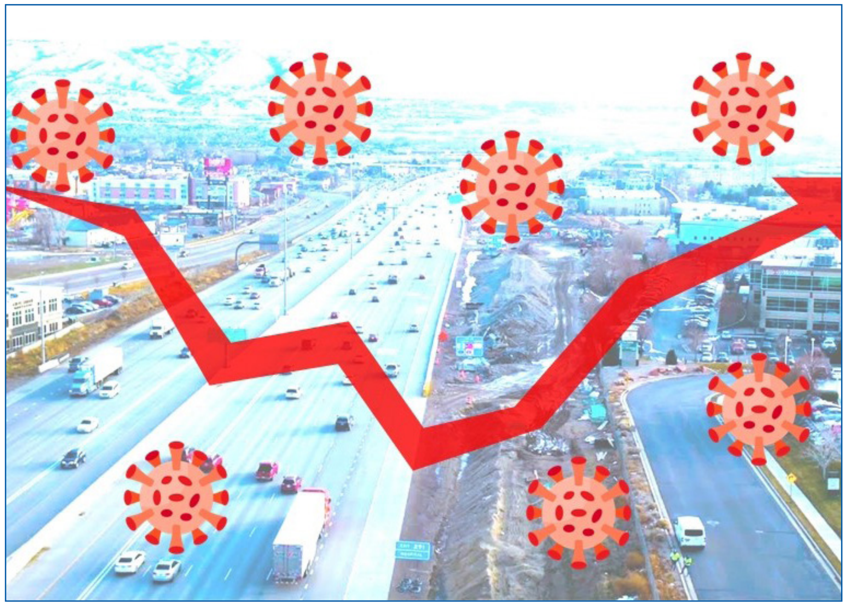
MOUNTAIN-PLAINS CONSORTIUM

MPC 23-502 | P. Lu, X. Yang and Y. Gong

KNOWLEDGE-BASED MACHINE LEARNING FOR FREEWAY COVID-19 TRAFFIC IMPACT ANALYSIS AND TRAFFIC INCIDENT MANAGEMENT



A University Transportation Center sponsored by the U.S. Department of Transportation serving the Mountain-Plains Region. Consortium members:

Colorado State University
North Dakota State University
South Dakota State University
University of Colorado Denver
University of Denver
University of Utah
Utah State University
University of Wyoming

Knowledge-based Machine Learning for Freeway COVID-19 Traffic Impact Analysis and Traffic Incident Management

Dr. Pan Lu
Associate Professor

North Dakota State University
Fargo, North Dakota

Pan.lu@ndsu.edu

Dr. Xianfeng (Terry) Yang
Assistant Professor

University of Utah
Salt Lake City, Utah

X.Yang@Utah.edu

Dr. Yaobang Gong (Postdoc Researcher)
Research Assistant

University of Utah
Salt Lake City, Utah

Yaobang.Gong@Utah.edu

August 2023

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
MPC-657		

4. Title and Subtitle	5. Report Date
Knowledge-based Machine Learning for Freeway COVID-19 Traffic Impact Analysis and Traffic Incident Management	August 2023
	6. Performing Organization Code
7. Author(s)	**8. Performing Organization Report No.**
Pan Lu, Ph.D. Dr. Xianfeng (Terry) Yang, Ph.D. Dr. Yaobang Gong (Postdoc Researcher)	MPC 23-502
9. Performing Organization Name and Address	**10. Work Unit No. (TRAIS)**
North Dakota State University NDSU Dept 2880 Fargo, ND 58108-6050 The University of Utah 201 President Circle Salt Lake City, UT 84112	11. Contract or Grant No.
12. Sponsoring Agency Name and Address	**13. Type of Report and Period Covered**
Mountain-Plains Consortium North Dakota State University PO Box 6050, Fargo, ND 58108	Final Report
	14. Sponsoring Agency Code

15. Supplementary Notes

Supported by a grant from the US DOT, University Transportation Centers Program

16. Abstract

The U.S. Department of Transportation (USDOT) needs to respond quickly and adapt to the coronavirus (COVID-19) to ensure continuation of critical infrastructure support and relief for the American people. Since early March 2020, the COVID-19 pandemic has had significant impact on traffic across the United States. It is clear to see that traffic patterns, traffic demands, and duration alter with COVID status. Therefore, there is a critical research need to study the impact of COVID on traffic patterns and to analyze the relationships among traffic demand patterns, daily confirmed cases/deaths, state policies, public perceptions, etc. In this research, we investigate the impact of COVID-19 on traffic safety in different stages, focusing on Salt Lake County, Utah. Statistical methods are employed to determine if there are any differences in the effects of the pandemic. Further, the effect of COVID-19 on traffic patterns in Salt Lake County and Utah County from January 2019 to July 2021 was analyzed. Different vehicle miles traveled (VMT) patterns in the pre-pandemic stage, early stage of the pandemic, and late stage of the pandemic are identified. Finally, a knowledge-based traffic prediction model utilizing an innovative approach that integrates machine learning with graph theory is proposed to forecast traffic patterns in the near future.

17. Key Word	18. Distribution Statement
COVID-19, demand, freeways, incident management, machine learning, predictive models, traffic incidents, traffic models, traffic volume, travel patterns	Public distribution

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	53	n/a

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Acknowledgments

This research was made possible with funding supported by the U.S. Department of Transportation though the Mountain-Plains Consortium (MPC) Transportation Center. The authors express their deep gratitude to U.S. DOT and MPC.

Disclaimer

ABSTRACT

The U.S. Department of Transportation (USDOT) needs to respond quickly and adapt to the coronavirus (COVID-19) to ensure continuation of critical infrastructure support and relief for the American people. Since early March 2020, the COVID-19 pandemic has had significant impact on traffic across the United States. It is clear to see that traffic patterns, traffic demands, and duration alter with COVID status. Therefore, there is a critical research need to study the impact of COVID on traffic patterns and to analyze the relationships among traffic demand patterns, daily confirmed cases/deaths, state policies, public perceptions, etc. In this research, we investigate the impact of COVID-19 on traffic safety in different stages, focusing on Salt Lake County, Utah. Statistical methods are employed to determine if there are any differences in the effects of the pandemic. Further, the effect of COVID-19 on traffic patterns in Salt Lake County and Utah County from January 2019 to July 2021 was analyzed. Different vehicle miles traveled (VMT) patterns in the pre-pandemic stage, early stage of the pandemic, and late stage of the pandemic are identified. Finally, a knowledge-based traffic prediction model utilizing an innovative approach that integrates machine learning with graph theory is proposed to forecast traffic patterns in the near future.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AIC	Akaike Information Criterion
COVID-19	Coronavirus Disease 2019
DOT	Department of Transportation
DUI	Driving Under the Influence
EMS	Emergency Medical Services
GCN	Graph Convolutional Networks
GNN	Graphical Neural Networks
LSTM	Long Short-Term Memory
MAPE	Mean Absolute Percentage Error
NB	Negative Binomial
NCEI	National Centres for Environmental Information
NHSTA	National Highway Traffic Safety Administration
PeMS	Performance Measurement System
ReLU	Rectified Linear Unit
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
UDOT	Utah Department of Transportation
VIF	Variance Inflation Factor
VMT	Vehicle Miles Traveled

EXECUTIVE SUMMARY

During the early stages of the pandemic, transportation networks experienced a significant decrease in traffic demand. However, as local businesses, schools, and other establishments began to reopen and daily COVID cases declined, traffic demands gradually increased over time, particularly with express delivery truck traffic. In some states, traffic demands have even returned to pre-pandemic levels. It is evident that COVID status has influenced traffic patterns, demands, and durations. Therefore, there is a pressing need for research to examine the impact of COVID on traffic patterns and analyze the relationship between traffic demand patterns, daily confirmed cases/deaths, state policies, public perceptions, and other factors. Despite numerous studies focusing on this issue, most have only investigated the earlier stages of the pandemic when health emergencies were still prevalent and vaccines were not widely available to the general public. As travel restrictions expire, non-pharmaceutical interventions ease, and public perceptions shift, it becomes crucial to investigate the pandemic's impact on traffic safety during the latter stages of the pandemic.

In this research project, researchers examined the influence of COVID-19 on traffic safety across various stages, with a particular focus on Salt Lake County, Utah. Statistical techniques are employed to determine whether the pandemic's effects vary over time. Negative binomial models are used to examine crash frequency, while binary logit models are used to investigate crash severity. These models take into account exposure, environmental factors, and human factors. Furthermore, Welch's t-test and pairwise t-test are utilized to investigate any potential secondary effects of the pandemic on factors not directly related to the pandemic in the statistical models.

Researchers then analyzed the impact of COVID-19 on traffic patterns in Salt Lake County and Utah County from January 2019 to July 2021. The study identified distinct vehicle miles traveled (VMT) patterns during the pre-pandemic, early pandemic, and late pandemic stages. The researchers discovered that VMT is significantly influenced by the severity of the pandemic (as measured by the number of new cases), policies, and individual/societal risk perceptions regarding travel during the pandemic.

Finally, a novel traffic prediction model that combines machine learning with graph theory to forecast traffic patterns in the near future is developed. The proposed approach incorporates human knowledge, resulting in improved model performance. The model's effectiveness is demonstrated through its high level of accuracy in predicting traffic patterns.

The traffic patterns identified in this research project can aid agencies in comprehending the effects of COVID-19 on traffic mobility, potentially assisting with long-term urban planning objectives during the post-pandemic period. To enhance traffic safety in the "new normal" phase, stakeholders should implement measures to discourage driving under the influence and reduce crashes involving commercial vehicles. State departments of transportation (DOTs) and other responsive agencies can utilize the developed prediction model to prepare for upcoming traffic demand patterns in the near future.

1. INTRODUCTION

1.1 Problem Statement

Since early March 2020, the COVID-19 pandemic has had significant impacts on U.S. traffic due to the quarantine rules of many states and the "work-from-home" style of many residents. The weekday VMT traffic pattern changed from a typical two-peak pattern (morning peak followed by a drop and then afternoon peak) to gradually increasing to a single afternoon peak (Skip Descant, 2020). One reason for the changes is that home-based work reduces the early morning travel needs and provides people with new flexibility for midday in-person errands.

Beside the commute traffic pattern changes, some researcher also observed a sharp increase in on-demand delivery and e-commerce (AASHTO Journal, 2020). Hence, at the early stage of the pandemic, a great reduction in traffic demands was observed on transportation networks. Later, with the process of re-opening local businesses, schools, etc., and the decrease in daily COVID confirmed cases, traffic demands have gradually increased over time, especially with express delivery truck traffic, and have even recovered to pre-pandemic levels in some states (Glaeser et al., 2020). It is clear to see that traffic patterns, traffic demands, and duration change with COVID status.

Therefore, there is a critical research need to study the impact of COVID on traffic patterns and analyze the relationships among traffic demand patterns, daily confirmed cases/deaths, state policies, public perceptions, etc. Existing studies show that the crash frequency has reduced from the start of the pandemic until June 2021, mainly due to the reduced traffic volume. The crash severity has increased with the increase of risky driving behavior, including speeding, driving under the influence (DUI), and not using seat belts.

Although many researchers have focused on this problem, most studies focus only on the earlier stage of the pandemic when health emergencies were in place and vaccines were not widely available to the public. As travel restrictions expired, nonpharmaceutical interventions relaxed, and public perceptions changed, studies (Glaeser et al., 2020; Gong et al., 2022; Mahmoudi & Xiong, 2022) found that the traffic volume started to recover to a level comparable to the pre-pandemic level. As mentioned earlier, existing studies have pointed out that the change in traffic safety is primarily related to the change in mobility. Therefore, critical research needs to investigate the pandemic's impact on traffic safety during the latter stage of the pandemic, which motivates the research in this project.

1.2 Objectives

The primary objective of this research project is to help DOTs and other responsive agencies to better understand the long-term impacts of COVID on transportation safety and vehicular traffic patterns in different time stages.

The secondary objective of this research project is to predict the near-future traffic demand patterns with the novel knowledge-based machine learning model, and help state DOTs prepare for near-future traffic demand patterns during the post-pandemic period.

1.3 Scope

Task 1: Literature review

Task 1 focuses on conducting a literature review on COVID-19's impact on transportation in other states and areas.

Task 2: Analysis of the effect of COVID-19 on traffic safety

This task aims to illustrate the correlations between COVID-related social factors and transportation safety performances on Salt Lake County, Utah, in different stages.

Task 3: Analysis of the effect of COVID-19 on traffic patterns

This task focuses on the analysis of the effect of COVID-19 on traffic patterns in Salt Lake County and Utah County from January 2019 to July 2021. Different vehicle miles traveled (VMT) patterns in the pre-pandemic stage, early stage of the pandemic, and late stage of the pandemic are identified.

Task 4: Propose a knowledge-based machine learning model for traffic prediction

For the task of incorporating human knowledge, a traffic prediction model based on an innovative approach integrating machine learning with graph theory is proposed to forecast traffic patterns in the near future.

1.4 Outline of Report

This report documents the findings of the research and proceeds with the following sections:

- Introduction
- Literature review
- Impact of COVID-19 on traffic safety
- Impact of COVID-19 on travel pattern
- Knowledge-based machine learning for near-future prediction
- Conclusions and key findings

2. LITERATURE REVIEW

2.1 Review of the Impact of COVID-19 on Traffic Safety

There have been several studies focusing on the effects of the COVID-19 pandemic. The National Highway Traffic Safety Administration (NHSTA) conducted a nationwide study (Wagner et al., 2020) in October 2020 to examine the traffic safety environment of the second quarter of 2020. The study employs data from emergency medical services (EMS) and hospital trauma centers to learn about the impact on motor vehicle crashes and fatalities. The results show that VMT reduced sharply due to the travel restrictions. Moreover, crashes and crash fatalities were also reduced, while the fatality rate was increased. The contributing factors to the increase in fatality rate include a decrease in seat belt use, an increase in speeding, and an increase in alcohol and drug (including marijuana and opioids) use. Three subsequent updates (Research, 2021a, 2021b; USDOT, 2021) to the aforementioned study were made to include the new data from July 2020 to June 2021. To a large extent, the increases in risky traffic safety behaviors, including speeding, alcohol and drug use, and the decrease in seat belt use, continued throughout 2020, resulting in an increase in severe injury rates. As for the first half of 2021, although the trip-taking rate rebounded, the severe injury rate was still higher than the pre-pandemic level but lower than 2020. Seat belt use remained low, and speeding remained prevalent. However, the data regarding alcohol and drug use were not available.

Many other studies investigate traffic safety during the pandemic's earlier "lockdown" stage. A study (Qureshi et al., 2020) was conducted in Missouri to investigate the relationship between the mandated societal lockdown and traffic crashes. Crash data from January 1, 2020, to May 15, 2020, were used. The modeling shows that traffic crashes resulting in minor or no injuries were significantly reduced but not in those resulting in severe or fatal injuries during the "lockdown" stage. In other words, the total number of crashes was reduced, but the severe injury rate was increased, which aligns with the NHTSA studies. Similar results can also be found in a case study in Connecticut (Doucette et al., 2021), as well as in a case study (Pathak et al., 2022) in Maharashtra, India. Another study (Jie Zhang et al., 2021) conducted in New York City also found that person miles traveled are positively related to the number of people involved in crashes while the stay-at-home policy has a negative impact. A study (Koloushani et al., 2021) of the spatiotemporal impact of COVID-19 on traffic crashes in Florida also shows the reduction in crash frequency and reductions are less in areas populated by the elderly. In general, studies show that during the earlier lockdown, the total number of traffic crashes was reduced, mainly due to the travel restrictions. The number of crashes that lead to severe injuries and deaths is unlikely to be impacted, which results in a higher severe injury rate.

Some other studies extend beyond the lockdown period toward the end of 2020. One study (Islam et al., 2022) employs real-time traffic parameters to investigate the change in road safety trends of a freeway in Florida during 2020. The traffic volume decreased in 2020 compared with the 2017-2019 average volume. The total number of crashes decreased in 2020. The rate per 100 million VMT of crashes leading to fatalities, incapacitating injuries, and non-incapacitating injuries is high in 2020 compared with the 2017-2019 average. In other words, the crash severity also increased. These findings are similar to the series of NHTSA studies. Interestingly, while drug-related crashes increased by 300% in 2020, alcohol-related crashes decreased by 22%. A time series analysis (Sekadakis et al., 2021) employs 10-year (January 2010 to August 2020) road crash, fatality, and slight injury data from Greece to investigate the pandemic's impact on traffic safety. The results show that the total number of road crashes decreased due to the traffic volume decrease, but fatality and slightly injured rates significantly increased. Another interesting finding is that the percentage reduction of crashes and traffic volume is disproportionate, which may indicate that more crashes occurred with regard to the prevailing traffic volume. Another study in New York City (N. Dong et al., 2022) employs survival analysis to explore the effects of the pandemic on different modes of transportation. Crashes that involved injuries and fatalities from March 1, 2020, to December 4, 2020, were used. The modeling results reveal that pedestrian and cyclist safety is improved owing to the increased percentage of people staying at home, while the likelihood of injuries for motor vehicle drivers rises. Another interesting finding is that nonpharmaceutical interventions implemented increased motor vehicle drivers' crash risk. To summarize, the reduction in crash frequency and the increase in crash severity increased throughout 2020, even after the lockdown period.

As the crash severity increased during the pandemic, some studies explored the contributing factors to crash severity, especially the change in driving behaviors. A statistical analysis (X. Dong et al., 2022) found that aggressiveness (speeding, use of alcohol, and improper passing) and inattentiveness (failure to use seat belts, distraction, and failure to signal) of drivers increased significantly during the pandemic, leading to a higher likelihood of severe crashes. A survey of U.S. and Canadian drivers (Vanlaar et al., 2021) suggests that the majority of respondents indicated their behavior did not change; however, notable proportions indicated they were more likely to engage in risky driving behaviors such as speeding, drunk driving, and distracted driving during the pandemic. Two other surveys regarding drunk driving during the lockdown period show different results for different countries. A U.S. survey (Watson-Brown et al., 2021) shows increased drunk driving, while an Australian survey (Manning et al., 2021) shows the opposite result. This may be due to cultural differences as well as different policies. A descriptive analysis (Katrakazas et al., 2020) of Greece and Saudi Arabia found that the pandemic increased speeds, extreme events, and the use of mobile phones. Another simulation study (Sekadakis et al., 2022) shows that pandemic response measures alter driving behavior by increasing mobile phone use and driving speed.

To summarize, existing studies show that crash frequency was reduced from the start of the pandemic until June 2021, mainly due to the reduced traffic volume. Crash severity has increased due to the increase of risky driving behavior, including speeding, DUI, and not using seat belts. However, most studies focus only on the earlier stage of the pandemic when health emergencies were in place and the vaccine was not widely available to the public. As travel restrictions expired, nonpharmaceutical interventions relaxed, and public perceptions changed, studies (Glaeser et al., 2020; Gong et al., 2022; Mahmoudi & Xiong, 2022) found that traffic volume started to recover to a level comparable to pre-pandemic levels. As mentioned earlier, existing studies have indicated that the change in traffic safety is primarily related to the change in mobility. Therefore, critical research needs to investigate the pandemic's impact on traffic safety during the latter stage of the pandemic.

2.2 Review of the Impact of COVID-19 on Vehicular Traffic

Since early March 2020, the global COVID-19 pandemic has placed pronounced impacts on various aspects of society. In addition to the loss of life and illness, the pandemic has resulted in a great impact on traffic across the U.S. Many studies found that at the early stage of the pandemic, traffic was reduced significantly due to the travel restrictions imposed by the government, fear of getting sick, lower levels of economic and social activity, and the work-from-home style of many residents (Katrakazas et al., 2020; Kim, 2021; Jinbao Zhang & Lee, 2021). Later, during the process of re-opening local businesses, schools, etc., and the decrease in daily confirmed COVID cases, traffic demands gradually increased (Glaeser et al., 2020). It is clear to see that traffic patterns, traffic demands, and duration change with COVID status.

As restrictions were relaxed due to fewer COVID cases and the rollout of the vaccines, a return to a post-pandemic "normality" is ongoing and will likely lead to a recovery of mobility to levels comparable to the past. However, there are few studies investigating how traffic patterns will appear at the latter stage of the pandemic or even post-pandemic. A policy analysis conducted by Rothengatter et al. (Rothengatter et al., 2021) discussed the impacts of COVID-19 on different travel modes. However, they ignored car travel. Another European long-term travel demand study by Christidis et al. (Christidis et al., 2021a) was conducted to investigate the post-pandemic recovery of transportation. They found that travel by car will likely return to the 2019 level around 2025. Unfortunately, their analysis was based on a 2018 travel survey rather than either recent traffic data or travel surveys. Therefore, there is a critical research need to study the impact of COVID-19 on traffic patterns in the latter stage of the pandemic based on recent data, and to analyze the relationships among traffic patterns, daily confirmed cases/deaths, government policies, economics, and other factors. Such research results will be valuable for responsive agencies such as state DOTs to better understand the long-term impacts of COVID-19 on transportation and prepare for the near-future traffic demand pattern.

3. IMPACT OF COVID-19 ON TRAFFIC SAFETY

3.1 Overview

In this section, we investigate the impact of COVID-19 on traffic safety in different stages, focusing on Salt Lake County, Utah. Statistical methods are employed to determine if there are any differences in the effects of the pandemic. Crash frequency and severity are studied using negative binomial models and binary logit models, respectively, while accounting for exposure, environmental factors, and human factors. Additionally, Welch's t-test and pairwise t-test are used to explore any potential indirect effects of the pandemic on non-pandemic-related factors in the statistical models.

The results reveal that crash frequency is significantly lower than that of the pre-pandemic period throughout the pandemic. However, during the latter stages, crash frequency significantly increases due to relaxed restrictions. The severity levels of crashes were higher in the earlier stages of the pandemic, resulting from increased traffic speed, DUI prevalence, reduced seat belt use, and increased presence of commercial vehicles. It later decreased to a level comparable to that of the pre-pandemic period due to the reduction of speed and increased seat belt use. For the "new normal" phase, stakeholders should take measures to deter DUI and decrease commercial vehicle-related crashes to improve traffic safety.

3.2 Methodology

The study aims to investigate the impact of COVID-19 on both crash frequency and severity. Therefore, two sets of models are employed for crash frequency and severity analysis. Statistical models were employed to study the direct impact of the pandemic while accounting for the impact of other confounding factors. Widely used negative binomial (NB) models are utilized for crash frequency modeling, and logit models are utilized for crash severity analysis.

Moreover, one primary objective of the study is to understand whether the impact of the pandemic changes during its entire course. To achieve the objective, three different types of models are developed:

1. Models verifying the existence of the pandemic impact: Although many existing studies have pointed out that the pandemic has impacts on traffic safety, it is still worth verifying the existence of the impact using local data. In these types of models, apart from other explanatory variables, a binary dummy variable is employed to indicate whether a specific day falls in the pandemic period (crash frequency analysis), or a particular crash occurred during the pandemic (crash severity analysis).
2. Models examining the difference of the impact between the earlier and the latter stages of the pandemic: In these types of models, apart from other explanatory variables, a trinary dummy variable is utilized to indicate whether a specific day (crash frequency analysis) or a specific crash (crash severity analysis) falls before the pandemic or in the earlier or the latter stages of the pandemic.
3. Models exploring the impact of different pandemic quantifiers: It should be noted that in the aforementioned two types of models, the dummy pandemic indicators are the only pandemic-related variables. Therefore, the objective of this type of model is to understand the impact of various pandemic quantifiers (such as pandemic-related policies or number of new cases).

In total, six statistical models (2 types of analyses × 3 types of objectives) are developed.

Apart from the statistical models, further analysis is conducted to investigate the possible indirect effect of the pandemic by influencing other non-pandemic-related explanatory variables (such as environmental and human factors) in the statistical models. The analysis aims to provide insights that can benefit transportation agencies and policymakers. Contributing factors derived from the statistical models are compared to see whether they are statistically different before and during the pandemic as well as at the earlier and latter stages of the pandemic in terms of crash frequency and severity.

Crash Frequency Modeling: Negative Binomial Model

The study employs the widely used NB model to model the impact of COVID-19 on daily crash frequency (Lord & Mannering, 2010). An NB modal can be specified as follows:

$$\lambda_i = \exp(\boldsymbol{\beta X_i} + \varepsilon_i) \tag{3.1}$$

$$P(y_i) = \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{\Gamma(y_i + 1)\Gamma\left(\frac{1}{\alpha}\right)} \left(\frac{\frac{1}{\alpha}}{\frac{1}{\alpha} + \lambda_i}\right)^{\frac{1}{\alpha}} \left(\frac{\lambda_i}{\frac{1}{a} + \lambda_i}\right)^{y_i} \tag{3.2}$$

where $P(y_i)$ is the probability of entity i having y_i crashes in a given time period and $\Gamma(\cdot)$ is the gamma function; λ_i is the Poisson parameter, which is the expected number of crashes in the given time period; $\boldsymbol{X_i}$ is a set of explanatory variables; $\boldsymbol{\beta}$ is the corresponding coefficient set; ε_i is the error term and $\exp(\varepsilon_i)$ is gamma-distributed with mean 1 and variance α. Akaike information criterion (AIC) and pseudo-R^2 are used as the goodness-of-fit measures.

Crash Severity Modeling: Binary Logit Model

As for the crash severity analysis, crashes were classified into two classes: 1) with visible injury (K: fatal, A: incapacitating injury, and B: non-incapacitating injury); and 2) without visible injury (C: possible injury and O: no injury), and the class acts as the dependent variable of the crash severity analysis. Therefore, a binary logic modal is used to investigate the probability of a crash leading to injuries (positive outcome) against no injury (negative outcome) (Sze & Wong, 2007). A negative binomial modal can be specified as follows:

$$g(x) = \boldsymbol{\beta X_i} \tag{3.3}$$

$$\pi(x) = \frac{\exp\big(g(x)\big)}{1 + \exp\big(g(x)\big)} \tag{3.4}$$

where $\boldsymbol{X_i}$ is a set of explanatory variables; $\boldsymbol{\beta}$ is the corresponding coefficient set; $g(x)$ is a latent variable; and $\pi(x)$ is the conditional probability of the positive outcome, i.e., a crash leads to injuries. Akaike information criterion (AIC) and pseudo-R^2 are also used as the goodness-of-fit measures.

Comparison: Welch's T-Test & Holm–Bonferroni Method

When the only single comparison between two groups is needed, Welch's t-test (Welch, 1947) is employed since two groups may have unequal sizes and/or possibly unequal variances. Welch's t-test defines the statistic t by:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_1}}} \tag{3.5}$$

where $\overline{x_i}$, s_i, N_i are the sample mean, standard deviation, and size of sample i. The degree of the freedom df associated is calculated as follows:

$$df \approx \frac{\left(\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}\right)^2}{\frac{s_1^4}{N_1^2(N_1 - 1)} + \frac{s_2^4}{N_2^2(N_2 - 1)}} \tag{3.6}$$

When multiple comparisons are conducted simultaneously, pairwise t-tests are employed. To control the possible family-wise error rate, p-values are adjusted by Holm–Bonferroni method (Holm, 1979). The Holm-Bonferroni method first sorts m p-values of the pairwise t-tests into order lowest-to-highest $p_1, \dots, p_m$, and their corresponding null hypotheses $H_1, \dots, H_m$. Starting from p_1, at step k, test whether $p_k < \frac{\alpha}{m+1-k}$. If so, reject H_k and continue to test the larger p-values. This ensures that the family-wise error rate is less than the preset significant level α. It should be noted that although this method could control the family-wise error rate, it could sacrifice statistical power.

3.3 Data

The study selects the most populous metropolitan county, Salt Lake County in the State of Utah, as the study area (Figure 3.1). Five datasets are used in the study: 1) crash data from January 2019 to April 2022; 2) factors related to the pandemic; 3) Traffic data, including vehicle miles traveled (VMT) and speed of freeways within the county; and 4) weather conditions.

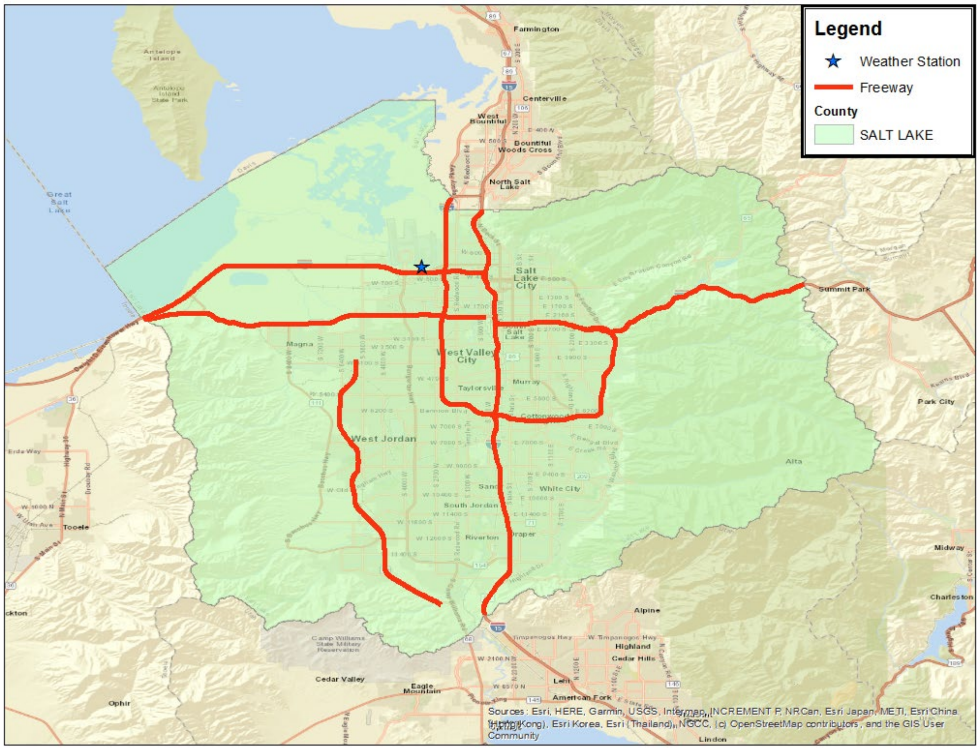


Figure 3.1 Freeways and the Weather Station in Salt Lake County of the State of Utah

Detailed crash data were collected from the Numetric system of the Utah Department of Transportation (UDOT), including crash time, injury severity, manner of collision, vehicles' characteristics, characteristics of people involved, and environmental conditions. Since the most important crash contributing factor, i.e., the exposure measure (such as VMT and traffic volume), is only available in detail for freeways during the study period, this study will only focus on crashes that occurred on the freeways. The daily number of crashes is used as the dependent variable of the crash frequency analysis. Several variables describing the crash's characteristics that are both available from the database and possibly related to the injury severity according to the existing studies were selected for the crash severity analysis. These include the manner of collision (AASHTO, 2010), light condition (H. Zhou et al., 2020), adverse roadway surface condition (Papadimitriou et al., 2019), adverse weather (J. Yuan et al., 2019), whether a commercial vehicle was involved in the crash (Chen et al., 2021), whether a driver was distracted (X. Dong et al., 2022), whether a driver was drowsy (X. Dong et al., 2022), whether a driver was driving under the influence (DUI), whether a motorcycle was involved in the crash (Chang & Wang, 2006), whether an older driver was involved in the crash (Yue et al., 2019), whether an overturn/rollover occurred (Conroy et al., 2006), and whether a driver was unrestrained (not wearing a seat belt) (USDOT, 2021). Further data cleaning was conducted for injury severity analysis to exclude crash records with missing or unknown values of these variables of interest.

The VMT of freeways in Salt Lake County, namely Interstate 15, Interstate 80, Interstate 215, State Road 85, and State Road 201, is used to quantify vehicular traffic. The average speed across the freeway is also collected as it may correlate to the crash severity (Wagner et al., 2020). It should also be noted that the crash data also refer to the same freeways. The VMT and speed data were collected from the UDOT Performance Measurement System (PeMS) (Utah Department of Transportation, 2019) during the study period. Weather conditions such as daily average temperature and total precipitation were collected from the nearest airport weather station (Salt Lake City International Airport, located at 40.77069°N, 111.96503°W, as shown in Figure 3.1) through the National Oceanic and Atmospheric Administration (*Find a Station | Data Tools | Climate Data Online [CDO] | National Climatic Data Center [NCDC]*, n.d.). The weather station is used because other weather stations in the study area are either located in mountainous areas far away from the freeways or unable to provide data every day.

Several factors related to the pandemic are collected as explanatory variables. The number of daily new COVID-19 confirmed cases and the percentage of deaths among the new confirmed cases were collected from the Utah Department of Health (*Coronavirus | Keeping Utah Informed on the Latest Coronavirus Updates*, n.d.). Many existing studies use the number of confirmed cases to quantify the severity of the pandemic. Those indicating the severity of diseases caused by the virus (i.e., deaths) are used to provide additional information. Note that using the absolute numbers of deaths may raise collinearity issues since they are highly correlated with the number of cases, and the rates are utilized instead. Pandemic-related policies (*Coronavirus | Keeping Utah Informed on the Latest Coronavirus Updates*, n.d.) were also reviewed. Two binary policy indicators, namely whether there were lockdown policies restricting travel directly and whether there were mask mandates that potentially influence people's willingness to travel, were then summarized. When a certain policy is effective on a specific day, the indicator is set to "1"; otherwise, it was set to "0".

Moreover, two dummy pandemic indicators were created. The first binary dummy variable, "During Covid," was employed in the first type of models to indicate the existence of the pandemic. The selection of the cutoff date is straightforward. March 12, 2020, which is when the first COVID-19 case was confirmed in the State of Utah, was chosen. If a crash occurred before March 12, 2020, the value of the dummy variable was assigned to be 1, otherwise, it was assigned to be 0. The second trinary variable, "Covid Stage," was employed in the second type of models to indicate the progression of the pandemic. It has three values: 0, 1, and 2, to indicate pre-pandemic, the earlier stage of the pandemic, and the latter stage of the pandemic. Besides March 12, 2020, the other cutoff date for earlier and latter stages of the pandemic is April 10, 2021, when the statewide mask mandate expired. The date was selected for several reasons. First, new statewide travel restrictions and mask mandates were never issued after that date (although the mask mandate was briefly placed in Salt Lake County during the outbreak related to the omicron variant), which could indicate the state government's policies have changed. Second, COVID vaccines were widely available, and the public started to be fully vaccinated after the date. Thus, the risk perception toward COVID-19 may be changed. Last, but most importantly, an earlier study by the authors (Gong et al., 2022) found that the VMT of Salt Lake County was restored to a level comparable to pre-pandemic levels around the date (see Figure 3.3), meaning the impact of the pandemic on mobility is decaying.

There were 17,038 crashes during the entire study period, and 16,748 crashes are used in the crash severity analysis after the data cleaning. Table 3.1 shows the number of crashes that occurred before and during each stage of the pandemic. The descriptive statistics of the variables used in the analysis can be found in Table A and Table B in the appendix.

Table 3.1 Number of Crashes Occurred During the Whole Study Period

Period		Number of Crashes	Crashes Used in Severity Models
Before Pandemic (During Covid = 0/Covid Stage =0)		7,295	7,221
During Pandemic (During Covid =1)	Earlier Stage (Covid Stage =1)	4,128	4,078
	Later Stage (Covid Stage =2)	5,565	5,449

3.4 Results and Discussion

3.4.1 Impact of COVID Pandemic on Crash Frequency

Trends

Figure 3.2 shows the number of crashes with the progression of the pandemic (note that although the statistical modeling uses daily data, weekly data are employed here for better illustration). Although crash frequency is significantly less than that of pre-pandemic levels during the whole course of the pandemic, it varied considerably between the earlier and latter stages. During the earlier stage of the pandemic, the number of crashes dropped dramatically when the lockdown was in place. Once the travel restrictions were relaxed, crash frequency gradually increased, but it remained low compared with the pre-pandemic period. However, at the latter stage of the pandemic, crash frequency gradually increased to slightly less but comparable to the pre-pandemic level. Another outbreak related to the omicron variant briefly reduced the crash frequency, and it is increasing to the previous level.

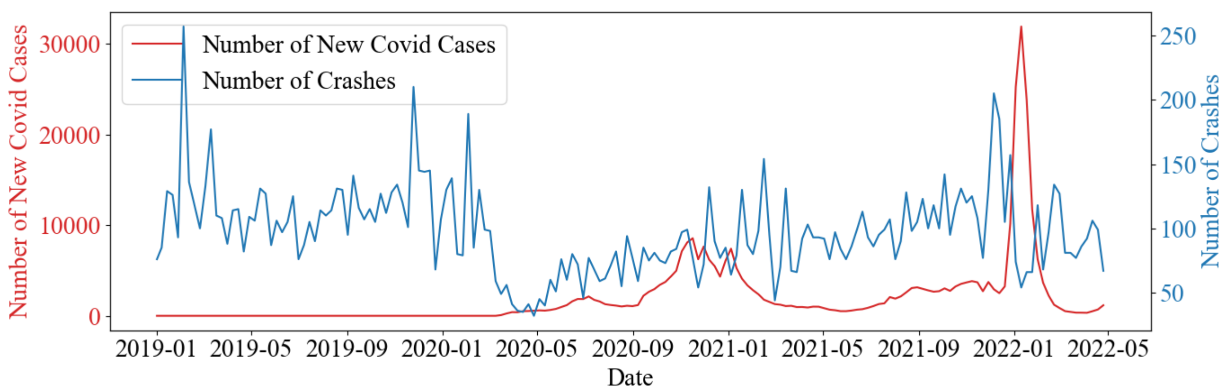


Figure 3.2 Weekly Number of Crashes versus Number of New COVID Cases

Modeling Results

The results of the three different types of NB models are shown in Table 3.2. Note that the variance inflation factor (VIF) of all variables in the three models was checked to avoid the collinearity issue. All VIFs are less than 5, which indicates that the collinearity issue should not be concerned (*43*).

Table 3.2 Estimates of Crash Frequency Models

Variable	Estimates	Std. Error	Z Value	P Value
With Only Binary Covid Indicator				
(Intercept)**	**-19.2407**	**1.3060**	**-14.7325**	**<0.0001**
Ln (VMT)**	**1.3834**	**0.0816**	**16.9581**	**<0.0001**
Average Temperature**	**-0.0066**	**0.0008**	**-8.1761**	**<0.0001**
Total Precipitation**	**1.7904**	**0.1193**	**15.0028**	**<0.0001**
During Covid: Yes**	**-0.1082**	**0.0319**	**-3.3950**	**0.0007**
Observations	1216			
AIC	7929.3			
Pseudo-R^2	0.474			
With Only Trinary Covid Indicator				
(Intercept)**	**-17.7010**	**1.3407**	**-13.2032**	**<0.0001**
Ln (VMT)**	**1.2875**	**0.0837**	**15.3756**	**<0.0001**
Average Temperature**	**-0.0064**	**0.0008**	**-8.0544**	**<0.0001**
Total Precipitation**	**1.7416**	**0.1187**	**14.6738**	**<0.0001**
Covid Stage: Earlier**	**-0.1997**	**0.0387**	**-5.1647**	**<0.0001**
Covid Stage: Later	-0.0429	0.0356	-1.2032	0.2289
Observations	1216			
AIC	7915.2			
Pseudo-R^2	0.487			
With Covid Quantifier				
(Intercept)**	**-16.8858**	**1.3409**	**-12.5931**	**<0.0001**
Ln (VMT)**	**1.2379**	**0.0837**	**14.7831**	**<0.0001**
Average Temperature**	**-0.0071**	**0.0008**	**-8.8887**	**<0.0001**
Total Precipitation**	**1.7112**	**0.1179**	**14.5093**	**<0.0001**
Number of New Covid Cases	-0.00002	<0.0001	-0.8347	0.4039
Death Rate	0.0469	0.0355	1.3214	0.1864
Lockdown: Yes**	**-0.5601**	**0.0901**	**-6.2153**	**<0.0001**
Mask Mandate: Yes**	**-0.1369**	**0.0389**	**-3.5178**	**0.0004**
Observations	1216			
AIC	7900.2			
Pseudo-R^2	0.502			

* Statistically significant at 0.05 level.
** Statistically significant at 0.01 level.

According to Table 3.2, the exposure measure VMT is significantly and positively related to crash frequency, which is expected (AASHTO, 2010). Precipitation is significantly positively associated with crash frequency since it may lead to adverse road surface conditions and low visibility (Sheather, 2009), and thus increase the crash risk. Although the effect is relatively low, daily average temperature is negatively related to crash frequency. A possible reason is that during the winter when the temperature is low, precipitation is likely to be in the form of snow, which leads to an even higher risk (Qiu & Nixon, 2008). The aforementioned variables are statistically significant in all three models.

As for the pandemic-related parameters, in the first model, the COVID dummy variable is statistically significant with a negative coefficient, meaning that crash frequency was reduced during the pandemic while accounting for other confounding factors. However, while accounting for the other factors, the second model reveals that crash frequency reduction is only statistically significant during the earlier stage of the pandemic, which is also found in many other existing studies (N. Dong et al., 2022; Doucette et al., 2021; Ebrahim Shaik & Ahmed, 2022; Islam et al., 2022; Katrakazas et al., 2020; Koloushani et al., 2021; Pathak et al., 2022; Qureshi et al., 2020; Research, 2021a, 2021b; Sekadakis et al., 2021; Wagner et al., 2020), but not during the latter stage of the pandemic.

The results of the third model reveal some possible reasons for the difference. Both pandemic-related policies are found to be significantly and negatively related to crash frequency, while the lockdown has a stronger impact. Note that the lockdown policy was only in place during the early pandemic stage, and the number of days when wearing masks is mandated is significantly higher early in the pandemic (see Table B in the appendix for more details). The differences in policies contribute to the different crash frequencies between the earlier and latter stages of the pandemic. While the lockdown is widely known to reduce the crash frequency (Islam et al., 2022; Koloushani et al., 2021; Qureshi et al., 2020; Wagner et al., 2020), the mask mandate's positive impact on traffic safety is found for the first time. We suspect that pandemic-related policies may be related to human factors that are not explicitly modeled. First, pandemic-related policies may impact the public's risk perception (Duan et al., 2020), which in turn impacts their travel and driving behaviors. Second, government policies directly alter travel behaviors. In addition to the lockdown orders that directly restrict traveling, the so-called "social distancing" policies encourage remote working during the earlier pandemic stage, and working from home is negatively related to crash frequency (Abdel-Aty et al., 2013).

Discussion

Admittedly, the pandemic may also indirectly impact crash frequency by influencing other factors, as suggested by earlier research (N. Dong et al., 2022; Doucette et al., 2021; Ebrahim Shaik & Ahmed, 2022; Islam et al., 2022; Katrakazas et al., 2020; Koloushani et al., 2021; Pathak et al., 2022; Qureshi et al., 2020; Research, 2021a, 2021b; Sekadakis et al., 2021; Wagner et al., 2020). Table 3.3 shows the results of Welch's t-test and pairwise t-test, while Figure 3.3 shows the number of crashes against the VMT. The significantly higher VMT during the latter stage of the pandemic could also be a contributing factor to the higher crash frequency compared with the earlier stage of the pandemic.

Table 3.3 Results from Welch's T-Test and Pairwise T-Test for Crash Frequency Analysis

Variable	Before/During		Stages (Adjusted P Values Only)#		
	T Value	P Value	Before/Earlier	Before/Later	Earlier/Later
Number of Crashes	**6.8209**	**<0.0001****	**<0.0001****	**0.0033****	**<0.0001****
Ln (VMT)	**7.7957**	**<0.0001****	**<0.0001****	**0.0341***	**<0.0001****
Average Temperature	**-4.0518**	**0.0001****	**0.0027****	**0.0007****	0.6141
Total Precipitation	**3.0739**	**0.0022****	**0.0006****	0.1251	0.1251
Mask Mandate: Yes	N/A		N/A	N/A	**<0.0001****

* Statistically significant at 0.05 level.
** Statistically significant at 0.01 level.
# The t values of pairwise t-tests may be misleading since the p values were adjusted. Therefore, they were omitted.

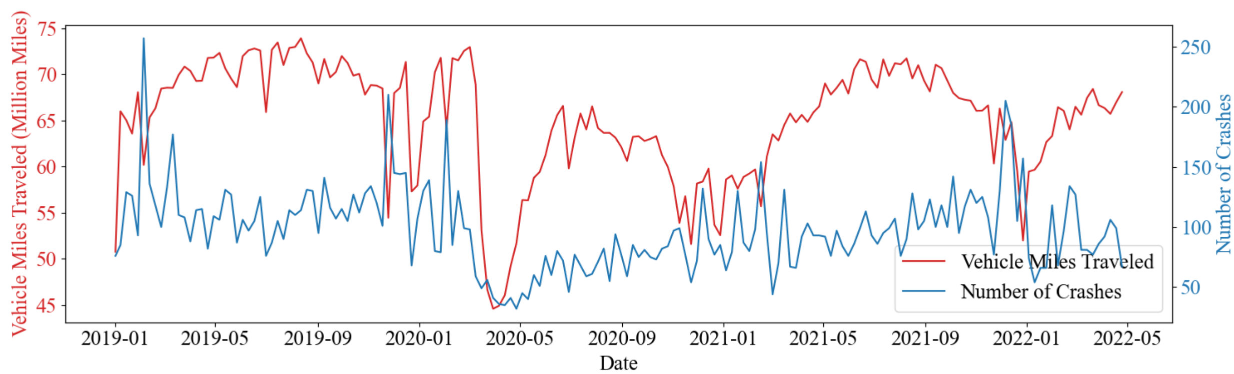


Figure 3.3 Weekly Number of Crashes versus VMT

3.4.2 Impact of COVID Pandemic on Crash Severity

Trends

Figure 3.4 shows the percentage of injury crashes ("Injury Rate") with the progression of the pandemic. Different from the crash frequency, the crash severity increased significantly during the earlier stage of the pandemic, but it generally reduced to a level comparable to the pre-pandemic during the later stage.

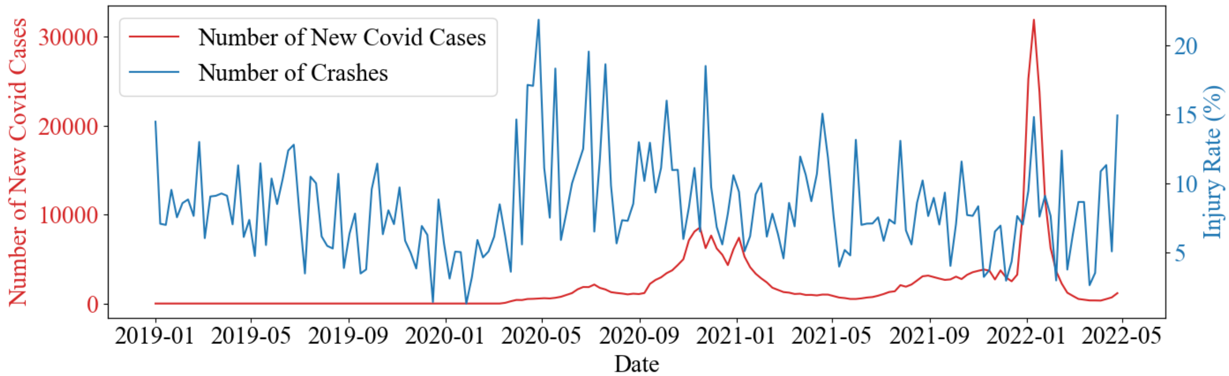


Figure 3.4 Weekly Average Injury Rate versus Number of New COVID Cases

Modeling Results

The results of the three different types of binary logit models are shown in Table 3.4. VIFs were also checked to clear the concern of the collinearity issue.

Table 3.4 Estimates of Injury Severity Models

Variable	Estimates	Std. Error	Z Value	P Value
With Only Binary Covid Indicator				
(Intercept)**	**-6.4105**	**0.8011**	**-8.0019**	**<0.0001**
Average Speed**	**0.0536**	**0.0123**	**4.3471**	**<0.0001**
During Covid: Yes	0.0496	0.0636	0.7802	0.4352
Manner of Collision: Angle**	**1.6022**	**0.0987**	**16.2384**	**<0.0001**
Manner of Collision: Head On**	**1.5014**	**0.2741**	**5.4767**	**<0.0001**
Manner of Collision: Single Vehicle	0.0369	0.0789	0.4683	0.6396
Manner of Collision: Parked Vehicle	0.3537	0.4362	0.8108	0.4175
Manner of Collision: Rear to Rear	-10.4082	199.9480	-0.0521	0.9585
Manner of Collision: Rear to Side	0.8487	1.0603	0.8004	0.4235
Manner of Collision: Sideswipe Opposite Direction	0.8689	0.5211	1.6673	0.0955
Manner of Collision: Sideswipe Same Direction**	**-0.6036**	**0.1022**	**-5.9083**	**<0.0001**
Daylight Condition	-0.1331	0.0681	-1.9526	0.0509
Commercial Vehicle Involved*	**0.2166**	**0.0975**	**2.2230**	**0.0262**
Distracted Driving Involved**	**0.5656**	**0.1085**	**5.2140**	**<0.0001**
Drowsy Driving Involved**	**0.8825**	**0.1666**	**5.2968**	**<0.0001**
DUI Involved**	**1.1897**	**0.1109**	**10.7312**	**<0.0001**
Motorcycle Involved**	**2.6310**	**0.1974**	**13.3286**	**<0.0001**
Older Driver Involved**	**0.4411**	**0.0955**	**4.6174**	**<0.0001**
Overturn or Rollover Involved**	**1.6846**	**0.1112**	**15.1462**	**<0.0001**
Unrestrained Involved**	**1.7704**	**0.1353**	**13.0840**	**<0.0001**
Observations	16748			
AIC	8023.1			
Pseudo-R^2	0.126			
With Only Trinary Covid Indicator				
(Intercept)**	**-6.2137**	**0.8169**	**-7.6061**	**<0.0001**
Average Speed**	**0.0505**	**0.0126**	**4.0238**	**0.0001**
During Covid Earlier Stage: Yes	**0.0986**	**0.0785**	**1.2565**	**0.2089**
During Covid Later Stage: Yes	**0.0120**	**0.0733**	**0.1639**	**0.8698**
Manner of Collision: Angle**	**1.5969**	**0.0988**	**16.1637**	**<0.0001**

Variable	Estimates	Std. Error	Z Value	P Value
Manner of Collision: Head On**	**1.4963**	**0.2737**	**5.4659**	**<0.0001**
Manner of Collision: Single Vehicle	0.0318	0.0791	0.4019	0.6878
Manner of Collision: Parked Vehicle	0.3504	0.4357	0.8041	0.4213
Manner of Collision: Rear to Rear	-10.4440	200.0260	-0.0522	0.9584
Manner of Collision: Rear to Side	0.8351	1.0602	0.7877	0.4309
Manner of Collision: Sideswipe Opposite Direction	0.8580	0.5220	1.6436	0.1003
Manner of Collision: Sideswipe Same Direction**	**-0.6075**	**0.1023**	**-5.9403**	**<0.0001**
Daylight Condition	-0.1319	0.0681	-1.9353	0.0530
Commercial Vehicle Involved*	**0.2187**	**0.0975**	**2.2434**	**0.0249**
Distracted Driving Involved**	**0.5653**	**0.1085**	**5.2106**	**<0.0001**
Drowsy Driving Involved**	**0.8857**	**0.1666**	**5.3159**	**<0.0001**
DUI Involved**	**1.1921**	**0.1108**	**10.7545**	**<0.0001**
Motorcycle Involved**	**2.6349**	**0.1974**	**13.3492**	**<0.0001**
Older Driver Involved**	**0.4436**	**0.0956**	**4.6418**	**<0.0001**
Overturn or Rollover Involved**	**1.6832**	**0.1113**	**15.1290**	**<0.0001**
Unrestrained Involved**	**1.7715**	**0.1352**	**13.0988**	**<0.0001**
Observations	16748			
AIC	8024			
Pseudo-R^2	0.127			
With Covid Quantifier				
(Intercept)**	**-6.3187**	**0.8127**	**-7.7753**	**<0.0001**
Average Speed**	**0.0523**	**0.0125**	**4.1954**	**<0.0001**
Number of New Covid Cases	0.0000	0.0001	-0.0152	0.9879
Death Rate	-0.0196	0.0853	-0.2300	0.8181
Lockdown: Yes	0.1432	0.2140	0.6690	0.5035
Mask Mandate: Yes	0.0594	0.0824	0.7200	0.4715
Manner of Collision: Angle**	**1.6000**	**0.0988**	**16.2020**	**<0.0001**
Manner of Collision: Head On**	**1.4953**	**0.2739**	**5.4589**	**<0.0001**
Manner of Collision: Single Vehicle	0.0359	0.0789	0.4545	0.6495
Manner of Collision: Parked Vehicle	0.3525	0.4358	0.8090	0.4185
Manner of Collision: Rear to Rear	-10.4314	200.0034	-0.0522	0.9584
Manner of Collision: Rear to Side	0.8442	1.0602	0.7962	0.4259
Manner of Collision: Sideswipe Opposite Direction	0.8555	0.5217	1.6398	0.1011
Manner of Collision: Sideswipe Same Direction**	**-0.6050**	**0.1022**	**-5.9176**	**<0.0001**
Daylight Condition	-0.1310	0.0683	-1.9171	0.0552
Commercial Vehicle Involved*	**0.2191**	**0.0975**	**2.2466**	**0.0247**
Distracted Driving Involved**	**0.5646**	**0.1085**	**5.2046**	**<0.0001**
Drowsy Driving Involved**	**0.8874**	**0.1667**	**5.3231**	**<0.0001**
DUI Involved**	**1.1905**	**0.1108**	**10.7429**	**<0.0001**
Motorcycle Involved**	**2.6337**	**0.1974**	**13.3388**	**<0.0001**
Older Driver Involved**	**0.4432**	**0.0956**	**4.6365**	**<0.0001**
Overturn or Rollover Involved**	**1.6826**	**0.1113**	**15.1215**	**<0.0001**
Unrestrained Involved**	**1.7725**	**0.1353**	**13.1018**	**<0.0001**
Observations	16748			
AIC	8028.7			
Pseudo-R^2	0.126			

* Statistically significant at 0.05 level.

** Statistically significant at 0.01 level.

The effects of the variables related to the manner of collisions, the characteristics of vehicles involved, and the drivers' behavior are in line with previous studies. Speed is positively and significantly related to crash severity(X. Dong et al., 2022; Wagner et al., 2020). Angle, head-on, overturn, and rollover crashes tend to be severer and sideswipe crashes are likely less severe (AASHTO, 2010). The severity of crashes with commercial vehicles (Chen et al., 2021) and/or motorcycles (Chang & Wang, 2006) involved tends to be severe. Crashes with the older driver (Yue et al., 2019) involved are likely to be severer. Risky driving behavior, including DUI (Watson-Brown et al., 2021), distracted driving (X. Dong et al., 2022), drowsy driving (X. Dong et al., 2022), and unrestrained (not wearing the seat belt) (X. Dong et al., 2022; Wagner et al., 2020) could increase the crash severity.

Interestingly but not surprisingly, all pandemic-related variables are not statistically significant even at the 0.05 level in all three models. Similar results can be found in an earlier study (X. Dong et al., 2022), which found that the conventional statistical model suggests an insignificant impact of the pandemic on crash severity.

Discussion

Another important conclusion of the aforementioned study (X. Dong et al., 2022) is that the pandemic does have indirect effects on driving behavior, e.g., increasing aggressiveness and inattentiveness of drivers, which leads to a higher likelihood of severe crashes. Thus, a plausible reason for the insignificance of the pandemic-related variables is that the impact of the pandemic can be well explained by the other variables including speed and crash characteristics.

Therefore, to investigate the possible contributing factors to different crash severity levels before and during the different stages of the pandemic, the mean values of explanatory variables were compared. Table 3.5 presents the results of the comparison.

Table 3.5 Results of Welch's T-Test and Pairwise T-Test for Crash Severity Analysis

Variable	Before/During		Stages (Adjusted P Values Only) [#]		
	T Value	P Value	Before/Earlier	Before/Later	Earlier/Later
Whether a Crash Leads to Injuries	**-2.9360**	**0.0033****	**0.0001**	0.3534	**0.0033**
Average Speed (mph)	**-4.0576**	**0.0001****	**<0.0001****	0.4612	**<0.0001****
Manner of Collision: Angle	-0.1120	0.9108	0.1478	0.1613	**0.0117***
Manner of Collision: Head On	**2.2980**	**0.0216****	0.4559	**0.0207***	0.4559
Manner of Collision: Sideswipe Same Direction	**-4.7202**	**<0.0001****	**0.0001**	**0.0004****	0.5093
Commercial Vehicle Involved	**-5.4457**	**<0.0001****	**<0.0001****	**<0.0001****	**<0.0001****
Distracted Driving Involved	**2.2421**	**0.0250***	0.2679	0.0789	0.6115
Drowsy Driving Involved	**-2.1976**	**0.0280***	0.0638	0.3475	0.3475
DUI Involved	**-6.2752**	**<0.0001****	**<0.0001****	**<0.0001****	**<0.0001****
Motorcycle Involved	**-1.9716**	**0.0487***	0.2229	0.2987	0.6345
Older Driver Involved	0.6186	0.4978	0.1269	0.4790	0.0536
Overturn/Rollover Involved	**-1.9865**	**0.0470***	**0.0023****	0.9624	**0.0030****
Unrestrained Involved	**-2.8160**	**0.0049****	**0.0225***	0.0997	0.3776

* Statistically significant at 0.05 level.
** Statistically significant at 0.01 level.
[#] The t values of pairwise t-tests may be misleading since the p values were adjusted. Therefore, they were omitted.

According to the results of pairwise t-tests and descriptive statistics (see Table B in the appendix for more details), during the earlier stage of the pandemic, the average speed is increasing, which is also found by other earlier studies (X. Dong et al., 2022; Research, 2021a, 2021b; USDOT, 2021; Vanlaar et al., 2021; Wagner et al., 2020). The increase in DUI and decrease in seat belt use are the major findings of a series of NHTSA's national studies (Research, 2021a, 2021b; USDOT, 2021; Wagner et al., 2020), which also be found in this study. The proportion of crashes that involves commercial vehicles increased, leading to an increase in crash severity according to the results of all three logic models. The change in crash types exhibits mixed effects. Crashes with overturn/rollover involved increased but the less severe sideswipe crashes also increased. Overall, the crash severity increased during the earlier stages of the pandemic.

However, the situation got improved with the progression of the pandemic. The use of the seat belt increased to a level that is comparable to the pre-pandemic. The percentage of overturn/rollover crashes also decreased to the pre-pandemic level, but the percentage of less severe sideswipe crashes is still higher than the pre-pandemic. Even the concerning DUI is decreasing, although it is still more frequent than the pre-pandemic. The only major issue is that the percentage of crashes that involve commercial vehicles remains higher than in the pre-pandemic and even higher than in the earlier stage of the pre-pandemic, which may be due to the increased truck traffic caused by the growth of online shopping and on-demand delivery (Gong et al., 2022). All these factors together lead to a decrease in crash severity during the later stages of the pandemic to a level comparable to the pre-pandemic.

There might be two possible reasons for the change in the crash severity between the earlier and the latter stage of the pandemic. First, during the earlier pandemic, when the public feared getting infected, those still on the road may have had higher degrees of risk acceptance (Vanlaar et al., 2021). Therefore, they may have had a higher probability of not using seat belts, which has gradually changed with the change in the public's risk perception toward COVID-19, especially when people were getting vaccinated. When more and more people started driving again during the latter pandemic stage, the average level of risk acceptance returned to the pre-pandemic level. Second, surveys indicate that people started, or increased, substance use to cope with pandemic-related stress or emotions (Wagner et al., 2020), which could have also increased the probability of DUI, even during the latter pandemic stage. However, with the relaxation of the restrictive measures in the latter stage of the pandemic, people are gradually returning to their normal lives, which may relieve the pandemic-related stress or emotions, leading to reduced drug and alcohol use.

3.4.3 Outlook and Policy Discussions

While crash frequency increased and crash severity decreased to the pre-pandemic levels, we could expect an overall safety performance is similar to, if not better than, the case before the pandemic. However, the new normal is never an "old normal." First of all, two issues regarding crash severity still need attention. First, although the percentage of crashes involving DUI is decreasing, it still remains higher than in pre-pandemic days. Effective law enforcement and public campaigns are needed to deter DUI (NHTSA, 2022; Wagner et al., 2020). Second, the percentage of commercial vehicle-involved crashes is increasing even during the latter stage of the pandemic because of the increased popularity of online shopping and on-demand delivery (AASHTO Journal, 2020; Christidis et al., 2021b). Educational programs for commercial vehicle drivers, internal monitoring of drivers' behavior by their companies, and external regulations (Chen et al., 2021; T. Zhou & Zhang, 2019) may help to reduce commercial vehicle-related crashes. Moreover, although living with COVID-19 is inevitable, there are still some uncertainties. Since the virus is evolving rapidly, a possible new variant could lead to widespread infection and/or a drastic increase in deaths, like what the omicron variant did during the 2021-2022 winter season. During the possible seasonal endemic, travel volumes may again decrease, seat belt use may again decrease, and policies such as mask-mandates may again be invoked. All of these can impact traffic safety. The impacts of possible seasonal endemics should be considered to make traffic-safety-related policies more robust. Nevertheless, the impact of the COVID-19 pandemic on traffic safety can serve as a valuable reference for policymakers in response to future pandemics.

3.5 Conclusion

The global COVID-19 pandemic has significantly impacted traffic safety across the U.S. However, few studies investigated the pandemic's impact on traffic safety during the latter stage of the pandemic. Therefore, this study employs several statistical methods to investigate whether the impact of COVID-19 on traffic safety differs during the different stages. Freeways of Salt Lake County, Utah, were selected as the study sites. Negative binomial models and binary logit models were utilized to study the effects of the pandemic on crash frequency and severity, respectively, while accounting for the exposure, environmental factors, and human factors. Welch's t-test and pairwise t-test are employed to investigate the possible indirect effect of the pandemic by influencing other non-pandemic-related explanatory variables in the statistical models. The results show that crash frequency is significantly less than that of the pre-pandemic during the entire course of the pandemic while accounting for the other factors. However, it is significantly higher during the latter stage due to the relaxed restrictions. Crash severity levels increased during the earlier pandemic stage due to increased speed, the prevalence of DUI, reduced seat belt use, and increased presence of commercial vehicles. But later it was reduced to a level comparable to pre-pandemic times, owing to reduced speed and increased seat belt use comparable to pre-pandemic levels. As for the incoming new normal, stakeholders may need to take action to deter DUIs and reduce commercial-vehicle-related crashes to improve traffic safety.

4. IMPACT OF COVID-19 ON TRAVEL PATTERN

4.1 Overview

In this section, the effect of COVID-19 on traffic patterns in Salt Lake County and Utah County from January 2019 to July 2021 is analyzed. Different vehicle miles traveled (VMT) patterns in the pre-pandemic stage, early stage of the pandemic, and late stage of the pandemic are identified.

The results indicate that VMT is significantly affected by the severity of the pandemic (measured by the number of new cases), policies, and individual/societal risk perceptions of traveling during the pandemic. In the early stage of the pandemic, vehicular traffic decreases due to government restrictions and individuals' risk perception. However, in the latter stage of the pandemic, with the relaxation of travel restrictions and increasing vaccine rates, vehicular traffic has either recovered to or exceeded pre-pandemic levels. Specifically, truck traffic is higher than the pre-pandemic level due to the growth of online shopping and on-demand delivery.

The traffic patterns revealed in this chapter can assist agencies in better understanding the impacts of COVID-19 on traffic mobility and potentially support long-term urban planning strategic goals during the post-pandemic period.

4.2 Data Description

4.2.1 Research Area

The study selects the major metropolitan counties, Salt Lake County and Utah County, in the State of Utah as the study area. These two counties are the most populous in Utah and account for more than 55% of the state's population (2020 estimate). The majority difference between the two counties is the political diversity. Salt Lake County is one of Utah's more politically diverse areas while Utah County is less diverse. This translates into differences in county policies toward the pandemic.

This study focuses on traffic patterns of freeways due to data availability. Figure 4.1 shows the freeways in the study area. The traffic dynamics of two main freeway corridors, Interstate-15 and Interstate-80, and other smaller freeways, including Interstate 80, Interstate 215, State Road 85, State Road 201, State Road 92, U.S. Route 6, and U.S. Route 189, are studied. There are 850 lane miles and 540 lane miles covered for Salt Lake County and Utah County, respectively.

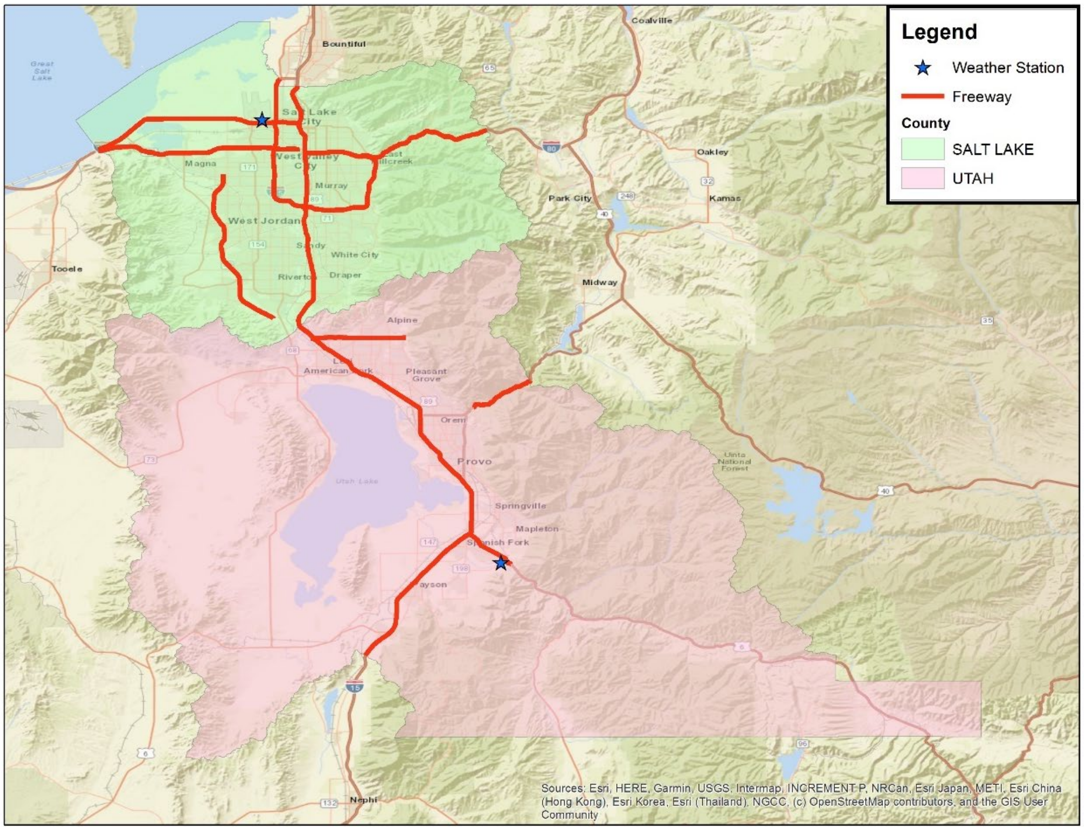


Figure 4.1 Freeways in Salt Lake County and Utah County in the State of Utah

4.2.2 Data

County-wide VMT of all freeways are used to quantify the vehicular dynamic. The VMT data were collected from the UDOT Performance Measurement System (PeMS) (Utah Department of Transportation, 2019) from January 2019 to the first week of July 2021, which gives the traffic patterns before and during the pandemic. VMT for all vehicle types and trucks are collected and analyzed separately.

Various factors related to the pandemic and vehicular traffic are also collected as explanatory variables. New COVID-19 confirmed cases and the percentage of fully vaccinated individuals over 12 years of age are received from the Utah Department of Health (The State of Utah, 2021a). The former is a direct quantifier of the pandemic's severity, which has been used by several existing studies (Kim, 2021); the latter impacts people's risk perception in traveling.

Traffic is intrinsically related to the economy. Economic factors are also used as explanatory variables to capture the unexplained heterogeneity by the pandemic-related factors. The monthly unemployment rate and daily news sentiment index are obtained from the Utah Department of Workforce Services (Utah Department of Workforce Services, 2021) and the Federal Reserve Bank of San Francisco (Shapiro et al., 2017). The unemployment rate can affect traffic as fewer people will be willing to drive due to economic hardship, especially commuting trips. The daily news sentiment index is a measure of economic sentiment based on a lexical analysis of economics-related news articles from 24 major U.S. newspapers. The developers of the index created a sentiment scoring model based on publicly available lexicons with a news-specific lexicon constructed by the developers. The scores of individual articles are then aggregated into a daily time-series measure of news sentiment that is statistically adjusted to account for changes in the composition of the sample across newspapers. Then the index is constructed as a trailing weighted average of time series, with weights that decline geometrically with the length of time since article publication. The index provides information regarding economic downturns and overall sentiment in the public eye.

Weather conditions could affect traffic operations. For example, traffic volumes could be reduced by rainstorms and from snowstorms (Maze et al., 2006). Weather-related parameters such as temperature, precipitation, and snow depth were collected from the National Centers for Environmental Information (National Centres for Environmental Information (NCEI), 2020). The weather observing station at the Salt Lake International Airport was selected for weather data in Salt Lake County, and another station at Spanish Fork Power House was selected for the weather data in Utah County (their locations are shown in Figure 4.1 as blue stars). These weather stations were selected since they are located near analyzed freeways. All aforementioned data are numeric.

Another important factor is pandemic-related policy. For instance, lockdown orders could significantly reduce the traffic volume (Kim, 2021). Important statewide and county-level policies (State of Utah, 2021b) are listed in Figure 4.2 and include deceleration of state of emergency, mask mandates, administration of vaccines, and others. Some policies restrict traveling directly while others influence people’s willingness to travel. Policy indicators are pre-processed as 0-1 dummy variables. When a certain policy is effective at a specific time, the dummy variable was set to 1; otherwise, it was set to 0.

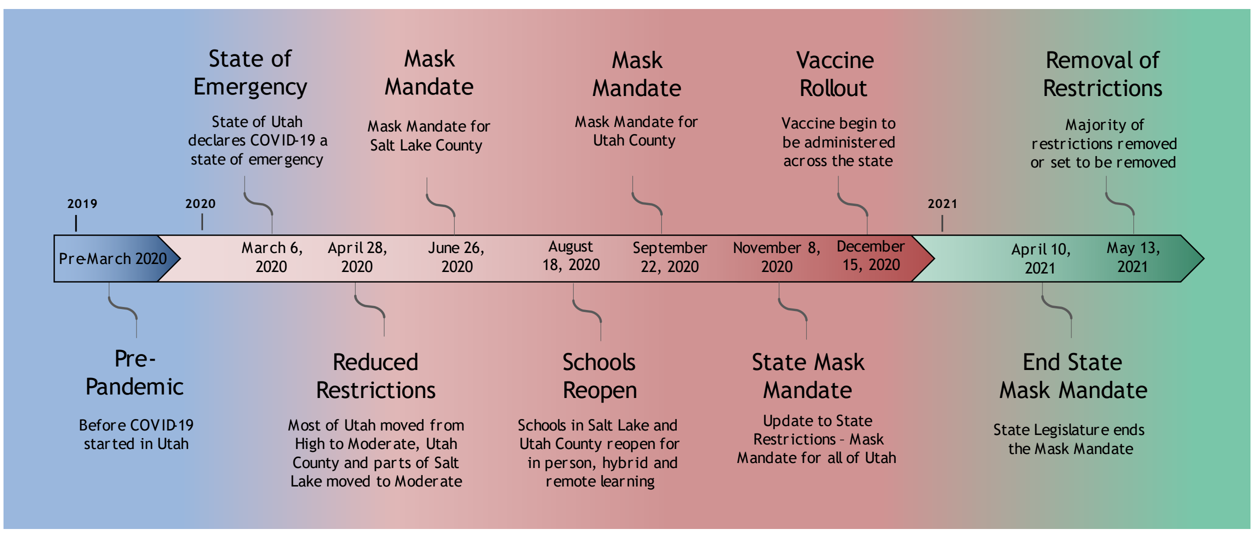


Figure 4.2 Milestones of Pandemic-Related Policies

All data are aligned with VMT data and aggregated by week. In other words, economic and weather data are collected from January 2019 to the first week of July 2021, while pandemic-related data and policies are collected once they are available until the first week of July 2021.

4.3 Results and Discussion

VMT is highly affected by the severity of the pandemic (quantified by the number of new cases), policies, and individual/societal risk perceptions of traveling during the pandemic. Figure 4.3 shows the VMT for all types of vehicles (Total VMT) and trucks for two counties with the progression of the pandemic. It should be noted that the Truck VMT shown in the figure is scaled up by 10 times for better illustration. The most prolonged decrease in VMT occurs during the initial phase of the pandemic. After travel restrictions regarding COVID-19 were announced (around the second week of March 2020), total VMT dropped significantly by around 38.9% for Salt Lake County (72.9 million miles vs. 44.6 million miles) and 36.7% for Utah County (37.7 million miles vs. 23.8 million miles) in one month. Truck VMT also dropped by 26.2% (4.6 million miles vs. 3.6 million miles) and 19.8% (3.4 million miles vs. 2.7 million miles) for Salt Lake and Utah Counties, respectively. Travel restrictions issued by the state government and public concerns regarding the virus both lead to this large and prolonged drop.

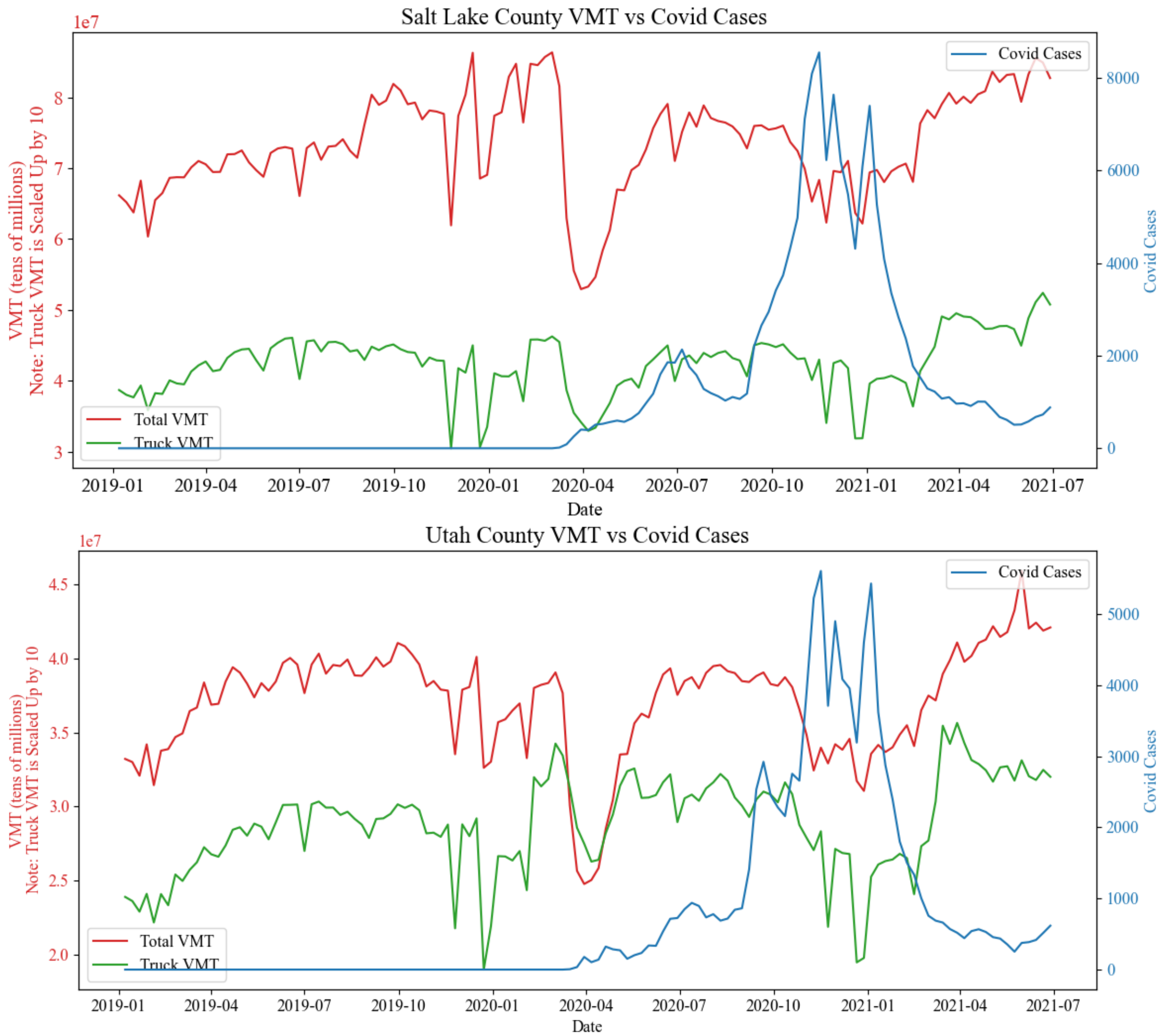


Figure 4.3 VMT versus Number of New COVID Cases (Truck VMT is Scaled up by 10 Times for Better Illustration)

Once restrictions were relaxed and public perceptions shifted, even during rising case counts, starting from the last week of April 2020, VMT began to recover to pre-pandemic levels. By mid-June 2020, total VMT recovered to 89.9% (65.6 million miles) and 99.6% (37.5 million miles) of pre-pandemic levels for Salt Lake and Utah Counties, respectively. Truck VMT shows a similar trend. Note that the VMT recovered faster and stronger for Utah County. To this date, there is not a huge difference in pandemic-related policies between the two counties; the differences are likely due to the different risk perceptions of the residents living/traveling in the two counties. Interestingly, implementing mask mandates did not significantly reduce VMT.

During the late months of 2020 and the beginning of 2021, VMT began to drop again. This decrease can partially be attributed to holidays during this time of year, especially for truck VMT. However, the total VMT of the first week in 2021 (52.6 million miles for Salt Lake County and 29.9 million miles for Utah County) is less than that of the first week in 2020 (57.9 million for Salt Lake County and 31.8 million miles for Utah County). The slight drop can be attributed to increased state restrictions and the highest case count totals of COVID-19. These high case counts and restrictions remained throughout early 2021, causing the decline in VMT to stagnate as the high case counts persisted.

During the latter stage of the pandemic, the vaccines plays an important role in the recovery of traffic. Figure 4.4 shows the changes in VMT with the percentage of fully vaccinated individuals over 12 years of age. COVID vaccines began to be administered in the State of Utah in December 2020 during the peak of the pandemic. In the ensuing months, as vaccines became more available, new confirmed cases began to decrease significantly. A Pearson correlation study reveals that total VMT is positively correlated to the percentage of fully vaccinated individuals with coefficients of 0.837 and 0.937 for Salt Lake and Utah counties, respectively (Figure 4.4). At the end of the study period, VMT gradually recovered to, if not exceeded, pre-pandemic levels. In late June 2021, the total VMT is 75.8 million miles for Salt Lake County and 32.5 million miles for Utah County, which respectively yields a 2.2% and 6.8% increase compared with total VMT of late February 2020. Again, VMT increased faster and more so for Utah County due to similar reasons. Truck VMT increased by more than 10% (4.6 million miles vs. 5.2 million miles) for Salt Lake County. A possible reason is the significant growth of online shopping and on-demand delivery during the pandemic (AASHTO Journal, 2020; Christidis et al., 2021a; Shamshiripour et al., 2020). These trends increased during the crisis as a response to limitations in retailing, risk aversion, and social distancing. The growth of online shopping requires more delivery truck trips, then in turn significantly increases the truck VMT.

Figure 4.4 VMT versus Percentage of Fully Vaccinated People (Truck VMT is Scaled up by 10 Times for Better Illustration)

In summary, the pandemic significantly impacts the traffic in Utah's Salt Lake County and Utah County. At the earlier stage, traffic drops significantly due to the direct travel restrictions. Once the restrictions have been reduced, traffic has gradually recovered while the recovering process may be impacted by the pandemic if it is severe. As the vaccinated population continuously grows, traffic in both counties has fully recovered to the pre-pandemic level. The recovery is much faster than the forecast by Christidis et al. (2021), which stated that vehicular traffic will return to the pre-pandemic level around 2025. Truck traffic even increased as a result of ever-increasing online shopping activities.

4.4 Conclusion

The global COVID-19 pandemic has had a great impact on the traffic across the U.S. However, there are few studies investigating the pandemic's impact on vehicular traffic at the latter stage of the pandemic. Therefore, this section studied the change of freeway traffic patterns in two metropolitan Utah counties, Salt Lake County and Utah County, during the pandemic. We conclude that:

1. Vehicular traffic is decreased during the early stage of the pandemic due to the government restrictions and individuals' risk perception in traveling.
2. With the relaxation of travel restrictions and COVID vaccines, vehicular traffic has recovered to, if not exceeded, pre-pandemic levels.
3. Truck traffic at the latter stage of the pandemic is higher than the pre-pandemic level due to the growth of online shopping and on-demand delivery.

The summarized traffic patterns at the latter stage of the pandemic could help transportation agencies better understand the impacts of COVID-19 on traffic mobility. These can also potentially support the long-term urban planning strategic goals during the post-pandemic periods. For example, relevant agencies need to prepare adequate facilities such as truck parking and rest facilities in response to increasing truck traffic.

5. KNOWLEDGE-BASED MACHINE LEARNING FOR TRAFFIC PREDICTION

5.1 Overview

In this chapter, a traffic prediction model based on an innovative approach integrating machine learning with graph theory is proposed to forecast traffic patterns in the near future.

The evaluation results show that the proposed prediction model has a highly desirable performance on root mean squared error (RMSE) and mean absolute percentage error (MAPE). The MAPE is between 0.38% and 1.74% for different jurisdictions. On average, the model outperforms the traditional long short-term memory model by 31.20% in terms of RMSE.

The model performance reassures that incorporating human knowledge helps to improve model performance. The developed prediction model could be used by responsive agencies such as state DOTs to prepare for near-future traffic demand patterns.

5.2 Methodology

Recently, graphical neural networks (GNN) have been used in various traffic forecasting studies, such as traffic flow (Tang & Zeng, 2021) and speed (Zhao et al., 2020) prediction, vehicular trajectory estimation (Li et al., 2021), and travel demand forecasting (Xiong et al., 2020). As traffic networks are naturally graphs, GNN-based models are able to capture spatial dependency of traffic data, and thus outperform previous forecasting models such as autoregressive integrated moving average model, support vector regression, and recurrent neural network (RNN) based models, such as long short-term memory (LSTM) (Jiang & Luo, 2021; J. Yuan et al., 2019; Zhao et al., 2020). In other words, the graphs used in almost all existing GNN-based traffic forecasting studies focus on obtaining spatial information. A typical traffic graph is defined as $G_t = (V, E, A)$, where V is the set of nodes such as roadway segment/traffic detectors for microscopic models or specific geographical areas for macroscopic models; E is the set of edges between nodes, which show the spatial connectivity; A is the adjacency matrix representing the "edge weight" such as distances (Jiang & Luo, 2021).

In this study, although the problem could be formulated into a time series forecasting, we are specifically interested in modeling the impacts of external factors on vehicular traffic. A preprint paper reveals that adding human knowledge as a form of "knowledge graph" to the existing GNN-based traffic forecasting model could improve the model performance (Zhu et al., 2020). Therefore, we adopted this idea and developed a knowledge graph depicting the relationships between the factors mentioned in the previous section. The directed knowledge graph $G_k = (V_k, E_k, A_k)$ is shown in Figure 5.1, where node set V_g consists of VMT and impact factors; edge set E_g represents the "possible" impact relations (the edge e_{ij} exists if the node i has possible impact on node j); the adjacency matrix A_g is a binary matrix showing the existence of edges only. The complex knowledge graph clearly demonstrates that these factors are highly intercorrelated, which might indicate that simple regression models may fail due to collinearity.

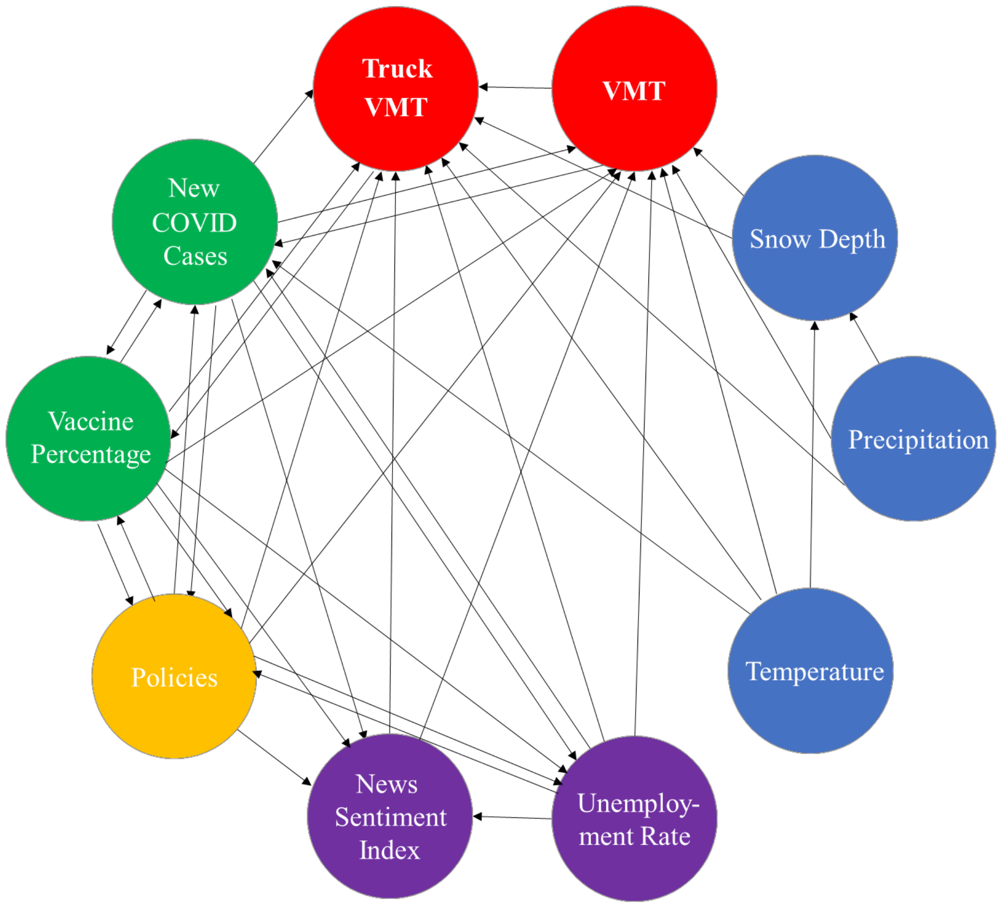


Figure 5.1 Knowledge Graph Depiction

Therefore, the forecasting problem is formulated as learning the mapping function f on the premise of knowledge graph G_k and the factor matrix X and calculate X_T in next T timestamps. In this study, a one-step forecast (one week ahead) is considered as longer-term forecasting and may not be valid due to the rapid change in pandemic and policy status:

$$X_{t+1} = f\big(G_k; (X_{t-n}, \dots, X_{t-1}, X_t)\big) \tag{5.1}$$

where X_{t+1} is the values of all factors at the timestamp $t+1$ although we are only interested in the VMT, and n is the length of historical time series, which is a tunable factor.

The model used to learn the mapping is a graph convolutional networks-long short-term memory (GCN-LSTM). It is a variant of the model proposed by Zhao et al. (2020). The model consists of two parts (Figure 5.2): the graph convolution network (GCN) (Defferrard et al., 2016), a popular GNN model used to obtain the relationships between factors from the knowledge graph, and the LSTM (Hochreiter et al., 1997) used to obtain the temporal dependency.

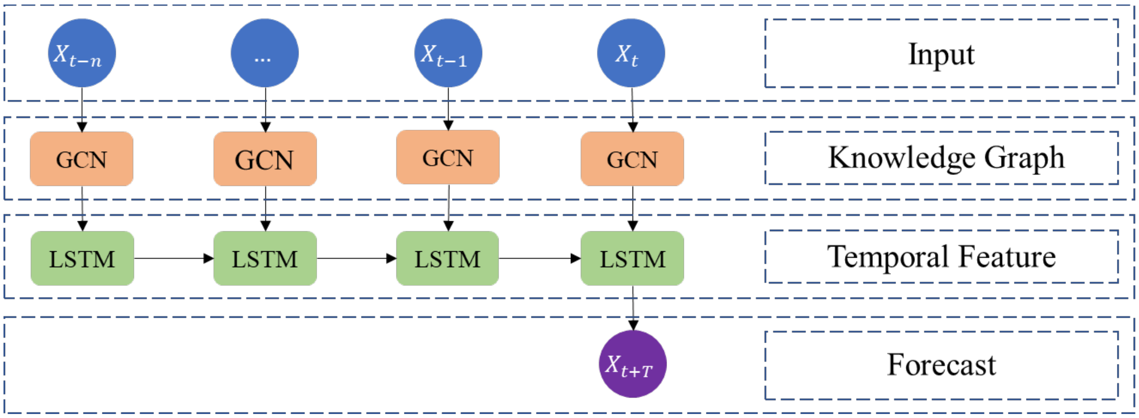


Figure 5.2 GCN-LSTM Model Structure

The basic concept behind the GCN model is using a filter to capture the feathers between a node and its first-order neighborhood. The GCN can then be built by stacking multiple convolutional neural network layers:

$$H^{(l+1)} = \sigma\left(\widetilde{D}^{-\frac{1}{2}}\tilde{A}\widetilde{D}^{-\frac{1}{2}}H^{(l)}\theta^{(l)}\right) \tag{5.2}$$

where $\tilde{A} = A_k + I_n$ is the adjacency matrix, including self-connections of the nodes; I_n is an n-degree identity matrix representing self-connections; $\widetilde{D} = \sum_j \tilde{A}_{ij}$is the degree matrix of the graph representing the neighborhood information; $H^{(l)}$ is the output matrix the layer l and the $\theta^{(l)}$ is the associated trainable parameters; $\sigma(\cdot)$ is the sigmoid function.

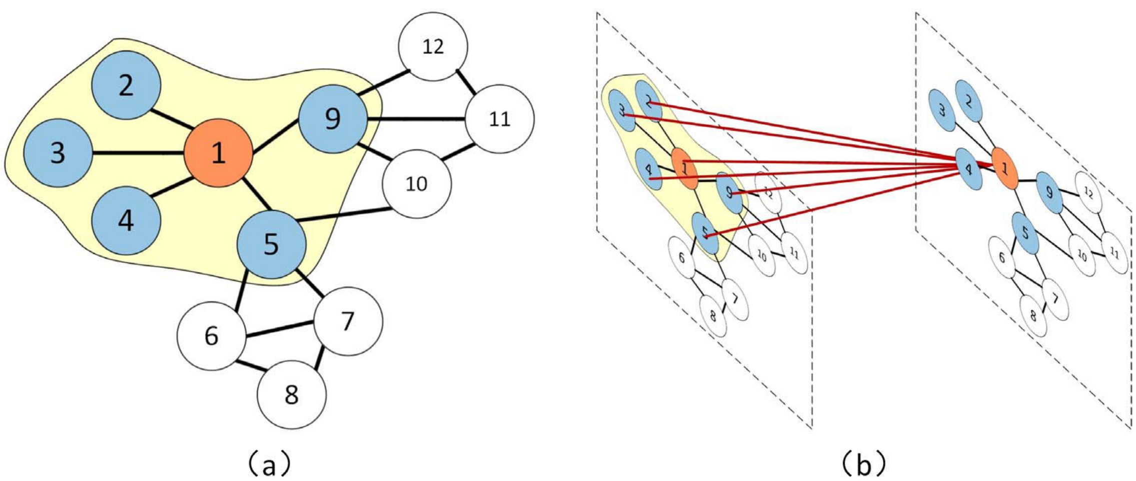


Figure 5.3 Illustration of a GCN Filter (Adopted from Zhao et al., 2020)

Incorporating GCN with LSTM, we get:

$$h_t = o_t ReLU(c_t) \tag{5.3}$$

$$o_t = \sigma(W_o[f(A_k, X_t), h_{t-1}] + V_o c_t) \tag{5.4}$$

$$c_t = f_t c_{t-1} + i_t ReLU(W_c[f(A_k, X_t), h_{t-1}]) \tag{5.5}$$

$$f_t = \sigma\left(W_f[f(A_k, X_t), h_{t-1}] + V_f c_{t-1}\right) \tag{5.6}$$

$$i_t = \sigma(W_i[f(A_k, X_t), h_{t-1}] + V_i c_{t-1}) \tag{5.7}$$

where h_t is the output of LSTM unit at the timestamp t, while the forecast is the output of the final LSTM layer (as shown in Figure 5.3); o_t is the "output gate" that modulates the amount of memory content exposure; Ws are trainable matrixes; $f(A_k, X_t)$ is the final output of the stacked GCN layers; Vs are diagonal matrixes; c_t is the "memory" maintained by the unit at t and is updated by partially forgetting the existing memory by factor gate f_t and adding a new memory content through input gate i_t. Note that normally the activation function used in LSTM is the hyperbolic tangent function (tanh). However, the tanh activation does not perform well in this forecasting problem according to extensive algorithm trainings done by the research team. Since the rectified linear unit (ReLU) activation functions could be used in RNNs with right initialization of the weights (Le et al., 2015), the ReLU activations are adopted.

The loss function used in training is mean square errors between the predicted factors and the observed ones. Adam optimizer was selected to minimize the loss.

Forecasting models are developed for two counties separately since their demographics are very different. For example, as stated in the last section, residents of two counties have different political views and, in turn, impact their risk perceptions on the virus and traveling during the pandemic. Such unobserved heterogeneity could not be modeled using the existing data. Therefore, in total, four models (2 counties × 2 VMT types) are developed.

The model is evaluated by two benchmark models, the persistence model and a fine-tuned LSTM. The persistence model is widely used as the benchmark for time series forecasting problems. A persistence model assumes that the future value of a time series is calculated under the assumption that nothing changes between the current time and the forecast time. Note that although GCN-LSTM is able to forecast all input factors at the future timestamps, only VMT is used to quantify the model performance. Two evaluation metrics are employed, which are root mean squared error (RMSE) and mean absolute percentage error (MAPE) as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(VMT_t - \widehat{VMT_t}\right)^2} \tag{5.8}$$

$$MAPE = \frac{100}{n}\sqrt{\sum_{t=1}^{N}\left(\frac{VMT_t - \widehat{VMT_t}}{\widehat{VMT_t}}\right)} \tag{5.9}$$

5.3 Results and Discussion

Table 5.1 shows the specifications and hypermeters used in the GCN-LSTM model and the benchmarking LSTM. As stated earlier, due to the rapid change of pandemic and policy status, only the data from the latter stage of the pandemic, i.e., after the rollout of vaccines, was selected to develop the forecasting model. Twenty weeks of data are used for training, and the last four weeks of data are used for testing. During the training, an early stopping technique was employed to prevent overfitting. For GCN-LSTM, the data from the past two weeks were used to construct the direct input according to the model tuning. The tuning process shows that adding previous timestamp data into the knowledge graph improves the model performance, possibly because some independent variables may have delayed impacts on the others. Therefore, $X_{t+1,g}$ is calculated as follows:

$$X_{t+1,g} = f_g\left(G_k; (X_{t-1}, X_t)\right) \tag{5.10}$$

However, for LSTM, the tuning process indicates that adding previous timestamp data may distort the memory since it does not have a structure that allows the interactions between independent variables. Thus, the $VMT_{t+1,l}$ is as follows:

$$VMT_{t+1,l} = f_l(X_{t-1}) \tag{5.11}$$

Table 5.1 Model Configuration

Model	GCN-LSTM	LSTM
Layer Configuration	GCN: (16,10) LSTM: (20, 20, 40, 40, 40, 40, 20, 20)	N/A 20
Learning Rate	0.001	0.01
Optimizer	Adam	Adam
Past Data Used	Two weeks	One week

Table 5.2 Model Performance

County	Vehicle Type	Persistence RMSE (10^6 Miles)	Persistence MAPE	LSTM RMSE (10^6 Miles)	LSTM MAPE	GCN-LSTM RMSE (10^6 Miles)	GCN-LSTM MAPE
Salt Lake	Total	2.3979	2.98%	0.9440 (-60.63%)	1.25% (-58.08%)	**0.6374** (-73.42%)	**0.72%** (-75.90%)
	Truck	0.2480	4.53%	0.1449 (-41.56%)	2.19% (-51.57%)	**0.1006** (-59.45%)	**1.74%** (-61.67%)
Utah	Total	1.8757	3.09%	0.9327 (-50.28%)	1.76% (-42.91%)	**0.5943** (-68.32%)	**1.39%** (-55.01%)
	Truck	0.0689	1.93%	0.0236 (-65.76%)	0.64% (-66.62%)	**0.0176** (-74.46%)	**0.38%** (-80.43%)
Average Performance Improvement		Persistence / LSTM RMSE	Persistence / LSTM MAPE	Persistence / GCN-LSTM RMSE	Persistence / GCN-LSTM MAPE	LSTM / GCN-LSTM RMSE	LSTM / GCN-LSTM MAPE
		-54.56%	-54.80%	-68.91%	-68.25%	**-31.20%**	**-31.48%**

The GCN-LTSM is developed using Python programming language with the support of machine learning packages StellarGraph (Data61, 2018), Keras (Chollet, 2015), and TensorFlow (Abadi et al., 2016).

Table 5.2 shows the performance of GCN-LSTM and other benchmark models for different scenarios (2 counties × 2 vehicle types). The table shows that GCN-LSTM obtains the best forecast performance for all four scenarios in terms of both evaluation metrics. On average, GCN-LSTM reduced RMSE by 31.20% and MAPE by 31.48% compared with traditional LSTM models. Thus, incorporating knowledge regarding the interrelationships between explanatory factors significantly improves the model's prediction ability, which is also confirmed by previous studies (Y. Yuan et al., 2021; Zhu et al., 2020).

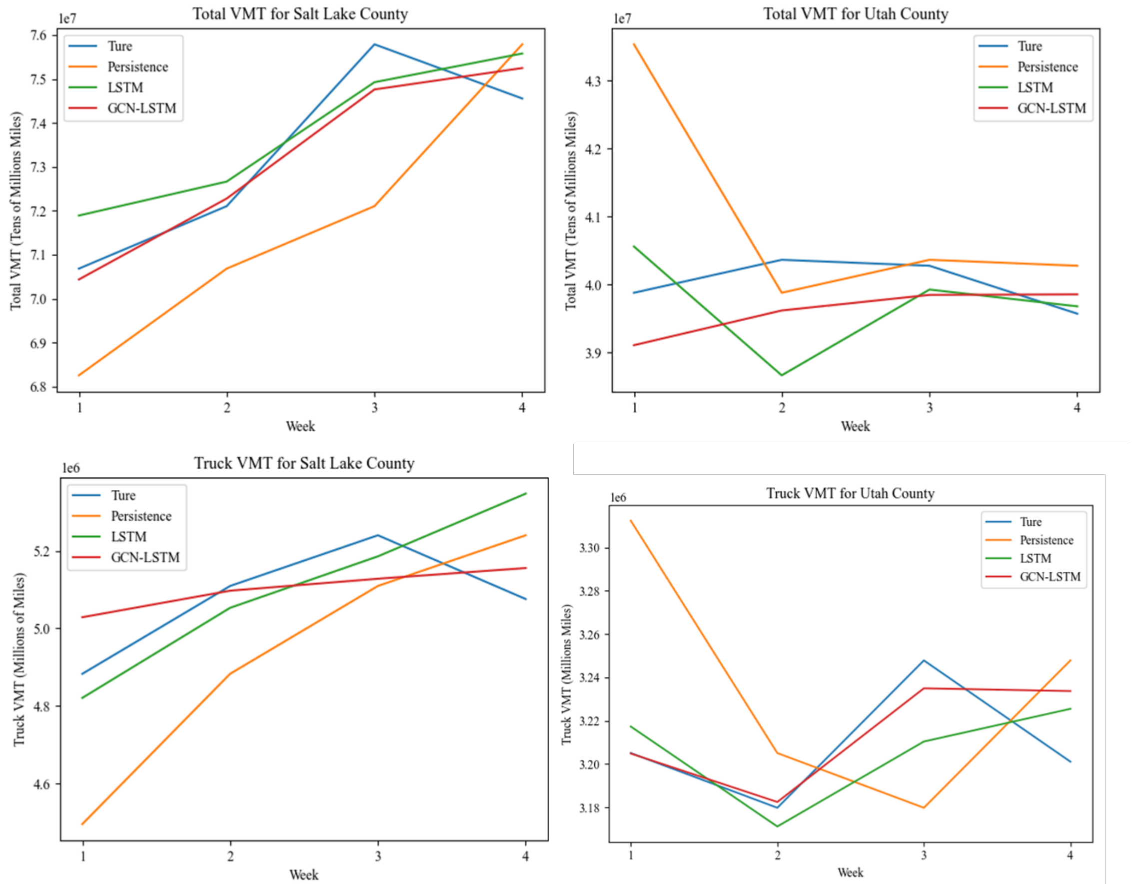


Figure 5.4 Prediction Results of the Models (Up Left: Total VMT for Salt Lake County; Up Right: Total VMT for Utah County; Down Left: Truck VMT for Salt Lake County; Down Right: Truck VMT for Utah County)

To better understand the GCN-LSTM model, the prediction results of the model and the benchmarks on testing data are visualized (Figure 5.4). The results show that:

1. Persistence models fail to forecast future VMTs due to the rapid change of the pandemic status. Take the total VMT of Salt Lake County as an example. The total VMT of the fourth week (74.5 million miles) increased by 5.4% (70.7 million miles) in three weeks. This also implies that long-term forecasting during the pandemic might not be valid.

2. Both LSTM and GCN-LSTM are able to capture the general increasing trend of VMT. GCN-LSTM has smaller prediction errors for almost all prediction points. This re-confirms that human knowledge helps to improve model performance.
3. However, a sudden drop of VMT exists from the third week to the fourth. Both LSTM and GCN-LSTM models have poor capability in predicting this drop. LSTM failed to predict drops for all scenarios while GCN-LSTM only captured the drop when predicting the truck VMT of Utah County. We speculate that the main cause is uncaptured randomness. Prediction models tend to make smoother predictions.

5.4 Conclusion

A prediction model based on innovative GCN-LSTM is developed to forecast traffic patterns in the near future. GCN-LSTM is able to capture the interrelations between the explanatory variables. The evaluation results show that the proposed prediction model has a highly desirable performance. The highest MAPE of the model among all four scenarios (2 counties × 2 vehicle types) is only 1.74% while the lowest is 0.38%. The model outperforms the benchmarking persistence models and LSTM models by -68.91% and -31.20% in terms of RMSE. This reassures that incorporating human knowledge helps to improve model performance. The developed prediction model could be used by responsive agencies such as state DOTs to prepare for the near-future traffic demand pattern.

6. CONCLUSION

6.1 Summary

In response to the COVID-19 pandemic, the U.S. Department of Transportation must quickly adapt to ensure the continuation of critical infrastructure support and relief for the American people. Since March 2020, the pandemic has significantly impacted traffic across the country, with traffic patterns, demands, and durations changing in response to COVID status. Thus, there is a crucial need for research to study the impact of COVID on traffic patterns and analyze the relationship among traffic demand patterns, daily confirmed cases/deaths, state policies, public perceptions, etc. This study focuses on Salt Lake County, Utah, to investigate the impact of COVID-19 on traffic safety in different stages using statistical methods. The study also analyzes the effect of COVID-19 on traffic patterns in Salt Lake County and Utah County from January 2019 to July 2021, identifying different vehicle miles traveled (VMT) patterns during the pre-pandemic stage, early pandemic stage, and late pandemic stage. Finally, the study proposes a knowledge-based traffic prediction model that integrates machine learning with graph theory to forecast traffic patterns in the near future.

6.2 Findings

According to the findings regarding traffic safety, crash frequency throughout the pandemic was significantly lower than that of the pre-pandemic period, even when considering other factors. However, during the latter stage of the pandemic, it increased significantly due to the relaxation of restrictions. In the early stages of the pandemic, crash severity levels increased due to higher speeds, increased DUI incidents, decreased seat belt usage, and more commercial vehicles on the roads. However, crash severity levels later decreased to levels comparable to the pre-pandemic period due to reduced speeds and increased seat belt use. To enhance traffic safety in the incoming “new normal,” stakeholders should take measures to prevent DUIs and reduce commercial vehicle-related crashes.

For traffic patterns, during the early stage of the pandemic, vehicular traffic decreased as a result of government restrictions and individuals’ risk perceptions regarding travel. However, as travel restrictions were relaxed and COVID vaccines became available, vehicular traffic gradually recovered and, in some cases, even exceeded pre-pandemic levels. Furthermore, during the latter stage of the pandemic, truck traffic was higher than pre-pandemic levels due to the rise in online shopping and on-demand delivery.

6.3 Limitation and Challenges

There are limitations in this project. First, the study was only conducted for freeways due to limited data availability. Subsequent studies may assess arterials' safety performance during the pandemic's different stages. Second, due to the need for detailed local survey data, human factors were not included in the statistical modeling of crash frequency. Further studies on the relationship between human factors and crash frequency during the latter pandemic stage are desirable. Third, the pandemic may have complicated impacts on traffic beyond the VMT. For example, during the pandemic, traffic patterns changed from the typical two-peak pattern (morning peak followed by a drop and then afternoon peak) to a gradually increasing to a single afternoon peak in some metropolitan areas (Loo & Huang, 2022; Skip Descant, 2020). The authors attempted to model the change in traffic patterns by introducing speed-related factors, but the resultant models suffered from multicollinearity issues. A good future direction could be conducting a real-time safety analysis (J. Yuan et al., 2019), which focuses on the occurrence of each crash. It can model the impact of real-time traffic and environmental factors closely preceding the crash. Fourth, the statistical models used in this study can be improved, e.g., by using random parameter models to consider the unobserved heterogeneity.

7. REFERENCES

AASHTO. (2010). *Highway Safety Manual | FHWA*. https://highways.dot.gov/safety/data-analysis-tools/highway-safety-manual

AASHTO Journal. (2020, October 9). *Study: COVID-19 Changing Commuter Traffic Patterns – AASHTO Journal*. AASHTO Journal. https://aashtojournal.org/2020/10/09/study-covid-19-changing-commuter-traffic-patterns/

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I. J., Harp, A., Irving, G., Isard, M., Jia, Y., Józefowicz, R., Kaiser, L., Kudlur, M., … Zheng, X. (2016). "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems." *CoRR*, *abs/1603.0*.

Abdel-Aty, M., Lee, J., Siddiqui, C., & Choi, K. (2013). "Geographical unit based analysis in the context of transportation safety planning." *Transportation Research Part A: Policy and Practice*, *49*, 62–75. https://doi.org/10.1016/J.TRA.2013.01.030

Chang, L. Y., & Wang, H. W. (2006). "Analysis of traffic injury severity: An application of non-parametric classification tree techniques." *Accident Analysis & Prevention*, *38*(5), 1019–1027. https://doi.org/10.1016/J.AAP.2006.04.009

Chen, T., Sze, N. N., Chen, S., Labi, S., & Zeng, Q. (2021). "Analysing the main and interaction effects of commercial vehicle mix and roadway attributes on crash rates using a Bayesian random-parameter Tobit model." *Accident Analysis & Prevention*, *154*, 106089. https://doi.org/10.1016/J.AAP.2021.106089

Chollet, F. (2015). Keras. In *GitHub repository*. GitHub.

Christidis, P., Christodoulou, A., Navajas-Cawood, E., & Ciuffo, B. (2021a). "The Post-Pandemic Recovery of Transport Activity: Emerging Mobility Patterns and Repercussions on Future Evolution." *Sustainability 2021,* Vol. 13, Page 6,359, *13*(11), 6359. https://doi.org/10.3390/SU13116359

Christidis, P., Christodoulou, A., Navajas-Cawood, E., & Ciuffo, B. (2021b). "The Post-Pandemic Recovery of Transport Activity: Emerging Mobility Patterns and Repercussions on Future Evolution." *Sustainability 2021,* Vol. 13, Page 6,359, *13*(11), 6359. https://doi.org/10.3390/SU13116359

Conroy, C., Hoyt, D. B., Eastman, A. B., Erwin, S., Pacyna, S., Holbrook, T. L., Vaughan, T., Sise, M., Kennedy, F., & Velky, T. (2006). "Rollover crashes: Predicting serious injury based on occupant, vehicle, and crash characteristics." *Accident Analysis & Prevention*, *38*(5), 835–842. https://doi.org/10.1016/J.AAP.2006.02.002

Coronavirus | Keeping Utah Informed on the Latest Coronavirus Updates. (n.d.). Retrieved March 22, 2023, from https://coronavirus.utah.gov/

Data61, C. (2018). StellarGraph Machine Learning Library. In *GitHub Repository*. GitHub.

Defferrard, M., Bresson, X., & Vandergheynst, P. (2016). "Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering." *Advances in Neural Information Processing Systems*, 3,844–3,852.

Dong, N., Zhang, J., Liu, X., Xu, P., Wu, Y., & Wu, H. (2022). "Association of human mobility with road crashes for pandemic-ready safer mobility: A New York City case study." *Accident Analysis & Prevention*, *165*, 106478. https://doi.org/10.1016/J.AAP.2021.106478

Dong, X., Xie, K., & Yang, H. (2022). "How did COVID-19 impact driving behaviors and crash Severity? A multigroup structural equation modeling." *Accident Analysis & Prevention*, *172*, 106687. https://doi.org/10.1016/J.AAP.2022.106687

Doucette, M. L., Tucker, A., Auguste, M. E., Watkins, A., Green, C., Pereira, F. E., Borrup, K. T., Shapiro, D., & Lapidus, G. (2021). "Initial impact of COVID-19's stay-at-home order on motor vehicle traffic and crash patterns in Connecticut: an interrupted time series analysis." *Injury Prevention*, *27*(1), 3–9. https://doi.org/10.1136/INJURYPREV-2020-043945

Duan, T., Jiang, H., Deng, X., Zhang, Q., & Wang, F. (2020). "Government Intervention, Risk Perception, and the Adoption of Protective Action Recommendations: Evidence from the COVID-19 Prevention and Control Experience of China." *International Journal of Environmental Research and Public Health 2020,* Vol. 17, Page 3,387, *17*(10), 3387. https://doi.org/10.3390/IJERPH17103387

Ebrahim Shaik, M., & Ahmed, S. (2022). "An overview of the impact of COVID-19 on road traffic safety and travel behavior." *Transportation Engineering*, *9*, 100119. https://doi.org/10.1016/J.TRENG.2022.100119

Find a Station | Data Tools | Climate Data Online (CDO) | National Climatic Data Center (NCDC). (n.d.). Retrieved March 22, 2023, from https://www.ncdc.noaa.gov/cdo-web/datatools/findstation

Glaeser, E. L., Gorback, C., & Redding, S. J. (2020). *How Much does COVID-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities*. https://doi.org/10.3386/W27519

Gong, Y., Isom, T., Lu, P., Yang, X., & Wang, A. (2022). "Modeling the impact of COVID-19 on transportation at later stage of the pandemic: A case study of Utah." *Journal of Intelligent Transportation Systems*. https://doi.org/10.1080/15472450.2022.2157212

Hochreiter, S., Schmidhuber, J., Gambardella, L. M., & Schmidhuber, J. (1997). "Long Short-Term Memory." *Neural Computation*, *9*(8), 1,735–1,780. https://doi.org/10.1162/neco.1997.9.8.1735

Holm, S. (1979). "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics*. https://doi.org/10.2307/4615733

Islam, M. R., Abdel-Aty, M., Islam, Z., & Zhang, S. (2022). "Risk-Compensation Trends in Road Safety during COVID-19." *Sustainability 2022,* Vol. 14, Page 5,057, *14*(9), 5057. https://doi.org/10.3390/SU14095057

Jiang, W., & Luo, J. (2021). *Graph Neural Network for Traffic Forecasting: A Survey*.

Katrakazas, C., Michelaraki, E., Sekadakis, M., & Yannis, G. (2020). “A descriptive analysis of the effect of the COVID-19 pandemic on driving behavior and road safety.” *Transportation Research Interdisciplinary Perspectives*, *7*, 100186. https://doi.org/10.1016/J.TRIP.2020.100186

Kim, K. (2021). “Impacts of COVID-19 on transportation: Summary and synthesis of interdisciplinary research.” *Transportation Research Interdisciplinary Perspectives*, *9*, 100305. https://doi.org/10.1016/J.TRIP.2021.100305

Koloushani, M., Ghorbanzadeh, M., Ozguven, E. E., & Ulak, M. B. (2021). “Crash Patterns in the COVID-19 Pandemic: The Tale of Four Florida Counties.” *Future Transportation 2021, Vol. 1, Pages 414-442*, *1*(3), 414–442. https://doi.org/10.3390/FUTURETRANSP1030023

Le, Q. V., Jaitly, N., & Hinton, G. E. (2015). *A Simple Way to Initialize Recurrent Networks of Rectified Linear Units*.

Li, Z., Lu, C., Yi, Y., & Gong, J. (2021). “A Hierarchical Framework for Interactive Behaviour Prediction of Heterogeneous Traffic Participants Based on Graph Neural Network.” *IEEE Transactions on Intelligent Transportation Systems*, 1–13. https://doi.org/10.1109/TITS.2021.3090851

Loo, B. P. Y., & Huang, Z. (2022). “Spatio-temporal variations of traffic congestion under work from home (WFH) arrangements: Lessons learned from COVID-19.” *Cities*, *124*, 103610. https://doi.org/10.1016/J.CITIES.2022.103610

Lord, D., & Mannering, F. (2010). “The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives.” *Transportation Research Part A: Policy and Practice*, *44*(5), 291–305. https://doi.org/10.1016/J.TRA.2010.02.001

Mahmoudi, J., & Xiong, C. (2022). “How social distancing, mobility, and preventive policies affect COVID-19 outcomes: Big data-driven evidence from the District of Columbia-Maryland-Virginia (DMV) megaregion.” *PLoS ONE*, *17*(2 February), 1–21. https://doi.org/10.1371/journal.pone.0263820

Manning, A. R., Romano, E., Diebold, J., Nochajski, T. H., Taylor, E., Voas, R. B., & Scherer, M. (2021). “Convicted drinking and driving offenders: Comparing alcohol use before and after the pandemic outbreak.” *Alcoholism: Clinical and Experimental Research*, *45*(6), 1,225–1,236. https://doi.org/10.1111/ACER.14613

Maze, T. H., Agarwal, M., & Burchett, G. (2006). “Whether Weather Matters to Traffic Demand, Traffic Safety, and Traffic Operations and Flow.” *Https://Doi.Org/10.1177/0361198106194800119*, *1948*(1), 170–176. https://doi.org/10.1177/0361198106194800119

National Centres for Environmental Information (NCEI). (2020). *Find a Station | Data Tools | Climate Data Online (CDO) | National Climatic Data Center (NCDC)*.

NHTSA. (2022). *Drunk Driving | Statistics and Resources | NHTSA*. https://www.nhtsa.gov/risky-driving/drunk-driving

Papadimitriou, E., Filtness, A., Theofilatos, A., Ziakopoulos, A., Quigley, C., & Yannis, G. (2019). "Review and ranking of crash risk factors related to the road infrastructure." *Accident Analysis & Prevention*, *125*, 85–97. https://doi.org/10.1016/J.AAP.2019.01.002

Pathak, A. A., Chandrasekaran, S., & Annamalai, B. (2022). "Analysis of Motor Vehicle Accidents: Comparison Between Before and During the COVID-19 Lockdown in Maharashtra, India." *Https://Doi.Org/10.1177/03611981221089936*, 036119812210899. https://doi.org/10.1177/03611981221089936

Qiu, L., & Nixon, W. A. (2008). "Effects of Adverse Weather on Traffic Crashes." *Https://Doi.Org/10.3141/2055-16*, *2055*, 139–146. https://doi.org/10.3141/2055-16

Qureshi, A. I., Huang, W., Khan, S., Lobanova, I., Siddiq, F., Gomez, C. R., & Suri, M. F. K. (2020). "Mandated societal lockdown and road traffic accidents." *Accident Analysis & Prevention*, *146*, 105747. https://doi.org/10.1016/J.AAP.2020.105747

Research, U. S. D. of T. N. H. T. S. A. O. of B. S. (2021a). *Update to Special Reports on Traffic Safety during the COVID-19 Public Health Emergency: Third Quarter Data: Third Quarter Data [Traffic Safety Facts]*. https://doi.org/10.21949/1525990

Research, U. S. D. of T. N. H. T. S. A. O. of B. S. (2021b). *Update to Special Reports on Traffic Safety during the COVID-19 Public Health Emergency: Fourth Quarter Data [Traffic Safety Facts]*. https://doi.org/10.21949/1526015

Rothengatter, W., Zhang, J., Hayashi, Y., Nosach, A., Wang, K., & Oum, T. H. (2021). "Pandemic waves and the time after Covid-19 – Consequences for the transport sector." *Transport Policy*, *110*, 225–237. https://doi.org/10.1016/J.TRANPOL.2021.06.003

Sekadakis, M., Katrakazas, C., Michelaraki, Eva, & Yannis, G. (2022). "Driving Behavior and Its Correlation with COVID-19 Response Measures: A Neural Network Forecasting Analysis." *Journal of Transportation Engineering, Part A: Systems*, *148*(10), 04022083. https://doi.org/10.1061/JTEPBS.0000735

Sekadakis, M., Katrakazas, C., Michelaraki, E., Kehagia, F., & Yannis, G. (2021). "Analysis of the impact of COVID-19 on collisions, fatalities and injuries using time series forecasting: The case of Greece." *Accident Analysis & Prevention*, *162*, 106391. https://doi.org/10.1016/J.AAP.2021.106391

Shamshiripour, A., Rahimi, E., Shabanpour, R., & Mohammadian, A. (Kouros). (2020). "How is COVID-19 reshaping activity-travel behavior? Evidence from a comprehensive survey in Chicago." *Transportation Research Interdisciplinary Perspectives*, *7*, 100216. https://doi.org/10.1016/J.TRIP.2020.100216

Shapiro, A. H., Sudhof, M., & Wilson, D. (2017). "Measuring News Sentiment." *Federal Reserve Bank of San Francisco, Working Paper Series*, 01–22. https://doi.org/10.24148/WP2017-01

Sheather, S. (2009). *A Modern Approach to Regression with R*. https://doi.org/10.1007/978-0-387-09608-7

Skip Descant. (2020, October 16). *Has COVID-19 Forever Changed Rush-Hour Traffic Patterns?* Government Technology. https://www.govtech.com/analytics/has-covid-19-forever-changed-rush-hour-traffic-patterns.html

Sze, N. N., & Wong, S. C. (2007). “Diagnostic analysis of the logistic model for pedestrian injury severity in traffic crashes.” *Accident Analysis & Prevention*, *39*(6), 1267–1278. https://doi.org/10.1016/J.AAP.2007.03.017

Tang, J., & Zeng, J. (2021). “Spatiotemporal gated graph attention network for urban traffic flow prediction based on license plate recognition data.” *Computer-Aided Civil and Infrastructure Engineering*. https://doi.org/10.1111/MICE.12688

The State of Utah. (2021a). *Case Counts: coronavirus*. Coronavirus.Utah.Gov.

The State of Utah. (2021b). *Special Orders | coronavirus*.

USDOT. (2021). *Continuation of Research on Traffic Safety during the COVID-19 Public Health Emergency: January – June 2021 [Traffic Safety Facts]*. https://doi.org/10.21949/1526036

Utah Department of Transportation. (2019). *PeMS @ UDOT*.

Utah Department of Workforce Services. (2021). *Seasonally Adjusted Unemployment Rates by County*.

Vanlaar, W. G. M., Woods-Fry, H., Barrett, H., Lyon, C., Brown, S., Wicklund, C., & Robertson, R. D. (2021). “The impact of COVID-19 on road safety in Canada and the United States.” *Accident Analysis & Prevention*, *160*, 106324. https://doi.org/10.1016/J.AAP.2021.106324

Wagner, E. K., Atkins, R., Berning, A., Robbins, A., Watson, C. E., & Anderle, J. (2020). *Examination of the Traffic Safety Environment During the Second Quarter of 2020: Special Report*.

Watson-Brown, N., Truelove, V., Parker, E., & Davey, J. (2021). “Drink driving during the COVID-19 pandemic.” *Transportation Research Part F: Traffic Psychology and Behaviour*, *78*, 369–380. https://doi.org/10.1016/J.TRF.2021.02.020

Welch, B. L. (1947). “The Generalization of ‘Student’s’ Problem when Several Different Population Variances are Involved.” *Biometrika*, *34*(1–2), 28–35. https://doi.org/10.1093/BIOMET/34.1-2.28

Xiong, X., Ozbay, K., Jin, L., & Feng, C. (2020). “Dynamic Origin–Destination Matrix Prediction with Line Graph Neural Networks and Kalman Filter.” *Https://Doi.Org/10.1177/0361198120919399*, *2674*(8), 491–503. https://doi.org/10.1177/0361198120919399

Yuan, J., Abdel-Aty, M., Gong, Y., & Cai, Q. (2019). “Real-Time Crash Risk Prediction using Long Short-Term Memory Recurrent Neural Network.” *Transportation Research Record*, 0361198119840611.

Yuan, Y., Zhang, Z., Yang, X. T., & Zhe, S. (2021). “Macroscopic traffic flow modeling with physics regularized Gaussian process: A new insight into machine learning applications in transportation.” *Transportation Research Part B: Methodological*, *146*, 88–110. https://doi.org/10.1016/J.TRB.2021.02.007

Yue, L., Abdel-Aty, M., Lee, J., & Farid, A. (2019). "Effects of Signalization at Rural Intersections Considering the Elderly Driving Population." *Transportation Research Record*, *2673*(2), 743–757. https://doi.org/10.1177/0361198119825834/ASSET/IMAGES/LARGE/10.1177_0361198119825834-FIG3.JPEG

Zhang, Jie, Feng, B., Wu, Y., Xu, P., Ke, R., & Dong, N. (2021). "The effect of human mobility and control measures on traffic safety during COVID-19 pandemic." *PLOS ONE*, *16*(3), e0243263. https://doi.org/10.1371/JOURNAL.PONE.0243263

Zhang, Jinbao, & Lee, J. (2021). "Interactive effects between travel behaviour and COVID-19: a questionnaire study." *Transportation Safety and Environment*, *3*(2), 166–177. https://doi.org/10.1093/tse/tdab003

Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., & Li, H. (2020). "T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction." *IEEE Transactions on Intelligent Transportation Systems*, *21*(9), 3848–3858. https://doi.org/10.1109/TITS.2019.2935152

Zhou, H., Yuan, C., Dong, N., Wong, S. C., & Xu, P. (2020). "Severity of passenger injuries on public buses: A comparative analysis of collision injuries and non-collision injuries." *Journal of Safety Research*, *74*, 55–69. https://doi.org/10.1016/J.JSR.2020.04.003

Zhou, T., & Zhang, J. (2019). "Analysis of commercial truck drivers' potentially dangerous driving behaviors based on 11-month digital tachograph data and multilevel modeling approach." *Accident Analysis & Prevention*, *132*, 105256. https://doi.org/10.1016/J.AAP.2019.105256

Zhu, J., Han, X., Deng, H., Tao, C., Zhao, L., Tao, L., & Li, H. (2020). *KST-GCN: A Knowledge-Driven Spatial-Temporal Graph Convolutional Network for Traffic Forecasting*.

8. APPENDIX

Table A Descriptive Statistics of Variables Used in Crash Frequency Analysis

Variable	Pandemic		Statistics			
			Mean	S.D.	Minimum	Maximum
Daily Number of Crashes	Before		16.6934	11.2159	0	110
	During	Total	12.5071	8.3181	1	77
		Earlier	10.4506	7.1871	1	77
		Later	14.6224	8.8637	1	76
Non-Covid-Related Independent Variables						
Ln (VMT) (The unit of VMT is mile)	Before		16.0861	0.1957	15.4264	16.3018
	During	Total	15.9944	0.1983	15.1075	16.2719
		Earlier	15.9316	0.2046	15.1075	16.1946
		Later	16.0591	0.1690	15.4753	16.2719
Average Temperature (°F)	Before		50.6133	18.6965	16	90
	During	Total	55.1887	19.2418	16	91
		Earlier	54.8456	19.1433	23	90
		Later	55.5417	19.3612	16	91
Total Precipitation (inch)	Before		0.0536	0.1313	0	0.8500
	During	Total	0.0311	0.1044	0	0.9100
		Earlier	0.0252	0.0852	0	0.8500
		Later	0.0372	0.1208	0	0.9100
Covid-Related Independent Variables						
Numerical Variable	Covid Stage		Descriptive Statistics			
			Mean	S.D.	Minimum	Maximum
Number of New Covid Cases	Earlier		368.1722	349.3825	0	1646
	Later		517.8047	903.7283	24	5810
Death Rate (%)	Earlier		0.1804	0.6748	0	6.3830
	Later		0.1014	0.3524	0	2.3474
Categorial Variables	Covid Stage		Yes (1)		No (0)	
			Count	%	Count	%
Lockdown	Earlier		52	13.1646	343	86.8354
	Later		0	0	384	100%
Mask Mandate	Earlier		287	72.6582	108	27.3418
	Later		42	10.9375	342	89.0625

Table B Descriptive Statistics of Variables Used in Crash Severity Analysis

Variable		Pandemic		Yes (1)		No (0)	
				Count	%	Count	%
Whether a Crash Leads to Injuries		Before		521	7.22%	6700	92.78%
		During	Total	804	8.44%	8723	91.56%
			Earlier	387	9.49%	3691	90.51%
			Later	417	7.65%	5032	92.35%
Categorical Independent Variables							
Manner of Collision	Angle	Before		352	4.87%		
		During	Total	468	4.91%		
			Earlier	231	5.66%		
			Later	237	4.35%		
	Front to Rear	Before		3312	45.87%		
		During	Total	3565	37.42%		
			Earlier	1298	31.83%		
			Later	2267	41.60%		
	Head On (front-to-front)	Before		46	0.64%		
		During	Total	36	0.38%		
			Earlier	19	0.47%		
			Later	17	0.31%		
	Single Vehicle	Before		2016	27.92%		
		During	Total	3209	33.68%		
			Earlier	1548	37.96%		
			Later	1661	30.48%		
	Parked Vehicle	Before		22	0.30%		
		During	Total	33	0.35%		
			Earlier	16	0.39%		
			Later	17	0.31%		
	Rear to Rear	Before		4	0.06%		
		During	Total	3	0.03%		
			Earlier	3	0.07%		
			Later	0	0.00%		
	Rear to Side	Before		6	0.08%		
		During	Total	4	0.04%		
			Earlier	2	0.05%		
			Later	2	0.04%		
	Sideswipe Opposite Direction	Before		17	0.24%		
		During	Total	14	0.15%		
			Earlier	8	0.20%		
			Later	6	0.11%		
	Sideswipe Same Direction	Before		1446	20.02%		
		During	Total	2195	23.04%		
			Earlier	953	23.37%		
			Later	1242	22.79%		
Daylight Condition		Before		5238	72.54%	1983	27.46%
		During	Total	6765	71.01%	2762	28.99%
			Earlier	2823	69.23%	1255	30.77%
			Later	3942	72.34%	1507	27.66%

Adverse Roadway Surface Condition	Before		1970	27.28%	5251	72.72%
	During	Total	1881	19.74%	7646	80.26%
		Earlier	778	19.08%	3300	80.92%
		Later	1103	20.24%	4346	79.76%
Adverse Weather	Before		1559	21.59%	5662	78.41%
	During	Total	1447	15.19%	8080	84.81%
		Earlier	622	15.25%	3456	84.75%
		Later	825	15.14%	4624	84.86%
Commercial Vehicle Involved	Before		681	9.43%	6540	90.57%
	During	Total	1147	12.04%	8380	87.96%
		Earlier	453	11.11%	3625	88.89%
		Later	694	12.74%	4755	87.26%
Distracted Driving Involved	Before		474	6.56%	6747	93.44%
	During	Total	545	5.72%	8982	94.28%
		Earlier	239	5.86%	3839	94.14%
		Later	306	5.62%	5143	94.38%
Drowsy Driving Involved	Before		117	1.62%	7104	98.38%
	During	Total	198	2.08%	9329	97.92%
		Earlier	92	2.26%	3986	97.74%
		Later	106	1.95%	5343	98.05%
DUI Involved	Before		204	2.83%	7017	97.17%
	During	Total	443	4.65%	9084	95.35%
		Earlier	219	5.37%	3859	94.63%
		Later	224	4.11%	5225	95.89%
Motorcycle Involved	Before		46	0.64%	7175	99.36%
	During	Total	86	0.90%	9441	99.10%
		Earlier	39	0.96%	4039	99.04%
		Later	47	0.86%	5402	99.14%
Older Driver Involved	Before		691	9.57%	6530	90.43%
	During	Total	890	9.34%	8637	90.66%
		Earlier	348	8.53%	3730	91.47%
		Later	542	9.95%	4907	90.05%
Overturn/Rollover Involved	Before		211	2.92%	7010	97.08%
	During	Total	330	3.46%	9197	96.54%
		Earlier	170	4.17%	3908	95.83%
		Later	160	2.94%	5289	97.06%
Unrestrained Involved	Before		113	1.56%	7108	98.44%
	During	Total	205	2.15%	9322	97.85%
		Earlier	94	2.31%	3984	97.69%
		Later	111	2.04%	5338	97.96%
Numerical Independent Variables						
Variable	Covid		Descriptive Statistics			
			Mean	S.D.	Minimum	Maximum
Average Speed (mph)	Before		64.3535	3.7825	47.2906	70.1399
	During	Total	65.5001	2.2703	55.4369	68.5575
		Earlier	66.4314	1.8597	55.4369	68.5575
		Later	64.8031	2.2999	55.8201	68.4569