R-E-S-P-E-C-T
(or why you shouldn't read the ESPN MLB power rankings)

Phillip G. Rogers

6/24/2022

What are power rankings for?

Power rankings are ubiquitous these days. You can find them on all the major sports news websites for every major sport. For many sports, they come in weekly installments, offering the sports writer frequent opportunities to discuss the progress of each team (or at least those that are actually contending for a title, for those power rankings that only include the top teams).

But what do these power rankings attempt to capture? Surely, it cannot just be a reflection of the standings. If that were the point, then the writers could just paste in a copy of the standings. Perhaps the power rankings are intended to be a sort of prediction going forward. This makes sense for preseason rankings. But if this is always the case, then why do they often come in weekly editions during the season? A couple of wins or losses in a row doesn't drastically change how we think a team will perform going forward, and advances in statistical analyses support this idea. In a sport like baseball, for example, flukey games and weeks happen all the time, and predictions based on larger sample sizes are much more reliable. So, if power rankings are about prediction, they will end up being pretty similar from week to week.

Most people understand that the primary purpose of frequent, in-season power rankings is to show which teams are 'hot' and which are not. In other words, they tell us about a team's recent success (or lack thereof). Now, this recent success must be balanced with a team's overall success or position in the standings. If not, we could just rank teams based on their winning percentage over the last week or month. Again, something like this is already found in the standings, usually in the form of a team's win-loss record over the past ten games. So what is the perfect balance between overall and recent success? There is no single right answer, and this makes power rankings interesting. They are subjective, reflecting the observations and opinions of the author(s). However, when we read power rankings, we should expect to see the influence of both overall success and recent success.

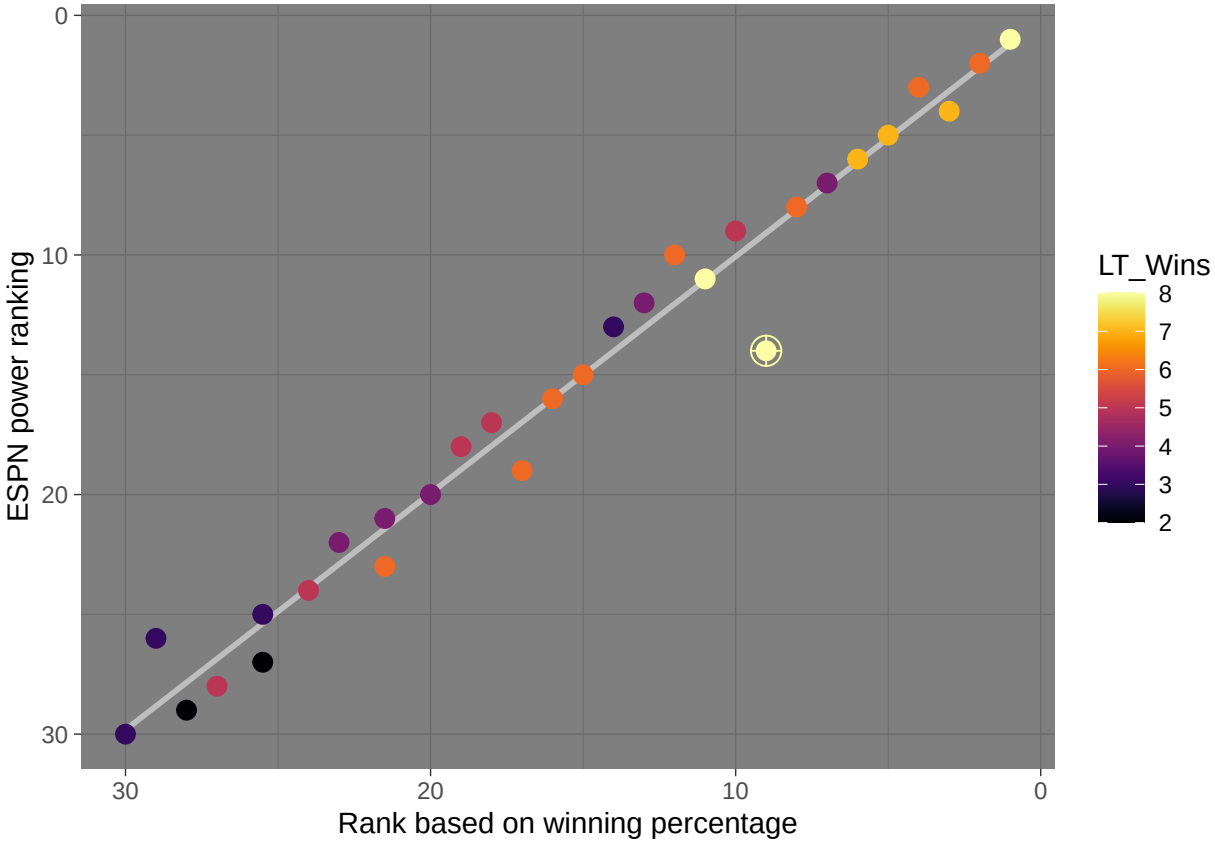
Checking in on the Guardians

The Cleveland Guardians have been much better this year than people expected. This offseason they dealt with the loss of superstar Francisco Lindor and other financially pragmatic roster changes that we've come to expect from a low-budget MLB team. And yet they've come on strong after a mediocre start to the season. As of Wednesday, June 22nd, the Guardians had won 8 of their last 9 games and 17 of their last 21. During this stretch, they turned a 5-game deficit to Minnesota in the division race to a 1-game division lead.

So, given their recent success, it was surprising to see them sitting at #14 in the most recent ESPN power rankings. It seemed to me that the Guardians should be getting a little more respect. It also made me wonder: Could this ranking be justified by their overall performance relative to other teams? Or could it be justified by recent performance? Perhaps, despite the Guardian's hot streak, there are other teams that have been even hotter?

Analysis

To answer these questions, I pulled some basic MLB data. In addition to each team's rank in the power rankings, I collected their overall win percentage (overall success) and wins in the past ten games (recent success). The overall winning percentage was converted to a winning percentage *rank*, with teams ranked from the best winning percentage to the worst. We can use these to try to understand what's motivating the power rankings.



The plot above shows the relationship between a team's winning percentage rank and their current ESPN power rank. Each of the 30 dots in the plot represent a team. Teams further to the right have a better winning percentage rank, and teams higher in the plot have a better power rank. For example, the dominant Yankees are the yellow dot at the top right corner of the plot; they have the best winning percentage in the majors and they are currently ranked #1 in the ESPN power rankings. Finally, the color of each dot represents the number of wins that team has in its last ten games. Orange and especially yellow dots represent teams that have won most of their recent games, while purple and especially black dots represent teams that have lost most of their recent games.

It's easy to see that the dots in the plot form a rather neat line from bottom left to top right. This is evidence that the power rankings are based largely on a team's overall success. In general, as a team's winning percentage goes up, so does its position in the power rankings. This relationship is represented by the gray line behind the dots. However, some teams fall above or below that line. These are teams whose power rank is better or worse than their winning percentage rank. Dots above the gray line are ranked *higher* in the power rank than in the winning percentage rank, and dots below the gray line are ranked *lower* in the power rank than in the winning percentage rank.

The outlier

The yellow dot with a target on it represents the Cleveland Guardians. They are further from the gray line than any other team, meaning that they have the largest difference between their winning percentage rank and power rank. The fact that they are below the gray line means that their power rank is very poor relative to their winning percentage rank. In fact, there are five teams that are ranked *higher* in the power rankings than the Guardians (higher in the plot) despite having *worse* records (further left in the plot). So, just based on overall success this season, one could argue that the Guardians are the least respected team in the power rankings.

However, I said earlier that power rankings should also reflect a team's recent performance to some degree. In theory, this should help to explain the discrepancy between winning percentage rank and power rank for some teams. Teams that have been hot lately should get a boost in the power rankings, and we should find these teams above the gray line. Meanwhile, teams that have played poorly lately should get poor power ranks relative to their winning percentage, and we should find these teams below the gray line. By this logic, the Guardians' current power rank would make sense if they had struggled lately, but they've been red hot! In fact, the Guardians have been as good as anyone over their last 10 games, one of three teams to win 8 of their last 10. It stands to reason that their power rank should be even *better* than their winning percentage rank, but instead it is much *worse*.

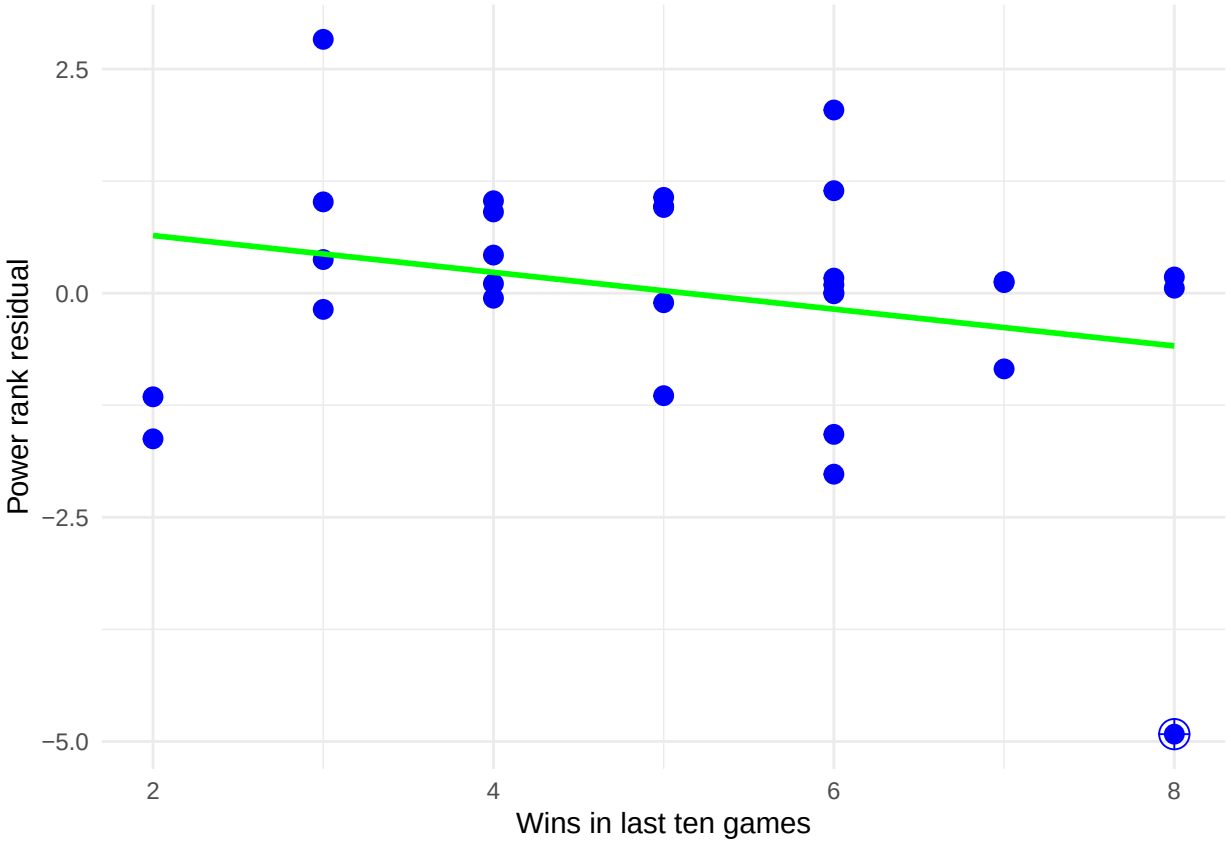
This egregious error in power rank is even worse when we consider those five teams ahead of the Guardians in the power rankings. I mentioned that these five teams have worse records than the Guardians, and only one has played as well as the Guardians recently. Like the Guardians, the Red Sox have won 8 of their last 10, but the other four teams have won 6, 5, 4, and 3 of their last 10, respectively. For example, the Rays have a considerably worse record than the Guardians and have only won 3 of their last 10–yet they sit one spot ahead of the Guardians in the power rankings! It seems as though neither overall success nor recent success can justify the decision to rank the Guardians so low in the power rankings.

There is one more possible explanation for the Guardians' low ranking. Could it be that their recent hot streak is the result of playing weak teams? This was partly true–for last week's power rankings! A week ago, the Guardians were wrapping up a series against the Rockies after winning series against the Rangers and Athletics. These teams range from being barely under .500 (Rangers) to having the worst record in MLB (Athletics). That week's edition of ESPN's power rankings made this exact case against the Guardians. They were ranked #15, writing, "We're about to find out if Cleveland can transition its success against lesser lights into wins against MLB's elite." Since that was published, the Guardians merely finished their sweep of the Rockies, won 2 of 3 against the mighty Dodgers, and then won 2 straight against the Twins to pass them for the division lead. So, how were the Guardians rewarded for continuing their strong play against playoff-calibur teams? They were bumped up just 1 spot to #14 in the current rankings, 2 spots behind a Twins team they just passed in the standings and who had won only 4 of their last 10 games.

Is it just the Guardians?

The plot above shows pretty clearly that the Guardians are an outlier in the most recent ESPN rankings. They are ranked way too low relative to both their overall record and recent performance. But is that the only team the power rankings got wrong? And more generally, are the power rankings even doing what they're supposed to do?

Earlier, I discussed how the linear trend in the plot reflects a strong relationship between the overall success of a team and its rank in the power rankings. I also described what we should see in the plot if the power rankings are influenced by recent success: teams that have played well recently should generally be found above the gray line, while teams that have played poorly of late should generally fall below it. The Guardians certainly do not follow this trend, but we can see if the rest of the MLB does using the residuals of the plot. The residual of each data point (e.g., team) is the distance above or below the prediction line (the gray line). We can test if these residuals correlate with the number of wins a team has in its last 10 games.



Our new plot shows the relationship between recent success and power rank residual among MLB teams, with each dot representing a team. Teams that have played well over the last ten games are on the right side of the plot, while teams that have not played well are on the left. The power rank residual is just the difference between a team's actual power rank and its expected power rank based on record. A positive residual (higher in the plot) means the team is power ranked high relative to its record, while a negative residual (lower in the plot) means the team is power ranked low relative to its record. For example, the Guardians are the dot in the lower right corner; they have 8 wins in their last 10 games, but their low residual value is due to their relatively poor rank in the power rankings.

Teams that have played well in their last ten games are 'hot', and more often than not we should expect them to have positive residuals. After all, they should get a boost in the power rankings due to their recent play. Similarly, we should expect teams that have not played well to have negative residuals. However, we find the *opposite* trend! There is some variation in the data, but overall the 'hot' teams on the right side of the plot show lower residuals than the slumping teams on the left side of the plot. The green line in the plot shows this trend, declining as you move left to right.

This finding calls into question the integrity of the ESPN power rankings as a whole. Overall, the rankings strongly reflect a team's overall success. A better overall record corresponds closely to a better rank in the power rankings. However, where teams deviate from that correspondence, it appears as though these power rankings are biased *against* teams that have played well recently. Teams with *more* wins in their last 10 games are likely to be ranked *lower* (worse) than their overall record suggests. In contrast, teams that have struggled lately are ranked *higher* (better) than one would expect.

What have you done for me long ago?

Power rankings are supposed to capture current trends by rewarding surging teams and punishing faltering ones. The question, "What have you done for me lately?", is what makes power rankings dynamic and relevant from one week to the next. However, I have shown how the latest ESPN power rankings have managed to do just the opposite! In these rankings, streaking teams are punished and slumping teams are rewarded. This approach is comically conservative, ignoring recent results in favor of what a team looked like *before* they went on their recent run (or slide). No team has borne the brunt of this misguided approach more than the Guardians. Despite continuing to be one of the hottest teams in the majors–even against tough competition–they are ranked below five teams with worse records.

All this begs the question: What's a team gotta do to get a little respect around here? Don't ask ESPN–they're too focused on your past.