Machine learning model to project the impact of COVID-19 on U.S. motor gasoline demand

Shiqi Ou[a], Xin He[b *], Weiqi Ji[c], Wei Chen[d], Lang Sui[b], Yu Gan[e], Zifeng Lu[e], Zhenhong Lin[a], Sili Deng[c], Steven Przesmitzki[b], Jessey Bouchard[b]

a. Energy and Transportation Science Division, Oak Ridge National Laboratory. 2360 Cherahala Blvd, Knoxville, Tennessee 37932, USA.
b. Aramco Services Company: Aramco Research Center – Detroit, 46535 Peary Ct, Novi, Michigan 48377, USA.
c. Department of Mechanical Engineering, Massachusetts Institute of Technology. 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.
d. Michigan Department of Transportation. 18101 W Nine Mile Rd., Southfield, MI 48075, USA.
e. Energy Systems Division, Argonne National Laboratory, Lemont, Illinois 60439, USA.

[*]Corresponding Author; xin.he@aramcoamericas.com, ORCID: 0000-0003-2492-4454

Citation:

Ou, S., He, X., Ji, W. *et al.* Machine learning model to project the impact of COVID-19 on US motor gasoline demand. *Nature Energy* (2020): 1-8. https://doi.org/10.1038/s41560-020-0662-1

Abstract

Owing to the global lockdowns that resulted from the COVID-19 pandemic, fuel demand plummeted and the price of oil futures went negative in April 2020. Robust fuel demand projections are crucial to economic and energy planning and policy discussions. Here we incorporate pandemic projections and people's resulting travel and trip activities and fuel usage in a machine-learning-based model to project the US medium-term gasoline demand and study the impact of government intervention. We found that under the reference infection scenario, the US gasoline demand grows slowly after a quick rebound in May, and is unlikely to fully recover prior to October 2020. Under the reference and pessimistic scenario, continual lockdown (no reopening) could worsen the motor gasoline demand temporarily, but it helps the demand recover to a normal level quicker. Under the optimistic infection scenario, gasoline demand will recover close to the non-pandemic level by October 2020.

Keywords: Gasoline demand, Projection, COVID-19 pandemic, Mobility behaviors, Machine Learning

Introduction

Since December 2019, the infectious coronavirus disease 2019 (COVID-19) quickly swept the world and reached pandemic status in just a few months [1,2]. Of the over 7.5 million cases reported worldwide, about 25% are in the United States (U.S.), making it the country with the largest number of confirmed cases in the world so far [3]. Bendavid et al.(2020) implies that the actual infections could be more widespread than indicated by the number of confirmed cases [4]. To cope with the pandemic, public health policies were implemented to slow the spread of COVID-19 by “flattening the curve” using “stay-at-home” policies. The unanticipated reduction in mobility, and therefore fuel demand, resulted in a glut of oil in the market and the West Texas Intermediate oil price plunged to an unprecedented negative value in April 2020 [5,6]. For the oil market, the U.S. is one of the largest energy consumers and oil producers in the world. Fluctuations and uncertainties in U.S. oil consumption impacts the petroleum supply chain and trends of the broader energy economy [7]. Beyond temporary and local oil price and demand shocks, U.S. gasoline demand alone is significant enough to impact longer-term investments in the global energy industry, which impacts the world economy as a whole.

While the value of oil has somewhat recovered since April, uncertainties in the U.S. economy persist because of the lingering pandemic. The International Monetary Fund estimated the annual change in real Gross Domestic Product (GDP) in the U.S. could be -5.9% in 2020 and the global economy could contract by 3%. To effectively reduce uncertainties, multiple models were developed by researchers to project the trend of the COVID-19 pandemic [8]. Projections of the evolution of COVID-19 pandemic trends show that lock-downs help reduce COVID-19 transmissions by 90% compared to the baseline in Austin, TX [9]. However, this unprecedented phenomenon could last for a few years: Kissler et al. (2020) suggested that, even after the pandemic peaked, COVID-19 surveillance should be continued since a resurgence in contagion could be possible as late as 2024 [8]. Therefore, beyond immediate economic responses, the longer-term impact on the U.S. economy may persist well beyond 2020. Therefore, an effective forecast or estimate of the pandemic impacts could help people well prepare and navigate around unknown risks. More specifically, reliably projecting the oil demand, a critical leading indicator of the state of the U.S. economy, is beneficial to related business activities and investment decisions.

There are studies that discuss the impacts of unexpected natural hazards/disasters on energy demand/consumption [10,11], or studies that evaluate impacts of occurred pandemic on tourism [12] and economics [13]. However, few studies have quantified and forecasted the oil demands under multiple pandemic scenarios, and this research is desperately needed. To date, studies focused on energy impacts of COVID-19 pandemic are limited to the short-term Energy Outlook released by the U.S. Energy Information Administration (EIA); this outlook uses a simplified evolution of COVID-19 pandemic to forecast the U.S. GDP, energy supplies, demands and prices until the fourth quarter of 2021 [14]. In this work, we develop a model that combines personal mobility with motor gasoline demand and uses a neural network to correlate personal mobility with the evolution of the COVID-19 pandemic, the government policies, and the demographic information.

In this study we extend the understanding of the COVID-19 impact on medium-term fuel demand as the pandemic evolves. This would be useful to both policy makers and stakeholders in balancing the COVID-19 spread and economic recovery. Some key findings are: under the reference infection scenario, the growth of motor gasoline demand in the U.S. is slow after a quick rebound in May, and it is unlikely that demand will recover to a non-pandemic level prior to October 2020; under both the reference and pessimistic infection scenarios, a continual lockdown (no-reopening) policy could worsen the motor gasoline demand temporarily, but it helps the demand recover to a normal level quicker due to its impact on infection rate; under the optimistic infection scenario, the projected trend of motor gasoline demand will recover to about 95% of the non-pandemic gasoline level (almost fully recover) by late September 2020; however, under the pessimistic infection scenario, the second wave of infections in mid-June to August could significantly lower the gasoline demand once more, but it will not be worse than it was in April. These results imply that government intervention, and its impact on infection rate, does influence infection rate which thereby impacts mobility and fuel demand.

Mobility and Pandemic

A result of policies aimed at “flattening the curve” is an unprecedented restriction to mobility. Beginning in March and through April, at least 316 million people in the U.S. lived in regions with some form of stay-at-home policy [15]. U.S. Department of Transportation data shows travel on all roads and streets in March 2020 decreased by 18.6%, an equivalent of 50.6 billion fewer vehicle miles compared with travel in March 2019 [16]. Subsequently, U.S. on-road transportation fuel consumption in April 2020 was 30% lower than it was in April 2019 [17].

Widespread availability of personal mobile devices has allowed us to quantitatively measure people’s confinement. Since the onset of the pandemic, Google and Apple released mobility reports or webpages using aggregated map data to help better quantify the correlation between reduction in mobility and the spread of COVID-19 [18–20]. Figure 1(a)-(c) shows changes in mobility in the U.S. reported by Google for the period between February 15 and June 4, 2020 for work, retail and recreation, and grocery and pharmacy [18]. The U.S. shutdown policies are governed at the state or more localized levels instead of a federal level, and there is considerable variation of mobility decrease among states. Although all three categories experience widespread decrease in mobility starting March 26, at the onset of many stay-at-home orders, the decrease is heterogeneous across the states and categories. Figure 1(d) shows the 7-day average for daily COVID-19 cases and deaths during the same period [21]. Although the daily number of cases of COVID-19 increased between March 16 and April 20, during a period of which much of the country was shut down, this increase is primarily due to the long incubation period of COVID-19 [22]. Likewise, an increase in the daily number of deaths occurred roughly two weeks after confirmed cases. The data clearly shows that the changes in mobility and pandemic data are highly non-linear. A sharp change in mobility is observed in the first two weeks after the pandemic reaches a threshold, then mobility plateaus and becomes insensitive to further increases in confirmed cases or deaths. We conduct a detailed correlation analysis between mobility and pandemic data to identify the key parameters being considered in the model development. The results of the correlation analysis are presented in Supplementary Figure 1 and Supplementary Note 1.

Since the change in mobility implies the change in vehicle miles traveled (VMT), the mobility data provides a solid foundation to analyze the fuel demand change in each travel category. This also means that a reliable fuel demand projection model can be created if the future mobility can be predicted based on the projected evolution of the pandemic.

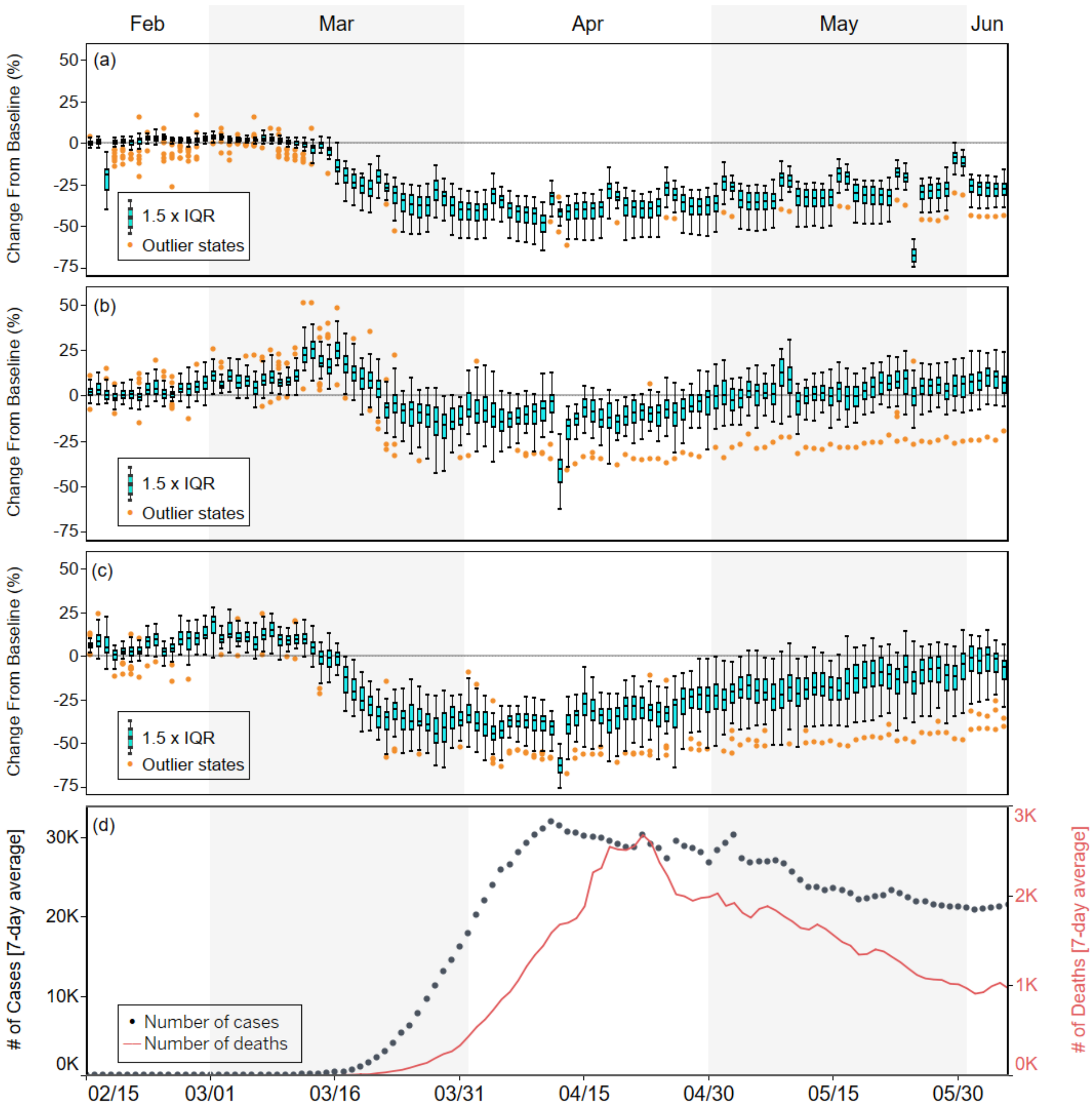

Figure 1. Changes in mobility in various sectors and number of cases. Percentage changes in Google mobility between February 15 and June 4, 2020 for (a) workplaces; (b) grocery and pharmacy; (c) retail and recreations [18]. (d) 7-day average of the number of COVID-19 cases and deaths in the U.S. [3] Please note the boxplots represent the interquartile range (IQR) where the lower and upper whiskers represent 1.5×IQR. Each yellow data point represents the mobility of outlier states that do not fall into the 1.5xIQR.

Pandemic Oil Demand Analysis (PODA) Model

Motor gasoline is the main U.S. transportation fuel, accounting for about 58% of total transportation sector energy consumption and 45% of total petroleum consumption [23]. In this work, a machine-learning based PODA model is developed to project the U.S. gasoline demand using COVID-19 pandemic data, government policies, and demographic information. As shown in Figure 2, the model contains two major modules: a Mobility Dynamic Index Forecast Module and a Motor Gasoline Demand Estimation Module. The Mobility Dynamic Index Forecast Module identifies the changes in travel mobility caused by the evolution of COVID-19 pandemic and government orders, and it projects the changes in travel mobility indices relative to the pre-COVID-19 period in the U.S. Notably, the change in travel mobility, which affects the frequency of human contact or the level of social distancing, can reciprocally impact the evolution of the pandemic to some extent, as the black dashed line shows in Figure 2. The Motor Gasoline Demand Estimation Module estimates VMT on pandemic days while considering the dynamic indices of travel mobility, and it quantifies motor gasoline demands by coupling the gasoline demands and VMT. The neural network model, which is the core of the PODA model, has 42 inputs, 2 layers, and 25 hidden nodes for each layer, with rectified linear units as the activation function (see the "Methods" section for details). The data sets used for model training, calibration, and validation are provided in Supplementary Note 2. In the PODA model, the potential induced travel demand due to the lower oil prices under the COVID-19 pandemic is not explicitly considered.

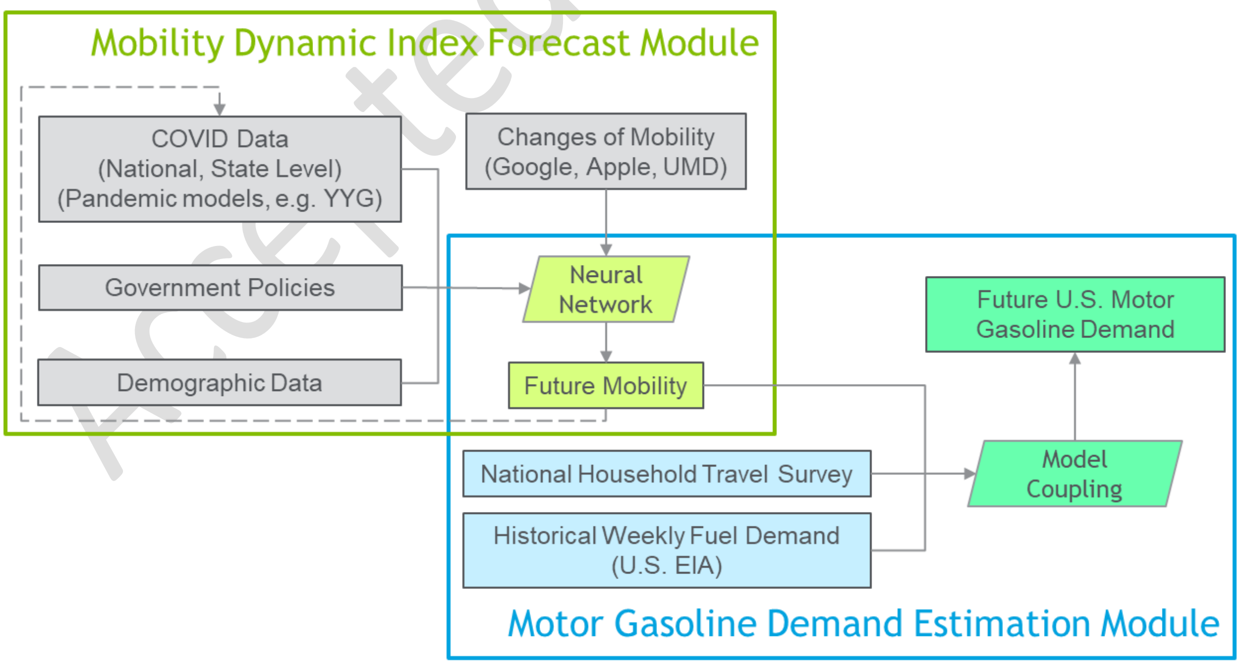

Figure 2. The structure of the PODA model.

The pandemic data used to project future motor gasoline demand is taken from the Youyang Gu (YYG) COVID-19 projection model, which is one of few models referenced by the U.S. Centers for Disease Control and Prevention (CDC) that offers a medium-term (three-month) forecast of infections and deaths for each state under multiple scenarios [8,24]. The model is updated daily providing new projections for the next three months. The YYG model generates three new infection cases (mean, lower, and upper) under a 95% confidence interval and uncertainties. Therefore, we divide our study into three scenarios that align with these projections: reference (mean), optimistic (lower) and pessimistic (upper). In addition, the pandemic data forecasted by the MIT model [8,25] is used to project the gasoline demand, which is compared against the projection based on inputs from the YYG model. The details of these scenarios are provided in Supplementary Note 3 and Supplementary Figures 5-7. We also provide the fuel demand of the "Non-pandemic" scenario for comparison. For this scenario, we use the four-year average from 2016 to 2019 to reflect the seasonal change in gasoline demand (see Supplementary Figures 8 and 9 for details).

Projections of Motor Gasoline Demand Under COVID-19 Pandemic

Figure 3 shows U.S. motor gasoline demand projections and state-level mobility projections up to September 21, 2020 under the reference pandemic scenario. The historical mobility and gasoline demand data for simulation in the model is from March 1 to June 5, and the pandemic data is from the YYG model projection as of June 10, 2020. Shown in Figure 3 the projected motor gasoline demands from the PODA model fit well with the historical motor gasoline demand from EIA as of June 5, 2020. As the reopening was successively implemented by states, the U.S. motor gasoline demand gradually recovered from its low point during the first weeks of April. By the week of June 5, 2020, the gasoline demand reached 7.9 million barrels per day (BPD), growing nearly 56% compared to the April 3 low point [17].

This study adopts the mobility trend data from Google [18] into the PODA model for projecting future motor gasoline demand, and uses the mobility data from Apple [19] as inputs for projected fuel demand comparison. The value deviations between the projected motor gasoline demands based on mobility trends from these two data sources are shown in Figure 3. For predictions after May, the projected gasoline demands based on Apple mobility data tend to be about 10% higher than the projections based on Google mobility data. Thus, though the mobility trends from Google and Apple are estimated using different methodologies—Google estimates mobility using location and visit duration data [18], and Apple uses route map requests [19]—these collection methods have a relatively limited effect on the projected motor gasoline demand. However, the Apple mobility data offers only the aggregated changes of mobility, which do not differentiate mobility among detailed trip activities/purposes and may not well represent some trip categories. Therefore, the gasoline demand projections based on Apple mobility data could result in a larger error. As a result, this study uses the inputs from Google mobility data for projection and scenario discussion.

The PODA model can also be used to project the mobility level. In addition to motor gasoline demand, Figure 3 also exhibits the projected mobility levels in each individual state from March to September 2020. Clearly, the mobility levels in all states are much lower in April than in March, and the differences are much larger for states that have more COVID-19 infections, such as New York and California. However, in May when many states start reopening, the mobility levels in most states increase dramatically. Especially in states from the Midwest and South, such as North Dakota and Alabama, mobility levels are already higher than they were in January 2020 when the baseline mobility level was low partially due to the cold weather. However, if this disease evolves as the reference scenario suggests—shown in Figure 5(a)—the mobility levels and gasoline demand in these states could stay low in mid-June to early August 2020 as the pandemic resurges. After that, the mobility and gasoline demand slowly recover as the daily confirmed cases gradually decline.

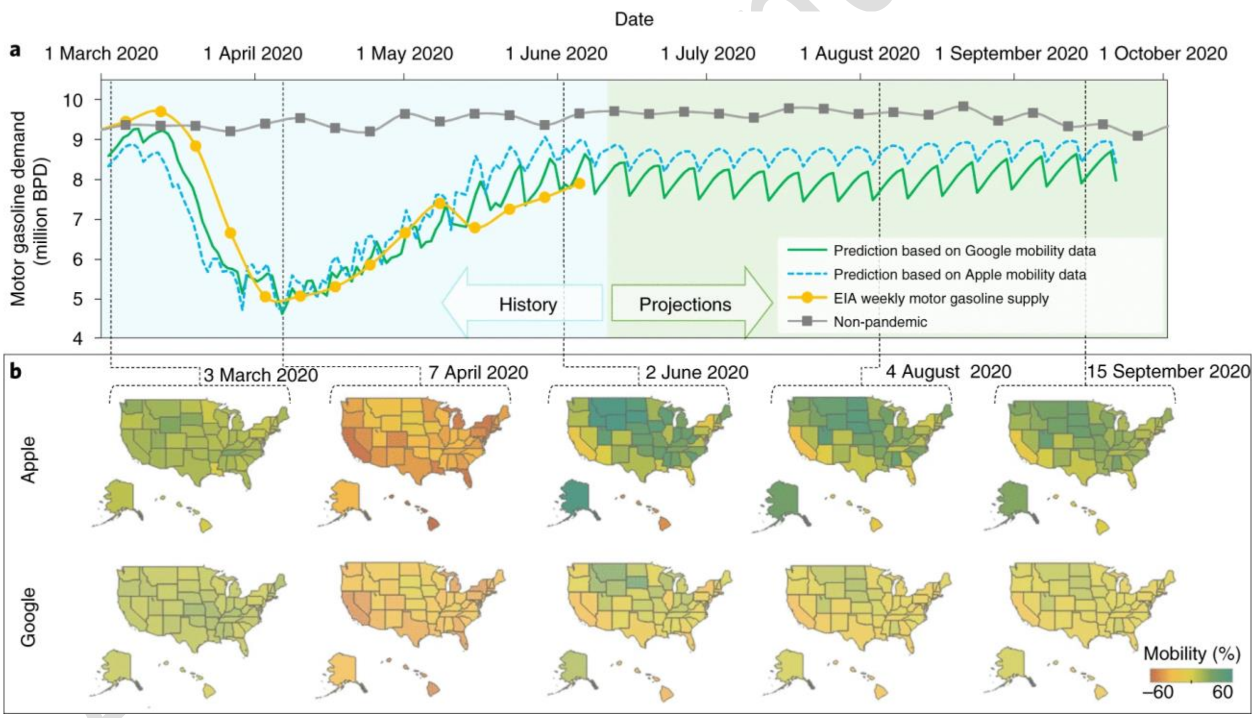

Figure 3. The projection of gasoline demand and mobility level in the U.S. under the reference pandemic scenario. (a) Comparison of gasoline demand projections based on the Apple mobility data and Google mobility data [18,19], EIA weekly motor gasoline supply and non-pandemic gasoline demands are provided for comparison [17]; (b) State-level mobility projections. "Apple" means the inputs for projection are from Apple, and "Google" means the inputs for projection are from Google. Mobility level at 0% in the legend bar for Apple mobility refers to the baseline on January 13, 2020 [19]. Mobility level at 0% in the legend bar for Google mobility refers to the baseline of the median value for the corresponding days of the weeks from January 3 to February 6, 2020 [18].

Dynamics of Future Motor Gasoline Demand

The knowledge deficit of this disease magnifies the uncertainty of the projections on motor gasoline demand. Figure 4 gives the corresponding motor gasoline demands under three different scenarios. To present the weekly changes to better show the trend and clear comparison between different scenarios, the projected values of gasoline demands shown in Figure 4 are all given as 7- day moving averages. The scenario details are described in Supplementary Note 3. Analyzing the medium-term projected motor gasoline demand to October 2020 under the reference scenario, it can be seen that, although the motor gasoline demand will grow rapidly after April recovery, it is expected to remain fairly flat from mid-May until October. The projected motor gasoline demand by the end of September will be about 8.342 million BPD. Compared to the demands under non-pandemic scenario, the projected gasoline demand in late September 2020 under the reference pandemic scenario is about 90% of the demand under the non-pandemic scenario. The motor gasoline demand is about 78% of the demand under the non-pandemic scenario on May 8, 2020; and it is about 82% on June 5, 2020 [17]. Therefore, the motor gasoline demand in the U.S. is unlikely to recover to a non-pandemic level in the medium-term (before October 2020) under the current reference pandemic scenario. This is primarily because the YYG model projects that the pandemic will continue to evolve, and a new smaller wave of infections could occur in mid-June to August, which could lessen people's desire to travel or prompt the government to adjust its intervention measures.

The projections of optimistic and pessimistic scenarios expand the probabilities of the motor gasoline demands in the medium-term. In the optimistic scenario, the reproduction values are believed to be smaller than at the beginning of the pandemic period, and people are assumed to be more willing to travel. Thus, the motor gasoline demand after April continues rapidly growing till October. The projected motor gasoline demand in late September 2020 under the optimistic scenario is about 98% of the demand under the non-pandemic scenario, which means the U.S. gasoline demand is almost normal by then. In the pessimistic scenario, the pandemic response measures are assumed to deteriorate, and a more serious wave of daily new infections are projected. This would result in more social distancing measures by governments, companies, and individuals, as well as increased concerns about travel. Thus, the motor gasoline demand after April declines to about 6.1 million BPD before it starts recovering by mid-August, but the demand will not be worse than it was in April. The projected motor gasoline demand in late September 2020 under the pessimistic pandemic scenario is about 78% of the demand projected for the non-pandemic scenario. Although results as of June 10, 2020, are presented in this work, the PODA model is updated regularly with the evolution of the pandemic (https://covid19-mobility.com/).

This study is also able to project the motor gasoline demands with pandemic inputs from other pandemic projection models. Considering the forecast period, and availability of model source codes, this study chooses the pandemic projections from the MIT model as an example [8,25]. The purple curve in Figure 4 shows the projected gasoline demands in the U.S. based on inputs from the MIT model, which is very close to the demand projected using the optimistic scenario from the YYG model. This is because the evolution of the COVID-19 pandemic projected by the MIT model is similar to that projected by the YYG model under the optimistic scenario [24,25], as shown in Supplementary Figure 5.

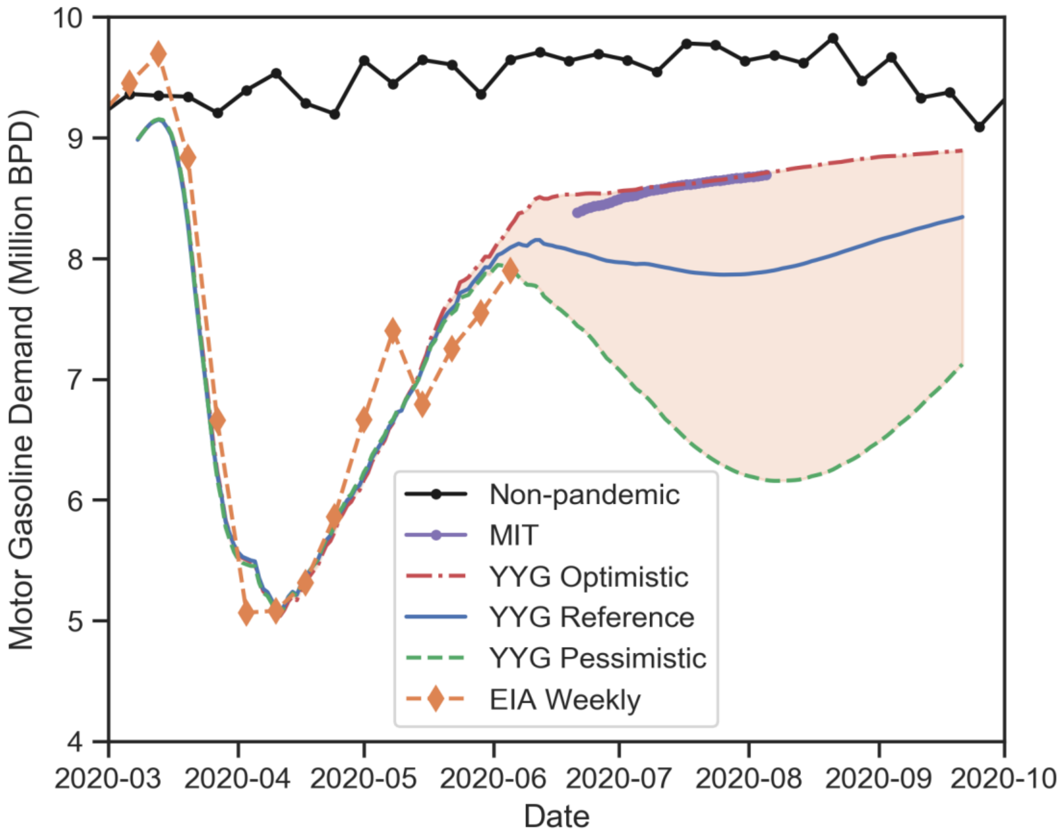

Figure 4. Gasoline demand under different scenarios (projected by June 10, 2020). Gasoline demand projections are based on the three scenarios of the YYG model and MIT model [24,25]. EIA weekly motor gasoline supply and non-pandemic gasoline demands are provided for comparison [17].

Gasoline Demand Under Reopening and No-Reopening Policies

With the reopening of many states in late April and early May 2020, the motor gasoline demand has clearly increased, even though the pandemic is far from over. Some researchers have warned that too early reopening in some states could result in more infections and deaths [26], while a continuation of shutdown measures, or postponing reopening, will mean continued reduction in mobility. To further explore the potential impact of reopening and no-reopening policies on motor gasoline demand, we create a hypothetical scenario (the “no-reopening” scenario) in which the reopening policy is postponed by four weeks (i.e., the reopening is not implemented until late May and early June 2020). Here, the travel mobility and motor gasoline demand on April 24 – May 21 (four weeks) is assumed to be postponed by four weeks to May 22 – June 18. Travel mobility and motor gasoline demand on April 24 – May 21 are assumed to be linearly extended from April 23. In addition, the study by Fowler et al. shows that, in the U.S., a strict stay-at-home order can help to a 30.2% reduction of confirmed infection cases after one week, a 40.0% reduction after two weeks, and a 48.6% reduction after three weeks [27]. These reduction rates are built into the no-reopening scenario in comparison to the scenarios with the YYG model under the facts with reopening [24]. In the no-reopening scenarios, the reciprocal effect of mobility on the evolution of the pandemic is not considered.

Figure 5 shows the differences in motor gasoline demand resulting from reopening and no-reopening policies under the reference, optimistic, and pessimistic scenarios. Under the reference scenario with a no-reopening policy, the four-week delay causes the motor gasoline demand to recover much slower than it would be with a reopening policy during this period. However, after the reopening in late May and early June, the rebound in gasoline demand is more prominent than it is in late April and early May. The cumulative motor gasoline demand during the period from April 23 to September 21, 2020, under both reopening and no-reopening policies are calculated in these three pairs of scenarios, as shown in Figure 5. Under the reference scenario, the cumulative motor gasoline demand with the no-reopening policy is approximately 6.4 million barrels more than the motor gasoline demand with the reopening policy. Under the optimistic scenario, the cumulative motor gasoline demand resulting from the no-reopening policy is about 40.0 million barrels less than that resulting from the reopening policy. Under the pessimistic scenario, the cumulative motor gasoline demand resulting from the no-reopening policy is about 109.5 million barrels more than that resulting from the reopening policy. In summary, under the reference and pessimistic scenarios, the no-reopening policy could worsen the motor gasoline demand in the short-term, but it might help the motor gasoline demand recover to a normal level sooner, and therefore the cumulative gasoline demand during this period would be higher. Comparatively, under the optimistic scenarios, the reopening policy could result in a quicker recovery of the motor gasoline demand in the medium-term.

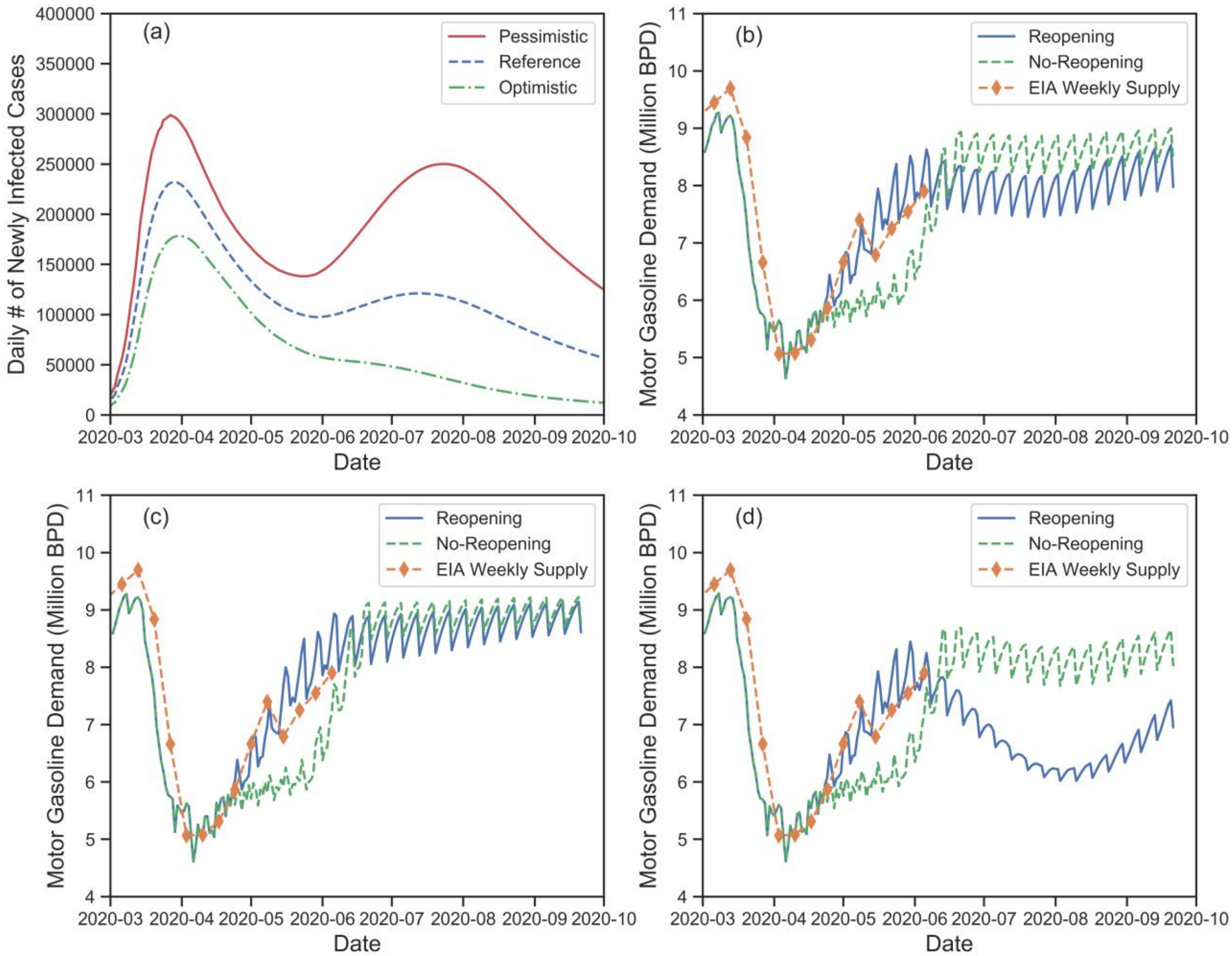

Figure 5. Projections of motor gasoline demands under reopening and no-reopening policies. (a) Three scenarios of daily number of newly infected cases projected by the YYG model [24]; (b) Gasoline demand projection comparing reopening and no-reopening policies under the reference YYG scenario; (c) Gasoline demand projection comparing reopening and no-reopening policies under the optimistic YYG scenario; (d) Gasoline demand projection comparing reopening and no-reopening policies under the pessimistic scenario.

Conclusions

A machine learning model, PODA, was developed to predict travel mobility in combination with motor gasoline demand. This model allows review of the mobility trends from Google and Apple, compares with the non-pandemic period, analyzes the historical weekly motor gasoline demand published by EIA, and projects the future motor gasoline demand. By projecting the motor gasoline demand under different pandemic and policy scenarios and combined with the pandemic models referenced by the CDC, this paper provides several insights of interest to stakeholders.

Under the current reference pandemic scenario, the growth of motor gasoline demand is slow from mid-May until August. It is unlikely that motor gasoline demand in the U.S. will reach a non-pandemic level in the medium-term (before October 2020). Under the optimistic pandemic scenario, the motor gasoline demand (using the PODA model with Google mobility as input) is expected to continue growing and will recover to about 98% of the demand under the non-pandemic scenario by late September 2020, which means the gasoline demand would be almost fully recover by then. Under the pessimistic infection scenario, the second wave of infections in mid-June to August could significantly lower the gasoline demand again, but it would not be worse than it was in April.

With a reopening policy, the mobility level in the states from the Midwest and South have rebounded near or back to a normal level already. However, under the reference pandemic scenario, if infections increase in these states, their mobility levels could still decrease by mid-June to early August 2020. Under the reference and pessimistic scenarios, a no-reopening policy could lower the motor gasoline demand temporarily, but it might also help demand increase to a normal level quicker. However, under the optimistic scenarios, a reopening policy recovers motor gasoline demand sooner.

The contribution of this study is the creation of a framework for investigating and projecting motor gasoline demand based on COVID-19 pandemic impacts, changes in mobility, and demographic information in each individual state. This study can contribute to evaluate the effects of pandemic on energy industry and economy, to outlook macroeconomics and social activities. However, the model has some limitations that should be addressed. First, the model assumes the gasoline demands from non-light-duty vehicles and other sectors, which accounts for 8% of the total gasoline consumption in the U.S., are constant during the pandemic [23,28]. Second, the model does not consider the dynamic impact of travel mobility on the evolution of the COVID-19 pandemic. As more is learned about both the pandemic model and the reciprocal effects of mobility on model results, the analysis and the model will be updated and improved.

The motor gasoline demand projections by the PODA model are updated and released regularly on a publicly available website (https://covid19-mobility.com). With appropriate modification, the PODA model will be applied to project other transportation-related fuel demand (such as for freight trucks) and to study fuel demand changes in other countries in the near future.

Methods

The Mobility Dynamic Index Forecast Module

For the Mobility Dynamic Index Forecast Module, we develop a method using the machine-learning neural network that uses pandemic data, policies, and demographic data as inputs to predict variations in mobility. The historical COVID-19 pandemic data, government policies, and demographic data are used as the model inputs, which are listed in Supplementary Note 2. These inputs are selected based on the correlation analysis between the pandemic and the mobility data [18–20], which are presented in Supplementary Figure 1. The neural network model has 5 outputs related to the Google and Apple mobilities. The complete list of data used as inputs and outputs in the Mobility Dynamic Index Forecast Module is provided in Supplementary Table 1.

Data preprocessing is applied to the datasets to accelerate training and improve the model performance: both the network inputs and outputs are centered by the mean values and normalized by the standard deviations. The loss function is specified as the mean square error of the standardized outputs. The neural network model has 42 inputs (listed in Supplementary Table 1 with descriptive statistics shown in Supplementary Figures 2-4), 2 hidden layers, and 25 hidden nodes for each hidden layer, with rectified linear units as the activation function. In addition, the mode is implemented with the deep learning framework of PyTorch [29]. The network is optimized using Adam [30], with a default learning rate of 1e-3 and a batch size of 64. To avoid potential overfitting, the L_2 regularization is applied with a weight decay factor of 1e-4. In addition, the rolling-window cross-validation is adopted to search the optimal network structure, which is detailed in Supplementary Note 5. Out-of-sample testing is also performed for the selected neural network structure to estimate the performance of the model in predicting future mobility. Specifically, the number of hidden layers and nodes are chosen to maximize the accuracy of both training and test datasets, the performance of different number of layers and nodes is compared in Supplementary Table 4. Normally, the network reaches a good performance after 5,000 epochs. The results of neural network model validation are provided in Supplementary Figures 10 and 11, and described by Supplementary Note 5.

The future evolution of COVID-19 pandemic is needed to project the future gasoline fuel demand. In this study, we adopt the pandemic projection of the YYG model and use the projection from the MIT model for comparison [8,24,25]. The YYG model parameters are constantly calibrated against newly reported data using machine learning; thus, they can capture the effect of state policies. In addition, this model predicts the number of newly infected people, newly recovered people, and new deaths for the coming three months. It should be clarified that the Mobility Dynamic Index Forecast Module uses the number of U.S. daily confirmed cases as an input rather than the daily newly infected cases, as the confirmed cases have more impacts on people's mobility decisions. However, the YYG model projects daily new infection cases, but not the daily new confirmed cases. And the confirmed cases is only a portion of the infection cases as a lot of asymptomatic cases are not tested. We then convert the number of newly infected cases to the number of newly confirmed cases by recognizing the following relationship:

$$N_{infected,d} = \eta * N_{confirmed,d+16} \tag{1}$$

where η is calculated so that it best fits the total U.S. confirmed cases with the historically projected total infected cases from the YYG model for the reference, lower, and upper scenarios, respectively. *d+16* means the confirmed cases value has a time delay of 16 days compared to the newly infected cases. Supplementary Figure 5 shows the calculated daily number of confirmed cases based on the newly infected cased projected by YYG on June 10, 2020.

The Motor Gasoline Demand Estimation Module

Travel mobility is fairly stable under normal conditions, and the estimated driving miles by trip activity for each individual state from the National Household Travel Survey (NHTS) 2017 are used as a benchmark [31]. Estimated miles of travel by trip activities from the NHTS are presented in Supplementary Table 2, and the calculation method is described by Supplementary Note 6. Normally, the gasoline demand from personal travel mobility is comparatively inelastic [7]. However, the motor gasoline demand decreases due to a decrease in out-of-home trips during the pandemic. Therefore, it is important to decompose personal out-of-home trips and connect them to the evolution of the pandemic for the Dynamic Mobility Index Forecast module. We denote dynamic mobility indices for people in state (*n*) on date (*d*) as a vector $Q_{n,d} \in \mathbb{R}^K$:

$$Q_{n,d} = \begin{bmatrix} q_{n,d}^1 \\ q_{n,d}^2 \\ \dots \\ q_{n,d}^K \end{bmatrix}, \qquad n \in \{1,2, \dots N\} \tag{2}$$

where *N* denotes the total number of states in the U.S. In the case where the historical mobility data in the PODA model are from Google, then *K* = 4, and the dynamic indices of mobility include "workplaces", "grocery and pharmacy," "retail and recreation," and "parks." In case where the historical mobility data are from Apple, then *K* = 1, and the dynamic index of mobility is "driving." This study uses the Google mobility data as an example. Due to the tracking errors of mobility (e.g., Google uses the locational and duration data instead of actual route distance data for quantifying the mobility trend) and the partial mismatching in definitions of trip activities between Google and NHTS. Using the Google data as an example, the indices of dynamic mobility include the activities: workplaces ($q_{n,d}^1$), grocery and pharmacy ($q_{n,d}^2$), retail and pharmacy ($q^3_{n,d}$), and parks ($q^4_{n,d}$). The changes of mobility by trip activity, based on NHTS trip activities and used for projecting the motor gasoline demand, include the activities: work ($m^1_{n,d}$), school/daycare/religious activity ($m^2_{n,d}$), medical/dental services ($m^3_{n,d}$), shopping/errands ($m^4_{n,d}$), recreational/social ($m^5_{n,d}$), transport someone ($m^6_{n,d}$), meals ($m^7_{n,d}$), and something else ($m^8_{n,d}$). The vector of the mobility adjustment factor $X = [x^1\ x^2\ \cdots\ x^9]$ must be used for better calibration and projection. The indices of dynamic mobility from Google - $q^k_{n,d}$ need adjustments with *X* before they are converted to $m^p_{n,d}$, which represents the mobility changes for trip activity (*p*) in state (*n*) on date (*d*). The changes in mobility by trip activity are presented in Eq. (3).

$$M_{n,d} = \begin{bmatrix} m^1_{n,d} \\ m^2_{n,d} \\ \cdots \\ m^P_{n,d} \end{bmatrix}, \qquad n \in \{1,2, \dots N\} \tag{3}$$

Through the mobility adjustment factor $X = [x^1\ x^2\ \cdots\ x^9]$, we can connect the changes of mobility ($M_{n,d}$) by trip activity and the indices of dynamic mobility ($Q_{n,d}$) based on their definitions of trip activities, respectively [18,31]. The conversion equations are shown in Eq. (4).

$$\begin{cases} m^1_{n,d} = q^1_{n,d} \cdot x^1 \\ m^2_{n,d} = q^1_{n,d} \cdot x^2 \\ m^3_{n,d} = q^2_{n,d} \cdot x^3 \\ m^4_{n,d} = q^2_{n,d} \cdot x^4 \\ m^5_{n,d} = \left[x^9 \cdot q^3_{n,d} + (1 - x^9) \cdot q^4_{n,d}\right] \cdot x^5 \\ m^6_{n,d} = q^3_{n,d} \cdot x^6 \\ m^7_{n,d} = q^3_{n,d} \cdot x^7 \\ m^8_{n,d} = q^3_{n,d} \cdot x^8 \end{cases} \tag{4}$$

Using the changes in mobility for trip activity (*p*) in state (*n*) on date (*d*) - $m^p_{n,d}$, we can estimate the miles traveled for different trip activities listed in Supplementary Table 2. They are calculated by Eq. (5).

$$t^p_{n,d} = S^p_n \cdot \left(1 + m^p_{n,d}\right), \qquad p \in \{1, 2, \dots, P\} \tag{5}$$

where *P* denotes the total number of NHTS trip activities; $t^p_{n,d}$ is the estimated miles traveled for an average household for trip activity (*p*) in state (*n*) on date (*d*); and S^p_n is the average miles traveled for each trip activity (*p*) in state (*n*) on normal days of a year. The baseline for mobility changes in Google is travel during the 5-week period from January 3 to February 6, 2020 [18].

Lastly, the estimated total miles traveled $T_{n,d}$ of people in state (*n*) on date (*d*) is shown in Eq. (6).

$$T_{n,d} = \sum_{p=1}^{P} t_{n,d}^{p} + C \tag{6}$$

where $\sum_{p=1}^{P} t_{n,d}^{p}$ denotes the total miles traveled by an average household in state (*n*) on date (*d*), and *C* is a constant that denotes the "miles" converted from non-household activities. Correspondingly, the motor gasoline demand in state (*n*) on date (*d*) should be roughly proportionally related to the total miles traveled by people in the same state on the same date, as shown in Eq. (7).

$$G_d = \sum_{n=1}^{N} \left(T_{n,d} \frac{G_B}{T} \cdot y_n \right) \tag{7}$$

where G_d denotes the nationwide motor gasoline demand in the U.S. on date (*d*); G_B is the medium value of motor gasoline demand in the U.S. during the 5-week period from January 3 to February 6, 2020; and *T* denotes the total miles traveled by people on normal days. This period is used as a benchmark for mobility change by Google [18]. Finally, y_n denotes the on-road demand share in state (*n*) in the U.S, which is shown in Supplementary Table 3.

We solve Eq. (8) to calibrate the values of the adjustment factor vector X and constant C.

$$\min \sum \left(\left| G_{EIA,W} - \overline{G_d} \right| / G_{EIA,W} \right)^2, \qquad s.t. X, C \in \mathbb{R}^+ \tag{8}$$

where Eq. (8) is used to minimize the sum of squared absolute percentage error. $G_{EIA,W}$ denotes the weekly motor gasoline demand in the U.S., which is published by the U.S. EIA every week. $\overline{G_d}$ is the average weekly value of daily estimated motor gasoline demand (G_d).

EIA provides values of weekly finished motor gasoline supplies, which are used to estimate gasoline demand in the U.S. [32]. For this study, we use this data to calibrate in the Motor Gasoline Demand Estimation Module. A quadratic equation is adopted to fit the seasonality of weekly motor gasoline demands in 2016–2019, as discussed in Supplementary Note 4, and it will be used as a "normal days" benchmark to compare with the motor gasoline demand in 2020 under the pandemic. In addition, the information on the shares of motor gasoline consumed by the transportation sector by state in 2017 and 2018 are presented in Supplementary Table 3. The state shares of motor gasoline consumption are pretty stable in 2017-2018 [33], and therefore it assumes that the state shares of motor gasoline consumption before the pandemic remain the same in 2020. What's more, the transportation sector accounts for 94%-98% of motor gasoline consumption [33]. Therefore, this study does not consider the motor gasoline consumed by the commercial and industrial sectors as the inputs in the model but calibrates the potential differences as a “constant” in the model.

Reference

1. Sohrabi, C. *et al.* World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *Int. J. Surg.* **76**, 71–76 (2020).

2. World Health Organization. WHO announces COVID-19 outbreak a pandemic. *WHO* (2020). Available at: http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic. (Accessed: 12th May 2020)

3. Johns Hopkins University. COVID-19 Dashboard by the Center for Systems Science and Engineering at Johns Hopkins University. *JUH* (2020). Available at: https://coronavirus.jhu.edu/map.html. (Accessed: 23rd June 2020)

4. Bendavid, E. *et al.* COVID-19 antibody seroprevalence in Santa Clara county, California. *medRxiv* 2020.04.14.20062463 (2020). doi:10.1101/2020.04.14.20062463

5. Egan, M. Oil prices turned negative. Hundreds of US oil companies could go bankrupt. *CNN* (2020). Available at: https://www.cnn.com/2020/04/20/business/oil-price-crash-bankruptcy/index.html. (Accessed: 12th May 2020)

6. Metz, T. *Coronavirus and energy, a sector challenged by geographical concentrations (working paper)*. (2020).

7. Kilian, L. The economic effects of energy price shocks. *J. Econ. Lit.* **46**, 871–909 (2008).

8. U.S. Centers for Disease Control and Prevention. COVID-19 forecasts. *U.S. CDC* (2020). Available at: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html#why-forecasting-critical. (Accessed: 12th June 2020)

9. Duque, D. *et al.* COVID-19: How to relax social distancing if you must. *medRxiv* (2020). doi:10.1101/2020.04.29.20085134

10. Taghizadeh-Hesary, F., Yoshino, N. & Rasoulinezhad, E. Impact of the Fukushima nuclear disaster on the oil-consuming sectors of Japan. *J. Comp. Asian Dev.* **16**, 113–134 (2017).

11. Doytch, N. & Klein, Y. L. The impact of natural disasters on energy consumption: An analysis of renewable and nonrenewable energy demand in the residential and industrial sectors. *Environ. Prog. Sustain. Energy* **37**, 37–45 (2018).

12. Kuo, H.-I., Chen, C.-C., Tseng, W.-C., Ju, L.-F. & Huang, B.-W. Assessing impacts of SARS and Avian Flu on international tourism demand to Asia. *Tour. Manag.* **29**, 917–928 (2008).

13. Meltzer, M. I., Cox, N. J. & Fukuda, K. The economic impact of pandemic influenza in the United States: priorities for intervention. *Emerg. Infect. Dis.* **5**, 659–671 (1999).

14. U.S. Energy Information Administration. *Short-term energy outlook*. (2020).

15. Swales, V., Lu, D. & Mervosh, S. See which states and cities have told residents to stay at home. *New York Times* (2020). Available at: https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html. (Accessed:

13th May 2020)

16. U.S. Department of Transportation Federal Highway Administration. March 2020 Traffic Volume Trends. (2020). Available at: https://www.fhwa.dot.gov/policyinformation/travel_monitoring/20martvt/. (Accessed: 12th June 2020)

17. U.S. Energy Information Administration. Weekly Petroleum Status Report. *U.S. EIA* (2020). Available at: https://www.eia.gov/petroleum/supply/weekly/. (Accessed: 12th June 2020)

18. Google. COVID-19 Community Mobility Reports. (2020). Available at: https://www.google.com/covid19/mobility/. (Accessed: 12th May 2020)

19. Apple. Mobility Trends Reports. (2020). Available at: https://www.apple.com/covid19/mobility. (Accessed: 12th May 2020)

20. Maryland Transportation Institute. University of Maryland COVID-19 Impact Analysis Platform. (2020). Available at: https://data.covid.umd.edu. (Accessed: 12th May 2020)

21. OurWorldData. Daily confirmed COVID-19 deaths per million, rolling 7-day average. *OurWorldData* (2020). Available at: https://ourworldindata.org/grapher/daily-covid-deaths-per-million-7-day-average. (Accessed: 13th May 2020)

22. U.S. Centers for Disease Control and Prevention. Healthcare Facilities: Preparing for Community Transmission. *U.S. CDC* (2020). Available at: https://www.cdc.gov/coronavirus/2019-ncov/hcp/guidance-hcf.html. (Accessed: 13th May 2020)

23. U.S. Energy Information Administration. Gasoline is the main U.S. transportation fuel. (2020). Available at: https://www.eia.gov/energyexplained/gasoline/use-of-gasoline.php. (Accessed: 12th June 2020)

24. Gu, Y. COVID-19 Projections Using Machine Learning. *covid19-projections.com* (2020). Available at: https://covid19-projections.com/about/. (Accessed: 12th June 2020)

25. Bertsimas, D. DELPHI Epidemiological Case Predictions. *MIT* (2020). Available at: https://www.covidanalytics.io/projections. (Accessed: 15th June 2020)

26. Yamana, T., Pei, S., Kandula, S. & Shaman, J. Projection of COVID-19 Cases and Deaths in the US as Individual States Re-open May 4,2020. *medRxiv* 2020.05.04.20090670 (2020). doi:10.1101/2020.05.04.20090670

27. Fowler, J. H., Hill, S. J., Obradovich, N. & Levin, R. The Effect of Stay-at-Home Orders on COVID-19 Cases and Fatalities in the United States. *medRxiv* (2020). doi:10.1101/2020.04.13.20063628

28. U.S. Energy Information Administration. *Annual Energy Outlook 2020*. (2020). doi:DOE/EIA-0383(2020)

29. Paszke, A. *et al.* PyTorch: An Imperative Style, High-Performance Deep Learning Library. in *Advances in Neural Information Processing Systems 32* (eds. Wallach, H. et al.) 8026–8037 (Curran Associates, Inc., 2019).

30. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv Prepr. arXiv1412.6980*

(2014).

31. U.S. Department of Transportation Federal Highway Administration. 2017 National Household Travel Survey. Available at: http://nhts.ornl.gov.

32. U.S. Energy Information Administration. Petroleum & Other Liquids - Weekly Supply Estimates. *U.S. EIA* (2020). Available at: https://www.eia.gov/dnav/pet/pet_sum_sndw_a_epm0f_vpp_mbblpd_w.htm. (Accessed: 13th May 2020)

33. U.S. Energy Information Administration. State Energy Data System (SEDS). *U.S. Department of Energy, Energy Information Administration* (2020). Available at: https://www.eia.gov/state/seds. (Accessed: 14th May 2020)

Acknowledgement

This research was financially supported by Aramco Services Company and used resources at the National Transportation Research Center at Oak Ridge National Laboratory and Systems Assessment Center of Energy Systems Division at Argonne National Laboratory. This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 and UChicago Argonne, LLC under Contract DE-AC02-06CH11357 with the U.S. Department of Energy. The authors would like to thank the developer of the YYG model, Mr. Youyang Gu, for his feedback on the COVID-19 projections used in this study. The authors are solely responsible for the views expressed in this study.

Author Contributions

PODA model concept and methodologies were structured by S.O., X.H., W.J., and W.C. Python codes were developed by X.H. and W.J. Data was collected by S.O., X.H., W.J., Z.L., L.S., and Y.G. Data Analyses were performed by S.O., X.H., L.S., W.J., W.C., J.B., Z.L., and S.D. The manuscript was written by the combination of S.O., L.S., S.P., X.H., W.J., and Z. L.

Data Availability

The data used for model development are either provided in Supplementary Information or publicly available at:

(1) EIA weekly gasoline demand: https://www.eia.gov/dnav/pet/pet_cons_wpsup_k_w.htm
(2) U.S. pandemic data: https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/
(3) YYG pandemic projection model: https://github.com/youyanggu/covid19_projections
(4) MIT pandemic projection model: https://github.com/COVIDAnalytics/DELPHI
(5) Apple Mobility: https://www.apple.com/covid19/mobility
(6) Google Mobility: https://www.google.com/covid19/mobility/
(7) Government policy: https://github.com/COVID19StatePolicy/SocialDistancing/blob/master/data/USstatesCov19distancingpolicy.csv
(8) National Household Travel Survey 2017: https://nhts.ornl.gov/
(9) EIA State Energy Data System: https://www.eia.gov/state/seds/

The results that support the findings of this study are provided in the main text and Supplementary Information. The source data underlying all figures in the main manuscript and

Supplementary Information are provided as a Source Data file. With the evolution of the pandemic, the U.S. motor gasoline demand and mobility level projection is updated periodically and posted at https://covid19-mobility.com and https://teem.ornl.gov/poda.shtml.

Code Availability

The code of the PODA model is deposited and managed on GitHub (https://github.com/jiweiqi/covid19-mobility).

Competing interests

We declare that none of the authors have competing financial or non-financial interests.

Figure Legends

Figure 1. Changes in mobility in various sectors and number of cases. Percentage changes in Google mobility between February 15 and June 4, 2020 for (a) workplaces; (b) grocery and pharmacy; (c) retail and recreations [18]. (d) 7-day average of the number of COVID-19 cases and deaths in the U.S. [3] Please note the boxplots represent the interquartile range (IQR) where the lower and upper whiskers represent 1.5×IQR. Each yellow data point represents the mobility of outlier states that do not fall into the 1.5xIQR.

Figure 2. The structure of the PODA model.

Figure 3. The projection of gasoline demand and mobility level in the U.S. under the reference pandemic scenario. (a) Comparison of gasoline demand projections based on the Apple mobility data and Google mobility data [18,19], EIA weekly motor gasoline supply and non-pandemic gasoline demands are provided for comparison [17]; (b) State-level mobility projections. "Apple" means the inputs for projection are from Apple, and "Google" means the inputs for projection are from Google. Mobility level at 0% in the legend bar for Apple mobility refers to the baseline on January 13, 2020 [19]. Mobility level at 0% in the legend bar for Google mobility refers to the baseline of the median value for the corresponding days of the weeks from January 3 to February 6, 2020 [18].

Figure 4. Gasoline demand under different scenarios (projected by June 10, 2020). Gasoline demand projections are based on the three scenarios of the YYG model and MIT model [24,25]. EIA weekly motor gasoline supply and non-pandemic gasoline demands are provided for comparison [17] .

Figure 5. Projections of motor gasoline demands under reopening and no-reopening policies. (a) Three scenarios of daily number of newly infected cases projected by the YYG model [24]; (b) Gasoline demand projection comparing reopening and no-reopening policies under the reference YYG scenario; (c) Gasoline demand projection comparing reopening and no-reopening policies under the optimistic YYG scenario; (d) Gasoline demand projection comparing reopening and no-reopening policies under the pessimistic scenario.