MassDOT IMPACT Phase II - Identification of Risk Factors for SHSP Emphasis Areas

Large Vehicle Crashes

PREPARED FOR

massDOT
Massachusetts Department of Transportation

PREPARED BY

vhb

REPORT DATE: FEBRUARY 2021

Purpose & Background

The Massachusetts Department of Transportation (MassDOT) was awarded a grant by the United States Department of Transportation (USDOT) under its Safety Data Initiative (SDI) competition. MassDOT's work under this grant includes the creation of a Safety Analysis Module in their online IMPACT tool. One feature in this module will be a mapping component which will include crash-based and systemic network screening maps. As part of this work, MassDOT is identifying focus crash types, facility types, and risk factors for their Strategic Highway Safety Plan (SHSP) Emphasis Areas. This report was developed under the SDI project and summarizes the risk factor analysis performed for large vehicle crashes. It also describes a method to identify risk factors using binary logistic regression, which is one potential method to identify risk factors under the SDI grant. Reports for other emphasis areas describe different methods used to adapt to the needs of those areas.

Focus Crash Types and Focus Facility Types

After 34 fatalities due to large truck-involved crashes between 2012 and 2016, MassDOT identified those crashes as an emphasis area in the 2018 Strategic Highway Safety Plan (SHSP)[1]. Based on discussions with MassDOT, VHB established large vehicle involved-crashes as the focus crash type.

VHB and MassDOT used two crash data fields to identify large vehicle involved-crashes: "Federal Motor Carrier Safety Administration (FMCSA) Reportable (All Vehicles)" and "Vehicle Configuration (All Vehicles)". A crash was selected as a focus crash type if the crash was reportable to the FMCSA or if the vehicle configuration field included one of the following vehicle configurations:

- "Bus (seats for 16 or more, including driver)".
- "Bus (seats for 9-15 people, including driver)".
- "Single-unit truck (2-axle, 6-tires)".
- "Single-unit truck (3-or-more axles)".
- "Tractor/doubles".
- "Tractor/semi-trailer".
- "Tractor/triples".
- "Truck tractor (bobtail)".
- "Truck/trailer".
- "Unknown heavy truck, cannot classify".

After querying the crash data in the MassDOT IMPACT tool, VHB identified 907 fatal and serious injury crashes (excluding crashes in the City of Boston due to known under-reporting issues) involving large-vehicles between 2013 and 2017.

VHB used crash trees to identify focus facility types for large vehicle crashes. VHB used fatal and serious injury (KA) crashes for the crash tree. Figure 1 includes the crash tree.

[1] https://www.mass.gov/doc/massachusetts-shsp-2018/download

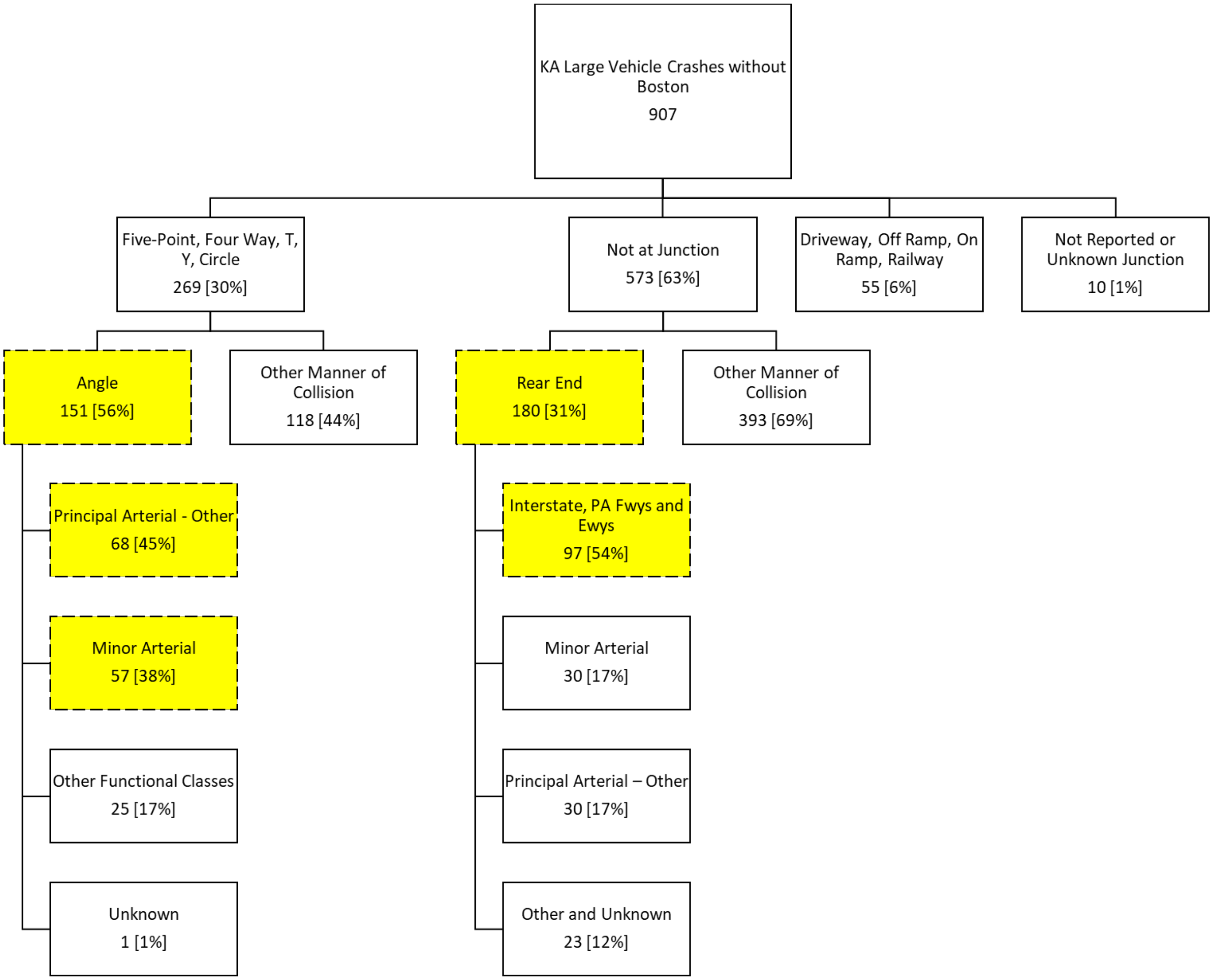

Figure 1. Crash tree to identify focus facility types for KA large vehicle crashes.

The crash tree showed that certain crash types are especially prominent on certain facility types. Intersection crashes comprise 30 percent of the KA large vehicle crashes. Of those 151 (56 percent) involve an angle collision, of which 125 (83 percent) occurred on principal arterial – other and minor arterial roadways.

Non-junction crashes comprised 63 percent of the KA large vehicle crashes. Of those, 180 (a 31 percent plurality) were rear end collisions. Of those rear end collisions, 97 (54 percent) occurred on interstates or principal arterial freeways and expressways. Based on these results, VHB and MassDOT selected two focus crash type-focus facility type combinations for further analysis:

1. Angle crashes at intersections on principal arterial – other and minor arterial roadways.
2. Rear end non-junction crashes on interstates and principal arterial freeways and expressways.

Unfortunately, MassDOT's intersection inventory is not complete as of the completion of this report, so a further risk factor analysis will be performed in the future for the first combination. This report simply provides an additional summarization of those crashes with some insights from that summary.

Risk Factor Analysis

After identifying focus crash type and focus facility type combinations, VHB proceeded with the risk factor analysis. The following sections describe the methodology, data, and results of this analysis for each focus crash/facility type combination.

Angle Crashes at Intersections

As of the completion of this analysis, MassDOT does not have a complete intersection inventory to use for this analysis. As such, this section describes a summary of the angle crashes at intersections on principal arterial – other and minor arterial roadways.

Methodology

The summary is a comparison of KA angle intersection crashes at the focus facility types against all intersection crashes on the same roadways. Where the proportion of focus crash types for a given attribute is statistically larger than the proportion for the comparison group, that attribute is flagged as a potential risk factor. Statistical overrepresentation is checked by building 95 percent confidence intervals around the proportion using sampling errors. Equation 1 and Equation 2 show how the lower and upper bounds, respectively, are calculated based on the proportion of crashes (*p*) and the number of crashes in the sample (*N*). If the lower bound of the large vehicle KA angle intersection crashes is larger than the upper bound of either comparison group, the attribute was considered "overrepresented" for the data.

Equation 1. Calculation of the lower bound of the 95 percent confidence interval for the proportion of crashes with an attribute.

$$95\%\ Confidence\ Interval, Lower\ Bound\ =\ p - 1.96 * \sqrt{\frac{p(1-p)}{N}}$$

Equation 2. Calculation of the upper bound of the 95 percent confidence interval for the proportion of crashes with an attribute.

$$95\%\ Confidence\ Interval, Upper\ Bound\ = p + 1.96 * \sqrt{\frac{p(1-p)}{N}}$$

Data

The data used for this summary is crash data queried from the IMPACT tool for the years 2013 through 2017. The data are limited to crashes at four-way intersections, t-intersections, y-intersections, traffic circles, and five-point or more intersections on principal arterial – other and minor arterial roadways. VHB processed the crash data using the "FMCSA Reportable (All Vehicles)" and "Vehicle Configuration (All Vehicles)" fields to identify large vehicle crashes as describe previously when querying IMPACT and using the "Manner of Collision" field to identify angle crashes.

Results

Table 1 summarizes the crash data attributes which were found to be statistically overrepresented for large vehicle KA Angle intersection crashes compared to either all large vehicle intersection crashes or all intersection crashes. Crashes with these attributes disproportionately result in a fatality or serious injury. Notable characteristics include collisions at four-way intersections, collisions at stop-controlled intersection, and collisions at night with lighting present.

Table 1. Over-represented crash data attributes for large vehicle KA angle intersection crashes.

Crash Data Field	Crash Data Attribute	Percentage of Large Vehicle KA Angle Intersection Crashes	Percentage of Large Vehicle Intersection Crashes	Percentage of Intersection Crashes
Crash Hour	4:00 AM to 4:59 AM	3.9%	0.4%	0.4%
First Harmful Event	Collision with motor vehicle in traffic	91.4%	80.0%	85.1%
First Harmful Event	Collision with pedalcycle (bicycle, tricycle, unicycle, pedal car)	5.5%	0.6%	1.6%
Roadway Junction Type	Four-way intersection	54.7%	46.2%	45.3%
Lighting Conditions	Dark - lighted roadway	18.8%	10.1%	22.0%
Traffic Control Type	Stop signs	39.8%	20.3%	24.5%

Table 2 summarizes the overrepresented driver contributing circumstances for angle intersection crashes for large vehicle drivers and other drivers involved in the crash. Notable circumstances include both drivers failing to yield right of way and disregarding traffic control devices, as well as the other driver (not large vehicle driver) operating the vehicle recklessly and speeding.

Table 2. Over-represented driver contributing circumstances for large vehicle KA angle intersection crashes.

Crash Data Field	Crash Data Attribute	Percentage of Large Vehicle KA Angle Intersection Crash Drivers	Percentage of Large Vehicle Angle Intersection Crash Drivers	Percentage of Intersection Crash Drivers
Large Vehicle Driver Contributing Circumstance 1	Failed to yield right of way	14.0%	8.8%	7.0%
Large Vehicle Driver Contributing Circumstance 1	Disregarded traffic signs, signals, road markings	7.0%	2.9%	2.5%
Other Vehicle Driver Contributing Circumstance 1	Failed to yield right of way	15.8%	13.7%	7.1%
Other Vehicle Driver Contributing Circumstance 1	Disregarded traffic signs, signals, road markings	7.1%	5.7%	2.6%
Other Vehicle Driver Contributing Circumstance 1	Operating vehicle in erratic, reckless, careless, negligent, or aggressive manner	4.4%	1.0%	1.4%
Other Vehicle Driver Contributing Circumstance 1	Exceeded authorized speed limit	3.3%	0.4%	0.3%

Rear End Crashes

The rear end crash analysis followed similar procedures to other emphasis areas, notably the Speeding, Pedestrian, and Bicycle analyses.

Methodology

Based on discussions with MassDOT, VHB used a modeling approach, previously used for the pedestrian, bicycle, and speeding safety analyses, to identify risk factors for large vehicle rear end crashes. Due to the binary nature of the crash severity outcome of interest, the project team used binary logistic regression. This probabilistic modeling technique assesses the probability that an event has occurred (i.e., a KA rear end crash) on a given segment based on the model inputs. Agresti (2007) provides more background information on this method.[2] In this context, odds ratios for variables greater than 1.0 indicate the independent variable increases the probability of a KA crash on the segment, while odds ratios less than 1.0 indicate a decrease in probability. With one focus crash types on one focus facility type, VHB estimated one risk factor model.

When modeling, VHB added variables one at a time, monitoring the coefficients to ensure the inclusion of a variable did not result in large changes in the magnitude of odds ratios for other variables. Additionally, VHB was willing to include variables with p-values upwards of 0.30 assuming the magnitude of the results made sense. VHB did not select a strict level of significance, as Hauer noted this could lead to misunderstanding or outright disregard for potentially noteworthy results[3].

Data

VHB used ArcGIS to manage and integrate data for this analysis. MassDOT provided VHB with various sources of data, as described in the following sections. As stated in the methodology section, the binary logit model was developed at the segment level. As such, VHB tied all data to roadway inventory segments.

Crash Data

MassDOT provided statewide geolocated crash data for the years 2013 through 2017. VHB used the Spatial Join tool in ArcGIS to assign crashes to roadway segments, using the Street Name fields in the crash and roadway data to verify the match is correct. VHB processed the crash data using the “FMCSA Reportable (All Vehicles)” and “Vehicle Configuration (All Vehicles)” fields to identify large vehicle crashes as describe previously when querying IMPACT, using the “Manner of Collision” field to identify rear end crashes, and the functional class field to identify crashes on interstates and principal arterial freeways and expressways.

Roadway Data

VHB downloaded the Massachusetts statewide roadway inventory as of July 2020, available at https://massdot.maps.arcgis.com/home/item.html?id=10a2766a607345928c6a66ffb479c937. Based on discussions with MassDOT, VHB filtered the roadway data in ArcGIS using mileage counted (equal to 1), jurisdiction (not equal to null), and facility type (less than 7) to identify unique segments that were counted for the Highway Performance Monitoring System (HPMS). Filtering the roadway inventory in this way prevented potential double-counting of mileage and VMT for divided roads and roads with overlapping route numbers. The roadway inventory included an estimate of annual average daily traffic (AADT) for each segment.

Horizontal Curve Data

MassDOT provided VHB with horizontal curve data consisting of horizontal curve radii. VHB assigned horizontal curves to roadway inventory segments using the Identity tool in ArcGIS. This allowed VHB to identify the sharpest curve radius present within each roadway inventory segment.

---

[2] Agresti, A. (2007). An Introduction to Categorical Data Analysis. Second Edition. John Wiley & Sons, Inc., New York.

[3] Hauer, E. (2004). The harm done by tests of significance. *Accident Analysis & Prevention*, *36*(3), 495-500.

Freight Data

MassDOT published the "Massachusetts Freight Plan"[4] in 2018 describing challenges present for each mode of freight transportation in the State. Notably, this plan included a list and map of highway freight bottlenecks on the Commonwealth's roadway network. The bottlenecks include the following interchanges:

- Exit 2 of Interstate 90, Lee.
- Interstate 90 and Interstate 91, West Springfield.
- Interstate 90 and Interstate 84, Sturbridge.
- Interstate 90 and Interstate 495, Hopkinton.
- Interstate 290 and Interstate 495, Marlborough.
- Interstate 290 and MA-146, Worcester.
- Interstate 93 and MA-3, Braintree.
- Interstate 95 and Interstate 93, Canton.
- Interstate 95 and Interstate 90, Weston.
- Interstate 95 and US-3, Burlington.
- Interstate 95 and Interstate 93, Reading.
- Central Artery Tunnel System (I-90 and I-93), Boston.
- Bell Circle (US-1), Revere.

VHB used the Spatial Join function in ArcGIS to identify the distance of each segment to the closest bottleneck, with the theory that rear end crashes are associated with back of queue collisions, and queues are consistently present at these bottlenecks.

MassDOT also provided VHB with a shapefile of the freight network, allowing VHB to identify whether a segment is considered part of the truck network or not. Finally, MassDOT provided a shapefile of land use across the Massachusetts, available at https://docs.digital.mass.gov/dataset/massgis-data-2016-land-coverland-use. This file from May 2019 shows the land use of each parcel of land, as codified by the Massachusetts Department of Revenue[5]. MassDOT and VHB reviewed this file to identify parcels which can be labelled as large vehicle trip generators. Table 3 summarizes the land use code designated as large-vehicle trip generators. Note the Department of Revenue lists two categories – commercial and industrial. These were spatially joined to the roadway inventory by identifying the closest parcel with one of these land uses, and the distance to that parcel.

[4] https://www.mass.gov/doc/2017-massachusetts-freight-plan/download

[5] https://www.mass.gov/doc/property-type-classification-codes-non-arms-length-codes-and-sales-report-spreadsheet/download

Table 3. Large vehicle trip generation land use codes.

General Category	Land Use Code	Land Use Description
Commercial	313	Lumber Yards
Commercial	314	Trucking Terminals
Commercial	315	Piers, Wharves, Docks, and Related Facilities
Commercial	316	Other Storage, Warehouse, and Distribution Facilities
Commercial	321	Facilities Providing Building Materials, Hardware and Farm Equipment, Heating, Hardware, Plumbing, Lumber Supplies, and Equipment
Commercial	322	Discount Stores, Junior Department Stores, Department Stores
Commercial	323	Shopping Centers/Malls
Commercial	324	Supermarkets (in excess of 10,000 square feet)
Commercial	325	Small Retail and Services Stores (under 10,000 square feet)
Commercial	326	Eating and Drinking Establishments
Commercial	350	Property Used for Postal Services
Commercial	354	Bus Transportation Facilities and Related Properties
Industrial	400	Buildings for Manufacturing Operations
Industrial	401	Warehouses for Storage of Manufactured Products
Industrial	410	Mining and Quarrying – Sand and Gravel
Industrial	411	Mining and Quarrying – Gypsum
Industrial	412	Mining and Quarrying – Rock
Industrial	413	Mining and Quarrying - Other
Industrial	425	Gas Production Plants

Results

This section describes the results of the risk factor model. The models were run using segments at least 0.05 miles in length. Additionally, segment length was included as a continuous variable in all models to account for exposure. It should not be included when extracting risk factors based on the model. As part of the modeling efforts, VHB used correlation matrices to verify correlation between variables was low. There were no cases where the absolute value of the correlation between explanatory variables exceeded 0.41.

Table 4 summarizes the binary logit regression model for rear end non-intersection crashes on interstates and principal arterial freeways and expressways. All variables (with the exception of length) are binary – meaning the variable is equal to 1 if the condition is true for the segment and 0 otherwise. This model excludes segments in the City of Boston due to known under-reporting issues with the crash data. Additionally, this model is only for non-intersection crashes.

Table 4. Binary logit regression model for rear end non-intersection crashes on interstates and principal arterial freeways and expressways.

Variable	Odds Ratio	Standard Error	z-value	P>|z|	95% Confidence Interval	
Natural log of segment length (miles)	1.68	0.22	4.01	<0.001	1.31	2.17
Pioneer Valley Planning Commission (PVPC)	2.51	0.88	2.63	0.008	1.27	4.97
Median Width ≤ 40 feet	1.67	0.42	2.02	0.043	1.02	2.73
Segment is part of the National Truck Network	1.51	0.58	1.08	0.281	0.71	3.20
Freight Bottleneck within 15 Linear Miles of Segment	1.33	0.43	0.88	0.381	0.70	2.51
Annual Average Daily Traffic (AADT) > 75,000 vehicles per day	1.84	0.56	2.01	0.044	1.02	3.34
Closest land use is industrial and within 4 miles of the segment	1.82	0.43	2.51	0.012	1.14	2.90
Constant	0.01	0.01	-10.25	1.31	0.01	0.03

Note: Number of observations = 3,181; Log likelihood = -330.53456; Pseudo R2 = 0.0694; LR chi2(7) = 49.27; Prob > chi2 < 0.0000.

The binary logit model in Table 4 consists of an interesting mix of characteristics. In terms of standard infrastructure characteristics, narrow medians (40 feet or less) and high traffic volumes (AADT exceeding 75,000 vehicles) both indicate an increased probability of a severe large vehicle rear end collision. Presence of the segment within the truck network and adjacent to an industrial trip generation parcel, truck volume surrogates, both also indicate an increased probability of a collision. The distance to a freight bottleneck indicates the potential presence of queues, in which large vehicles can encounter rear end collisions, especially in the back of the queue. Finally, the presence in the PVPC simply indicates roadways within that MPO are more likely to have a crash, which is likely reflecting unobserved effects or omitted variables. Figure 2 shows, with the PVPC highlighted, a cluster of rear end target crashes within the MPO, specifically south of the city of Springfield. Potential reasons for this cluster include congestion near Springfield or changes in geometric design or pavement design and maintenance between Connecticut and Massachusetts. This cluster contributes to the statistical significance of the PVPC indicator variable.

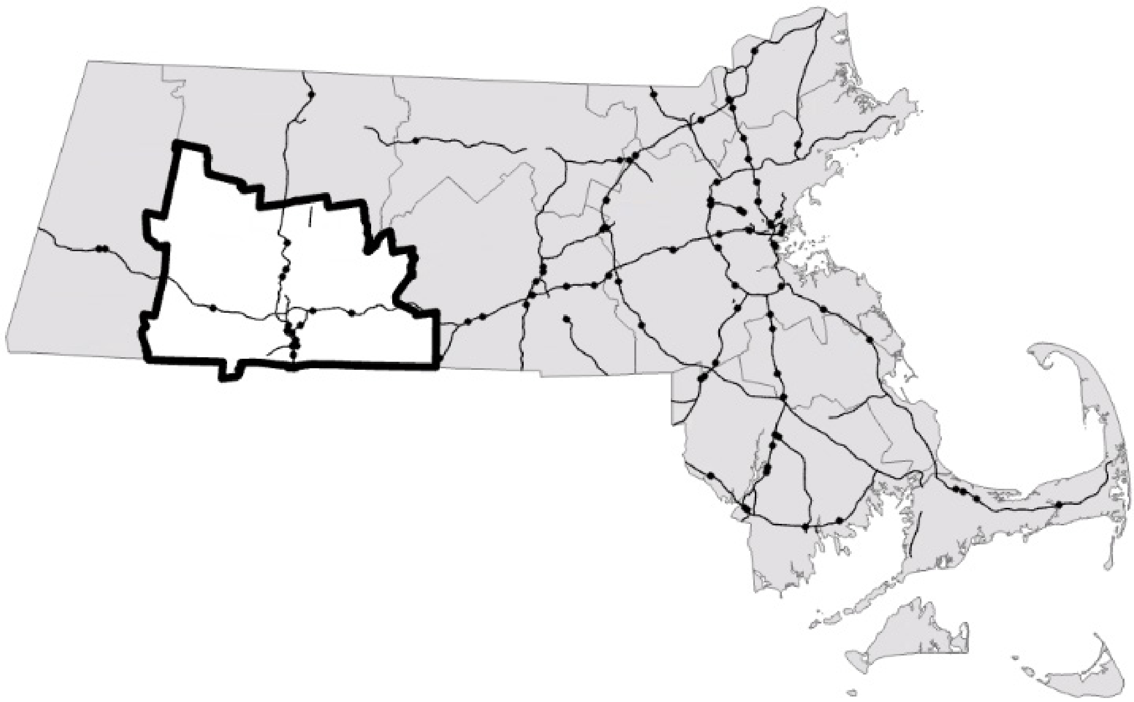

Figure 2. Map of rear end focus crashes in Massachusetts, with the PVPC highlighted.

Conclusions and Recommendations

The purpose of this report is to summarize the systemic analysis of large vehicle crashes on MassDOT highways. VHB and MassDOT identified two combinations of focus crash type and focus facility type:

1. Angle crashes at intersections on principal arterial – other and minor arterial roadways.
2. Rear end non-junction crashes on interstates and principal arterial freeways and expressways.

The angle crash analysis was superficial and based solely on crash data due to the current lack of a complete intersection inventory for Massachusetts. Based on overrepresentation of data, notable crash attributes associated with an increased probability of a large vehicle angle KA crash at an intersection include crashes in which drivers disregard traffic control devices, fail to yield right of way, strike a pedalcycle, occur at a four-way intersection, occur under dark-lighted conditions, and occur at STOP-control intersections.

The analysis of rear end crashes followed a similar approach to other emphasis areas. MassDOT and VHB integrated roadway, crash, and freight data to estimate segment-level binary logit models, estimating the probability that a KA rear end large vehicle collision occurred on the segment or not. Table 5 summarizes the risk factors identified in this analysis. VHB used binary variables when estimating the models, setting up MassDOT to use a binary approach when applying the risk factors.

Table 5. Summary of risk factors for non-intersection rear end crashes.

Rear End Interstates and Principal Arterials Freeways and Expressways	Scoring
Located in PVPC	1 if true; 0 otherwise.
Median Width ≤ 40 feet	Risk Score = -0.025 * Median Width + 1 if less than or equal to 40 feet; 0 otherwise.
Segment is part of the National Truck Network	1 if true; 0 otherwise.
Freight Bottleneck within 15 Linear Miles of Segment	Risk Score = -0.033*Distance to Freight Bottleneck + 1 if Distance to Freight Bottleneck is within 15 miles; 0 otherwise.
AADT > 75,000 vehicles per day	1 if true; 0 otherwise.
Segment is within 4 miles of an industrial land use.	1 if true; 0 otherwise.

VHB recommends MassDOT disregard the Odds Ratio results from Table 4 and assign risk scores of 1 if a characteristic is present on a focus segment. If four characteristics are present, a segment should receive a score of 4, regardless which are present. Table 6 provides an example of how to calculate a risk factor using the model results. In this example, the segment has 3 risk factors present; therefore, it receives a risk score of 1.96 using the scoring schemes in Table 5. MassDOT can normalize the score using the total number of potential risk factors, for example assigning a segment a risk score of 100 percent if all risk factors for the facility type are present. Under this approach, the risk score for the example segment in Table 6 is 33 percent.

Table 6. Example risk score calculations for large vehicle rear end KA crashes on a principal arterial interstate segment.

Variable	Segment Characteristic	Risk Factor	Risk Score
MPO	Central Massachusetts Regional Planning Commission	Located in PVPC	0
Median Width	35 feet	Median Width ≤ 40 feet	0.125
National Truck Network	Yes	Segment is part of the National Truck Network	1
Distance to Nearest Freight Bottleneck	5	Freight Bottleneck within 15 Linear Miles of Segment	0.835
AADT (vehicles per day)	50,000	AADT > 75,000	0
Distance to Trip Generation Land Use	3 miles to Commercial	Closest land use is industrial and within 4 miles of the segment	0
		Total Risk Score:	1.96
		Risk Percent Score:	33%

The segments were then ranked at both the Statewide and MPO levels using the normalized risk score and the percentile of score ranking (rank kind equal to weak) function in ArcGIS. For each normalized risk score, a percentile rank for the given score was computed relative to all the normalized risk scores. If there are repeated occurrences of the same normalized risk score, then the percentile rank corresponds to values that are less than or equal to the given score. The advantage of the weak ranking approach is that it guarantees that the highest normalized score will receive a percentile rank of 100%. The risk categories were then determined using the computed ranks. For example, sites ranked in the top 5 percentile (95 through 100) were categorized as "Primary Risk Site," sites ranked in the next 10 percentile (85 through 95) were categorized as "Secondary Risk Site," and the remaining sites were not categorized. In instances where there are large repeated occurrences of the same normalized risk score, the percentage of segments computed for top 5% or next 10% may not be equal to 5 or 10%. This is a byproduct of the weak ranking approach used. Table 7 and 8 show the distribution of focus facility type segments with the normalized risk score (presented as percentages) across these categories for Statewide and MPO rankings, respectively.

Table 9. Statewide risk categories.

State	Risk Category	Minimum Normalized Risk Score Percentage	Maximum Normalized Risk Score Percentage	Number of Segments	Percent of Scored State Segments
MA	Primary Risk Site	81.26%	98.20%	882	5.02%
	Secondary Risk Site	66.66%	81.26%	1759	10.01%

Table 10. MPO risk categories.

MPO	Risk Category	Minimum Normalized Risk Score Percentage	Maximum Normalized Risk Score Percentage	Number of Segments	Percent of Scored MPO Segments
Berkshire Regional Planning Commission	Primary Risk Site	59.84%	60.00%	12	5.22%
	Secondary Risk Site	58.92%	59.76%	24	10.43%
Boston Region MPO	Primary Risk Site	81.95%	83.33%	277	5.09%
	Secondary Risk Site	80.10%	81.94%	540	9.92%
Cape Cod Commission	Primary Risk Site	33.33%	50.00%	89	9.48%
	Secondary Risk Site	25.00%	29.17%	70	7.45%
Central Massachusetts Regional Planning Commission	Primary Risk Site	80.50%	81.92%	93	5.01%
	Secondary Risk Site	74.85%	80.48%	187	10.08%
Franklin Regional Council of Governments	Primary Risk Site	33.33%	33.33%	288	86.23%
	Secondary Risk Site	N/A	N/A	0	0.00%
Martha's Vineyard Commission	Primary Risk Site	N/A	N/A	0	0.00%
	Secondary Risk Site	N/A	N/A	0	0.00%
Merrimack Valley Planning Commission	Primary Risk Site	67.90%	74.56%	66	5.04%
	Secondary Risk Site	60.24%	67.89%	132	10.08%
Montachusett Regional Planning Commission	Primary Risk Site	57.63%	63.10%	63	6.05%
	Secondary Risk Site	44.34%	57.62%	94	9.02%
Nantucket Planning and Economic Development Commission	Primary Risk Site	N/A	N/A	0	0.00%
	Secondary Risk Site	N/A	N/A	0	0.00%
Northern Middlesex Council of Governments	Primary Risk Site	66.82%	76.81%	43	5.07%
	Secondary Risk Site	63.53%	66.77%	86	10.14%
Pioneer Valley Planning Commission	Primary Risk Site	76.54%	78.54%	43	5.04%
	Secondary Risk Site	64.79%	76.52%	86	10.07%
Old Colony Planning Council	Primary Risk Site	95.33%	98.20%	93	5.04%
	Secondary Risk Site	91.43%	95.32%	184	9.98%
Southeastern Regional Planning and Economic Development District	Primary Risk Site	65.00%	65.83%	170	5.91%
	Secondary Risk Site	50.00%	64.58%	429	14.91%