

John Incantalupo

MTH-320 Statistical Methods for Data Science

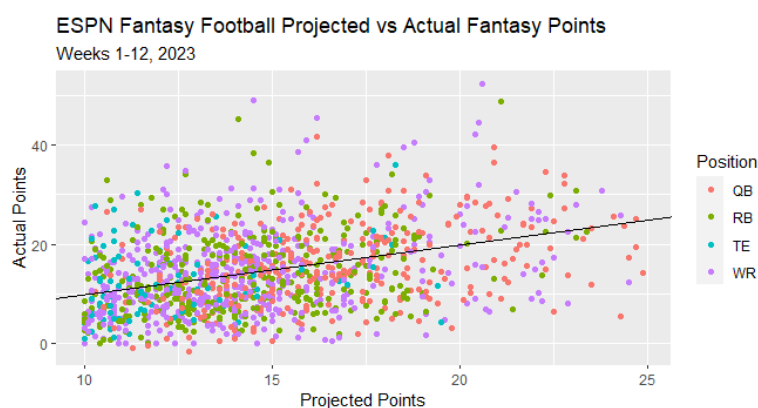
December 14, 2023

### How Accurate Are ESPN's Fantasy Football Projections?

Over the past decade, ESPN has provided projections for their fantasy football game to assist users in making the right moves for their team. Naturally, given the unpredictable nature of football, I have always wondered what goes into determining the projections as well as how accurate they actually are. Although I was unable to find the factors that contribute to these projections, I decided to test their accuracy and see if there was a way to improve upon them. To do this, I have gathered data from the first 12 weeks of the 2023 NFL season on every player projected to score at least 10 fantasy points in ESPN's standard points-per-reception format. Important attributes that I included are the player's name, position, and opponent for that week, as well as their projected points, the actual number of points that they scored, and if the player sustained an injury during the game and was ruled out. One of the main statistics that I will be measuring during this project will be the difference between projected points and actual points, which I will be referring to as the projection error. Ideally, we want this metric to be as close to zero as possible, but there are many outside factors that will affect it. The main obstacle is injuries, which are unfortunately inevitable in such a physical sport. ESPN creates their projections with the assumption that players will remain healthy throughout the entire game, which can lead to some large negative projection errors when injuries do occur. That is why I plan to filter out the instances of players who left a game early due to injury in order to decrease the number of these outliers in the data. I am also interested to see if there are any significant

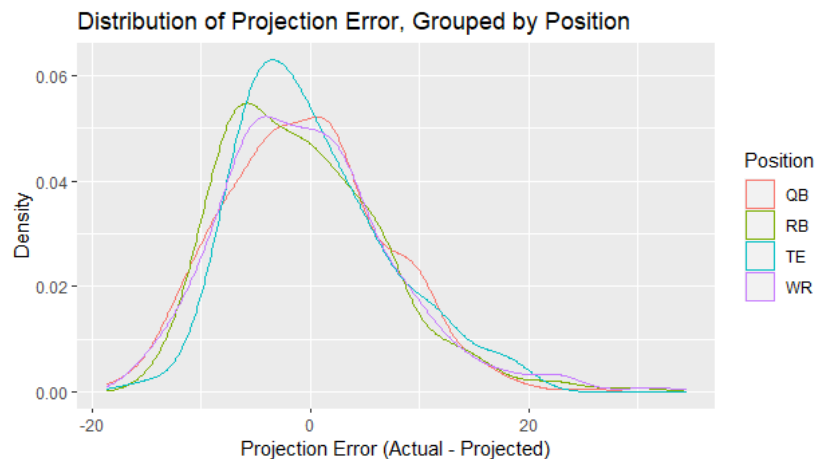
week-to-week shifts in the data. Perhaps there were some weeks that were much more high-scoring than others or had a larger variance of projection errors.

Upon implementing the dataset into R, I began the explanatory data analysis by plotting projected points against actual points and found a weak correlation between the two variables with a coefficient of about 0.3480. Even when partitioning the data into the four different positions of quarterback, running back, wide receiver, and tight end, the correlation does not improve much, topping out at around 0.3620 for quarterbacks. Below is a scatterplot of the data, with the black line representing where the projection error is zero. Dots above the line represent players who exceeded their fantasy projections, while the dots below the line represent players who underperformed their projections.



Then, I decided to look at the distribution of the projection errors themselves. Overall, this distribution is skewed to the right, as the occasional 40-point and 50-point performances led to some large projection errors that contributed to the skew. Once again, I grouped this data by position and found that the quarterbacks and tight ends had less of a skew than the running backs and wide receivers. The latter two positions typically have the best upside in fantasy football, as the top seven performances from the dataset were all from running backs and wide receivers. Quarterbacks, on average, have higher expectations in terms of fantasy points, which limits their

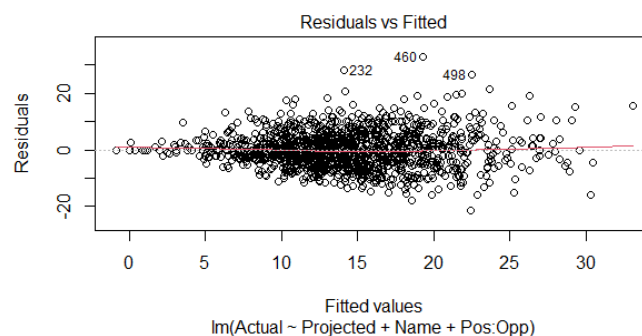
projection errors when they have big performances. Tight ends typically do not have big performances, and this is reflected in their low expectations. Overall, this position had the lowest variance of projection errors.



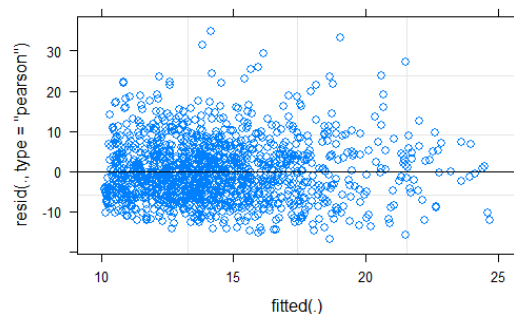
Just looking at projected and actual fantasy points does not tell the whole story, especially considering the weak correlation between the two. There are other contributing factors that can potentially explain the large variations in projection error, such as the player's team, their opponent, and their position. While the graph above does not show any significant differences in projection error between different positions, I figured that the week number and the different opposing defenses are two important factors that can help us to explain the output. I individually tested all three variables to see if there were any significant differences between the different groups' average projection error. Since we have previously established that our data is non-normal, I decided to use the Kruskal-Wallis test rather than a one-way ANOVA. When performing this test, I found that both the week number and the opponent had significant p-values at 0.008159 and 0.03426, respectively. On the other hand, the test on the different positions did not provide any evidence of differing means of projection errors, given that the p-value was 0.5728. The week number and the opponent are factors that will definitely assist our

model, while the position probably will not be as useful. Still, I plan to incorporate the position attribute into the model anyway to see if it can give us slightly better predictions.

To start, I tested a simple linear regression model using projected points, position, and opponent. I found that this model's predictions were heavily reflected by the projected number of points, as each prediction was within five points of the projected points. The correlation between this model's predictions and the actual fantasy output slightly increased to 0.3933018, and the RMSE and MAE were about 7.45 and 5.90, respectively. When tweaking this model, I decided to include a few interaction terms between the week number and the opponent, as well as the opponent and the position, in order to improve our RMSE and MAE. My reasoning was that opposing defenses can improve or fall apart as the season goes along, and these defenses can also be better at defending certain positions, so these interaction terms made sense to include in this model at the time. In doing so, we were able to significantly improve our RMSE to 6.15. However, when this model was used on Week 13's projections, which will be our testing set for this project, we got a much worse RMSE of 10.29, which is a sign that this model was an overfit of the original dataset. I tried a few more different combinations of linear models, and the best one I was able to find was similar to the previous model discussed except without the week-position interaction term. The fitted values of this model are plotted below against their residuals. Even then, this model was still an overfit, as the RMSE for the training set from the first 12 weeks was 6.32, while the RMSE for Week 13 was 9.38.



After that, I tested a few different mixed models to see if I could get a better fit. I played around with random slopes and intercepts, with projected points being the fixed effect and the player's name being the random effect. I ended up with four different mixed models, and they all surprisingly produced near identical RMSEs and MAEs for both the training and testing data. When compared to the linear models, the RMSE and MAE of the training set were slightly worse, while the RMSE and MAE of the testing set were a slight improvement. Then, I wanted to see the values that each model predicted, and so I compiled each model's predictions into a single table to compare them to the actual output. This was when I noticed that while the linear models have a large spread of predicted points that greatly exceeds ESPN's range of projected points, the mixed models play it safer with their predictions. As a final attempt at finding a better model, I added the week variable to one of the mixed models, and yet the same issues that plagued the linear models once again appeared. While the training set's RMSE improved, the testing set's RMSE suffered as a result. Once again, I have plotted the fitted values of the best mixed model against their residuals below.



Unfortunately, I was not able to find a suitable model to predict future fantasy production. Overall, there is probably a lot more data that goes into these projections, such as a player's previous stats, their current target share, their offense's scheme, and much more. Also, this dataset is incomplete given that I had to limit it to position players who were projected at least 10

points. Even if we do have a complete dataset at our disposal, including the unquantifiable metrics, any predictive model can still fail due to the randomness of fantasy football. This has especially been true this NFL season, as there have been so many surprises just within the first 14 weeks. However, I still found analyzing the data to be very interesting. Although I did not mention this when I discussed my explanatory data analysis of the dataset, I looked at the mean and variance of projection errors when grouped by different variables such as week number, opponent, and even by each player. Below is a grid of density curves for each opposing team's projection errors. Through this process, I was able to learn a lot more about this dataset, such as how the Cleveland Browns were the best team in terms of preventing fantasy points and how Week 5 has the highest variance of projection errors from the first 12 weeks of the season. The player stats were also interesting. Keenan Allen of the Los Angeles Chargers overperformed his projections the most out of all players in this dataset with at least a 10-game sample, while Patrick Mahomes of the Kansas City Chiefs surprisingly underperformed his projections the most out of this same sample of players. Despite the recent movement to integrate advanced analytics into the world of football, the results of this project prove that there will always be random effects that cannot be accounted for. This volatility can be best described in one of football's most iconic phrases: "Any given Sunday."

