Supplementary Materials for

"openWAR: An Open Source System for Evaluating Overall Player Performance in Major League Baseball"

A openWAR Package

Code for the openWAR package is available for download on GitHub at https://github.com/beanumber/openWAR.

B Previous Implementations of WAR

The major components of each existing implementation of WAR are summarized in Table 1. The details of how each of these components is calculated are beyond what we can present here. Instructions for how to reproduce these numbers are illustrative, but not rigorous (Slowinski, 2010; Forman, 2010, 2013a,b; Tango, 2008; Lichtman, 2010; Wyers, 2013). At best, the authors may provide a step-by-step example calculation, but never specify a statistical model in formal notation, nor do they include code that would unambiguously reveal the algorithms used. The Baseball Info Solutions (BIS) data set, which is used to compute the fielding component of $rWAR$ and $fWAR$, is proprietary, and thus cannot be part of a reproducible piece of scholarship in which the results (as opposed to the models or algorithms) are the primary contribution. The high cost of obtaining this data (tens of thousands of dollars per year) prevents all but a few persons from verifying any results that stem from its use. The fielding metrics used by those two implementations, Defensive Run Saved (DRS) and Ultimate Zone Rating (UZR) are also proprietary. While extensive descriptions of each have been published (Fangraphs Staff, 2013; Lichtman, 2010), they too are illustrative—rather than specific—and none include source code. Occasionally, these organizations publish "bug" fixes (Hamrahi, 2013) or updates (Appelman, 2010) that change previously published point estimates. Baseball Prospectus has announced plans to include more uncertainty and transparency in $WARP$ (Wyers, 2013), but it is not known if this will include a release of source code[1].

C MLBAM data set

There are two main open sources of baseball data. Lahman (2013) maintains a database of seasonal data that has also been packaged for R by Friendly (2013). However, this data does not contain play-by-play information, making it insufficiently granular for WAR-type calculations, especially

[1] Incidentally, no further uncertainty updates to WARP have been published since Wyers left Baseball Prospectus to joined the Houston Astros in November 2013.

	$rWAR$	$fWAR$	$WARP$
Data Source	BIS	BIS	Retrosheet
Batting	modified $wRAA$	$wRAA$	Linear Weights
Baserunning	Baserunning Runs	BsR	Baserunning Runs Above Average
Fielding	DRS	UZR	Fielding Runs Above Average
Pitching	Runs Allowed	FIP	Pitching Runs Above Average

Table 1: Comparison of WAR implementations Forman (2013b). The Baseball Info Solutions (BIS) data source is proprietary. Defensive Runs Saved (DRS) is a proprietary fielding metric developed by BIS. Ultimate Zone Rating (UZR) is a proprietary fielding metric developed by Lichtman (2010) and licesensed to Fangraphs.

with respect to fielding. Retrosheet (Smith (2013)) is an excellent source of free play-by-play data, but the batted ball locations are discrete, rather than continuous. That is, each batted ball is reported as falling into one of several dozen pre-defined polygonal zones. This level of detail is sufficient for some sophisticated defensive metrics, such as Humphreys (2011), but not others, such as UZR or SAFE (Jensen et al., 2009). Both of these data sources are updated periodically (usually at the end of the season).

As noted in the paper, *openWAR* uses data obtained from MLBAM. This data is not *libre*, but it does reside on a publicly-available web server, making it *gratis*. Furthermore, it is updated in real-time, and contains (x, y)-coordinates for each batted ball in every major league game. The R package (Baumer and Matthews, 2013) which has been developed simultaneously, will retrieve all data necessary to compute *openWAR*. The package contains simple R functions that will enable any user with an Internet connection to download the data of their choice.

The data available through this package is generally accurate. For example, summary statistics aggregated by team from all 184,739 observations in 2012 are shown in Table 2 in the Appendix, next to the corresponding figures available through the Lahman database (Lahman, 2013). The agreement between the numbers presented in Table 2 is over 99.8%[2], indicating that the data collected and processed by *openWAR* is of high fidelity.

In Table 3, we list the 31 event types in the MLBAM data set, along with their frequencies of occurrence in 2012.

Nevertheless, there are some significant limitations to this data set (Fast, 2009). It is important to note that these data are collected for the purposes of entertainment (e.g. feeding the MLBAM web application) and not for the purposes of data analysis.

[2]Specifically, the ratio of the Frobenius norm of the difference between the two sets and the Frobenius norm of the Lahman set is very small.

team	G	PA	AB	R	H	HR	BB	K	G	PA	AB	R	H	HR	BB	K
ana	162	6121	5537	766	1517	187	449	1112	162	6120	5536	767	1518	187	449	1113
ari	162	6152	5466	734	1414	165	539	1266	162	6148	5462	734	1416	165	539	1266
atl	162	6126	5427	699	1339	149	567	1289	162	6125	5425	700	1341	149	567	1289
bal	162	6160	5562	712	1374	214	480	1315	162	6158	5560	712	1375	214	480	1315
bos	162	6200	5636	737	1460	166	430	1204	162	6166	5604	734	1459	165	428	1197
cha	162	6111	5518	748	1409	211	461	1202	162	6111	5518	748	1409	211	461	1203
chn	162	5967	5411	613	1295	137	447	1235	162	5967	5411	613	1297	137	447	1235
cin	162	6115	5477	669	1375	172	481	1266	162	6115	5477	669	1377	172	481	1266
cle	162	6196	5526	667	1385	136	555	1087	162	6195	5525	667	1385	136	555	1087
col	162	6183	5584	758	1525	166	450	1213	162	6176	5577	758	1526	166	450	1213
det	162	6119	5477	726	1465	163	511	1103	162	6119	5476	726	1467	163	511	1103
hou	162	6014	5409	583	1276	146	463	1364	162	6012	5407	583	1276	146	463	1365
kca	162	6151	5638	676	1492	131	404	1032	162	6149	5636	676	1492	131	404	1032
lan	162	6091	5438	637	1367	116	481	1156	162	6091	5438	637	1369	116	481	1156
mia	162	6059	5440	610	1329	137	484	1228	162	6056	5437	609	1327	137	484	1228
mil	162	6226	5559	776	1443	202	466	1240	162	6224	5557	776	1442	202	466	1240
min	162	6209	5562	701	1446	131	505	1069	162	6209	5562	701	1448	131	505	1069
nya	162	6231	5524	803	1462	245	565	1175	162	6231	5524	804	1462	245	565	1176
nyn	162	6091	5454	650	1356	139	503	1250	162	6089	5450	650	1357	139	503	1250
oak	162	6187	5532	714	1317	195	550	1386	162	6183	5527	713	1315	195	550	1387
phi	162	6174	5546	684	1413	158	454	1094	162	6172	5544	684	1414	158	454	1094
pit	162	6014	5412	651	1311	170	444	1354	162	6014	5412	651	1313	170	444	1354
sdn	162	6112	5425	651	1336	121	539	1237	162	6112	5422	651	1339	121	539	1238
sea	162	6061	5499	621	1285	149	466	1259	162	6057	5494	619	1285	149	466	1259
sfn	162	6200	5559	718	1492	103	483	1097	162	6200	5558	718	1495	103	483	1097
sln	162	6326	5624	765	1524	159	533	1192	162	6326	5622	765	1526	159	533	1192
tba	162	6106	5401	697	1289	175	571	1324	162	6103	5398	697	1293	175	571	1323
tex	162	6216	5592	808	1523	200	478	1103	162	6214	5590	808	1526	200	478	1103
tor	162	6137	5525	723	1353	200	478	1255	162	6093	5487	716	1346	198	473	1251
was	162	6221	5615	729	1467	194	479	1325	162	6221	5615	731	1468	194	479	1325

Table 2: Cross-check between MLBAM data collected by openWAR (left) and Lahman data (right), 2012. These data are aggregated by team from 187,739 observations.

D Converting Runs to Wins

As changes in the run expectancy matrix are measured in *runs*, but the units of WAR are *wins*, it is necessary to convert runs to wins. A common convention used by all providers of WAR is that 10 runs is equivalent to 1 win (Cameron, 2008). This value can thought of as a slope in the relationship between runs and wins at a point representing the average team. More specifically, this value can be derived as the partial derivative of Pythagorean Win Expectation evaluated at a specific point.

Consider the general form of James' formula for *expected winning percentage*, which is derivable if run scoring follows independent Weibull distributions (Miller, 2007). That is, with p equal to a parameter (originally 2), then

$$WPct_p(RS, RA) = \frac{1}{1 + \left(\frac{RA}{RS}\right)^p}$$

where RS and RA are the runs scored and allowed by a team, respectively. The gradient of this

Event Type	N	Frequency
Strikeout	36286	0.196
Groundout	35266	0.191
Single	27954	0.151
Flyout	24890	0.135
Walk	13660	0.074
Pop Out	9072	0.049
Double	8221	0.045
Lineout	6666	0.036
Home Run	4937	0.027
Forceout	3984	0.022
Grounded Into DP	3613	0.020
Field Error	1705	0.009
Hit By Pitch	1494	0.008
Sac Bunt	1478	0.008
Sac Fly	1213	0.007
Intent Walk	1056	0.006
Triple	927	0.005
Double Play	494	0.003
Runner Out	463	0.003
Bunt Groundout	410	0.002
Fielders Choice Out	352	0.002
Bunt Pop Out	209	0.001
Strikeout - DP	146	0.001
Fielders Choice	114	0.001
Fan interference	46	0.000
Batter Interference	35	0.000
Catcher Interference	23	0.000
Sac Fly DP	11	0.000
null	5	0.000
Bunt Lineout	4	0.000
Triple Play	3	0.000
Sacrifice Bunt DP	2	0.000

Table 3: Frequency of Events in MLBAM data set (2012)

function is

$$\nabla WPct_p(RS, RA) = \left\langle \frac{\partial WPct_p}{\partial RS}, \frac{\partial WPct_p}{\partial RA} \right\rangle = \frac{p \cdot (RA/RS)^p}{(1 + (RA/RS)^p)^2} \cdot \left\langle \frac{1}{RS}, -\frac{1}{RA} \right\rangle .$$

Thus, if $r = RS = RA$ (as it will be for an average team), then this becomes:

$$\nabla WPct_p(r, r) = \frac{p}{4} \left\langle \frac{1}{r}, -\frac{1}{r} \right\rangle = \frac{p}{4r} \cdot \langle 1, -1 \rangle .$$

The gradient points in the direction of scoring more runs and allowing fewer, and from the magnitude we recover that the number of runs associated with one win over a 162 game season is:

$$\text{Runs per Win}_p(r) = \left(\frac{p}{4r}\right)^{-1} /162 = \frac{2r}{81p}.$$

The optimal choice of the parameter p may depend on the run-scoring environment. While James originally chose $p = 2$ for convenience, better fits for Major League Baseball have been obtained using $p = 1.83$ (Davenport and Woolner, 1999) and $p = 1.86$ (Tung, 2010). The average number of runs scored per 162 games has been approximately $r = 714$ since 1901, and $r = 761$ since the league expanded to 30 teams in 1998. Reasonable choices for p and r will yield conversion factors in the neighborhood of 10.

References

Appelman, D. (2010), "UZR Updates!" http://www.fangraphs.com/blogs/uzr-updates/.

Baumer, B. and Matthews, G. J. (2013), *openWAR: An Open Source System for Overall Player Performance in Major League Baseball*, http://github.com/beanumber/openWAR/.

Cameron, D. (2008), "Win Values Explained: Part Five," http://www.fangraphs.com/blogs/win-values-explained-part-five/.

Davenport, C. and Woolner, K. (1999), "Revisiting the Pythagorean Theorem," http://baseballprospectus.com/article.php?articleid=342.

Fangraphs Staff (2013), "DRS," http://www.fangraphs.com/library/defense/drs/.

Fast, M. (2009), "Confessions of a DIPS apostate," http://www.hardballtimes.com/confessions-of-a-dips-apostate/.

Forman, S. (2010), "Player Wins Above Replacement," http://www.baseball-reference.com/blog/archives/6063.

— (2013a), "Position Player WAR Calculations and Details," http://www.baseball-reference.com/about/war_explained_position.shtml.

— (2013b), "WAR Comparison Chart," http://www.baseball-reference.com/about/war_explained_comparison.shtml.

Friendly, M. (2013), *Lahman: Sean Lahman's Baseball Database*, r package version 2.0-3.

Hamrahi, J. (2013), "Replacement Level and 10-Year Projections," http://www.baseballprospectus.com/article.php?articleid=19910.

Humphreys, M. (2011), *Wizardry: Baseball's All-time Greatest Fielders Revealed*, Oxford University Press.

Jensen, S. T., Shirley, K. E., and Wyner, A. J. (2009), "Bayesball: A Bayesian hierarchical model for evaluating fielding in major league baseball," *The Annals of Applied Statistics*, 3, 491–520.

Lahman, S. (2013), "Sean Lahman's Baseball Database," http://www.seanlahman.com/baseball-archive/statistics/.

Lichtman, M. (2010), "The Fangraphs UZR Primer," http://www.fangraphs.com/blogs/the-fangraphs-uzr-primer/.

Miller, S. J. (2007), "A derivation of the Pythagorean Won-Loss Formula in baseball," *Chance*, 20, 40–48.

Slowinski, S. (2010), "What is WAR?" http://www.fangraphs.com/library/index.php/misc/war/.

Smith, D. (2013), "Retrosheet," http://www.retrosheet.org/.

Tango, T. (2008), "How to calculate WAR," http://www.insidethebook.com/ee/index.php/site/article/how_to_calculate_war/.

Tung, D. D. (2010), "Confidence Intervals for the Pythagorean Formula in Baseball," http://vixra.org/abs/1005.0020.

Wyers, C. (2013), "Reworking WARP," http://www.baseballprospectus.com/article.php?articleid=21586.