Unlocking the value added by ball carrier speed on rushing plays in the NFL

Matthew J. Worley[1]

[1]University of California, San Diego
La Jolla, CA 92093
mworles@gmail.com

Abstract

The NFL is now known as a "passing league" and many now question the value of talent at the running back position. However, the rushing game remains an integral component of the NFL offense, and player-tracking data presents a new opportunity to identify valuable aspects of ball-carrier performance that may contribute to rushing success. This project uses speed of ball-carriers and machine learning to predict the outcomes of rushing plays in the NFL, to evaluate the importance of ball carrier speed on rushing play outcomes and identify speed-related features of ball carriers most closely tied to play success.

1 Introduction

Description of the problem

In today's NFL, opinions on the value of running backs (RBs) and the impact of the running game have become more polarized than ever. Take, for example, the New York Giants' selection of Saquon Barkley with the 2nd overall pick in the 2018 NFL draft. Barkley was hyped up as a "generational talent" by his trainer, and his NFL combine performance metrics were off the charts. However, despite the acquisition of a "can't-miss" prospect, the Giants' decision was panned in the media [1] and questioned by many NFL team scouts [2], who pointed to the Giants' needs at other arguably more important positions (e.g., quarterback, edge rusher), the growing recognition that RBs have a short shelf-life and are largely "interchangeable". Barkley would seem to have silenced these critics by becoming just the 3rd rookie ever to eclipse 2,000 yards from scrimmage, but his team finished 5-11. In contrast, two of the four conference championship teams featured strong playoff performances from RBs who were either signed off the couch in week 16 (the Rams' C.J. Anderson) or buried on the pre-season depth chart (the Chiefs' Damien Williams). These strong performances by unheralded backups support the emerging, analytics-driven view that strong RB performances typically have more to do with the offensive system than the specific player who totes the rock [3]. The NFL is also now widely recognized as a "passing league", with team success much more reliant on elite QB performance and offensive passing efficiency than rushing efficiency [4].

And yet, other than the quarterback, no single offensive position touches the ball more than the RB, and in the 2018-2019 season rushing plays still comprised 42% of offensive plays [5]. Because rushing unquestionably remains an integral component of the NFL game, **if teams can identify specific characteristics of ball carriers that predict rushing success, this knowledge will help them create strategic advantages through better personnel decisions, offensive play design, or both**. This report describes an initial attempt to identify such characteristics using player tracking data from NFL games.

General approach

This project tests the assumption that **the trajectory of speed for the ball carrier is a relevant and determining factor for the success of rushing plays in the NFL**. Using both player-tracking data and external play-by-play data from *nflscrapR,* I isolate player-tracking data to rushing plays, and for each rushing play, I create a snap-to-whistle speed time-series for the ball carrier. I then use a feature extraction module to extract meaningful characteristics from the speed time series data, referred to as "speed features". I also use player tracking and play-by-play data to extract *within-play* and *situational* variables for a "baseline model" expected to relate to rushing success. To assess the unique and incremental effect of ball-carrier speed on success of rushing plays, I compare the "speed model" (the baseline model + speed features) to the baseline model on prediction accuracy rushing play success. I test several metrics of rushing success, including yards, yards after contact, success rate, and expected points added. Using supervised machine learning, the baseline model and speed model were trained to predict each metric, with models tuned and compared on cross-validation performance. Final generalization performance was assessed on held-out data. To derive meaningful interpretation and insight from the final speed model, the most "important" speed characteristics for rushing outcomes were identified and described with a combination of feature selection, dimension reduction, and data visualization.

2 Data

Rushing play selection

Given the 91 games of player-tracking data available, we consider all rushing plays as those marked as a rush attempt in *nflscrapR* play-by-play data. Runs where the ball carrier was tackled behind the line scrimmage were excluded from the dataset, with the assumption that these runs were unsuccessful due to factors outside the ball-carrier's control, such as poor blocking[1]. I also excluded rushes with obvious errors in tracking data, such as the ball carrier having zero speed well into the play or located in front of the ball at the snap. The final set of rushing plays included 3,533 distinct plays, at a mean of 39 per game.

Tracking data reduction

For all plays in the final set of rushing plays, I selected the raw player tracking data for further reduction and extraction of relevant variables. For each play the tracking "event" variable was used to limit the tracking data to include only the segments occurring from the snap of the ball to the end of the play, with the end marked by one of four events: tackle, out-of-bounds, touchdown, or fumble. Tracking data was further reduced down to the 2-dimensional location and speed of the ball carrier, and location of all defenders on the field for each play, to compute the play-level features and ball-carrier speed time series as described below.

Features for baseline model

For a baseline model of rushing play success, I attempt to identify situational and within-play factors that might impact the success of any given rushing play. The player-tracking and play-level data were used to compute the following features:

Early contact	A yes/no indicator of whether first instance of defender contact occurred behind the line of scrimmage

[1] While I believe the ability to create positive gains from poor blocking is definitely a valuable ball-carrier skill, these plays were seen as clear examples of situations where both ball-carrier speed and success of the play are interrupted prematurely.

Closest defender at line of scrimmage	Distance from ball-carrier to closest defender when ball-carrier reaches line of scrimmage. Included as an in-play indicator of the relative readiness of the defense to prevent ball-carrier from gaining positive yards
Defenders in the box	From play-level tracking data, described as “number of defenders in close proximity to the line of scrimmage”
Down	1st, 2nd, 3rd, or 4th down
Distance to go	Yards to gain for a first-down or touchdown
Offensive formation	Indicator of shotgun, single-back, or I-formation

Metrics of rushing success

From the player-tracking and play-by-play data, the following four metrics of rushing play success were obtained for each play:

Yards –Number of yards gained by the offense on the rush, from play-by-play data.

Yards after contact –Defined as the yards gained by the ball-carrier after the moment of first defender contact, as marked by the “first_contact” event code in player-tracking data.

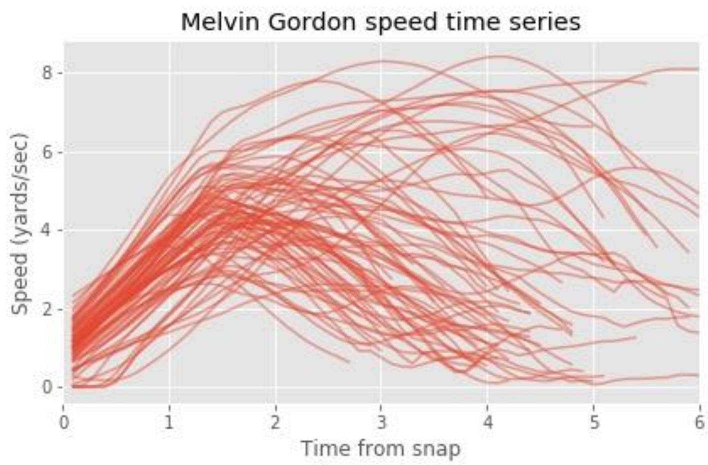
Expected points added (EPA) –The intent of EPA is a provide single metric describing the value of each individual play. In this context, EPA represents the expected points for the offensive possession created or lost by the rushing play. The EPA for each play is available in the *nflscrapR* data.

Success rate – A yes/no label of whether the play was successful or not, as calculated from the yards gained adjusting for the down and distance [6]. Success is marked by gaining 40% of distance-to-go on first down, gaining 60% of distance-to-go on second down, and gaining a first down on 3rd or 4th down. Computed from play-by-play data.

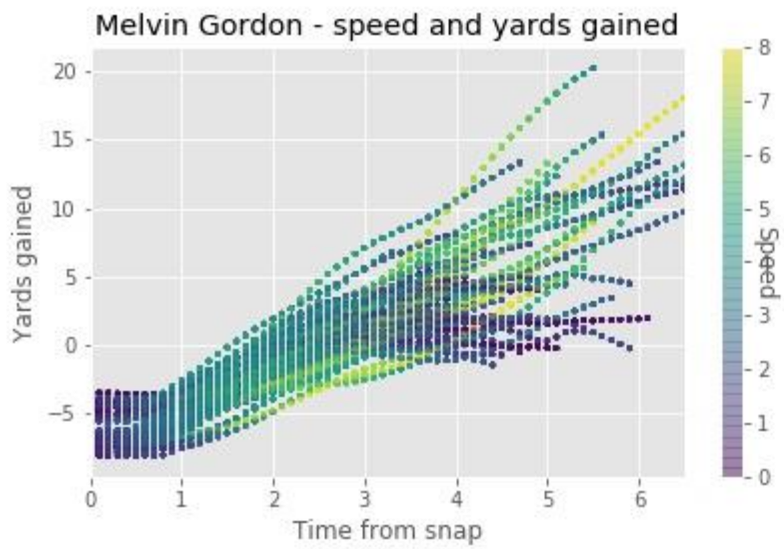
3 Using ball-carrier speed to predict rushing success

Time series of ball-carrier speeds

For each rushing play, consider that the ball carrier's speed can be viewed as a time-series, with time on the x axis, and speed on the y axis. The chart below shows all examples of speed time series in the dataset for Melvin Gordon.

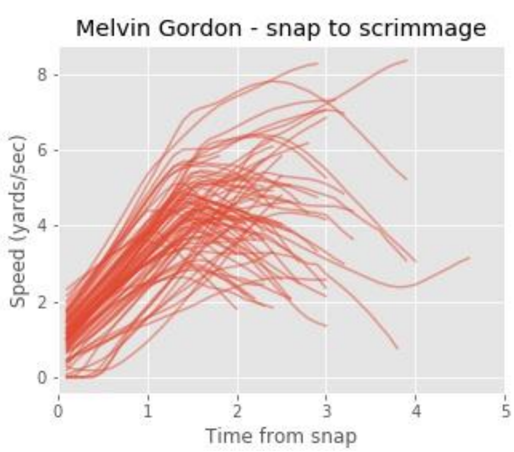


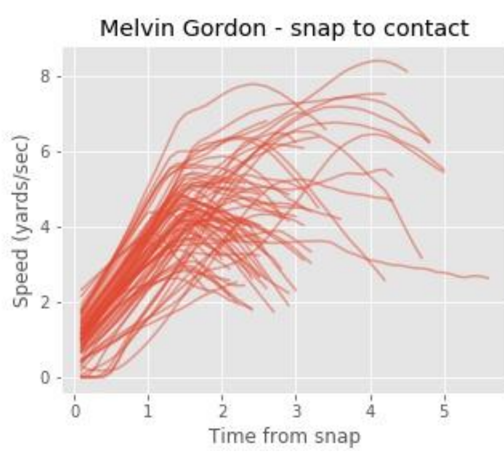
The ball-carrier time-series contain an average of 43 moments per play, while the metrics of rushing success are assessed at the play level. One major challenge of this project was translating the time-series into a reasonable set of characteristics that could be used to predict success at the play level. Another challenge was presented by the strong likelihood of bias between the speed time series and rushing outcomes. The concern of bias is present because when comparing successful rushing plays to unsuccessful ones, the successful rushing plays are likely to have at least some differences in speed trajectories that are more likely a *product* and not a *cause* of the play's success. Below I show Gordon's runs again, with the y axis displaying yards gained, data points colored by speed:



On Gordon's most successful runs, he has more data points, and he also achieves higher max speed, but his top speed usually is not achieved early in the run, meaning his top speed is not *generating* the play's success but instead occurs *after* the play has already become successful.

In order to have confidence my analysis would identify characteristics of ball-carrier speed responsible for *generating* the success of a run, separate sets of ball-carier speed time-series were created for two time segments that occur early in the time course of each play: snap to line of scrimmage (snap2LOS) and snap-to-first-contact (snap2contact). The restricted time series for Melvin Gordon are shown below:





Extraction of speed features from time series

To extract meaningful characteristics from each ball-carrier time-series, I used an open-source module for time-series feature extraction, *tsfresh* [7]. Using the dataset of time-series as input, the module extracts hundreds of features and assigns feature values to each time series example. These features were further cleaned for missing data and selected for relevance with each rushing success metric using the *tsfresh* filtering function, resulting in a variable number of speed features for each segment and target variable as outlined below:

Number of ball-carrier speed features, filtered by rushing success metric

	Yards	Yards after contact	EPA	Rush success
snap2LOS	113	96	115	109
snap2contact	123	99	117	115

Exploration of speed time series features

Use of supervised machine learning to predict rushing play success

Prior to model fitting, 20% of the rushing plays were held out, to provide a sample completely removed from the model training and tuning process for later testing of generalized classification/regression performance.

The remaining 80% of plays were used as a training set to train supervised learning algorithms to predict each rushing success metric, with all features scaled prior to model fitting. Regularized linear models were used for model training and testing, with logistic regression for classifying rushing success, and linear regression for predicting yards, yards after contact, and EPA. For scoring classification model predictions of "rush success", I used accuracy score. For scoring regression model prediction of the three numeric rushing metrics, I used root mean squared error and R^2 .

Models were initially trained using 5-fold cross-validation. Over 5 repetitions, the algorithm was trained on 4/5 of the training set, and used to predict rushing outcomes for the remaining 1/5, with scores averaged across all 5 repetitions. Using a random search for algorithm hyperparameter values, repeated cross-validation predictions were obtained, scored, and compared to identify optimal algorithm hyperparameter settings. This process was repeated for the baseline and speed model to achieve optimal predictions for each model type.

Comparison of baseline and speed models

A summary of scores for cross-validation and out-of-sample testing are shown below. For the RMSE score, values are positive with a minimum of 0 and highly dependent on the scale of the target variable (hence the lower values for EPA, then yards after contact, then yards). Lower values indicate model predictions are closer to the actual values. For R^2 and accuracy score, values range from 0 to 1 and larger values indicate more accurate predictions, but the scores are interpreted quite differently (accuracy score of .50 is usually comparable to random guessing, while R^2 of 0.50 roughly indicates half of the variance in the outcome measure is explained by the model). Because the three score types cannot be interpreted on the same scale, comparisons should only be made between models within each column.

Scores for baseline model and speed model predictions on four measures of rushing play success.

	Yards		Yards after contact		EPA		Rush success
	RMSE	R^2	RMSE	R^2	RMSE	R^2	Accuracy
Play segment							
snap2LOS							
Baseline model							
Cross-validation	27.74	0.13	5.65	0.09	0.46	0.12	0.74
Speed model							
Cross-validation	17.75	0.44	4.40	0.29	0.35	0.34	0.82
Held-out 20% set	16.47	0.49	4.21	0.32	0.33	0.39	0.84
snap2contact							
Baseline model							
Cross-validation	12.02	0.21	5.65	0.09	0.37	0.13	0.74
Speed model							
Cross-validation	7.71	0.49	4.02	0.35	0.32	0.27	0.79
Held-out 20% set	7.30	0.52	3.54	0.43	0.31	0.30	0.81

The speed model offered vastly improved predictions of rushing success over the baseline model for every metric of rushing success examined. In particular, the R^2 measure was quite low for the baseline model across most of the rushing metrics, and the speed model offered 2-3x better R^2 for some metrics for the snap2LOS and snap2contact time segments. Finally, when predicting outcomes for the 20% of held-out rushes, performance of the speed models was comparable to the performance on cross-validation.

Interpretation of model comparison

From the performance of supervised learning models trained on speed data from time segments occurring early in rushing plays, the results do suggest that **on a per-play basis, the speed characteristics of the ball carrier are a relevant and important factor in the success of individual rushing plays.** Predictive models using a host of speed features were much more accurate in predicting the success of rushing plays than baseline models that did not include speed features. The results also suggest that the "shape" of a player's speed, even when restricted to an early interval such as the snap to the line of scrimmage, may be an important factor in determining seemingly unrelated rushing outcomes, such as yards after contact and expected points added.

4 Identifying useful speed characteristics

This knowledge provides motivation to identify which specific facets of speed early in a rushing play bear the strongest relationship to successful rushing outcomes.

Data reduction and interpretation of speed characteristics

From over 90 descriptive characteristics of the ball-carrier's early speed trajectory included in the supervised learning models, the challenge lies in finding the best metrics of ball-carrier speed that can be either built or found in individual ball carriers, or created through play design.

In an initial attempt to uncover such insights, the full set of speed features from each trained model was filtered down using a low coefficient selection threshold of 1^{-5}.

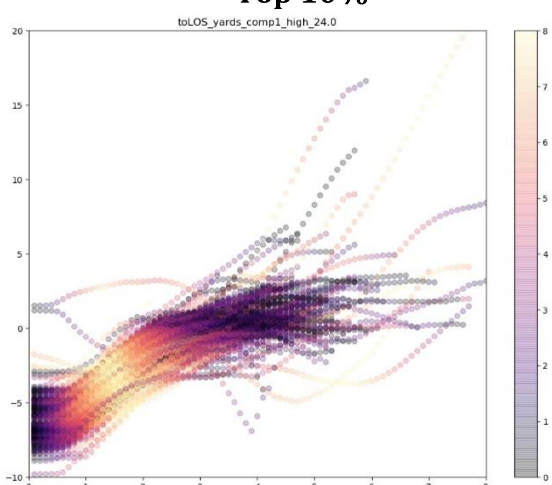
Essentially, this removed the speed features that had been reduced to irrelevance by regularization during training of the linear models. Then, principal components analysis was used to extract new summary characteristics that summarize and capture commonality across the full set of speed features, with three principal components extracted from the set of speed features used in each individual model.

Inspecting principal components to identify effective ball-carrier speed

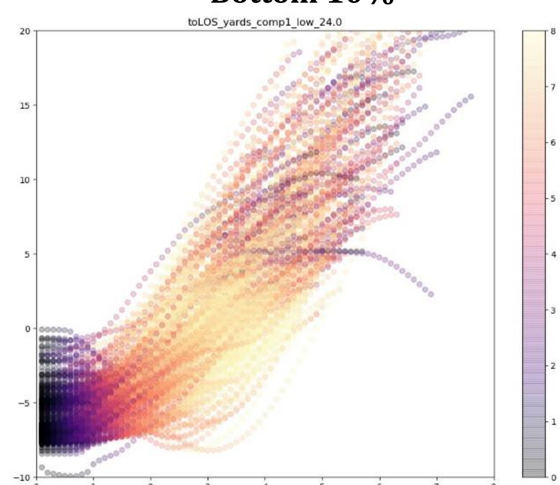
Below I display the first principal component extracted from the rush2LOS segment, using the unique set of speed features used to predict total yards gained. To illustrate the effect of individual differences in the component on rushing success, I display speed and yards gained for rushing plays when ball carriers were either in the bottom 10% or top 10% of component values. Although the chart displays the full course of the run to illustrate the relationship between the component and total yards gained, *the component was derived only from speed assessed from snap to scrimmage of each run.*

For this component, lower values are better, as ball-carriers with the bottom 10% of component values (displayed on the right) were much more successful. These plays were mostly characterized by slower buildup of speed that did not reach the top end until approaching or past the line of scrimmage, but top-end speed was maintained until at least 5 yards downfield. In contrast, rushing plays on the left reached top speed more rapidly, but maintenance of top speed was very short, with ball carriers slowing rapidly as they approached the line of scrimmage and few rushes continuing beyond 2-3 yards downfield. *Time on the x axis, yards gained on the y axis, with data points colored by speed.*

Top 10%



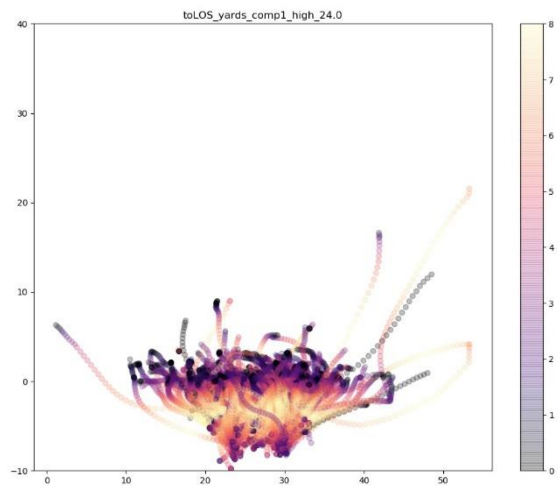
Bottom 10%



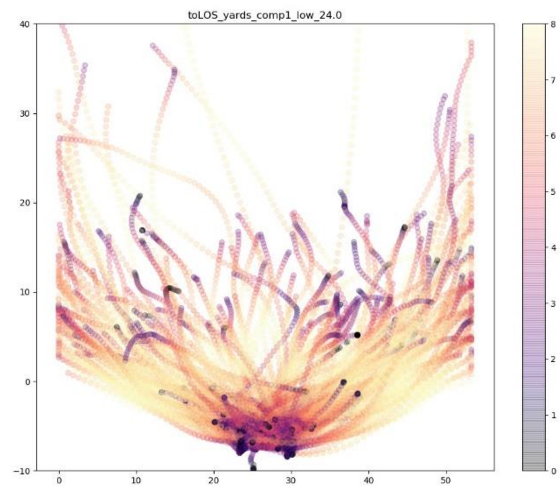
When viewed on the 2-dimensional space of the field, it becomes clear that the bottom 10% of this component mostly involve outside runs, while the top 10% in the component occur inside. This visualization suggests that although rushing plays on the right grid were characterized by slower onset of speed prior to the line of scrimmage, the open space afforded by rushing outside may have allowed for better maintenance of speed across the line of scrimmage and beyond.

Time on the x axis, yards gained on the y axis, with data points colored by speed.

Top 10%



Bottom 10%



5 Conclusions and implications

The results of this project suggest that characteristics of the trajectory of ball carrier speed early in rushing attempts are related to results of rushing attempts, across several distinct measures of rushing success. It is possible that factors other than the ball carrier himself are responsible for shaping the trajectory of the ball carrier's speed, such as the size of the gap or the duration of successful blocks. Future research needs to be conducted to identify these other situational factors and their direct impact on player speed. Additionally, although this project confirmed suspicions that the overall trajectory of ball carrier speed was related to rushing success, I have yet to drill down to identify clear and identifiable metrics that could be used by teams to make strategic decisions, such as creating plays or acquiring players with the most potent speed profiles. However, this research provides a solid starting point to identify how and why individual ball carriers still exert impact on the outcomes of the modern NFL rushing game.

References

[1] https://www.washingtonpost.com/news/fancy-stats/wp/2018/04/26/saquon-barkley-is-a-risk-at-the-top-of-the-2018-nfl-draft/, 2018

[2] http://www.espn.com/blog/new-york-giants/post/_/id/56798/scouts-giants-would-be-wise-not-to-select-saquon-barkley-at-no-2, 2018

[3] https://fivethirtyeight.com/features/todd-gurley-is-in-the-right-system-at-the-right-time/, 2018

[4] https://thepowerrank.com/2018/09/24/the-surprising-truth-about-passing-and-rushing-in-the-nfl/, 2018

[5] Data obtained from nflscrapR, January 2019.

[6] https://www.footballoutsiders.com/stat-analysis/2004/introducing-running-back-success-rate, 2018

[7] https://github.com/blue-yonder/tsfresh, 2019