2023 Risk Mitigation Model Guide

Table of Contents

Application Overview

The purpose of this application is to identify athletes that are at increased risk for hamstring injury due to workload and identify specific interventions that can be applied in the near term to reduce their risk. The risk of injury for any specific athlete on any specific day of training camp is low, but there are very large differences in the risk levels between athletes on a given day, even within the same position. To simplify the expression of risk for the NFL population, the risk probabilities are expressed in terms of **normalized risk factors**, which reflect the difference in risk between a specific athlete and the typical player both within the NFL population as a whole and the population of NFL athletes at their specific position. The definitions of the two normalized risk values are:

- **NFL Risk Factor:** An NFL athlete at the 50th percentile for daily risk of injury across the NFL preseason period has an NFL Risk Factor of 1.0. An athlete with an NFL Risk Factor of 1.5 has a 50% higher risk of injury than the player with an NFL Risk Factor of 1.0. An athlete with an NFL Risk Factor of 2.0 has double the risk of injury of an athlete with an NFL Risk factor of 1.0. NFL Risk Factors are capped at 5.0, so a risk factor of 5.0 reflects a player who has at least 5 times the risk of a typical player.

- **Positional Risk Factor:** An NFL CB at the 50th percentile for daily risk of injury across all cornerbacks during the NFL preseason period has a Positional Risk Factor of 1.0. A CB with a Positional Risk Factor of 1.5 has a 50% higher risk of injury than the CB with a Positional Risk Factor of 1.0. Positional Risk Factors are also capped at 5.0, so a Positional Risk Factor of 5.0 reflects a player who has at least 5 times the risk of a typical player at their position.

Due to the large difference in risk for injury between positions, positional risk factors can be very different than NFL risk factors. For example, a CB with a Positional Risk Factor of 1.0 has an NFL Risk Factor of approximately 1.7, reflecting that the average cornerback is at about 70% higher risk for hamstring injuries than the average NFL player. Conversely, an OL player with a Positional Risk Factor of 1.0 has an NFL Risk Factor of only 0.5, reflecting that lineman have a greatly reduced risk of hamstring injury.

NFL Risk Factors provide the ability to quickly identify the athletes most at risk across your team (largely supported with the Athlete Roster Table workflow described below), regardless of position. Positional Risk Factors provide

the ability to quickly identify the athletes on your roster at each position that are at increased risk relative to their peers at the same position (largely supported with the Positional Risk Workflow), facilitating discussions with positional coaches.

The app operates in two different "modes" depending on the time within the season:

- Preseason mode: In this mode, an athlete's risk factor for a given day is calculated to express the athlete's risk **tomorrow,** given the training load they have experienced during the season **up to and including today.**

- Regular Season mode: In this mode, an athlete's risk factor for a given day is calculated to express the athlete's risk for **the following seven (7) days**, given their load up and to the current day.

Guide continues on next page…

Athlete Roster Table Workflow

The Athlete Roster Table workflow is designed to quickly review the status of all athletes on the roster for a particular day. Every athlete recorded on your active roster (the official NFL roster as populated in the NFL databases) will appear in the table. In the situation that the NFL's roster database was not updated in a timely fashion for the day, the previous day's roster will be used to populate this table until your team updates the status (thus there are cases where athletes that have participated may not appear). The diagram below describes a workflow for quickly identifying athletes most at risk and recommended interventions for reducing their risk. If the performance data for your team has not yet been populated into the database (this should occur less than 1 hour after you sync your local systems with your performance vendor), then the daily load values will be shown as 0.

1. Select a date to review records for athletes on your team. If you select a date that is not populated in the database (a date in the future or a date in the off-season), no table will be displayed. The default date is today.

2. If needed, filter by position or NFL Risk Percentile. Your applied filters will be stored in your web browser and automatically applied the next time you log in from the same computer.

3. If needed, search for a specific athlete by name or roster number (roster number search to be activated in regular season).

4. Click here to customize the columns shown in the table. Your customization will be stored in your web browser and automatically applied each time you log in.

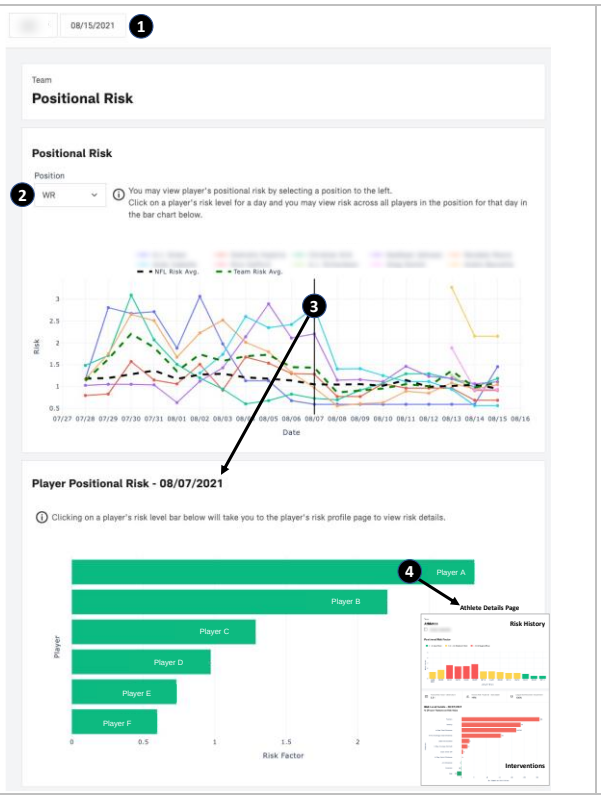
5. Export the current table view to Excel.

6. Sort the athletes by any column. The default sort is in decreasing order of NFL Risk Factor.

7. Clicking on the "Select" column will open the Athlete Details page for the selected athlete. We recommend opening with a Control-Click (Command-Click on Mac) to open the Athlete Details page in a new browser window to facilitate quickly reviewing multiple athletes.

Positional Risk Workflow

The Positional Risk page provides the opportunity to review the risk history for all athletes at a given position. This facilitates quickly identifying athletes at each position that are at increased risk as well as identifying athletes whose risk is trending up.



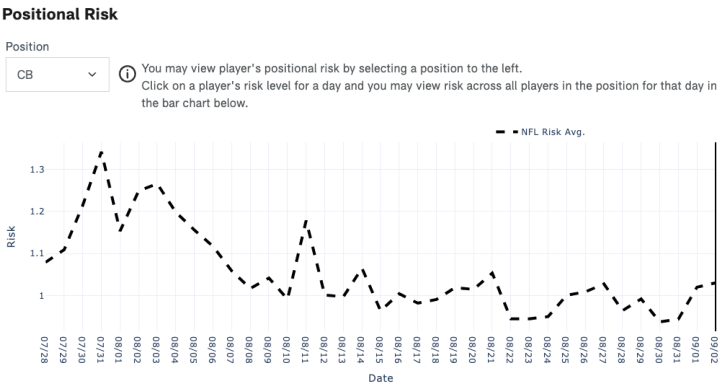
1. The date selector at the top controls the temporal window. Positional risk histories up to the selected date are displayed.

2. Select a position to view. All athletes on the roster at that position are displayed.

3. Clicking on an athlete's risk value for a given day updates the barplot at the bottom to reflect the relative risk between athletes on that specific date. The default date when the page loads is the date selected at the top of the page.

4. Clicking on the bar depicting an athlete's daily Positional Risk Factor loads the Athlete Details page for the athlete (see the **Athlete Details Page** section below).

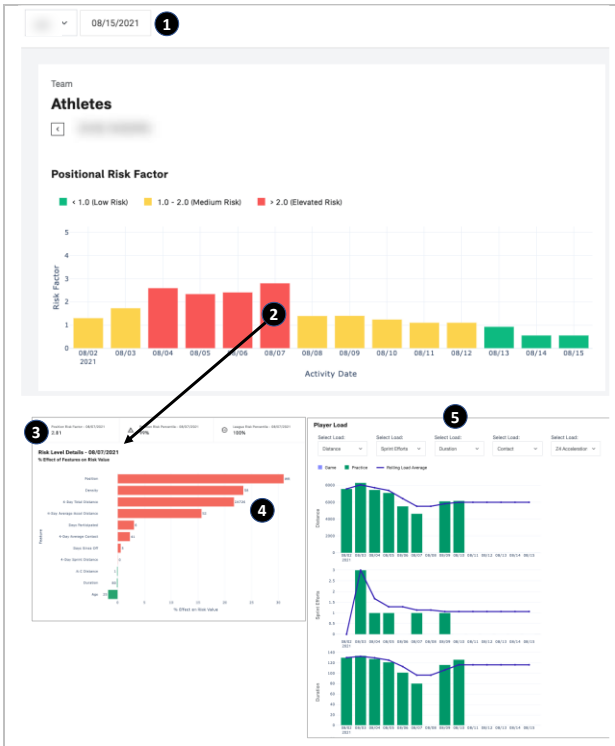
It is important to note that a positional risk value of 1.0 denotes the average risk value over the entire Preseason. As the graphic below illustrates, early in training camp the risk is much higher at most positions (note the changes in risk after acclimation, during the first week of contact integration, and right around the first preseason game). For this reason, the positional risk plots provide the average positional risk value for the NFL position cohort for each day (this is calculated for each team based on days since vet report date to account for varying team schedules). Athletes above this line are at higher than expected risk for the time period. Those that are below are at lower risk than average for their position at that time period.



Athlete Details Page

The Athlete Details page provides a detailed overview of a given athlete's load and risk history. Specifically, it provides the following information:

- The daily risk profile for an athlete for the selected period.
- Detailed explanation for the load features that contribute to the athlete's risk value for the day, providing opportunities for intervention to reduce risk.
- Load history for the athlete for the selected period.



1. The athlete's history is displayed up to the date selected. If a date is selected for which there is no time history, available dates will be plotted (and the risk level details plot will be blank since there is no data for that date).

2. Clicking on a bar in the Positional Risk Factor history chart will update the date of the Risk Level Details Plot (item 3). The Risk Level Details chart defaults to the date in the date selector (unless there is no data for that day).

3. The Risk Level Details chart provides the position risk factor, position risk percentile, and the NFL risk percentile for the athlete-day.

4. The Risk Level Details chart provides opportunities for interventions by plotting the % importance that each model feature contributes to the difference between this athlete and the average NFL athlete. See the user guide section

5.

6. **Explanation of Risk Mitigation *Plots*** for more details

7. Player load statistics for the selected date range are also included. Up to five different player load features can be plotted simultaneously. Your customization will be stored in your web browser and automatically applied each time you log in. Without customization, the load statistics default to those used by the model.

Guide continues on next page…

Risk Mitigation Model Mode

This section describes the two modes that the risk model operates in:

- **Preseason Mode:** The app operates in this mode from the start of training camp through the final week of preseason games. While in this mode the application predicts the athlete's probability of injury **tomorrow**. Due to the consistent loading of players throughout the preseason this target is suitable for managing an athlete's load.

- **Regular Season Mode:** In the week prior to the first regular season games, the app will switch to operating in this mode and predictions will be made for an athlete's probability of injury for the **following seven (7) days.** This change is necessary due to the cyclical nature of the regular season. In a typical week where a team has a game the following Sunday, this allows users to see an athlete's risk through that game.

Risk Mitigation Model Overview

This section describes the structure of the machine learning model, the feature selection process used to populate the model, a brief visual summary of the process by which probabilities are assigned, and a guide on interpreting the risk mitigation (daily risk details) plots.

Daily risk estimates are calculated each day using a machine learning model that predicts an athlete's probability of injury tomorrow (during preseason), **and for the following seven (7) days** (during the regular season), **given the training load they have experienced during the season up to and including today.** An athlete's load history starts from the 1st day of preseason early report or the athlete's first day of participation. The selected machine learning model is an extension of a decision tree model that employs numerous decision trees to handle the highly non-linear and complex task of assigning probabilities to these rare events. The ***Pre-Season Feature Effects Plots*** section below provides a visual explanation that summarizes how the model assigns probabilities for athlete risk each day. These probabilities are converted into normalized risk factors for ease of interpretation as discussed in the **Application Overview** section above.

Model Feature Selection

Model Feature Selection Summary (Selected Feature Types Indicated in Gray)

Type	Base Feature	Temporal Window				
		Fixed	Cumulative	Daily	4 Day Window	Acute:Chronic
Non-Modifiable	Position	■				
Non-Modifiable	Age	■				
Schedule	Training Days		■			
Schedule	Sequential Days		■			
Duration	Duration			■		
Intensity	Density (y/m)			■		
Speed	Speed Zones				■	
Contact	Contact				■	
Acceleration	Acceleration Zones				■	
Distance	Distance				■	■
Deceleration	Deceleration Zones					

To select the features used in the predictive model, we first consulted a focus group of NFL athletic trainers, sports scientists, and strength & conditioning coaches to identify a set of load features that teams routinely track on a daily basis. Critically, we asked them to identify load features they tracked daily for which, when flagged for adding risk, there was a known intervention that could be employed to reduce the athlete's risk. Those features were all of a type

listed in the Base Feature column of the table above. Note that within a base feature description, there were potentially multiple metrics considered. For example, for Speed Zones, teams may track Zone 4 (High Speed Running), Zone 5 (Sprints), or the combination of both zones. They may also track either sprint efforts or sprint distance. All of these various metrics were considered for inclusion in the model.

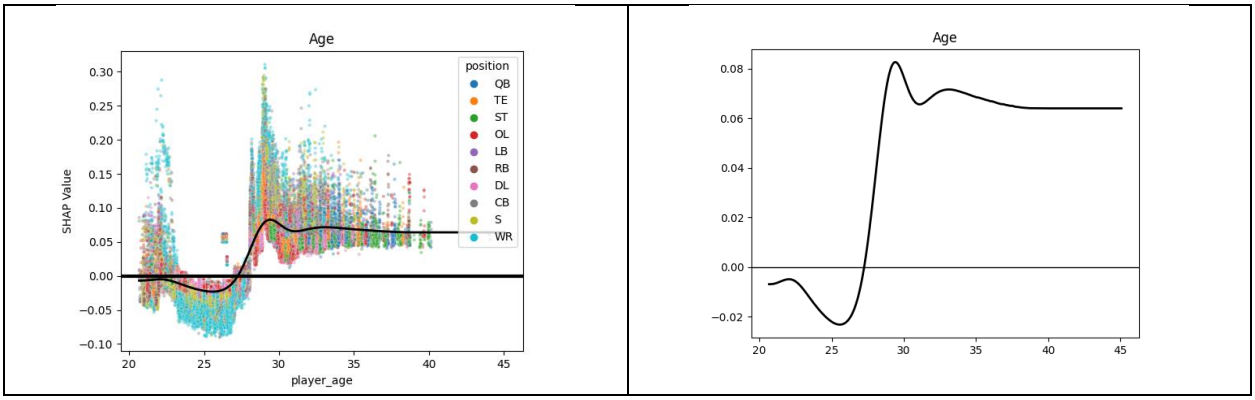
Once a comprehensive list of base features was assembled, we applied those features over various types of temporal windows as depicted in the table above. The Daily features represent the acute load for the day's training period. The 4 Day Window features included the sum and average (when training) for each metric over the immediately preceding 4-day training block. The Acute:Chronic features represent the ratio of the athlete's acute load to the athlete's average daily load when training (computed using an exponentially decaying average so that changes in athlete capacity over time due to adaptation were considered). We then ran a machine learning feature selection process to identify an ideal combination of features for daily intervention. The selected feature types are depicted in gray in the chart above. Definitions for each of the selected features is available in the **Feature and Load Statistic Definitions** section below.

When in regular season mode, the model leverages additional long range acute to chronic features. These features provide a ratio of the athlete's prior 7-day average load relative to the prior 28 days. These features provide a view into if an athlete is being worked above or below their typical loading.

Pre-Season Feature Effects Plots

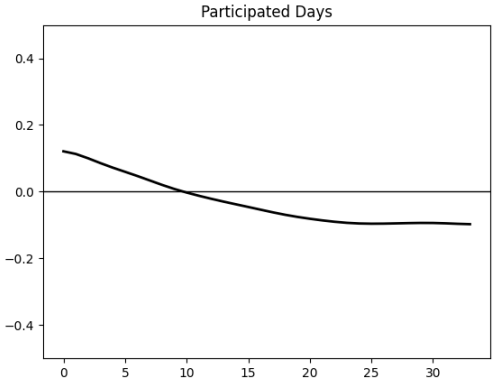
This section provides a visual explanation for how the machine learning model assigns risk. The selected machine learning model is designed to perform well for decisions involving features that have complex interactions. It is important to note that model does not consider the features for a given athlete by independently adding risk feature by feature as in a regression model, but rather assigns risk by simultaneously considering all of the features for a given athlete using a large ensemble of decision trees. However, it is possible to use visualizations of model assignments to understand the decision-making process.

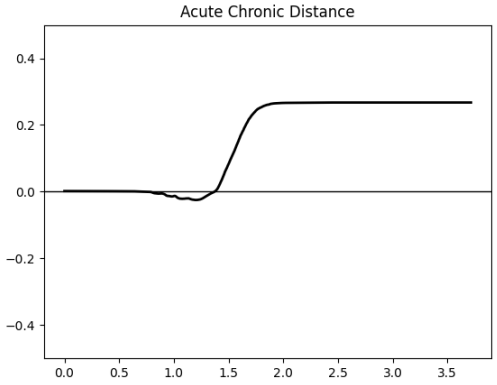
Feature Effects Plot for Age for NFL Cohort



Consider the example above, in which the panel on the left plots the effect assigned to every athlete in the NFL based on their age over multiple seasons (each point on the graph represents an athlete-day). Note that there is variation in the assigned risk due to interaction with other features. The panel on the right depicts the smoothed average value at each observed age (the y axis depicts the change in log-odds). When this line is above 0, it indicates that the model (on average) assigns higher risk to the athlete, and when below 0 the model (on average) assigns reduced risk at that age. The ability of the model to handle non-linear effects is clearly illustrated above. The interpretation of the effect from these plots is that athletes under the age of 24 (i.e., rookies etc.) are at the highest risk. Athletes with experience between the ages of 24 and 28 are at reduced risk. Athletes above the age of 28 are also at increased risk.

Average Feature Effects for the NFL Cohort for Selected Features



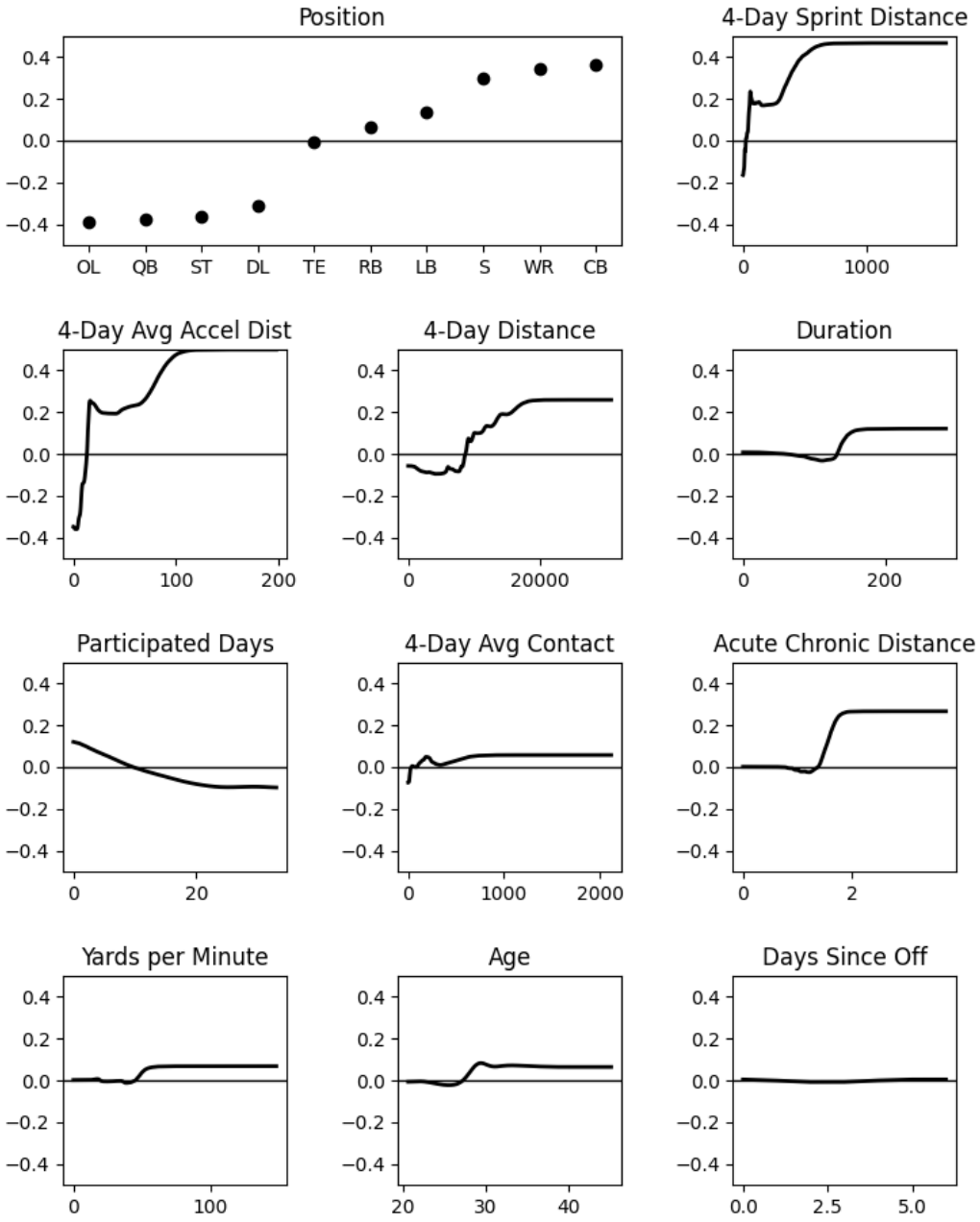


Two more example plots of feature effects are provided above, plotted on the same scale. At high values, these features contribute more to a player's risk value than player age (note the y-axis scale difference). It can be seen that the maximum effect of these features is similar, and the overall non-linear pattern is evident. The interpretation of these plots is that athletes are at increased risk during the 1st 10 days of their participation (for most athletes this encompasses the acclimation period), with risk decreasing steadily by day. The pattern for Acute:Chronic Distance is somewhat different. Risk remains unchanged for Acute:Chronic Distance values up to 1. Athlete-days with A:C Distance ratios between 1 and 1.25 are at slightly reduced and then risk begins to climb as an athlete's A:C Distance ratio increases.

The chart above depicts the average effect of all of the model features over the domain of values in the NFL cohort plotted on the same scale. Features at the top of the chart are more influential in affecting risk with position and acceleration features providing the most influence. The features at the bottom of the chart may contribute little to an athlete's risk *on average*, but remain in the model because they improve the model's performance by contributing meaningfully in certain situations that improve overall model performance.

Guide continues on next page…

Average Feature Effects for the NFL Cohort



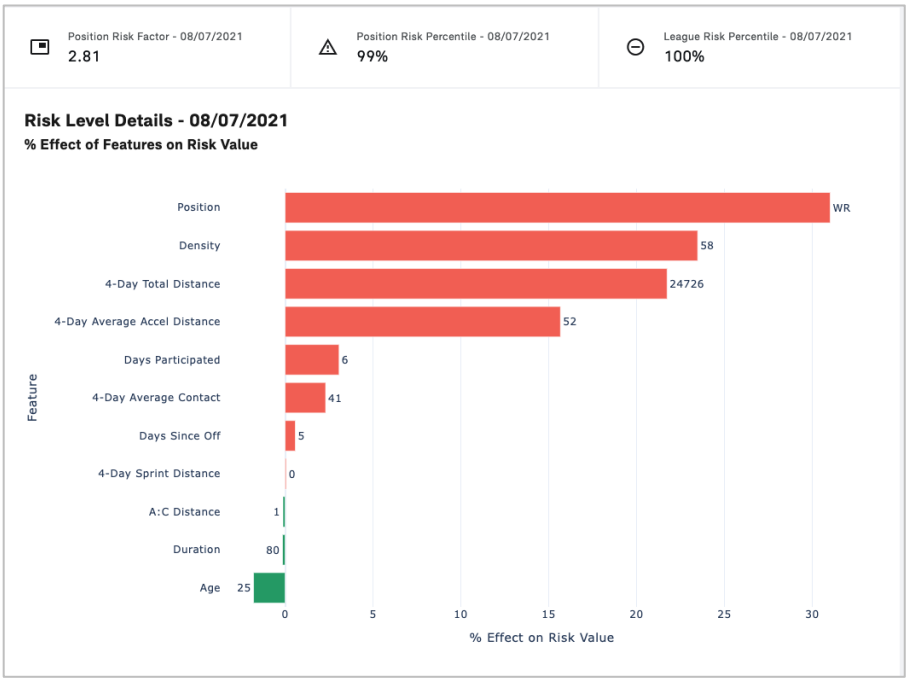
Feature Importance and Feature Effects Plots for the Regular Season model are available as an Appendix at the end of this guide.

Explanation of Risk Mitigation Plots

The values in the risk mitigation plots (i.e., Risk Level Details) represent the risk effects discussed in the last section applied to a specific athlete-day. The risk effects are plotted as a % of the overall difference between this athlete-day and the 50th percentile NFL athlete (an athlete with an NFL Risk Factor of 1.0). Bars in red indicate contribution to increased risk. Bars in green indicate contributions towards reduced risk. For an athlete with an NFL Risk Factor of exactly 1.0 (note that Risk Factors are rounded to 1 decimal place), the sum of the red and green bars will add up to 0, representing an athlete that is exactly average for the NFL cohort. It is important to note that the bars represent the relative contribution to that player's deviation from average risk. Therefore, a large value does not necessarily indicate that the feature significantly impacts the player's overall risk.

Large effects that add risk provide opportunities for intervention. The athlete depicted below has a high risk factor due to large values for density (yards per minute), 4 day total distance, and 4 day average acceleration. This indicates that the model believes this athlete is in an overloaded state and could benefit from reduced overall distance and acceleration in the next few practices.

Example Risk Mitigation Plot for an NFL Athlete at Increased Risk



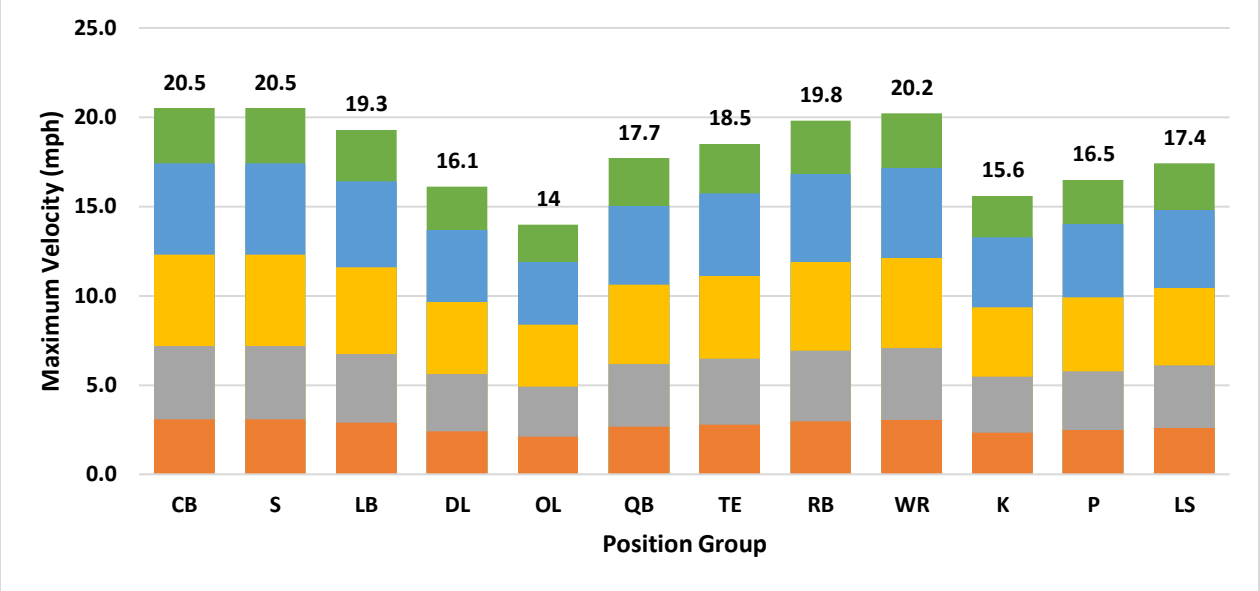
Guide continues on next page…

Feature and Load Statistic Definitions

The following table documents definitions for all of the load statistics and model features contained in the application. The model column identifies those features used in the risk model (either Pre-Season or Regular Season). The load plot column denotes load statistics available for daily plotting in the Athlete Details page. All features and load statistics are available for display in the Athlete Roster table.

Table	Load Plot	Pre Model	Reg Model	Feature Name	Definition
				Position	Athlete's assigned position in NFL database
				Age	Athlete age in years
				Days Participated	The cumulative number of days of participation by athlete in current season.
				Days Since Off	The cumulative number of days of participation in a row.
				4-day (21 Day) Total Distance	Total distance travelled over the prior 4 (21) days.
				4 Day Sprint Distance	Sum of sprint distance over last 4 days.
				Duration	Total duration of tracking session
				21 Day Total Duration	Total duration of tracking sessions over 21 day window
				7:28 A:C Duration	The ratio of the average daily duration over the last 7 days over the average daily duration over the last 28 days.
				Density	Total distance divided by total duration
				4-Day Average Contact	The average contact load over the last 4 days.
				4-Day Average Accel Distance	Average distance travelled while accelerating between 3.5 and 5.5 m/s/s calculated over the prior 4 days
				Distance	Total distance travelled during a tracking session
				A:C Total Distance	The ratio of today's load to the players average daily load
				21 Day Total Distance	The sum of distance over last 21 day window.
				7:28 A:C Distance	The ratio of the average daily distance over last 7 days over average daily distance over last 28 days.
				Max Speed	Maximum velocity achieved during a tracking session
				21 Day Total Sprint Distance	Sum of distance in Zone 5 over previous 21 days.
				21 Day Total High Speed Distance	Sum of distance in Zone 4 and Zone 5 over previous 21 days.
				High-Speed Distance	Distance travelled at speeds > 60% positional maximum velocity
				7:28 A:C High Speed Distance	The ratio of the average daily high speed distance over last 7 days over average daily high speed distance over last 28 days.
				Sprint Distance	Distance travelled at speeds > 85% positional maximum velocity
				Sprint Efforts	Count of times a player reaches speeds above 85% positional maximum velocity and sustains that speed for 0.5s
				Contact	Total contact during a tracking session
				21 Day Total Contact (11 on 11)	Total contact over 21 day window
				7:28 A:C Contact	The ratio of the average daily contact over previous 7 days over the average daily contact over previous 28 days
				Distance Under Contact	Distance travelled while under contact
				7:28 A:C Distance Under Contact	The ratio of the average daily distance under contact over previous 7 days over the average daily value over previous 28 days
				Time in Contact	Total time spent while under contact
				Contact (11on11)	The contact load from 11on 11 periods (only functions if periods tagged)
				Distance under Contact (11on11)	Distance travelled during 11on11 periods while under contact
				Time in Contact (11on11)	Total time spent in 11on11 periods while under contact
				Z4 Acceleration Distance	Distance travelled while accelerating between 3.5 and 5.5 m/s/s
				Z5 Acceleration Distance	Distance travelled while accelerating above 5.5 m/s/s
				Acceleration Distance	Distance travelled while accelerating above 3.5 m/s/s
				Deceleration Distance	Distance travelled while decelerating below -3.5 m/s/s
				21 Day Z5 Deceleration Distance	Total Z5 deceleration distance over 21 day window
				Acceleration Efforts	Count of times a player accelerates above 3.5 m/s/s and maintains that acceleration for 0.3s
				Deceleration Efforts	Count of times a player decelerates below -3.5 m/s/s and maintains that deceleration for 0.3s
				Lateral Accel Efforts	Count of times a player accelerates to change direction above 3.5 m/s/s and maintains that acceleration for 0.3s. This is the component of total acceleration used to change directions and not change speed.
				Lateral Accel Distance	Distance travelled while a player accelerates to change direction above 3.5 m/s/s. This is the component of total acceleration used to change directions and not change speed.

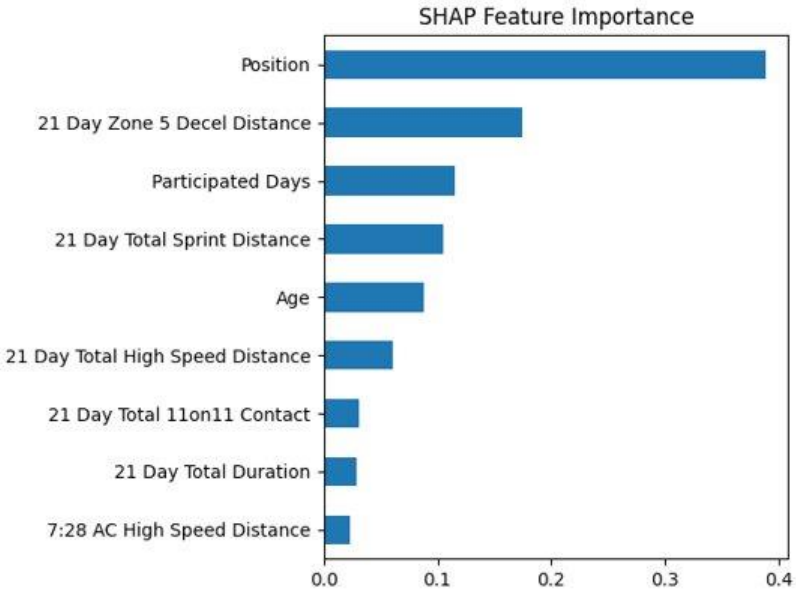
Positional Velocity Bands



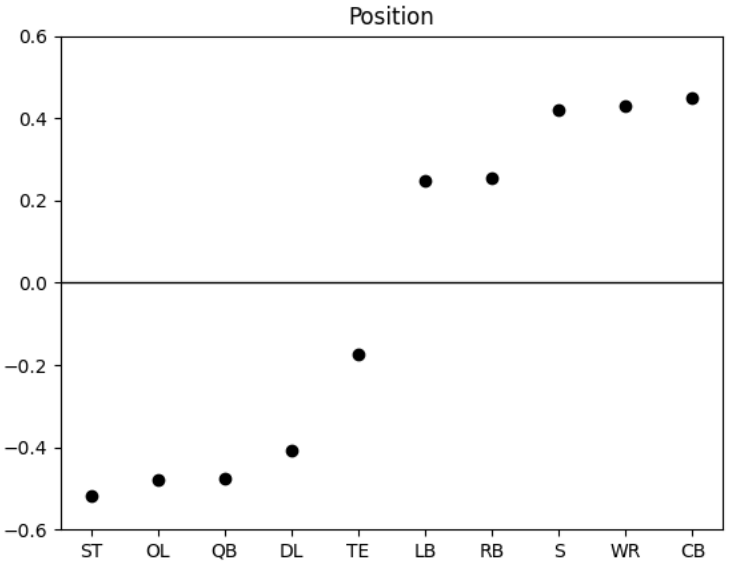
Guide continues on next page…

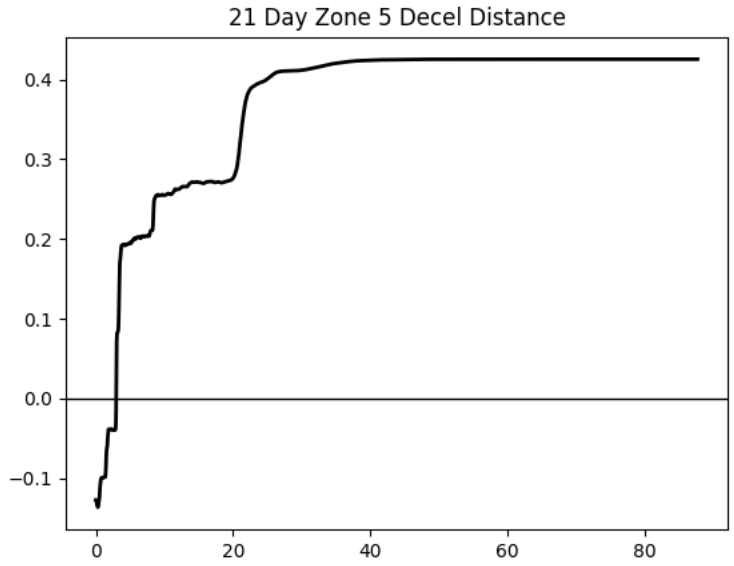
Feature Importance and Feature Effects Plots for Regular Season Model

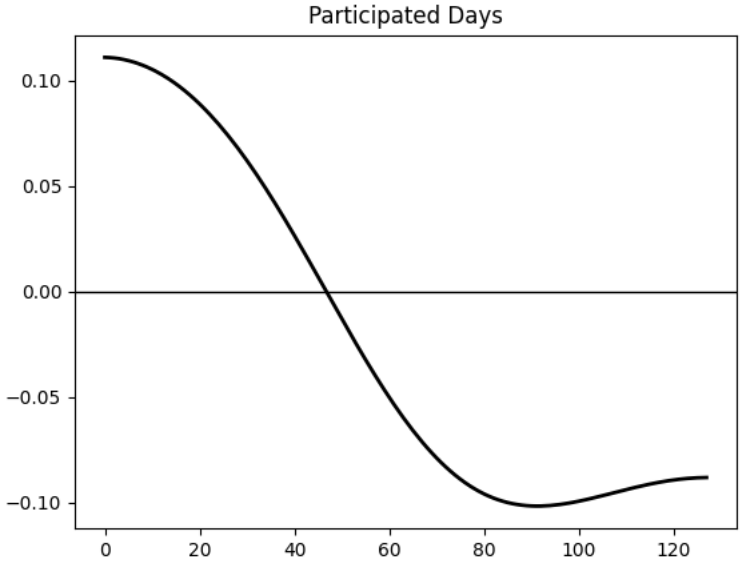
The plot below illustrates the relative importance of the features in the regular season model when applied to the entire NFL population, with features at the top of the plot contributing more to the risk scores of athletes than those at the bottom.

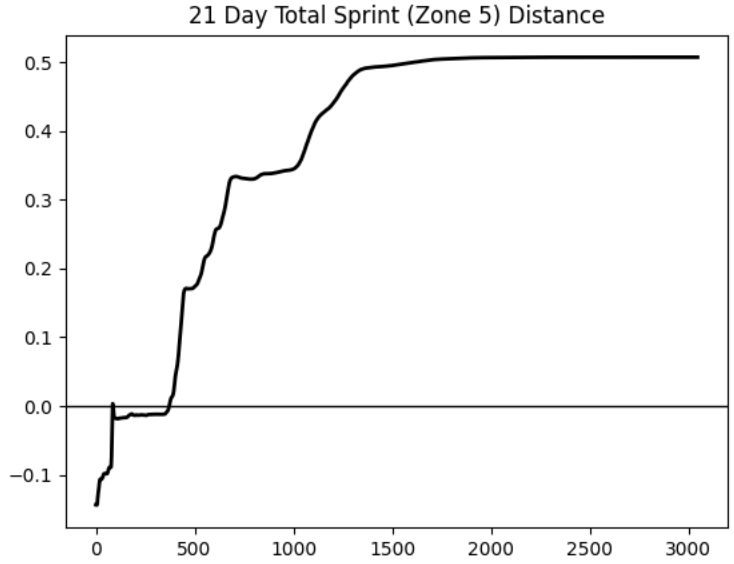


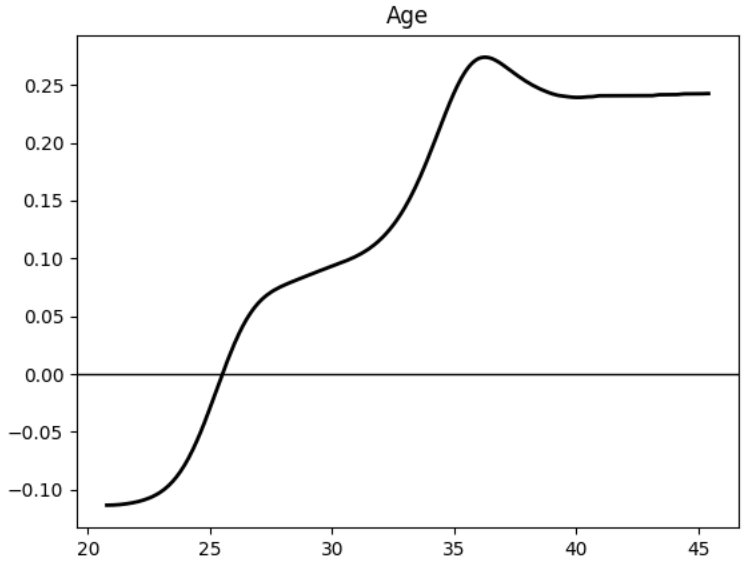
The plots below illustrate the average effects of each feature for each value across the NFL population.

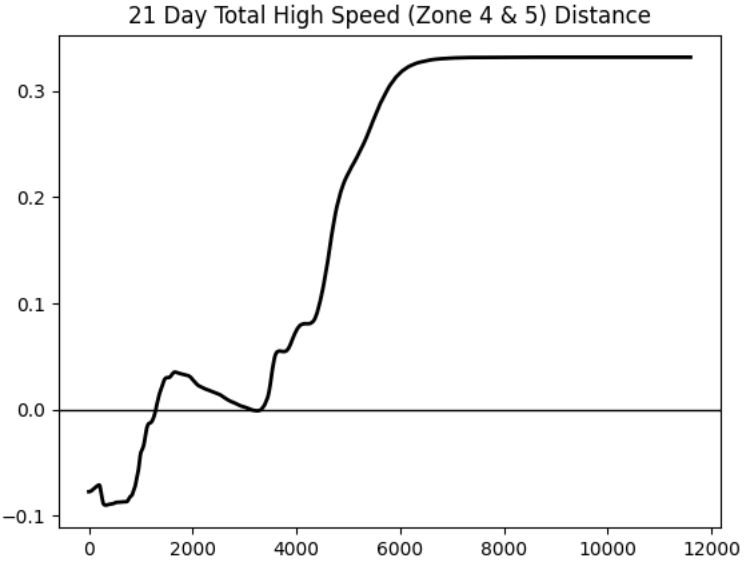


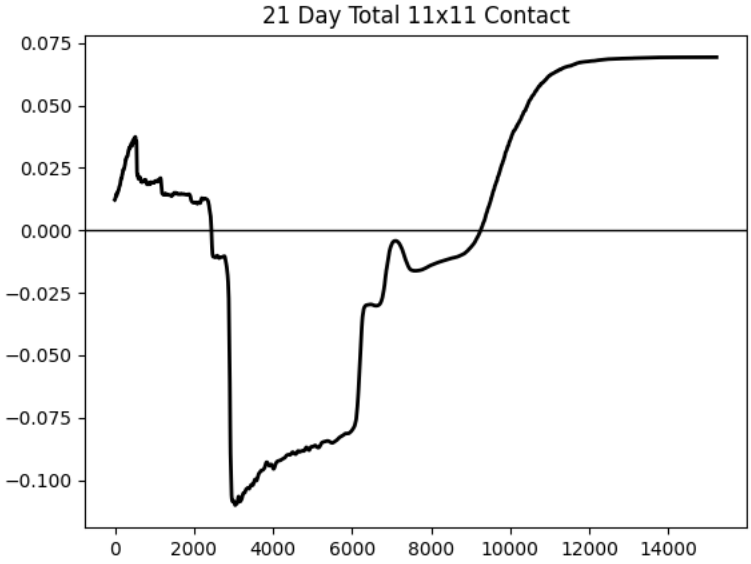

21 Day Zone 5 Decel Distance
0.4
0.3
0.2
0.1
0.0
−0.1
0
20
40
60
80

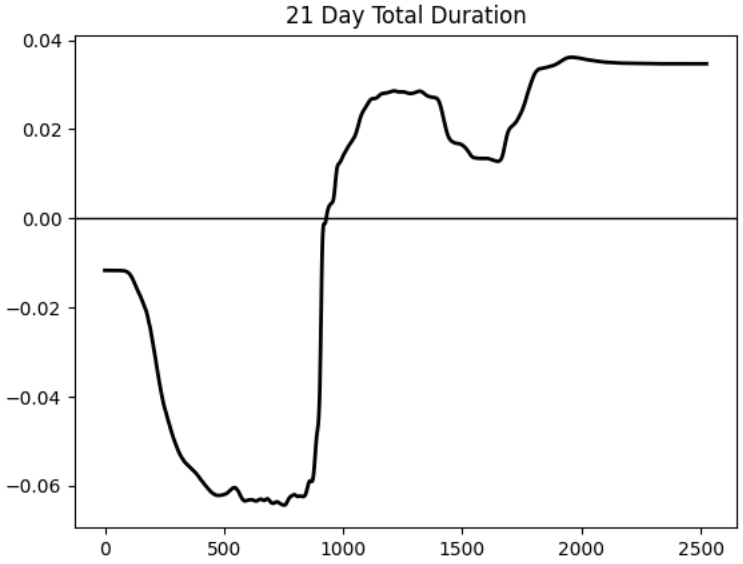

Participated Days
0.10
0.05
0.00
−0.05
−0.10
0
20
40
60
80
100
120

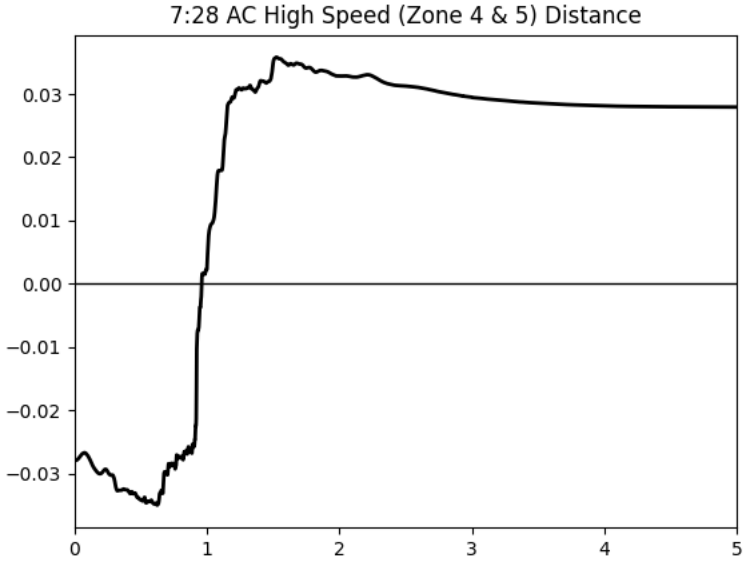

21 Day Total Sprint (Zone 5) Distance
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
0
500
1000
1500
2000
2500
3000


Age
0.25
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
20
25
30
35
40
45


21 Day Total High Speed (Zone 4 & 5) Distance
0.3
0.2
0.1
0.0
−0.1
0
2000
4000
6000
8000
10000
12000


21 Day Total 11x11 Contact
0.075
0.050
0.025
0.000
−0.025
−0.050
−0.075
−0.100
0
2000
4000
6000
8000
10000
12000
14000


21 Day Total Duration
0.04
0.02
0.00
−0.02
−0.04
−0.06
0
500
1000
1500
2000
2500


7:28 AC High Speed (Zone 4 & 5) Distance
0.03
0.02
0.01
0.00
−0.01
−0.02
−0.03
0
1
2
3
4
5