The Association between Precipitation Intensity and Traffic Speeds

Abstract

This project seeks to develop an understanding of the relationship between traffic conditions and precipitation intensity as well as the time of day.

1. Introduction

It is well established that there is a strong interaction between weather and traffic speeds. According to the Federal Highway Administration (FHWA) there are roughly 5.8 million crashes every year, 21% of which are weather related. Of this 21%, 46% are directly attributable to rain, this works out to be roughly 560 thousand crashes every year. Many can relate to the experience of driving in inclement weather conditions when speeds drop drastically. It would be ideal to flesh out the relationship between many weather variables such as temperature, dewpoint, wind speeds, wind direction, and precipitation intensity but locating high temporal resolution weather data is particularly difficult. A typical choice for weather data is the Federal Aviation Administration's (FAA) Automated Surface Observing Systems (ASOS). This dataset was downloaded and experimented with but proved to be missing entirely too much data to be useful. Due to the amount of missingness in the data, imputation wasn't even an option. An ideal choice for this analysis would be the High-Resolution Rapid Refresh (HRRR) weather model by the National Oceanic & Atmospheric Administration (NOAA) but there are no existing archives of the 15-minute data, only the hourly data. Hourly data, for the purposes of traffic weather interaction analysis, will not do as smoothing the traffic speed data over an hour would cause the data to regress to the mean. The interest of this study is on short term interactions; thus, a high-resolution dataset, both temporally and spatially, is necessary.

This study will focus on the relationship between traffic speeds and precipitation intensity. Spatially, the analysis will be focused on the North bound corridor of I-65 from Zionsville, IN to Lafayette, IN. Temporally, the analysis has been confined to the month of June 2018. This study aims to answer the following questions. What impact does precipitation intensity have on traffic speeds? Can a palpable percent reduction be calculated between normal conditions and those impacted by precipitation? Are there any temporal or spatial relationships noticeable in the data?

2. Traffic Speed Data

The traffic speed data is sourced from the company known as INRIX and provided courtesy of Purdue’s Joint Transportation Research Program (JTRP). INRIX specializes in real time crowd sourced traffic data and thus provide an extremely high-resolution dataset. INRIX boasts an astounding 300 million data sources for their platform (Inrix). The traffic speed data consists of road segments, roughly 1 km in length, with average speed values at a 1-minute temporal resolution with units of miles per hour. The speeds data works as follows; INRIX has defined segments based on an in-house algorithm, these segments are treated as geo-fenced locations from which GPS speed measurements are taken from their various data sources, the resulting data is averaged over one minute intervals and is saved to the database. Conveniently, INRIX ensures no missing values will occur in the dataset as they handle missing value imputation on the backend and provide a confidence score to the user indicating what type of imputation was used. The confidence scores are as follows; 30 is a high confidence score and indicates that the data is a real time measurement from the segment in question, 20 is a medium confidence score and indicates that the data is a composite of surrounding real time segments, and finally 10 is a low confidence score and indicates that the data was imputed using historical data (Kim, S., & Coifman, B.).

For the purposes of this study, I’ve restricted my analysis to the north bound corridor of I-65 from Zionsville, IN to Lafayette, IN. This section of interstate consists of 69 discrete road segments and can be seen in Figure 1.

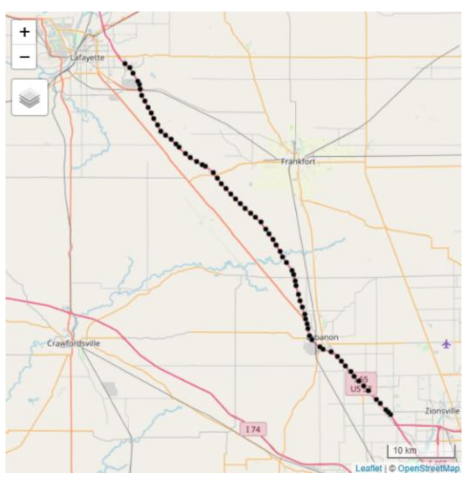

Figure 1: The 69 Segment used in this study along the north bound corridor of I-65 from Zionsville, IN to Lafayette, IN.

A review of the confidence scores for the speed data in June 2018 along this section of interstate reveals that the data consists of 98.75% high confidence data, 1.25% medium confidence data, and 0% low confidence data. This lends a high amount of credibility to the speeds data leveraged in this analysis. The traffic speed data provides a native temporal resolution of 1-minute but to match it up with the precipitation data, it has been down-sampled to a 2-minute resolution via a simple average.

3. Precipitation Intensity Data

The precipitation intensity data is sourced from the National Severe Storms Laboratory's (NSSL) Multi-Radar Multi-Sensor (MRMS) system and is provided courtesy of the Iowa Environmental Mesonet (IEM). The MRMS is considered a Level IV weather data product, that is to say it is highly processed. The MRMS combines data from multiple sources, such as upper-air observations, satellite observations, radar, and forecasts, into a convenient 1 km x 1 km grid. Various variables are available in the MRMS dataset but the NSSL does not archive any of this data. The IEM does maintain an archive of a select few variables, of which precipitation intensity is one.

Precipitation intensity is provided at a 1 km x 1 km grid spacing, a 2-minute temporal resolution, and is provided in units of mm/hr. A few time steps were missing from the archive, these will be addressed in the methodology section.

4. Methodology

The first step in this analysis required interpolating the precipitation data down to the road segment level and handling missing time steps from the precipitation data. It was decided that the missing data could be determined from a naïve forecast of the previous timestep after interpolation was carried out. Various interpolation methods were considered in this analysis. Ordinary Kriging interpolation was the original choice for interpolation in this project as it provides a way to give confidence values to the resulting interpolation in addition to the way in which it utilizes the full dataset. Kriging proved to be unrealistic given that a 2-minute temporal resolution in the month of June has 21,600 timesteps that would require interpolation. This could be accomplished but would require an embarrassingly parallel approach to accomplish in a reasonable timeframe and is thus outside the scope and time available for this project.

Nearest neighbor interpolation was considered but was ultimately discarded as it didn't feel sufficient for the nature of precipitation data. Precipitation data often goes from high to low values in a gradual manner which nearest neighbor would be incapable of representing. Ultimately, bilinear interpolation from the Akima package in R was chosen as it essentially uses a weighted average in the X and Y directions to determine the point of interest. Bilinear interpolation for the full precipitation dataset required roughly six hours to complete. The resulting precipitation data was merged with the traffic speeds data and saved to csv.

Upon completion of the interpolation for the precipitation data, it was merged with the traffic data by segment ID and timestamp. At this point, an analysis of the traffic data in light of precipitation intensity could be carried out. To carry out the analysis, some terms needed to be defined. Normal traffic conditions were defined in this data as traffic speeds greater than 60 miles per hour. Traffic events where precipitation was suspected to play a role and simply referred to "Events" from this point forward were defined as times during which speeds were less than 60 miles per hour and precipitation intensity was greater than 0 mm/hr. A temporal component was also suspected to play a prominent role and thus day and night were defined according to civil twilight hours. Civil twilight is defined as the time before sunrise and after sunset in which there is enough light for most outdoor activities (Civil twilight). In June 2018, daylight was defined as 6:00am to 8:30pm (Sunrise, USA).

Four combinations of criteria were available as a result of the above definitions; night/normal conditions, night/event, day/normal conditions, day/event. With these criteria in mind, counts for each criterion were calculated in addition to the average speed for each segment under various conditions. A percentage was calculated for each criteria according to Equation 1.

Equation 1: Percentage calculation of segments.

$$\frac{\sum count_{segment}}{T_{timesteps} \times N_{segments}} \times 100\ ,\ Where\ T = 21{,}600\ and\ N = 69$$

Finally, a percent reduction from normal was calculated according to Equation 2, where x_i is the speed under event conditions at segment i, y_i is the speed under normal conditions at segment i, and n is the number of segments and is equal to 69.

Equation 2: Percent reduction in mean speed from normal by segment.

$$Percent\ Reduction\ in\ Mean\ Speed = 100 - \left(\frac{\frac{1}{n}\sum x_i}{\frac{1}{n}\sum y_i} \right) * 100$$

5. Results & Discussion

A count was carried out of the events per each set of criteria and the result can be seen in Table 1.

Table 1: Count and percentage under each criterion in the traffic/precipitation dataset.

	Sum Normal Conditions Count – Night	**Sum Normal Conditions Count – Day**	**Sum Event Conditions Count – Night**	**Sum Event Conditions Count – Day**
Count	591,308	896,393	712	1,987
Percentage	39.67%	60.14%	0.04%	0.13%

It became apparent that the proportion of the data that this study is particularly interested in, events, are exceedingly rare in comparison to normal conditions. Normal conditions make up over 99% of the full dataset. This leaves a paltry 0.17% of the data that this study is interested in. This problem is far from unique and is quite common in weather study applications elsewhere. Expanding this analysis to further months could increase the counts for event conditions but the overall percentages would likely be relatively unchanged.

Counts were subsequently mapped to look for spatial patterns in the dataset. The sum of day and night events can be seen in Figure 2a and b. Upon review of this figure, it quickly becomes apparent that nighttime events display a high amount of spatial clustering. Day time events display a similar clustering but are far more homogenous.

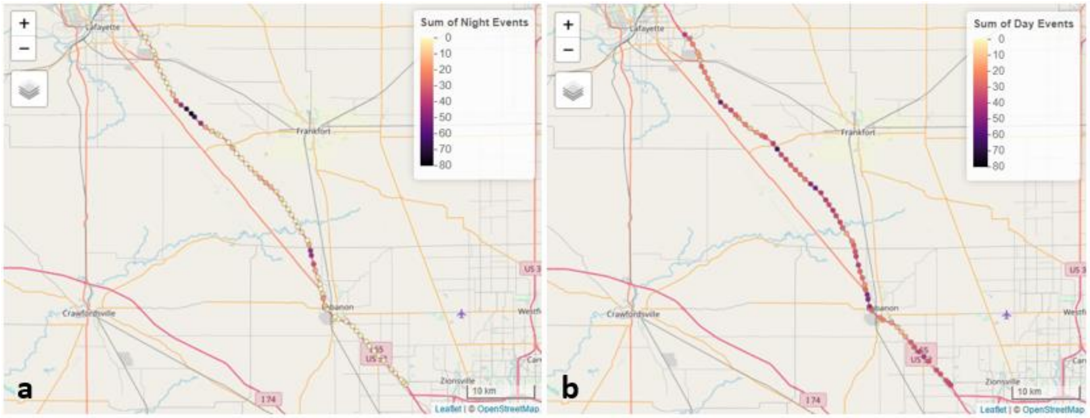

Figure 2a,b: Counts of events for day and night conditions at each segment.

Zooming in on the southern-most cluster of Figure 2a, it was discovered that this particular location is an on/off ramp for I-65 as seen in Figure 3, suggesting that weather is likely not the primary contributor to the speeds in this location.

Figure 3: Zoomed in view of on/off ramp during nighttime event.

A review of the northerly most clustered events in Figure 2a reveals another interesting external factor, an underpass, that seems to have more influence than weather impacts alone. Further investigation leveraging Google Street View, revealed construction occurring in this location in August of 2018, two months after the data used in this study as seen in Figure 4a and b. Interestingly enough this interaction is particular strong at night. Many interstate construction crews work primarily during the night to capitalize on cooler temperatures, this points to construction areas as another primary contributor to traffic slowdowns.

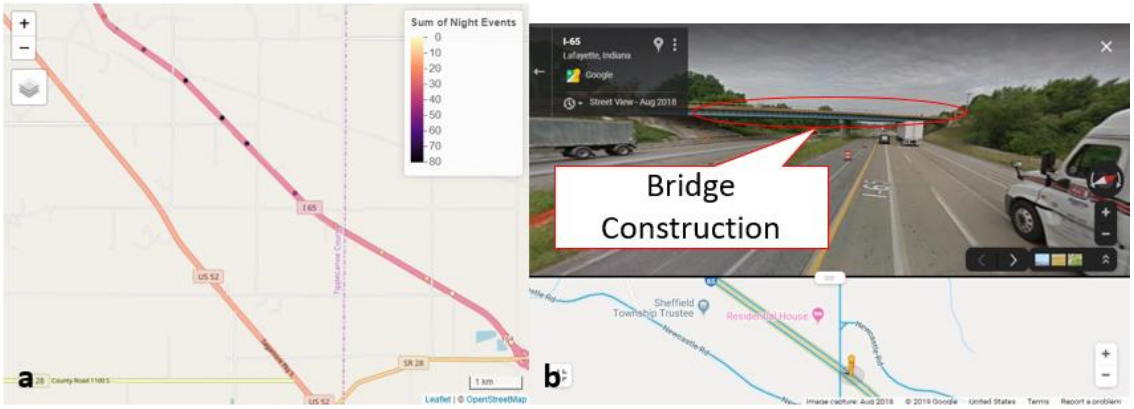

Figure 4a,b: Zoomed in view of construction along with Google Street View image to prove bridge construction during time frame.

To develop an understanding of the speeds occurring during normal and event conditions at day and night mean speeds were calculated for each segment in June 2018. These speeds can be seen in Figure 5a and b.

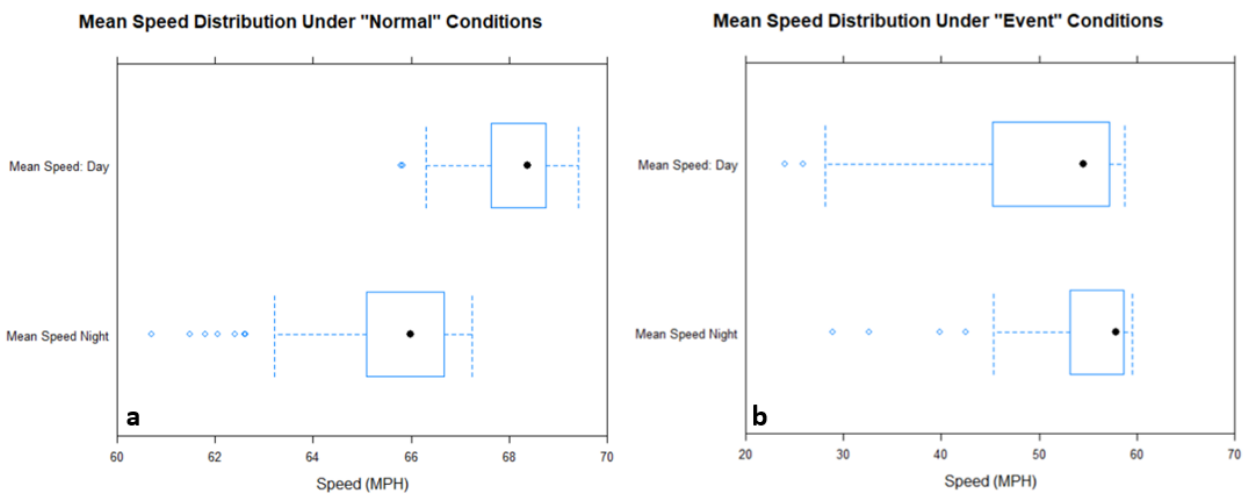

Figure 5a, b: Distribution of mean speeds for all segments in study under normal and event conditions during day and night.

When reviewing Figure 5a, under normal conditions daytime speeds tend to be higher than nighttime speeds. In stark contrast though, daytime speeds are significantly lower than nighttime speeds under event conditions as seen in Figure 5b. This seems to suggest daytime precipitation events have a greater impact on speeds than nighttime events. This is somewhat intuitive due to higher traffic volumes during the day.

To quantify the impact of precipitation intensity on traffic speeds a percent reduction from normal was calculated for each segment according to Equation 2. The results are somewhat striking and can be seen in Figure 6a and b. What appears to be an apparent clustering of the data is circled in red in both figures.

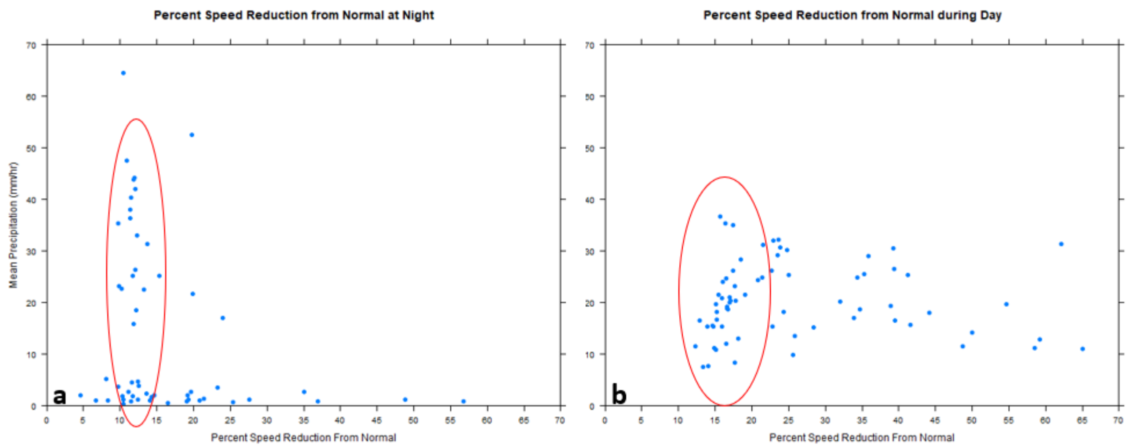

Figure 6a, b: Percent speed reduction from normal conditions during the night and day.

Utilizing the percent speed reductions from normal, it appears that there is a 10 – 15% reduction in speeds at night from normal when precipitation is 15 mm/hr or greater and a 13 – 22% reduction speeds during the day from normal when precipitation is 8 mm/hr or greater. The nighttime data tends to more concentrated around this 10 – 15% mark whereas daytime speeds seem to suffer significantly greater drops in speed more often but at lower precipitation intensities than at night. This points to daytime traffic speeds being more sensitive to precipitation intensity. It seems likely that this is a byproduct of higher traffic volumes during the day.

6. Conclusions

This study set out to determine what the relationship was between traffic speeds and precipitation intensity. This was accomplished by applying bilinear interpolation to precipitation intensity data to get values at the road segment level and then statistics were calculated for normal and event conditions during the day and night. The results suggest that precipitation does have a significant impact on traffic speed though there are more variables involved than just precipitation intensity, such as construction or on/off ramps. A temporal relationship in the day and night cycles was confirmed by this study. It was determined that the daytime events are more sensitive to precipitation with 8 mm/hr or greater resulting in speed reductions of 13 – 22%. Nighttime conditions seemed to show a 10 – 15% reduction in speed for precipitation intensities of 15 mm/hr or greater. Further research would do well to include traffic volumes as well as construction zones to further isolate the impact of precipitation intensities.

Acknowledgement

- **Traffic data was provided courtesy of Purdue's Joint Transportation Research Program.**
- **Precipitation data was provided courtesy of the Iowa Environmental Mesonet.**
- **A large thanks to Professor Jie Shan who provided some much appreciated advice in completing this project.**

References

1. Civil twilight. In The Merriam-Webster.com Dictionary. Retrieved December 7, 2019, from https://www.merriam-webster.com/dictionary/civil%20twilight
2. Inrix. "Research." Inrix, 11 July 2017, inrix.com/research/.
3. Kim, S., & Coifman, B. (2014). Comparing INRIX speed data against concurrent loop detector stations over several months. Transportation Research Part C: Emerging Technologies, 49, 59–72. https://doi.org/10.1016/j.trc.2014.10.002
4. "Sunrise, USA - Sunrise, Sunset, and Daylength, June 2018." Timeanddate.com, www.timeanddate.com/sun/usa/sunrise?month=6&year=2018.