

A Primer on the Signal WARP Ratings

Alex Wang

December 2025

1 Introduction

This primer documents the process behind the creation of the Signal WARP Ratings. It is not meant to be a full explanation of the algorithm.

The Signal WARP Ratings are a **PREDICTIVE** power rating for all 265 **Division I** football teams, including **FBS** and **FCS** teams. The lack of reliable play-by-play data for Division II and Division III football teams precludes me from fully applying the algorithm to them. A similar "score-based" variation of my algorithm exists and is used to fill in when data is insufficient, but sacrifices accuracy in prediction.

In short, the system contains two components: **points scored** through raw margin and **deserved points scored** through statistical analysis of play-by-play data. Opponent adjustments are iteratively factored in for a power rating of each team.

The accuracy of this rating comes not through its structure, but its **recency weighting** and unique **deserved margin** analysis.

2 What is Considered

The Signal WARP ratings consider both surface-level and play-level pieces of data.

2.1 Scores of Games

The scores of games are utilized in a manner similar to **Massey's Method**, in which linear least squares are used to create a simple rating for each team. However, linear least squares sacrifices granularity regarding the true skill level for a certain team.

In 2025, **Memphis** defeated **Arkansas** 32-31. The scores of games would indicate that Memphis outplayed Arkansas, but an analysis of the fumbles, individual plays, and yards per run or pass would indicate that **Arkansas** was the better team on the field.

2.2 Plays in Games

To alleviate the aforementioned issue, every **FIELD GOAL ATTEMPT**, **PUNT**, and **REGULAR PLAY** is considered. These three categories of plays are considered separately, as they encode different signals regarding the strength of a certain team. For example, a field goal kick on **4th and 1** from 50 yards is an exceptionally strong kick, but furthermore indicates an inability for the offense to gain a first down in short yardage situations. Punts encode similar signals.

These plays are translated into a **deserved margin**, indicating a score of the game if it had been played without luck factors. For example, California defeated SMU 38-35 in 2025. However, the deserved margin, derived from individual plays in the game, demonstrated that SMU was the better team, by 5 points.

These two actual and deserved margins are weighted separately based on their consistency with each other, and using a Gauss-Seidel iterative solving, creates power ratings for each team.

2.3 Locations of Games

It is well-known that **home-field advantage** plays a role in the points scored within a football game. This number has often been quoted as **2.5-3.5** points.

A unique strength of the Signal WARP Ratings is that home-field advantage is calculated per-season, based on the errors between past predictions without this advantage and the results of the games.

For the 2025 season, as of 2 January 2026, home-field advantage is **2.80 points**.

This per-season home-field advantage calculation also allows us to note decreases in attendance or trends in fan behavior. In 2021, home-field advantage was **1.99 points**, indicating a decrease in attendance after the COVID-19 pandemic.

2.4 Recency of Games

Exponential smoothing is used extensively, both in home-field advantage calculations and weighting recent team performance. By using exponential smoothing, later observations can be weighted higher than early-season observations, in order to create a rating which most accurately reflect the **current ability** of a team.

3 What is Not Considered

3.1 Recruiting Data and Returning Talent Projections

Although recruiting data is a powerful tool that can be used to estimate a team's power rating before games are played, it is derived from human judgment or proprietary metrics, particularly from sources such as On3, 247Sports, or ESPN. All calculations for power ratings are tuned purely by my own optimization. As such, I do not include player talent projections or recruiting data.

However, performance from the previous two seasons is considered.

3.2 Garbage-Time Points and Plays

Similarly to ESPN's Bill Connelly and his SP+ ratings, all plays during garbage time, defined as when margin exceeds a certain amount by a certain quarter within a game, are discarded.

Often, an insurmountable point deficit or surplus indicates that both sides will pull their starting players, and playing style shifts drastically from what is expected of a team.

Furthermore, margin of victory is given a slight smoothing factor to decrease the influence of blow-out wins or losses.

4 Ratings and the Prediction Method

4.1 Solving the System

Recall that for each game for each team we have a **true point margin** P and a **deserved point margin** \hat{P} .

The **total margin** for team i in a game between i and j is

$$M_{i,j} = \frac{a \cdot P + b \cdot \hat{P}}{a + b}$$

which is, in effect, a weighted average of these two. A team's rating r_i is modelled by

$$r_i = \sum w_{i,j} (r_j + M_{i,j})$$

where j consists of all opponents for team i . This is essentially **Massey's Method**; but $w_{i,j}$ is a recency weight and the inputs are not solely point margins. The point of this is to create opponent adjustments.

To prevent oscillation when iteratively solving, we ensure that $\bar{r} = 0$. An "average team" should have a rating of 0.

4.2 Ratings

Consider the following three team ratings.

Navy	16.89
Army	13.64
Cal Poly	-6.49
Home Field Advantage =	02.80

Ratings are presented in the form of a **point margin**. A team's rating r_i means that team i will be expected to beat a hypothetical average team by r_i points. A team's rating being negative indicates that they are below-average.

Navy would be expected to beat an average Division I team by 16.89 points.

4.3 Point Spreads

Given HFA , the home-field advantage, and r_i and r_j , the projected margin of i over j is

$$r_i - r_j \pm HFA$$

depending on the site of the game. HFA is omitted if the game is neutral-site.

In the example provided above, if the Army-Navy Game were to be played at Northwest Stadium, a neutral site, the line should be set at **Navy -3.25**.

5 Limitations and Improvements

5.1 Limitations

Firstly, this rating system only functions for the 265 NCAA Division I teams. Play-by-play data is only available on a regular basis for such teams, and this cannot extend to Division II, III, and the NAIA.

Secondly, although minimal, games between Division I teams and non-Division I opponents must be omitted for the above reason.

This is in contrast to systems like Bill Connelly's **SP+** or Jeff Sagarin's **Predictor**, as they use less granular statistics at the game-level such as Success Rate and Point Margin.

5.2 Future Improvements

Certain systems attempt to create a data-driven analysis of team-specific home-field advantages. However, a naive single-season analysis of this is not conducive to accuracy. I intend

to analyze long-term trends of team attendance to create more accurate HFA ratings.

I am also not considering more advanced factors such as **injury**, **proximity of neutral-site games to a certain team** (see the Independence Bowl in Shreveport, Louisiana, which played host to the Louisiana Tech Bulldogs), and per-game predictions separate from the subtractable power rating. These factors are being considered and their effects analyzed.

6 Accuracy of the Ratings, 2025

This section engages primarily with the **second-half** accuracy of game predictions (games from weeks 8 to 15). Second-half accuracy is used because first-half season accuracy primarily relies on recruiting projections, which is data that I do not use.

RMSE	14.926
MAE	12.007
Bias	00.418
Accuracy	00.713

In comparison to ESPN's Football Power Index:

RMSE	15.278
MAE	12.230
Bias	-0.585
Accuracy	00.744

In comparison to ESPN BET's opening line:

RMSE	14.858
MAE	12.059
Bias	-0.043
Accuracy	00.717

My model's performance, based on subtractable power ratings and a unifying home-field advantage, is comparable to the performance of the betting industry and ESPN's proprietary algorithm.

7 Acknowledgments

I would like to thank **Wes Colley** and the Colley Matrix Method for inspiring me to get into sports ratings.

I would like to thank **Bill Connelly** for getting me interested in advanced, predictive statistics.

I would like to thank **Jeff Sagarin** for inspiration with the design of my website.