

Examining the Effects of Winter Road Maintenance Operations on Traffic Safety through Visual Analytics

Bryce Hallmark and Jing Dong*, *Senior Member, IEEE*

Abstract—Many past efforts have been exerted towards describing and quantifying the effects of winter maintenance operations on traffic conditions and safety. As highly granular data on snowplow activity become available, many agencies are becoming interested in incorporating these data in their decision-making processes. However, due to its sheer volume, the processing of snowplow automatic vehicle location (AVL) data has been challenging. In addition, adverse weather conditions are usually accompanied by higher crash rates and also correlate with an increase in maintenance operations. Thus, improper model and variable selection can produce misleading results that indicate maintenance operations lead to a higher crash rate. This paper presents simple visualization tools and analysis methods that examine the effects of winter road maintenance operations on traffic safety by combining various data sources including weather, traffic, snowplow AVL, and crash data. Such intuitive tools and results can help agencies better understand the relationship between winter road maintenance activities and traffic safety.

Keywords—Automatic vehicle location (AVL), winter road maintenance, traffic safety, data dashboard

I. INTRODUCTION

Winter road maintenance is a major concern for many state agencies. Over 70 percent of the U.S. population and roadway networks receive 5 or more inches of snow per year. As a result, about 1,300 people die each year and 166,000 injuries occur due to winter-weather-related crashes [1]. Winter road maintenance operations, such as salting and plowing, are some of the most widely adopted methods to combat winter weather events. Resultantly, winter maintenance operations account for approximately 20 percent of the maintenance budget on average for state agencies [1].

Previous studies have quantified the safety benefit of winter road maintenance using maintenance operations data collected by the personnel, such as winter maintenance report logs [2, 3]. Due to the obvious correlation between road weather conditions and winter maintenance activities, however, some conflicting results have been reported. For example, Qin et al. used material-spreading measurements on various Wisconsin highways from snowstorm reports to develop a crash frequency model incorporating both snowplow operations and weather data [2]. Unexpectedly, their model inferred a negative correlation between crashes and de-icing material and a positive correlation between crashes and salting material. Such conflicting results demonstrate the complex interactions amongst winter

weather and traffic conditions and winter maintenance activities, as well as the difficulty of incorporating multiple data sources to generate informative results.

Visual analytics have been widely used to examine the distribution and structure of traffic data for data-driven intelligent transportation systems [4, 5]. Visualization helps investigators better understand the spatial and temporal pattern of large datasets as well as discover hidden patterns and correlations among multiple variables.

To monitor and improve the efficiency of maintenance operations, many agencies have implemented automatic vehicle location (AVL) systems to track the locations and operational status of their snowplow trucks. Visual analytics incorporating crash reports and AVL data along with the highly granular traffic speed and volume data collected from roadway sensors allow us to examine the effects of snowplow activities on traffic safety.

To visualize the interactions amongst weather, traffic, and winter maintenance operations, a data dashboard was developed in this study, allowing decision-makers to examine the spatial and temporal relationships among multiple crash factors. Specifically, focusing on the spatial and temporal proximity of snowplow passes and crashes, the safety benefit of winter road maintenance is explored in this study using statistical analysis combined with visual display. Such visualization tools can help agencies better understand complex datasets and shed light on how to mitigate the adverse effects of winter weather on traffic operations and safety.

II. DATA COLLECTION

This study uses data collected from two winter seasons, namely from November 2016 to May 2017 and from November 2017 to May 2018. The Iowa Department of Transportation (DOT) operates and maintains over 24,000 lane-miles of the primary highway system covering the Interstates, US highways, and Iowa highways across the state. Only data pertaining to these Iowa DOT maintenance routes are analyzed in this study.

A. Automatic Vehicle Location

The Iowa DOT has been collecting fine-grained snowplow operations data since 2011. The snowplow automatic vehicle location system records for each vehicle date and time, longitude and latitude, traveling speed, plow position (up vs. down), and material-spreading rates at approximately a 10-second refresh rate. The Iowa DOT has over 900 snowplow trucks spread throughout 101 garages. Three

types of spreading rates are recorded, namely a solid rate, prewet rate, and liquid rate. Four plow wing positions are recorded, including front plow, left wing, right wing and underbelly plow. Each snowplow truck's capacity is 12,000 lbs. for single-axle trucks and 24,000 lbs. for tandem-axle trucks. Their spreading rate is approximately 200 lbs. per lane mile for solid material and 60 gallons per lane mile for liquids. Their travel speed when plowing and spreading material is about 30 miles per hour and their deadheading speed can be as high as the speed limit.

B. Weather

The weather data used in this study were obtained from the Iowa Environmental Mesonet system. Their Multi-Radar/Multi-Sensor (MRMS) project combines information from many sources and radar systems to generate precise weather information. The weather variables utilized in this study include air temperature, wind speed, hour- and minute-based precipitation, daily snowfall depth, precipitation type, and so forth. The precipitation and precipitation type were recorded at 5-minute intervals.

C. Roadway

This study's roadway information was obtained from the Iowa DOT's Roadway Asset Management System (RAMS). In RAMS, each roadway segment is associated with unique route ID and mile marker (MM) information. RAMS also provides information such as the number of lanes and the route type at specified geographic locations. The Iowa DOT's Linear Reference System (LRS) in conjunction with RAMS was used in this study to merge data from various sources.

D. Traffic

The Iowa DOT has deployed over 900 Wavetronix sensors throughout the state. These sensors collect traffic speed, occupancy and volume data that are archived at 5-minute aggregation intervals. Most of the sensors are located in urban areas. In this study, vehicle speeds were compared to the speed limits (obtained from RAMS) to create a "Relative Speed" variable. That is, the difference between the observed speeds and the speed limit. Negative relative speeds signify slower speeds than the posted speed limit.

E. Crash

This study's crash data were obtained from the Iowa DOT crash database. These crash data include information such as the location, time, crash severity, direction of travel, lighting conditions, and weather conditions at the time of the crash. All crashes that occurred in the presence of a winter weather event were included in the initial analysis. A total of 5,089 winter-weather-related crashes occurred along the Iowa-DOT-maintained roadways from 2016 to 2018. To examine the impact of the snowplow operations, only the crashes that occurred within 2 hours of a snowplow pass were included in the final analysis. In total, 1,372 crashes were linked to snowplow passes via the LRS system.

III. METHODOLOGY

The traffic, weather and snowplow AVL data were linked based on their temporal and spatial attributes. In particular, the LRS assigned a mile reference point to each AVL record, crash location, and Wavetronix sensor location. The mile marker references were used to group relevant data together.

A. Overview of the Prototype Safety Data Dashboard

Using the resulting database, a data dashboard was developed to display the traffic conditions (e.g., speed), weather conditions (e.g., snowfall, intensity, and visibility), and snowplow operations (e.g., solid rate or liquid rate). A snapshot can be seen in Fig. 1. In order to create an intuitive display, a heat map template was employed. For example, to display the speeds measured by the Wavetronix sensors, the heat map is constructed to show the speed difference (relative to the speed limits) at different locations in 5-minute intervals. Each bin of the heat map is populated by a color scheme that describes the relative intensity of that variable. Because the heat map data display is stacked both spatially and temporally, users can observe differences across sensor locations for a single time frame by viewing down a single column or any temporal trend for a single location by viewing the corresponding row across columns. In addition, a dynamic filter process was created. This filter allows for the selection of specific time ranges, time aggregate levels, and roadway networks. This provides an effective way to conduct location- and time-based analysis.

In fact, the data dashboard includes 3 dynamic and identically styled heat maps that are stacked and aligned on top of one another. Each heat map can be set to display any single variable from the combined dataset (such as precipitation, wind speed, or plowing data). This provides users the ability to simultaneously observe 3 different variables for the same location and across the same time window. Additionally, to mitigate visual overload and aid interpretability, these combined heat maps have a dynamic highlight function. When a given time interval is selected, a highlight bar appears across the relevant time interval data on each heat map. This design provides an easy way to identify trends and patterns across various datasets.

In addition, the data dashboard's roadway network map that incorporates the crash data provides a geographic reference for the analysis area. Each crash is color-coded based on its severity. The crash icons in this part of the dashboard are also interactive. By selecting a crash icon, users can open another dashboard window displaying the proximity of snowplow passes to the time of the crash.

In order to help identify storm conditions more easily, a calendar heat map is also available. This uses a similarly styled color scheme as the data dashboard's other heat maps to represent the variables, such as traffic conditions and snowplow operations. This calendar heat map also provides dynamic variable selection that allows users to select the type of events they desire.

B. Data Dashboard Analysis of Snowplow Passes

In addition to visualization, the data was extracted and used to perform statistical analysis to examine the impact of snowplow passes on crash events. For each crash, we calculated the nearest snowplow passes before and after the crash and the total number of snowplow passes within a four-hour window (that is, 2 hours before and 2 hours after the crash). Observing the conditions both before and after the crash time helps identify how the snowplow maintenance operations might have impacted traffic safety.

Because each crash event was associated with a winter storm event, a control group dataset could be created to reflect the a breakdown of the actual storm-wide conditions, or the expected conditions. The conditions present throughout the storm were compiled by starting at the beginning of the storm and then shifting every 30 minutes until the end of the storm, recording the conditions observed at each 30-minute interval. By comparing the conditions at the time of the crash to the conditions present throughout the storm, the differences associated with snowplow operations in particular can be identified.

The difference of proportions test was employed to compare the observed and the expected crash condition datasets. In this test, a significant difference is determined by the resultant Z score in Equation 1 [6]. An absolute value Z-score of 1.96 or higher indicates a significant difference at a 95% confidence level.

$$Z = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\hat{p} * (1 - \hat{p}) * (\frac{1}{n_1} + \frac{1}{n_2})}} \quad (1)$$

The \tilde{p} , or p-tilde, is an adjusted sample proportion and can be calculated following Equation 2 [7].

$$\tilde{p} = \frac{x + 2}{n + 4} \quad (2)$$

where x is the number of bins and n is the total number of samples.

The \hat{p} , or the p-dot, is the pooled proportion between the samples and is determined by Equation 3 [8].

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} \quad (3)$$

The difference of proportions test was applied in this study to identify whether snowplow operations contributed to a higher or lower crash frequency when compared to the expected crash conditions.

IV. RESULTS

A screenshot of the data dashboard is shown in Fig. 1. The righthand side contains the 3 stacked heat maps. Each heat map is broken down into 5-minute aggregate bins, with the rows being the sensor location and the columns being the 5-minute timestamp, and their color representing the corresponding aggregated data. The righthand side of each heat map contains the heat map legend as well as a dropdown window to allow users to select which variable to display. In this screenshot, the time filter is set to 2 days, meaning that a 48-hour window is displayed.

Below the heat maps are the time filter controls. These allow the user to select which day(s) and week(s) are displayed. In addition, adjusting the time range will alter how many hours are displayed (i.e., the total number of 5-minute data points).



Fig. 1. A Data Dashboard Showing the Interactions and Trends Amongst Variables

The Wavetronix sensor locations and crash locations are shown on the map on the left-hand side, with the highway route filters placed below. The blue circles indicate the Wavetronix sensor locations while other shapes and colors depict crash events. For example, the orange triangle indicates locations of the property-damage-only crashes.

This data dashboard tool provides agencies the ability to visually pinpoint the conditions that lead to a breakdown in traffic operations and to observe, record, and evaluate mitigation strategies. This helps agencies get an idea of what prevailing conditions lead to breakdowns in traffic operations without their having to manually sift through and analyze large amounts of data or perform in depth statistical analysis. One observation from this study's data dashboard is the link between visibility and crash counts. In particular, several storms across Iowa were observed to have resulted in a rapid decrease in visibility that led to crashes occurring before snowplows were in operation. Accordingly, proactive steps and timely warnings can be designed to mitigate the adverse impact when visibility drops quickly. Another observation is the high degree of variability in speeds during winter events. Past research has shown that high variability in speeds increases the crash rate [9].

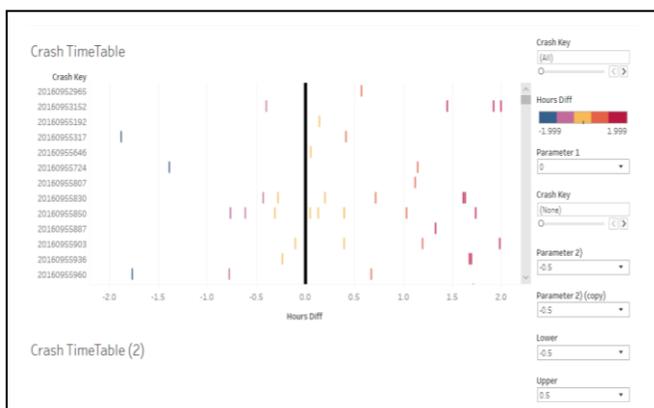


Fig. 2. Snowplow Pass Proximity Visualization Tool

An extension for the data dashboard also provides a way to examine the interactions between snowplow operations and crashes. By clicking on a crash event on the geographic map in the dashboard tool on the left-hand side of Fig. 1 will redirect users to Fig. 2. This is a dashboard of the relevant AVL and crash data. In this dashboard view, each crash event is associated with any AVL data point overlapping its LRS mile reference point. Each row represents a crash and each column is a time bin ranging from 2 hours before to 2 hours after the crash event. A colored bar indicates the presence of a snowplow pass at that time and location, and the black bar represents the time of the crash.

A statistical analysis of the relationship between winter crashes and snowplow activity can be performed. To further examine the impact of snowplow pass proximity on traffic safety in our data, the snowplow passes for each crash were categorized into 30-minute bins, ranging from 2 hours before to 2 hours after the crash, with the column data indicating the nearest snowplow pass before the crash

occurred and the row data the nearest snowplow pass after the crash occurred. (Fig. 3). The proportion of crashes in each category were then computed and displayed in a heat map of the total number of crashes occurring within 2 hours of a snowplow pass (i.e. 1,372 crashes). In Fig. 3, the dark blue areas represent the categories where the highest proportion of crashes occurred, with each number representing the proportion of crash events in that bin. Clearly, the main trend is that most crashes occur within close temporal proximity to a snowplow pass. Specifically, as shown in the top right of the figure, most crashes occur within one hour of a pass either before or after the crash.

To better understand the observed crash frequencies, we

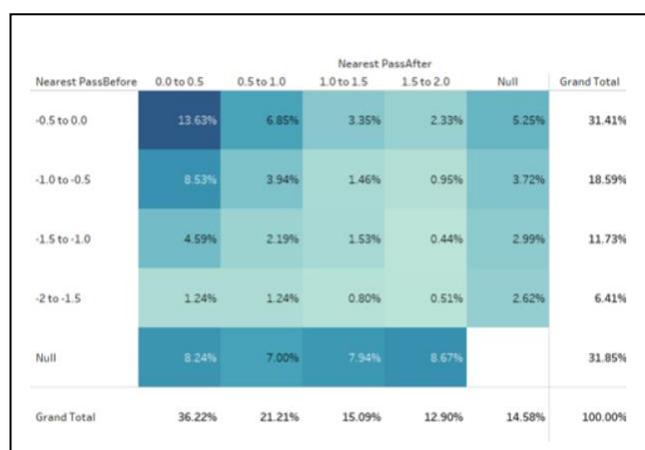


Fig. 3. Nearest Snowplow Pass Proximity Heat Map of Observed Crash Conditions

need to determine whether more crashes occurred in these categories specifically because of snowplow operations or rather because other weather and road conditions associated with crashes are more predominant during winter storms. Therefore, the breakdown of observed crash conditions were compared with the breakdown of stormwide expected crash proportions. That is, by assuming crashes should occur uniformly for the entire storm duration, the storm-wide expected crash proportions can be calculated for each 30-minute bin. This provides a control group with which to compare the breakdown of observed crash proportions for each category bin.

A difference of proportions test was employed to identify the degree to which these crashes coincide with the winter maintenance operations in each dataset. If there were no discernable differences between the observed and the expected crash conditions, this would suggest that snowplow operations had little impact on safety. If, however, the observed crash proportions differ from the expected stormwide crash proportions, this suggests that snowplow operations have either a positive or negative effect on crash frequency, depending on their respective proportions. Fig. 4 shows a heat map of the p-values for each category's Z-score. A p-value below 0.05 is considered statistically significant. In addition, the green shading indicates that there was a greater proportion of expected crashes than the observed crash proportions for that bin. The

red indicates a greater proportion of observed crashes than the expected.

Before \ After (hr)	0.0 to 0.5	0.5 to 1.0	1.0 to 1.5	1.5 to 2.0	No After	Grand Total
0.0 to 0.5	0.017	0.076	0.125	0.876	0.053	0
0.5 to 1.0	0.624	0.657	0.045	0.229	0.035	0.037
1.0 to 1.5	0.442	0.844	0.496	0.098	0.034	0.266
1.5 to 2.0	0.011	0.795	0.857	0.982	0.039	0.004
No Before	0.137	0.005	0	0	NaN	0
Grand Total	0.23	0.984	0.017	0	0	NaN

Fig. 1. Difference of Proportions with Regard to Nearest Snowplow Pass Proximity

Fig. 4 implies that the presence of a snowplow pass results in better safety. For example, the top-left cell that contains the closest temporal pass of snowplows, accounts for 13.6 percent of observed crashes in contrast to the 16.8 percent storm-wide expected proportions. This trend is present for several other categories as well as for the grand totals for close-temporal-proximity plow passes. This suggests that the closer the temporal proximity of plow passes, the less likelihood there is for a crash. Additionally, the “0.0 to 0.5” row has multiple green boxes. This signifies that with before-plow passes, fewer crashes were observed compared to the stormwide expected proportions.

Furthermore, the frequency of snowplow passes before and after the crash time was analyzed. The count of snowplow passes was divided by the number of lanes to obtain a normalized scale, that is, “passes per lane”. The normalized scale allows us to compare various crash locations. A difference of proportions test was deployed, as shown in Fig. 5, with the same color scheme and p-value threshold as described in Fig. 4. The majority of the top row (with no snowplow passes before crashes happen) have a p-value of less than 0.05 and are shaded red. This signifies

that more crashes occurred than expected when comparing the observed crash proportions to the expected storm-wide proportions. Meanwhile, the presence of snowplow pass frequency lowered the likelihood of crash events. This is demonstrated in the far right “Grand Total” column that has a green shading in every row category for a “before” pass frequency. This offers evidence that treating the roads early in a storm might result in safer driving conditions.

V. CONCLUSION

Understanding how winter weather and winter road maintenance operations impact traffic conditions is a goal that many state agencies share. One of the major barriers for state agencies to adopt data visualization is the lack of simple, usable tools [10]. As a result, the creation and implementation of visual analytics tools has become increasingly desired as the transportation industry moves towards data-driven decision-making.

This paper has therefore presented a data dashboard with tools and analysis methods to help agencies better derive from large datasets intuitive results that they can incorporate in their decision-making processes.

Secondly, this paper’s example data dashboard investigation has demonstrated that snowplow activities have a positive impact on traffic safety. The closer the temporal proximity to snowplow passes, the lower the proportion of crashes is. That is, the more snowplow passes that occur before a crash, the lower the proportion of crashes that occur. This suggests an actionable item that agencies can take to mitigate the dangerous winter weather driving conditions, namely deploying road maintenance operations earlier in winter storms and designing traffic control that reflect weather and road conditions and are specifically targeted at the critical locations.

Third, this research incorporates various lessons learned in the analytics of large datasets. Many agencies and organizations have large quantities of data from which they are attempting to derive meaningful results. When the term

Before \ After	0	0.0 to 0.5	0.5 to 1.0	1.0 to 1.5	1.5 to 2.0	2.0 to 2.5	2.5 to 3.0	3.0 to 3.5	3.5 to 4.0	4.0 >	Grand Total
0		<0.000	0	0.027	0	0.001	0.826	0.403	0.192	0.192	0
0.0 to 0.5	0.033	0.046	0.075	0.009	0.071	0.826	0.217	0.67	0.368	0	0
0.5 to 1.0	0.049	0.042	0.213	0.05	0.652	0.035	0.265	0.33	0.74	0.28	0
1.0 to 1.5	0.349	0	0.044	0.127	0.588	0.896	0.143	0.112	0.949	0.67	0
1.5 to 2.0	0.42	0.21	0.002	0.064	0.067	0.254	0.071	0.654	0.885	0.192	0
2.0 to 2.5	0.149	0.233	0.209	0.012	0.616	0.423	0.067	0.524	0.677	0.256	0
2.5 to 3.0	0.89	0.505	0.18	0.035	0.067	0.811	0.58	0.873	0.241	0.524	0.031
3.0 to 3.5	0.252	0.569	0.309	0.289	0.087	0.513	0.368	0.375	0.397	0.411	0.02
3.5 to 4.0	0.721	0.288	0.218	0.267	0.995	0.74	0.808	0.512	0.384	0.784	0.271
4.0 >	0.321	0.353	0.787	0.763	0.2	0.667	0.687	0.811	0.293	0.699	0.122
Grand Total	0	0.88	0.125	0	0.458	0.652	0.004	0.675	0.147	0.556	

Fig. 5. Difference of Proportions with Regard to Snowplow Pass Frequency

“big data” is mentioned, many assume that complicated models and high-level statistics must be involved in order to understand and analyze such data. As many agencies lack funding as well as the availability of qualified individuals to perform intricate analysis, this assumption brings various problems. In the current paper, meaningful results were obtained by performing a simple difference of proportions test on large combined data sources. By starting with a simple and defined analysis, effective results can be obtained, with more intricate analysis following only as more is understood on the subject. This paper’s simple-to-complex methodology can therefore serve as a template for future researchers in performing big data analytics.

By using the data dashboard tool, the interaction amongst winter weather conditions, snowplow operations, travel speeds and crashes can be examined. In other words, by using a simple and straightforward statistical analysis, this paper has demonstrated the benefit of winter road maintenance in conjunction with visual analytics in that a clear relationship between winter maintenance operations and traffic safety can be observed despite working with complex data and conflicting statistical models. In sum, by incorporating visual analytics, we can gain a better understanding of winter maintenance operations and their interactions with traffic safety, traffic operations, and weather conditions. Such straightforward methods of analysis can help agencies make more informed and therefore better decisions.

REFERENCES

- [1] Federal Highway Administration. *How Do Weather Events Impact Roads*. United States Department of Transportation, Washington, D.C., 2014. http://ops.fhwa.dot.gov/Weather/q1_roadimpact.htm.
- [2] Qin, X., Noyce, D. A., Lee, C., and Kinar, J. R. 2006. *Snowstorm Event-Based Crash Analysis*. Transportation Research Record: Journal of the Transportation Research Board, No. 1948, pp. 135–141.
- [3] Usman, T., Fu, L., & Miranda-Moreno, L. F. (2010). *Quantifying safety benefit of winter road maintenance: Accident frequency modeling*. *Accident Analysis & Prevention*, 42(6), pp. 1878-1887.
- [4] Hans, Z., N. Hawkins, P. Savolainen, and E. Rista. 2018. *Operational Data to Assess Mobility and Crash Experience during Winter Conditions*. Center for Weather Impacts on Mobility and Safety, Institute for Transportation, Ames, IA.
- [5] Chen, W., Guo, F., & Wang, F. Y. (2015). A survey of traffic data visualization. *IEEE Transactions on Intelligent Transportation Systems*, 16(6), pp. 2970-2984.
- [6] Vardeman, S. B., Jobe, M. J., *Basic Engineering: Data Collection and Analysis*. Pacific Grove, California. DUXBURY. 2001. pp. 334-442
- [7] Snedecor G., and Cochran, W. (1989), "Statistical Methods", Eighth Edition, Iowa State University Press, pp. 125-128. "Lecture Notes 5: Estimation with Confidence Intervals." *Statistics 13 Elementary Statistics*.
- [8] Ryan, T. (2008), "Modern Engineering Statistics", Wiley, pp. 124-126.
- [9] Choudhary, Pushpa, et al. *Impacts of Speed Variations on Freeway Crashes by Severity and Vehicle Type*. *Accident Analysis & Prevention*, vol. 121, 2018, pp. 213–222., doi:10.1016/j.aap.2018.09.015.
- [10] *Visualizations and Analytics as a Tool for Policy-Making*. Government Information Quarterly. June 2019.