Draft version December 7, 2018
Typeset using LaTeX default style in AASTeX61

PREDICTIVE ANALYTICS TO BEAT AN EXPANDING MARKET

Andrew C Guyader, PhD, Michael Feldmann, PhD, and Ryan Black

ABSTRACT

The recent US Supreme Court decision to overturn the Professional and Amateur Sports Protection Act now gives states the decision to authorize and legislate sports gambling. The increasing NFL point spread offerings from the betting market expansion will yield more opportunities to attack market imbalances. This paper presents the critical elements of a successful game score differential predictive model founded on a proprietary performance metric. This data-driven model creates a successful baseline for betting edge using only drive data produced in NFL games. The drive performance metric is used as the foundational piece in a novel 32 degree-of-freedom non-linear model. A generalized reduced gradient algorithm is surrounded with empirically tuned factors. Analyzed over eleven NFL seasons, the model wins on game accuracy, unit bet returns and also by means of a mutual fund betting concept. The results are transparently validated.

1. EXECUTIVE SUMMARY

Sports gambling is a multi-billion dollar industry that engages about 15-percent of a fan base who invest heavily with their time and money. Betting on athletic events provides fans an increased level of active participation. This paper presents a highly successful predictive model that can be deployed to attack the rapidly expanding legal NFL betting market. The point differential model for NFL games evaluates drive performance and applies it to game and team analysis. Proprietary performance metrics create a transformation of drive data into a success continuum evaluating team performance. The metrics incorporate starting field position, net yardage, drive result and game context. Real-time graphical game summaries display every possession, drive result and drive success in precise time history, increasing analysis insights.

Game-specific and team-specific endogenous covariates are incorporated in a novel non-linear multi-dimensional model. A generalized reduced gradient solver produces an optimal solution over the 32 team space, uniquely taking into account the quality of opponents. The model prediction and point spread are compared to determine if a game triggers a bet at 1, 2 or 3 units.

The point prediction model was trained on only four years of complete regular-season NFL data and executed in real-time during the 2017 season, as shown in Figure 9 and archived on twitter @aguyader17. Model version 1.0 in 2017 recorded a 26-12 record and a mutual fund style betting model produced a 36-percent Return On Investment (ROI), including a moderate accounting for the vigorish. In preparation for the 2018 season, training was implemented on all 11 years of data from 2007 thru 2017 resulting in ROI improvements while triggering more bets. Historical analysis and week-to-week fund value visualizations are shown in Figure 5. The 2018 season has been executed in real-time with results comparable to the mean values from historical analysis.

2. INTRODUCTION

Football is game built around emotion, anticipation, and drama. The NFL delivers elite skill, extreme athleticism, speed, and power to create a spectacle unlike any other sport. Year after year, the NFL continues to successfully retain its position as America's most popular sport. The business of football is *big* and getting bigger. Revenue for the NFL exceeded $13 billion in 2015, up 50 percent from 2010 (Belzer 2016). Additionally, the NFL continues to market its product internationally as regular season games are consistently played on foreign soil.

For the many entities emotionally and financially invested in the game outcomes there exists a need for useful and accurate tools to help with the reflection, interpretation and tracking of on-field football performance. An improved understanding and unifying interpretation of the game can benefit these groups. The descriptive and predictive analytics presented in this manuscript focus on the fans and will increase their engagement opportunities with the action on the field. Moreover, data being used by the league and its teams has trickled down to the fans who can access and interact with football's many data sets.

Since the 1980's the popularity of the NFL has increased along with its fan engagement offerings. Many fan engagement offerings that have achieved popular zeitgeist have been centered on game outcomes such as pick'em contests, survivor pools, daily and season-long fantasy football, and gambling. The most popular form of gambling is point spread betting. Different from the money line where the bet is on the win/lose outcome and odds vary, the point spread bet is often $11 to win $10 from the final score adjustment. This requires a bet winning percentage of 52.4% for the bettor to break even. As the gambling market expands due to the passage of PASPA, more opportunities will arise to identify and attack market imbalances. This paper will present a completely data-driven approach to game score predictions with proof of consistently beating the point spread market.

Much work exists in several areas of predictive models for NFL game outcomes. Extensive work has been done on win/loss predictive models and overall point differential minimization approaches. Most of the published work involves analyzing all games within a season and presenting results over the entire suite of games. Little has been published or presented on the creation of betting schemes. Highlighting prediction models used for NFL point spread bets leads to a small collection of previous work.

Warner used two years of game predictions to compile a 51% prediction success against the spread (ATS). The effort included years 2008 and 2009 with one year being a winner at near 55% and the other a loser near 47% (Warner 2010). Sinha, et al. combined traditional statistics with twitter information to assist in predicting games ATS at a 52% success rate (Sinha 2013). Schlicht presented interesting work focusing on exploiting oddsmaker bias and entropy models that have shown promise in ATS success. (Schlicht 2017). Historic work from Gray investigated a betting strategy using a subset of games mainly with home underdogs and focusing on biases in recent performance (Gray 1997).

3. BUILDING THE DRIVE PERFORMANCE METRIC

The predictive model is founded upon the successful evaluation of team performance from analyzing game data only at the drive level. A performance metric introduces the transformation of drive data into a single success continuum. The metric incorporates starting field position, net yardage, drive result and game context accentuating the importance of drives initiated at the extreme ends of the field. The evaluation of drive data was composed of the empirical evaluation of the data within the lense of successful performance and winning football games. The first step was to create something accurately descriptive of the flow and feel of the football game.

The initial focus of this project was to create a meaningful football drive performance metric and a visualization platform to graphically display the analysis. This venture into descriptive analytics for football drive data incorporates over 15 years of playing and coaching experience by the primary author at the Division-1 college level. The concept was to apply the knowledge base of what wins football games from a drive possession perspective. Field position and scoring opportunities are foundational elements to the success of a football team in a contest and the creation of an algorithm to accurately reflect these beliefs.

The performance metric was applied to game data and visualized in chronological order according to the game clock. The presentation is explained in Section **??**. Feedback for the drive metric and the data presentation occurred anecdotally as the primary author used games that he coached in and extensively shared the analysis with several colleagues in the profession that they coached in. Several iterations of tuning the metric until it was thought that an accurate description and analysis of any football game resulted.

3.1. *Drive Data*

The dataset built for this study reveals that an NFL game has nearly 24 distinct possessions. Touchdowns and turnovers account for roughly one-third of all drives. Slightly more than five drives should end in a touchdown while three drives should end in a turnover. But what about the information produced in each of the remaining drives? It will be shown that carefully tracking all drive information leads to valuable insights about team strength and strongly links to the probability of winning the game. The extension from these descriptive analytics of meaningful drive performance to game predictions is an important validating marker of the performance metric algorithm. The drive performance metric utilizes information on starting field position, drive yardage, drive result and game situation. An example of the data is shown in Figure 8. Eleven seasons of NFL games (2007 thru 2017) are analyzed in this study. Over 100,000 lines of drive data were captured from espn.com and cbssports.com drive charts and game logs. The data were corroborated with inferences from NFL JSON play-by-play files.

Drive data for several 30-yard drives is shown in Table 1. To facilitate the success discussion, consider the game situation presented is *not* near the end of the half or game where clock and score have a strong influence on strategy. The amount of success for the drive is from the perspective of the offense. The quality of the special teams play either before or after the drive is not part of the drive success evaluation.

Yards	Starting FP	Drive Result	Success
30	30	TD	YES
30	97	Punt	Yes
30	75	Punt	neutral
30	55	FG attempt	no
30	35	FG attempt	No
30	60	Interception	NO

Table 1. Drive data for several 30 yard drives.

In a game, 30-yard movements of the ball on each offensive possession yields just over 300 yards for the game, as an offense has an average of 11 possessions. The fourth column in Table 1 displays the amount of success relates to the amount of capitalization on the words yes, neutral and no. This column introduces a degree of subjectivity infused into the objective data. The idea put forth is that yardage can be analyzed within the context of starting field position and that will lead to a better understanding of performance and also a better metric to use as a predictor for future performance.

3.2. *Drive Endings*

The possible drive results or outcomes at the conclusion of any possession are the following: touchdown (TD), field goal (FG), missed FG, punt, safety, turnover (by fumble or interception), downs (a failed 4^{th} down conversion), end-of-half and end-of-game. Although rare, several possessions may occur within a single play. For instance, a fumble returned for a touchdown would be two possessions - one for each team with a zero play drive for the scoring team. In all cases, the possession of the football is determined by the officials. Here forward, each possession is termed a *drive*.

A summary for drive endings in NFL games is shown in Figure 1. This analysis reveals that punts occur on almost 40% of possessions. Point scoring possessions - touchdowns, field goals and safeties - occur on approximately one out of every three drives. Drives end with a turnover, fumble or interception on just under 15% of possessions.

Drive endings by starting field position are shown in Figure 2. The horizontal axis represents the starting field position (SFP) in 5 yard increments. SFP is the yardage to the endzone at the start of the drive. The bin labelled *0* covers SFP from 1 to 4 yards, the bin labelled *5* covers 5 to 9, etc, the bin labelled *90* covers 90 to 94, the bin labelled *95* covers 95 to 99 yards. The drive ending options are color coded showing percentage of occurrence in each bin. The large count in the 80 bin reflects the large number of possessions that start after a touchback. Note that before the 2016 season, the touchback line was changed from the 20-yard line to the 25-yard line.

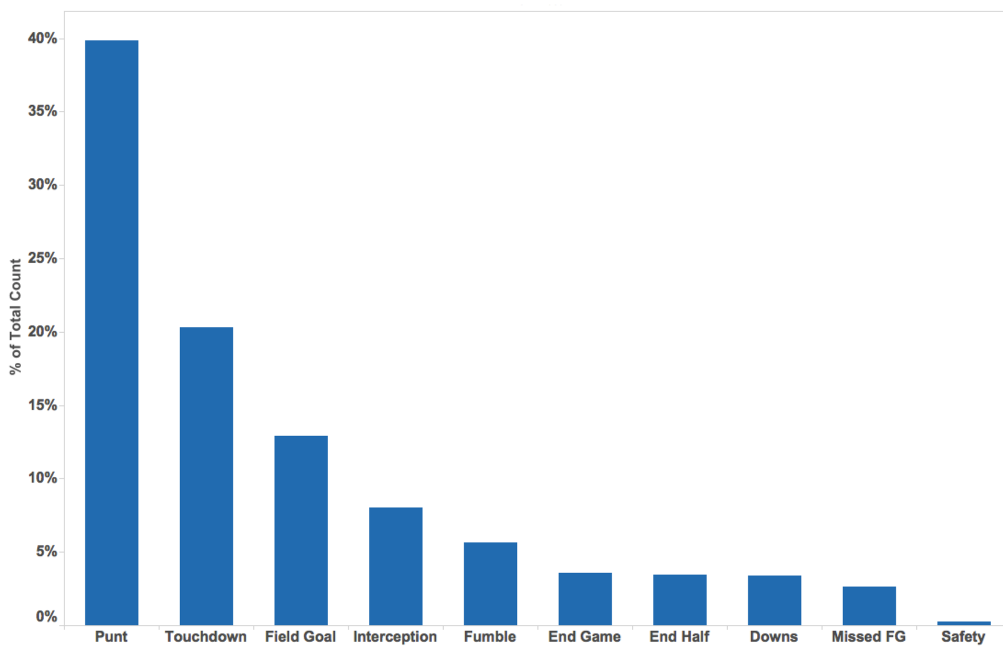


Figure 1. Drive results.

3.3. *Starting Field Position*

Drives starting at the extreme ends of the field are accentuated in the analysis due to the critical nature of the drive. Points are imminent in the plus territory and a defense senses a game changing opportunity when their opponent starts a drive backed up in their own territory. If you can make a team punt the ball with the punter near the end zone, great field position should result.

3.4. *Performance Metric Calculation on a success continuum*

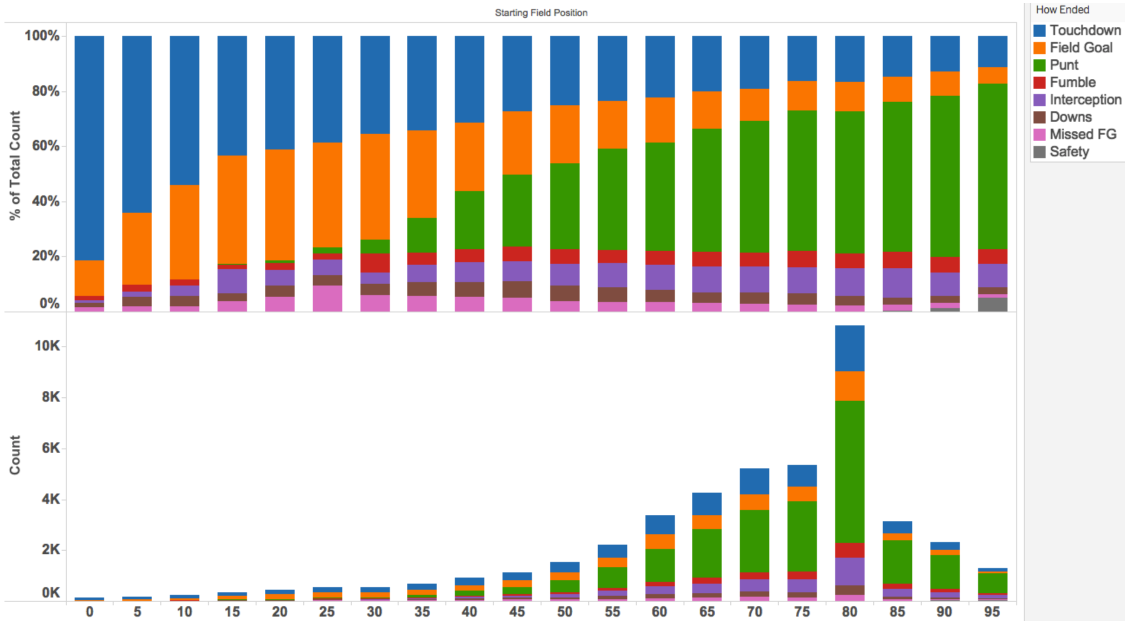


Figure 2. Drive results by starting field position excluding end of half or game.

The performance metric calculation for a drive involves the normalization of both the field position functional value and the drive success base value. The normalized functions are multiplied together and scaled to -500 to +500 - the range of the success continuum. The concept accentuates drives initiated at the extreme ends of the field.

3.5. *Metric Benchmarks*

The performance metric values for all possible drive endings are presented in the following sections. Analyses are visualized in three-dimensional surface plots with starting field position (StartingFP) defined as distance to the scoring end zone for the offense on one horizontal axis and ending field position (EndingFP) on the other horizontal axis. The vertical scale shows the performance metric value. The performance metric value range was chosen to span -500 to +500 for an individual drive. This scale is present on the right edge of the GameMap canvas described in Section **??**.

Note that in the current metric, all drives are analyzed without regard to the score of the game or the clock time left in the 2_{nd} or 4_{th} quarter. The first drive of the game is analyzed with the exact same algorithm as the last drive of the game even if the game is a blowout or if it will be decided on the final possession. This is discussed further in Section 3.6.

3.5.1. *Drives Ending in Punt*

Almost 40 percent of drives end in a punt. A punt is a special teams change of possession play where yardage is covered by the kicked ball. The punt play itself is not part of the analysis. The metric uses the yardage covered by the drive before the punt play occurs. Figure 3 shows the metric values for drive ending in a punt. The high point of the surface is for drives that have starting field position at the extreme far end of the field and end past the 50 yard line. Drives starting with good field position but ending in a punt are harshly evaluated by the metric as the defense has *risen to the occasion*. Elements of this analysis have shown to be valuable and proprietary due to their use as the foundational piece within the predictive model created. For this reason, only partial details are presented here.

3.5.2. *Drives Ending in Touchdown*

An NFL game expects 5.2 touchdowns to be scored. The performance value for drives ending in touchdowns maxes out the metric scale when the drive originates at the extreme ends of the field. The long field touchdown is perhaps more intuitive than the short field touchdown. Long drives are a show of great strength for an offense. But what about a 40 yard touchdown drive? The defensive reaction to a shorter offensive field position is key. These short drives for a TD are equally impressive because not scoring a TD would be a huge boost for the other team. The defense has been explicitly tasked to rise to the occasion so anything less than a TD is a big success for the defense. If an offense gets

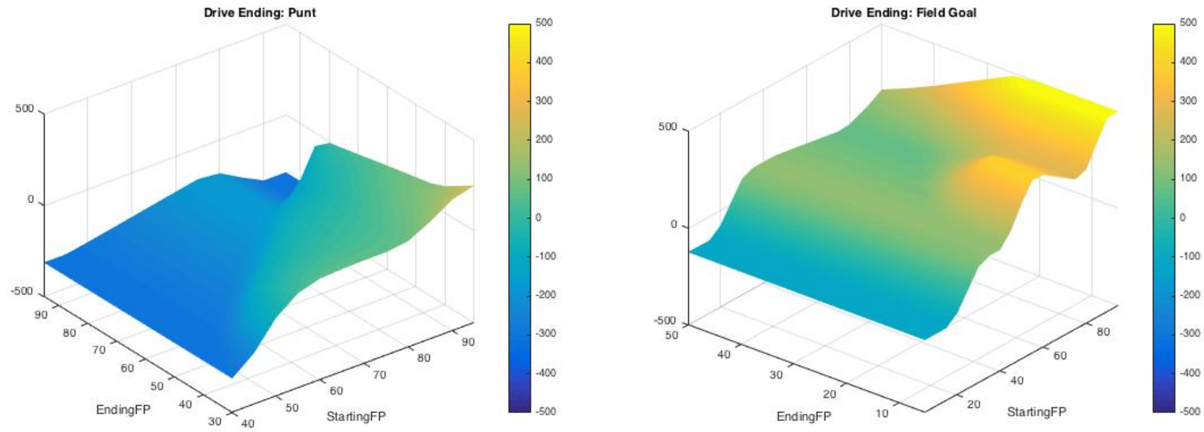


Figure 3. Metric values for drives ending in a Punt and Field Goal .

the ball on the plus-10 yard line and they don't score a touchdown, that drive is a huge failure. Again, we choose to leave out the rare instance where the field goal wins the game on the last play or moves the game from 6 to 9 points making it a two-possession game with just minutes left on the clock.

3.5.3. *Drives Ending in Field Goal*

Outside of the extremely memorable touchdowns and turnovers in a game, the field goal is a unique drive ending as it could be either successful or unsuccessful depending on the starting FP. A drive starting on the minus-5 yard line and moving down the field for a field goal is a strong show of strength by the offense. They successfully overcame unsuccessful field position.

Starting a drive with great field position perhaps on the plus-20 yard line and ending by kicking a field goal shows that great field position was squandered by the offense. This is a bad drive and although specific clock situations exist where a FG could win a game at the end. Figure 3 shows how starting field position and yards covered (ending field position) combine to create a performance metric value for drives ending in a field goal.

3.6. *Game Metric (GM)*

The average of all drive performance values from a game is defined as the Game Metric. Calculations of the Game Metric for both teams has a negative value relationship. Each team's Game Metric is the negative value of the other. Linear regression analyses reveal the association between the Game Metric and the probability of winning (more specifically, the odds of winning), as well as the association between the Game Metric and winning team point differential. Both analyses revealed significant positive associations.

In the predictive model, GM values are capped at the corresponding 24-point value so that huge blowout victories will not be unfairly bias future predictions. This cap is meant to limit the effect of extreme performances in a single game. If a team ends the game with a GM value above 104.3, the number is reduced to 104.3 for analysis in the predictive model. Future refinements to the metric will account for late game analysis in a more fashion accounting for the score in the game and clock time remaining.

3.7. *Association with Winning Point Differential*

Additional analyses were conducted to examine the association between the Game Metric and the point differential (winning team score minus losing team score) among all 1,912 winning teams from 2009 to 2015. It was of particular interest to determine whether the Game Metric is a significant predictor of the point differential. The results of the regression model is shown in Figure 4 are displayed in Table 2. The model is significant with an F-stat = 3697.27 and p-value < 0.001.

Ignoring any other potential predictors, the Game Metric explains approximately 66 percent of the variability in the point differential, and Game Metric is significantly positively associated with point differential (p-value < 0.001). Each six-unit increase in the Game Metric is associated with an approximate one point increase in the point differential on average.

	Coefficient	Standard Error
Constant	3.233*	0.191
Game Metric	0.158*	0.003

Table 2. Linear regression results for Game Metric (*p-value < 0.001).

4. PREDICTIVE MODEL FORMULATION

The point prediction model will analyze and predict performance by means of the Game Metric. A functional relationship between GM and score differential is needed in the model. A constrained least-squares linear regression through the origin results in the following equation:

$$PointDiff = 0.23 * GM \tag{1}$$

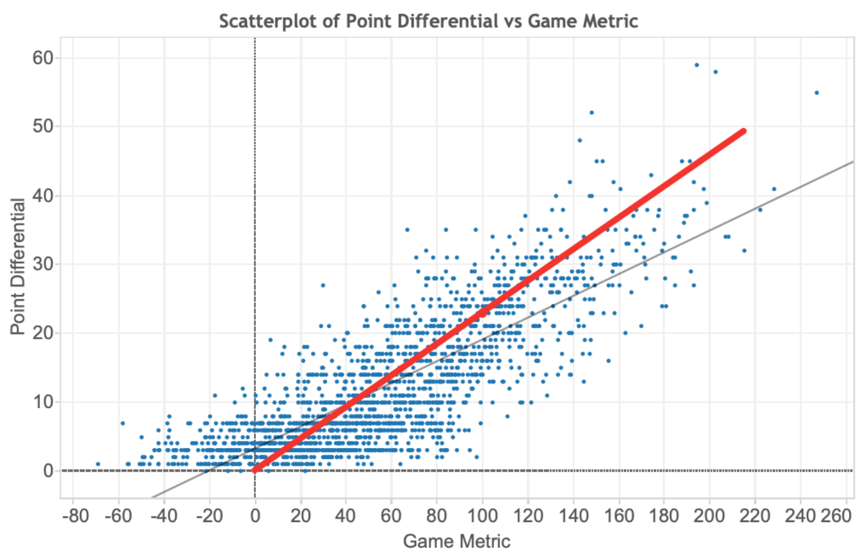


Figure 4. Point Differential vs Game Metric (GM) with unconstrained and constrained (red) linear regressions.

From week "K", the metric value for each team (i) is denoted $recentQ_{i,K}$. To solve for this team metric value for each team in week "K", the following expression is minimized:

$$min[\sum_{i=1}^{32}\sum_{j=K-3}^{K}\alpha_j(recentQ_{(i,K)} - recentQopp_{(j,K)} - GM_{(i,j)})^2] \tag{2}$$

where $recentQopp_{(j,K)}$ is the recentQ calculated in week "K" of the opponent of team "i" from week "j". This analysis spans the previous 4 games which is why the model does not begin to execute until week 5. This is a solution for the games played in the most recent 4 weeks, neglecting weeks when a bye occurred. It uses a weighting percentage of 30% for the most recent two weeks and 20% for the next two weeks. Therefor $\alpha_K = .3$, $\alpha_{K-1} = .3$, $\alpha_{K-2} = .2$, $\alpha_{K-3} = .2$. As mentioned in section 3.6, GM values are capped at 104.3 which corresponds to a 24 point differential, limiting the effects of extreme performances.

A generalized reduced gradient solution algorithm was used with the following constraint:

$$\sum_{i=1}^{32} recentQ_{(i,K)} = 0 \tag{3}$$

The minimization solution of most recent four games uniquely accounts for the quality of the teams involved. The most recent four weeks give a very good indicator of the accurate evaluation of a team. Weeks further back than 4 weeks are still relevant and are accounted for in a year-to-date calculation. In addition to the *recentQ* calculation, a more traditional year-to-date analysis is used without an accounting for the quality of opponent. This is referred to as the *seasonQ* and is calculated by:

$$seasonQ_{(i,K)} = \frac{1}{K}\sum_{k=1}^{K} GM_{(i,k)} \tag{4}$$

After the recentQ is calculated for each team, the current team evaluation, referred to as the TeamQ, will be used in the upcoming week's matchup. The TeamQ value for each team after games completed in week K is calculated as:

$$TeamQ_{(i,K)} = 0.75 * recentQ_{(i,K)} + 0.25 * seasonQ_{(i,K)} \tag{5}$$

The TeamQ values are used to create a matchup prediction for the upcoming week, week "K+1". The predicted score differential for each upcoming matchup is calculated from:

$$PredictedPTDIFF_{(i,K+1)} = 0.23 * [TeamQ_{(i,K)} - TeamQopp_{(i,K)}] \tag{6}$$

where TeamQopp is the TeamQ value for the opponent of team "i" in the upcoming week. The magnitude of $PredictedDIFF_{(i,K+1)}$ is compared to the spread for the game (the betting line).

$$Trigger_{(i,K+1)} = PredictedPTDIFF_{(i,K+1)} - Spread_{(i,K+1)} \tag{7}$$

Depending on the difference between the prediction and the spread, a 0, 1, 2 or 3-unit bet is implemented. This is based on a pre-determined tiered threshold. The run of the model spans from week 5 to week 16. Four previous games are required for predictions and week 17 is excluded due to possible roster adjustments for teams slotted or eliminated from the playoffs.

4.1. *Model version 1.0 evaluated in a mutual fund betting model*

.

The first version of the model was created by training on years 2009, 2011, 2013 and 2016. This left validation years of 2010, 2012, 2014 and 2015. The closing lines from historical seasons were captured from www.footballlocks.com. The betting model used a mutual fund approach with a unit of bet as 2% of the existing week's fund value. Parameters in the model were empirically tuned to maximize ROI on the mutual fund betting scheme. The fund was initiated at $100 and the friction applied to all winnings is 6%. This friction is not an exact representation of the vigorish and further refinements of the betting model will include precise representation of any defined betting odds.

The 2017 season was run with a real-time verification on twitter @aguyader17. The results as seen in Figure 9 represent the first pass at executing the model in a betting simulation and the results were extremely encouraging. The model record was 26 wins versus only 12 losses (68%) The model maxed at $155 and settled at $136 after week 16. The model applies the same betting rules in week 16 as it does back in week 5. No winnings are protected as the season progresses. This is a true *let it ride* betting scheme.

4.2. *Home Field and other covariates*

It is generally known that home field is worth somewhere between 2-3 points for a game in a betting line. In 2017, model version 1.0 triggered on only 38 games. It is in these games that an adjustment for home field would be considered. Analysis of the model results - both versions 1.0 and 2.1 - do not account for a home field advantage across all games. The conclusion is not that home field advantage does not exist but that its inclusion could not be justified in this model triggering only a subset of games.

The model also does not know if injuries have occurred from one game to the next. Information about teams coming off a bye or a team on a short week is also not used in the predictive model. Also the recent specific history of the matchup is not included. All of these covariates could be tested and included if proven to improve the model. This and other future work is discussed in section **??**.

5. PREDICTIVE MODEL RESULTS FOR MODEL VERSION 2.1

At the end of the 2017 season, the model parameters after the minimization step were empirically optimized with the objective to increase the number of triggered betting opportunities while maintaining the mutual fund model returns. This was successfully achieved using all eleven years of existing data and point spread information. Model version 2.1 used closing lines for 2007-2017 is using early lines for 2018 thru week 12. The game accuracy sits at 54.1% with a 4.0% standard deviation. The betting units results differ slightly from the game accuracy as the units improve to 55.6% winning percentage. Noting the 52.4% mark needed to turn a profit, three years fall below that mark while eight years sit above as winning seasons. Model version 2.1 as compared to version 1.0 shows an increase in the amount of action produced. The triggered games has increased from somewhere around 40 to near 70 games in a season.

The results also apply the same mutual fund betting scheme previously described. Figure 5 shows graphical week-by-week values of the mutual fund model and also tabular values for bet accuracy, unit return and year end ROI. The historical analysis of the model yields an average ROI of 25% with a standard deviation of just under 29%. The T-stat value is 2.88.

The analysis reveals that three of the eleven years were losers. One a big loser at -25% and two others near -10%. The other eight seasons ended with five season bunched around the overall mean value of 25% and three years of large returns. High point and low points of the mutual fund values are included in the table as is the column labelled *Actual $ Used* which calculates the amount of the original $100 needed to execute the model. This assumes that the model action first uses winnings to fill wagering requirements. The whole $100 is never completely at risk and winning early in the season protects the original $100 fund investment.

The model for the 2018 season after week 12 sat just under $120 or a 20% ROI as seen in Figures 5 and 10. If this paper is invited to present at the conference, the 2018 season weeks 13-16 will be completed.

5.1. *A note on using early lines*

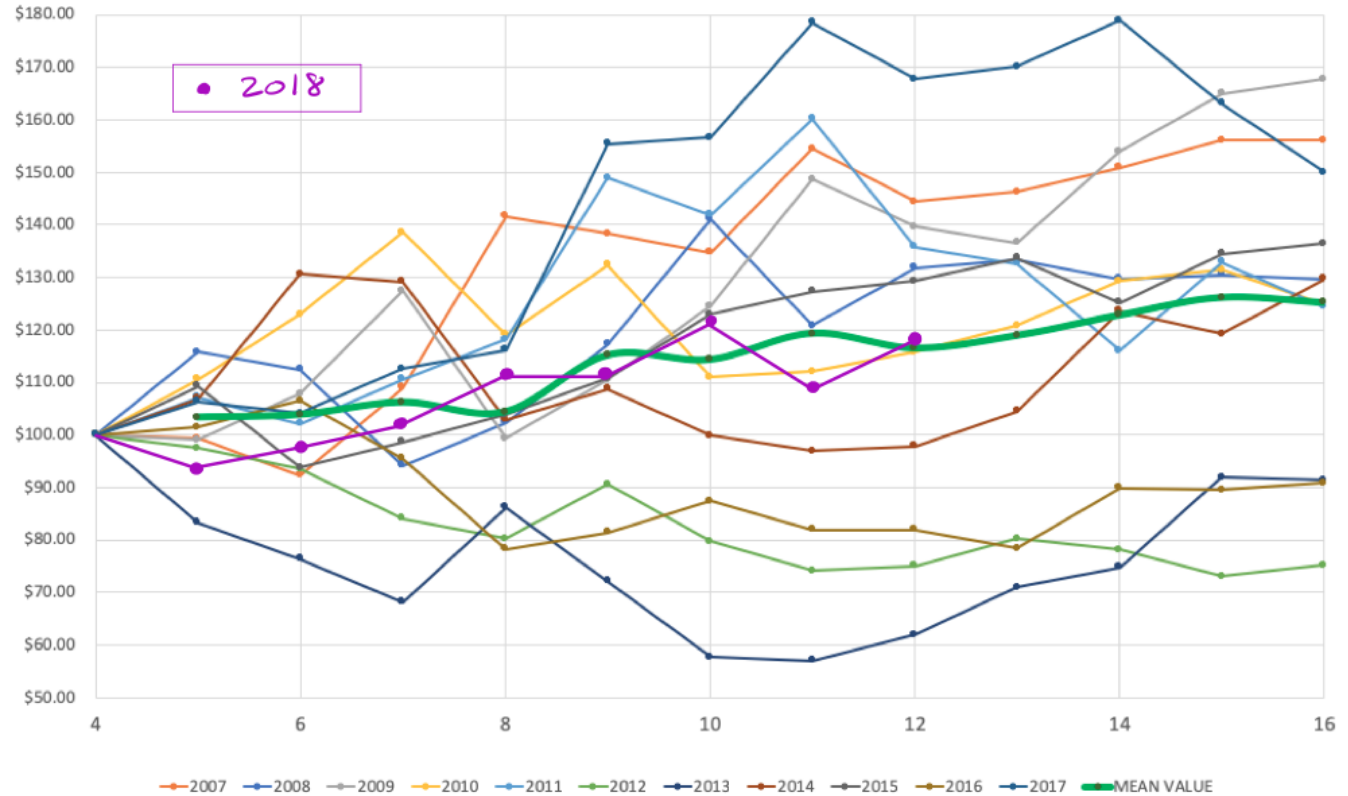
Years 2007-2016 use only closing lines in the analysis with values found at www.footballlocks.com. The 2017 season was analyzed in version 2.1 using both early lines and closing lines. Early lines outperformed closing lines in the point spread model with a bet record of 39-30 versus 36-31. The same advantage existed in the units measure. The model analysis allows for the attacking lines minutes at the end of the Monday Night Football contest as data is complete for a week of games. Due to early line advantage in 2017, the 2018 model simulation has occurred with early lines present at the end of the MNF game. All early line information is captured from www.footballlocks.com by screen shot picture. The early lines will be used in an attempt to capture more edge on the dynamic betting market.

5.2. *Margins*

Tracking the margin of bet success and failure gives a nice visualization of the margins and variability of the betting market. The plot on the right in Figure 6 reads well with bins of 5 point width. Comparing the losing 1-5 points on a spread bet to winning 1-5 points on a spread bet reveals the margins involved in a qualitative sense. Comparing successive bins moving away from the zero line continues to show the qualitative representation of the bet margin performance over the eleven years analyzed.

For each comparative bin, the positive points side has more occurrences than the negative points side. More games win by 1-5 points than lose by 1-5 points. The same is true for 6-10 point winners versus 6-10 point losers. The margin is positive on 54.1% of the bets in the analysis. Interesting to see a normal-shaped distribution when looking at the margins of victory and defeat. The most likely outcome is to win the bet by 1 to 5 points when looking at the plot on the right. When looking at the plot on the left, the most likely outcome is to win the bet by a single point. Also the plot on the left shows the second most common outcome is a bet winning by 6 points and the third most likely is a blowout win. Although 30 games were blowout wins of more than 25 points, 23 games were blowout losses of more than 25 points.

The plots qualitatively reveal some of the sensitivity and stability/instability of the betting markets. Also, the end margin is subject to late scores in a game. Some of the late scores - say a late inconsequential touchdown can push this margin to an end value that does not reflect how it felt. For instance, a two point win could come from a 10 point cushion and a TD scored with a few seconds left versus a 1 point cushion dissolved with a last second field goal. The unique stories behind the games in the -10 to 10 point range make the betting excitement a reality.



	W/L outcomes					Betting Units			Mutual Fund	%	Avg last 3	Avg last 5	High Pt. -	Low Pt -	*Actual $*
Year	Bets	Over 50%	wins	losses	Bet %	IN	OUT	Diff	$100	return	yrs	yrs	Year	Year	*Used*
2007	64	4	34	30	53.1%	131	161	30	$ 156.11	56%			$ 156.11	$ 92.32	*-$48.30*
2008	82	6	44	38	53.7%	160	180	20	$ 129.77	30%			$ 141.17	$ 94.24	*-$37.80*
2009	67	11	39	28	58.2%	141	175	34	$ 167.72	68%	$ 151.20		$ 167.72	$ 99.04	*-$35.28*
2010	71	7	39	32	54.9%	151	170	19	$ 124.95	25%	$ 140.81		$ 138.63	$ 100.00	*-$42.08*
2011	68	4	36	32	52.9%	157	174	17	$ 124.47	24%	$ 139.05	$ 140.60	$ 160.19	$ 100.00	*-$29.08*
2012	64	-2	31	33	48.4%	133	124	-9	$ 75.17	-25%	$ 108.20	$ 124.42	$ 100.00	$ 73.12	*-$53.20*
2013	72	0	36	36	50.0%	154	157	3	$ 91.34	-9%	$ 96.99	$ 116.73	$ 100.00	$ 57.03	*-$62.36*
2014	78	14	46	32	59.0%	177	197	20	$ 129.68	30%	$ 98.73	$ 109.12	$ 130.66	$ 96.95	*-$36.01*
2015	54	10	32	22	59.3%	116	134	18	$ 136.52	37%	$ 119.18	$ 111.44	$ 136.52	$ 93.72	*-$37.67*
2016	52	-2	25	27	48.1%	116	113	-3	$ 90.92	-9%	$ 119.04	$ 104.73	$ 106.51	$ 78.29	*-$47.54*
2017	69	9	39	30	56.5%	144	172	28	$ 150.06	50%	$ 125.83	$ 119.70	$ 178.93	$ 100.00	*-$44.00*
sum	741	61	401	340		1580	1757	177	$ 1,376.72						
averages	67.4	5.5	36.5	30.9	54.1%	143.6	159.7	16.1	$ 125.16	25%					
Standard Dev					4.0%			13.66	$ 28.94						
				betting units win %:			*55.6%*	T-stat:	2.88						

Figure 5. Model version 2.1 at 2% bet per unit trained on previous 11 seasons

6. EXPECTED POINTS ADDED FORMULATION BASED ON DRIVE DATA

Pioneering work developing expected points models traces back to a paper by Virgil Carter and Robert Machol (Carter and Machol 1971). In that paper, an expected points model was produced using 1^{st} and 10 situations from NFL play-by-play data from the first 56 games of the 1969 NFL season. The next points calculation was performed using ten field position buckets each with a span of 10 yards. The data set included all 1^{st} and 10 situations including at the beginning of drives and when 1^{st} downs occur within the possession or drive of the football.

6.1. *Expected Points Regression and Expected Points Added*

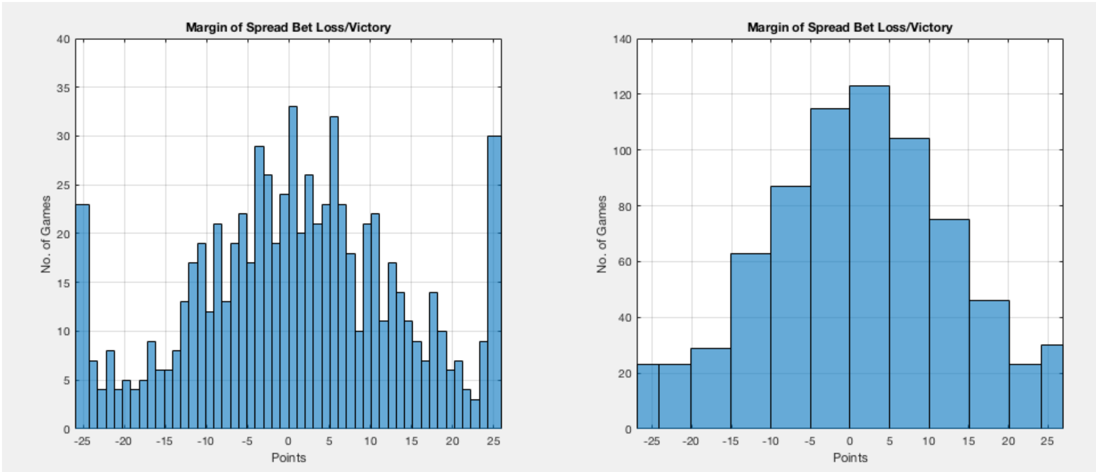


Figure 6. Distribution of bet win/loss margins for version 2.1 analyzing seasons 2007-2017

A linear regression on the average next points calculation yields the expected points (EP) equation. This analysis uses eight years of NFL data (the odd years from 2000-2016) and the resulting equation is:

$$EP(i) = -0.0646 * SFP(i) + 5.590 \tag{8}$$

The application of the EP model to individual lines of data within the context of a game further generates the need to track the team in possession of the football. The expected points derived thus far is from the perspective of the team in possession of the football and this perspective will continue throughout.

Expected Points Added (EPA) quantifies the performance of a drive by taking into account the drive result, points scored and starting field position values. The index i references the team in possession of the ball on drive i. If no points are scored on a drive, the calculation for EPA on drive i is:

$$EPA(i) = -[EP(i+1) + EP(i)] \tag{9}$$

If Points are scored on drive i by the offense then the EPA on drive i is calculated as

$$EPA(i) = Points(i) - EP(i) \tag{10}$$

Equation 10 is applied only when points are scored by the offense from a scrimmage play on a drive where possession was initiated after a punt or turnover. The possible values for $Points(i)$ are 7 for a touchdown, 3 for a field goal and -2 for a safety.

6.2. *Predictive model using expected points metric*

The Game Metric in the predictive model was replaced with the Expected Points Added game analysis. The regression used to relate point differential to GameEPAavg is shown in Figure 7. The minimization step was recalculated and the predictions evaluated. The use of the GameEPAavg did not create a successful predictive model. The predictions had an accuracy around 47% using in-sample analysis and the mutual fund simulation lost substantial money. Although the EPA analysis for plays and drives is widely used to evaluate the individual effect of plays, coaching decisions and player contributions, it was not successfully used as a game-level analysis tool used in this specific predictive model.

An attempt was also made to use the final score differential (the scoreboard) as the game metric used in the predictive model. This also was not successful and created a bad prediction model. It did not show preliminary results that warranted further study beyond the two seasons already investigated.

7. IMPLICATIONS

The performance of the metric within the framework of the predictive model further validates the performance metric and its formulation. The accuracy of predictions based only on drive data justifies the effort to expand the data

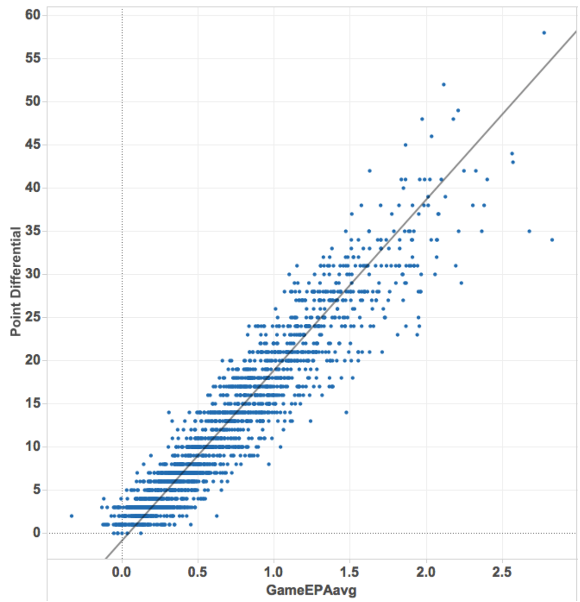


Figure 7. Point differential vs GameEPAavg

and include more covariates in the prediction model. Inclusion of a play performance metric should be thoroughly investigated. This would quickly expand the modeling possibilities to include offense-defense match-up, special teams match-up and scheme comparisons. This expansion would make over/under bets possible.

Parameters in the performance metric and predictive model could be made adjustable by each unique user to create a betting tool that reflects the philosophy of the bettor. With enough infrastructure and programming, the model would become personalized and each user could create their own specialized predictive model. User-defined parameter values would give control of the model tuning and the drive performance metric to allow for in-sample analysis of historical data. A user could accentuate elements of the game and analysis that according to their personal beliefs as to what successful football performance looks like.

Also, the expansion into other sports is justified. A similar analysis could be explored in any sport where possession data is produced. This would make sports such as rugby, soccer and basketball as candidates for the creation of performance metrics to be used in a similar predictive model.

The real-time analysis can occur which would open up in-game betting opportunities. Also, the graphical display of the analysis is worthwhile as it can enhance fan experience in several ways. The graphical game summary displays every possession, drive result and drive success in precise time history, replacing multiple clicks on summary statistics, play-by-play logs, drive charts and analysis articles. Progress in this direction has already occurred and an example of the GameMap is shown in Figure Figure 8. Drive length ticks on the horizontal game clock axis and metric values on the vertical axis marked at drive-ending locations. Distinctly colored labels define all possible possession outcomes, except punts, which are unmarked. The performance metric values are on the vertical axis, qualitatively defined on the left side of the graph and numerically represented on the right side. The horizontal line across the middle represents a performance metric value equal to zero or neutral to both teams. Above the neutral line is success for the focus team and below the neutral line is the opposite. The further from the neutral line, the more extreme the performance in either a successful or unsuccessful manner.

A quick study of this visualization by spectators or participants is meaningful in both real-time or reflection. The inclusion of hover-over technology showing drive data specifics of plays, run-pass, duration and play log will make GameMaps interactive and full of explorable data. User-defined metric parameters will be instantly applied to the metric calculations at both the drive level and play level. Direct links to play clips will allow users to explore particular footage from a game while the GameMap provides full context to game situation. Users will be able to create highlights of games or seasons from search criteria of drive and play situations and performance thresholds.

The predictive model creates a unique opportunity to build fan engagement for the majority of a sports fan base. The predictive algorithm used in the model can be reapplied using player's biometric data as the inputs. This reapplication

Drive	Team	Qtr	Yard Line	Time	HowEnded	Plays	Yds
1	NE	1	N18	15:00	Punt	6	17
2	SEA	1	S16	11:34	Punt	3	8
3	NE	1	N32	9:20	Interception	13	58
4	SEA	1	S14	1:39	Punt	5	12
5	NE	2	N35	13:57	Touchdown	9	65
6	SEA	2	S20	9:47	Punt	3	2
7	NE	2	N28	8:05	Punt	3	7
8	SEA	2	S30	7:07	Touchdown	8	70
9	NE	2	N20	2:16	Touchdown	8	80
10	SEA	2	S20	0:31	Touchdown	5	80
11	SEA	3	S20	15:00	Field Goal	7	72
12	NE	3	N20	11:09	Interception	6	12
13	SEA	3	S50	8:07	Touchdown	6	50
14	NE	3	N35	4:54	Punt	3	1
15	SEA	3	S20	3:15	Punt	4	33
16	NE	3	N14	0:55	Punt	3	8
17	SEA	4	S36	14:17	Punt	3	-5
18	NE	4	N32	12:10	Touchdown	9	68
19	SEA	4	S20	7:55	Punt	3	5
20	NE	4	N36	6:52	Touchdown	10	64
21	SEA	4	S20	2:02	Interception	7	79
22	NE	4	N1	0:20	End of Game	2	17

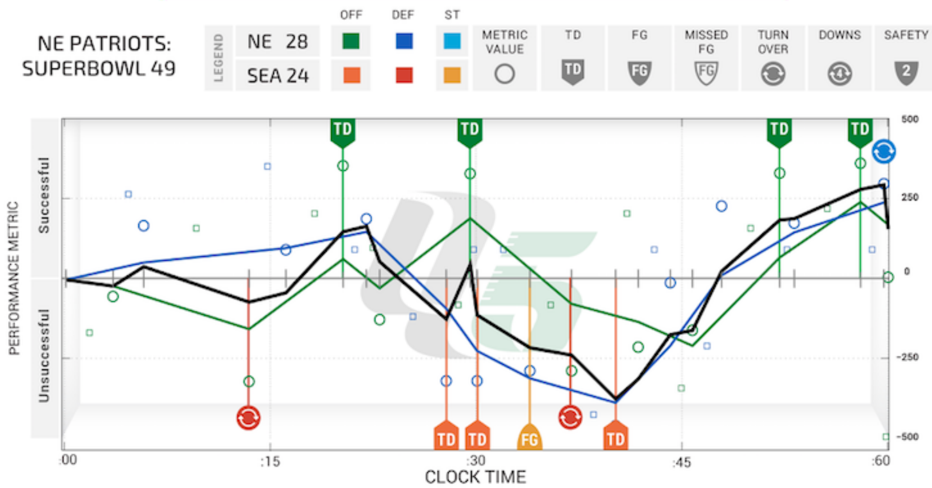


Figure 8. Data and visualization for Super Bowl XLIX.

creates a unique fan engagement opportunity. Just as various wagers engage fans in various outcomes, our predictive model allows us to build engagement opportunities around game highlights, and play-by-play outcomes.

Successful prediction of biometric data points and their occurrence presents a massive, low-risk opportunity to engage causal fans by rewarding their participation. Similar to a Las Vegas slot-machine, the low risk, low variance, and high volume provides fans constant engagement opportunities to win without the time and investment required for betting. Utilizing the biometric data of athletes provides an on-demand experience where fans can respond to queries on player performance at any point during a game thus allowing fans to invest their time as they choose.

The success of the predictive model presents an untapped opportunity to create and design an experience capable of engaging the growing and technology-inclined millennial audience. At present, 80 percent of fans are already on their phones while watching sports and 38-percent of all millennials identify as sports fans who prefer non-traditional engagement options. Our predictive framework applied to fan engagement will be tailored to those who prefer mobile, on-demand, low cost options and will fill the void left between the gambler/fantasy player and the casual fan.

7.1. *Super Bowl XLIX*

The GameMap for SuperBowl XLIX is shown in Figure 8. The visualization uses a trend line calculation using the most recent four-drive performance metric values. The weighting function was chosen to be 40 percent for the more recent metric value; and 30 percent, 20 percent and 10 percent for the next three drives in re- verse chronological order. The selection of this weighting function is arbitrary but reflects the philosophy that the residual effect of past drives decrease as time marches forward.

The New England Patriots defeated the Seattle Seahawks, 28 to 24, to secure their fourth championship. After the teams were tied 14 to 14 at halftime, Seattle built a 10-point lead to end the third quarter. New England, however, rallied to take a 28 to 24 lead with roughly two minutes left in the game. Seattle threatened to score in the final moments, driving the ball to New England's one yard line. With 26 seconds remaining in the game, Seattle decided to pass the ball in what has become a highly scrutinized call that resulted in New England's undrafted rookie Malcolm Butler making a game-saving interception in the end zone.

8. CONCLUSIONS

The model works without any human intervention. Including qualified human intervention could only improve bet accuracy.

The drive performance metric has been further validated as an accurate and useful measure of on-field performance using only drive data. The GameMap visualization canvas accurately reflects the action on the field in a largely qualitative format. The metric has been implemented as the foundational element to a highly successful predictive score differential model. The game accuracy averaged 54.1% but the unit levels successfully increased the win percentage to 55.6%. The average ROI for the mutual fund model version 2.1 is just over 25% over the 11 years analyzed with a standard deviation of 28.9%.

The 2018 season executed in real time thru week 12 using lines available immediately upon the completion of the MNF game each week. The game accuracy sits at 59.1% (26 wins vs 18 losses) and a mutual fund model ROI just under 20% which matches the week 12 mean value of model version 2.1 over the 11 previous years.

There is more meat left on the bone. This can still be improved - several legitimate opportunities are low hanging fruit.

This is a legitimate investment opportunity as in a mutual fund model. Different from other mutual fund models - there is entertainment value in the investment/bet.

This model creates a unique opportunity to assist and guide NFL bettors as the gambling market expands in the coming years. The analysis procedure allows for an opportunity to remove emotion from the gambling process of picking sides in NFL games against the spread. The analysis allows for wagering to be made minutes after the conclusion of Monday Night Football.

REFERENCES

Alamar, Benjamin. (2013) *Sports Analytics*, Colombia University Press, New York.

Bassett, G.W. (1997). *Robust Sports Ratings Based on Least Absolute Errors*, The American Statistician, 39, 152-155.

Belzer, Jason. *Thanks To Roger Goodell, NFL Revenues Projected To Surpass $13 Billion In 2016* Taken from: http://www.forbes.com.

Balreira, Eduardo Cabral , Brian K. Miceli and Thomas Tegtmeyer. (2014). *An Oracle Method to Predict NFL Game*, Journal of Quantitative Analysis in Sports. Vol. 10, Issue 2.

Bhandari, Inderpohl, et al. (1997). *Advanced Scout: Data Mining and Knowledge Discovery in NBA Data*, Data Mining and Knowledge Discovery, 1,121-125.

Boulier, Bryan and H.O. Stekler (2003). *Predicting Outcomes of National Football League Games*, International Journal of Forecasting, 19(2003), 257-270.

Burke, Brian. (2012). *Expected Points and Expected Points Added Explained*, Taken from: http://www.advancedfootballanalytics.com.

Carroll, Bob, Pete Palmer and John Thorn. (1988). *The Hidden Game of Football*, Warner Books, New York.

Carter, Virgil and Robert Machol. (1971). *Technical Note - Operations Research on Football*, Operations Research. 19(2):541-544.

David John A., Pasteur R. Drew, Ahmad M. Saif and Janning Michael C. (2011). *NFL Prediction using Committees of Artificial Neural Networks*, Journal of Quantitative Analysis in Sports, De Gruyter, vol. 7(2), pages 1-15, May.

Dixon, M. J. and S. G. Coles. (1997). *Modelling Association Football Scores and Inefficiencies in the Football Betting Market*, Applied Statistics. 46:265280.

Fry, Michael J and F. Alan Shukairy. (2012). *Searching for Momentum in the NFL*, - Journal of Quantitative Analysis in Sports, Manuscript 1362.

Glickman, M. E. and H. S. Stern. (1998). *A State-Space Model for National Football League Scores*, Journal of the Amercan Statistical Association. 93:2535.

Goldner, Keith. (2012). *A Markov Model for Football: Using Stochastic Processes to Model a Football Drive*, Journal of Quantitative Analysis in Sports. Manuscript 1400.

Testing Market Efficiency: Evidence From The NFL Sports Betting Market.

Harville, D. (1980). *Predictions for National Football League Games via LInear Model*, Journal of the American Statistical Association. 75:516-524.

Jackson, D. A. (1994). *Index Betting on Sports.* Journal of the Royal Statistical Society. Series D 43(2):309-315.

Kaslis, D. and I. Ntzoufras. (2003). *Analysis of Sports Data by using Bivariate Poisson Models*, The Statistician 52(3):381393.

Lock, Dennis and Dan Nettleton. (2014) *Using random forests to estimate win probability before each play of an NFL game* Journal of Quantitative Analysis in Sports. Vol.10, Issue 2.

Macdonald, Brian and WIlliam Pulleyblank. (2014). *Realignment in the NHL, MLB, NFL, and NBA*, Journal of Quantitative Analysis of Sports. Vol. 10, Issue 2.

Palmer, Pete. (1989). *Hidden Game of Football*, .

Pattani, Alok. Expected points and EPA explained. http://www.espn.com/nfl/story/_/id/8379024/nfl-explaining-expected-points-metric
2012

Quenzel, Jared and Paul Shea. (2016). *Predicting the Winner of Tied National League Football Games: Do the Details Matter?*, Journal of Sports Economics. 17, 661-671.

Romer, David. (2006). *Do Firms Maximize? Evidence from Professional Football*, Journal of Politcal Economy. 114:2, 340-365.

Schlicht, Erik. (2017) *Exploiting oddsmaker bias to improve the prediction of NFL Outcomes.*

Silver, Nathan. (2012). *The Signal and the Noise*, Penguin Group, USA.

Predicting the NFL Using Twitter.

Skinner, Brian. (2012). *Scoring Strategies for the Underdog: A General, Quantitative Method for Determining Optimal Sports Strategies*, Journal of Quantitative Analysis in Sports. Vol. 7, Issue 4.

Stern, H. S. (1991). *On the Probability of Winning a Football Game*, The American Statistician. 45:179183.

Stern, H. S. (1994). *A Brownian Motion Model for the Progress of Sports Scores*, Journal of the American Statistical Association. 89:11281134.

Warner, Jim (2010) *Predicting Margin of Victory in NFL Games: Machine Learning vs the Las Vegas Line.*

2017 Bet Summary

	Team	Line	Bet	W/L	OUT $	Fricton 6% on winnings	$ 100.00	RECORD	
Week 5	BUF	3	$ 2.00	L	$ -	$ -			
	JAX	8	$ 2.00	W	$ 4.00	$ (0.12)			
	KC	-1	$ 2.00	W	$ 4.00	$ (0.12)			
	LAC	3	$ 2.00	W	$ 4.00	$ (0.12)			
	MIN	-2.5	$ 2.00	W	$ 4.00	$ (0.12)			
	NYJ	PK	$ 2.00	W	$ 4.00	$ (0.12)			
	TEN	-2.5	$ 2.00	L	$ -	$ -			
			$ 14.00		$ 20.00	$ (0.60)	$ 105.40	5W - 2L	71%
Week 6	TB	-1.5	$ 2.11	L	$ -	$ -			
			$ 2.11		$ -	$ -	$ 103.29	5W - 3L	63%
Week 7	KC	-3	$ 6.20	L	$ -	$ -			
	JAX	-3.5	$ 4.13	W	$ 8.26	$ (0.25)			
	PHI	-4.5	$ 2.07	W	$ 4.13	$ (0.12)			
			$ 12.40		$ 12.40	$ (0.37)	$ 102.92	7W - 4L	64%
Week 8	NYJ	6	$ 2.06	W	$ 4.12	$ (0.12)			
			$ 2.06		$ 4.12	$ (0.12)	$ 104.86	8W - 4L	67%
Week 9	BAL	5.5	$ 4.19	W	$ 8.39	$ (0.25)			
	CAR	-2	$ 2.10	W	$ 4.19	$ (0.13)			
	DAL	PK	$ 2.10	W	$ 4.19	$ (0.13)			
	JAX	-5	$ 2.10	W	$ 4.19	$ (0.13)			
	LAR	-3.5	$ 4.19	W	$ 8.39	$ (0.25)			
	NO	-7	$ 6.29	W	$ 12.58	$ (0.38)			
	PHI	-7.5	$ 4.19	W	$ 8.39	$ (0.25)			
			$ 25.17		$ 50.33	$ (1.51)	$ 128.51	15W - 4L	79%
Week 10	MIN	-2	$ 5.14	W	$ 10.28	$ (0.31)			
	NE	-7.5	$ 7.71	W	$ 15.42	$ (0.46)			
	NO	-2	$ 2.57	W	$ 5.14	$ (0.15)			
	JAX	-4	$ 2.57	L	$ -	$ -			
	DAL	3	$ 5.14	L	$ -	$ -			
			$ 23.13		$ 30.84	$ (0.93)	$ 135.30	18W - 6L	75%
Week 11	NE	-7	$ 8.12	W	$ 16.24	$ (0.49)			
	TB	PK	$ 5.41	W	$ 10.82	$ (0.32)			
			$ 13.53		$ 27.06	$ (0.81)	$ 148.01	20W - 6L	77%
Week 12	JAX	-4.5	$ 8.88	L	$ -	$ -			
			$ 8.88		$ -	$ -	$ 139.13	20W - 7L	74%
Week 13	NE	-8	$ 2.78	W	$ 5.57	$ (0.17)			
	NYJ	3.5	$ 8.35	W	$ 16.70	$ (0.49)			
			$ 11.13		$ 22.26	$ (0.66)	$ 149.61	22W - 7L	76%
Week 14	GB	-3	$ 2.99	W	$ 5.98	$ (0.18)			
	OAK	4	$ 2.99	L	$ -	$ -			
	BAL	6	$ 5.98	W	$ 11.97	$ (0.36)			
			$ 11.97		$ 17.95	$ (0.54)	$ 155.05	24W - 8L	75%
Week 15	ARI	4.5	$ 3.10	L	$ -	$ -			
	BAL	-7.5	$ 6.20	W	$ 12.40	$ (0.37)			
	LAC	-1	$ 9.30	L	$ -	$ -			
	PHI	-8	$ 9.30	L	$ -	$ -			
	SF	2	$ 3.10	W	$ 6.20	$ (0.19)			
			$ 31.01		$ 18.61	$ (0.56)	$ 142.09	26W - 11L	70%
Week 16	LAR	-6.5	$ 5.68	L	$ -	$ -			
			$ 5.68		$ -	$ -	$ 136.41	26W - 12L	68%
			total bet		total won	total friction			
			$ 161.06		$ 203.56	$ (6.10)			

Figure 9. Model version 1.0 game bet history for 2017

Week 5	GB	-1.5	$ 4.00	L	$ -	$ -		
	WSH	6.5	$ 2.00	L	$ -	$ -		
			$ 6.00		$ -	$ -	**$ 94.00**	**0W-2L**
Week 6	ARI	10.5	$ 3.76	W	$ 7.52	$ (0.23)		
	CHI	-3	$ 5.64	L	$ -	$ -		
	DAL	3	$ 5.64	W	$ 11.28	$ (0.34)		
	KC	3	$ 5.64	Push	$ 5.64	$ -		
	LAC	-1.5	$ 3.76	W	$ 7.52	$ (0.23)		
	LAR	-7	$ 5.64	L	$ -	$ -		
	NYG	3	$ 3.76	L	$ -	$ -		
	PIT	2.5	$ 3.76	W	$ 7.52	$ (0.23)		
	SEA	-3	$ 5.64	W	$ 11.28	$ (0.34)		
	TB	3.5	$ 3.76	L	$ -	$ -		
			$ 47.00		$ 50.76	$ (1.35)	**$ 96.41**	**5W-6L**
Week 7	BUF	6.5	$ 5.78	L	$ -	$ -		
	CHI	3.5	$ 1.93	L	$ -	$ -		
	DET	1	$ 1.93	W	$ 3.86	$ (0.12)		
	HOU	4.5	$ 5.78	W	$ 11.56	$ (0.35)		
	NO	2.5	$ 1.93	W	$ 3.86	$ (0.12)		
	NYG	6	$ 3.86	W	$ 7.72	$ (0.23)		
			$ 21.21		$ 27.00	$ (0.81)	**$ 101.39**	**9W-8L**
Week 8	BAL	-1	$ 4.06	L	$ -	$ -		
	HOU	-7	$ 2.03	W	$ 4.06	$ (0.12)		
	IND	-1.5	$ 2.03	W	$ 4.06	$ (0.12)		
	PIT	-7.5	$ 2.03	W	$ 4.06	$ (0.12)		
	SEA	2.5	$ 4.06	W	$ 8.12	$ (0.24)		
	WSH	-1	$ 6.08	W	$ 12.16	$ (0.36)		
			$ 20.29		$ 32.46	$ (0.97)	**$ 112.58**	**13W-9L**
Week 9	KC	-8	$ 4.50	W	$ 9.00	$ (0.27)		
	PIT	3	$ 2.25	W	$ 4.50	$ (0.14)		
	NYJ	3	$ 2.25	L	$ -	$ -		
	CHI	8.5	$ 4.50	W	$ 9.00	$ (0.27)		
	WSH	-1.5	$ 6.75	L	$ -	$ -		
	DAL	-4	$ 2.25	L	$ -	$ -		
			$ 22.50		$ 22.50	$ (0.68)	**$ 111.91**	**16W-12L**
Week 10	KC	-16.5	$ 4.48	L	$ -	$ -		
	LAC	-9.5	$ 6.71	W	$ 13.42	$ (0.40)		
	NO	-4.5	$ 2.24	W	$ 4.48	$ (0.13)		
	NYG	3	$ 2.24	W	$ 4.48	$ (0.13)		
	WSH	2.5	$ 4.48	W	$ 8.96	$ (0.27)		
			$ 20.15		$ 31.34	$ (0.94)	**$ 122.16**	**20W-13L**
Week 11	CAR	-4	$ 7.33	L	$ -	$ -		
	CHI	-3	$ 2.44	W	$ 4.88	$ (0.15)		
	TEN	2.5	$ 7.33	L	$ -	$ -		
	DEN	7	$ 4.89	W	$ 9.78	$ (0.29)		
	ARI	-4	$ 2.44	L	$ -	$ -		
	PIT	-5.5	$ 7.33	L	$ -	$ -		
	KC	2.5	$ 4.89	W	$ 9.78	$ (0.29)		
			$ 36.65		$ 24.44	$ (0.73)	**$ 109.21**	**23W-17L**
Week 12	CHI	-4	$ 6.55	W	$ 13.10	$ (0.39)		
	LAC	-11.5	$ 2.18	W	$ 4.36	$ (0.13)		
	NYG	6	$ 6.55	W	$ 13.10	$ (0.39)		
	SF	3.5	$ 4.37	L	$ -	$ -		
			$ 19.65		$ 30.56	$ (0.92)	**$ 119.21**	**26W-18L**

Figure 10. Model version 2.1 in 2018 game bet history thru week 12