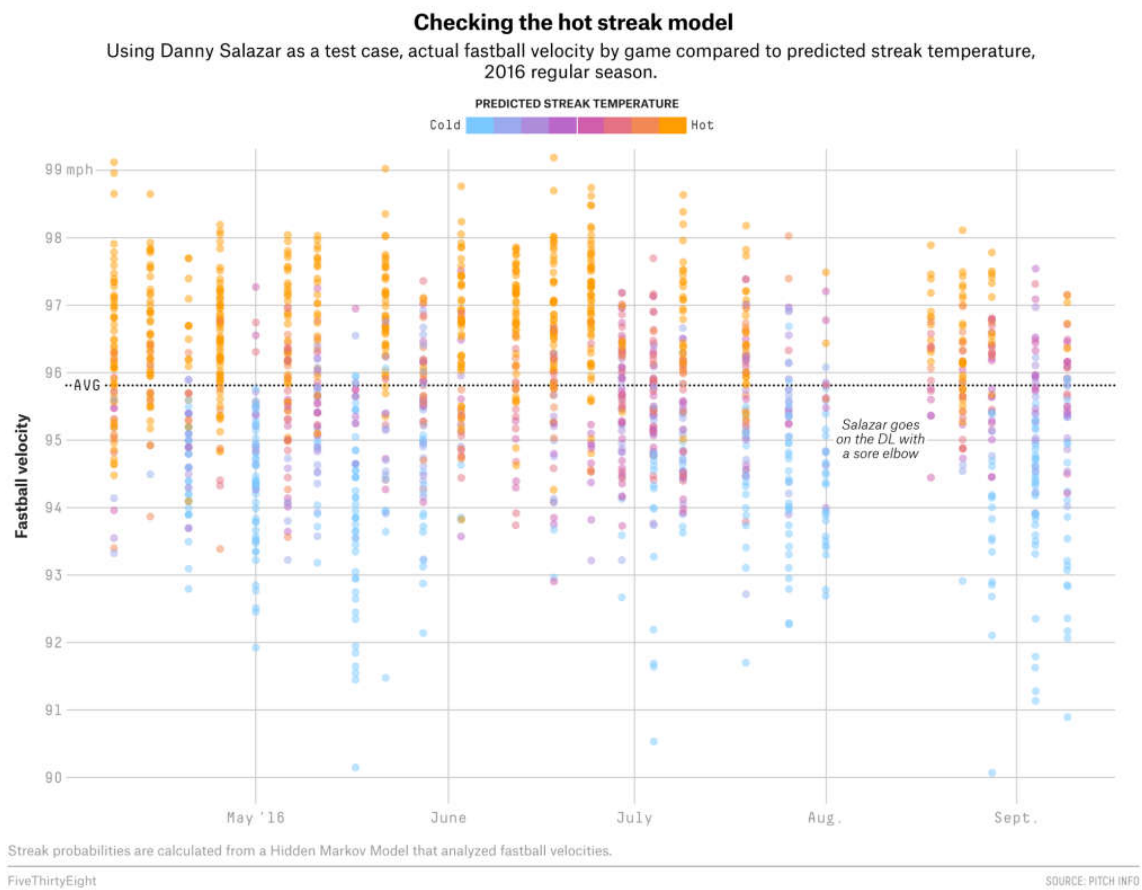
Every starting pitcher in our data shows a noticeable pattern of switching between hot and cold states.[5] Some pitchers' streakiness manifests in a pronounced downside when they're cold, like what happens to Texas Rangers lefty Cole Hamels. When he's on, Hamels' fastball is just a tick faster than average for him. But when he's off, he loses about two and a half miles per hour compared to his average fastball. The impact is massive: That almost 4 mph difference in heat translates to a 1.03-run difference in projected runs allowed per nine. If you apply those numbers to Hamels' 2016 season, he had ace-level stats when hot (3.41 RA/9, 18th among 73 qualified starters), and mediocre ones when off (4.44 RA/9, 44th).



What's more, the differences in Hamels' performance seem to be steady from year to year. Running the same analysis on his 2014 and 2015 seasons shows that Hamels always fluctuates between about 2.5 mph down when cold and 1 mph up when hot. He's not alone: Players who appeared in all three seasons we studied tended to show the same hot and cold effects from year to year, suggesting that we picked up on some of each pitcher's true characteristics rather than just noise.

Unlike Hamels, some pitchers tended to gain and lose velocity in roughly equal measure. Masahiro Tanaka and Carlos Rodon, for instance, both gained about 1.5 mph when on and lost the same amount when off. But in general, pitchers in cold stretches lost more velocity than they gained when they were hot.

Although our approach can tell us when pitchers are on or off, it can't tell us *why*. For some throwers, it's obvious. Cleveland Indians' star Danny Salazar suffered an elbow injury midway through the 2016 season, and his velocity tumbled, creating a cold streak, in his last two starts before he took a trip to the disabled list. It seems likely that the elbow was affecting Salazar's mechanics even before the issue became bad enough to take him out of the game.



Salazar's situation is not unusual. Many transitions between hot and cold streaks correspond with injuries. Pitchers on our list took 59 trips to the DL in 2016, and in 28 of those cases, they went through a notable frigid period in the two weeks before their injury took them out of the game.[6] The chance of that many pitchers going cold by chance just before a DL stint is very slim, so it's likely that our method can also detect injuries. In particular, we found evidence that clusters of several slow pitches in a row are associated with a hurt pitcher.

In other cases, it's not so clear what's driving a streak. Hamels, for example, has stayed relatively healthy over the last few years, but that hasn't stopped his fastball from fluctuating wildly. Any number of factors could explain Hamels' performance: Maybe he is especially reliant on getting good rest between starts, or maybe he's vulnerable to lingering, unreported medical issues — ones that never get bad enough to send him to the disabled list, but that diminish his performance nevertheless. Whatever the cause, his velocity does fluctuate in a meaningful way, so there must be more to it than noise.

Our approach isn't just backwards-looking, either — we can also predict whether a pitcher will be hot or cold in the future. Using just the first two months' worth of 2016 data, we tried to predict every pitcher's subsequent fastball velocity.[7] Our model was able to predict how hard the next pitch would be better than a guess based on the pitcher's season-long average would be able to, suggesting that it's able to pick up on when a pitcher is hot or cold at any point in the season after June 1. That's useful information for managers who are deciding when a starter might be getting cold and should be pulled from the game or which bullpen arm to turn to in a jam.

And the predictions don't just affect fastball velocity. The hotter a pitcher was projected to be on a given pitch, the greater the probability was that the next fastball would yield a swinging strike, even accounting for the increased speed of that heater. Similarly, a pitch was significantly less likely to go for a hit, and less likely to yield extra bases, if a pitcher was considered hot.[8] It's no surprise that good velocity is correlated with better results, but it reinforces that these hot streaks aren't just quirks in the speed readings — they also indicate that something extra that happens when a pitcher is sharp.

If hot streaks are real, as our results suggest, they probably apply to more than just pitchers. We started out looking at fastball velocity because it is one of the most consistent and important performance characteristics in baseball, but our method could be applied to anything, from how hard a batter is hitting the ball to how often he swings the bat.

As evidence mounts that player performances do vary from hot to cold, it's probably time to revisit the long-held sabermetric belief that such streaks are a fallacy. Whether a player is in the zone or in a funk, it's not just magical thinking — it actually tells us something real about how he will play in the near future. And although our eyes might not be especially reliable indicators of a pitcher's state, fastball velocity doesn't lie.

Special thanks to Harry Pavlidis, Pitch Info and Rob McQuown.

Footnotes

1. Specifically, the data we used relied on PITCHf/x, the system MLB used to track pitches before it switched to Statcast this season.
2. We used Pitch Info data to discern which fastballs are which. In rare cases in which a pitcher tended to use a cutter as his dominant fastball, we analyzed cutters instead of or in addition to four-seam fastballs.
3. Here's a more technical explanation of our methodology: We begin by running a generalized linear model that predicts a pitch's velocity based on who's pitching and the number of pitches they've thrown. Then we took the residual from that model and controlled for external factors such as whether the bases were occupied, stadium elevation, park effects and game-time temperature, all of which can affect pitch velocity. We also looked at a number of other contextual elements, such as the difference in the teams' scores, but we found that they did not significantly impact velocity. Finally, we applied a Hidden Markov Model to the residuals of the second model, searching for two hidden states — corresponding to hot and cold — within each pitcher.
4. We limited it to pitchers who had thrown at least 800 fastballs in the 2016 season.
5. We guarded against overfitting by scrambling the order of the fastball velocities for each pitcher and re-running the Hidden Markov Model. These randomized, fictitious sets of velocities produced smaller spreads between hot and cold states, as well as more random transition matrices than we saw in the real data, thus confirming that our model picked up on real streakiness in the authentic data.
6. Specifically, their cold-state probability spiked above the 99th percentile of what you would expect to see if you randomly selected from their pitches.
7. We ran our model on the first two months of pitches for each pitcher, and then used the Viterbi algorithm to generate updated state estimates for every subsequent pitch in the season.
8. We used generalized linear models that examined the hot-state probability of the previous pitch, the count, and the current pitch's velocity as predictors. All p-values were less than p=.05.