

Predicting Coverages on NFL Third Down Plays

Classification Modeling and Factor Exploration

Erik Fretland

Predicting Coverages on NFL Third Down Plays

Erik Fretland

May 2018

1 Abstract

This study examines the feasibility of using a classification tree model to predict the types of coverage faced by NFL receivers on 3rd down plays, in addition to examining the factors that are most predictive of these coverages.

2 Introduction

If you know the enemy and know yourself, you need not fear the result of a hundred battles.

- Sun Tzu

Thinking of the game of American football brings to mind images of intense struggles, barely controlled violence, and rare feats of physical strength and speed. However, to win football games on a consistent basis requires winning a mental battle even more so than winning a physical one. For this reason, football coaches spend countless hours gameplanning for opponents, studying their tendencies, identifying their strengths and weaknesses as well as self-scouting their own, and preparing ways to leverage their own team's strengths against the other team's weaknesses. This level of preparation carries over to the players on the team, as players spend time every week watching film of opponents so that they know what to expect when Friday, Saturday, or Sunday comes around. At

its most fundamental level, all this preparation serves the purpose of learning as much about the opponent as possible in order to know what they will do in a given situation. In football, like in war, knowing the opponent gives one a key advantage.

With the availability of data and access to film, as well as the increasing usage of predictive modeling, it seems almost inevitable that coaches at all levels of football will look to take a more data-heavy approach to many aspects of their job, from pre-game preparation to even in-game playcalling. However, even if coaches are not willing to go quite so far as to let a program call plays, the information that predictive modeling can provide them can give them a significant advantage. For example, if an offense is facing third and long and the defense lines up in a Cover-0, blitz-heavy look, having the ability to know if they actually will send a heavy blitz or instead drop back into coverage could be invaluable. Although this study does not approach quite that level of predictive complexity, there are still useful advantages to be obtained from knowing if a receiver will be facing man or zone coverage, which is instead what this study focuses on.

3 Literature Review

Many of the concepts that this study touches on have been addressed, or at least discussed, by other researchers in a variety of fields ranging from academic studies to being employed by a competitive football organization. Fortunately, there are both a wide enough variety of preexisting research that can provide guidelines and suggestions for future research, as well as numerous enough applications for applied statistics and machine learning models to optimize football decision making that even with the large quantity of preexisting research, there are countless new avenues to explore. The majority of literature that was found that was relevant to the subject of my search was related to either optimizing offensive playcalling relative to expected values of league-wide playcalls, predicting offensive playcalls (run or pass), or identifying less straightforward ways in which machine learning could be used to give a team a competitive advantage. None that I found specifically addressed the possibility of identifying defen-

sive playcalls, at least in part because to do so requires the type of intensive play-by-play data collection that was used in this study.

One of the most useful pieces of literature that was found was Karson Ota's 2016 study on predicting whether an NFL offensive play would be a run or a pass. Ota approached his study in a similar way to how this study was approached, with the objective of taking on-field data points and predicting a per-play outcome. However, his data usage was much broader, as he used Pro Football Reference statistics for play-by-play data from all plays from scrimmage in the NFL during the 2016 season, for a total of over 30,000 data points. As a result, his models had much more data to work with, but this was offset by not having as specific of indicators as my study did. He used multiple different types of predictive models and finally settled on a logistic regression when looking at situational factors rather than the entire sample size. His final results were, among others, a slightly less than 3 percent increase in predictive capability over the naïve statistical prediction for ability to predict all playcalls across all downs and distances.

A work similar to the previously mentioned one was done by the Harvard Sports Analytics Collective, but this study was much larger in scale. The authors used Pro Football Reference's database to pull all plays from scrimmage that occurred in the NFL between the 2002 and 2014 seasons, meaning they had almost 450,000 individual plays and about 100 independent variables. The goal of the study was similar to Ota's – to determine the factors most important in predicting run or pass, and figuring out how best to use that information to inform decisions. They settled on using a gradient boosted model and were able to obtain a success rate of about 0.703 for their original test set. An added bonus of using gradient boosting is that they were able to identify which of the factors was most impactful to their overall prediction, and the results they found were pleasantly intuitive to what an educated football fan might expect – the three biggest indicators of run or pass were time elapsed in the half, proportion of offensive plays that were passes for that team the previous year, and the score differential.

In his article for X's and O's Lab, Steve Steele (former NCAA offensive coordinator) discusses the different ways in which analytics can be used to

self-scout, opponent-scout, and prescribe courses of action going forward. He identifies areas in which informed decision-making can have the largest impact on a game, including 3rd and 4th down tendencies, improving the speed of playcalling when an offense wants to go fast, and doing scouting leading up to the week to determine what the other team did well and poorly (and thus could be exploited) as well as scouting the team's own success rates to determine what worked and what could be tweaked for improvement. In Tom Taylor's 2016 article for Wired, he discussed the potential for machine learning to overtake human beings in the decision-making process for playcalling, or at the very least, the various ways in which predictive analytics could augment the decisionmaking process. He discussed a history of AI and machine learning, starting with a history of human-vs-computer chess matches and then making the parallel for the far more complex chessboard that is the football field. However, the focus of his article was on the vast potential for machine learning to add to the knowledge base of the human coaches- for example, physiological analysis of the other team's players in order to determine who was most tired and who would be vulnerable to being targeted. The final component of his article addresses the value of sometimes choosing suboptimal strategies in order to catch the opponent off guard, or to break tendencies.

Brian Burke wrote two articles for Advanced Football Analytics that influenced the thought processes followed in this study. The first of these was an overview of game theory applied to NFL playcalling, specifically when considering opponent tendencies and the ability to partially predict the opponent's playcall. Burke argues that over time, an equilibrium will be reached in NFL offensive vs. defensive playcalls based off of historical tendencies as well as the expected values of each playcall relative to what the opposing team may call to counter a given play. Certain plays may be safer, but have a lower total expected value than others. Ultimately, the best playcalling strategy is to have a mixed strategy that takes into account the opponent's preferred strategy mix.

In his second article for Advanced Football Analytics, Burke makes the case that NFL coaches have actually approached an equilibrium, at least as far as 3rd down playcalling in some situations goes. In his research analyzing NFL plays from 2002-2007, Burke found that NFL teams obtained approximately the same expected value on runs in 3rd and medium situations as they did

from passing, although this equivalency is of course in large part due to the “suboptimal”, unexpected play of running in those situations.

4 Data

The data used in this study was play by play data of the Jacksonville Jaguars’ defensive third down pass plays throughout the 2017 regular season. There were 201 observations with 72 variables upon final completion of the data collection process, but some of the observations were eliminated from being used in the final model because they fell outside the Score Differential parameters.

The data was obtained through a play-by-play data collection process using NFL Gamepass, a video service provided by the NFL that allows viewers to replay all NFL games, both in the TV broadcast view as well as the far more useful “Coaches’ Film”. Coaches’ Film provides a higher angle and a continuous view of all 22 players on the field, allowing for analysis to be done of players who aren’t in the immediate vicinity of the ball. A series of variables to collect was decided upon, with over 100 factors being noted, ranging from game situation characteristics such as Opponent, Home vs. Away, Down and Distance, Time Left, etc., to player and formation data such as Formation type, Defensive Alignment, specific player alignments, etc., to actions carried out, such as where blitzes came from, defensive line stunts, and the various coverage types and actions. Data was manually input into the dataset for each play, after which it underwent multiple reviews to confirm accuracy and eliminate as much subjectivity as possible.

Because the objective of the study was to build a model that could predict the type of coverage faced by the outermost receivers in the offensive formation, the variables used were narrowed down to ones that described defensive players’ positioning, the overall alignment of the defense, and positioning on the football field. Specifically, the dependent variable was the coverage type over the rightmost receiving threat in the offensive formation (CovrRWR). The independent variables were LCBBody, LCBDepth, LCBAlign, Matchup, DeepestAtSnap, MOFOPre, Distance, and YdLine.

5 Methods

The objective of this study was to predict the correct outcome out of two potential outcomes based on a variety of independent variables. As such, the method selected to be used was a classification tree model, due to the categorical nature of the outcome variable. Although briefly discussed earlier, it's worth mentioning again that expectations were relatively low for the predictive ability of the model due to the relatively small dataset size (less than 200 observations).

5.1 Data Preparation

Throughout the data collection process, it became increasingly apparent what factors were correlated to one another and what might be used to predict various dependent variables. It also became apparent that with such a small dataset, some factors had too many levels and would not be useful in a classification model due to protections against overfitting, even if the relationships that they indicated in the sample can be extrapolated successfully out-of-sample. Therefore, a large degree of the pre-modeling process involved transforming the data into usable factors, reducing levels, and narrowing down the best potential predictors.

5.2 Other Classification Models Considered

The original goal of this study was to be able to predict the overall defensive playcall for the entire defensive unit rather than for one specific offense-vs-defense matchup. However, this proved impractical for several reasons, including the wide variety of defensive playcalls (as well as the possibility of combined coverages, i.e. the defense playing Man on the left of the field and Zone on the right) leading to the need for a multiclass predictor, and the ambiguity of understanding what the exact playcall was (often, the only way to tell what the defense's responsibilities were was to see how they reacted to the offensive players, and if the offensive players did not move in certain ways, then it would be impossible to see what the defense would do in a different situation). Ultimately,

the decision was made to focus on the type of coverage over a specific receiver, due to the higher degree of certainty in the dependent variable, the lower number of classes, and the still relatively high usefulness of that information.

5.3 Classification Tree Model

CovgRWRTree = CovgRWR (predicted by) Dist + YdLine + LCBBBody + LCBDepth + LCBLev + Matchup + MOFOPre + DeepestAtSnap

The target value was the defensive coverage over the rightmost receiving threat, specifically, if he would face man or zone coverage. Data was also collected for coverage over the leftmost receiver, but the focus for this study was arbitrarily decided to be the rightmost receiving threat. The predictors used were:

- Distance - How many yards the offense required to gain a first down (Significance: Jacksonville more frequently employed man coverage in distances under 5 yards)
- YdLine - How many yards the offense was away from their ultimate target, the opposing goal line (Significance: Jacksonville played zone coverage more frequently when in their own red zone)
- LCBBBody - Whether the LCB was aligned in a Press, Pedal, or Open position (Significance: Press was heavily indicative of Man, Open with an outside leverage was heavily indicative of zone)
- LCBDepth - How many yards away from the line of scrimmage the LCB was at the snap (Significance: This, combined with LCBLev, established an easy distinguishing element between man and zone alignments)
- LCBLev - How many yards inside or outside the rightmost receiver the LCB was at the snap
- Matchup - Whether or not specific defensive players lined up in ideal positions to defend specific offensive players (Significance: Matchup was heavily indicative of Man)

- MOFOPre - Whether or not there was a deep coverage player lined up in the middle of the field at the snap (Significance: MOF Open was a strong indicator of Cover 2 or Cover 4 zone)
- DeepestAtSnap - How many yards from the line of scrimmage the deepest defender was at the snap (Significance: For man coverage plays, the modal distance was 16, for zone, it was 14)

5.4 Measures of Predictive Ability

The two primary measures of predictive ability for this model were accuracy and balanced error rate. Accuracy simply measures the percent of the test set that was correctly predicted, while the balanced error rate is the mean of the two misclassification rates of each of the two predictor variable classes. For the final iteration of the classification model, the accuracy rate was 0.8421 and the BER was 0.12. The accuracy rate was a notable improvement over the naïve prediction of Man on every play, while the fact that the BER was lower than the overall misclassification rate tells us that a larger amount of mistakes were made classifying the less frequent of the two response variable classes. This is in itself useful because it allows us to be more confident that when the model predicts the more common of the two classes, it's even more accurate than the overall accuracy rate would indicate.

6 Findings

The classification model used allowed for an accurate prediction rate of approximately 0.82. Over 10 iterations, the average correct prediction rate was 0.8263, which would be an almost 18 percent improvement over the Naive prediction, which would have a 64.74 percent accuracy rate if Man (the most common class) was predicted for all 190 instances in the sample size. Additionally, the classification model resulted in a Balanced Error Rate of approximately 12 percent. This tells us that the majority of the misclassification comes from the more well-represented class being incorrectly classified, which gives us more confidence when the model predicts that more common class. In this case, because

Man is the more common result, when the model predicts Man we can be more confident that it is correct because the BER indicates that most of the mistakes occur when the model has predicted Zone.

6.1 Indicator Variables

This study showed that certain characteristics of defensive alignment and game situation are more strongly correlated to defensive playcall than others. Cornerback alignment was extremely indicative, namely if the cornerback was lined up within 2 yards of the LOS in a press position, he was very likely to be playing Man and if he was lined up outside the receiver in an "open" position, he was very likely to be playing Zone. Defensive matching to specific offensive players was also indicative of Man coverage, as was Middle-of-Field Closed, while Middle-of-Field Open and a generally lower DeepestAtSnap number were more indicative of Zone coverage. Lastly, on-field position had a strong effect, as shorter distances to go to convert a third down increased the probability of Man, while being closer to the Jacksonville redzone increased the probability of Zone.

7 Conclusion

In what often seems like a primarily physical competition, having a mental or strategic edge in football can often be what pushes a team over the edge to victory. One way to do that is to be able to consistently anticipate what the opponent is trying to do to counter your moves, which introduces an element of game theory to the playcalling process.

7.1 Game Theory Element

When a team (player) has a set of choices available to them, there may be one that stands out as a more optimal choice than others. However, the opposing player also has a set of choices to counter the first team. With both teams

aware of the choices available to the other, a frequent play is to implement a suboptimal strategy in hopes of catching the other player off guard. However, if the other player can predict when a new or suboptimal strategy will be played, then all advantage to that team is lost and the suboptimal strategy can be exploited [4]. To use a football example, if a defense is trying to disguise a blitz and is successful, then they may catch the quarterback by surprise and sack him before he can get the ball to an open receiver. However, if the quarterback recognizes that a blitz may be coming, he has the option to check to a screen pass, which is the perfect play to take advantage of the defense's non-standard playcall.

It is in situations like this that machine learning can be applied to football. If, historically, a team likes to send that type of blitz in a certain situation, and this play can be recognized by a subtle indicator of one defensive player lining up in a specific way, then a machine learning model could recognize that and convey that information to a coach in real-time, allowing him a strategic advantage.

Ultimately, the more information that a decision-maker has available to them in a format that they can understand and use, the more of an advantage they will have over a less aware opponent. Many football programs have resisted implementing practices that take away responsibility from coaches' traditional duties, but as time goes on, those programs will be at more and more of a disadvantage. Machine learning can reveal valuable information both about one's own team and about an opponent, and while Sun Tzu's claim that this will lead to victory in a hundred battles may be exaggerated, successful utilization of ML may just lead to the second undefeated season in NFL history.

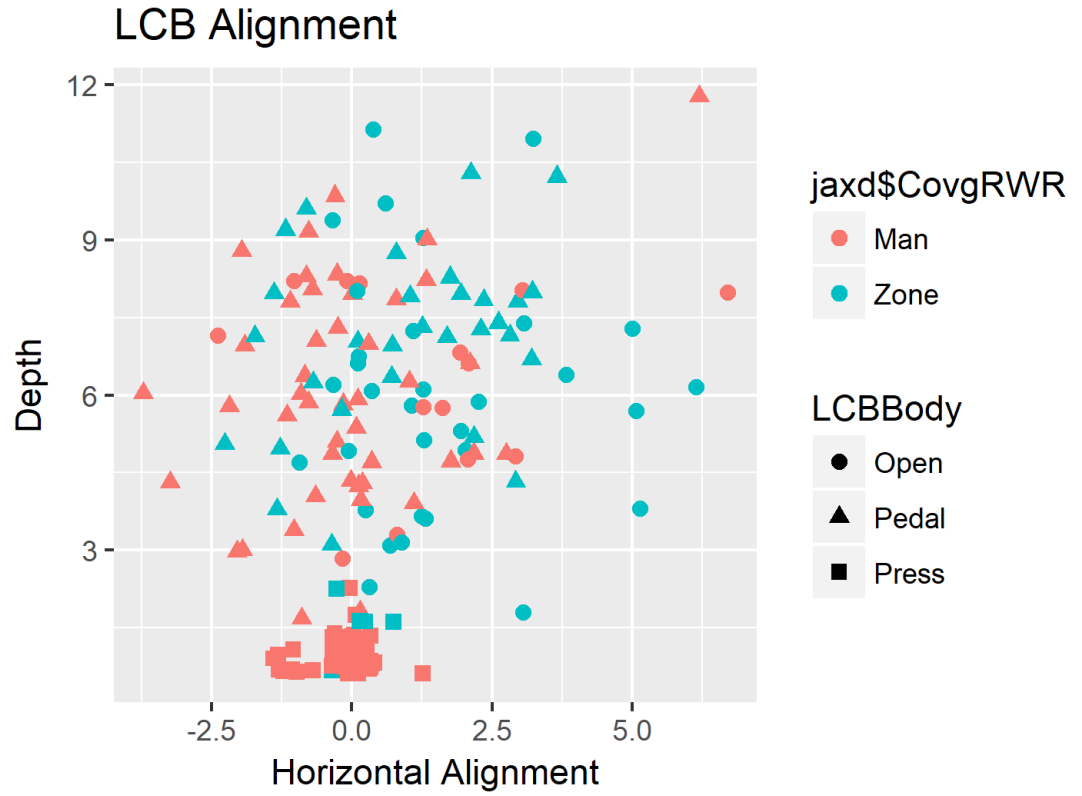
References

- [1] Matt Goldberg, Adam Gilfix, Steven Rachesky, Nathaniel Ver Steeg: Predicting Offensive Playcalling in the NFL.
<http://harvardsportsanalysis.org/2016/03/predicting-offensive-play-calling-in-the-nfl/>.
March 2, 2016
- [2] Steve Steele: Using Analytics to Guide Offensive Playcalling.
<https://www.xandolabs.com/index.php:using-analytics-to-guide-offensive-play-calling-2-cat>
August 19, 2017
- [3] Tom Taylor: Football Coaches are Turning to AI for Help Calling Plays.
<https://www.wired.com/2016/01/football-coaches-are-turning-to-ai-for-help-calling-plays/>
January 20, 2016
- [4] Brian Burke: Game Theory and Run/Pass Balance.
<http://archive.advancedfootballanalytics.com/2008/06/game-theory-and-runpass-balance.html>
June 13, 2008
- [5] Brian Burke: Playcalling on 3rd and Short.
<http://archive.advancedfootballanalytics.com/2008/08/play-calling-on-3rd-and-short-part-1>
August 8, 2008
- [6] Karson Ota. *Football Play Type Prediction and Tendency Analysis*. May 25, 2017

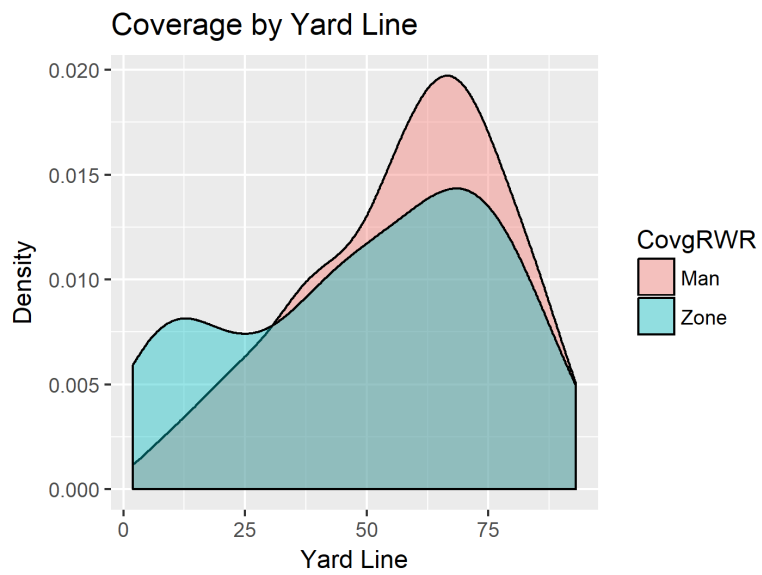
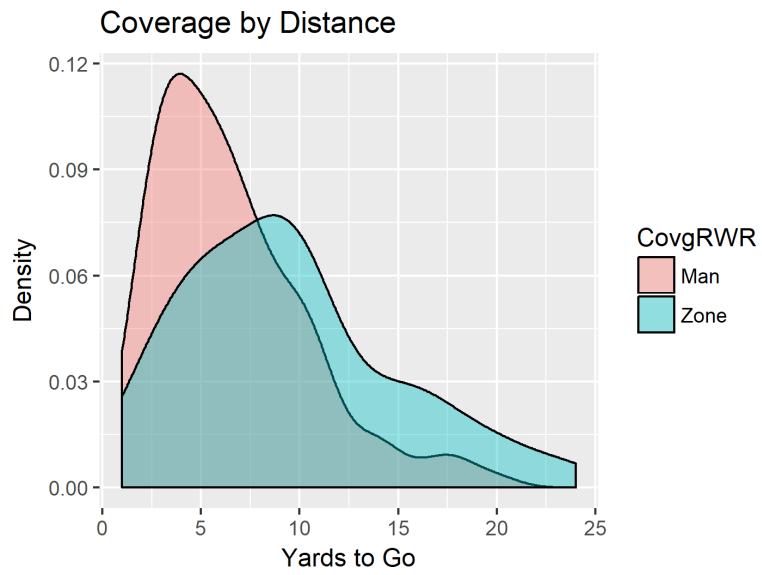
8 Figures and Tables

Table 1: Cornerback Positioning Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
jaxd.LCBLev	190	0.447	1.585	-4	7
jaxd.LCBDepth	190	4.405	3.020	1	12
jaxd.RCBLev	190	0.221	1.565	-4	13
jaxd.RCBDepth	189	4.000	2.893	5	10



Alignment of Left Cornerback relative to rightmost receiving threat.



Density distributions of types of coverage for yards to go until a first down, and distance away from the goal line, respectively.