Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Projecting daily travel behavior by distance during the pandemic and the spread of COVID-19 infections – Are we in a closed loop scenario?

Dothang Truong\textsuperscript{a,},*, My D. Truong\textsuperscript{b}

\textsuperscript{a}Embry-Riddle Aeronautical University, United States
\textsuperscript{b}University of Central Florida, United States

\begin{abstract}
Understanding the future development of COVID-19 is the key to contain the spreading of the coronavirus. The purpose of this paper is to explore a potential relationship between United States residents’ daily trips by distance and the COVID-19 infections in the near future. The study used the daily travel data from the Bureau of Transportation Statistics (BTS) and the COVID-19 data from the Centers for Disease Control and Prevention (CDC) in the United States. Time-series forecast models using Autoregressive Moving Average (ARIMA) method were constructed to project future trends of United States residents’ daily trips by distance at the national level from November 30, 2020, to February 28, 2021. A comparative trend analysis was conducted to detect the patterns of daily trips and the spread of COVID-19 during that period. The results revealed a closed loop scenario, in which the residents’ travel behavior dynamically changes based on their risk perception of COVID-19 in an infinite loop. A detected lag in the travel behavior between short trips and long trips further worsens the situation and creates more difficulties in finding an effective solution to break the loop. The study shed new light on efforts to contain and control the spread of the coronavirus. The loop can only be broken with proper and prompt mitigation strategies to reduce the burden on hospitals and healthcare systems and save more lives.
\end{abstract}

1. Introduction

COVID-19 pandemic has been a primary concern for the whole world in 2020. The first case of COVID-19 in the United States was reported on January 21, 2020, and the first COVID-19 related death was reported on March 1, 2020. Since then, the pandemic has spread at an exponential rate, threatening the health of residents all around the world. The United States is the country with the most COVID-19 cases and deaths. As of December 8, 2020, there have been a total of 14,823,129 COVID-19 cases and 282,785 deaths in the United States (CDC, 2020a). Fig. 1 shows daily new COVID-19 cases in the United States from March 1 to November 29, 2020. Due to a spike of new cases in late March and early April 2020, many states had declared state emergencies, increased the test capacity, shut down non-essential businesses, executed stay-home orders, and restricted travels (Mervosh et al., 2020). As the case number started to stabilize and decrease, in early May 2020, many states started the reopening phase to gradually bring businesses and travel back to normal (Washington Post, 2020). We saw more people on the road, in stores, and in restaurants. While major events were still prohibited, grocery stores and restaurants were full of customers, and beaches were swamped with people during the summertime. It was not a surprise that spiking cases and deaths took place in mid-July 2020, with more than 70,000 daily new cases due to the reopening. Then, new daily cases seemed to reduce to about 40,000 cases per day in August 2020. When everyone thought the situation was getting better, we experienced a much higher surge of new cases in November 2020, when more than 180,000 cases were reported per day. That was the time of the Presidential election and Thanksgiving holiday. States with more populations and attractions, such as California, Texas, Florida, and New York, are among the top states with the highest number of cases and deaths (CDC, 2020a).

It appears that travel restrictions and fear of the pandemic have negatively affected residents’ long-distance travel. The airline industry has been affected significantly, with much fewer flights and emptier airports (Jacus et al., 2020; TSA, 2020). However, whether people have decided to stay home or not is a different story. As shown in Fig. 2, before COVID 19, the population staying home, i.e., making

\footnotesize{*} Corresponding author at: Embry-Riddle Aeronautical University, 1 Aerospace Blvd., Daytona Beach, FL 32114, United States.
E-mail address: truongd@erau.edu (D. Truong).

https://doi.org/10.1016/j.trip.2020.100283
Received 16 September 2020; Revised 13 December 2020; Accepted 13 December 2020
Available online 18 December 2020
2590-1982/© 2020 Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
no trips daily, was about 20%. It is considered the baseline for our comparison. Since the pandemic, the percentage of residents staying home increased to almost 30% in early March 2020, during statewide stay-home-orders. In mid-May 2020, when most states decided to reopen, the population staying home reduced to lower than our baseline of 20%. Then, the number slightly increased again in mid-July 2020, when we noticed a spiking number of COVID-19 cases. Finally, given the surge of COVID-19 in late October and mid-November, the population staying home increased to almost 30%. Thus, during the pandemic time, the population not staying home fluctuates from 70% to 82% (BTS, 2020). That raises questions about the daily travel behavior of this population during the pandemic.

Fig. 3 shows the number of trips that U.S. residents made daily during the years 2019–2020. Before the pandemic, this number ranges from 3.5 billion trips to 4.5 billion trips daily, including various distances, from one mile to more than 500 miles. During the pandemic, the trips decreased to the lowest in mid-March 2020, during stay-home orders, which was about 2.5 billion trips per day. Then, this number gradually increased to the highest of about 3 billion trips per day in mid-May 2020, when the stay-home-orders were lifted in most states. In mid-July 2020, when we observed a spiking number of COVID-19 cases, this number started to decrease again. Finally, the number of trips increased in November 2020, when more people travel for the Presidential Election and Thanksgiving.

Those statistics show how U.S. residents’ travel behavior changes as they received state stay-home-orders and observed changes in COVID-19 reported cases and deaths. Research in daily travel in the United States in relation to the COVID-19 infection is still limited. Several studies indicated that mobility could lead to a higher risk of COVID-19 infection, which poses a higher risk to the health of residents (Bruinen de Bruin et al., 2020; Gostic et al., 2020; De Vos, 2020). Accordingly, Centers for Disease Control and Prevention (CDC) recommends extra precautions with travel during the pandemic, especially wearing masks and keeping physical distancing (CDC, 2020b). Without proper control, infected travelers can create a new exponential outbreak (Linka et al., 2020; De Vos, 2020). Using the travel data in China, Kraemer et al. (2020) and Chinazzi et al. (2020) found that travel restrictions issued by the Chinese government have reduced the spread of COVID-19 in that country. Regarding travels in the United States, a study by Karaer et al. (2020) showed that COVID-19 reduced vehicle travels in all Florida counties, which resulted in a significant reduction in traffic. Finally, Yilmazkuday (2020) focused on
inter-county travels and found that strict restrictions on travels across counties in the United States reduced the number of new cases and new deaths.

One thing that is still unknown is whether the travel distances have any effect on the spread of COVID-19 infection and how the fear of the pandemic may affect residents’ daily travel behavior. The pandemic has certainly reduced long-distance travel significantly, especially by air, but we cannot say the same for shorter trips, especially by road. Even during the state stay-home-orders, residents were still allowed to go to work, stores, or take-out restaurants. Thus, their daily travel behavior, especially after the state reopening, may not change much during the pandemic. Beck and Hensher (2020) found that while travel risk perception increased with age, it decreased with travel frequency, i.e., the more trips someone makes, the lower the risk perception of the virus infection. In most cities, if we go outside, we do not see anything unusual in terms of cars on the road or people in stores, except that more people are wearing masks now. This situation raises a serious question on whether this travel behavior contributes to the continued spread of COVID-19 or vice versa. The more trips people make, the more challenging it is to keep physical distancing, which may lead to a faster spread of COVID-19 infections. On the other hand, the more reported cases and deaths could cause safety concerns for some residents, which would lead to fewer daily trips. So, the overarching question is: Are we in a closed loop of COVID-19 infections and daily trips? In other words, as the number of COVID-19 cases fluctuates and is at an all-time high in November 2020, will we see the repeated travel behavior which leads to further infections, or will we see the light at the end of the tunnel to get out of this serious situation?

The purpose of this study is to explore the future trends of residents’ daily trips by distance in the United States and how they are related to the continued spread of COVID-19. Time series projections of future daily trips were conducted to show how residents’ travel behavior changes in the near future and how this behavior may increase the risk of COVID-19 spreading, which also leads to further changes in daily travel. In other words, the study explores possible inter-relationships between these variables in a possible loop. The trip distances are selected based on the Bureau of Transportation Statistics’ recommendation with ten categories: less than one mile, 1–3 miles, 3–5 miles, 5–10 miles, 10 – 25 miles, 25–50 miles, 50–100 miles, 100–250 miles; 250 – 500 miles, and greater than 500 miles.

The study has several delimitations. First, it focuses on the daily travel behavior of residents in the United States at the national level. A follow-on project could focus on travel behavior in specific states. Second, it is important to note that it is not the purpose of this study to examine or test causal relationships between daily trips and COVID-19 infection. Since this is a two-way relationship with a possibility of inter-correlations in a loop, the paper is exploratory in its nature and focuses on forecasting the future values of the daily trips using a time-series method. It assumes that daily travel is temporal dependent. Third, the demographics and characteristics of residents are also not included in the study due to the unavailability of the data. This delimitation should not affect the outcomes of the study, which focuses on the time dependency of the trips. Finally, the effects of various intervention strategies on daily trips and COVID-19 infection are also not examined since it is not in the scope of this study and would be a relatively large project. Nonetheless, discussions on scenarios for daily travel and COVID-19 infection based on different intervention strategies were provided.

The paper is organized as follows. In Section 2, the research methodology is presented with data sources and data analysis methods. Section 3 presents descriptive statistics of daily trips, relationships between daily trips and COVID-19 infection, and forecast of future daily trips in relation to the spread of COVID-19. Finally, discussions of the findings and conclusions are presented in Section 4. Four scenarios based on different intervention strategies are discussed in this section.

2. Data collection and methodology

2.1. Data sources

In this paper, the Daily Travel During the COVID-19 Pandemic data provided by BTS was used. BTS is part of the Department of Transportation (DOT) and is the major source of statistics on commercial aviation, multimodal travel activity, and transportation economics (BTS, 2019). The BTS Daily Travel data in the United States was collected by the Maryland Transportation Institute and Center for Advanced Transportation Technology Laboratory (CATTL) at the University of Maryland. In this data, trips are defined as movements that include a stay of longer than 10 min at an anonymized location away from home; a movement including multiple stays of longer than 10 min before returning home is counted as multiple trips. The daily travel data were collected from a mobile device data panel that merges multiple data sources. In order to ensure the quality of the data, the data collector considers “temporal frequency and spatial accuracy of anonymized location point observations, temporal coverage and representativeness at the device level, and spatial representativeness at the sample and county level”. The data panel only includes mobile devices that meet strict data quality standards regarding anonymized location in order to further ensure the quality and consistency of the collected
data. A multi-level weighting method was used to extend the sample to the underlying population at the county and state levels (BTS, 2020).

This data source collects and updates data frequently, which allows access to the most updated daily travel data. Given the dynamic change of the pandemic, using the most up-to-date data is critical in providing meaningful results to capture real-life situations. In order to examine U.S. residents’ daily trips amid COVID-19, March 1, 2020, was selected as the beginning point since it was the time when we started receiving reliable data on COVID-19 cases and deaths.

Hence, daily trips from March 1, 2020, to November 29, 2020, was selected as the beginning point since it was the time when we started receiving reliable data on COVID-19 cases and deaths. The COVID-19 dataset includes three numeric variables, which are numeric in the ratio scale.

Table 1 presents the variables, variable description, and variable type. Due to the focus on residents’ daily trips at the national level, the dataset was consolidated for the entire country, which includes 274 observations with ten major variables, excluding the date. All variables are numeric in the ratio scale.

For the COVID-19 data, CDC COVID Data Tracker was used to extract the data, including daily new cases and deaths in the United States. CDC is one division of the Department of Health and Human Services with the responsibility to protect the United States from health, safety, and security threats (CDC, 2020c). COVID-19 cases and deaths are reported to the CDC by states (CDC, 2020a). In addition, daily test data was collected from the COVID tracking project (COVID Tracking Project). The data from these two sources were merged by date, for the period from March 1, 2020, to November 29, 2020. The COVID-19 dataset includes three numeric variables, new daily cases, new daily deaths, and new tests. It is worth noting that the COVID-19 dataset does not capture the actual number of new cases and new deaths in real-time. Due to the lead time in the testing procedure, it usually takes several days to two weeks to receive the test results. Similarly, there are also some delays in reporting daily deaths since it requires time to file the death report and find the cause. Accordingly, these numbers mainly represent reported cases and deaths on a given date, not the cases and deaths actually occurring on that date. However, the lapse between reported cases and actual occurrence does not affect the results of this study because it examines residents’ travel behavior, which tends to be based on the reported cases and deaths they see on the Internet and media.

2.2. Data analysis methods

This study aims at exploring a potential two-way relationship between U.S. residents’ daily travel and COVID-19 infection. Due to the exploratory nature of the study, several data analysis methods were used in this paper to find the answers to the research question. First, a two-tailed Pearson’s correlation test was conducted to explore correlations between daily trips by distance and daily new cases and deaths. Since Pearson’s correlation results only show the potential correlation and effect size between two variables without specifying which one impacts another, it is an appropriate method to explore a two-way relationship. In other words, a significant correlation indicates a relationship is possible in any direction, and the correlation coefficient shows the effect size of this relationship. The results provide some initial insight into how the reported new cases and deaths have any correlations with trips that U.S. residents make daily. However, due to the simplicity and limited extent of this test, it does not provide any detailed results of the relationships holistically nor projections of trips in the near future. Hence, this test was primarily used as an initial exploration of the data.

Second, in order to forecast the future values of daily trips, a time-series forecast method was used. Time-series forecast is a statistical method that predicts future values of a variable based on time-series data and trend analysis. This method establishes the dependence of two observations of the same variable at two different time points (Box and Jenkins, 1976; Brockwell and Davis, 1991). It can detect seasonal patterns in the data, estimate the trend and growth of a variable over time, and produce reliable results. A big advantage of this method is that it does not require predictors to forecast the future values of the target variable. The only assumption is that the target variable is dependent on time. Given the uncertainty of residents’ daily trips during the pandemic, there are limitations with using traditional predictive modeling methods such as multiple regression analysis, which requires up-to-date data for various predictors to produce meaningful results. Such data are not available at this time since COVID-19 is a novel situation and changes dynamically every day. Additionally, since the relationship between COVID-19 infection and daily trips could be two-way, such a method could lead to a false interpretation of the outputs due to the presumed relationships between predictors and the target variable. This study aims at exploring this two-way relationship objectively; hence, time-series forecast is the appropriate choice because it can produce a forecast of future daily trips primarily on the time factor.

In this study, ARIMA (Auto Regressive Integrated Moving Average) method was used to build forecasting models for trips by distance. AR stands of Autoregression, I for Integrated, and MA for moving average. It is a robust time-series forecast method that predicts future values based on the dependent relationship between an observation and the number of lagged observations, differencing of raw observation and dependency between an observation and a residual error from a moving average model (Box and Jenkins, 1976; Box et al., 2016). There are three major parameters in the ARIMA model: p, d, and q. p is autoregressive order, d is differencing order, and q is moving average order. The general forecasting formula using ARIMA method is expressed in the following equation

\[
\hat{y}_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q}
\]

where

- \(\hat{y}_t\): the forecast value of the target variable at the time point \(t\)
- \(y_{t-i}\): the value of the target variable in the prior time point \(t-i\)
- \(\mu\): intercept
- \(\phi_i\): autoregression coefficients; \(i\) autoregressive order, \(i = 1, \ldots, p\)
- \(\theta_i\): moving average coefficients; \(i\) moving average order, \(i = 1, \ldots, q\)
- \(e_t\): the error value at time point \(t\)
\( \theta \): moving average parameter; \( j \): moving average order, \( j = 1, \ldots, q \)

\( \varepsilon_t \): lag error at the time \( t-i \)

Finding appropriate parameters is key to this method, and it requires an iterative process, which includes trials and errors to fine-tune these values. First, autocorrelation analysis was conducted to examine autocorrelations for all lags, based on which a value of \( p \) can be decided. Then, values of \( d \) and \( q \) were determined through an iterative process of trials and errors until a satisfactory model fit and forecast accuracy was achieved. Residual autocorrelation function charts were examined to ensure only white noise exists. Model fit statistics include \( R^2 \) and Ljung-Box Q statistic test. \( R^2 \) of 0.5 or higher is generally considered acceptable (Hair et al., 2010). In addition, a non-significant Ljung test indicates there is no evidence of lacking model fit. In other words, the residuals are merely white noise (Box et al., 2016). The forecast accuracy is evaluated by the Mean Absolute Percent Error (MAPE). MAPE of 10% or lower is considered acceptable, indicating that the forecast model is about 90% accurate. Ten ARIMA models were constructed and tested for ten trip categories, as described in Table 1. Those models produced forecast values for trips by distance from November 30, 2020, to February 28, 2021.

Third, a comparative trend analysis was conducted to compare trends of future COVID-19 metrics and future daily trips by distance. The comparison allows us to understand the potential two-way relationships among those trends. This comparison requires forecast values for COVID-19 metrics. Since there have been many available forecast models, the author reviewed and selected validated models instead of developing a new one. According to the CDC (2020a), CDC (2020b), CDC (2020c), CDC (2020d), there are two groups of COVID-19 forecast models. The first group models future cases and deaths based on assumptions on how levels of social distancing will change in the future. Examples of such models include the ones developed by Johns Hopkin University, Columbia University, Georgia Institute of Technology, Institute for Health Metrics and Evaluation, and Yougyang Gu. The second group makes assumptions that existing social distancing measures will continue through the projected period. They include models developed by institutions, such as Carnegie Mellon University, London School for Hygiene and Tropical Medicine, Notre Dame University, Texas Tech University, University of Texas, Austin, University of Michigan, and Walmart Labs Data Science Team (CDC, 2020d). In this paper, the forecast model for new daily cases and deaths by COVID-19 Simulator Consortium was selected. This model is in the second group and developed by a team of experts from Massachusetts General Hospital, Harvard Medical School, Georgia Institute of Technology, and Boston University School of Medicine. The model uses the Susceptible, Exposed, Infectious, and Recovered compartments (SEIR) method to predict how COVID-19 is transmitted from one to another by moving the population through a set of connected compartments based on rates of incubation, infection, and recovery. Two key assumptions of this model are the current intervention will continue and all cases can be isolated to stop the transmission of COVID-19 when the rate of active cases falls below ten per million (COVID-19 Simulator Consortium, 2020). These assumptions make this model a good fit for the purpose of this study, which focuses on the current trend of the pandemic rather than the impacts of different interventions (COVID-19 Simulator Consortium, 2020). This model produces forecasts of new daily cases and deaths until March 31, 2021.

### 3. Results

#### 3.1. Descriptive statistics

The pandemic has been spreading fast in the United States in the past eight months. Fig. 4 presents the overall trends of some critical nationwide COVID-19 metrics, including daily tests, daily cases, daily hospitalized, and daily deaths. Daily tests started low at around 100,000 tests in early April 2020 and then increased linearly to approximately 900,000 tests per day in mid-July 2020, when it started to slow down to about 700,000 tests per day in August 2020. Finally, the daily tests reached the highest number in early December 2020.

![Nationwide COVID-19 key metrics](COVID Tracking Project, 2020)
with more than 1.8 million tests on December 7. Daily cases, somewhat dependent on daily tests, increased in April 2020 to a high peak of about 35,000 new cases per day, and then decreased in May 2020, during stay-home-orders, to the lowest of about 20,000 new cases per day. Then the number increased again at a higher rate in early June 2020, during state reopening, and reached a very high number of about 75,000 new cases in mid-July 2020 before it started decreasing again to about 40,000 new cases in August 2020. The highest number of daily new cases was reported in early December 2020, with more than 180,000 new cases on December 7. Daily deaths follow a similar pattern, with the highest in April 2020, when we observed more than 2500 deaths per day, and then a decrease to the lowest of about 600 deaths per day in late June 2020, before it increased again to about 1500 deaths per day in mid-July 2020. The daily deaths continued to increase toward the end of November 2020, with almost 2500 new deaths. Similarly, the number of hospitalized reached a peak number of about 60,000 patients per day in mid-April and mid-July 2020. The lowest number of hospitalized of less than 30,000 patients occurred in June 2020, and the highest number of hospitalized was reported in early December 2020, with more than 100,000 hospitalized cases on December 7.

As for U.S. residents’ daily trips during 2020, Fig. 5 presents the percentages of trips by distance, and Table 2 shows key descriptive statistics of trips by distance at the national level, including mean, min, max, and standard deviation. From March 1 to November 29, 2020, short trips, including trips less than one mile and trips between 1 and 3 miles, account for about 50% of the total daily trips. The next frequent ones are the trips between 5 and 10 miles (about 20%) and trips between 10 and 25 miles (about 12%). Interestingly, trips between 3 and 5 miles only account for less than 10% of the total trips. For trips longer than 25 miles, the longer the distance, the fewer the trips. The least frequent trips are trips greater than 500 miles, followed by trips between 250 and 500 miles and trips between 100 and 250 miles.

### 3.2. Correlational analysis results

Bivariate correlational analysis was conducted using two-tailed Pearson correlation coefficients to examine potential correlations between trips by distance and COVID-19 metrics. The results are shown in Table 3, indicating that New Tests have a significant correlation with new cases, which matches current findings since the higher the test number, the more cases are detected. In addition, there is a significant correlation between new cases and new deaths, i.e., more cases could lead to more deaths. However, the correlation coefficient of this relationship (0.257) is much lower than the correlation coefficient between new tests and new cases (0.817). As for trips by distance, it is worth noting that eight of ten daily trip variables have significant negative correlations with the reported New Deaths. That means the number of reported deaths can have a negative effect on the number of trips people choose to make daily. Additionally, Pearson’s correlation coefficient indicates that the effect size is largest for the trips between 100 and 250 miles (−0.483). On the other hand, shortest trips, including trips less than one mile and trips between

![Fig. 5. Percentages of number of trips by distance (BTS, 2020).](image-url)
residuals are mainly white noise. For the other three models with

data mining that there is no evidence of lacking model

major. The autocorrelation analysis results were examined to deter-
mine the appropriate p value for the model. Then, the iterative process
for the model, the autocorrelation analysis results were examined to deter-
mine the appropriate p value for the model. Then, the iterative process
for the model, the autocorrelation analysis results were examined to deter-

1 and 3 miles, have no significant correlations with New Deaths. As for
New Cases, only six of ten correlations are significant. Among signifi-
cant correlations, most of them are negative, indicating a possible neg-


datiﬁcation (−0.44), while the trips between 3 and 5 miles have the lowest corre-
lation coefﬁcient (−0.121). Trips between 5 and 50 miles have no signi-
ficant correlations with New Cases. An interesting ﬁnding is that trips
between 50 and 100 miles have a positive correlation with
New Cases. Thus, the negative effects of new deaths on daily trips are
more apparent than the effects of new cases. These ﬁndings are
the initial exploration of the data and do not seem to provide a clear
pattern of relationships between COVID-19 infection and daily trips.
The result is not unexpected due to the uncertainty of COVID-19 infec-
tion and unpredictable U.S. residents’ travel behavior during the pan-
demic. The next sections will examine the forecast results and
comparative trend analysis to shed further light on this relationship.

3.3. Forecast results

Ten non-seasonal ARIMA models were constructed with a 95% con-
fidence interval to forecast future daily trips by distance. For each
model, the autocorrelation analysis results were examined to deter-
mine the appropriate p value for the model. Then, the iterative process
was used to ﬁnd the appropriate values for q and m to ensure the model
ﬁt. The best ﬁt parameters found for ARIMA models are
p = 7, d = 1, and q = 1. Table 4 presents the forecast results for
ten ARIMA models, including R², Ljung-Box Q statistics, signiﬁcance
values, and MAPEs. Overall, most models achieved a good model ﬁt
with R² greater than 0.7. In addition, seven models have non-
signiﬁcant Ljung-Box Q statistics at the 0.01 signiﬁcance level, indicat-
ing that there is no evidence of lacking model ﬁt. In other words, the
residuals are mainly white noise. For the other three models with

Table 3

Bivariate correlation result – Pearson’s correlation coefﬁcients.

| Trips by distance | New Cases | New Deaths |
|-------------------|-----------|------------|
| Trips < 1 mile    | −0.440**  | −0.062     |
| Trips 1–3 miles   | −0.177**  | −0.074     |
| Trips 3–5 miles   | −0.121*   | −0.145*    |
| Trips 5–10 miles  | −0.118*   | −0.174*    |
| Trips 10–25 miles | −0.092    | −0.141*    |
| Trips 25–50 miles | 0.048     | 0.174*     |
| Trips 50–100 miles| 0.319**   | −0.337**   |
| Trips 100–250 miles| −0.019   | −0.483*    |
| Trips 250–500 miles| −0.144*  | −0.363*    |
| Trips greater than 500 miles | −0.131* | −0.139* |
| New Cases         | 1         | 0.275**    |
| New Deaths        | 0.275**   | 1          |
| New Tests         | 0.817**   | 0.043      |

**Correlation is signiﬁcant at the 0.01 level
*Correlation is signiﬁcant at the 0.05 level

1 and 3 miles, have no signiﬁcant correlations with New Deaths. As for
New Cases, only six of ten correlations are signiﬁcant. Among signiﬁ-
cant correlations, most of them are negative, indicating a possible neg-
ative effect of reported new cases on daily trips. The trips less than one
mile have the highest negative correlation coefﬁcient with New Cases
(−0.44), while the trips between 3 and 5 miles have the lowest corre-
lation coefﬁcient (−0.121). Trips between 5 and 50 miles have no signiﬁ-
cant correlations with New Cases. An interesting ﬁnding is that trips
between 50 and 100 miles have a positive correlation with
New Cases. Thus, the negative effects of new deaths on daily trips are
more apparent than the effects of new cases. These ﬁndings are
the initial exploration of the data and do not seem to provide a clear
pattern of relationships between COVID-19 infection and daily trips.
The result is not unexpected due to the uncertainty of COVID-19 infec-
tion and unpredictable U.S. residents` travel behavior during the pan-
demic. The next sections will examine the forecast results and
comparative trend analysis to shed further light on this relationship.

3.4. Comparative trend analysis results

A comparative trend analysis was conducted by comparing the cur-
rent and future trends for COVID-19 metrics and daily trips by dis-
tance. Fig. 6 shows the current and future trends of COVID-19 new
cases and new deaths until March 31, 2021. The current and future
trends for trips by distance for the same period are presented in
Figs. 7–10. The trend analysis revealed the following results.

3.4.1. Travel behavior and COVID-19 infections

Future comparative trend analysis results are summarized in
Table 5, focusing on the current trends since August 2020 and future
trends since December 2020. The results show that new COVID-19
cases and deaths slightly decreased in late August and increased again
since early October 2020. These numbers are projected to increase
exponentially until they reach the highest in late-January or mid-
February 2021 and will start to decrease. Thus, the future trends for
new cases and deaths seem very consistent with a small lag for new
daths.

As for daily trips, it is worth noting that trips less than 10 miles
have very similar trends and patterns. In addition, trips between 10
and 50 miles seem more ﬂuctuating, but the overall trends share simi-
larities with trips less than 10 miles. Accordingly, these trips are
grouped together as short trips. The trend analysis for those trips
shows a peak in mid-September, which could be a result of the
decrease of new cases and deaths in late August 2020. Then the trips
decreased due to the increased cases in October. However, these trips
increased again in mid-November 2020, possibly due to the Presiden-
tial Election and Thanksgiving. The decrease in those trips afterward
could be due to the awareness of the high surge of COVID-19 cases
and deaths in November 2020. These trips are projected to increase
in December 2020, which could be due to the Christmas holiday
despite the current pandemic. This increase can explain the continued
increase in new cases and deaths. In January 2021, these short trips
are projected to decrease, followed by a stable decrease in February
2021. The stabilized decrease in daily short trips can explain the
projected decrease of new cases in late January and new deaths in

Table 4

ARIMA time series forecast results.

| Models | R-squared | Ljung-Box Q Statistics | Sig. | MAPE |
|--------|-----------|-------------------------|------|------|
| Trips < 1 mile - Model 1 | 0.803 | 16.447 | 0.088 | 4.781 |
| Trips 1–3 miles – Model 2 | 0.728 | 11.279 | 0.336 | 4.078 |
| Trips 3–5 miles - Model 3 | 0.720 | 16.097 | 0.097 | 4.045 |
| Trips 5–10 miles - Model 4 | 0.856 | 13.918 | 0.003 | 3.140 |
| Trips 10–25 miles - Model 5 | 0.907 | 17.963 | 0.000 | 2.810 |
| Trips 25–50 miles - Model 6 | 0.911 | 31.874 | 0.000 | 2.972 |
| Trips 50–100 miles - Model 7 | 0.907 | 10.195 | 0.017 | 3.289 |
| Trips 100–250 miles - Model 8 | 0.746 | 20.529 | 0.025 | 7.160 |
| Trips 250–500 miles - Model 9 | 0.759 | 14.569 | 0.017 | 3.289 |
| Trips > 500 miles – Model 10 | 0.886 | 13.452 | 0.199 | 14.419 |
mid-February 2021. Thus, residents’ travel behavior seems to change dynamically depending on their risk perception of the coronavirus along with holiday times. An interesting finding is that a change in their travel behavior, such as increased daily trips, contributes to rising infections and deaths.

3.4.2. A lag in travel behavior between short and long trips

Different trends are observed with trips greater than 50 miles. These numbers are projected to decrease to a certain point before increasing again. More specifically, these trips decreased in mid-September and then increased in mid-October through early November 2020. These numbers seem inconsistent with the short trips described above. While short trips have an increase in mid-September, this increase can only be seen a month later in long trips over 50 miles. Similarly, the short trips have a decrease in early October before an increase in late November. At the same time, long trips only show an increase in mid-October, followed by a decrease in mid-November. Finally, short trips are projected to have an increase in December 2020 and start decreasing slowly in January and February 2021. On the other hand, longer trips are projected to only increase slightly in December 2020 but will continue to increase in January and February 2021. This mismatch requires further investigation. By comparing the patterns between short and long trips in the past, we can see that trips above 50 miles increased in April 2020, after the reopening, and fluctuated in July, but did not decrease in August despite spiking new cases observed in July 2020, as seen with trips below 50 miles. Actually, long trips increased to reach the highest in mid-August 2020, especially trips over 250 miles, which showed a really big jump in mid-August. We can only see a decrease in mid-September, as described above. This comparison indicates a month-lag in the travel behavior between short trips and long trips, i.e., residents’ travel behavior for long trips seems to react slower to new cases and new deaths than for short trips. This lag also explains the increase of these long trips in mid-October 2020 and the possible increase in January and February 2021. The lag is a very interesting new finding, which is not discovered in the correlation analysis results or current observation of coronavirus infection and residents’ daily trips.

4. Discussion and conclusions

4.1. Discussion: a closed loop scenario of COVID-19 spreading based on travel behavior

As the future development of the COVID-19 pandemic is still an unknown factor, understanding the future trend of residents’ travel behavior is critical to creating awareness of the coronavirus spreading
and developing proper mitigation policies to contain the spread. The past data shows that U.S. residents became less aware of the actual risk of COVID-19 during the reopening period and Thanksgiving holiday, which lead to a significant rise in new cases and deaths in the past few months. This paper explores the relationship between the trends of residents’ daily trips by distance and COVID-19 metrics. The correlation analysis results indicate that residents’ daily trips longer than 3 miles are negatively affected by daily new deaths, but the effects of new cases are much less apparent. The time-series forecast for daily trips and the trend analysis results reveal that the patterns are not that simple. This study’s results show new interesting findings and shed more light on the spread of coronavirus in our community.

The results of the trend analysis revealed a possible closed loop scenario between residents’ travel behavior and COVID-19 infections in the United States. As shown in Fig. 11, the loop starts with a decrease of reported new cases and deaths, which creates an impression in the community that it is safer now to go out since the risk of contracting COVID-19 is lower. Residents will start making more short trips to stores, works, beaches, visiting families and friends, and other activities. Due to the lag in travel behavior, longer trips will start increasing one or two months later. As more short trips are made, it is harder to keep physical distancing and ensure the safety protocol due to the density of population in the areas. As a result, the risk of COVID-19 infections will increase, causing the further spread of coronavirus in the community. We will start to see more infections daily, which will likely lead to more daily deaths. With the increase of reported new cases and deaths, the residents start feeling less safe and decide to reduce the number of daily trips. The decreased trips by distance mean fewer people on the road, in stores, or restaurants, which will eventually lead to fewer reported cases and deaths, which, in time, will make residents feel safer; and we will go back to the beginning of the loop. The lag in travel behavior between short and long trips is the reason for some lag in reported cases and deaths. This lag is a big challenge since some longer trips still occur during the surge of COVID-19, which makes it harder to mitigate the risk.

The important question is how to break this loop. Fig. 12 shows the Institute for Health Metrics and Evaluation’s (IHME) prediction for daily infections in four different scenarios: current projection, when mandates are easing, when mask mandate is made universal, and when we have rapid vaccine rollout to high risk populations. The
projection shows that when the current mandates are easing, infections will rise exponentially and uncontrollably until late February 2021 before it starts decreasing. The other hand, a universal mask mandate can reduce daily infections substantially. Based on the projection, with the universal mask mandate across states, new cases can start decreasing immediately and continue to fall. The rapid rollout of vaccines to high risk populations is certainly important but will not have much effect in the beginning due to the time it takes for the vaccine to be effective. It is projected that if vaccines become available in late December 2020, new cases will still increase until late January 2021 and then start decreasing at a higher rate compared to the regular situation due to the effect of vaccines.

Based on these projections, there are several scenarios for U.S. residents’ daily travel behavior and COVID-19 infection based on the intervention strategies:

- **Current intervention:** If nothing changes, daily trips will increase in December 2020 due to the Christmas and New Year holidays despite the continued increase of new cases and deaths. The spike of COVID-19 will make people feel less safe and, therefore, make fewer trips, which will likely lead to fewer reported cases. The lag in travel behavior will result in more long-distance trips though, as described above. The loop seems to start again since people will start making more trips when they notice the decrease of new cases and deaths. Basically, this is the scenario with a possible infinite closed loop.

- **Universal masks mandate:** If most states mandate facial masks outdoor or when there are contacts, new cases and deaths will decrease immediately. Residents will feel safer and make more daily trips. While daily trips will increase, the mask mandate, along with other restrictions some states may issue, will help contain the spread of COVID-19; thus, new cases and deaths will continue to decrease. Eventually, we can break this loop and reduce the burden on hospitals and healthcare systems. When vaccines are widely available to the public, possibly in June 2021, the pandemic can be contained. This is possibly the best case scenario.

- **Rapid vaccine rollout to high risk populations:** As shown in the projection above, this scenario will not lead to an immediate decrease of new cases and deaths. That means short trips will start decreasing after the holidays due to the increase of new cases and deaths,
as in the first scenario (we may not see a decrease in long trips due to the lag in travel behavior). However, when vaccines become effective, new cases and deaths will decrease at a higher rate. Then, daily trips will increase again because residents start feeling safer.

Fig. 9. Forecast trends for Trips 50–100 miles, Trips 100–250 miles, and Trips 250–500 miles.

Fig. 10. Forecast trends for Trips > 500 miles.
How soon the pandemic can be contained will depend on how quick the vaccines will become available to the public. If we have vaccines for a large portion of the population in June 2021, we may still experience the closed loop situation between January and June.
2021 because the vaccines are only available for high-risk populations during this time. Nonetheless, the situation will get worse. We have noticed high surges in several states during the Presidential Election and Thanksgiving holiday due to the lack of restrictions. More people travel for Christmas and New Year holidays, and the mandates easing will contribute to a high surge of new cases and deaths in December 2020. It is likely that short trips may start decreasing in late December since some residents will feel less safe. But due to the lag in travel behavior, long trips will continue to increase. That means more people continue to travel to different cities or even different states for business or personal purposes, which will likely lead to more cases and deaths. The closed loop will expand in state and out-of-state residents. Even with vaccine rollout for high-risk populations, the pandemic would still be much harder to contain, given the continued spread of COVID-19 at a higher rate. It may take another year or more to get the pandemic in control due to uncertainty in residents’ travel behavior. Thus, this could be considered the worst-case scenario for our community.

4.2. Conclusions

This paper provides major contributions to COVID-19 literature and sheds new light on efforts of containing the spread of coronavirus by exploring the COVID-19 infections from the residents’ daily travel perspective. The results show correlations between trips by distance and new cases and deaths. The forecast and trend analysis revealed a closed loop scenario, a very dangerous situation, in which residents’ travel behavior dynamically changes based on their risk perception of COVID-19, which further contributes to the spread of COVID-19 in the community without an end. The situation becomes more tricky with a lag, one to two months, in travel behavior between short trips and long trips. This lag further worsens the closed loop situation since it creates more uncertainty and unexpected daily travels during the pandemic.

This closed loop is a new situation that we have not been aware of nor prepared to deal with. If there is no proper and immediate mitigation, it will not be possible to break this loop any time soon, and we continue to see more innocent people suffer and even die from the coronavirus. Any easing in restrictions and safety mandates will lead to a shorter life cycle of the loop, which will put more burden to hospitals and healthcare systems and likely cause faster spreading and more deaths. The data shows that our attempts to flatten the curve have failed, which is, in essence, one of the reasons leading to this closed loop situation. Four scenarios with different intervention strategies are discussed in this paper to show how resident’s daily travel changes in each situation and the likelihood of containing COVID-19 infections. The closed loop can only be broken with proper interventions and a rapid rollout of vaccines to the public. Various safety mitigation strategies should be used, including increasing awareness among residents about the danger of COVID-19, providing accurate updates of COVID-19, requiring mask covering in public areas, enforcing physical distancing, producing vaccine and treatment and delivering them quickly to the public, and controlling misinformation in the media. These strategies must be implemented in a multi-level process to be effective.

It is important to note that the study is exploratory rather than confirmatory. The findings of this study are based mainly on residents’ travel behavior at the national level, assuming time is the primary factor and the effects of other factors remain equal. The closed loop situation and how it changes in different scenarios are outlined and discussed based on the comparative analysis of future trends for daily travels and COVID-19 infection. Future research should examine the effects of demographic variables to understand better travel behavior by gender, age, ethnicity, education, travel purpose, modality, and location. The more we understand about the travel behavior, the more likely we can develop effective mitigation strategies to break the loop, or, at least, to lengthen its life cycle to reduce the burden to the healthcare systems and save more lives.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Beck, M.J., Hensher, D.A., 2020. Insights into the impact of COVID-19 on household travel and activities in Australia – the early days of easing restrictions. Transport Policy 99, 95–119. https://doi.org/10.1016/j.tranpol.2020.08.004.

Box, G.E.P. & Jenkins, G.M., 1976, Time Series Analysis: forecasting and control. Oakland, California: Holden-Day.
Box, G.E.P., Jenkins, G.M., Reinsel, G.C., Ljung, G.M., 2016. Time Series Analysis: Forecasting and Control. John Wiley & Sons, Inc, Hoboken, NJ.

Brockwell, P.J., Davis, R.A., 1991. Time Series: Theory and Methods. Springer-Verlag, New York.

Bruinen de Bruin, Y., Lequarre, A.-S., McCourt, J., Clevestig, P., Pigazzani, F., Zare Jeddi, M., Colosio, C., Goulart, M., 2020. Initial impacts of global risk mitigation measures taken during the combatting of the COVID-19 pandemic. Safety Sci. 128, 104773. https://doi.org/10.1016/j.ssci.2020.104773.

Bureau of Transportation Statistics (BTS), 2019. About the Bureau of Transportation Statistics. Bureau of Transportation Statistics, February 16. Retrieved from https://www.bts.gov/about-BTS.

Bureau of Transportation Statistics (BTS), 2020. Daily Travel During COVID-19 Pandemic. Bureau of Transportation Statistics, December 10, 2020. Retrieved from https://www.bts.gov/browse-statistical-products-and-data/trips-distance/daily-travel-during-COVID-19-pandemic.

Centers for Disease Control and Prevention (CDC), 2020. CDC COVID Data Tracker. Centers for Disease Control and Prevention, September 8. Retrieved from https://covid.cdc.gov/COVID-data-tracker/#cases.

Centers for Disease Control and Prevention (CDC), 2020. COVID-19 model FAQs. Institute for Health Metrics and Evaluation, December 8. Retrieved from http://www.healthdata.org/COVID/faqs#Scenario.

Kraemer, M.U.G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D.M., du Plessis, L., Faria, N.R., Li, R., Hanage, W.P., Brownstein, J.S., Layan, M., Vespiignani, A., Tian, H., Dye, C., Pybus, O.G., Scarpino, S.V., 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 368 (6490), 493–497. https://doi.org/10.1126/science.abb4218.

Linka, K., Rahman, P., Goriely, A., Kuhl, E., 2020. Is it safe to lift COVID-19 travel bans? The Newfoundland story. Comput. Mech. 66 (5), 1081–1092. https://doi.org/10.1007/s00466-020-01899-x.

Mervosh, S., Lu, D., & Swales, V., 2020. See Which States and Cities Have Told Residents to Stay at Home. New York Times, April 20. Retrieved from https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html.

Washington Post, 2020. Where states reopened and cases spiked after the United States shut down. Washington Post, September 9. Retrieved from: https://www.washingtonpost.com/graphics/2020/national/states-reopening-coronavirus-map/.

Yilmazkuday, H., 2020. COVID-19 spread and inter-county travel: Daily evidence from the U.S. Trans. Res. Interdisciplinary Perspectives 8, 100244. https://doi.org/10.1016/j.trip.2020.100244.