

Using Tracking Data to Build Offensive Line Development Tools

American Football
193949

1. Introduction

American football has in recent years made drastic shifts towards the quantitative. What was once a sport that was dominated by appeals to the intangible, unquantifiable aspects of winning football games, football is now a sport where many analysts and even fans are aware of some cornerstone models that frame the way we look at the game. From Ben Baldwin's fourth down bot on Twitter/X [2], to Pro Football Focus [15] grades on Sunday Night Football and other league-sanctioned broadcasts [1], to Amazon Prime's broadcast of the game – which includes blitz probabilities and other, tracking-data derived metrics [13] – football is inundated with attempts to quantify the impacts of all 22 players on the field, as well as the coaches that guide them.

One aspect of quantifying American football that has been left relatively dormant in the public sphere has been player development. There are likely a few reasons for this; firstly, public analysts don't have a big incentive to build tools to understand player development – aside from predicting it in markets like betting or fantasy football. Public analysts cannot intervene on behalf of their models and make players better the same way a coach or analyst on a team can. Secondly, player development in the NFL is less important than it had been in the past. Prior to the 2011 Collective Bargaining Agreement, and the pandemic-induced 2020 version that largely carried over the same principles, top draft picks were paid at a level comparable to the highest-paid veterans at their position, especially at non-premium positions. For example, when Eric Berry was selected as the fifth pick in the 2010 NFL Draft by the Kansas City Chiefs, his first contract, worth \$60 million over six seasons, was the richest in the history of the safety position [18]. The first-overall pick in the 2011 draft, for reference, earned a contract of \$22 million over four years [19], with a fifth-year option at the end. Previously, high-end draft picks were not cheap, and hence there was a premium on being able to develop those players to make good on the investment over time.

With early-career players being relatively inexpensive, especially relative to the second contracts of similar players after they've been through development, player improvement has taken a backseat to other team-building elements. Would that be the case if there were, like there are for the salary cap, free agency, and other aspects of roster building, analytical tools and frameworks used to measure development, and intervene when prescribed?

In this paper, we use data from the 2023 NFL Big Data Bowl [17] to build a framework for analyzing the movement patterns of offensive linemen in pass protection. Like shot maps in basketball [12] or pitch visualizations in baseball [3] – the path dynamics of an NFL offensive linemen, adjusted for game context, can provide a lot of information from which to work in the realm of not only player evaluation, but also development. We chose offensive linemen because the benefits of an increase in efficiency regarding development has a sizable impact. Firstly, it's a weak-link system, where the play of the worst player has an outsized impact on the play of the unit specifically, and the offense generally [7]. Being able to develop down-roster offensive linemen will have a non-linear impact on

the best-laid team-building plans of NFL front offices. Additionally, since the 2011 CBA offensive line play has been in decline for several reasons related to practice time [10]. Analytically driven developmental tools can multiply the impact of each repetition a young player takes in a game or practice, which can help curb the issues associated with the CBA and fewer repetitions from which to improve.

We find that a player's deviance from his average pass blocking path is very stable from one part of the season to the next, and having such an average path allows us to look at how similar two pass protectors are. We find anecdotal evidence of teams selecting linemen with similar pass blocking path profiles as the players they replaced. We find that the rate at which players reach expected depth – a feature we built using play- and tracking data-level metrics, was also stable across different halves of the data.

Lastly, using the functional clustering method of [5], we derived clusters for pass blocking paths for all five offensive line positions. For most offensive line positions, there are clear, football interpretations for the clusters – deep pass sets versus shallower ones, for example. We examine win rates within each of these clusters, and how stable they are across different positions.

2. The Data

The data for this analysis is the first eight weeks of the 2021 season for all NFL offensive linemen and their opponents. The data consists of two components: (x,y) tracking data – the location, orientation, and direction of each player 10 times a second for the duration of the play, along with some seconds before and/or after the play. This data is courtesy of NFL Next Gen Stats [16] and is supplemented by event data from the NFL and Pro Football Focus [15]. The event data includes basic play-by-play data like down, and distance, as well as some subjective data like whether the player was beaten on a block, gave up a pressure, etc. Additional contextual elements include which coverages were played, whether there was play action on a play, etc. Table 2.1 shows which features our machine learning models use. For team success at the play level, we use data from nflfastR [2] and their EPA/success rate models.

Features	Source	Model(s)
Time (10x per second)	NGS	Curve Similarity, Clustering
(x,y) data 10x per second	NGS	Curve Similarity, Clustering
Vertical Depth at 3.5 Seconds	NGS	Expected Depth
Down	NGS	Expected Depth
Distance	NGS	Expected Depth
Field Position	NGS	Expected Depth
Offense is Home Team	NGS	Expected Depth
Time to Throw	NGS	Expected Depth
Was the Pass Thrown by 3.5 Seconds?	NGS	Expected Depth
(x,y) Position at the Snap	NGS	Expected Depth
Pass Set	NGS	Expected Depth
Max Depth of OL	NGS	Expected Depth
Expected Points Added	nflfastR	Expected Depth
Successful Play	nflfastR	Expected Depth

Block Beaten	PFF (through Big Data Bowl)	Expected Depth
Position (Play Level)	PFF (through Big Data Bowl)	Expected Depth
Defenders in the Box	PFF (through Big Data Bowl)	Expected Depth
Play Action	PFF (through Big Data Bowl)	Expected Depth
Coverage Type	PFF (through Big Data Bowl)	Expected Depth
Dropback Type	PFF (through Big Data Bowl)	Expected Depth
Offensive Formation	PFF (through Bid Data Bowl)	Expected Depth

Table 2.1: the features used in the models built in Section 3

Many of the features above are raw from the Big Data Bowl feed (time, spatial coordinates, coverage type), while others are engineered by us (the set taken by an offensive lineman – inside or outside, for example). In [9] we showed how charting and tracking data, when used together, can offer predictive power that is significantly better than either is individually.

3. Models

3.1. Average Curves and Comparison Scores

The simplest analysis one can do with this data is looking at the average path taken by a player at a position and look at deviations between that player and his average path, or the difference between that player's average path and that of others at his position.

The average over a sample like eight games in the NFL is going to be a little noisy. Thus, we smooth out the average player's path over all his pass-blocking using Bezier curves as in [5].

Once the average curve for each player at each position is estimated, we compute the average curve distance from the average curve for each player, as well as between each pair of players. In the former case we can look at how volatile a player's pass sets are, and how stable that volatility is. In the latter case we can look at different types of players by looking at the similarities in their pass sets. Similarity is a big part of what teams are looking for in free agency and the draft, as they are often trying to plug and play a new player with the same role as departed player. Curve similarity was measured using Frechet distance [11], which considers that path of each player when determining distance. The final curve similarity number is the inverted Frechet distance, plus one unit to deal with singularities.

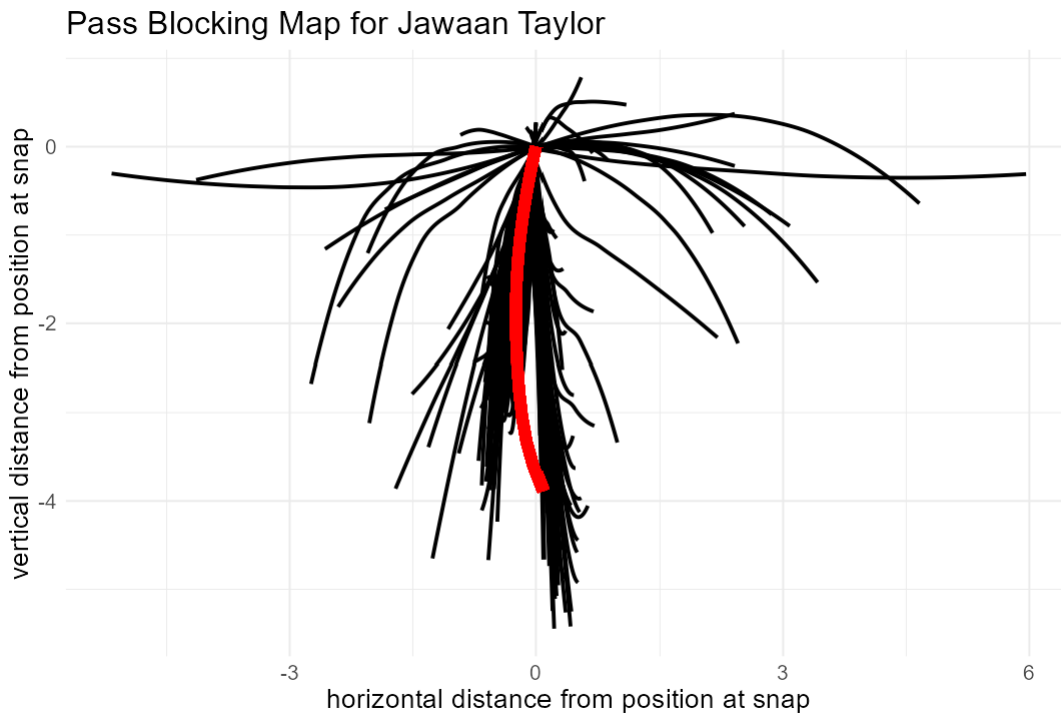


Figure 3.1.1: An example pass blocking path map for Jawaan Taylor of the Jacksonville Jaguars. The red curve is the smoothed average curve for Taylor.

3.2. Expected Depth

After building similarity measures for pass-blocking paths, we moved onto player evaluation by looking at the depth a player gets in pass protection. Former NFL All Pro offensive linemen, Mitchell Schwartz, when on the panel at the 2022 Sloan Sports Analytics Conference, discussed the process of getting depth as one of the most important aspects of pass protection, and credited his ability to meticulously get to his depth, without having the best athleticism in the world, as one of his edges as a pass protector [22].

Thus, we set out to find the “expected depth” for every pass protector on every play, given a set of play-level features about the play. The expectation is taken at 3.5 seconds after the ball was snapped, or when the pass was released, whichever is earlier. These features, which are listed in Table 2.1, include down and distance, field position, whether the game was at home or away, what was the time to throw, was the ball released when we measured depth, what was the maximum depth across the entire offensive line on the play, where the player was at the snap of the ball, what position they were playing, what was the score differential in the game, whether there was play action on the play, what pass coverage shell was being run by the defense, what was the dropback type, and what offensive formation was the team using on the play. Most of these are self-explanatory, but some are not. Whether the offense is the home or away team figures in when considering the lineman’s ability to get a jump on the snap, which is harder on the road in most cases. The maximum depth along the entire offensive line gives an idea as to what kind of pocket was intended by the coaching staff, while how many players that are in the box – along with the

offensive formation it's going up against, sets the stage for the pre-play expectations for how quickly or slowly the play is expected to develop.

We fit expected depth using an XGBoost model [21] in R [20]. The root mean square error of our model with all of the play-level data (down, distance, etc.) and none of the tracking data or charting data (e.g. dropback type, set type) was 1.38 yards, versus the full model, which was 1.10 yards.

Once we know the expected depth of an offensive lineman's pass set, we can then look at the rate at which linemen meet that depth, and if there's a signal in terms of that rate. We can also look at their win rate (as measured by PFF) when they do or do not reach expected depth, as well as the stability of win rates within each of those subsets.

3.3. Curve Clustering

In Sections 3.1 and 3.2 we took crude approaches to understanding offensive line movement in pass protection, while in this section we'll use the approach of [5] to cluster pass blocking curves using functional clustering.

Clustering is the most-common unsupervised learning technique, where objects are grouped together using a distance metric by similarity. In the method of [5], a predetermined number of clusters are used, in a process that requires some subject matter expertise to execute. In this paper, we balance the urge to generate enough clusters to span the set of all possible pass sets, while keeping few enough clusters so that there is a representable sample within each cluster from which we can study statistical properties. For all five offensive line positions, the most reasonable cluster structure included two clusters. For left and right tackles, along with left and right guards, there was a cluster for deep pass sets (see Figures 4.3.1 and 4.3.2) and another cluster for plays where the linemen spent more time near the line of scrimmage (e.g. quick game).

4. Results

4.1. Average Curves and Comparison Scores

Firstly, we took the average Frechet distance between a player's pass blocking paths and his smoothed average pass blocking path. We then look at these distances during the first four weeks of 2021 and the second four weeks of 2021 to get an idea of the stability of this measure. We compared this with the rate at which each player won their pass-rushing rep – per the charting at PFF.

A player's Frechet distance from their average pass blocking path was stable at a rate of $r = 0.36$ for all players with 100 pass-blocking snaps during both weeks 1-4 and 5-8 ($n = 83$, Figure 4.1.1). This compares favorably to the same measure of stability for win rate, which was $r = 0.32$ (Figure 4.1.1). Both values are pretty good for football and the timeframes considered, but it's encouraging that a player's average deviance from his average pass blocking path had better stability than his play-for-play performance, which we know to be stable over different time windows [6].

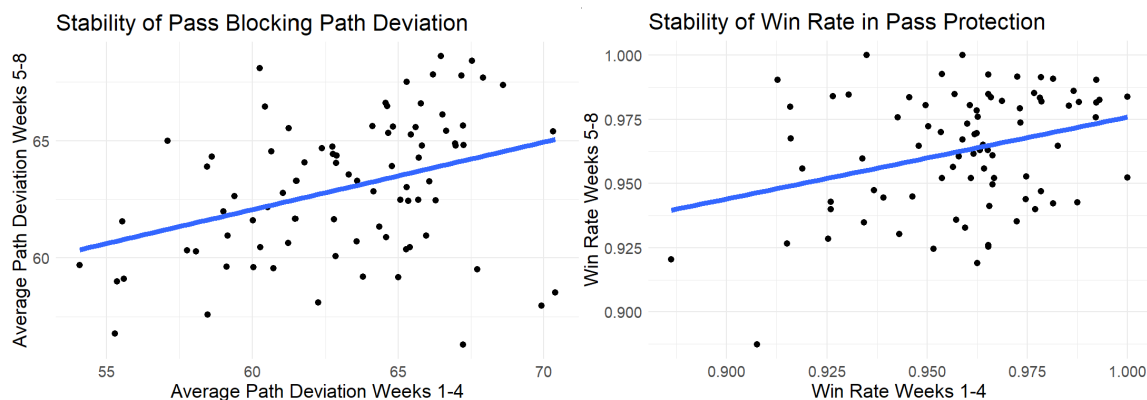


Figure 4.1.1: A comparison of pass blocking path deviation (left) and win rate (right) for the first four weeks of 2021 against the next four weeks.

Anecdotally, the similarity scores show promise in terms of grouping players. In 2021 the four all pro tackles were Trent Williams, Rashawn Slater, Tristan Wirfs, and Lane Johnson. We looked at their five closest comparisons (subsetting the data down to passes thrown between 2 and 3.5 seconds, and no play action, and players with 50 or more snaps in such situations). Williams has top five comps in Donovan Smith, Cameron Erving, Jake Matthews, Cam Robinson, and Yosuah Nijman. Slater has top five comps in Tyron Smith, Taylor Lewan, Andrew Whitworth, Elgton Jenkins, and Robinson, Wirfs has top five comps in Daryl Williams, Mike McGlinchey, Bobby Massie, Johnson, and Chukwuma Okorafor. Johnson has top five comps in Storm Norton, Wirfs, Massie, Kaleb McGary, and Rob Havenstein. Some overlap, some distinction in the way each wins. Slater, who was a rookie in 2021, had some of the movement characteristics of some great tackles in Smith, Lewan, and Whitworth, while Williams had a play style that some very athletic tackles also tried to pull off, but with much less success than he had.

It's even more compelling to go back and see some players that changed teams, and whether they fit into the mold of the player they replaced (with hopefully more production). The 2021 Kansas City Chiefs opened the season with Orlando Brown at left tackle, and Lucas Niang at right tackle. These two replaced Schwartz and Eric Fisher, after the two long-term tackles for the Chiefs were injured in 2020. In 2023 they acquired, through free agency, Donovan Smith to play left tackle and Jawaan Taylor to play right tackle. Smith's data from 2021 is with the Tampa Bay Buccaneers, while Taylor is from Jacksonville.

What's interesting to see is that in Smith's top 10 comps from a pass blocking path perspective, multiple players have Chiefs ties. His third-closest player, Cam Erving, was on the Chiefs from 2017-2019, including starting multiple games at left tackle for the 2019 Super Bowl-winning team. The man that Smith replaced, Orlando Brown, rang in as his seventh-closest comp among players with 50 or more reps, and fifth-closest among players with 100 or more reps. For Brown, Smith was his fifth-closest comp, behind Elgton Jenkins, Taylor Lewan, Andrew Whitworth, and Erving. He appears to have fit what the Chiefs wanted stylistically.

Taylor's closest comps included former Chiefs tackle Mike Remmers (fifth), and 2021 starting right tackle Niang (seventh). Niang's third-closest comp was Taylor, who was also Remmers' third-closest comp. As with Smith, the styles aligned for Taylor to join Kansas City's offensive line in 2023.

4.2. Expected Depth

The rate at which a player reached expected pass blocking depth from weeks 1-4 and weeks 5-8 were correlated at a rate of $r = 0.53$, which is larger than the metrics discussed in Section 4.1, and is about as good of a stability measure as you're going to get in football on a sample of that size.

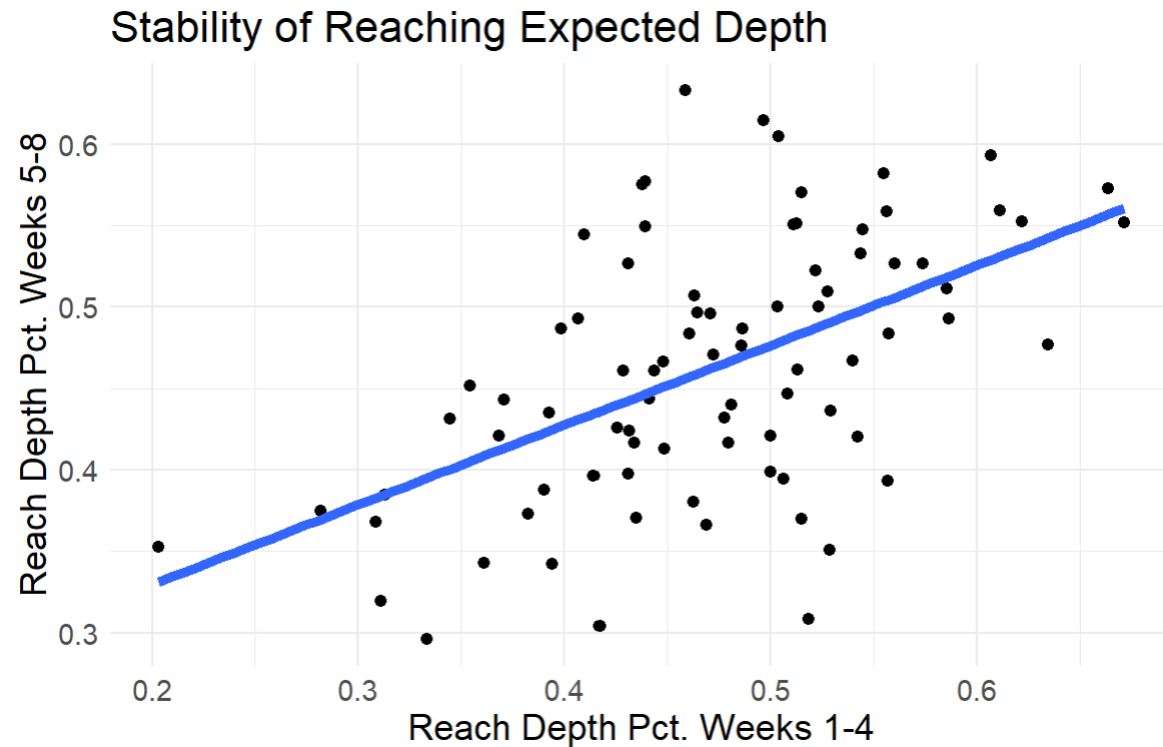


Figure 4.2.1: Percentage of pass blocking reps that reached expected depth, first four weeks of 2021 against the next four weeks of 2021.

Interestingly, whether a player reaches the expected depth on a play is positively related to whether he is beaten on a play ($p < 0.01$), while the degree to which the player deviates from his average pass blocking path is not. The former relationship persists even when considering other contextual features like down, distance, field position, etc., and even when one uses as a feature the propensity for the player to be beaten in the first place.

The in-season correlation between a player's win rate when reaching depth and not reaching depth is $r = 0.26$ (minimum 50 snaps of each, $n = 33$), which is lower than expected. There are only 26 players with 50 or more snaps in weeks 1-4 and 5-8 with reaching depth and 26 without reaching depth. It's interesting that for the former the correlation between the first four weeks and the second four was $r = 0.10$, while for the second it was $r = -0.12$.

Thus, the process of reaching expected depth is much more of a "how" variable than a "how good" variable at best and may be a detriment to efficient play as a pass protector. Table 4.2.1 gives the best left tackles in terms of winning when reaching expected depth, while 4.2.2 gives the same values for players who did not reach expected depth. These lists largely coincide with what offensive line experts believe about these players. For example, for Armstead the data matches how he plays. He doesn't reach depth all the time because he will aggressively pass set frequently. His

technique can almost never be copied because he crosses over his feet, but he repeatedly wins doing it. It's an approach that works for him, and it's easy to discern in the data.

Player	Team	Win Rate (Reached Depth)	Win Rate (Failed to Reach Depth)
Andre Dillard	PHI	98.8%	97.2%
Laremy Tunsil	HOU	98.3%	93.8%
Dion Dawkins	BUF	98.1%	95.8%
Charles Leno	WAS	97.5%	96.0%
Terron Armstead	NO	97.3%	100%
Tyron Smith	DAL	96.9%	98.4%

Table 4.2.1: Best left tackles in terms of win rate when they reached expected depth during the first eight weeks of the 2021 season.

Player	Team	Win Rate (Failed to Reach Depth)	Win Rate (Reached Depth)
Trent Williams	SF	100.0%	92.6%
Terron Armstead	NO	100.0%	97.3%
Andrew Thomas	NYG	98.9%	94.6%
Elgton Jenkins	GB	98.6%	91.4%
Tyron Smith	DAL	98.4%	96.9%
Jordan Mailata	PHI	98.3%	94.6%

Table 4.2.2: Best left tackles in terms of win rate when they failed to reach expected depth during the first eight weeks of the 2021 season.

4.3. Curve Clustering

Because of the distinct nature of each of the five offensive line positions in the NFL, we underwent separate clustering for each position, and hence the results and conclusions are going to be a little different for each.

Figure 4.3.1 shows the pass blocking paths for the two clusters for both left and right tackles. Notice that each of these positions includes two broad cluster types: one where the player takes a deep drop and one where he takes a shallower drop. This is largely what we expected, so this is a good result. Notice that for both tackle positions, there is a positive correlation between performance (as measured by win rate) from the first half of the data set (weeks 1-4) and the second (weeks 5-8, Table 4.3.1).

For interior offensive line positions, the guard clusters are like the tackle clusters, with one cluster being the plays in which the player opens outside and reaches depth, whereas the other is basically

that associated with quick game – eliciting shorter depths. For center, it's a bit different, where the two clusters correspond to whether the player opened left (possibly to help the left guard) or right (possibly to help the right guard, Figure 4.3.2). For centers, the stability between the first half of the data set and the second was substantial, but for guards it was not.

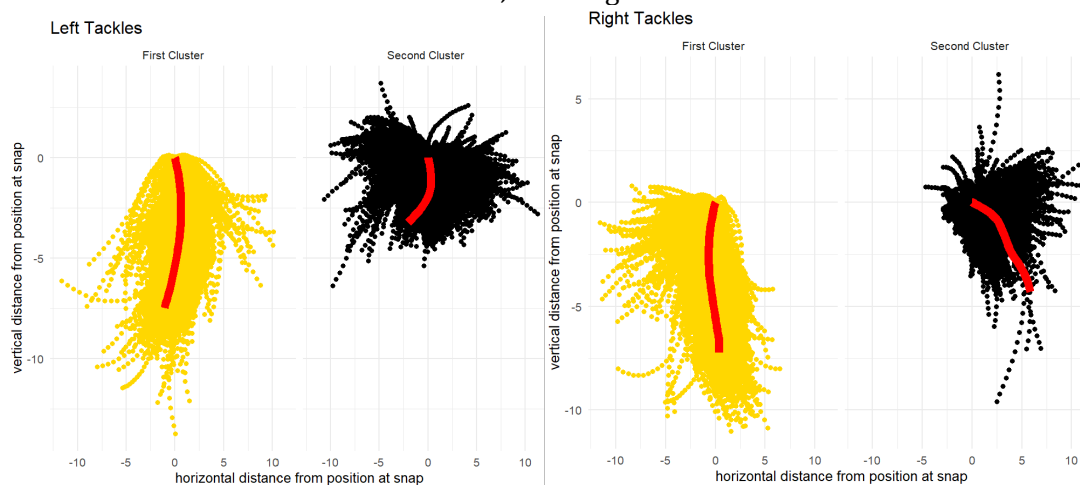


Figure 4.3.1: Curve clusters for left tackles (left) and right tackles (right). Red curve the average pass blocking path in each cluster for each position.

Position	Cluster	Win Rate Cor (Weeks 1-4/5-8)
Left Tackle	1 (Deep)	0.07 (n = 25)
Left Tackle	2 (Shallow)	-0.12 (n = 25)
Left Guard	1 (Deep)	-0.08 (n = 26)
Left Guard	2 (Shallow)	-0.04 (n = 25)
Center	1 (Left)	0.12 (n = 26)
Center	2 (Right)	0.42 (n = 26)
Right Guard	1 (Deep)	-0.39 (n = 26)
Right Guard	2 (Shallow)	-0.02 (n = 26)
Right Tackle	1 (Deep)	0.54 (n = 21)
Right Tackle	2 (Shallow)	0.09 (n = 16)

Table 4.3.1: Stabilities for win rates in each cluster for each position; first four weeks of 2021 against next four weeks of 2021.

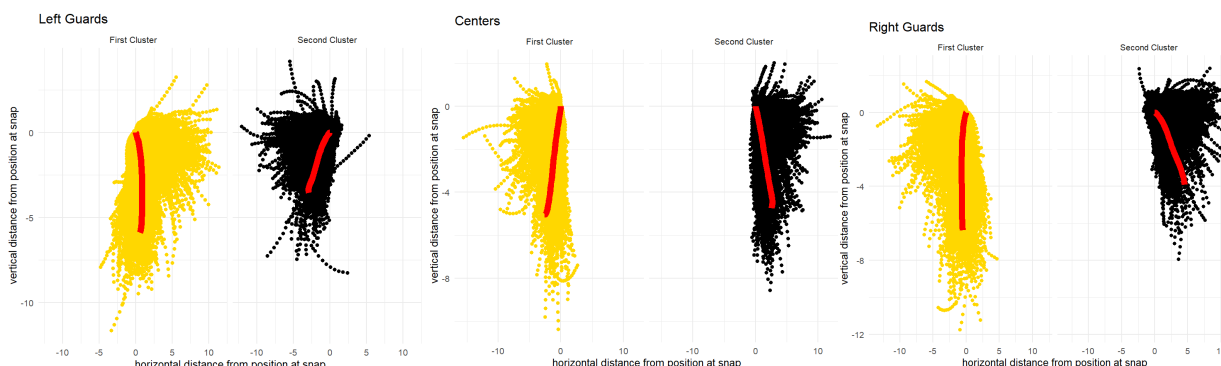


Figure 4.3.2: Curve clusters for left guards (left), centers (middle), and right tackles (right). Red curve the average pass blocking path in each cluster for each position.

While the sample sizes are not big enough to get a real definitive result regarding player ability within a cluster, it's instructive to look at some lists of players and their performances in these different clusters. Table 4.3.2 gives some of the highest-win-rate left tackles in the NFL and their win rates in different clusters.

Player	Team	Win Rate (Deep Cluster)	Win Rate (Shallow Cluster)
Terron Armstead	NO	97.2%	100.0%
Andre Dillard	PHI	99.0%	96.5%
Tyron Smith	DAL	96.2%	99.1%
Andrew Thomas	NYG	96.2%	98.3%
Dion Dawkins	BUF	96.7%	96.6%
Taylor Lewan	TEN	94.5%	98.1%
Charles Leno	WAS	97.6%	95.2%
Trent Williams	SF	92.8%	100.0%
Garett Bolles	DEN	96.3%	96.6%

Table 4.3.2: Top left tackles during the first eight weeks of the 2021 season in terms of overall win rate, separated by win rates in the two clusters. The overall win rate for the deep cluster is 93.3% and the shallow cluster 96.7%

Notice that, for the elite guys, all these numbers are high – individual failure in pass protection is relatively rare – but for all but Andre Dillard, Dion Dawkins, and Charles Leno, the win rate in the shallow cluster was higher than that in the deep cluster. This is true on average, as the win rate in the shallow cluster for left tackles is 96.7%, versus 93.3% for the deep cluster.

Different numbers of clusters were explored, and the results fruitful from a narrative/football standpoint, but not necessarily compelling from a data standpoint due to small sample sizes arising from just having eight weeks of publicly available data. For example, for centers if three clusters were the choice, then one cluster would be the center opening left, one the center opening right, and one like the shallow cluster for the other four positions. While not in the scope of this paper, due to public data availability, tests with more data, both at the NFL and NCAA (college) level have shown promise when going up from two to more than two clusters.

5. Use Cases

In this paper we've developed tools for better understanding offensive linemen in pass protection, and in so doing have provided a framework for using tracking data – supplemented with charting data – to create a rich context within which the NFL's most anonymous players can develop. In this section, we'll talk about some use cases that can serve as examples for football programs.

5.1. Efficient Film Study

One of the most easily seen value adds in analytics is the efficiency gain realized by having a richer data set to accompany film study. Being able to filter plays by the features like curve cluster,

expected depth, whether the player reached depth, etc. can aid in old fashioned player development – the interaction between position coach/coordinator and player. While companies like PFF, through their enterprise tools, have made substantial strides in this area, for offensive linemen in pass protection they are often graded pass-fail, which is not necessarily rich enough for some use cases. This work, along with its offshoots, help solve that problem.

5.2. Performance Monitoring

One obvious use of this analysis is to use it to further contextualize performance in pass protection. For example, one player may be struggling during the season at one cluster of pass-blocking paths, one can create a practice program to address these movements.

Furthermore, if certain pass-blocking clusters occur during certain plays – as in plays or concepts the team’s coaching staff calls, these plays can be called strategically to get an evaluation on a player one way or another. For example, in a preseason game or a practice, a coach can run a play that stresses the difficulties of a lineman so that they can monitor progress. In a regular season game, especially when the game is in the balance, a team can avoid plays that require the lineman to move along paths where they struggle.

5.3. Athleticism/Fitness Monitoring

Another way to use this information to benefit NFL teams is to use it to monitor the player’s athletic prowess over time, or after a long absence due to an injury, suspension, or benching. Decreasing depth in plays that require depth, for example, can be easily discerned using this data, and can be compared to similar declines in other players who are in similar schemes or are of similar ages.

5.4. Optimal Team Building

As we wrote about in [9] team success in pass blocking, where the probability of success can often be approximated as the product of the results of five (or more) independent battles between offensive and defensive linemen, is mathematically the problem of maximizing the product of these probabilities. This product is maximized when each of the elements are equal (assuming a fixed sum of the probabilities, which is a rough estimation of what the salary cap gives rise to in the NFL [4]).

In a broad sense, one can build an offensive line more efficiently by using this data to a) determine what are the most likely pass-blocking paths that the play calls on offense will induce, and select offensive linemen at the different positions that will optimize their win rates in such conditions, or b) given the offensive linemen on the team, alter the playbook in such a way that the product of the probabilities will be the highest (see below). Additionally, if the front office is trying to acquire a player, determining whether they fit into the existing group of linemen from an optimization perspective is within the potential of this work.

5.5. Optimal Game Planning

One thing that can be useful when it comes to development and team building along the offensive line is making sure that the plays called – which is private data to which only the team itself has

access, is such that each lineman is playing within a movement cluster of his that maximizes the offensive line's ability to play mistake free. This can also be paired with opposition scouting: As defensive linemen have their own curve clusters that they generate as pass rushers.

Thus, one can look at the curve clusters that are the most likely to arise when calling a play, both on offense and on defense, and estimate the likelihood that each lineman wins within that curve cluster and determine whether that is an optimal play from the perspective of the offensive line versus defensive line. Additionally, if a team is searching for an offensive lineman in free agency, or determining which backup to start if a starter goes down with an injury or suspension, finding the player that performs the best in the curve cluster that matches the one in which the rest of the offensive line thrives against a given opponent, is something in which this work can aid.

6. Discussion

In this paper we looked at tracking and charting data for offensive linemen in pass protection and devised a few starting points for building a tool to better develop these players at one of the most important positions in the sport – one that has been thought to be neglected since the most-recent collective bargaining agreement. We started with average paths, before deriving expected depth, and then curve clusters. Potential use cases were constructed using these tools, whose applications are only limited by data availability and imagination.

Additional modeling items that can be used to better tailor this system to the players are athleticism scores – which can be used to tune the depth expectation model for a player of interest, or as an extra feature in a clustering algorithm. This way, instead of looking at a player relative to league standards, one would be examining him relative to a standard his athleticism profile helped craft (his standard). Similarly, things like draft capital, which is a crude measure of how much the league thought of a player prior to him entering the league, could be beneficial as well. Additionally, there's nothing special about one set of (x,y) coordinates versus a set of five, six, or seven of them, and hence one can also extend this to classifying the movements of the whole group, examining the interplay between individual and group dynamics in search of efficiencies.

As tracking data continues to make a bigger impact on football – both at the NFL and NCAA level, traits-based analysis like that in this paper and in [8] can help in an automated way evaluators and coaches work together in an efficient manner to help grade and develop players into assets with positive expected value. In a salary cap league, where the margins are thin, this is of utmost importance.

For additional visualizations of our model, please see https://laplacefootball.shinyapps.io/sumer_ol. For our github repo, please see https://github.com/ericeager/ol_map.

References

[1] Awful Announcing (2017) Cris Collinsworth defends usage of Pro Football Focus data on NCA telecasts. <https://awfulannouncing.com/nfl/cris-collinsworth-pro-football-focus-defends-data-coaches.html>.

- [2] Baldwin, B. (2023) nfl4th <https://www.nfl4th.com/>
- [3] Baseball Savant (2023) 2023 Pitcher visualization report. https://baseballsavant.mlb.com/player-scroll?player_id=621244#pitchTypes
- [4] Bhatia, R., Kittaneh, F.(2000) Notes on matrix arithmetic-geometric mean inequalities. Linear Algebra and Its Applications. 308 (1-3): 203-211.
- [5] Chu, D., Reyers, M., Thomson, J., Wu, L.Y. (2020) Route identification in the National Football League. An application of model-based curve clustering using the EM algorithm. Journal of Quantitative Analysis in Sports. 16(2): 121-132.
- [6] Eager, E. (2020) True pass sets and their importance in player evaluation. <https://www.pff.com/news/nfl-true-pass-sets-and-their-importance-to-player-evaluation>
- [7] Eager, E. (2021) The value of perfectly blocked runs and passes. <https://www.pff.com/news/nfl-weak-links-offensive-line-2021>
- [8] Eager, E., Brown, B., Chahrouri, G., Riske, T., Spielberger, B., Sze Yui, L., Drapkin, Z., Seth, T., (2022) Using tracking and charting data to better evaluate NFL players: A review. MIT Sloan Sports Analytics Conference Research Paper. <https://www.sloansportsconference.com/research-papers/using-tracking-and-charting-data-to-better-evaluate-nfl-players-a-review>.
- [9] Eager, E. (2023) The weak-link nature of football. <https://sumersports.com/the-zone/the-weak-link-nature-of-football/>
- [10] Farrar (2015) Mythbusters: Are spread offenses the cause of offensive line decline? <https://www.si.com/nfl/2015/09/07/mythbusters-declining-offensive-lines-spread-offense>
- [11] Frechet, M. (1906) Sur quelques points du calcul fonctionnel. Rendiconti del Circolo Matematico di Palermo, 22: 1-74.
- [12] Goldsberry, K. (2019) Sprawlball: A visual tour of the new era of the NBA. Mariner Books.
- [13] Hayes, D. (2023) Amazon prime video execs prep data-rich update of 'Thursday Night Football', share plans for improving on weekly live streams that 'hadn't been done before'. <https://deadline.com/2023/09/thursday-night-football-amazon-prime-video-nfl-streaming-2023-1235545583/#!>
- [14] PFF (2018) How we grade offensive lineman. <https://www.pff.com/news/pro-how-we-grade-offensive-and-defensive-linemen>.
- [15] PFF (2023) <https://pff.com>.
- [16] NFL (2023) Next Gen Stats. <https://nextgenstats.nfl.com/>.
- [17] NFL (2023) NFL Operations. 2023 Big Data Bowl. <https://www.kaggle.com/competitions/nfl-big-data-bowl-2023/>.
- [18] OverTheCap (2023) Eric Berry. <https://overthecap.com/player/eric-berry/504>
- [19] OverTheCap (2023) Cam Newton. <https://overthecap.com/player/cam-newton/1147>
- [20] R Core Team (2023) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- [21] XGBoost (2021) <https://xgboost.readthedocs.io/en/latest/>
- [22] Sloan Sports Analytics Conference (2022) In the trenches of football analytics panel. <https://youtu.be/ZBa5bik6tHY>.
- [23] Yurko, R., Ventura, S.L., Horowitz, M. (2019) nflWAR: a reproducible method for offensive player evaluation in football. Journal of Quantitative Analysis in Sports. Journal of Quantitative Analysis in Sports. 15(3): 163-183.
- [24] Yurko, R., Matano, F., Richardson, L.F., Granered, N., Pospisil, T., Pelechris, K., Ventura, S.L. (2020) Going deep: models for continuous-time within-play valuation of game outcomes in American football with tracking data. Journal of Quantitative Analysis in Sports. 16(2): 163-182.