**How is snow impact New York State Highway network**

**CIE500 Final project report**

**Dongyang Wu**

**5/12/2025**

**Abstract**

This study aims to quantify the impact of snow accumulation on Annual Average Daily Traffic (AADT) across the New York State highway network using a combination of Ordinary Least Squares (OLS) regression and PyTorch-based neural networks. Inspired by Roh's work on traffic modeling in cold regions, we adapt a categorical approach to model nonlinear snow effects and evaluate model performance across geospatial road data.

**Introduction**

In cold climate regions such as New York State, winter weather conditions—particularly heavy snowfall—substantially influence roadway operations and reduce traffic flow. Accurately estimating winter traffic conditions is crucial for maintaining highway functionality and informing maintenance and resource allocation. Previous studies in Canada and the U.S. have highlighted how snowstorms and temperature extremes reduce traffic volumes due to changes in driver behavior, service levels, and road conditions.

**Lecture review**

Roh (2024) proposed a categorical approach to model cold-weather conditions by dividing temperature ranges into distinct cold categories (CC1 to CC6). The study showed that such classification improves model interpretability and transferability across road segments. Other studies, such as Hanbali and Kuemmel (1993), observed a 10–20% reduction in traffic volume during snow events, while Stern et al. (2003) emphasized the variability of traffic response depending on location, severity, and time of day.

Building on this foundation, our study introduces a snowfall-based classification scheme to model its impact on AADT. We integrate this with road-specific features (speed limit and segment length) to build predictive models using both OLS and PyTorch-based neural networks.

**Data analysis**

Traffic Data: AADT shapefile from NYSDOT containing segment-level traffic volumes, road geometry, speed limits, and road lengths.

Weather Data: NOAA GeoTIFF raster providing 48-hour snowfall accumulation data for January 15, 2025.

Geospatial Tools: QGIS and Python to spatially join snow depth data with road segments using centroid-based extraction.

Snowfall Categorization

| Baseline | Snow depth (inches) |
| --- | --- |
| Baseline | 0 |
| CC1 | 0-5 |
| CC2 | 5-10 |
| CC3 | 10-15 |
| CC4 | 15-20 |
| CC5 | 20-25 |
| CC6 | 25-30 |
| CC7 | 30-35 |

**Methodology**

This study employed two primary modeling approaches to evaluate the effect of snowfall on traffic: traditional linear regression (OLS) and a deep learning method using PyTorch. Both methods utilized the same set of input features derived from the integrated spatial and weather datasets.

The snowfall depth was categorized into seven discrete intervals (CC1 through CC7), following a structure inspired by Roh's cold category classification framework. Each category represented a 5-inch increment, starting from 0–5 inches (CC1) up to 30–35 inches (CC7). Road segments with snowfall depth below 0 inches served as the baseline category. Dummy variables were generated for each category (excluding the baseline) to allow the models to learn the marginal effect of each snow level on AADT.

In addition to snow depth categories, two continuous features were included: speed limit (in miles per hour) and road segment length (in feet). These were added to control for geometric and functional roadway differences across the network.

For the linear model, an Ordinary Least Squares (OLS) regression was performed using the statsmodels library. The model specified AADT as the dependent variable and the snow categories, speed limit, and road length as independent variables. Multicollinearity was minimized by dropping one dummy variable (baseline category) and checking for inflated standard errors.

The machine learning model was a fully connected feedforward neural network built in PyTorch. It featured two hidden layers with 64 and 32 neurons, respectively, each using ReLU activation. The output layer returned a single value representing the predicted (standardized) AADT. Training used Smooth L1 Loss (Huber) to reduce the influence of outliers, and the Adam optimizer was used for gradient updates. The input features were standardized, and the dataset was split into 80% training and 20% testing subsets. The model was trained for 100 epochs, with loss values recorded to monitor convergence.

**Results**

The OLS regression model revealed that snowfall has a measurable and statistically significant impact on traffic volume. The model achieved an R-squared value of 0.104, indicating that approximately 10.4% of the variability in AADT could be explained by the included variables. While this may seem modest, it is consistent with traffic forecasting studies involving environmental variables. Importantly, the F-statistic was significant (p < 0.001), confirming the model’s validity.

Among the snow depth categories, CC2 through CC6 showed statistically significant negative coefficients, ranging from -2,700 to -3,300 AADT. These results indicate that even moderate snowfall (5–10 inches) can reduce traffic volumes meaningfully. The strongest reductions occurred within the 20–30 inch range (CC5 and CC6), while CC7 (30–35 inches) had a negative coefficient but lacked statistical significance, likely due to fewer samples under extreme conditions.

Speed limit was positively correlated with AADT (coefficient: +709), supporting the hypothesis that higher capacity roads carry more traffic. Road length showed a small but statistically significant negative coefficient (-0.108), possibly due to longer segments often representing rural or lower-demand corridors.

The PyTorch neural network also confirmed the predictive value of the input features. Training loss decreased from 0.21 to approximately 0.19 over 100 epochs, demonstrating consistent model improvement. The final test loss was 0.2116, showing good generalization on unseen data. The scatter plot of predicted versus true AADT values showed a close-to-linear distribution, further supporting the model’s validity.

**Discussion**

The results of this study reinforce the understanding that snowfall significantly influences traffic volume in cold-weather regions. The consistent decline in AADT with increasing snow depth is aligned with expectations and confirms that both statistical and machine learning models are capable of detecting meaningful patterns in winter traffic behavior.

When comparing our findings with those of Roh (2024), several parallels emerge. Both studies identified specific weather categories that result in reduced traffic mobility. While Roh applied temperature thresholds to define cold categories, we implemented a snowfall-based framework. Despite this difference in meteorological variables, both approaches yielded similar conclusions: as weather conditions worsen, traffic volumes decline in a statistically significant manner. Moreover, while Roh focused on spatial transferability across different road types in Alberta, our study concentrated on predictive accuracy within a geographically fixed network in New York. This presents an opportunity for future work to bridge both directions—model generalizability and precision.

From a modeling standpoint, the OLS regression offered strong interpretability and allowed for straightforward evaluation of individual variables. The significant coefficients for CC2–CC6 suggest that even moderate snowfall (5–10 inches) triggers behavioral changes that reduce road usage. The inclusion of speed limit and road length provided control for functional differences among road segments, enhancing the robustness of our findings.

The PyTorch-based neural network model, while less interpretable, successfully captured non-linear relationships and interactions that may not be evident in a linear framework. The decrease in loss over 100 training epochs and strong alignment between predicted and actual AADT values demonstrate that neural networks can serve as effective tools in traffic prediction when properly trained. This finding supports the increasing interest in integrating deep learning into transportation analytics, particularly when data complexity or volume exceeds the limits of traditional regression.

Nevertheless, several limitations must be acknowledged. The lack of temperature and wind speed data restricts the model’s ability to distinguish between snow-induced reductions and other meteorological influences such as ice or visibility. Additionally, the use of aggregated AADT data limits temporal resolution, preventing the analysis of peak/off-peak behavior or storm-phase-specific impacts. Moreover, CC7, representing extreme snowfall, lacked sufficient samples to reach statistical significance, indicating the need for longer-term or broader datasets to evaluate rare but critical conditions.

Despite these limitations, the strengths of this study lie in its reproducible methodology, integration of spatial and environmental datasets, and successful adaptation of a cold-region modeling framework to the U.S. Northeast. The consistent findings across both modeling strategies suggest that transportation agencies can confidently adopt similar approaches for short-term forecasting and winter operations planning.

**Future extension**

While the current study provides a strong foundation for understanding how snow depth impacts traffic, there are several avenues for expanding this research in future work.

First, incorporating additional weather variables such as temperature, wind speed, humidity, and visibility could enhance the model’s explanatory power. In Roh’s original framework, temperature-based cold categories (CC1–CC6) played a key role in capturing behavioral and operational responses to freezing conditions. Adapting a multivariable meteorological model using high-resolution atmospheric datasets could help distinguish snowfall impacts from other winter hazards like ice and wind.

Second, the current study aggregates traffic data as AADT, which does not reflect temporal dynamics. Expanding the model to hourly or daily traffic counts would allow for dynamic traffic impact modeling under varying snow conditions. This would require integration of short count or continuous count station data from NYSDOT, enabling more detailed simulation of traffic behavior during the onset, peak, and recovery phases of snow events.

Third, the current analysis is limited to one geographic region. A valuable extension would be to evaluate spatial transferability, as Roh did, by testing whether a snow-impact model trained on one region of New York (e.g., Western New York) generalizes to others (e.g., Capital Region or Adirondacks). This would involve developing GNN-based models capable of capturing topological and regional heterogeneity.

Fourth, a Graph Neural Network (GNN) framework could be used to explicitly model the interconnectivity of road segments. Unlike the MLP used here, GNNs propagate feature information across connected roads, which could reveal cascading effects of snow on adjacent segments or routes. This would be especially valuable for incident response and rerouting applications during snowstorms.

**Conclusion**

This study confirms that snowfall exerts a statistically significant, negative effect on highway traffic in cold regions. Using a categorical approach adapted from Roh (2024), we demonstrated consistent reductions in AADT during snow events of 5–30 inches. Both OLS and machine learning models successfully captured this trend, with neural networks offering potential for nonlinear enhancement.

My methodology can be adapted by transportation agencies for snow event planning, traffic forecasting, and winter maintenance prioritization.

**References**

1. Roh, H. J. (2024). Modelling severe weather impact on traffic mobility and evaluating cross-regional applicability to functionally distinct segments in cold region highway networks. Cold Regions Science and Technology, 213.
2. Hanbali, R. M., & Kuemmel, D. A. (1993). Traffic volume reductions due to winter storm conditions. *Transportation Research Record*, 1387.
3. Stern, A., et al. (2003). Effects of weather on freeway traffic flow characteristics and facility capacity. Transportation Research Record, 1856.
4. NOAA National Weather Service. (2025). 48-hour snowfall accumulation GeoTIFF for New York.
5. NYSDOT (2023). Statewide AADT Shapefile and Road Inventory.