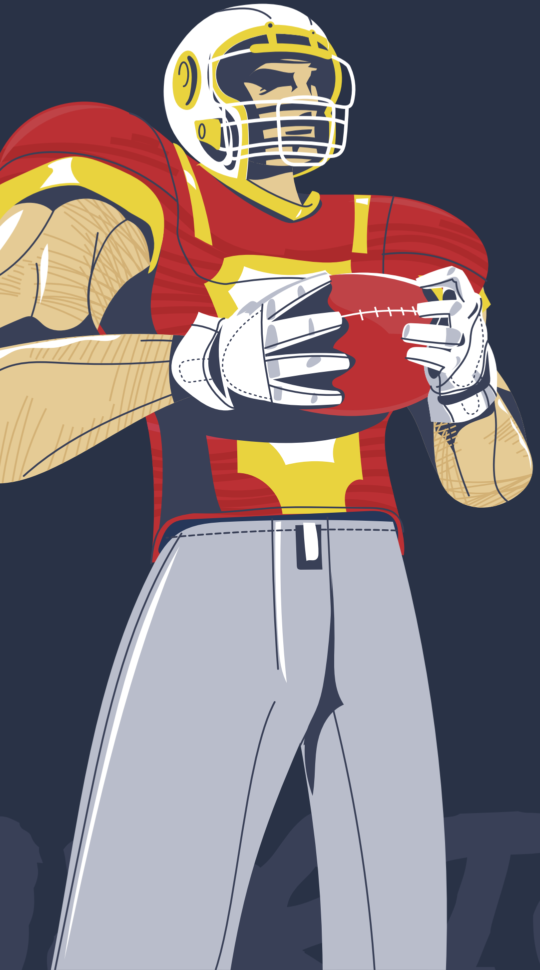


nflWAR: A Reproducible Method for Offensive Player Evaluation in Football

Sophy Huang, Nancy Li, Siyao Li

Motivation

Reproducibility of evaluation methods

Novel WAR measure on football games

Limited statistical metrics for football

Evaluation framework for NFL plays and players

01
WAR (Wins Above Replacement)

WAR (Wins Above Replacement)

Definition: WAR is a statistic used in baseball to summarize a player's total contributions to their team. It represents the number of additional wins a player would contribute compared to a replacement-level player at the same position.

Significance

- Aggregates a player's offensive, defensive, and baserunning contributions into a single value.
- Normalizes performance across different leagues and eras, allowing for meaningful comparisons between players from different seasons or leagues.

Applications

- **Major League Baseball (MLB):** Utilized for player evaluation and historical comparison.
- **Fantasy Baseball:** Helps in making informed decisions for drafting and trading players based on their overall contribution rather than just isolated statistics.

WAR (Wins Above Replacement)

Total Contribution

$$\text{WAR} = \frac{\text{Player Runs Above Average + Replacement Level Runs}}{\text{Runs Per Win}}$$

Player Runs Above Average (RAA): Measures how many runs the player contributed, above or below an average player, adjusted for ballpark and era.

Replacement Level Runs: The number of runs a replacement-level player is expected to contribute, which is typically set below the league average.

Runs Per Win (RPW): The number of additional runs that contribute to an extra win in the standings, varying slightly each season based on scoring environment.

Calculating WAR

Scenario: We have a player named Joe. We will calculate Joe's WAR based on his performance over a season compared to a replacement-level player.

Step 1 Calculate Player Runs Above Average (RAA)

- **Batting Runs:** Joe scored 20 runs above average with his batting.
- **Fielding Runs:** Joe saved 10 runs above average with his fielding.
- **Baserunning Runs:** Joe contributed 5 extra runs through base running.
- **Positional Adjustment:** Since Joe is a catcher, a position that's typically harder, we add 5 runs.
- **League Adjustment:** Joe plays in a league where the average run value is slightly higher, adding 3 runs.

Total RAA = 20 (batting) + 10 (fielding) + 5 (base running) + 5 (positional) + 3 (league) = 43 runs

Step 2 Determine Replacement Level Runs

- The typical replacement player would perform 20 runs below an average player.

Total Contribution = 43 RAA + 20 Replacement Level = 63 runs

Step 3 Calculate Runs Per Win (RPW)

- In Joe's league, approximately 10 runs translate to an additional win.

WAR = Total Contribution / RPW = 63 / 10 = 6.3

02

nflWAR (NFL Wins Above Replacement)

Calculating WAR for NFL

Step 1 **Estimate the value of each play**

- Expected Points
- Win Probability
- EPA & WPA

Step 2 **Estimate the effect of each player on play value added**

- Multilevel Modeling Framework
- Passing & Rushing Models

Step 3 **Evaluate relative to replacement level**

- Separate replacement level for positions and actions

Step 4 **Convert to a wins scale**

- Translate the estimated points and win probability to Wins Above Replacement (WAR).

Variable	Description
Possession Team	Team with the ball on offense (opposing team is on defense)
Down	Four downs to advance the ball ten (or more) yards
Yards to go	Distance in yards to advance and convert first down
Yard line	Distance in yards away from opponent's endzone (100 to zero)
Time Remaining	Seconds remaining in game, each game is 3600 seconds long (four quarters, halftime, and a potential overtime)
Score differential	Difference in score between the possession team and opposition

Step 1: Expected Points

Covariates:

Variable	Variable description
Down	The current down (1st, 2nd, 3rd, or 4th
Seconds	Number of seconds remaining in half
Yardline	Yards from endzone (0 to 100)
log(YTG)	Log transformation of yards to go for a first down
GTG	Indicator for whether or not it is a goal down situation
UTM	Indicator for whether or not time remaining in the half is under two minutes

Response Variable:

$$Y \in \{\text{Touchdown } (\mathbf{7}),\ \text{Field Goal } (\mathbf{3}),\ \text{Safety } (\mathbf{2}),\ \text{No Score } (\mathbf{0}),\ -\text{Touchdown } (\mathbf{-7}),\ -\text{Field Goal } (\mathbf{-3}),\ -\text{Safety } (\mathbf{-2})\}$$

Model:

A multinomial logistic regression that predicts the probability of each type of scoring event occurring given the current situation on the field.

$$\log\left(\frac{P(Y = Touchdown|\mathbf{X})}{P(Y = No\ Score|\mathbf{X})}\right) = \mathbf{X} \cdot \boldsymbol{\beta}_{Touchdown},$$
$$\log\left(\frac{P(Y = Field\ Goal|\mathbf{X})}{P(Y = No\ Score|\mathbf{X})}\right) = \mathbf{X} \cdot \boldsymbol{\beta}_{Field\ Goal},$$
$$\vdots$$
$$\log\left(\frac{P(Y = -Touchdown|\mathbf{X})}{P(Y = No\ Score|\mathbf{X})}\right) = \mathbf{X} \cdot \boldsymbol{\beta}_{-Touchdown},$$

$$EP = E[Y|\mathbf{X}] = \sum_y y \cdot P(Y = y|\mathbf{X}).$$

Step 1: Win Probability

Covariates:

Variable	Variable description
$E[S]$	Expected score differential = $EP + S$
s_g	Number of seconds remaining in game
$E[\frac{S}{s_g+1}]$	Expected score time ratio
h	Current half of the game (1st, 2nd, or overtime)
s_h	Number of seconds remaining in half
u	Indicator for whether or not time remaining in half is under two minutes
t_{off}	Time outs remaining for offensive (possession) team
t_{def}	Time outs remaining for defensive team

Model

$$\log\left(\frac{P(\mathrm{Win})}{P(\mathrm{Loss})}\right) = s(E[S]) + s(s_h)\cdot h + s\left(E\left[\frac{S}{s_g+1}\right]\right) + h\cdot u\cdot t_{off} + h\cdot u\cdot t_{def},$$

s is a smooth function while h, u, t_{off} , and t_{def} are linear parametric terms.
By taking the inverse of the logit we arrive at a play’s W P.

Step 1: EPA & WPA

In order to arrive at a comprehensive measure of player performance, each play in a football game must be assigned an appropriate value $\delta_{f,\ i}$ that can be represented as the change from state i to state f.

$$\delta_{f,i} = \boldsymbol{V}_f - \boldsymbol{V}_i,$$

Expected Points Added(EPA):

EPA = the expected points at the start of a play (**EPi**) - the expected points at the start of the subsequent play (**EPf**)

Win Probability Added (WPA):

WPA = win probability at the start of a play (**WPi**) - the win probability at the start of the following play (**WPf**).

Step 2: Multilevel Modeling Framework

Multilevel Model:

Embraces positional group structure and accounts for the observation dependence in NFL.

Example: Modeling δf,i with varying-intercepts for quarterbacks and receivers.

- Q_q a eceivers, respectively.

$$\delta_{f,i} \sim Normal(Q_{q[i]} + C_{c[i]} + X_i \cdot \beta,\ \sigma_\delta^2),\ for\ i = 1,\ldots,n \text{ plays},$$

- X_i are e, with β as their coefficients.

$$Q_q \sim Normal(\mu_Q,\ \sigma_Q^2), \text{ for } q = 1,\ldots,\#\text{ of QBs},$$
$$C_c \sim Normal(\mu_C,\ \sigma_C^2), \text{ for } c = 1,\ldots,\#\text{ of receivers}.$$

Step 2: Passing & Rushing Models

Purpose: To dissect and quantify the contributions of individual players to passing and rushing plays within the NFL.

Passing Models

1. Air Yards (AIR) Model
2. Yards After Catch (YAC) Model

Rushing Model

1. Unified Rushing Models
   - Combines metrics for initial contact yards and post-contact yards into a single model.

Step 3: Comparing to Replacement Level

- **Defining by Position**:

  Capture the unique responsibilities and skill sets of each position.

  Example: For quarterbacks, assess metrics like completion rate, passing yards, and touchdowns. For linemen, focus on blocking success rates and penalties.

- **Defining by Receiving & Rushing**:

  Recognize the distinct contributions within the roles of versatile positions like running backs and wide receivers.

  Examples:

  Rushing: For running backs, benchmark rushing yards, touchdowns, and yards after contact against replacement-level performance.

Step 4: Conversion to Wins

EPA-based

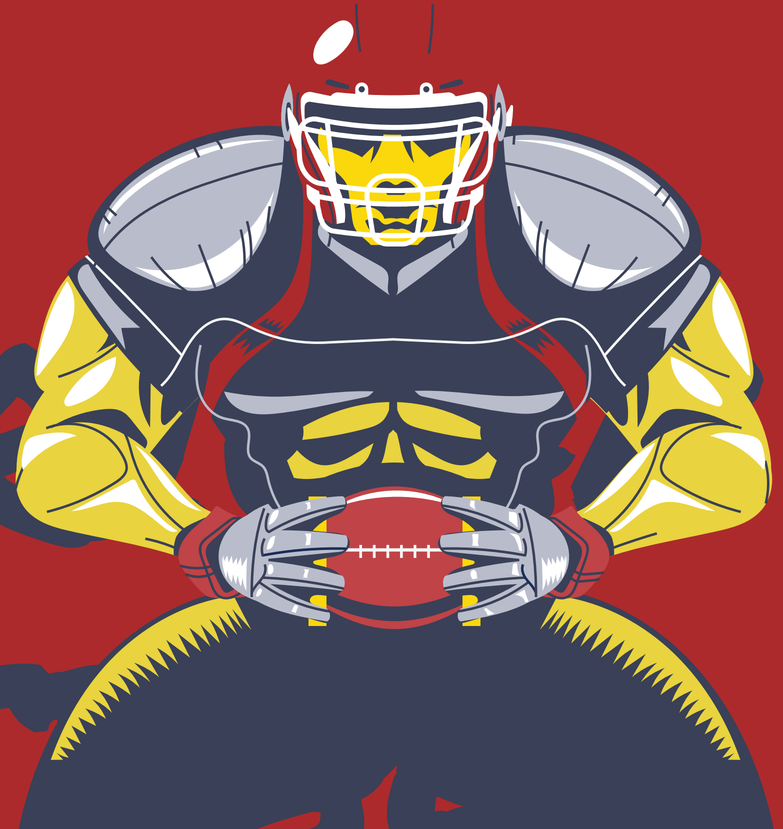
- Linear regression model to establish a relationship between a team's regular season win total and their score differential(*S*).

$$\text{Wins}_t = \beta_0 + \beta_S S_t + \epsilon_t, \quad \text{where } \epsilon_t \sim N(0, \sigma^2) \text{ (iid)}$$

- Points add up to a win: 1/ β*s*

WPA-based

- The final estimation values are an individual’s win probability added above replacement, which is equivalent to their wins above replacement (WAR).



03

Analysis on NFL Data

Data Analysis

Data

- ***nflscrapR***: an R package for analyzing data from the National Football League (NFL) API.
- **Data**: 2017 NFL season play-by-play data, scraped using *nflscrapR*
- **Size**: ~45K plays, 103 variables
- **Example variables**: date, play_id, AirYards, YardsAfterCatch, Passer, Rusher, Receiver, etc.
- **Players**: 71 QBs, 148 RBs, 201 WRs, and 109 TEs

Analysis

- **Four separate models for four player positions**: quarterback (QB), running back (RB), wide receiver (WR), tight end (TE)
- **Two approaches to evaluate plays**:
  - Expected points added (EPA)
  - Win probability added (WPA)
- **Three separate WAR components to evaluate players**:
  - total_WAR = air_WAR + yac_WAR + rush_WAR

Results: WAR, QB Example

EPA-Based WAR, QB

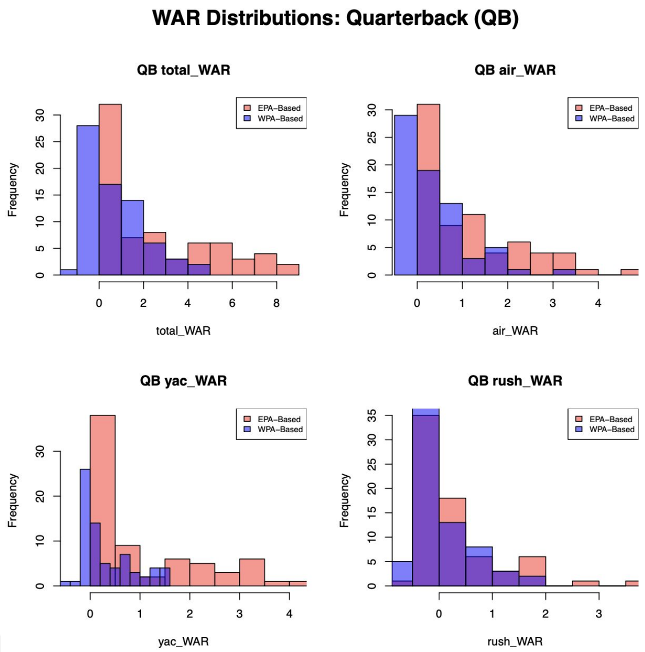
Player_ID_Name	A.Dalton-00-0027973	A.McCarron-00-0031288
Pass_Attempts	474	14
Rush_Attempts	20	0
Sacks	39	1
Position	QB	QB
Total_Plays	533	15
Replacement_Level	0	1
Player_Model_ID	A.Dalton-00-0027973	Replacement_QB
air_iPA	-0.0409008	-0.1489969
yac_iPA	0.02280526	-0.13068000
rush_iPA	-0.2975750	-0.2436734
air_iPAA	-19.386979	-2.085957
yac_iPAA	10.80969	-1.82952
rush_iPAA	-17.5569275	-0.2436734
air_iPAR	51.23755	0.00000
yac_iPAR	72.75201	0.00000
rush_iPAR	-3.180195	0.000000
total_iPAR	120.8094	0.0000
air_WAR	1.504021	0.000000
yac_WAR	2.135553	0.000000
rush_WAR	-0.09335105	0.00000000
total_WAR	3.546223	0.000000

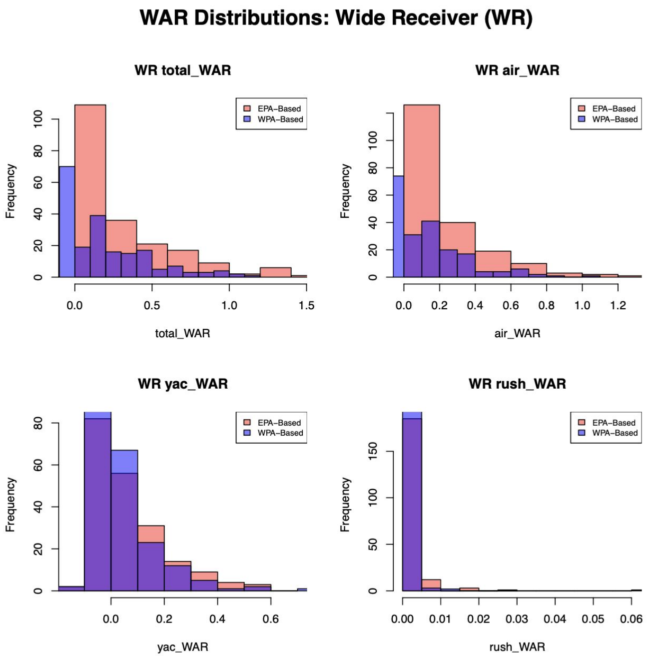
WPA-Based WAR, QB

Player_ID_Name	A.Dalton-00-0027973	A.McCarron-00-0031288
Pass_Attempts	474	14
Rush_Attempts	20	0
Sacks	39	1
Position	QB	QB
Total_Plays	533	15
Replacement_Level	0	1
Player_Model_ID	A.Dalton-00-0027973	Replacement_QB
air_iPA	-0.0007491534	-0.0018589553
yac_iPA	0.0006482179	-0.0013491094
rush_iPA	-8.888879e-03	-6.695802e-06
air_iPAA	-0.35509869	-0.02602537
yac_iPAA	0.30725527	-0.01888753
rush_iPAA	-5.244439e-01	-6.695802e-06
air_iPAR	0.5260461	0.0000000
yac_iPAR	0.9467331	0.0000000
rush_iPAR	-0.5240488	0.0000000
total_iPAR	0.9487304	0.0000000
air_WAR	0.5260461	0.0000000
yac_WAR	0.9467331	0.0000000
rush_WAR	-0.5240488	0.0000000
total_WAR	0.9487304	0.0000000

Results: WAR Distributions, QB & WR

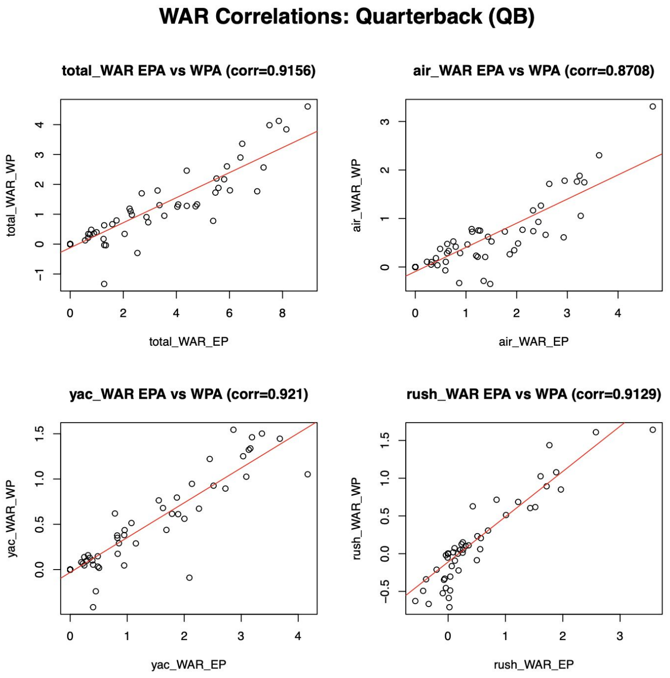
Comparing the WAR distributions, EPA-based WAR values tend to be higher than WPA-based values for all the positions. QB players have higher WAR values compared to other positions.

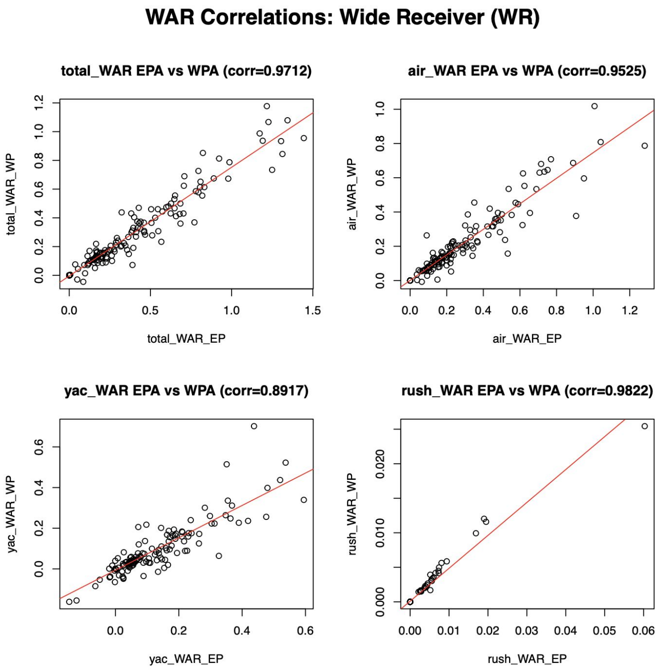




Results: WAR Correlations, QB & WR

The correlations between EPA-based and WPA-based WAR values are high for all positions, and all WAR components, showing consistency in the two approaches.

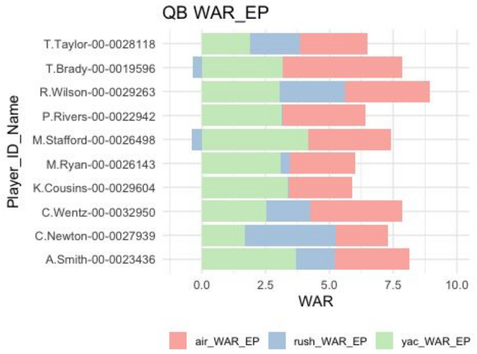


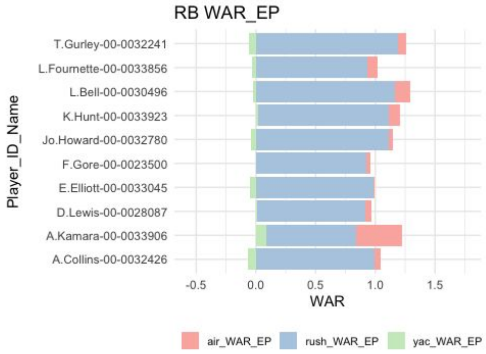


Results: Top 10 Players By Total WAR

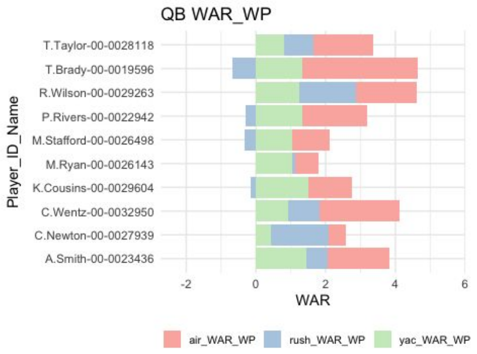
For QBs, WARs are mainly from passing (i.e., air yards and yards after catch). For RBs, WARs are mainly from rushing.

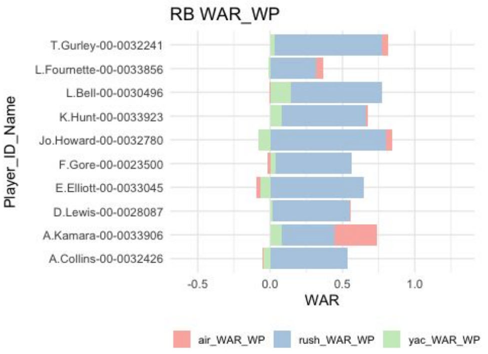
Top 10 Players: EPA-Based Total WAR





Top 10 Players: WPA-Based Total WAR





Summary

Contribution

- **Data and software development**: provide an R package *nflscrapR* to enable easy access to NFL play-by-play data
- **Play evaluation**: novel approaches to estimate expected points (EP) and win probabilities (WP)
  - EP: multinomial logistic regression
  - WP: generalized additive model
- **Player evaluation**: WAR in terms of air yards, yards after catch, rushing

Limitation & Extension

- **No player participation data from NFL**:
  - No information about which players are on the field for each play
  - Analysis is limited to players directly involved in the play and contextual information
- **WAR for players at all positions**:
  - Extend the framework to estimate WAR for players of any position

Thank You!

Why nflWAR?

Gap in Football Analytics: Unlike baseball, NFL lacks a unified, publicly available metric that can evaluate player contributions comprehensively.

Innovation in Data: The development of nflWAR utilizes publicly available NFL play-by-play data to estimate the wins contributed by individual players.

Importance for Teams: Provides NFL teams and analysts with a tool to assess player impact beyond traditional statistics, enabling better team management and player evaluation decisions.

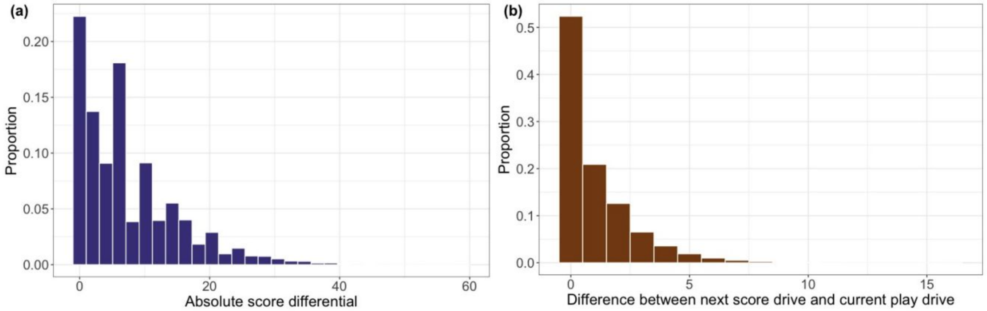
Variable	Description
Possession Team	Team with the ball on offense (opposing team is on defense)
Down	Four downs to advance the ball ten (or more) yards
Yards to go	Distance in yards to advance and convert first down
Yard line	Distance in yards away from opponent's endzone (100 to zero)
Time Remaining	Seconds remaining in game, each game is 3600 seconds long (four quarters, halftime, and a potential overtime)
Score differential	Difference in score between the possession team and opposition

NFL play by play data

Observation Weighting

Applied on **Expected points (EP) model** to ensure that certain plays don't **disproportionately influence** the model's predictions due to their unique circumstances.

Score Differential Weighting **Drive Distance to Next Score Weighting**



$$w_i = w(S_i) = \frac{\max_i(|S_i|) - |S_i|}{\max_i(|S_i|) - \min_i(|S_i|)}.$$

Model for NFL

Characteristics

- **Group Hierarchies**: NFL players belong to different positional groups which dictate their roles and the frequency of their involvement in plays.

- **Repeated Measures**: Players repeatedly involved in plays create a scenario where observations (plays) are not independent of each other.

Solution

- **Multilevel Model**: Embraces this positional group structure and accounts for the observation dependence.

- The model includes varying-intercepts for player groups involved in a play. This approach allows the intercepts to vary by player group rather than being fixed across all groups.

Model Formulation

$$\Delta_{air} \sim Normal(Q_{air,q[i]} + C_{air,c[i]} + F_{air,v[i]} + \mathbf{A}_i \cdot \boldsymbol{\alpha},\ \sigma_{\Delta_{air}}) \text{ for } i = 1, \ldots, n \text{ plays},$$
$$Q_{air,q} \sim Normal(\mu_{Q_{air}}, \sigma^2_{Q_{air}}), \text{ for } q = 1, \ldots, \text{\# of QBs},$$
$$C_{air,c} \sim Normal(\mu_{C_{air}}, \sigma^2_{C_{air}}), \text{ for } c = 1, \ldots, \text{\# of receivers},$$
$$F_{air,v} \sim Normal(\mu_{F_{air}}, \sigma^2_{F_{air}}), \text{ for } v = 1, \ldots, \text{\# of defenses},$$

$$\Delta_{yac} \sim Normal(Q_{yac,q[i]} + C_{yac,c[i]} + F_{yac,v[i]} + \mathbf{B}_i \cdot \boldsymbol{\beta},\ \sigma_{\Delta_{yac}}) \text{ for } i = 1, \ldots, n \text{ plays},$$
$$Q_{yac,q} \sim Normal(\mu_{Q_{yac}},\ \sigma^2_{Q_{yac}}), \text{ for } q = 1, \ldots, \text{\# of QBs},$$
$$C_{yac,c} \sim Normal(\mu_{C_{yac}},\ \sigma^2_{C_{yac}}), \text{ for } c = 1, \ldots, \text{\# of receivers},$$
$$F_{yac,v} \sim Normal(\mu_{F_{yac}},\ \sigma^2_{F_{yac}}), \text{ for } v = 1, \ldots, \text{\# of defenses},$$

Covariate vector Bi contains the same set of indicator variables in Ai but also includes the AirYards and interaction terms between AirYards and the various RecPosition indicators.

Why Passing & Rushing Models?

Tailored Evaluation

- **Position-Specific Metrics**: By using separate models, the analytics can capture the distinct skills required for different positions, such as a quarterback's passing accuracy or a running back's ability to gain yards after contact. This leads to more accurate and meaningful evaluations of player performance.

Clarity and Precision in Attribution

- **Clear Attribution of Contributions**: Separate models allow teams to clearly understand how much each player contributes to the success of specific types of plays. For example, in passing models, contributions can be separated into how well the quarterback threw the ball (air yards) and how effectively the receiver managed to gain additional yards (yards after catch).

Passing Models

The model separates the valuation of passing plays into two distinct components:

1. **Air Yards** ($\delta_{f,i,air}$) :

   The yards a ball travels through the air before being caught.

$$\Delta_{air} = \delta_{f,i,air} \cdot \mathbf{1}(\text{completion}) + \delta_{f,i} \cdot \mathbf{1}(\text{incompletion}),$$

2. **Yards After Catch ($\delta_{f,i,yac}$):**

   The yards gained by a receiver after catching the ball

$$\Delta_{yac} = \delta_{f,i,yac} \cdot \mathbf{1}(\text{completion}) + \delta_{f,i} \cdot \mathbf{1}(\text{incompletion}),$$

Passing Models

1. The passing model for **Δair**:

$$
\begin{aligned}
\Delta_{air} &\sim Normal(Q_{air,q[i]} + C_{air,c[i]} + F_{air,v[i]} + \boldsymbol{A}_i \cdot \boldsymbol{\alpha},\ \sigma_{\Delta_{air}}) \text{ for } i = 1,\ldots, n \text{ plays},\\
Q_{air,q} &\sim Normal(\mu_{Q_{air}}, \sigma^2_{Q_{air}}), \text{ for } q = 1,\ldots,\text{\# of QBs},\\
C_{air,c} &\sim Normal(\mu_{C_{air}}, \sigma^2_{C_{air}}), \text{ for } c = 1,\ldots,\text{\# of receivers},\\
F_{air,v} &\sim Normal(\mu_{F_{air}}, \sigma^2_{F_{air}}), \text{ for } v = 1,\ldots,\text{\# of defenses},
\end{aligned}
$$

Covariate vector Ai contains a set of indicator variables for Home, Shotgun, NoHuddle, QBHit, Location, RecPosition, as well as the RushStrength value

2. The passing model for **Δyac**:

$$
\begin{aligned}
\Delta_{yac} &\sim Normal(Q_{yac,q[i]} + C_{yac,c[i]} + F_{yac,v[i]} + \boldsymbol{B}_i \cdot \boldsymbol{\beta},\ \sigma_{\Delta_{yac}}) \text{ for } i = 1,\ldots, n \text{ plays},\\
Q_{yac,q} &\sim Normal(\mu_{Q_{yac}},\ \sigma^2_{Q_{yac}}), \text{ for } q = 1,\ldots,\text{\# of QBs},\\
C_{yac,c} &\sim Normal(\mu_{C_{yac}},\ \sigma^2_{C_{yac}}), \text{ for } c = 1,\ldots,\text{\# of receivers},\\
F_{yac,v} &\sim Normal(\mu_{F_{yac}},\ \sigma^2_{F_{yac}}), \text{ for } v = 1,\ldots,\text{\# of defenses},
\end{aligned}
$$

Covariate vector Bi contains the same set of indicator variables in Ai but also includes the AirYards and interaction terms between AirYards and the various RecPosition indicators.

Results: WPA-Based WAR, QB Example

QB WPA-Based WAR Table

Player_ID_Name	A.Dalton-00-0027973	A.McCarron-00-0031288
Pass_Attempts	474	14
Rush_Attempts	20	0
Sacks	39	1
Position	QB	QB
Total_Plays	533	15
Replacement_Level	0	1
Player_Model_ID	A.Dalton-00-0027973	Replacement_QB
air_iPA	-0.0007491534	-0.0018589553
yac_iPA	0.0006482179	-0.0013491094
rush_iPA	-8.888879e-03	-6.695802e-06
air_iPAA	-0.35509869	-0.02602537
yac_iPAA	0.30725527	-0.01888753
rush_iPAA	-5.244439e-01	-6.695802e-06
air_iPAR	0.5260461	0.0000000
yac_iPAR	0.9467331	0.0000000
rush_iPAR	-0.5240488	0.0000000
total_iPAR	0.9487304	0.0000000
air_WAR	0.5260461	0.0000000
yac_WAR	0.9467331	0.0000000
rush_WAR	-0.5240488	0.0000000
total_WAR	0.9487304	0.0000000

QB WPA-Based WAR – Passing Model (Air)

```
Linear mixed model fit by REML ['lmerMod']
Formula: airWPA_Result ~ Home_Ind + Shotgun_Ind + No_Huddle_Ind + QBHit +
    Receiver_Position + PassLocation + Rush_EPA_Att + (1 | Passer_ID_Name) +
    (1 | Receiver_ID_Name) + (1 | DefensiveTeam)
REML criterion at convergence: -57586.93
Random effects:
 Groups           Name        Std.Dev.
 Receiver_ID_Name (Intercept) 0.002816
 Passer_ID_Name   (Intercept) 0.002339
 DefensiveTeam    (Intercept) 0.000000
 Residual                     0.044128
Number of obs: 16980, groups:
Receiver_ID_Name, 291; Passer_ID_Name, 47; DefensiveTeam, 32
Fixed Effects:
       (Intercept)             Home_Ind            Shotgun_Ind           No_Huddle_Ind
        -1.014e-03           -4.768e-04             -2.898e-03               3.364e-03
             QBHit  Receiver_PositionFB    Receiver_PositionRB    Receiver_PositionTE
        -5.562e-03           -1.746e-02             -1.680e-02              -3.643e-03
 PassLocationmiddle    PassLocationright           Rush_EPA_Att
         5.609e-03            4.035e-05              8.770e-04
```