Implementation of Arterial Bottleneck Characterization Methods in Virginia

http://www.virginiadot.org/vtrc/main/online_reports/pdf/21-r9.pdf

MO ZHAO, Ph.D.
Research Scientist

RAMKUMAR VENKATANARAYANA, Ph.D., P.E.
Senior Research Scientist

Final Report VTRC 21-R9

VIRGINIA TRANSPORTATION RESEARCH COUNCIL
530 Edgemont Road, Charlottesville, VA 22903-2454

vtrc.virginiadot.org

Standard Title Page—Report on State Project

Report No.: VTRC 21-R9	Report Date: September 2020	No. Pages: 41	Type Report: Final	Project No.: 114686
			Period Covered: January 2019-September 2020	Contract No.:

Title:
Implementation of Arterial Bottleneck Characterization Methods in Virginia

Key Words: arterial, bottleneck, congestion, performance, mobility, conflation

Author(s):
Mo Zhao, Ph.D., and Ramkumar Venkatanarayana, Ph.D., P.E.

Performing Organization Name and Address:
Virginia Transportation Research Council
530 Edgemont Road
Charlottesville, VA 22903

Sponsoring Agencies’ Name and Address:
Virginia Department of Transportation
1401 E. Broad Street
Richmond, VA 23219

Supplementary Notes:

Abstract:

In 2019, the Virginia Transportation Research Council (VTRC) published *Improving the Identification and Characterization of Arterial Congestion Bottlenecks* by Zhao and Venkatanarayana. The report described a study recommending that the Virginia Department of Transportation and the Office of Intermodal Planning and Investment pursue development of data conflation and bottleneck ranking tools and explore the implementation of the arterial bottleneck identification methods in rural areas and for before-after studies.

The current study was initiated in January 2019 to implement the two recommendations of the original study. The study consisted of two phases. In Phase 1, VTRC researchers adapted and applied the previously developed method to select intersections in the Culpeper District. Based on the promising results and validation feedback from field experts, Phase 2 was initiated to explore the possibility of developing a conflation tool and to adapt further and apply the bottleneck algorithm and visualizations to all signalized intersections in Virginia. Between February and July 2020, the VTRC research team collaborated with a research team from Old Dominion University to develop new methods and scripts to conflate statewide INRIX XD, annual average daily traffic, posted speed limit, and signalized intersection data wherever they were available. These conflation results were then used by the VTRC research team to analyze arterial bottlenecks across the state. Project stakeholders deemed the statewide bottleneck ranking and before-after study results to be reasonable and useful. The stakeholders are highly interested in further adapting and using the developed methods and scripts for project prioritization, for monitoring of the operational performance of intersections, and to make decisions in the field. The methods, results, and conclusions from these two implementation efforts (the Culpeper District and statewide) are documented in this report.

FINAL REPORT

IMPLEMENTATION OF ARTERIAL BOTTLENECK CHARACTERIZATION METHODS IN VIRGINIA

Mo Zhao, Ph.D.
Research Scientist

Ramkumar Venkatanarayana, Ph.D., P.E.
Senior Research Scientist

Virginia Transportation Research Council
(A partnership of the Virginia Department of Transportation
and the University of Virginia since 1948)

Charlottesville, Virginia

September 2020
VTRC 21-R9

DISCLAIMER

ABSTRACT

In 2019, the Virginia Transportation Research Council (VTRC) published *Improving the Identification and Characterization of Arterial Congestion Bottlenecks* by Zhao and Venkatanarayana.[1] The report described a study recommending that the Virginia Department of Transportation and the Office of Intermodal Planning and Investment pursue development of data conflation and bottleneck ranking tools and explore the implementation of the arterial bottleneck identification methods in rural areas and for before-after studies.

The current study was initiated in January 2019 to implement the two recommendations of the original study. The study consisted of two phases. In Phase 1, VTRC researchers adapted and applied the previously developed method to select intersections in the Culpeper District. Based on the promising results and validation feedback from field experts, Phase 2 was initiated to explore the possibility of developing a conflation tool and to adapt further and apply the bottleneck algorithm and visualizations to all signalized intersections in Virginia. Between February and July 2020, the VTRC research team collaborated with a research team from Old Dominion University to develop new methods and scripts to conflate statewide INRIX XD, annual average daily traffic, posted speed limit, and signalized intersection data wherever they were available. These conflation results were then used by the VTRC research team to analyze arterial bottlenecks across the state. Project stakeholders deemed the statewide bottleneck ranking and before-after study results to be reasonable and useful. The stakeholders are highly interested in further adapting and using the developed methods and scripts for project prioritization, for monitoring of the operational performance of intersections, and to make decisions in the field. The methods, results, and conclusions from these two implementation efforts (the Culpeper District and statewide) are documented in this report.

FINAL REPORT

IMPLEMENTATION OF ARTERIAL BOTTLENECK CHARACTERIZATION METHODS IN VIRGINIA

Mo Zhao, Ph.D.
Research Scientist

Ramkumar Venkatanarayana, Ph.D., P.E.
Senior Research Scientist

INTRODUCTION

Identifying and characterizing traffic bottlenecks are important steps to mitigate congestion and optimize improvement investments. Data-driven methods have the potential to provide better accuracy in bottleneck ranking than a conventional modeling framework,[2] and transportation agencies are increasingly using them. One major limitation of data-driven bottleneck analysis tools currently used by the Virginia Department of Transportation (VDOT) for an arterial network application is that they analyze the road network as a series of links, ignoring the impacts of the side streets and turning movements, which are significant. In 2019, the Virginia Transportation Research Council (VTRC) published *Improving the Identification and Characterization of Arterial Congestion Bottlenecks* by Zhao and Venkatanarayana.[1] That report presented a new bottleneck ranking method for arterial intersections using a node-link approach as a proof of concept that examined all the intersection approaches using multiple data sources, which included high-definition probe vehicle speeds. A visualization tool was developed for summary and interactive drill-down analyses. A case study was conducted for a network grid consisting of 245 intersections in urban Northern Virginia. The results of that study showed that the link-node approach effectively complemented the link-based approaches, which are useful to analyze corridors, and the new method and that the visualization tool can help transportation agencies rank intersection bottlenecks based on quantitative measures to support data-driven decision making. Two major recommendations from that study were as follows:

1. VTRC should continue to stay abreast of progress made by VDOT vendors and consultants working on spatial data conflation. Once commercial conflation solutions are available or when VDOT decides to proceed ahead independently, VDOT should pursue development of an arterial bottleneck identification and characterization tool based on the lessons learned from the proof-of-concept study.

2. VTRC should conduct a pilot study of this new bottleneck analysis method with more stakeholders and explore the implementation of this method in rural areas and for before-after studies; some stakeholders expressed an interest to implement the results to a large network such as the Corridors of Statewide Significance (CoSS) and the entire state.

The study documented in this report was initiated in January 2019 to implement these recommendations. This implementation study consisted of two distinct phases. In Phase 1, researchers adapted and applied the newly developed arterial bottleneck identification and ranking method to select intersections in the Culpeper District, both for bottleneck ranking and for before-after analysis. Based on the results and validation feedback from field experts, the method was improved to filter probe speed data using quality thresholds and was deemed very promising to be expanded to larger geographical networks, possibly the entire state.

The technical review panel (TRP) then directed the research team to initiate Phase 2 to explore the possibility of developing a conflation tool and to adapt further and apply the bottleneck algorithm and visualizations to either the Corridors of Statewide Significance or to all the signalized intersections in the state. In January 2020, VTRC executed a 6-month contract with a research team at Old Dominion University (ODU) to collaborate on developing new methods and scripts to assist data conflations to support the statewide arterial bottlenecks analysis and other VDOT business analyses. The developed conflation scripts[3] were then used by the VTRC research team to analyze arterial bottlenecks across the state. The tasks performed in this implementation study, their results, and the conclusions drawn from these results are documented in this report.

BACKGROUND

The method and the major results of the original study formed the framework for this implementation study.[2] The original study is summarized here to provide the context for this study.

Bottleneck Identification

The bottleneck identification method developed for arterial intersections is based on the spatiotemporal traffic matrix concept. Each cell of the spatiotemporal traffic matrix is an analysis time interval of an analysis segment. The cell is associated with multiple attributes such as speed, volume, segment length, etc. A cell is considered congested when the average speed for the cell is lower than a defined threshold speed. The congested cells constitute the bottlenecks. Bottlenecks are analyzed at the intersection and approach levels. An intersection approach includes all analysis segments between the current intersection and the immediate upstream intersection. This definition of the intersection approach is designed to allow study of the spillbacks at intersections. In Figure 1, Approach A of Intersection N1 has three segments. Each segment is assigned an "order" value to represent its relative location on the approach. The segment closest to the intersection center is assigned an order of 1 (order =1), and the "orders" are ascending from the downstream segment to the upstream segment. An approach bottleneck was defined using two methods. Finding no significant difference in the rankings from the two methods, the computationally faster method (Method 1) was recommended and is used in this study. In this method, an approach is an active bottleneck when at least one segment on the approach is congested during the time interval, no matter the order of segment on the approach.

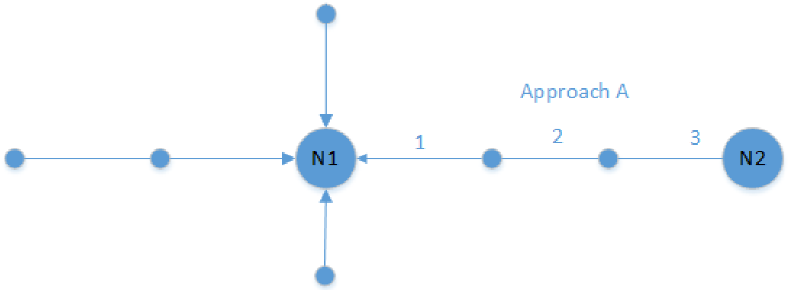


Figure 1. Example of Intersection Layout

The procedures used to identify bottlenecks and calculate performance measures are shown in Figure 2 and explained here.

1. *Step 1: Collect data.* To identify a bottleneck and calculate its performance measures, the following data are required: traffic speeds, traffic volumes, length and location of analysis segments, and threshold speeds. High spatial resolution probe data are used in this study. Traffic volume during each analysis interval is needed, but such data are often not available. In those cases, alternatives such as annual average daily traffic (AADT) and volume profiles may be used. Segment length and location are necessary to estimate queue length and identify the segments on each intersection approach.

2. *Step 2: Conflate data.* This step conflates probe speed data with traffic volume and speed thresholds for the segments. This step could be challenging as data from multiple sources are often saved in different formats and aggregated at different levels.

3. *Step 3: Identify bottlenecks.* Threshold speeds for each segment are required. The threshold speed selected after sensitivity analyses was the light traffic speed, which is the average vehicle speed between the nighttime hours of about 10 P.M. to 5 A.M. The substeps to identify approach and intersection bottlenecks included (1) classifying all segments into congested and uncongested segments for each time interval based on the threshold speeds; (2) for each timestamp, identifying congested approaches using Method 1; and (3) marking the parent intersection as congested during the analysis interval if any of the approaches is congested.

4. *Step 4: Calculate performance measures.* Performance measures are computed to quantify the duration, intensity, variability, and extent of intersection and approach bottlenecks. Performance measures calculated for Method 1 alone are presented here for context.

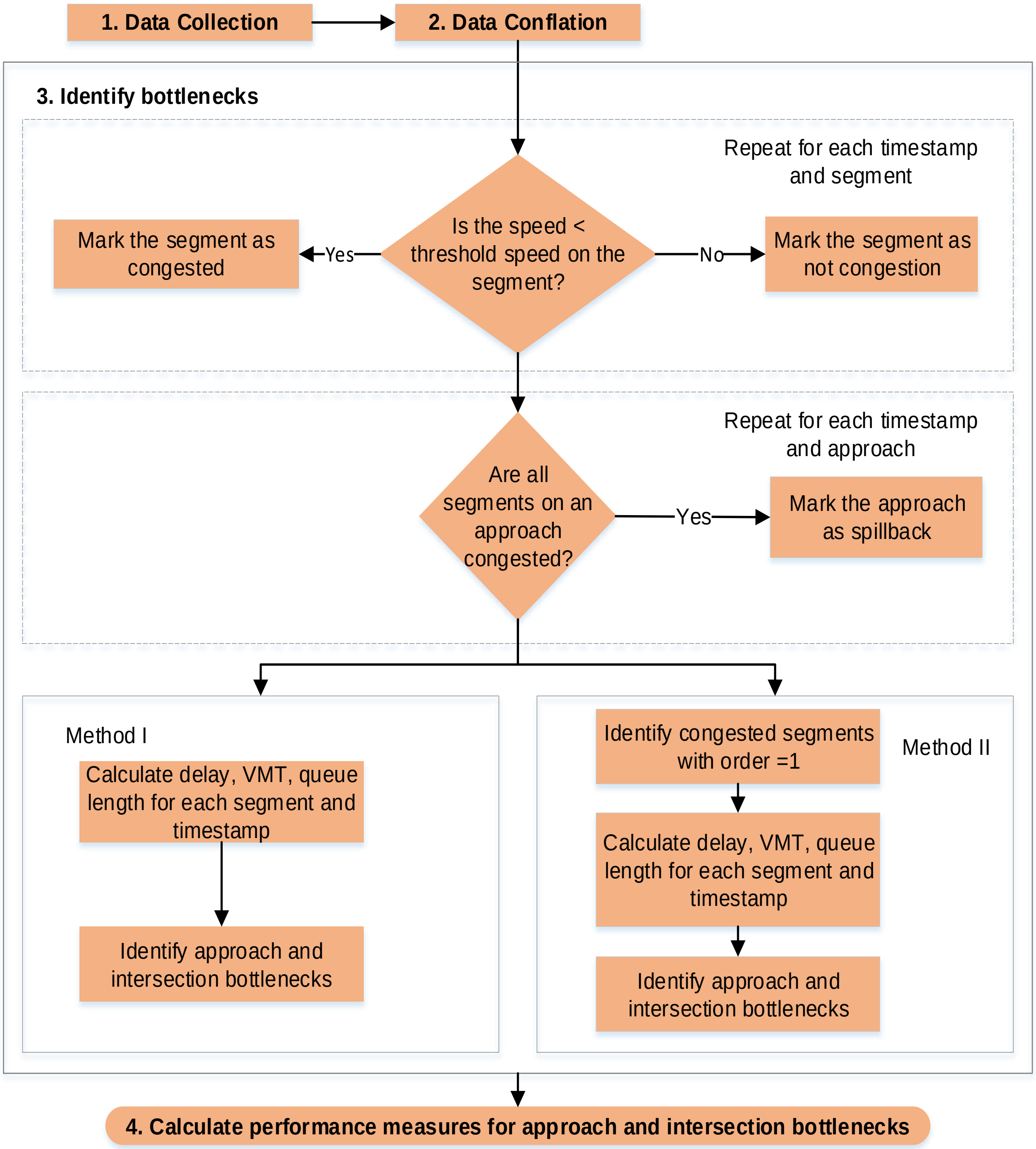


Figure 2. Overview of the Calculation Steps

- *Delay*. The total delay in vehicle-hours is simply the sum of the individual XD segment delays for the entire duration and spatial extent of the bottleneck. (XD segments are short roadway stretches, less than 1 mile, developed by INRIX to publish traffic speed data from vehicle probes.) The delay is calculated with respect to reference speed (60 percent of the light traffic speed, in this study). Delay at an intersection is the sum of delays on all its approaches. Since the number of segments and their lengths are not comparable across approaches, the total delay is not always the best comparative measure. Therefore, delay per vehicle mile traveled (VMT) and delay per mile are also calculated.

- *Queue length.* Queue length is the sum of segment lengths of all bottlenecked segments. Queue length is calculated for each time interval, and the maximum queue length during the bottleneck is used to quantify the spatial extent of the bottleneck. The intersection queue length is the sum of queue lengths on all the approaches.

- *Spillback.* If all XD segments on an approach are congested during a time interval, then that approach is considered to experience spillback. Spillbacks are tracked and summarized for each approach.

- *Bottleneck duration.* Bottleneck duration measures how long the bottleneck was active, i.e., the time elapsed between the bottleneck start and end times on that approach or intersection.

- *Average confidence score.* The average confidence score is an indicator of probe speed data quality. The confidence score is provided by the probe data vendor to determine the source of speed data. Score 30 represents real-time data, Score 20 indicates historical average data, and Score 10 is for reference speed. The approach or intersection average confidence score is calculated as the average of confidence scores of all segments on an approach or at an intersection.

Visualization Tool

All the performance measures calculated in the bottleneck identification process were summarized into various tables and visualizations, including maps, heatmaps, and cumulative distribution functions (CDFs). The visualization tool includes multiple dashboards and provides drill-down capabilities. Users can customize the results in the visualization tool by determining several input parameters including the following:

- the start/end dates, days of week, and time of day periods for the analyses

- the number of intersections to be included in the ranking list

- the intersections of interest by intersection type, corridor, city/county name, the data completeness on intersection approach, or the intersection displayed on the map

- the performance measures for bottleneck ranking, heatmaps, and CDFs.

In addition to simple sum and average, such as total delay, a number of normalization factors were used, including number of days, number of bottlenecks, VMT, and length of roadway. Together these measures characterized the following dimensions: duration (duration; selected measure per day), intensity (most performance measures), and extent (e.g., average queue length per day). CDFs characterized day-to-day variations in congestion. Heatmaps also characterized the intensity and the variability dimensions for the measure chosen (across dates and times of day; days of week and times of day) for both the intersections and the individual approaches. Spatial maps characterized the extent and intensity through the selected measures.

Validation of Results Through Expert Review

Given the data sources, the bottleneck identification method, some performance measures, and the visualizations are all relatively new, an expert panel was employed to evaluate the value of the tools and the validity of the bottleneck ranking results compared to field observations. Through a process of demonstrations and interviews, the research team also identified new performance measures and additional metadata of interest to the experts for making informed decisions. Finally, the experts were asked to identify use cases and implementation concerns.

Northern Virginia District Case Study

A total of 245 intersections in the highly urban Tyson's Corner and Seven Corners area were analyzed in detail. Nine months of INRIX XD data aggregated at 15-minute intervals were used. The XD segments were manually conflated with VDOT AADT shapefiles. The expert panel consisted of VDOT engineering staff with extensive experience and knowledge of traffic operations in the study areas. In general, the panel concluded that the results were consistent with field observations. They expressed some concerns about the quality of probe data and the availability or completeness of the various datasets. The manual conflation was very time- and labor-intensive.

METHODS

The implementation of the recommendations of the original study was carried out in a two-phase approach using two case studies:

1. Culpeper District case study
2. statewide case study.

Compared to the highly urban Tyson's Corner and Seven Corners area in the Northern Virginia District investigated in the original study,[2] the relatively more rural Culpeper District was the focus of the Phase 1 case study. Phase 1 consisted of three main tasks, namely:

1. Select sites and collect and conflate data.
2. Calculate, visualize, and analyze performance measures.
3. Review and validate results.

Based on the promising findings from the Phase 1 case study, the Phase 2 case study expanded implementation to as many intersections across the state as possible. This statewide case study consisted of the following four tasks, with the introductory numbers corresponding to the task numbers:

4. Develop conflation scripts.
5. Collect and conflate data.
6. Calculate, visualize, and analyze performance measures.
7. Review and validate results.

The TRP played a significant role in guiding the study direction and evaluating the results of each phase. For the Phase 1 Culpeper District case study, the TRP was composed of Troy Austin (District Traffic Engineer), Charles Proctor (District Planner), and Rowes Hanna (Assistant District Traffic Engineer) from the Culpeper District. The composition of the TRP for Phase 1 was focused on VDOT personnel with extensive knowledge of traffic operations in the district. The overall project TRP was composed of Mena Lockwood (Assistant State Traffic Engineer, VDOT's Traffic Engineering Division, champion of the original study), Chad Tucker (Program Manager, Office of Intermodal Planning and Investment) (OIPI), Troy Austin (District Traffic Engineer, VDOT's Culpeper District), and Chien-Lun Lan (Research Scientist, VTRC). They represented different stakeholders and oversaw the entire study from various business perspectives, including statewide congestion monitoring, planning, and project evaluation and selection.

Task 1: Phase 1—Select Sites and Collect and Conflate Data

For the Culpeper District implementation case study, the TRP helped identify 40 intersections for bottleneck ranking and 10 intersections with recently completed projects for before-after analyses. Data collection involved downloading traffic speed data from the INRIX XD Version 19.1 data downloader at 15-minute aggregation intervals, AADT data from VDOT shapefiles, and posted speed limit (PSL) data from VDOT shapefiles. For the before-after studies in the Culpeper District, project data were obtained from field staff and the VDOT project pool website. Data conflation for the Culpeper District case study was performed manually following the same procedure used for the Northern Virginia District case study in the original study.

For both the ranking and before-after analyses, INRIX XD Map Version 19.1 was used, which was the latest/current version in effect during this case study. INRIX Roadway Analytics[4] was used as the data source for the XD speed data, which allowed the research team to download all the years of data using the selected XD Map. For bottleneck ranking analyses, nearly 3.5 years of data from January 1, 2016, through July 31, 2019, were downloaded. For before-after analyses, nearly 5.5 years of data from January 1, 2014, through July 31, 2019, were downloaded. It should be noted that the AADT data used for each link were the most recent data available. These data were used as-is in the current study without any growth factor adjustments as a proof-of-concept application.

Task 2: Phase 1—Calculate, Visualize, and Analyze Performance Measures

Wherever possible, the algorithms, codes, and visualizations developed in the original study were used predominantly as-is to calculate and present the performance measures for the Culpeper District case study. Visualizations developed in Tableau were again used both for internal analyses to support the research effort and for broader dissemination of results to the business users. Although the bottleneck ranking visualizations developed for the Northern Virginia District case study were largely adapted for the Culpeper District case study, new visualizations that provided intersection and approach level comparisons of different study periods were developed for the before-after analyses.

Task 3: Phase 1—Review and Validate Results

Mirroring the original study, results for the Culpeper District case study were presented to field experts at the end of Phase 1 in October 2019 to solicit their feedback. Since no other similar tool or results were currently available, validation was predominantly a qualitative exercise, with the field experts verbalizing their opinions about the results in the context of their personal knowledge of the roadway system and traffic. The expert review was used as feedback to update the method and the visualizations. Analyses were re-run, and the final result dataset and visualizations were provided to field users

Task 4: Phase 2—Develop Conflation Scripts

Data conflation for the Northern Virginia District and Culpeper District case studies (including fewer than 300 intersections) were manual processes that were labor- and time-intensive efforts. Since such an effort is not scalable for the entire state, which has an estimated 5,000 to 7,000 intersections, scripts were developed for the statewide case study in close collaboration with the ODU research team.

Task 5: Phase 2—Collect and Conflate Data

For the statewide implementation case study, around 6,000 intersections were identified from three distinct sources, namely, the VDOT Highway Maintenance and Management System (HMMS) signals shapefile, OpenStreetMap (OSM), and select manual identification. The same data sources used in the Culpeper District case study were used here for traffic speeds, i.e., AADT and PSL. The latest available INRIX XD Map, Version 20.1, was used. Three years of data from January 1, 2017, through December 31, 2019, were downloaded.

Task 6: Phase 2—Calculate, Visualize, and Analyze Performance Measures

The output files from the computing codes in the Northern Virginia District and the Culpeper District case studies contained detailed performance measures for each bottleneck at each intersection and approach. These details allowed the users to select any combination of month, day of week, and time of day for summarizing the final results in the visualizations. However, the final Tableau file size was about 22 MB for 40 intersections. Such file sizes are not scalable to about 6,000 intersections in the statewide case study. Further, several data fields such as road name, direction, missing roads, etc., that were conflated manually in the two smaller case studies were not available from the automated conflation performed for the statewide datasets. Therefore, summary bottleneck results by month, weekday/weekend, and pre-specified time periods of the day (A.M. peak, P.M. peak, All Day) were used in the statewide case study. Accordingly, new input and output data structures, codes, and visualizations were developed in an attempt to maintain the balance between high usefulness of the results and manageable result dataset sizes. Visualizations developed in Tableau were again used both for internal analyses to support the research effort and for broader dissemination of results to the business users.

Task 7: Phase 2—Review and Validate Results

Mirroring the two previous case studies, the statewide case study results were presented to field experts at the end of Phase 2 in September 2020 to solicit their feedback. Validation was again predominantly a qualitative exercise to determine the usefulness of the results, to learn of any comments that required updates to the method or visualizations, and to document the feedback. The expert review resulted in minimal updates to finalize the visualizations.

RESULTS AND DISCUSSION

Task 1: Phase 1—Select Sites and Collect and Conflate Data

Figure 3 shows the intersection sites selected by the Phase 1 TRP for the case study (40 for bottleneck ranking and 10 for before-after analyses).

Two challenges were encountered in this Phase 1 implementation study that were not present in the original study. First, the intersections were spread out across the district, rather than being concentrated in one area, and the length of XD segments was generally longer than in the urban areas. Therefore, for each intersection, the spatial extent of analysis (i.e., selection of XD segments on each approach) had to be tailored for a fair comparison. In order to capture the entire bottleneck queue, several XD segments were needed at some locations; at some other sites, the long XD segment extended much farther upstream of the intersection and included other access points along it. With the TRP’s input and concurrence, a maximum approach length of 1.5 miles was used. Second, the data quality indicated by the average confidence score was relatively low for some intersections. Details of how the low data quality was addressed are presented in subsequent sections on Tasks 2 and 3.

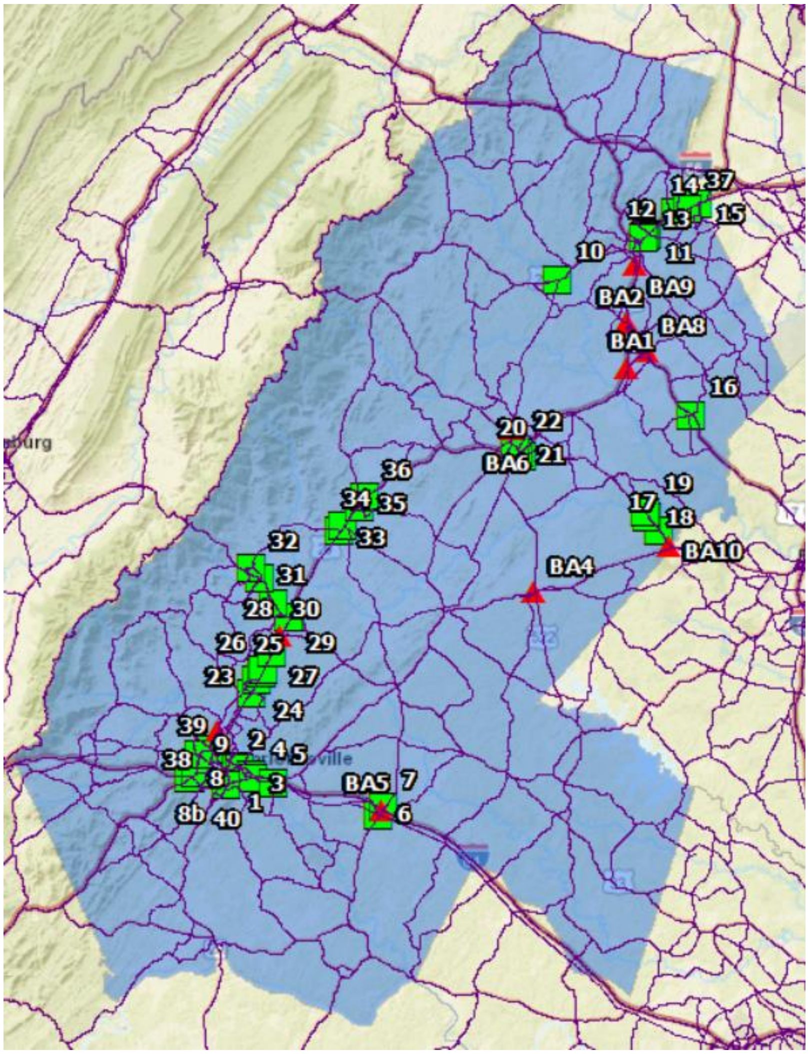


Figure 3. Phase 1 Intersection Sites. Green squares represent ranking sites; red triangles represent before-after analysis sites; and purple lines represent the INRIX XD19.1. The Culpeper District is shaded in blue.

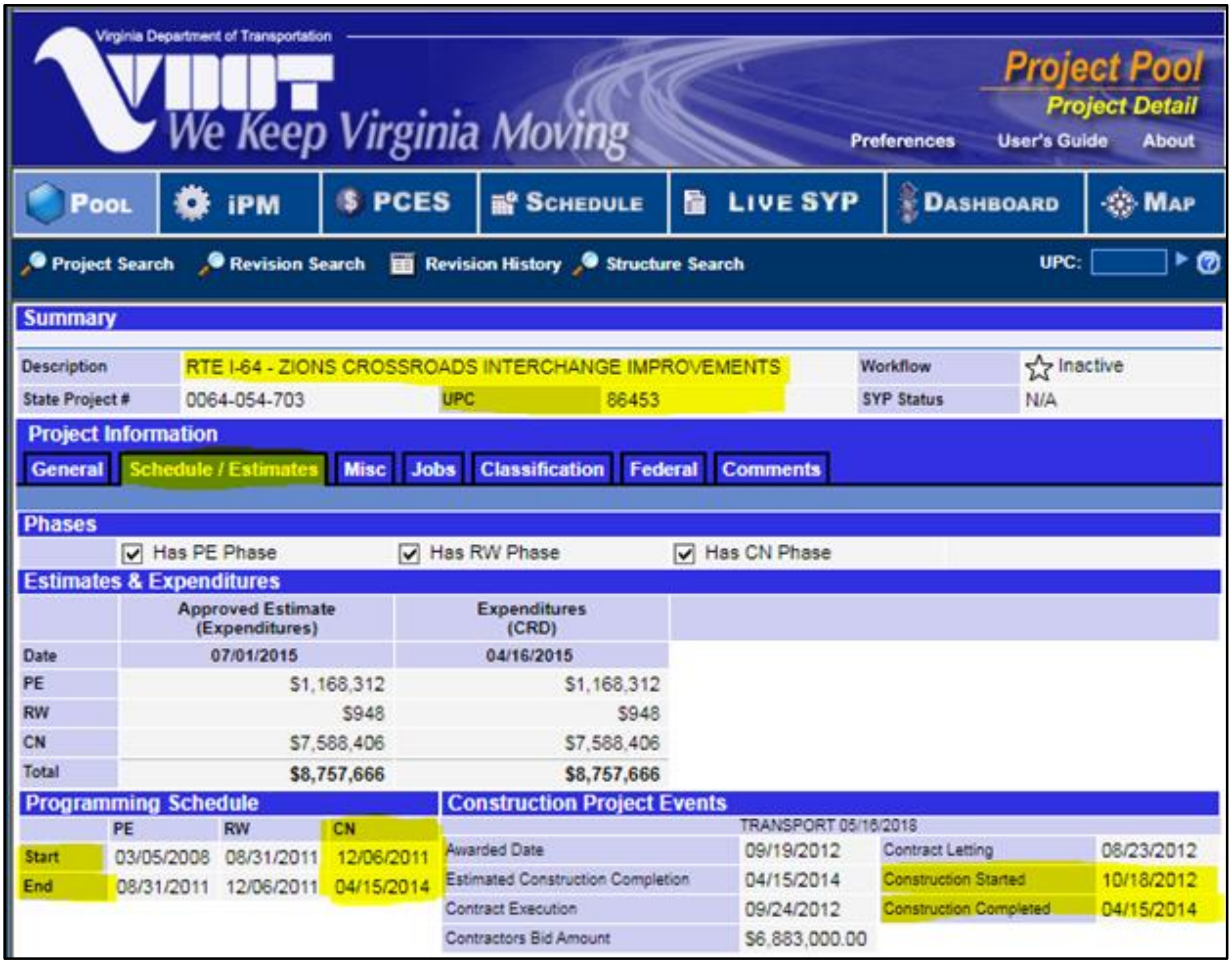
The before-after analyses required one additional dataset, namely, the start and end dates of the relevant project at each site. The VDOT project pool website[5] was used wherever possible to identify the construction start and end dates. For example, the diverging diamond interchange project at Zions Crossroads on I-64 is identified in this database with the Universal Project Code (UPC) 86453. Its schedule is presented in Figure 4. It should be noted that a number of start and end dates are presented on this webpage, and the actual construction period was confirmed with the TRP for this study. For a number of construction projects, the work was performed in-house and there was no UPC code, and hence the construction period was not available from this database. In these cases, the TRP provided the project start and end dates.

To account for driver behavior adjustments after the end of a project, the research team excluded from the analysis 1 to 2 full months of data after the project end date. Wherever a full year of data was not available for both the before and after periods, the same months of different years were analyzed to minimize seasonality effects on the performance measures. These preset before and after periods provided one set of tables and visualizations. A second method of comparison used annual comparisons of performance measures. A third method allowed the user to select one intersection and one set of before and after dates for analysis.

Task 2: Phase 1—Calculate, Visualize, and Analyze Performance Measures

Similar to the original study, all results, tabulations, and visualizations were developed in Tableau and contained drill-down capabilities. These files were used by the research team to identify and address anomalies in data and to enhance the computing codes. Some aspects of the development were in response to the Phase 1 TRP validation and feedback.

All before-after analysis options resulted in a number of new comparative visualizations, as shown in Figure 5 using the example of intersection tables. Similar updates were made to all relevant visualizations in the Tableau files. It should be noted that data quality was generally not an issue for the before-after studies, perhaps because the project intersections had good quality data across the years.



Description	RTE I-64 - ZIONS CROSSROADS INTERCHANGE IMPROVEMENTS			Workflow	☆ Inactive
State Project #	0064-054-703	UPC	86453	SYP Status	N/A

Phases

☑ Has PE Phase ☑ Has RW Phase ☑ Has CN Phase

Estimates & Expenditures

	Approved Estimate (Expenditures)	Expenditures (CRD)
Date	07/01/2015	04/16/2015
PE	$1,168,312	$1,168,312
RW	$948	$948
CN	$7,588,406	$7,588,406
Total	**$8,757,666**	**$8,757,666**

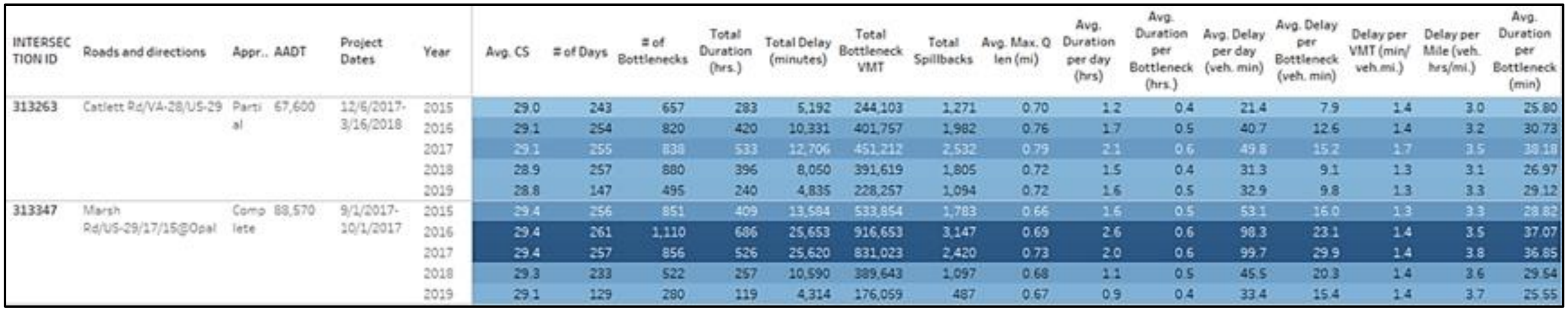
Programming Schedule

	PE	RW	CN
Start	03/05/2008	08/31/2011	12/06/2011
End	08/31/2011	12/06/2011	04/15/2014

Construction Project Events

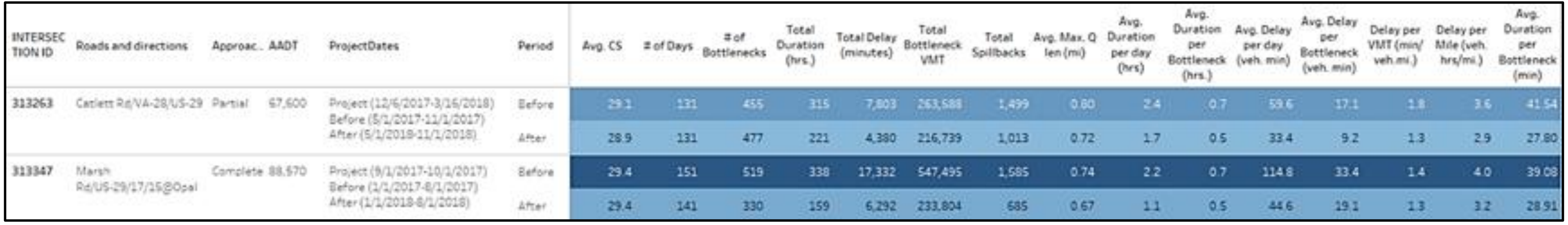
	TRANSPORT 05/16/2018		
Awarded Date	09/19/2012	Contract Letting	08/23/2012
Estimated Construction Completion	04/15/2014	Construction Started	10/18/2012
Contract Execution	09/24/2012	Construction Completed	04/15/2014
Contractors Bid Amount	$6,883,000.00		

Figure 4. Construction Dates for a Project in the VDOT Project Pool Website. Yellow highlights were added by the research team to emphasize specific fields important to this study.



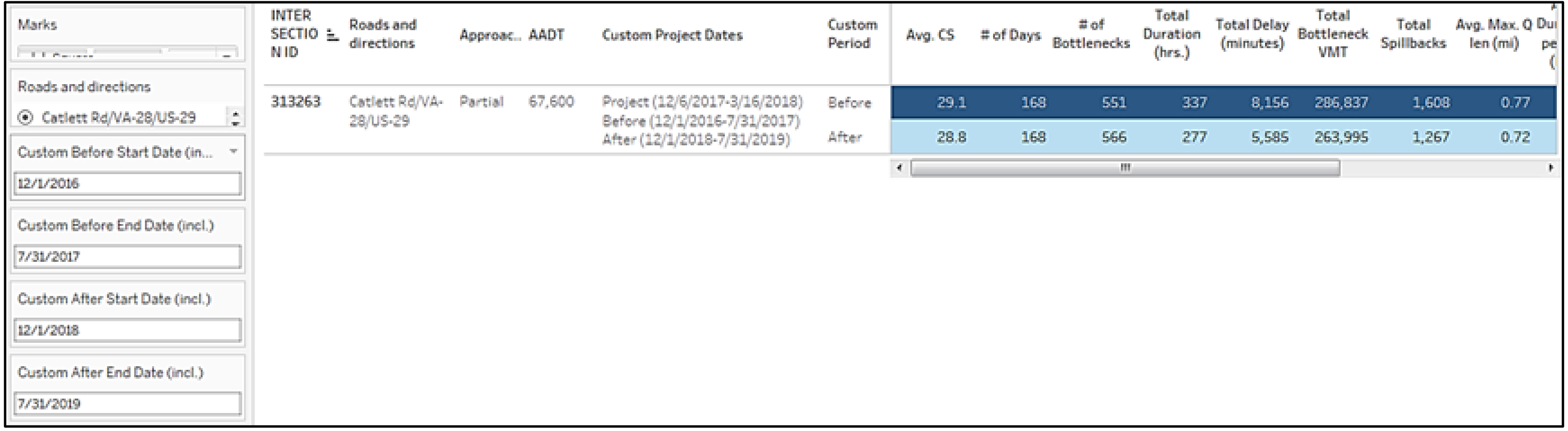
INTERSECTION ID	Roads and directions	Appr..	AADT	Project Dates	Year	Avg. CS	# of Days	# of Bottlenecks	Total Duration (hrs.)	Total Delay (minutes)	Total Bottleneck VMT	Total Spillbacks	Avg. Max. Q len (mi)	Avg. Duration per day (hrs)	Avg. Duration per Bottleneck (hrs.)	Avg. Delay per day (veh. min)	Avg. Delay per Bottleneck (veh. min)	Delay per VMT (min/ veh.mi.)	Delay per Mile (veh. hrs/mi.)	Avg. Duration per Bottleneck (min)
313263	Catlett Rd/VA-28/US-29	Partial	67,600	12/6/2017-3/16/2018	2015	29.0	243	657	283	5,192	244,103	1,271	0.70	1.2	0.4	21.4	7.9	1.4	3.0	25.80
					2016	29.1	254	820	420	10,331	401,757	1,982	0.76	1.7	0.5	40.7	12.6	1.4	3.2	30.73
					2017	29.1	255	838	533	12,706	451,212	2,532	0.79	2.1	0.6	49.8	15.2	1.7	3.5	38.18
					2018	28.9	257	880	396	8,050	391,619	1,805	0.72	1.5	0.4	31.3	9.1	1.3	3.1	26.97
					2019	28.8	147	495	240	4,835	228,257	1,094	0.72	1.6	0.5	32.9	9.8	1.3	3.3	29.12
313347	Marsh Rd/US-29/17/15@Opal	Complete	88,570	9/1/2017-10/1/2017	2015	29.4	256	851	409	13,584	533,854	1,783	0.66	1.6	0.5	53.1	16.0	1.3	3.3	28.82
					2016	29.4	261	1,110	686	25,653	916,653	3,147	0.69	2.6	0.6	98.3	23.1	1.4	3.5	37.07
					2017	29.4	257	856	526	25,620	831,023	2,420	0.73	2.0	0.6	99.7	29.9	1.4	3.8	36.85
					2018	29.3	233	522	257	10,590	389,643	1,097	0.68	1.1	0.5	45.5	20.3	1.4	3.6	29.54
					2019	29.1	129	280	119	4,314	176,059	487	0.67	0.9	0.4	33.4	15.4	1.4	3.7	25.55

(a)



INTERSECTION ID	Roads and directions	Approac..	AADT	ProjectDates	Period	Avg. CS	# of Days	# of Bottlenecks	Total Duration (hrs.)	Total Delay (minutes)	Total Bottleneck VMT	Total Spillbacks	Avg. Max. Q len (mi)	Avg. Duration per day (hrs)	Avg. Duration per Bottleneck (hrs.)	Avg. Delay per day (veh. min)	Avg. Delay per Bottleneck (veh. min)	Delay per VMT (min/ veh.mi.)	Delay per Mile (veh. hrs/mi.)	Avg. Duration per Bottleneck (min)
313263	Catlett Rd/VA-28/US-29	Partial	67,600	Project (12/6/2017-3/16/2018) Before (5/1/2017-11/1/2017) After (5/1/2018-11/1/2018)	Before	29.1	131	455	315	7,803	263,588	1,499	0.80	2.4	0.7	59.6	17.1	1.8	3.6	41.54
					After	28.9	131	477	221	4,380	216,739	1,013	0.72	1.7	0.5	33.4	9.2	1.3	2.9	27.80
313347	Marsh Rd/US-29/17/15@Opal	Complete	88,570	Project (9/1/2017-10/1/2017) Before (1/1/2017-8/1/2017) After (1/1/2018-8/1/2018)	Before	29.4	151	519	338	17,332	547,495	1,585	0.74	2.2	0.7	114.8	33.4	1.4	4.0	39.08
					After	29.4	141	330	159	6,292	233,804	685	0.67	1.1	0.5	44.6	19.1	1.3	3.2	28.91

(b)



Roads and directions

⦿ Catlett Rd/VA-28/US-29

Custom Before Start Date (in...

12/1/2016

Custom Before End Date (incl.)

7/31/2017

Custom After Start Date (incl.)

12/1/2018

Custom After End Date (incl.)

7/31/2019

INTERSECTION ID	Roads and directions	Approac..	AADT	Custom Project Dates	Custom Period	Avg. CS	# of Days	# of Bottlenecks	Total Duration (hrs.)	Total Delay (minutes)	Total Bottleneck VMT	Total Spillbacks	Avg. Max. Q len (mi)
313263	Catlett Rd/VA-28/US-29	Partial	67,600	Project (12/6/2017-3/16/2018) Before (12/1/2016-7/31/2017) After (12/1/2018-7/31/2019)	Before	29.1	168	551	337	8,156	286,837	1,608	0.77
					After	28.8	168	566	277	5,585	263,995	1,267	0.72

(c)

Figure 5. Before-After Analysis Options for Intersections Using (a) Each Year, (b) Preset Before-After Dates, and (c) Custom Before-After Dates

In response to TRP feedback, a comparison of the A.M. peak intersection ranks based on total bottleneck delays calculated from (1) all data and (2) only high-quality data (average confidence score ≥ 25) was conducted. The ranks and the corresponding average confidence scores are presented in Figure 6. Three intersections that ranked high in A.M. peak without data quality thresholding (average confidence score = 25) were ranked considerably lower with the thresholding. Based on the TRP's field observations, the ranks without data quality thresholding were overestimated for these three intersections and the ranks with thresholding were more consistent with the reality. Data quality thresholding did not greatly change the rankings at other intersections in the A.M. peak and hardly affected the rankings or average confidence scores in the P.M. peak.

Both the Exit 118 interchange and US-29/VA-3 intersections were affected by long XD segments that extended past neighboring intersections. Therefore, the data characteristics for the entire XD segment were an inherent feature of the data that could not be improved through data filtering or analyses alone.

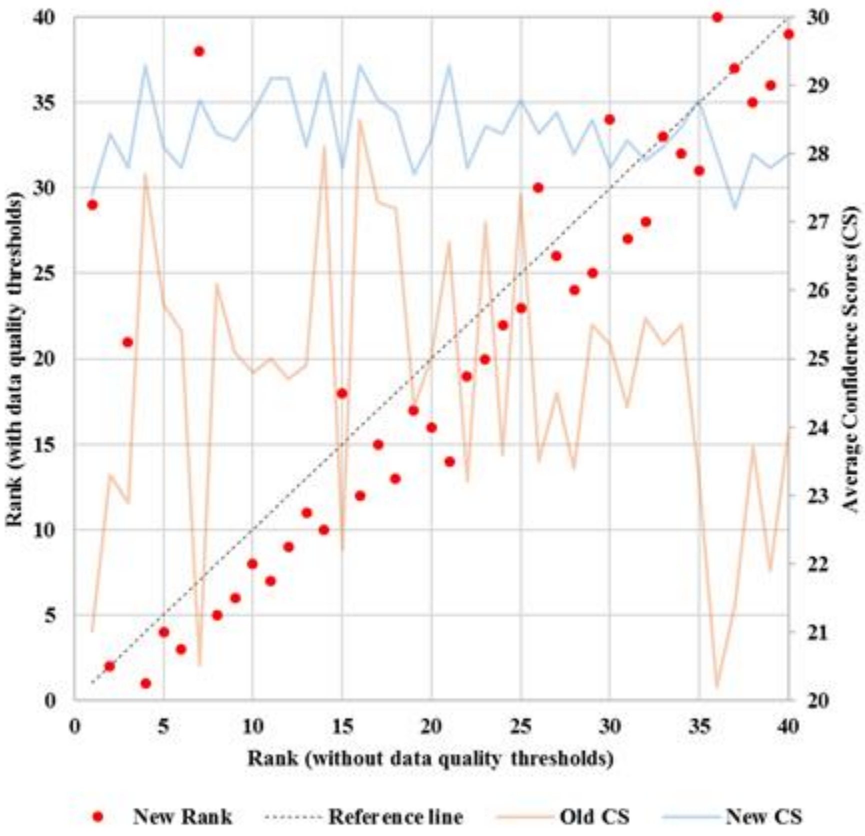


Figure 6. Comparison of A.M. Peak Intersection Bottleneck Ranks and Average Confidence Scores With (New) and Without (Old) Confidence Score Thresholding. Each red dot corresponds to one intersection. CS = confidence score.

The lack of probe speed data on some approaches was an issue familiar to the TRP. In spite of that limitation, they wished to see more ranking analyses of intersections with those approaches that did have data. They stated explicitly that not all approaches at an intersection are often important to analyze from a bottleneck congestion perspective. If some low volume and low delay approach data were not available, they still wanted to analyze and report the delays for the main approaches, especially to facilitate sketch planning, network screening, and communication with stakeholders in other agencies.

To quantify the impacts of not having data for some minor approaches, the research team identified all the intersections where the major and minor roads could be identified clearly (based on the differences in the number of lanes, amount of traffic carried, and intersection type) and where data for all the approaches were available. For those nine intersections, the same method was applied by removing the minor street approaches. Results of total delay and delay per VMT are shown in Figure 7. The observed differences between the two sets of results are indeed small. It should be noted that interchanges and intersections with roads carrying comparable traffic volumes are expected to be much more affected if data for some approaches are missing.

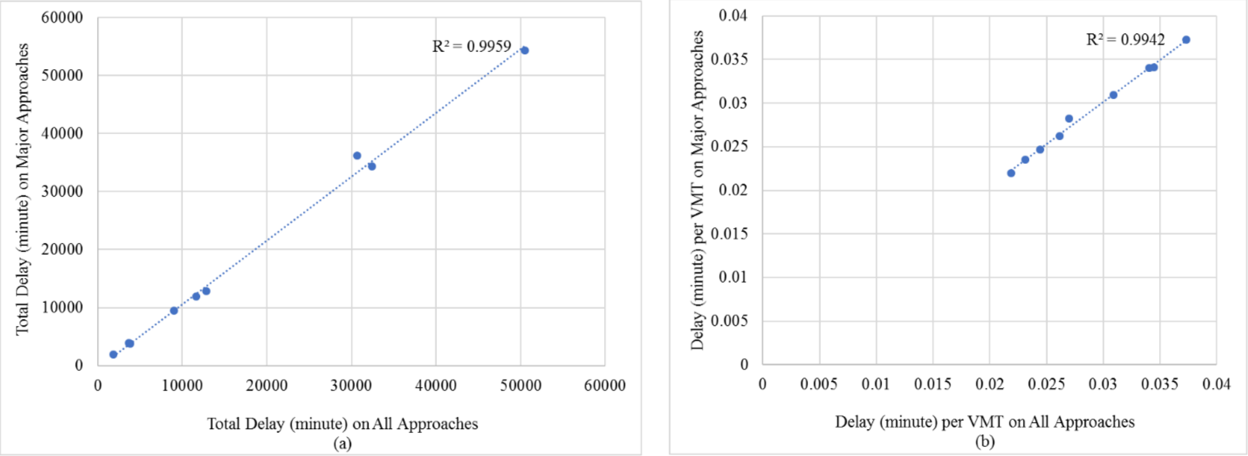


Figure 7. Comparison of Analysis Results When All Approaches at an Intersection Have Data and Only the Major Approaches Have Data Using Two Measures: (a) Total Delay (minutes), and (b) Delay (minutes) per VMT (vehicle miles traveled). Each blue dot corresponds to one intersection.

Task 3: Phase 1—Review and Validate Results

Culpeper District Intersection Ranking

The Tableau files were used to present the results to the Phase 1 TRP at a meeting in the Culpeper District on Friday, August 23, 2019. Major observations from the results files and the TRP feedback are presented and discussed in this section. The TRP validated most of the rankings and the relative magnitudes of the measures to be in line with their field observations. The TRP noted a few exceptions, such as the following:

- The interchange off Exit 121 on I-64, Proffit Road/Airport Road/US-29, and US-29/US-29 Business North of Madison—all in the A.M. peak (6 A.M. to 10 A.M.)—

seemed in their opinion to have been ranked higher (i.e., more congested) than in reality.

- The interchange off Exit 118 in the A.M. peak and the intersection of the US-29 ramp to VA-3 in the P.M. peak (3 P.M. to 7 P.M.) seemed in their opinion to have been ranked with much smaller delay than in reality.

Both sets of observations were highly correlated with a number of known data issues such as the following:

- low average confidence scores at some of these sites and time periods
- long XD segments that extended across multiple intersections, leading to some unexpected high or low total intersection delays
- variable lengths of the approach XD segments selected for analyses at different sites, leading to some unexpected high or low total intersection delays
- lack of XD segments and AADT data for some approaches, leading to partial performance measures at the corresponding intersections
- lack of lane-by-lane speed and volume data, leading to some observed field improvements for some lane movements not being reflected in the data.

An example of the first issue was the interchange at I-64, Exit 121, which was ranked No. 1 by total bottleneck delays in the A.M. peak in 2018 with an average confidence score of only 20.3. Toward mitigating this issue, the TRP and the research team decided to explore thresholding the data using a minimum average confidence score of 25 for a 15-minute period, which corresponds to 50 percent or more real-time data during the daytime. Given that 20 (historic) and 30 (real-time) are the possible values for daytime confidence scores for a given XD segment and timestamp, an average of 25 for a 15-minute period ensures that at least one-half of the available data are real time.

The TRP determined that normalized delays such as delay per mile or delay per VMT were more relevant to rural intersections with variable XD segment and approach lengths. The TRP argued that although predefining the approach length to one low value such as 250 ft would underestimate the total queues and delays on some highly congested approaches, long approaches are more susceptible to delays from accesses and not the intersection itself. Further, high volume interchanges will have far more total delays than other intersections. For the last reason, the TRP further wished to separate the ranking of freeway interchanges and surface street intersections for some discussion and communications purposes. The research team included the option in Tableau to include or exclude the freeway interchanges in the ranking process. The details about performance measure normalization and approach length are documented here for consideration by future users.

Culpeper District Before-After Analysis

Before-after analysis results generally seemed to match with the Phase 1 TRP's expectations and field knowledge; however, it should be noted that only six projects were analyzed. Given that different intersections are not compared, the TRP considered the performance measures equally favorably with or without normalization; however, since delay, VMT, queue length, and number of bottlenecks all increase or decrease together at the project sites, performance measures with normalization may not show any differences between the before and after periods. One specific intersection project (Figure 8) and its before-after results are presented here as an example. Even after the interchange to carry traffic from US-29 South to US-17 (Marsh Road) was built and opened, many vehicles continued to make the left turn at the signalized intersection. To remedy this situation, VDOT closed US-17 South near Pomeroy Lane in September 2017, as shown in Figure 8. Although VDOT has noticed mobility improvements at the intersection after this modification, they had not been studied or quantified.

Based on the information provided, the before period was selected as January 1 through August 1, 2017, and the after period was selected as January 1 through August 1, 2018. The total intersection bottleneck delays for these periods were 17,332 and 6,292 vehicle-minutes, respectively. Annual intersection bottleneck delays in 2017 and 2018 were 916,653 and 389,643 minutes, respectively, showing an improvement of around 60 percent. The TRP noted that the magnitude and direction of these quantified delay improvements seemed consistent with their qualitative experience. As mentioned previously, the normalized measures, delay per mile and delay per VMT, using the before and after periods, respectively, changed from 4.0 to 3.2 hours/mile and 1.4 to 1.3 minutes/vehicle/mile.

At one intersection (VA-20/US-522), the analyzed project filled in a median island on US-522 and the pavement was restriped to have dedicated left-turn lanes where there previously had been shared left/through lanes. For this project, the TRP observed that the left turning movement saw considerable improvement in the after period compared to the before period in the field; however, such nuances are not currently discernible in the analysis results because lane-by-lane data are not available.

Figure 8. Project Location for Signalized Intersection 313347 Showing (a) Interchange From US-29 South to US-17 South, the Intersection Analyzed (Marked in Yellow Box), and the Road Closure Site; and (b) the Road Closure Site

In spite of this drawback, the TRP expressed high interest in using the quick and approximate results to improve understanding and communication of mobility benefits of new projects (especially with alternative/innovative intersections) to their stakeholders to complement the safety analyses that they are already performing regularly.

Task 4: Phase 2—Develop Conflation Scripts

For Phase 2, as many intersections as possible across the entire state were identified from the available data sources, resulting in about 6,000 intersections. Conflation of data with different spatial scales and schemas was one of the challenges during the original study and the Culpeper District study. Manual conflation of data was time- and resource-consuming. Automating the data conflation process was identified as necessary for the statewide study.

Working closely with the VTRC research team, the ODU research team developed a set of open source scripts to assist in data preparation for the statewide bottleneck ranking study and other VDOT business efforts. The scripts create a browser-based application (Figure 9) that can efficiently assemble multisource data including road network geospatial data, traffic data (e.g., AADT), and third-party data (e.g., INRIX speed data).[3] The inputs for the scripts are shapefiles with data to be conflated, and the outputs are data tables in csv format. The scripts can fulfill two functions:

1. *Line to line conflation.* This function conflates polyline segment-based data from multiple map layers with different referencing systems. For example, to find the AADT for the XD segments, this function will locate each XD segment on the INRIX map and identify the corresponding AADT in the VDOT AADT map layer. This function can be widely used for VDOT business analyses that need to conflate segment-based data from multiple map layers.

2. *Line to point conflation.* For each intersection on a user-provided map layer, this function will identify the polyline segments along each intersection approach from a polyline map layer. For example, the function can locate an intersection from a VDOT HMMS map layer and identify the XD segments on each intersection approach from the INRIX map layer and their orders on the approach from the intersection center. This function was designed to prepare data for the intersection bottleneck analysis.

Details of these conflation functions such as the buffer distance between line and line entities, line and point entities, angle between the lines, use of attributes to improve accuracy, identification of the first segment, and order of segments on each approach, etc., are detailed in the user manual and the scripts themselves.[3] Since these are subject to change over time and extensive validation was not performed, those details are not presented here.

The scripts were coded with JavaScript (using node.js) and can be executed on a standard VDOT laptop; advanced users can further modify the scripts to customize the functions. The scripts, user manual, and tutorial are available to VDOT users upon request.

```
495    line_to_line_num:496
497    line_to_line_num:496
499    line_to_line_num:496
500    line_to_line_num:496
501    line_to_line_num:496
502    line_to_line_num:496
503    line_to_line_num:496
505    line_to_line_num:496
506    line_to_line_num:496
507    line_to_line_num:496
508    line_to_line_num:496
509    line_to_line_num:496
510    line_to_line_num:496
>
```

Figure 9. Browser-Based Conflation Application

Task 5: Phase 2—Collect and Conflate Data

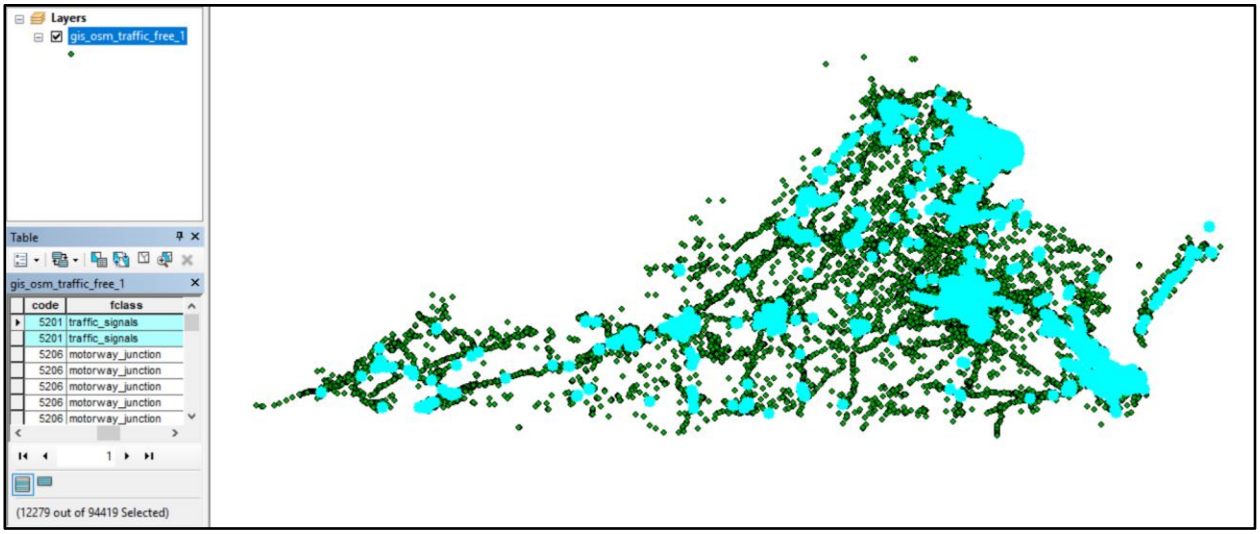
VDOT's HMMS traffic signals shapefile[6] included all the signalized intersections on all VDOT roads across the state; they contained 3,125 points, of which 3,079 were marked as traffic signals (the rest were intersection flashers). As this HMMS shapefile did not include any signals in cities or towns, OSM was explored as a data source.[7] OSM's virginia-latest-free.shp.zip contained a traffic node layer named "gis_osm_traffic_free_1" that included 12,279 traffic signal nodes (denoted by the field "fclass" value "traffic_signals") (Figure 10).

Although the traffic signals on city- and town-owned road systems were included in the OSM layer as desired, these nodes were associated with three issues of concern: (1) they also included VDOT-owned roads already represented in the HMMS layer; (2) most intersections were represented by multiple node points; and (3) there were some errors and omissions. To overcome these issues, the following steps were taken:

1. The ODU research team joined the OSM layer with the HMMS layer. If an intersection is present in the HMMS layer, irrespective of whether corresponding intersection(s) are present in the OSM layer, then the HMMS intersection is considered as the reference for further work and the OSM intersection(s) is ignored. At all other locations, OSM layer intersections are considered the main references.

2. The ODU research team clustered all the nodes on OSM and considered only one point for each cluster in the final results.

3. The VTRC research team manually identified signalized intersections on the CoSS roadways in the Northern Virginia and Hampton Roads districts using Google Earth satellite images and Google StreetView.



code	fclass
5201	traffic_signals
5201	traffic_signals
5206	motorway_junction
5206	motorway_junction
5206	motorway_junction
5206	motorway_junction
5206	motorway_junction

(12279 out of 94419 Selected)

Figure 10. OSM Traffic Nodes Layer. Signals are represented by the cyan dots, and all other nodes are represented by the green dots. OSM = OpenStreetMaps.

A total of 6,012 intersections were identified, including 3,125 from HMMS, 2,699 from OSM, and 188 from manual processing. Using the line to point conflation function of the scripts, the XD segments on each intersection approach within 1.5 miles were identified. INRIX XD Map Version 20.1 was used in this study. Among the 6,012 intersections, 5,443 had XD data for at least one approach. A quality assurance screening was conducted to identify conflation errors. One of the current limitations of the scripts is that the line to point function cannot handle complicated geometries. This and other limitations are discussed later in the results for Task 3.

As with the Culpeper District study, AADT data were obtained from VDOT shapefiles. The line to line conflation function of the scripts was used to find the AADT for each XD segment, and it took about 2 to 3 hours to process a total of 108,961 XD segments in the state, which was a significant efficiency improvement over manual conflation. After the intersections with conflation errors and those with no AADT on any of the approaches were removed, a total of 4,937 intersections were used in the statewide study; 18 percent had data for four or more approaches, 22 percent had data for three approaches, 50 percent had data for two approaches, and 10 percent had data for one approach.

A total of 22,107 XD segments were used for the intersection approaches. Speed data at 15-minute aggregation intervals on these segments were downloaded from INRIX Roadway Analytics for 2017 through 2019. The raw data were about 150 GB. About 60 percent of these segments had complete data for the entire 3 years, and 79 percent had some data for each year. Because of the INRIX map updates, some XD segments were added or removed every half year, which was the major cause of data incompleteness. Speed data, though not fully complete, were available for 94 percent, 91 percent, and 84 percent of the segments in 2019, 2018, and 2017, respectively.

Metadata for the statewide intersections, approaches, and XD segments were largely derived from the conflation scripts and some calculations in a Microsoft Excel spreadsheet for practical scalability. The “road list” field from the INRIX XD metadata was used as the road names/numbers (which contains all of the road names and numbers concatenated by "|"), and the "bearing" field was used in place of the cardinal direction. Although these two fields do not provide consistent, ideal details about the approaches, no other data were available at the statewide scale. These were supplemented by the "major road" and "minor road name" fields from HMMS wherever they were available. The researchers retained a field in each metadata table to control whether or not they should be visible and used in the Tableau visualizations to account for situations when manual review indicated that the conflation results or data are not useful. Because of the non-availability of information regarding the total number of approaches at an intersection, the final statewide results metadata simply presented the total number of approaches conflated.

Task 6: Phase 2—Calculate, Visualize, and Analyze Performance Measures

The size of the Tableau file for the Culpeper District case study was around 22 MB for 40 intersections. Using the same data structure for the entire state with thousands of intersections would have resulted in unmanageable datasets. Even if the analyses could be completed in a reasonable time frame, the output datasets and Tableau files would not be usable because of file size and the slow response. Therefore, the following decisions were made regarding the data and file structures:

- All of the intersections were retained because there was no easy way to decide which intersections to omit. Further, projects could be deployed at any intersection and their before-after analyses would be desired by the field personnel.

- All of the performance measures were also retained, again because there was no easy way to decide which measures would be more useful for which intersections.

- Three years of data was considered to enable users to observe performance trends and to perform some before-after analyses.

- Instead of retaining the detailed performance measures at each bottleneck (including date and time), results were summarized by intersection/approach, year, month, day of week (Sunday, Monday, etc.) and preset time of day (A.M. peak [5 A.M. to 10 A.M.); P.M. peak [3 P.M. to 8 P.M.); and All Day [5 A.M. to 8 P.M.]). This approach would still afford several perspectives with regard to the intersection and approach performance both at the detailed level for districts and at the summary level for the entire state while keeping the data size manageable at the same time.

Python scripts were developed in the original study and the Culpeper District study to identify bottlenecks and calculate performance measures. Those scripts were rewritten for the statewide analyses to better fit the large amount of input data and calculate performance measures at the new temporal aggregation levels. Whereas the quality of probe speed data was deemed to be relatively high and acceptable in the urban sites in the Northern Virginia District, additional consideration of data quality was identified as a need for rural intersections from the

Culpeper District case study. Accordingly, only speed data with an average confidence score of 25 or higher were used.

The calculations were performed on one node of a high performance computing cluster. The node had a maximum of 20 cores available, and the maximum memory available for each core was 9 GB. For a large area like the Northern Virginia District with 1,827 intersections and 7,426 XD segments, it took about 17 hours to process the 3-year raw data, execute the bottleneck identification algorithm, and output performance measures for developing Tableau visualization. The data processing and calculation for statewide data took about 50 hours.

Compared to the Northern Virginia District and Culpeper District case studies, the Tableau file(s) for the statewide case study (Figures 11 through 15) was updated as follows:

- *Rather than producing one Tableau file for all the results, different districts were grouped together to reduce the dataset sizes and improve response speeds to user actions.* Four district-wide files and one statewide file were created as described in Table 1. All of the district-wide files presented ranking results for both the intersections (Figure 11) and the approaches. They all also contained options for before-after analyses (Figure 12), observation of trends in performance measures over time (Figure 13), and a map view (Figure 14). Statewide files included only intersection ranking in order to manage file size.

- *All of the Tableau files (with the district files containing the district codes) were uploaded to the Commonwealth of Virginia (COV) server.* The URLs are shown in Table 1.

- *As with the previous case studies, these Tableau files contained notes tabs to provide user instructions; filters for users to select study sites, time period, and performance measures; heatmaps (Figure 15) to show the intensity and variability of select measures across day of week, month, and year; and detailed metadata (including maps).*

- *Because of the summarization of the results, CDFs were not meaningful.* The dashboard of various tabs used in the previous case studies was also removed because the details could not be discerned on standard issue VDOT laptop screens.

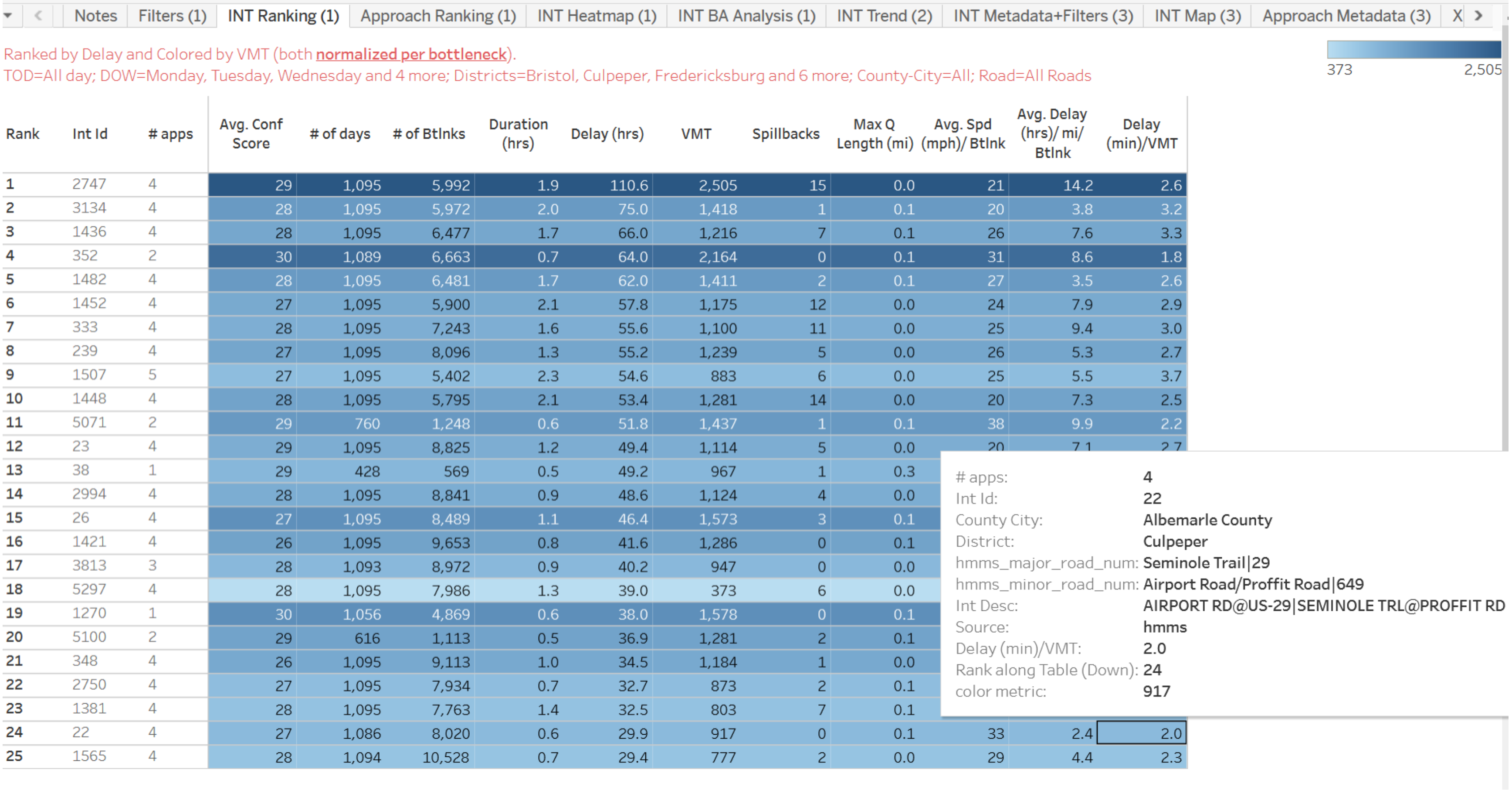
- *A number of smaller improvements were also made, including (1) the ability to normalize performance measures by number of bottlenecks or number of days, (2) the ranking and colorizing of the results tables by different measures, and (3) the moving of a number of intersection/approach descriptions (or attributes) to the tooltip, improving the readability of the tables on a standard laptop screen.*

- *Since project-related data were not available, preset before and after periods could not be specified for analyses.* Instead, users were given the option to select custom year-month values for each period.

Table 1. District and Statewide Results Files and Sizes

Group No.	District Code	District Name	No. of Signals Identified	Tableau Link	File Size (MB)		
					Signal	Approach	Tableau
1	1	Bristol	88	https://tableau.cov.virginia.gov/#/views/ArtBtlnx12368_v3/Notes	40	83	45
	2	Salem	284				
	3	Lynchburg	161				
	6	Fredericksburg	194				
	8	Staunton	280				
2	4	Richmond	692	https://tableau.cov.virginia.gov/#/views/ArtBtlnx47_v3/Notes	32	73	41
	7	Culpeper	131				
3	5	Hampton Roads	1,023	https://tableau.cov.virginia.gov/#/views/ArtBtlnx5_v3/Notes	39	86	45
4	9	Northern Virginia	1,624	https://tableau.cov.virginia.gov/#/views/ArtBtlnx9_v3/Notes	65	146	74
5	N/A	All districts	4,477	https://tableau.cov.virginia.gov/#/views/ArtBtlnx_state_v3/Notes	177	N/A	66

N/A = not applicable.

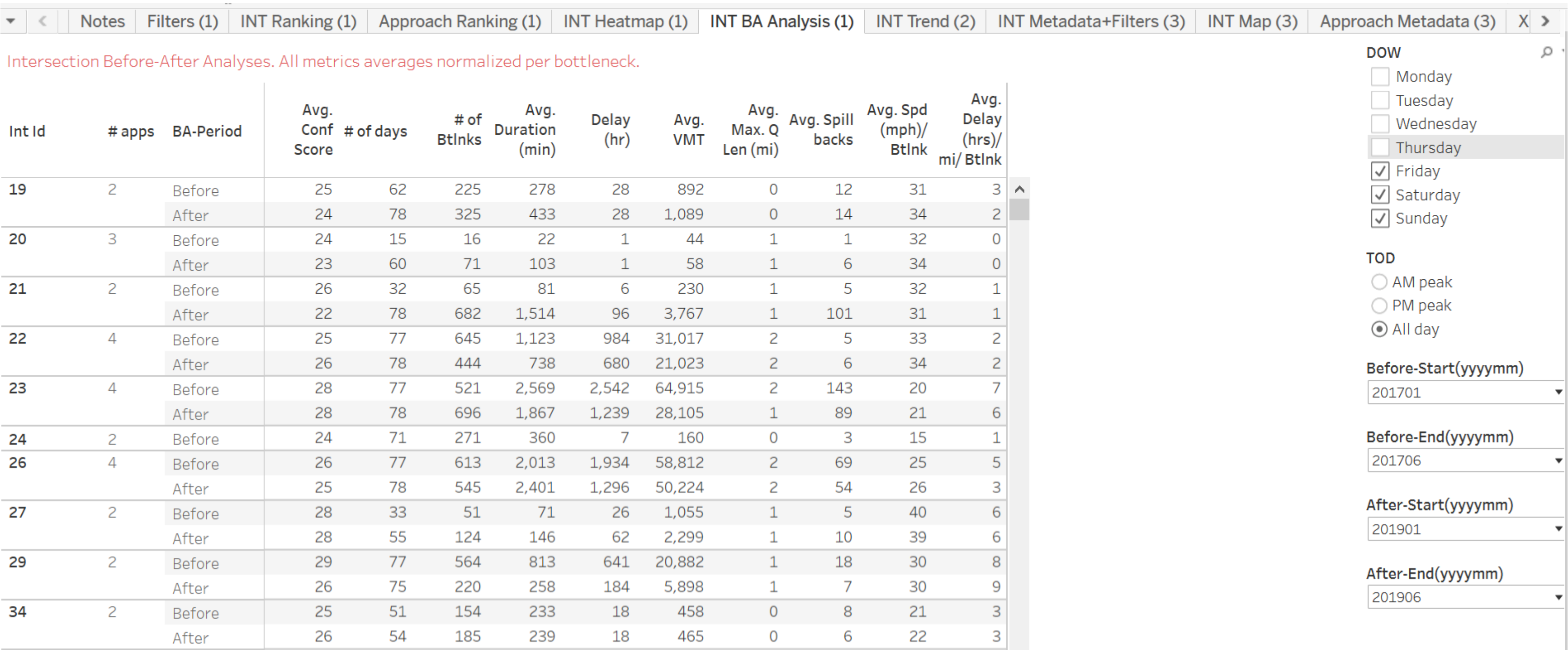


Ranked by Delay and Colored by VMT (both normalized per bottleneck).
TOD=All day; DOW=Monday, Tuesday, Wednesday and 4 more; Districts=Bristol, Culpeper, Fredericksburg and 6 more; County-City=All; Road=All Roads

Rank	Int Id	# apps	Avg. Conf Score	# of days	# of Btlnks	Duration (hrs)	Delay (hrs)	VMT	Spillbacks	Max Q Length (mi)	Avg. Spd (mph)/ Btlnk	Avg. Delay (hrs)/ mi/ Btlnk	Delay (min)/VMT
1	2747	4	29	1,095	5,992	1.9	110.6	2,505	15	0.0	21	14.2	2.6
2	3134	4	28	1,095	5,972	2.0	75.0	1,418	1	0.1	20	3.8	3.2
3	1436	4	28	1,095	6,477	1.7	66.0	1,216	7	0.1	26	7.6	3.3
4	352	2	30	1,089	6,663	0.7	64.0	2,164	0	0.1	31	8.6	1.8
5	1482	4	28	1,095	6,481	1.7	62.0	1,411	2	0.1	27	3.5	2.6
6	1452	4	27	1,095	5,900	2.1	57.8	1,175	12	0.0	24	7.9	2.9
7	333	4	28	1,095	7,243	1.6	55.6	1,100	11	0.0	25	9.4	3.0
8	239	4	27	1,095	8,096	1.3	55.2	1,239	5	0.0	26	5.3	2.7
9	1507	5	27	1,095	5,402	2.3	54.6	883	6	0.0	25	5.5	3.7
10	1448	4	28	1,095	5,795	2.1	53.4	1,281	14	0.0	20	7.3	2.5
11	5071	2	29	760	1,248	0.6	51.8	1,437	1	0.1	38	9.9	2.2
12	23	4	29	1,095	8,825	1.2	49.4	1,114	5	0.0	20	7.1	2.7
13	38	1	29	428	569	0.5	49.2	967	1	0.3			
14	2994	4	28	1,095	8,841	0.9	48.6	1,124	4	0.0			
15	26	4	27	1,095	8,489	1.1	46.4	1,573	3	0.1			
16	1421	4	26	1,095	9,653	0.8	41.6	1,286	0	0.1			
17	3813	3	28	1,093	8,972	0.9	40.2	947	0	0.0			
18	5297	4	28	1,095	7,986	1.3	39.0	373	6	0.0			
19	1270	1	30	1,056	4,869	0.6	38.0	1,578	0	0.1			
20	5100	2	29	616	1,113	0.5	36.9	1,281	2	0.1			
21	348	4	26	1,095	9,113	1.0	34.5	1,184	1	0.0			
22	2750	4	27	1,095	7,934	0.7	32.7	873	2	0.1			
23	1381	4	28	1,095	7,763	1.4	32.5	803	7	0.1			
24	22	4	27	1,086	8,020	0.6	29.9	917	0	0.1	33	2.4	2.0
25	1565	4	28	1,094	10,528	0.7	29.4	777	2	0.0	29	4.4	2.3

apps: 4
Int Id: 22
County City: Albemarle County
District: Culpeper
hmms_major_road_num: Seminole Trail|29
hmms_minor_road_num: Airport Road/Proffit Road|649
Int Desc: AIRPORT RD@US-29|SEMINOLE TRL@PROFFIT RD
Source: hmms
Delay (min)/VMT: 2.0
Rank along Table (Down): 24
color metric: 917

Figure 11. Culpeper and Richmond Districts' Tableau Visualizations for Intersection Ranking



Intersection Before-After Analyses. All metrics averages normalized per bottleneck.

Int Id	# apps	BA-Period	Avg. Conf Score	# of days	# of Btlnks	Avg. Duration (min)	Delay (hr)	Avg. VMT	Avg. Max. Q Len (mi)	Avg. Spill backs	Avg. Spd (mph)/ Btlnk	Avg. Delay (hrs)/ mi/ Btlnk
19	2	Before	25	62	225	278	28	892	0	12	31	3
		After	24	78	325	433	28	1,089	0	14	34	2
20	3	Before	24	15	16	22	1	44	1	1	32	0
		After	23	60	71	103	1	58	1	6	34	0
21	2	Before	26	32	65	81	6	230	1	5	32	1
		After	22	78	682	1,514	96	3,767	1	101	31	1
22	4	Before	25	77	645	1,123	984	31,017	2	5	33	2
		After	26	78	444	738	680	21,023	2	6	34	2
23	4	Before	28	77	521	2,569	2,542	64,915	2	143	20	7
		After	28	78	696	1,867	1,239	28,105	1	89	21	6
24	2	Before	24	71	271	360	7	160	0	3	15	1
26	4	Before	26	77	613	2,013	1,934	58,812	2	69	25	5
		After	25	78	545	2,401	1,296	50,224	2	54	26	3
27	2	Before	28	33	51	71	26	1,055	1	5	40	6
		After	28	55	124	146	62	2,299	1	10	39	6
29	2	Before	29	77	564	813	641	20,882	1	18	30	8
		After	26	75	220	258	184	5,898	1	7	30	9
34	2	Before	25	51	154	233	18	458	0	8	21	3
		After	26	54	185	239	18	465	0	6	22	3

DOW

☐ Monday
☐ Tuesday
☐ Wednesday
☐ Thursday
☑ Friday
☑ Saturday
☑ Sunday

TOD

○ AM peak
○ PM peak
● All day

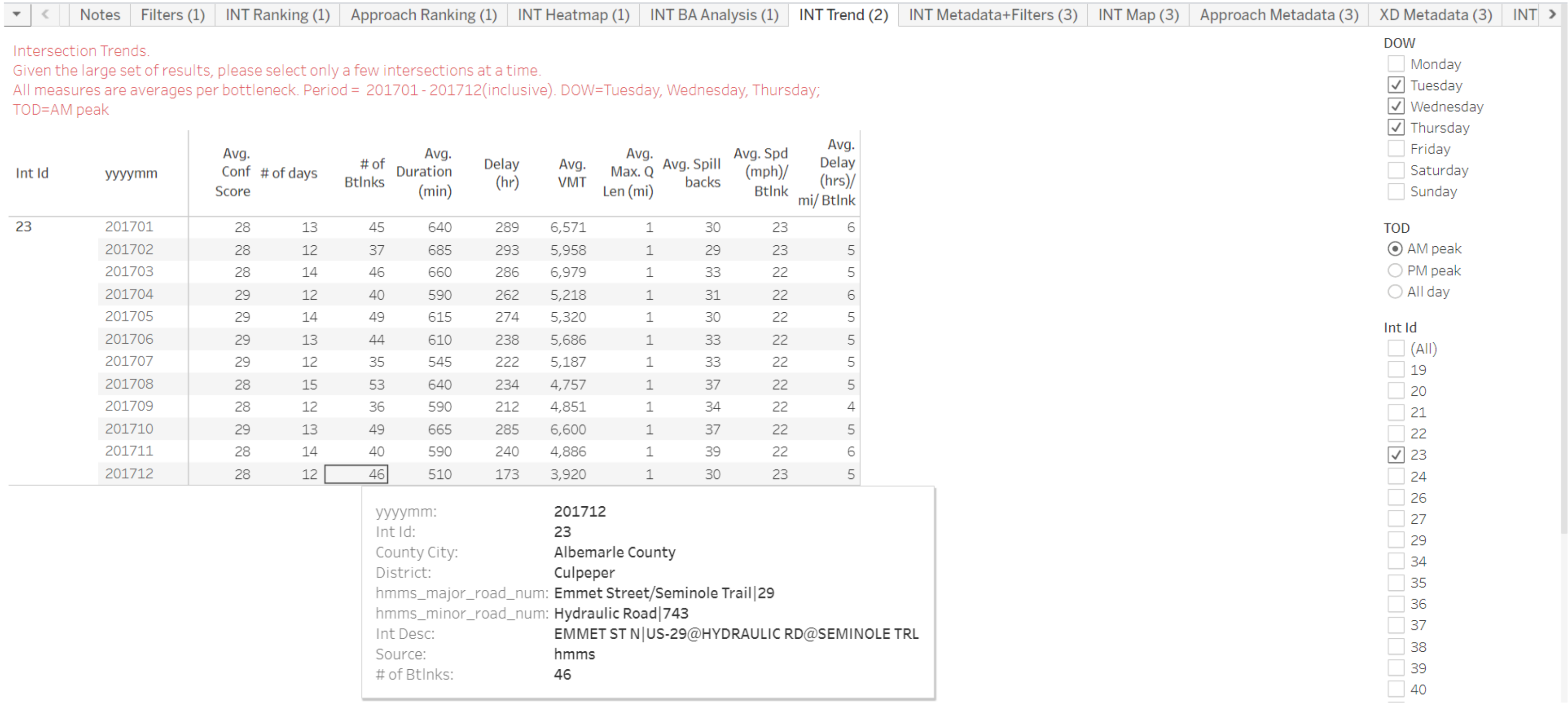
Before-Start(yyyymm) 201701

Before-End(yyyymm) 201706

After-Start(yyyymm) 201901

After-End(yyyymm) 201906

Figure 12. Culpeper and Richmond Districts' Tableau Visualizations for Before-After Analyses



Intersection Trends.
Given the large set of results, please select only a few intersections at a time.
All measures are averages per bottleneck. Period = 201701 - 201712(inclusive). DOW=Tuesday, Wednesday, Thursday;
TOD=AM peak

Int Id	yyyymm	Avg. Conf Score	# of days	# of Btlnks	Avg. Duration (min)	Delay (hr)	Avg. VMT	Avg. Max. Q Len (mi)	Avg. Spill backs	Avg. Spd (mph)/ Btlnk	Avg. Delay (hrs)/ mi/ Btlnk
23	201701	28	13	45	640	289	6,571	1	30	23	6
	201702	28	12	37	685	293	5,958	1	29	23	5
	201703	28	14	46	660	286	6,979	1	33	22	5
	201704	29	12	40	590	262	5,218	1	31	22	6
	201705	29	14	49	615	274	5,320	1	30	22	5
	201706	29	13	44	610	238	5,686	1	33	22	5
	201707	29	12	35	545	222	5,187	1	33	22	5
	201708	28	15	53	640	234	4,757	1	37	22	5
	201709	28	12	36	590	212	4,851	1	34	22	4
	201710	29	13	49	665	285	6,600	1	37	22	5
	201711	28	14	40	590	240	4,886	1	39	22	6
	201712	28	12	46	510	173	3,920	1	30	23	5

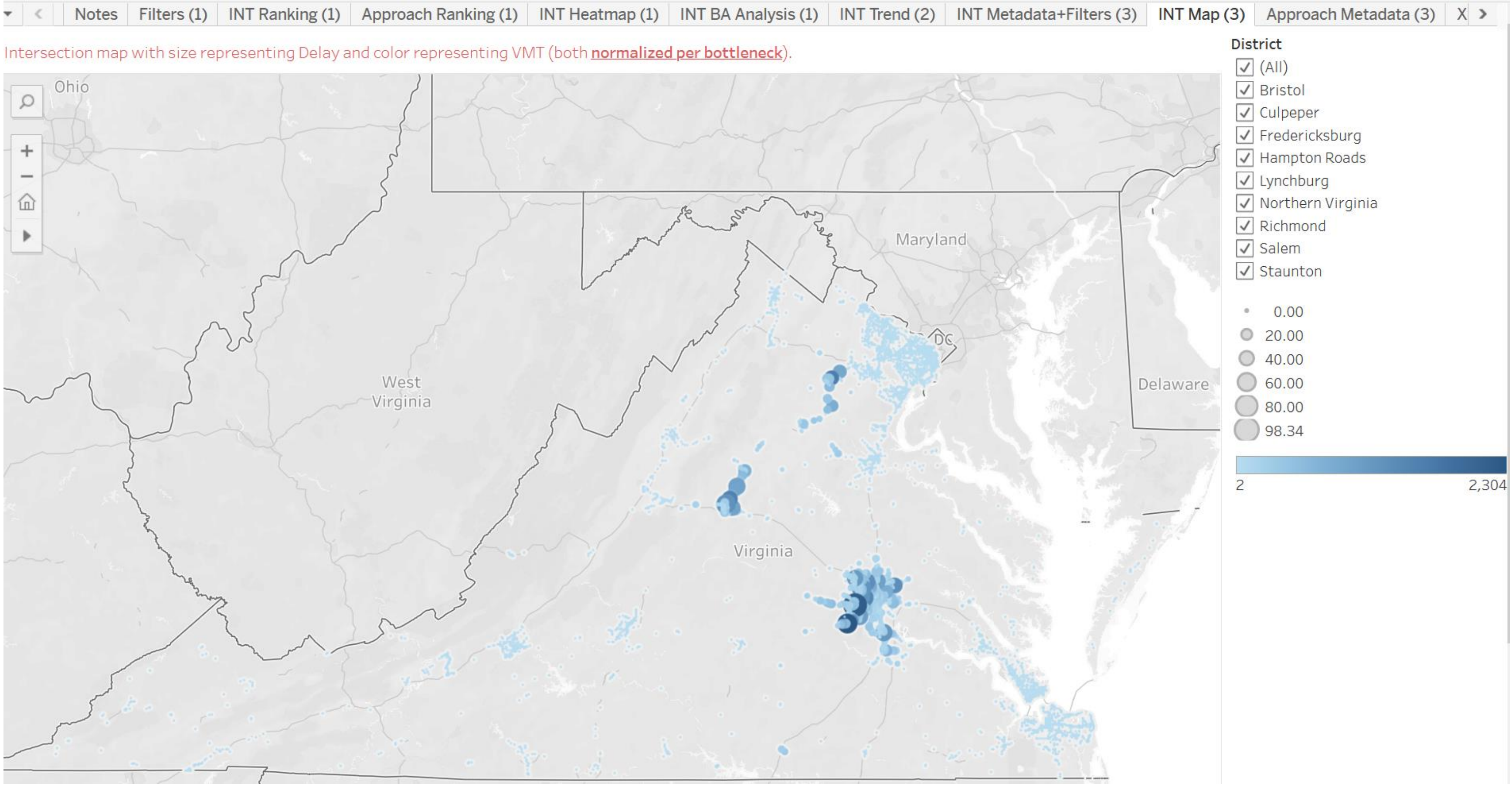
yyyymm:	201712
Int Id:	23
County City:	Albemarle County
District:	Culpeper
hmms_major_road_num:	Emmet Street/Seminole Trail|29
hmms_minor_road_num:	Hydraulic Road|743
Int Desc:	EMMET ST N|US-29@HYDRAULIC RD@SEMINOLE TRL
Source:	hmms
# of Btlnks:	46

DOW
☐ Monday
☑ Tuesday
☑ Wednesday
☑ Thursday
☐ Friday
☐ Saturday
☐ Sunday

TOD
● AM peak
○ PM peak
○ All day

Int Id
☐ (All)
☐ 19
☐ 20
☐ 21
☐ 22
☑ 23
☐ 24
☐ 26
☐ 27
☐ 29
☐ 34
☐ 35
☐ 36
☐ 37
☐ 38
☐ 39
☐ 40

Figure 13. Culpeper and Richmond Districts' Tableau Visualizations for Intersection Trend

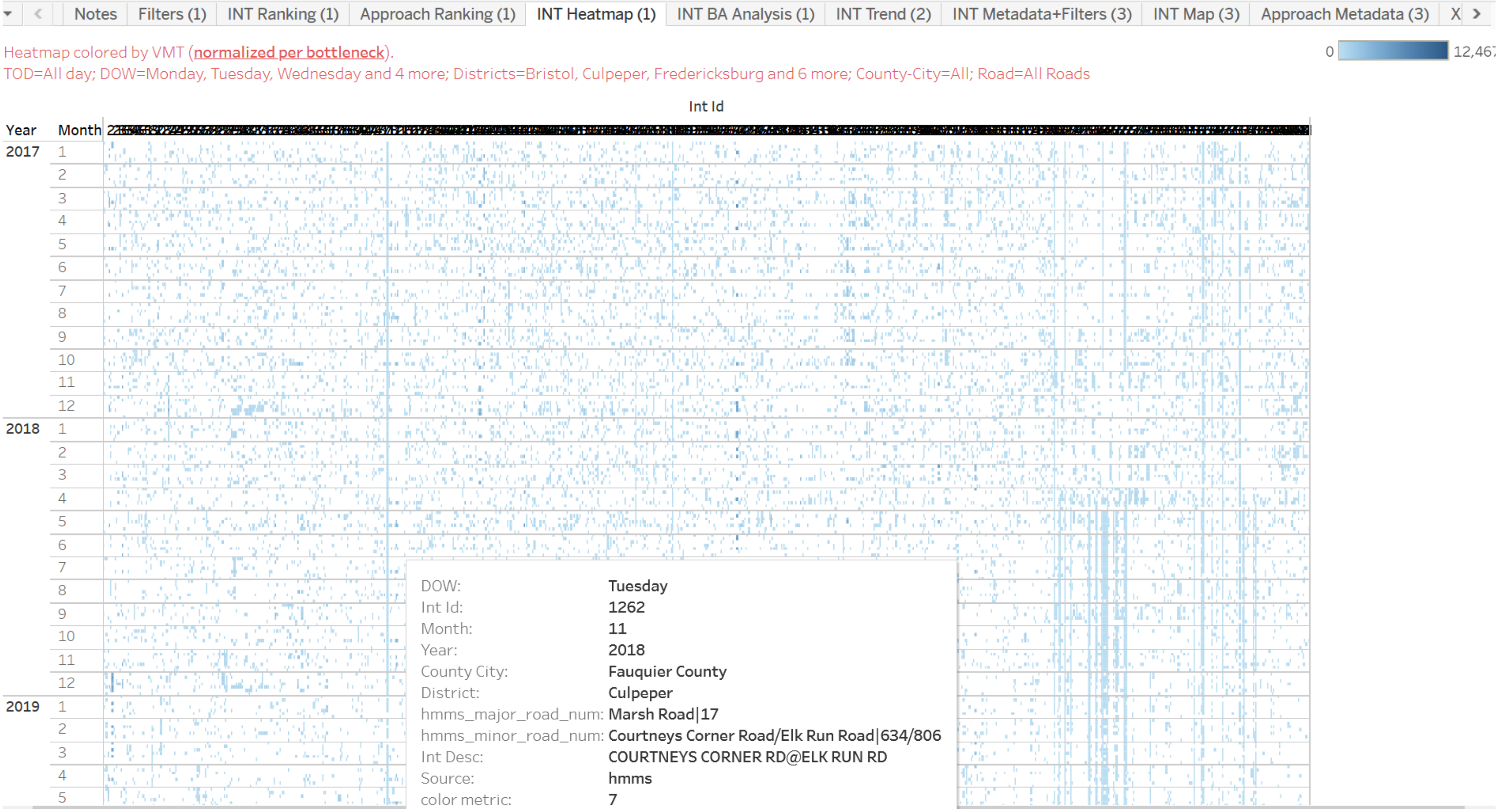


Intersection map with size representing Delay and color representing VMT (both **normalized per bottleneck**).

District

☑ (All)
☑ Bristol
☑ Culpeper
☑ Fredericksburg
☑ Hampton Roads
☑ Lynchburg
☑ Northern Virginia
☑ Richmond
☑ Salem
☑ Staunton

Figure 14. Culpeper and Richmond Districts' Tableau Visualizations for Intersection Map With Delay and VMT. VMT = vehicle miles traveled.



Heatmap colored by VMT (normalized per bottleneck).
TOD=All day; DOW=Monday, Tuesday, Wednesday and 4 more; Districts=Bristol, Culpeper, Fredericksburg and 6 more; County-City=All; Road=All Roads

DOW: Tuesday
Int Id: 1262
Month: 11
Year: 2018
County City: Fauquier County
District: Culpeper
hmms_major_road_num: Marsh Road|17
hmms_minor_road_num: Courtneys Corner Road/Elk Run Road|634/806
Int Desc: COURTNEYS CORNER RD@ELK RUN RD
Source: hmms
color metric: 7

Figure 15. Culpeper and Richmond Districts' Tableau Visualizations for Intersection Heatmaps. The color represents the performance measure (average VMT per bottleneck in this example) of intersection bottlenecks; each column shows the average VMT across day of week, month, and year for one intersection, and each row shows the average VMT at individual intersections during the same time interval. VMT = vehicle miles traveled.

Task 7: Phase 2—Review and Validate Results

Statewide Bottleneck Rankings and Before-After Analysis Results

The statewide case study results were reviewed by members of the overall project TRP. They were generally pleased with the implementation efforts, the tools, and the results. They mentioned that they saw several benefits from and use cases for these tools and results, including the following:

- SMART Scale applications
- understanding and communication of the mobility impacts of land use changes
- identification of intersections for projects
- before-after studies
- tools for localities maintaining their own signals to identify and prioritize intersections for operational improvements.

The names of all of the intersection approaches and the major/minor approach delineations were not available in the input data for all intersections. So the TRP asked the research team to present any available details in a list and on the map and make the intersections searchable. All of these features were subsequently added in Tableau, and for all of the HMMS intersections, the major and minor roads were added to the metadata.

Thinking of the next steps for continued implementation, the TRP asked the research team to explain the resource needs for annual updates of the results. They were informed that the annual updates include conflating the INRIX XD data (to newer map versions), downloading the INRIX data, analyzing the bottlenecks, updating the Tableau files, and performing any other incidental work (such as updating of the code and metadata; data quality checks; resolution of software, driver, hardware, and network issues; documentation and communication; etc.). Other comparable annual agency efforts include AADT publication, crash tools updates, and Potential for Safety Improvement (PSI) Top 100 lists. The research team believes that all of these efforts take several weeks or months to complete and often involve several staff members. More details on resource needs such as computing hardware and time needs are presented at the end of this subsection.

For longer term research and implementation considerations, the TRP shared the following feedback:

- VDOT has an ongoing interest to create super nodes for physical intersections from the existing Linear Reference System (LRS) nodes, which can be used for this ranking and analysis tool. However, that work is currently in progress.

- There needs to be more discussion within VDOT, perhaps with the Learning Center, about the need for more data munging and analysis skill sets, such as using SQL, Python, etc., and the use of tools such as Hadoop and Google Big Query for analyzing very large datasets.

Conflation Scripts and Data

The data conflation scripts significantly improved the efficiency of data conflation. Instead of about 6,000 person-hours for the statewide conflation at the rate of 1 person-hour per intersection for manual conflation, the conflation scripts took just about 8 person-hours. This time was needed for data pre-processing and post-processing and for changing the scripts to perform each computer run. It took about 3 hours to conflate all of the XD segments to the AADT data on a standard issue VDOT computer. Processing large data files without such additional pre-processing and post-processing efforts requires high computing power. It should be noted that the scripts are not perfect in producing results, mainly because of missing data and some inherent challenges in processing spatial features. These limitations and some potential solutions are outlined in the Appendix.

Missing data was a common issue for input data including AADT and HMMS from VDOT, intersection data from OSM, and speed data from INRIX. AADT were not available for 8,875 XD segments matched to intersection approaches, so these segments were not included in the bottleneck analysis. Based on the limited assessment in a small area, these segments with no AADT data were mostly on minor roads where the probability of bottlenecks is generally deemed to be very low. As a result, the impact of excluding these segments was minimal for the bottleneck analysis results. Performing a comprehensive data validation was outside the scope of this proof-of-concept implementation.

A complete dataset for all signalized intersections in the state did not exist at the time of this study. Although the HMMS is a good source for VDOT signals, data for municipality signals were not available. Currently, VDOT districts do not have such information for their jurisdictions, but some districts such as the Salem District are building that dataset (Nathan O’Kane, personal communication, June 2020). This study used OSM to locate municipality signals but the OSM signal layer was not up to date and no known large scale validations have been done on this dataset.

The semiannual updates of INRIX XD maps require tracking the changes of XD segments and updating the conflation results to make sure that the bottleneck ranking results for each intersection are comparable over time.

Considering the large speed dataset size (150 GB for 3 years of data), infrastructure planning for data storage and computing are necessary for ongoing field implementation of statewide bottleneck ranking. Cloud storage and computing would be a potential solution. Because of the complexity of the entire process, updating the bottleneck ranking and before-after study datasets would require a good knowledge of the input data and coding skills. Based on the experience in this case study, approximately 3 to 5 days of a full-time equivalent is estimated to be needed for annual updates including collecting data, processing data and computing performance measures, updating Tableau files, and performing any other incidental work such as updating of the Python scripts and metadata, data quality checks, resolution of software issues, documentation, and communication. This is based on the assumption that most data conflation work is completed during the initial implementation. The annual updates will fix network changes and add new analysis results, so they will be much less labor-intensive.

CONCLUSIONS

- *VDOT practitioners indicated that the rural and statewide arterial intersection bottleneck ranking results were generally acceptable.* Some data limitations such as coverage, availability, and quality were more pronounced for the rural intersections analyzed in this Phase 1 study compared to the original study; however, data quality thresholding improved some results. Although the reduced coverage and data availability as compared to urban areas are concerns, the results were still deemed very useful for sketch planning, network screening, and communication with stakeholders in other agencies. Normalized performance measures such as delay per VMT or delay per mile were deemed more appropriate for rural intersection ranking. Delay measures of intersections with and without data missing on some minor approaches were often comparable.

- *VDOT practitioners found that before-after bottleneck analysis results were generally acceptable and useful.* Field staff currently do not have any other tools to evaluate and communicate quickly the mobility changes at intersections stemming from projects. These analyses results will complement the before-after safety analyses already being conducted for major projects and will be especially useful in the case of innovative, alternative intersections.

- *Statewide conflation results, arterial intersection bottleneck rankings and before-after analysis results, and the developed tools are very promising and useful.* The TRP is considering the next steps for disseminating these results and tools to field personnel. The conflation results contain known errors and omissions, mainly because of the issues in the underlying datasets, including missing data and differences in base map layers. However, the final analysis results are still considered largely useful for the following reasons:

  - The major bottlenecks often occur at high traffic intersections, which are more likely to be well defined on all approaches and are covered by both XD speed data and AADT data.

  - Since the probability of bottlenecks occurring on minor approaches is low, the lack of minor road AADT at some intersections did not significantly affect the prioritization results.

- *The data conflation tool is promising and widely useful to many VDOT business analyses.* The automation of data conflation using this tool can significantly reduce the manual effort that is often needed for many studies where data must be conflated, and the tool can achieve at least 75 percent accuracy based on qualitative estimation. Multiple VDOT divisions and OIPI are very interested in using and further developing this tool for their business needs.

ESTIMATED BENEFITS

Two distinct types of benefits are estimated from this implementation effort, namely:

1. benefits from the results, code, and visualization tools
2. benefits from the conflation tool.

Benefits From the Results, Code, and Visualization Tools

TRP members have repeatedly mentioned to the research team that these results will be very useful for them. Three years (2017-2019) of analysis results were made available to VDOT and OIPI for all known intersections across the state that can be used immediately for various potential applications. Applications mentioned by the TRP included the following:

- SMART Scale applications
- understanding and communication of the mobility impacts of land use changes
- identification of intersections for projects
- before-after studies
- tools for localities maintaining their own signals to identify and prioritize intersections for operational improvements.

Such studies either were not performed in the past, missing out on the quantification and clear communication of benefits of projects, or were time-consuming to conduct. The ready availability of these results will support field staff to perform quick analyses of mobility performance at intersections. The code and the visualization tools, along with the documentation in the form of reports, form a proof of concept for performing such analyses and also provide estimates of the effort, skill sets, time, and computing resource needs. Many more details of the issues concerning the underlying input datasets are also better understood and documented, and resolved in some instances.

Benefits From the Conflation Tool

Several VDOT field experts have also been waiting eagerly for the conflation tool to become available. The line to line conflation tool is useful for conflating any two line shapefiles, including the VDOT LRS, INRIX TMC (Traffic Message Channel), the National Performance Measurement Research Data Set (NPMRDS), newer versions of XD, etc., with other datasets such as AADT, PSL, etc. Even with the various concerns and issues identified with the underlying datasets resulting in reduced accuracy or completeness of the conflation tool results, the automation of conflation of even 75 percent (qualitative estimation) of the links or segments with greater than 75 percent accuracy (qualitative estimation) will significantly reduce the amount of manual conflation that is often needed for many studies. As a number of these maps evolve over time, analysts have traditionally had to redo the conflation every time a map is updated.

The research team estimates it will take about 45 minutes to conflate manually all necessary data for an intersection and about 15 minutes to validate the conflation results from the tool. With these time estimates, manual conflation would require about 4,000 person-hours to conflate and validate all intersections in the state. With the conflation tool, it is estimated that 75 percent of the intersections will need only to be validated whereas the remaining 25 percent will have to be manually conflated and validated. As a result, the total required time would be 1,750 person-hours of staff time using the conflation tool, which constitutes more than a 50 percent reduction in labor time. This represents significant cost and labor savings attributable to the use of the conflation tool.

ACKNOWLEDGMENTS

The authors thank the TRP members (Mena Lockwood of VDOT's Traffic Engineering Division; Chad Tucker of OIPI; Troy Austin, Charles Proctor, and Rowes Hanna of VDOT's Culpeper District; and Chien-Lun Lan of VTRC); Andrew Pike, Beverly Quinlan, Chris Wichman, Jitender Ramchandani, Jonathan Robbins, and Margit Ray of OIPI; Hamlin Williams, In-Kyu Lim, Mike Clements, Ralph Jones, Sanhita Lahiri, Simona Babiceanu, and Wenling Chen of VDOT's Traffic Engineering Division; Jungwook Jun of VDOT's Traffic Mobility and Planning Division; Katie McCann of VDOT's Operations Division; Krishna Singh of Performance Strategic Planning; Mike Fontaine of VTRC; Nathan O'Kane of Strategic Innovation; Hong Yang and Qingyu Ma of ODU; and the VDOT district traffic engineers for supporting the study with ideas and brainstorming, data access and knowledge, feedback, and coding.

REFERENCES

1. Zhao, M., and Venkatanarayana, R. *Improving the Identification and Characterization of Arterial Congestion Bottlenecks.* VTRC 19-R20. Virginia Transportation Research Council, Charlottesville, 2019.

2. Hale, D.K; Hajbabaie, A., Ma, J., Hu, J., Park, H., and Bared, J. Proposed Data-Driven Performance Measures for Comparing and Ranking Traffic Bottlenecks. Presented at the 96th Annual Meeting of the Transportation Research Board, Washington, D.C., January 8-12, 2017.

3. Yang, H., and Ma, Q. Arterial Bottleneck Analysis Implementation Phase 2a: Development of Conflation Tool. Unpublished report. Norfolk, Va., July 2020.

4. INRIX. Roadway Analytics. https://analytics.inrix.com/roadway_analytics. Accessed September 11, 2019.

5. Virginia Department of Transportation. Project Pool Website. http://pool/Projects.aspx. Accessed September 11, 2019.

6. Virginia Department of Transportation. VDOT HMMS Traffic Signals Shapefile. https://www.virginiaroads.org/datasets/hmms-trafficsignals-flat. Accessed September 6, 2020.

7. OpenStreetMap. Virginia Shapefiles. https://download.geofabrik.de/north-america/us/virginia.html. Accessed September 7, 2020.

APPENDIX
DETAILED REVIEW NOTES ON CONFLATION SCRIPTS

The following observations from the research team on the performance of the scripts for statewide conflation are based on prior experience with the underlying datasets and manual conflations performed for the Northern Virginia District and Phase 1 Culpeper District case studies. It should be noted that (1) many of these limitations were identified during the analysis phase, after the development contract had ended; many other concerns identified during the development phase were addressed within the contract; (2) many of these limitations can be improved with future work by improving the underlying datasets; and (3) these limitations affect less than 10 percent of the total intersections, based on the limited, qualitative assessment; comprehensive evaluation of the results for the entire state was not feasible. Although the conflation tool has some limitations, it should be emphasized that it provides significant efficiency improvements for conflating the data for most intersections.

- The algorithm cannot recognize nearby parallel frontage roads or on-ramps near the intersection and process them as an intersection approach, mainly because such metadata often do not exist. The ability of the algorithm to separate different approaches depends mainly on the physical distance between the facilities and the buffer radius in use.

- The line to point conflation algorithm cannot handle complicated geometry such as grade separation, roundabouts, etc. For a two-dimensional spatial analysis, grade-separated roads and intersections look the same. Roundabouts are not identified as stopping points for approaches currently. So for any intersection approaches with a roundabout upstream within 1.5 miles from the intersection center, the algorithm considers the roundabout as part of an intersection approach. An example is shown in Figure A1 for the intersection of N. 7th Street and E. Duval Street in Richmond. The eastbound approach is highlighted. As the algorithm keeps searching for upstream segments along the polyline, the algorithm loops around the roundabout until the total lengths searched reach 1.5 miles, resulting in duplicate XD segments on an approach (this duplication can be easily fixed in the future).

- The algorithm cannot identify diverging points, which results in the same XD segments being assigned to different approaches at one intersection. Figure 11 shows the intersection on Richmond Road in Charlottesville near the I-64 Exit 124 interchange. The algorithm considered the ramp (highlighted in Figure A2) as one of the approaches of the intersection and searched for upstream segments onto Richmond Road. Further, those segments on Richmond Road upstream of the ramp were included on both the ramp approach and the Richmond Road eastbound approach, resulting in the same XD segments considered as part of two different approaches to the same intersection. Although the algorithm cannot figure out which of these duplicated XD segments are the right conflations for the final results, the algorithm has to be improved in the future to flag such duplicates. Further, programmatic identification of the diverging points can help the algorithm to stop the ramp approach at the diverging point.

Figure A1. Intersection of N. 7th Street and E. Duval Street in Richmond

Figure A2. Intersection of Richmond Road and I-64 Ramp Near Exit 124 Interchange

- The algorithm currently uses a fixed threshold of 1.5 miles if there is no signalized intersection upstream within this distance. Figure A3 shows an example for an intersection (at Huntington Avenue and 49th Street in Newport News) where such an approach is not appropriate. The upstream signal was 1.4 miles away. The algorithm included all segments toward 49th Street westbound between the two intersections (highlighted segments in Figure 12) as one approach of this intersection, which was far beyond the intersection impact area. As the XD segment definitions improve over time to more logical blocks of traffic and if more information about side streets is available, this limitation can be better addressed.

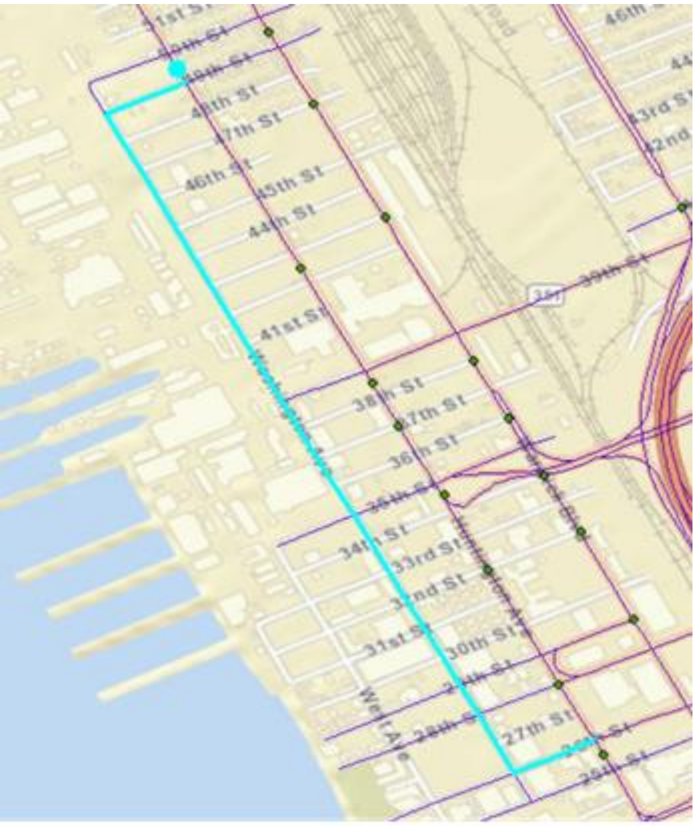


Figure A3. Intersection of Huntington Avenue and |49th Street in Newport News

- The algorithm cannot identify the intersection approaches that are not in XD coverage. Comparison between the XD map and a reference map with all roadways is needed to achieve this function. The research team tried to use VDOT LRS, US Census TIGER (Topologically Integrated Geographic Encoding and Referencing) road files, and OSM road files as a reference to identify the approaches that were not in the XD coverage. However large differences between the XD map and reference maps precluded such work. Figure A4 shows the coverage of VDOT LRS and XD for two intersections. Even the polylines representing the same road on different base maps do not overlap completely, making the comparison of the polylines within an intersection buffer area difficult or not meaningful. If the roadway polylines perfectly matched, it would still be difficult to identify automatically the approaches without XD coverage because of the limited information in the maps and the large differences of roadway geometries. The points representing the intersections on the map are also not always at the center of the intersections. Finally, the roadway geometries are very different at different intersections. So it is not reasonable to apply a fixed threshold for the buffer radius for all the intersections to define the intersection area to identify the roads crossing the intersections. Exploring varying buffer radii is an option for future research.

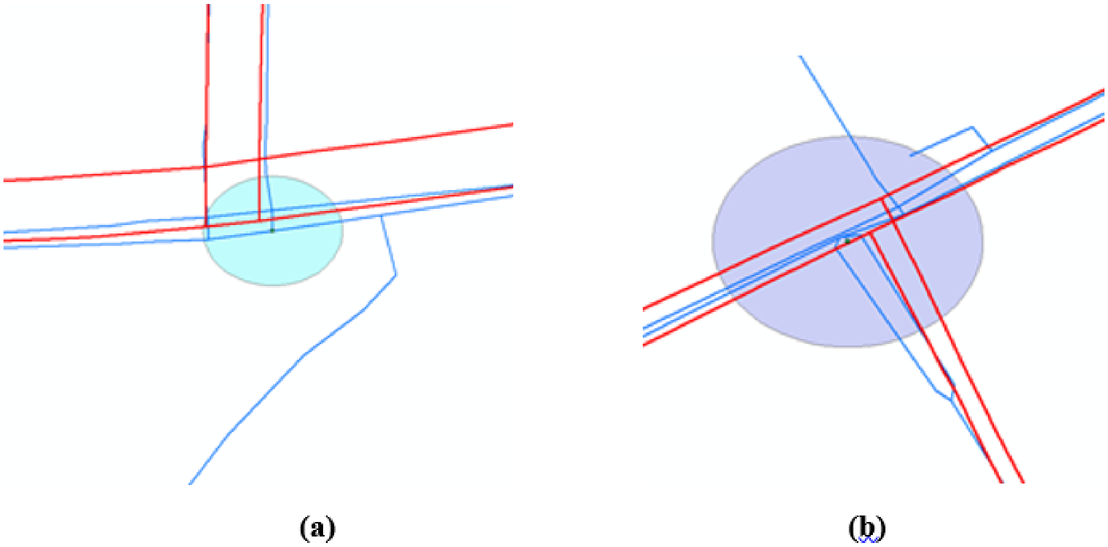


Figure A4. VDOT LRS (blue lines) and XD (red lines) Coverages at Two Different Intersections With Different Buffer Radii: (a) 100 ft, and (b) 200 ft. LRS = Linear Reference System.

- It should be noted that the conflation algorithms use the relative locations between polylines and points. If one layer is based on a geographic coordinate system (GCS) and the other layer uses a local coordinate system, the script would not work properly because the false origin of a local coordinate system can be anywhere on earth. If all input layers use a local coordinate system but the false origin in these systems is different, the tool will not work either. However, a local coordinate system is not used to produce standard maps. All maps published by VDOT use a GCS, as do the third-party maps such as maps from INRIX, NPMRDS (National Performance Measurement Research Data Set), and OSM. So far the research team has not experienced any issues with a coordinate reference system, even though the coordinate reference systems for the different input maps were different. For example, the AADT and PSL maps use GCS_WGS_1984, but INRIX and OSM use WGS_1984 Web Mercator Auxiliary Sphere.