Rethinking WAR for Pitchers: Grid WAR

Ryan Brill, Justin Lipitz, Emma Segerman, Ezra Troy, Adi Wyner

January 14, 2021

Abstract

Traditional methods of computing WAR (wins above replacement) for pitchers have serious flaws. Specifically, Fangraphs and Baseball Reference compute a pitcher's WAR as a function of his performance averaged over the entire season, which is problematic because it ignores a pitcher's game-by-game variance. Hence we propose a new framework for computing a pitcher's season-long WAR as the summation of the WAR of each of his individual game performances.

1 Introduction

Traditional methods for computing WAR (wins above replacement) for pitchers, notably by Baseball Reference and Fangraphs, calculate WAR as a function of a pitcher's average performance. Specifically, in computing WAR, Baseball Reference averages a pitcher's performance over the course of a season via its xRA statistic [CITE]. xRA, or "expected runs allowed", denotes a pitcher's average number of runs allowed per out. Additionally, Fangraphs averages a pitcher's performance over the course of a season via its $ifFIP$ statistic [CITE]. $ifFIP$, or "fielding independent pitching (with infield flies)", is defined by

$$ifFIP := \frac{13 \cdot HR + 3 \cdot (BB + HBP) - 2 \cdot (K + IFFB)}{IP} + ifFIPconstant,$$

which involves averaging some of a pitcher's statistics over his innings pitched. In this paper, we argue that using a pitcher's average performance to calculate his WAR is a bad way to measure his value on the mound, and so we derive an alternative approach to computing WAR.

Consider Max Scherzer's six game stretch from June 12, 2014 through the 2014 all star game [CITE]. His performance over those six games is shown in table (1).

game	1	2	3	4	5	6	total
earned runs	0	10	1	2	1	1	15
innings pitched	9	4	6	7	8	7	41

Table 1: Max Scherzer's performance over six games prior to the 2014 all star break.

In Scherzer's six game stretch, he averages 15 runs over 41 innings, or 0.366 runs per inning. So, on average, Scherzer pitches 3.3 runs per complete game, which isn't great. If we look at each of Scherzer's individual games separately, however, we see that he has five dominant performances and one blowup, which is essentially equivalent to five wins and one loss. Pitching 3.3 runs per complete game is certainly not as stellar as five nearly-guaranteed wins and one loss. On this view, a WAR statistic based upon averaging a pitcher's performance significantly devalues Scherzer's contributions during this six game stretch. Because "you can only lose a game once", it makes more sense to give Scherzer zero credit for his one blowup game than to distribute his one poor performance over all his other games via averaging. Hence we should not compute WAR as a function of a pitcher's average performance, and should compute season-long WAR as the summation of the WAR of his individual games.

We mathematically ground this intuition as follows: WAR should be a nonlinear function. In other words, let's think of WAR as a function of R and I, which refers to a pitcher allowing R runs over I innings. Then, letting R_j denote a pitcher's number of runs allowed in inning $j \in \{1,...,n\}$ of the season, we expect

$$WAR\left(\sum_{j=1}^{n} R_j, n\right) \neq \sum_{j=1}^{n} WAR(R_j, 1), \tag{1}$$

which is to say that WAR is a nonlinear function. In English, equation (1) tells us that the WAR of a pitcher's performance aggregated over a season is not equal to the aggregation of the WAR of a pitcher's individual game performances. As discussed previously, traditional WAR statistics compute a WAR function reminiscent of the left side of equation (1), whereas we aim to develop a WAR function based on the right side of equation (1).

2 Computing Grid War

2.1 Overview

We wish to compute a pitcher's season-long WAR as the sum of the WAR of his individual game performances. We begin by computing a pitcher's "win probability added" during game i, which we define as the difference in his team's win probability just after his last pitch and just before his first pitch of the game. We use statistical techniques to estimate this difference in win probability as a function of the pitcher's number of earned runs, starting inning number, ending inning number, base-runner configuration at the time he starts and finishes pitching, and number of outs at the time he starts and finishes pitching. Then, we estimate the expected win probability added if the pitcher had been replaced a hypothetical replacement-level pitcher. We then define a pitcher's WAR in game i is as the difference between his win probability added and that of the hypothetical replacement level pitcher. A pitcher's WAR over the whole season is thus defined as the sum of the WAR of his individual games.

2.2 Win Probability Added

In order to define the win probability added by pitcher p during game i, we need to first define two helper functions. First, we define the function $f = f(I,R)$ which computes a team's probability of winning a baseball game after giving up R runs through I complete innings (averaged over all other confounders such as the performance of his team's batters). Second, we define the function $g = g(R,S,O)$ which computes the probability that, starting midway through an inning with $O \in \{0,1,2\}$ outs and men on base given by $S \in \{000,100,010,001,110,101,011,111\}$, a team scores exactly R runs through the end of the inning. Because f is a function of two discrete variables $I \in \{1,2,...,9\}$ and $R \in \{0,1,...27\}$ (where 27 serves as the maximum number of runs allowed to be pitched in a game), f is fully specified by a 2D grid. Similarly, if we combine S and O into a 24-state tuple, then we may also view g as a function of two discrete variables R and (S,O), and so g is also fully specified by a 2D grid. As we end up defining a pitcher's WAR in terms of these helper functions f and g, which are specified by 2D grids, we decide to name our metric *Grid WAR*.

Now, using these helper functions f and g, we wish to compute a pitcher's win probability added during a game. We begin by defining a pitcher's win probability added during a stretch of complete innings pitched. Suppose a pitcher begins pitching at the start of inning I_0, and that his team has already given up R_0 runs. Then suppose he pitches a number of complete innings, and exits the game at the end of inning I_1 having thrown R_1 runs. Then we define his win probability added during this time as

$$f(I_1, R_0 + R_1) - f(I_0 - 1, R_0).$$

We use the convention that $f(I = 0, \cdot) = 1/2$.

We now define a pitcher's win probability added during a partial inning. Suppose a pitcher begins pitching at the start of inning I, and that his team has already given up R_0 runs. Then suppose he gives up R_1 runs, and exits the game midway through inning I with base-state S and number of outs O. Then we define his win probability added during the beginning-of-this-inning (*boi*) as

$$boi(I, R_0, R_1, S, O) := \sum_{r \geq 0} g(r, S, O) f(I, R_0 + R_1 + r) - f(I - 1, R_0).$$

Alternatively, suppose a pitcher begins pitching midway through inning I, starting with base-state S and number of outs O. Then suppose he exits the game at the end of inning I having given up R_1 runs. Suppose also that his team gives up R_0 runs through inning I, and R_0' runs during inning I but before he begins pitching. Then we define his win probability added during the end-of-this-inning (*eoi*) as

$$eoi(I, R_0, R_0', R_1, S, O) := f(I, R_0 + R_0' + R_1) - f(I - 1, R_0) - boi(I, R_0, R_0', S, O).$$

Lastly, suppose a pitcher begins pitching midway through inning I, starting with base-state S_0 and number of outs O_0. Then suppose he gives up R_1 runs, and exits the game with base-state S_1 and number of outs O_1. Suppose also that his team gives up R_0 runs through inning I, and R_0' runs during inning I but before he begins pitching. Then we define his win probability added during the middle-of-this-inning (*moi*) as

$$\begin{aligned} moi(I, R_0, R_0', R_1, S_0, O_0, S_1, O_1) :=& f(I, R_0 + R_0' + R_1) - f(I - 1, R_0) \\ & - eoi(I, R_0, R_0', R_1, S_1, O_1) - boi(I, R_0, R_0', S_0, O_0). \end{aligned}$$

Now, we have defined the win probability added by a pitcher during any inning, whether it is a complete or partial inning. Then, we define the win probability added by a pitcher over the course of the season as the sum of the win probability added over all of his innings pitched. Now, we need only estimate the helper functions f and g from the data in order to compute each pitcher's win probability added over a season.

2.3 Wins Above Replacement & the Cascading Replacement Paradigm

We have formulated a method to compute a pitcher's win probability added over the course of a season, and we shall use this to define Grid WAR (GWAR). A WAR metric is intended to compare a pitcher to a *replacement*-level pitcher, which is someone who would substitute for a pitcher if he got injured. We wish to define a pitcher's GWAR by the difference in his win probability added over the course of a season and that of a replacement-level player. Hence we need to quantify the performance of a replacement-level pitcher.

In general, the replacement for a starter is a low-tier starter, the replacement for a closer is a late-inning-middle-reliever, the replacement for a late-inning-middle-reliever is a mid-game-middle-reliever, and the replacement for a mid-game-middle-reliever is a low-tier mid-game-middle-reliever. Naturally, we call this system of replacement the *cascading replacement paradigm*.

3 Computing the Grid Functions

3.1 Estimating f

3.2 Estimating g

3.3 Replacement-Level Computations

4 Examples

5 Conclusion

Future Work: potentially modify win probability added to include other confounder's like the pitcher's team's runs, opposing fielder quality, etc

much better way to compute WAR!