# Modeling Fantasy Football Quarterbacks

Kyle Zeberlein
*Augustana College, Rock Island Illinois*

Myles Wallin
*Augustana College, Rock Island Illinois*

# MODELING FANTASY FOOTBALL QUARTERBACKS

Myles Wallin and Kyle Zeberlein

AUGUSTANA COLLEGE

# *Introduction*

At the end of August 2015, in our Mathematical Modeling class, we were posed with a project to analyze and model something math related. Through our shared interest in season long fantasy football, we decided it would be an interesting to try to create a game-by-game projection for each starting quarterback, and model as close to Yahoo's projection system as possible.

We started collecting career data on each starting quarterback. Based on our previous experience, we divided the quarterback into four different categories: True rookies, quarterbacks with less than 16 games started in the NFL, quarterbacks with 16-47 games started, and quarterbacks with 48 or more games started. Each quarterback category had a different formula for their projection. Each category has certain parameters that make up each formula. All the categories share one common parameter, Opponent points per game (PPG) allowed in 2014, weighted differently in each formula. This parameter is the average number of fantasy points a defense allowed to the quarterbacks they played against in the 2014 NFL season. This is important because the opponent's defensive strength usually factors strongly on how well the quarterback will perform that game.

This parameter is common between all the categories because every type of quarterback has to play a different defense each week. For our first two categories (rookie and less than 16 games), the parameter weight was 60% of the formula because these quarterbacks are less experienced in the NFL, which causes the defense to have more of an impact on the quarterback's performance as they have yet to settle themselves into the league as a predictable quarterback such as Aaron Rodgers. Our other two categories (Greater than 16 games and our general formula), had the parameter weighted at just 50% of the formula. This is due to the quarterbacks playing in the league for at least one full year, which allows us to have more historical data on them, causing them to be more predictable.

## *Original Projections*

For our first three categories listed below, our coefficients of the parameters added up to one. We did this because these weightings best mirrored a reasonable prediction based on historical stats. This allowed us to specifically identify which parameter should have more of an impact on each different formula, depending on the experience the quarterback has in the NFL. This is how to calculate each parameter:

- Opponent PPG allowed in previous year - the average number of fantasy points a defense allowed to the quarterbacks they played against in the previous NFL season
- own team's QB stats 2014 - all the quarterbacks that played for team A divided by the number of games Team A plays in a season
- Total Career Points/# of games - this calculates the average points that specific quarterback has produced per game throughout his career. This does not include the games where the quarterback comes in and just takes a knee, or plays the last couple minutes of the game and produces close to no points in that game. This is calculated by figuring out the total points the quarterback has his career divided by the total number of games that quarterback has started in his career.
- Average PPG in career - take the total number of points in their career and divide by the number of games that quarterback has played in for their career.
- Average PPG in the last season they played - take the total number of points scored in the previous season and divide it by the # of games they played in the previous season.

- Average PPG against opponent - take the total number of points the quarterback has scored against the team they are about to play and divide it by the number of games they have played that opponent.
- Average PPG rest of career - take the remaining points scored in their career and divide it by the remaining amount of games they have played.

Our initial formula for the true rookie category is:

$$0.6(Opponent\ PPG\ allowed\ 2014) + 0.4(own\ team's\ QB\ stats\ 2014)$$

For our rookie quarterbacks, we used only two parameters as shown above. The reason behind choosing the 'own team's QB stats for 2014' is due to the rookies having no NFL experience. We decided that, for the most part, each team's offensive players do not change drastically. This allows the teams to plug a new, inexperienced, quarterback into the already established system that they are running. As a rookie, you do not have the expectations like other quarterbacks in the league do. Although the franchise that the rookie plays for does expect them to have the same, if not better, production that they had from their quarterbacks in the previous year.

Our initial formula for the quarterbacks with less than 16 games started is:

$$0.6(Opponent\ PPG\ allowed\ 2014) + 0.3(own\ team's\ QB\ stats\ 2014)$$
$$+ 0.1(Total\ Career\ Points\ /\ \#\ games\ started)$$

For this category, the three parameters shown above are used. We decided to split up the remaining 40% of formula into two parameters. We used the same parameter as we did for rookies with a coefficient of .3 instead of .4 because the quarterbacks under this category have had some NFL experience. With the remaining 10% of the formula, we decided to use 'Total Career Points/ # games started'. We used this parameter because the quarterbacks do have some experience, even though it is minimal. This category mainly has quarterbacks who are second

year QB's who did not start every game their rookie season or has veteran quarterbacks who have been backups their entire life and are finally getting the opportunity to start. The 10% of the last parameter allows the minimum experience of the quarterback to have a slight impact on the formula. This slight weighting will not over project a quarterback, if the quarterback has been dominant in their first 15 or less games. Also, it will not under project them if they have been doing horrible as they are adjusting to the system their offense is running.

Our initial formula for the quarterbacks with 16-47 games started is:

$$0.5(Opponent\ PPG\ allowed\ 2014)\ +\ 0.5(avg.\ PPG\ in\ career)$$

For this category, we decided to only use two parameters. The second parameter is the 'avg. PPG in career'. We used this parameter because these quarterbacks have played an entire season in the NFL and have a good amount of experience. We have enough data from these quarterbacks to be able to use their general PPG for their career. We do not have enough data to be able to specifically have a parameter for the quarterback playing that specific team like we do in our general formula because, in order for a quarterback to play every team, they have to have played in 3 to 4 seasons in the NFL. They do have enough of an impact on how the offense is run now to have a major impact on how well the team's offense will do, which correlates to how well the quarterback plays in almost every game.

Our initial formula for the rest of the quarterbacks is:

$$([part\ 1] + [part\ 2] + [part\ 3] + [part\ 4])/2$$

- Part 1: (# games 2014) / (# games career) $*$ (avg. PPG 2014)
- Part 2: (# games vs opponent) / (# games career) $*$ (avg. PPG vs opponent)
- Part 3: (# games leftover) / (# games career) $*$ (avg. PPG rest of Career)
- Part 4: Opponent PPG allowed 2014

For our general formula, our coefficients differ compared to the previous three formulas. In this formula, we have four parameters, listed above, that each have a weight of 50%. The coefficients add up to two instead of one like the previous three ones. We had the coefficients add up to two because we weighted part four as one and the parts one through three have separate coefficients that add up to one. This is due to the number of games in the previous season plus the number of games versus the opponent plus the remaining number of games played all add up to the total number of games played in the quarterbacks career. The reason we have more parameters in this formula is because each one of these quarterbacks have more historical data than the rest of the categories. The quarterbacks in our general formula have started enough games to have played in the NFL for three years. We used the same common parameter, which is part 4 in this case, as we did for the first three. The other three are more detailed than what even the third formula has. With these experience quarterbacks, we had enough data to create Parts 1-3 above.

As shown above, the first three parts are each proportional to that category. For example, the first part is the number of games they played in the previous year divided by the total number of games which is then multiplied by what their points per game average was in the previous year. Then that whole part 1 is weighted 0.5. Part two and three are very similar but with different statistics. We chose these three parameters because each quarterback has enough historic data to show statistics versus a specific opponent. Parts 1 - 3 contain information that has to do with all the games the quarterback has ever played. For example, part 1 has is based on the last season they played, part 2 is based off of how well they have played against the opponent they are facing that week in their career, and part 3 is everything else after parts 1 and 2 have been calculated.

## *Data Analysis*

The fantasy football point breakdown is as followed:

- .04 points per yard passing
- 4 points per touchdown thrown
- -2 points per interception/fumble
- .1 points per yard rushing
- 6 points per touchdown rushed

We use these values because they are considered standard fantasy football scoring Yahoo leagues, with the exception of the interception category [1]. Yahoo's standard has -1 points for each interception thrown, but we felt that it is more common to play with -2 points for each interception thrown. However, both cases are accepted in the world of fantasy football.

We calculated a projection for each game of the season, and then summed these projections to get a season total for each individual quarterback.

In order to organize the statistics for our formulas, we had to extract each quarterback's career statistics with respect to each category that makes up their projection. We used player profiles from NFL.com [2] and profootballreference.com [3] to collect the information. We found profootallreference.com particularly useful because it broke down a quarterback's statistics against each team throughout their career, which was helpful for our general formula. We found most of the other statistics from NFL.com. The one category that was not specifically documented in the quarterback's career against each team were fumbles lost. For this, we had to go through the game log for each season the quarterback was in the league and manually count how many fumbles were lost playing against a certain team which was found from NFL.com.

To come up with the defensive projections, we looked at each team's defensive statistics against quarterbacks for the previous season. For the passing yards allowed, touchdowns pass,

and interceptions categories, we were able to take directly off of NFL.com team statistics site. For the rush yards allowed, rushing touchdowns allowed, and Quarterback fumbles forced, we had to go through NFL.com game statistics site, and add up the quarterback's contribution in those three categories against each particular team, for the entire season. Then we repeated the process for all 32 teams.

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Games | Passing yds | Passing Tds | INT | Rushing Yds | Rushing TD | Fumbles | avg ppg | |
| 2 | 6 | 1626 | 11 | 3 | 25 | 0 | 1 | 17.25666667 | |
| 3 | **144** | **35365** | **224** | **155** | **400** | **7** | **17** | 14.22638889 | |
| 4 | | Week 1 | Week 2 | Week 3 | Week 4 | Week 5 | Week 6 | Week 7 | Week 8 |
| 5 | 1 | 0.7190277778 | 0.7190277778 | 0.7190277778 | 0.7190277778 | 0.7190277778 | 0.7190277778 | 0.7190277778 | 0.7190277778 |
| 6 | 2 | 0.5604166667 | 0.54875 | 0.4918055556 | 0.4077777778 | 0.5277777778 | 1.266527778 | 1.298611111 | 1.425972222 |
| 7 | 3 | 14.22638889 | 14.22638889 | 14.22638889 | 14.22638889 | 14.22638889 | 14.22638889 | 14.22638889 | 14.22638889 |
| 8 | 4 | 18.41066667 | 20.27733333 | 15.26533333 | 15.02266667 | 14.26933333 | 18.37066667 | 16.49066667 | 15.63466667 |
| 9 | **Projection** | **16.95825** | **17.88575** | **15.35127778** | **15.18793056** | **14.87126389** | **17.29130556** | **16.36734722** | **16.00302778** |
| 10 | | | | | | | | | |
| 11 | Yahoo Projections | | | | | | | | |
| 12 | Actual week 1 | | | | | | | | |
| 13 | | Games | Passing yds | Passing Tds | INT | Rushing Yds | Rushing TD | Fumbles | avg ppg |
| 14 | Saints | 5 | 1330 | 9 | 5 | 15 | 0 | 0 | 16.14 |
| 15 | | **144** | **35365** | **224** | **155** | **400** | **7** | **17** | 14.22638889 |
| 16 | Bears | 4 | 888 | 12 | 2 | -5 | 0 | 0 | 19.755 |
| 17 | | **144** | **35365** | **224** | **155** | **400** | **7** | **17** | 14.22638889 |
| 18 | 49ers | 4 | 1268 | 7 | 4 | 1 | 0 | 0 | 17.705 |
| 19 | | **144** | **35365** | **224** | **155** | **400** | **7** | **17** | 14.22638889 |
| 20 | Rams | 5 | 1378 | 4 | 5 | -4 | 0 | 1 | 11.744 |
| 21 | | **144** | **35365** | **224** | **155** | **400** | **7** | **17** | 14.22638889 |
| 22 | Lions | 5 | 1270 | 9 | 5 | 12 | 0 | 1 | 15.2 |
| 23 | | **144** | **35365** | **224** | **155** | **400** | **7** | **17** | 14.22638889 |
| 24 | Steelers | 14 | 3032 | 21 | 14 | 11 | 1 | 1 | 13.02714286 |
| 25 | | **144** | **35365** | **224** | **155** | **400** | **7** | **17** | 14.22638889 |
| 26 | Ravens | 14 | 3570 | 17 | 12 | 62 | 0 | 3 | 13.35714286 |
| 27 | | **144** | **35365** | **224** | **155** | **400** | **7** | **17** | 14.22638889 |

Fig 1. Sample of the data collection and projection

We organized our data with spreadsheets, and linked formulas to the data to come up with the game-by-game projections. We linked our formulas within the specific spreadsheet and linked different spreadsheets together. Each category that influenced a quarterback's projection formulas had a line in the spreadsheet. We created skeleton templates for each quarterback category so that we could easily add the data to the new sheet when adding a new quarterback's data. The only tedious part of this process was connecting the defensive fantasy points per game

allowed with the week the quarterback was playing against that defense. The data collection portion of the project was slow and inefficient, but the depth of it allowed for us to do a great amount of data analysis and testing of our formulas.

During the season we were able to enter the actual scores that each quarterback produced. This task was not hard as it was just a week by week basis entry. This simple task allowed us to be able to look at how close our predictions were to the actual results as the season went on. Some weeks we were way over, others we were under and sometimes even spot on. In the end we were able to see why the weekly projection is not as accurate as the end of the year projection in which each week's projection is summed. This occurs because we could predict a quarterback to have 40 points in two games and let us say we projected them to have 18 points the first game and 22 the second. In reality, the quarterback did score the 40 points in two games but in the first game he scored a 2 and the second he scored 38. Our ending result was spot on but our week by week was off by huge margin. This is one of the reason while we waited until after the season to analyze our statistics.

## *Results*

After the season was over, we looked into interpreting our results and improving them for the upcoming year. One of the first things we did was compare our results to Yahoo results along with comparing ours and Yahoo's to the actual results. When looking over our results, we noticed that some of our quarterbacks' predictions were off significantly from their actual results. These were quarterbacks who did not participate in the entire year. With this information, we decided to create a grouping of quarterbacks who did not start 10 regular season games. We will refer to this set as the outlier grouping. These quarterbacks include the following: Ryan Mallett, Andrew Luck, Peyton Manning, Tony Romo, and Colin Kaepernick.

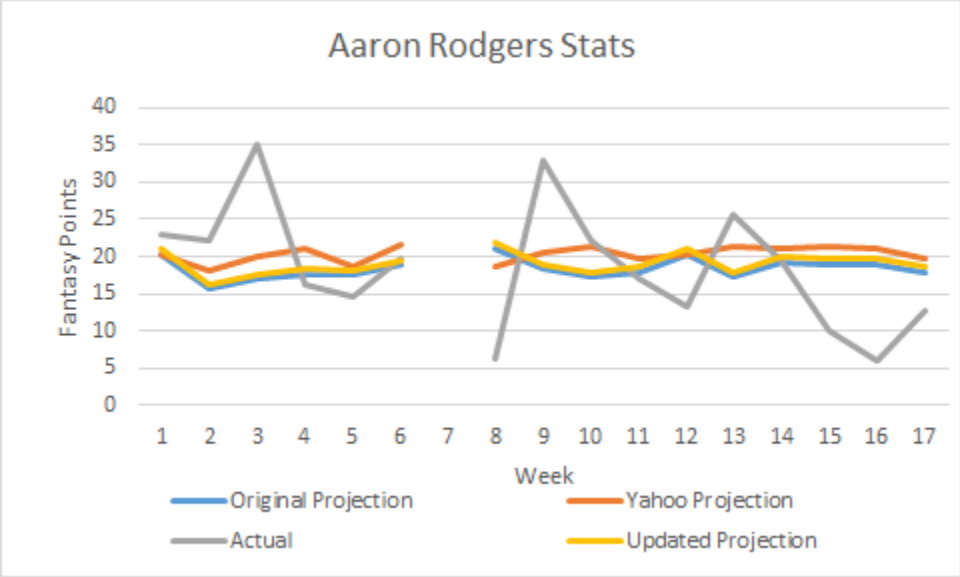


Fig 2. Plot of fantasy points by week of Aaron Rodgers.

Quarterbacks that did not fall in the outlier category were much easier to predict for the season as a whole. Although there are games where the actual output may have not fit the trend of the projections, the differences usually would average out as a whole. This is a good example of what was mentioned above, about the projection of the entire season being better than the week by week projection.

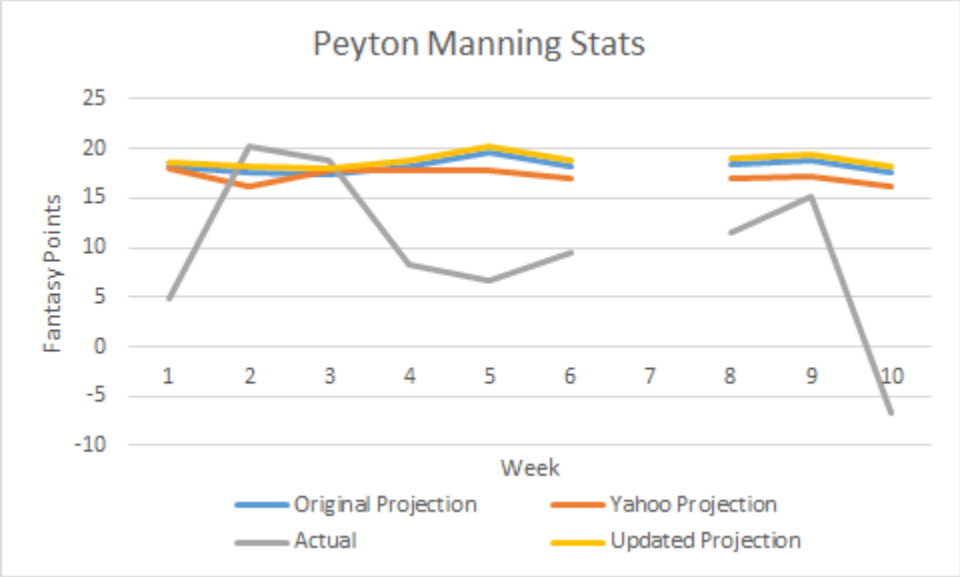


Fig 3. Plot of fantasy points by week of Peyton Manning. This plot only includes Games where Peyton Manning started.

It is much more difficult to project quarterbacks that fell into the outlier category. Injuries often affect performance, which is directly visible in the statistics. As seen above, Peyton Manning only outperformed his projections in two games.

We then looked at the differences for our quarterback predictions vs Yahoo vs Actual. When comparing the results of (our predictions - actual results) vs (Yahoo - actual), we found something very intriguing in the results. When we kept the outliers in, we over projected the actual results for the entire season by 1,063.30 points. This means we were 2.076 points off per game per quarterback. With outliers, Yahoo! Was 489.67 points off from the actual results (0.956 points off per game per quarterback). When we took the outliers out, we did better than Yahoo!. Without outliers, we were 18.299 points over the actual results (0.0357 points off per game per quarterback). Without outliers, Yahoo! Was 351.49 points under the actual results (0.6865 points off per game per quarterback).

We also analyzed our data using the square root of the sum of differences squared. We use this because it will take away the errors above or below that would cancel when just summing the differences. This analyzes the difference between the data and the actual results. While including the outlier quarterbacks, we projected a difference of 595.35 fantasy points for all of the quarterbacks, which is 17.48 points per quarterback. Yahoo! projected a difference of 535.06 points, which is 16.72 points per quarterback. Although Yahoo!'s projection is better, when you factor out the outliers, we ended having a smaller projection difference. We had a difference of 296.32 points, which is 10.98 points per quarterback. Yahoo!'s difference was 347.32 points, which is 12.86 points per quarterback. These point differences are the differences throughout the entire year. When the outlier quarterbacks are taken out of consideration, our projection is still shown to be significantly better.

After we found out that we had successfully completed our first goal of having created a more accurate projection system than Yahoo for the 2015 NFL season, we decided to create a new goal. This was to optimize our formulas in order to get as close to the actual results as possible. We knew that our weights for each category were already good, but we knew they could be better. We decided that with the data we had collected, we could perform a linear regression to optimize our projections. This linear regression allowed us find an error between the quarterback's actual fantasy football statistics and our projections for them.

Once we saw the error, we sought to optimize our formulas so that they would match their performance. Our goal for this is to adjust the weighting we have set to each parameter, so that we can improve our projections. Our thinking is that if we can tweak our formulas, they will be able to better project quarterbacks in the upcoming season. Doing this also allowed us to see if any of the categories we used to come up with projections ended up being unneeded. We decided

to hold constant the parameters and how the quarterbacks were grouped by formula, and only adjust the weighting of each formula.

## *Updated Projections*

In order to analyze the different formulas, we summed each quarterback's actual fantasy points for the season, summed each quarterback's projection for the season, and found the difference between the two. We then adjusted the weight by hand of each parameter until we could minimize the difference. If a quarterback did not play a certain week for whatever reason, we set their projection for that week to zero. This way, we only make our new projection based off of actual games played.

Starting with the rookie category, which was originally weighed 40% their own team's fantasy football quarterback statistics in 2014 and 60% the opposing defense's weight, transformed into:

$$0.8476(Opponent\ PPG\ allowed\ 2014)\ +\ 0.1524(own\ team's\ QB\ stats\ 2014)$$

With our original formula, we over projected by 24.0972 points. With our optimized formula, we now under project by 0.001625 points. In other words, our projection formula predicted slightly less points than the sum of all the fantasy points the quarterbacks in this category scored throughout the 2015 NFL season.

This shows that on average per game, per rookie quarterback, our new projection was .857 points closer to the actual result. We calculated this number by taking the original formula projection minus the optimized formula, then divide that answer by the average amount of games the rookies played and finally divide the previous by the number of rookie quarterbacks.

Although this projection is more accurate for this season, this could end up being changed as we get more data for quarterbacks in the true rookie category. For the 2015 season,

there were only two quarterbacks for this category, Marcus Mariota and Jameis Winston. Lack of substantial data could be at fault for the projection system inaccuracy moving forward.

Moving on to the less than 16 games started category, originally, the projection system valued the opposing defense's fantasy points per game allowed at 60%, their own team's quarterback fantasy production in 2014 at 30%, and that quarterback's previous fantasy points per game started at 10%. Our updated projection system changed to: somehow you've got to get this to fit on the page:

$$0.6(Opponent\ PPG\ allowed\ 2014) - 0.1813(own\ team's\ QB\ stats\ 2014) + 0.5813(Total\ Career\ Point\ /games\ started)$$

This new formula is particularly interesting because we found that the parameter which was previously weighed at 30% now contributes negatively to the projection. This would lead us to believe that if the last starter did poorly, the projections would be then higher for this quarterback, and vice versa. Although this train of thought could have some merit, this thinking would not align with the philosophy with the projection system that we created.

Another interesting piece of this new formula is how that quarterback's career production is weighed nearly as much as the opposing defense's weight. It is apparent that the quarterbacks that fall into this category are more closely related to the quarterbacks who have started 16-47 games than to the rookies.

With four quarterbacks falling into this category, the new formula will most likely be more trustworthy for future predictions than the rookie category. With the updated projection formula, we projected 0.0095 fantasy points below the actual output of all the quarterbacks who qualified for this category. The original projections under predicted by 99.43 fantasy points. This

averages out to 0.40 fantasy points per quarterback per game. Our new projections come out to be 0.000038 per quarterback per game.

The next category is quarterbacks with 16-47 games started in the NFL. Originally, our formula weighs evenly between the defense's fantasy PPG allowed and the quarterback's past fantasy performance in the league. The new formula is now:

$$0.0163(Opponent\ PPG\ allowed\ 2014)\ +\ 0.9837(avg.PPG\ in\ career)$$

This goes against our typical thinking when it comes to projecting the performance of the quarterbacks. We usually factor the defense's statistics over half of the weight but with this new projection system, it is essentially not a factor. This is most likely due to the fact that there were only two quarterbacks that fell into this category this year. The projection formula will improve as more data is collected. We would predict that these weights would eventually even out as we have more data on quarterbacks in this part of their career. This new projection formula under predicts by 0.0074 fantasy points overall, which breaks down to 0.00014 points per quarterback per game. The original prediction formula over predicts by 42.51 points, which is 0.78 points per quarterback per game.

If we continue this project further, for this category we could consider adding another parameter to the projection and seeing if that can factor into a more accurate projection that still includes the defense's performance. We want to keep that category in the projection system because it has a large impact on all of the other categories projections.

The final category ended up having the least change when comparing the original and final formulas. This is most likely due to the fact that this is the most populated category, we have data on how a quarterback performed against each team, and that quarterbacks usually are

the most predictable out of all the categories. Once again, we kept the parameters constant when we adapted the new formula. It is now:

$$0.6[part\ 1] + 0.5485[part\ 2] + 0.49[part\ 3] + 0.532[part\ 4]$$

*Reference each part's definition from page 5

We decided to no longer tie the sum of the weights to 2, since this works out better to match the quarterbacks in the group's performance. Originally, the formula under projected the 23 quarterbacks in the category by 220.16 fantasy points. This ends up being 0.031 points per quarterback per game started. We were very pleased by the precision of this estimate but strived to optimize it like the other formulas. The new formula under projected by 0.0081 fantasy points, which is 0.00000114 fantasy points per quarterback per game.

What is different about this category is that the sum of the weights is a little more than two, 2.1705 to be precise. Since this formula has four parameters shaping its projection, during the formula optimization process, we decided to put precedence on a near perfect projection for all of the quarterbacks, and did not want to limit ourselves to the constraint of summing the coefficients to two.

## *Conclusion*

This project has been very successful. With our new coefficients that have been shown above, we plan on using the updated formulas to create a new projection for the upcoming NFL season (2016). We then will be able to test our new formula against our old one and see which one is a better projection model for starting quarterbacks. We will collect Yahoo!'s weekly projections for each quarterback. This allows us to have a consistent comparison for our projections through the entire season. We will also be able to compare how much Yahoo changes each week based off the previous week. This will be interesting as Yahoo changes every day

throughout the entire season. With all this new data, we can make more conclusions and it will allow us to either gain confidence in our new projection models, or it will give us a reason to make adjustments for the next season. Along with gaining new information, we have been able to look back on the process of how this project was done.

There are some things we could have done better throughout this process. One main thing was the timing. We did not start this project until approximately week 2 of the NFL season. This caused us to have slightly different projections for Yahoo! Projections compared to theirs at the very beginning of the year before any game was played. Along with that, we could have created a system where we could change one thing and it would change the projection on its own. We did a lot brute work and did not develop a good system until about halfway through our analysis process. This could have been avoided in the beginning. We combatted this later by refactoring our spreadsheets the best we could without losing data. One thing we would have liked to figure out, but never got around to was to find a way to expedite the data collection process by creating a program which could automatically collect each player's statistics for the parameters that we had set for our projections. Some of these things were just trial and errors in which we learned many things from.

This project has allowed us to learn many things. One big thing we learned was to be confident in our work along with being able to explain the work. When we started this project, we knew that we would be presenting to people who did not understand fantasy football, let alone football in general. As we were going through the analysis process, we came across days where our professor who overlooked our project did not understand what we meant. From there on, we started documenting anything that could be 'confusing' to our audience and to remind

Works Cited:

Bales, Jonathan. Fantasy Football for Smart People: How to Dominate Your Draft. North Charleston, SC: CreateSpace, 2012. Print.

"ESPN." Fantasy Football. ESPN, n.d. Web. 26 Oct. 2015.

"Fantasy Sports." Yahoo Sports. N.p., n.d. Web. 26 Oct. 2015.

Kimbrell, Charles E. Fantasy Football Guide Tips on How to Draft and Play Fantasy Football. Place of Publication Not Identified: Sold by Amazon Digital Services, 2011. Print.

"Pro Football Reference." Pro-Football-Reference.com. N.p., n.d. Web. 26 Oct. 2015.