Adapting WinShares for Universal Player Evaluation in American Football

September 2019

Abstract

Uniform player evaluation in the sport of American football is severely lagging behind other professional sports like baseball and basketball. While these sports have advanced metrics like Wins Above Replacement and Win Shares to measure each player's contribution to his team in a fair way, there exists no equivalent metric in football. While there have been attempts in recent years to create a WAR statistic for the NFL similar to the one in baseball and basketball (Yurko et al., 2018), these mostly quantify the contribution of offensive players because the nature of the play-by-play data collected in the NFL makes it difficult to analyze offensive linemen and defensive players. Pro Football Focus provides a way to compare players by giving them grades between +2 and -2 on a per-play basis for all positions but the data is not publicly available. Currently there is no advanced metric that can objectively determine an offensive and a defensive player's contribution to his team. For this reason, we have created a new WinShares metric similar to the one used in baseball and basketball. We use the nflscrapR package and data from Pro-Football-Reference to calculate how many wins a player earned for his team over the course of a season (James et al., 2002). In this paper, we present the data and code needed to fully reproduce the WinShares statistic. We discuss the potential applications of this metric and also its limitations. Finally, we propose possible improvements and recommendations on how to adjust the calculation of WinShares if more relevant play-by-play data becomes available.

1 Introduction

The main idea behind win shares involves dividing a team's wins between its players based on contribution. The metric was invented in 2002 by statistician Bill James for use in the sport of baseball and has seen multiple iterations since then. In James' version of the statistic, each win share counts as 1/3 of a team win and players cannot receive negative win shares. Separate implementations of win shares have allowed negative values and have changed the method of calculation, but the main idea behind win shares remains the same: accurately depict how many wins a player earned for his team. Since its initial creation, the concept of win shares has expanded to hockey and basketball as a way to measure a player's value. In this paper, we will present a method for which win shares can be adapted to football. In our version of win shares, we decided the sum of all players' win shares for any given game should be equal to 1 for a win, 0.5 for a tie, and 0 for a loss. From there, win shares can be allocated to players accordingly.

This allocation is an open problem in the sport of football due to the unique and specific goals of each position. Since certain positions do not have much impact on the boxscore of a game, there are currently very few metrics that evaluate players at all positions. Pro Football Focus' grades are assigned to players at all positions based on their per-play performance (Pro Football Focus). Madden overall ratings are given out yearly and also show player performance. The problem with these metrics is that they are based on human judgement and their method of calculation is not public. Pro Football Reference's approximate value (AV) uses available box score data and introduces additional information like multipliers for all-pro voting and offensive/defensive performance relative to league average (Drinnen, 2013). It's method of calculation is public and it is assigned to all players at every position. For these reasons it was chosen as the starting point for calculating win shares.

2 Calculating WinShares

The next task was to convert a seasonal metric for player performance, AV, into a weekly metric that shows each player's contribution to his team's performance. We started by calculating a player's relative importance, which we defined as a player's AV multiplied by the percentage of snaps for his offense, defense, or special teams divided by the total relative importance of the remaining players in the offense, defense, or special teams.

For offense, defense, and special teams:

$$X = \frac{AV_{player}}{\sum_{players} AV_p} * \text{SnapCount}_{player} \tag{1}$$

$$\text{Relative Importance} = \frac{X}{\sum_{players} X_p} \tag{2}$$

For example, in week 2 of the 2018 NFL Season, Luke Kuechly, with a season-long AV of 14, played 100% of defensive snaps for the Carolina Panthers. The Total AV of the defense was 78. We divided Kuechly's AV by the total defensive AV and multiplied by his snap percentage to get 0.179. This number represents the player's relative value to the team by incorporating their performance over the course of the season relative to the rest of the team and their value on that specific game through their snap count. Kuechly's number (0.179) was summed with every other defensive player's number to get 0.737. Then the final relative importance was scaled back to 1 and Kuechly was given a relative importance of 0.244.

```
for k in ['Off', 'Def', 'ST']:
    #df is the roster with snap count percentage of given game
    temp = df[(df[f'{k}Pct'] != 0.0) & (~np.isnan(df[f'{k}Pct']))]
    #sum of all relative importances
    bottom = sum([(av * pct) / sum(temp['AV']) for av, pct in temp
    [['AV', f'{k}Pct']].values])
    #relative importance
    temp[f'{k}RelImp'] = [(av * pct) / (sum(temp['AV']) * bottom)
    for av, pct in temp[['AV', f'{k}Pct']].values]
```

This number was the 12th highest mark in 2018, showing how important Luke Kuechly was to the Panther's defense in week 2. This relative importance, however, says nothing about Kuechly's performance during week 2, it only shows that he was an integral part of the defense. To determine the performance of a player, we must introduce the next critical element to win shares: win probability.

In order to calculate individual performance, we must take into account team performance. We split team performance into offensive performance, defensive performance, and special teams performance and took a sum of a team's win probability added (WPA) for each of the 3 phases of the game. WPA shows the change in win probability on each play for the possession team. It provides context to football plays that other metrics, like yards gained, cannot. For example, a 5 yard gain on third-and-4 is better than a 5 yard gain on third-and-7 and will rightfully have a greater WPA. A team's win probability also depends on other factors like time remaining, score differential, and field position. There are various win probability models, and we chose the model included with the nflscrapR play-by-play package (Yurko et al., 2018). Then, for each game, we grouped together plays from each phase of the game and obtained the total WPA. This gave us a measure of team performance on offense, defense, and special teams. Due to the nature of the sport, however, most plays are positive WPA for the offense and negative WPA for the defense. In order to account for this, we adjusted the WPA for each phase to the league average WPA for that phase. This essentially created a WPA relative to league average for offense, defense, and special teams.

To get the player's WPA, we multiplied a player's relative importance with the corresponding WPA of the offense, defense, or special teams. The player's WPA was then summed across the 3 phases to create a Total WPA for each player. The player's win shares value was calculated by dividing each player's Total WPA with the Total WPA of the whole team and multiplying by the win value (1 for a win, 0.5 for a tie, 0 for a loss).

This calculation had two separate issues, however. First, for some teams in certain games, each player's win shares was inflated because the team's total WPA was too low. Secondly, good players playing on bad teams were being penalized. For example, in 2015, guard Marshal Yanda was a 1st-Team All-Pro selection and Pro Bowler for the 5-11 Baltimore Ravens, who finished with the 25th ranked scoring offense in the NFL. His AV was 10, and he amassed a seasonal relative importance of 3.49, one of the highest marks in the past 5 years for his contribution. However, even in wins the Ravens often had a negative WPA offense, so Yanda, despite being one of the best players on the offense, received much of the blame according to win shares. One of the reasons for this is that in the current calculation there is a negative and positive WPA for each play. By resetting the negative WPA's to 0, the team that effectively 'won' the down is the only side that gets credit, instead of also taking credit away from the other side of the field. With this solution, players can only earn win shares in victories and in games where their offense, defense, or special teams earned positive WPA. This adjustment moved Yanda to the top of the Raven's offense in win shares.

```
for wk in np.unique(df[df.Team == tm]['Week']):
    temp = df[(df.Team == tm) & (df.Week == wk) & ((~np.isnan(df.TeamOffWPA)) | (~np.isnan(df.TeamDefWPA)) | (~np.isnan(df.TeamSTWPA)))]
    for k in ['Off', 'Def', 'ST']:
        temp[f'{k}WPA'] = temp[[f'{k}RelImp']] * temp[[f'TeamAdj{k}WPA']].mode().values.item(0)
        temp[f'{k}WPA'] = temp[f'{k}WPA'].apply(lambda x: x if x > 0 else 0)
    temp['TotWPA'] = temp[['OffWPA', 'DefWPA', 'STWPA']].sum(axis = 1)
```

3 Results/Future Work

	CB	DE	DT	FB	K	LB	LS	OL	P	QB	RB	S	TE	WR
1	Josh Norman 2015 **WS: 1.6**	J.J. Watt 2014 **WS: 1.75**	Fletcher Cox 2018 **WS: 1.05**	Spencer Ware 2016 **WS: 0.39**	Stephen Gostkowski 2014 **WS: 0.19**	Thomas Davis 2015 **WS: 1.59**	Don Muhlbach 2018 **WS: 0.07**	Mitchell Schwartz 2018 **WS: 1.45**	Pat McAfee 2015 **WS: 0.11**	Tom Brady 2017 **WS: 1.74**	DeMarco Murray 2014 **WS: 1.09**	Landon Collins 2016 **WS: 1.16**	Rob Gronkowski 2017 **WS: 1.04**	Antonio Brown 2014 **WS: 1.16**
2	Kyle Fuller 2018 **WS: 1.38**	J.J. Watt 2015 **WS: 1.66**	Kawann Short 2015 **WS: 1.04**	Rod Smith 2017 **WS: 0.16**	Wil Lutz 2018 **WS: 0.14**	Chandler Jones 2017 **WS: 1.4**	Reid Ferguson 2017 **WS: 0.06**	Mitchell Schwartz 2017 **WS: 1.28**	Marquette King 2016 **WS: 0.1**	Patrick Mahomes 2018 **WS: 1.44**	Le'Veon Bell 2014 **WS: 1.02**	Harrison Smith 2018 **WS: 1.07**	Travis Kelce 2018 **WS: 0.93**	Brandin Cooks 2017 **WS: 1.13**
3	Chris Harris 2015 **WS: 1.31**	Danielle Hunter 2018 **WS: 1.29**	Ndamukong Suh 2014 **WS: 0.99**	Jamize Olawale 2016 **WS: 0.14**	Dustin Hopkins 2018 **WS: 0.14**	Von Miller 2015 **WS: 1.31**	Morgan Cox 2015 **WS: 0.05**	Andrew Whitworth 2015 **WS: 1.27**	Matthew Bosher 2014 **WS: 0.09**	Dak Prescott 2016 **WS: 1.4**	Ezekiel Elliott 2016 **WS: 0.99**	Kevin Byard 2017 **WS: 1.06**	Rob Gronkowski 2015 **WS: 0.68**	Tyreek Hill 2018 **WS: 1.09**
4	Richard Sherman 2014 **WS: 1.26**	Olivier Vernon 2016 **WS: 1.24**	Aaron Donald 2017 **WS: 0.98**	Mike Tolbert 2015 **WS: 0.12**	Stephen Hauschka 2017 **WS: 0.14**	D'Qwell Jackson 2014 **WS: 1.3**	Jake McQuaide 2017 **WS: 0.05**	Travis Frederick 2016 **WS: 1.23**	Bryan Anger 2015 **WS: 0.08**	Carson Palmer 2015 **WS: 1.29**	Le'Veon Bell 2017 **WS: 0.98**	Derwin James 2018 **WS: 1.03**	Jason Witten 2014 **WS: 0.66**	Michael Thomas 2018 **WS: 1.04**
5	Stephon Gilmore 2018 **WS: 1.11**	Cameron Jordan 2017 **WS: 1.19**	Damon Harrison 2016 **WS: 0.93**	Kyle Juszczyk 2017 **WS: 0.12**	Stephen Gostkowski 2015 **WS: 0.12**	Zach Orr 2016 **WS: 1.29**	Jon Dorenbos 2014 **WS: 0.04**	Zack Martin 2016 **WS: 1.23**	Pat McAfee 2016 **WS: 0.08**	Matt Ryan 2016 **WS: 1.24**	Alvin Kamara 2018 **WS: 0.82**	Mike Adams 2014 **WS: 1.02**	Jason Witten 2016 **WS: 0.66**	Antonio Brown 2017 **WS: 1.0**

Figure 1: Top 5 Season Performances by Position According to Total Season WinShares

In order to see the effectiveness of win shares beyond the top 5 at each position, we decided to test it against another metric. We chose Madden overall ratings and ran correlation tests for each season in the past 5 years and a test including all 5 seasons.

Year	All Overalls	80+ Overalls
2014	0.63	0.49
2015	0.61	0.50
2016	0.62	0.47
2017	0.6	0.46
2018	0.62	0.44
All 5	0.61	0.47

We found that, when grouped together over 5 seasons, win shares show a strong correlation with Madden overalls. While Madden overalls are subjective and far from perfect, they do generally show how a player performed in a given season and his value compared to other players in the league. Win shares, however, is most useful when comparing players on the same team since a player's win shares relies heavily on the number of games a team won. For example, a player with more win shares may not be a better player, but simply an equal or lesser player on a better team. In order to accurately compare players' contributions from different teams on a game-by-game basis, win shares would need to be adjusted for the total number of wins for each player.

4 Conclusion

In this paper, we have discussed a possible method for calculating win shares in the NFL (from current available data). The main idea behind the calculation was to multiply the WPA of each phase (offense, defense, or special teams) with each player's relative importance (their AV multiplied by their snap count by phase) and scale back down to the win value (1 for a win, 0.5 for a tie, and 0 for a loss). The win shares statistic was shown to be well-correlated when tested against Madden Overalls. Further analysis could explore win shares above replacement or win shares per dollar to see how players compare to others without positional bias or to determine the most cost-effective players in the NFL.

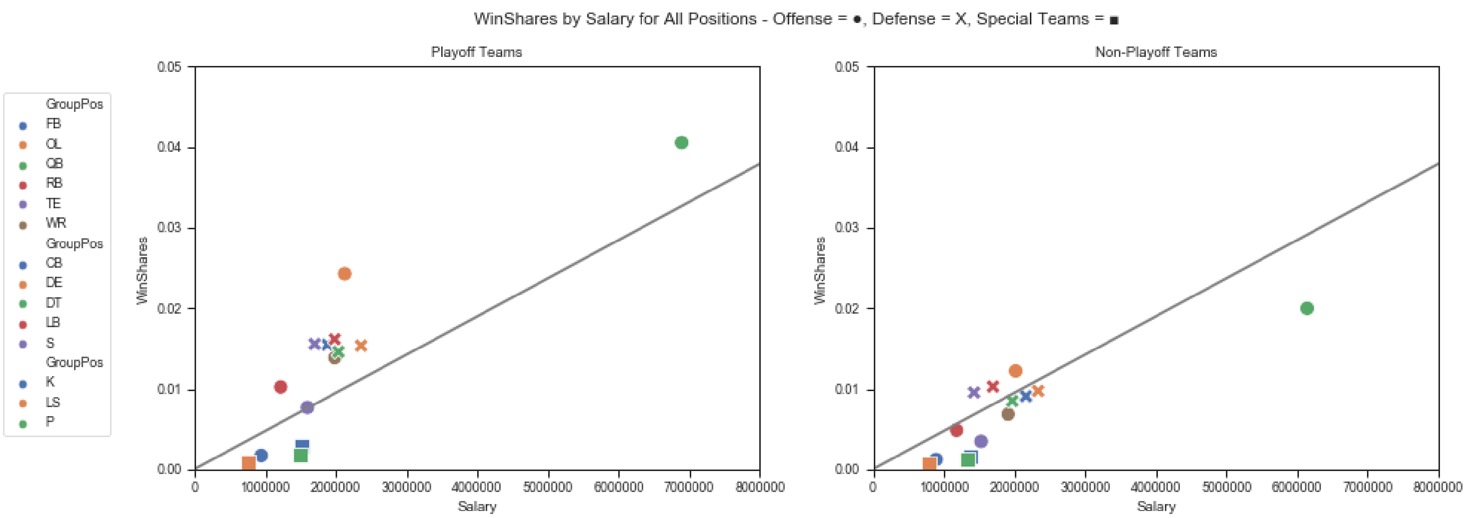

Figure 2: WinShares per game by position and salary as an example of possible analyses using WinShares

Another thing to note about win shares is that, since it is based upon snap count, it may not properly account for special teams players. Kickers do not get any more win shares for making or missing kicks, and any change in their win shares relies on the calculation of AV. A solution could potentially introduce weights to certain special teams positions like long-snappers, punters, or kickers. This could be expanded to have weights for different positions based on play type. For example, targeted receivers on pass plays, ball carriers on running plays, and tacklers or players who force turnovers on defense could be given more importance on a per-play basis. Another solution would include replacing AV in the relative importance calculation. While AV is currently the best possible metric for calculating win shares, it struggles in some respects. It uses arbitrary constants in its calculation, which can be tweaked to produce different results. More accurate win shares would be possible once data starts being collected on which players are on the field for each play. Relative importance could be calculated on a per-play basis and would represent players who only come in on certain downs or in certain substitution packages. In conclusion, while the current calculation of WinShares serves as a promising starting point, further research is needed to provide an accurate and useful metric.

References

[1] James, B and Henzler, Jim (2002): *Win Shares*

[2] NBA Win Shares, URL
`https://www.basketball-reference.com/about/ws.html`

[3] Calculating Point Shares, URL https://www.hockey-reference.com/about/point_shares.html

[4] Pro Football Focus, URL https://www.pff.com/

[5] Drinen, D. (2013): "Approximate value: Methodology," URL https://www.sportsreference.com/blog/approximate-value-methodology/

[6] Yurko, R., Ventura, S., and Horowitz, M (2018): "nflWAR: A Reproducible Method for Offensive Player Evaluation in Football," *New England Symposium on Statistics in Sports*.