



Machine learning model to project the impact of COVID-19 on US motor gasoline demand

Shiqi Ou¹, Xin He²✉, Weiqi Ji³, Wei Chen⁴, Lang Sui², Yu Gan², Zifeng Lu², Zhenhong Lin¹, Sili Deng¹, Steven Przesmitzki¹ and Jessey Bouchard²

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.

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, which makes it the country with the largest number of confirmed cases in the world so far³. Sood et al. implies that the actual infections could be more widespread than indicated by the number of confirmed cases⁴. 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 United States is one of the largest energy consumers and oil producers in the world. Fluctuations and uncertainties in US oil consumption impact the petroleum supply chain and trends of the broader energy economy⁷. Beyond temporary and local oil price and demand shocks, US gasoline demand alone is substantial enough to impact longer-term investments in the global energy industry, which impacts the world economy as a whole.

Although the value of oil has somewhat recovered since April, uncertainties in the US economy persist because of the lingering pandemic. The International Monetary Fund estimated the annual change in real gross domestic product in the United States could be -8.0% in 2020 and the global economy could contract by 4.9%⁸. To effectively reduce uncertainties, multiple models were developed by researchers to project the trend of the COVID-19 pandemic⁹. Projections of the evolution of COVID-19 pandemic trends show that lockdowns help to reduce COVID-19 transmissions by as much as 90% compared with the baseline without any social distancing in Austin, Texas¹⁰. However, this unprecedented phenomenon could last for a few years: Kissler et al. suggested that, even after the pandemic peaked, COVID-19 surveillance should be continued as a resurgence in contagion could be possible as late as 2024⁹. Therefore, beyond the immediate economic responses, the longer-term impact on the US economy may persist well beyond 2020. An effective

forecast or estimate of the pandemic impacts could help people to well prepare and navigate around unknown risks. More specifically, reliably projecting the oil demand, a critical leading indicator of the state of the US economy, is beneficial to related business activities and investment decisions.

There are studies that discuss the impacts of unexpected natural hazards and/or disasters on energy demand and/or consumption^{11,12}, and studies that evaluate the impacts of previous pandemics on tourism¹³ and economics¹⁴. However, few studies have quantified and forecast the oil demands under multiple pandemic scenarios, and this research is desperately needed. To date, studies focused on the energy impacts of the COVID-19 pandemic are limited to the short-term energy outlook released by the US Energy Information Administration (EIA); this outlook uses a simplified evolution of the COVID-19 pandemic to forecast the US gross domestic product, energy supplies, demands and prices until the fourth quarter of 2021¹⁵. 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, government policies and demographic information.

In this study, we extend the understanding of the COVID-19 impact on medium-term fuel demand as the pandemic evolves. This will be useful to both policymakers 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 US is slow after a quick rebound in May 2020, 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 lock-down (no-reopening) policy could worsen the motor gasoline demand temporarily, but it helps the demand recover to a normal level more quickly 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

¹Energy and Transportation Science Division, Oak Ridge National Laboratory, Knoxville, TN, USA. ²Aramco Services Company: Aramco Research Center—Detroit, Novi, MI, USA. ³Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. ⁴Michigan Department of Transportation, Southfield, MI, USA. ⁵Energy Systems Division, Argonne National Laboratory, Lemont, IL, USA. ✉e-mail: xin.he@aramcoamericas.com

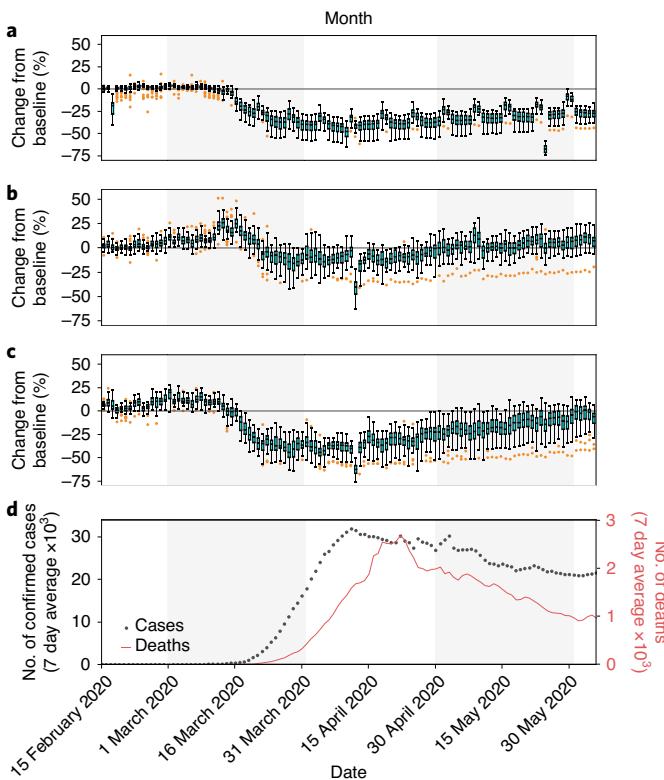


Fig. 1 | Changes in mobility in various sectors and number of cases.

a–c, Percentage changes in Google mobility between 15 February and 4 June 2020 for workplaces (**a**), grocery and pharmacy (**b**) and retail and recreation (**c**)¹⁹. **d**, Seven-day average of the number of COVID-19 confirmed cases and deaths in the United States³. The box plots represent the interquartile range (IQR) and the lower and upper whiskers represent $1.5 \times \text{IQR}$. Each yellow data point represents the mobility of outlier states that do not fall into the $1.5 \times \text{IQR}$.

substantially lower the gasoline demand once more, but it will not be worse than it was in April 2020. These results imply that government intervention does impact the 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 2020, at least 316 million people in the United States lived in regions with some form of stay-at-home policy¹⁶. US Department of Transportation data show that 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 2019¹⁷. Subsequently, US on-road transportation fuel consumption in April 2020 was 30% lower than it was in April 2019¹⁸.

The widespread availability of personal mobile devices has allowed us to quantitatively measure people’s confinement. At the onset of the pandemic, Google and Apple released mobility reports or webpages that used aggregated map data to help better quantify the correlation between the reduction in mobility and the spread of COVID-19^{19,20}. Figure 1a–c shows changes in mobility in the United States reported by Google for the period between 15 February and 4 June 2020 for work, retail and recreation, and for grocery and pharmacy¹⁹. The US shutdown policies are governed at the state or more localized levels instead of at a federal level, and there is considerable variation of mobility decrease among states. Although all three categories experienced a widespread decrease in mobility from 26

March 2020, the onset of many stay-at-home orders, the decrease is heterogeneous across the states and categories. Figure 1d shows the seven-day average for daily COVID-19 cases and deaths during the same period²¹. Although the daily number of cases of COVID-19 increased between 16 March and 20 April 2020, during a period of which much of the country was shut down, this increase is primarily due to the long incubation period of COVID-19²². Likewise, an increase in the daily number of deaths occurred roughly two weeks after confirmed cases. The data clearly show 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, and then mobility plateaus and becomes insensitive to further increases in confirmed cases or deaths. We conducted a detailed correlation analysis between mobility and pandemic data to identify the key parameters to consider in the model development. The results of the correlation analysis are presented in Supplementary Fig. 1 and Supplementary Note 1.

As the change in mobility implies a change in vehicle miles travelled, the mobility data provide a solid foundation to analyse 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.

Pandemic oil demand analysis model

Motor gasoline is the main US transportation fuel as it accounts for about 58% of the total energy consumption by the transportation sector and 45% of the total petroleum consumption²³. In this work, a machine-learning-based model of pandemic oil demand analysis (PODA) was developed to project the US gasoline demand using COVID-19 pandemic data, government policies and demographic information. As shown in Fig. 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 the COVID-19 pandemic and government orders, and it projects the changes in travel mobility indices relative to the pre-COVID-19 period in the United States. 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 grey dashed line shows in Fig. 2. The Motor Gasoline Demand Estimation Module estimates vehicle miles travelled on pandemic days while it considers the dynamic indices of travel mobility, and it quantifies motor gasoline demands by coupling the gasoline demands and vehicle miles travelled. 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 Methods 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.

The pandemic data used to project future motor gasoline demand is taken from the Youyang Gu (YYG) COVID-19 projection model (<https://covid19-projections.com/about/>), which is one of few models referenced by the US Centers for Disease Control and Prevention that offers a medium-term (three-month) forecast of infections and deaths for each state under multiple scenarios⁹. The model is updated daily to provide 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 divided our study into three scenarios that aligned with these projections: reference (mean), optimistic (lower) and pessimistic (upper). In addition, the pandemic data forecast by the Massachusetts Institute of Technology (MIT) model^{9,24} was used to project the gasoline demand, which was

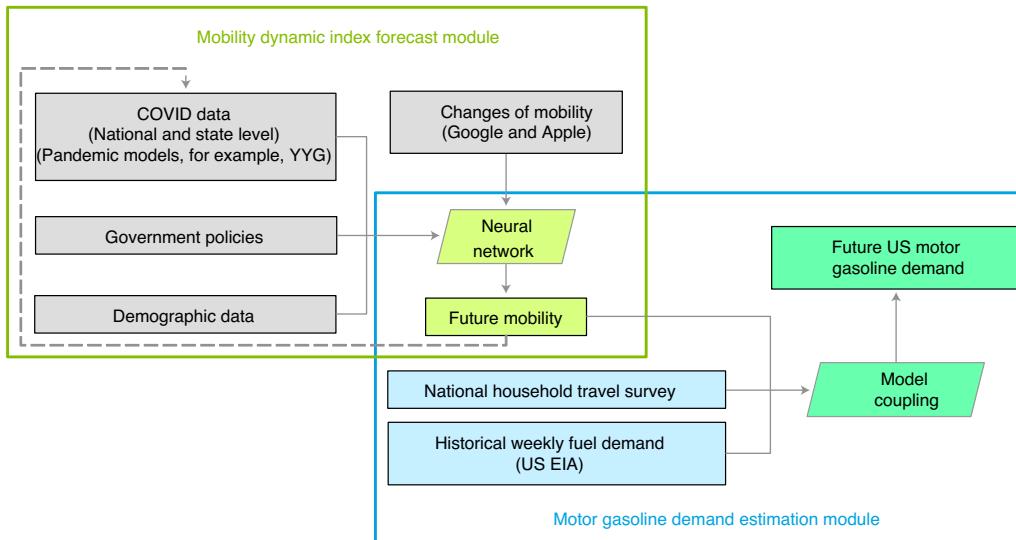


Fig. 2 | The structure of the PODA model. The PODA model is a machine-learning-based model to project the US gasoline demand using COVID-19 pandemic data, government policies and demographic information. The Mobility Dynamic Index Forecast Module identifies the changes in travel mobility caused by the evolution of the COVID-19 pandemic and government orders. The Motor Gasoline Demand Estimation Module quantifies motor gasoline demands due to the changes in travel mobility.

compared against the projection based on inputs from the YYG model. The details of these scenarios are provided in Supplementary Note 3 and Supplementary Figs. 5–7. We also provide the fuel demand of the non-pandemic scenario for comparison. For this scenario, we used the four-year average from 2016 to 2019 to reflect the seasonal change in gasoline demand (see Supplementary Figs. 8 and 9 for details).

Projections of motor gasoline demand under COVID-19 pandemic

Figure 3 shows US motor gasoline demand projections and state-level mobility projections up to 21 September 2020 under the reference pandemic scenario. The historical mobility and gasoline demand data for simulation in the model are from 1 March to 5 June 2020, and the pandemic data are from the YYG model projection as of 10 June 2020. Shown in Fig. 3, the projected motor gasoline demands from the PODA model fit well with the historical motor gasoline demand from EIA as of 5 June 2020. As the reopening was successively implemented by states, the US motor gasoline demand gradually recovered from its low point during the first weeks of April. By the week of 5 June 2020, the gasoline demand reached 7.9 million barrels per day (BPD), growing by nearly 56% compared with the April 3 low point¹⁸.

This study adopts the mobility trend data from Google¹⁹ into the PODA model to project future motor gasoline demand, and uses the mobility data from Apple²⁰ as input for the 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 Fig. 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, although the mobility trends from Google and Apple are estimated using different methodologies—Google estimates mobility using location and visit duration data¹⁹, and Apple uses route map requests²⁰—these collection methods have a relatively limited effect on the projected motor gasoline demand. However, the Apple mobility data offer only the aggregated changes of mobility, which do not differentiate mobility among detailed trip activities and/or 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 the projection and scenario discussion.

The PODA model can also be used to project the mobility level. In addition to motor gasoline demand, Fig. 3 also exhibits the projected mobility levels in each individual state from March to September 2020. Clearly, the mobility levels in all the 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 started to reopen, the mobility levels in most states increased dramatically. This was especially so 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 Fig. 4a—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.

Dynamics of future motor gasoline demand

The knowledge deficit of this disease magnifies the uncertainty of the projections on motor gasoline demand. Figure 5 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 Fig. 5 are all given as seven-day moving averages. The scenario details are described in Supplementary Note 3. Analysis of the medium-term projected motor gasoline demand to October 2020 under the reference scenario shows that, although the motor gasoline demand will grow rapidly after the April 2020 recovery, it is expected to remain fairly flat from mid-May until October 2020. The projected motor gasoline demand by the end of September 2020 will be about 8.342 million BPD. For comparison, 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%

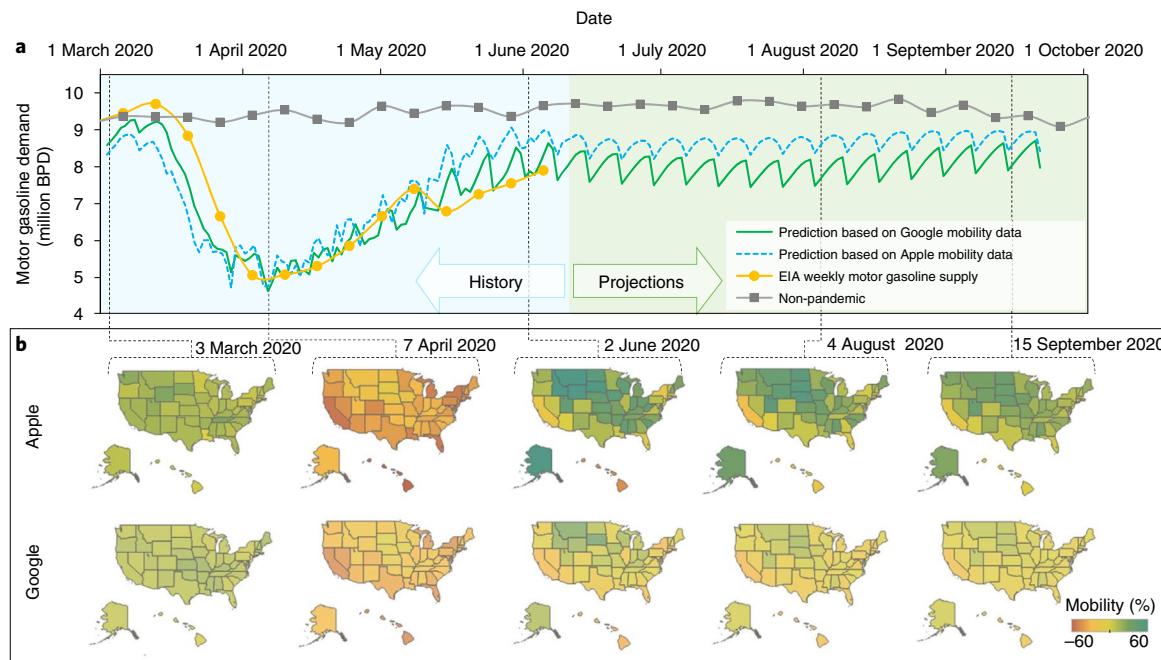


Fig. 3 | The projection of gasoline demand and mobility level in the United States under the reference pandemic scenario. **a**, Comparison of gasoline demand projections based on the Apple mobility data and Google mobility data^{19,20}. EIA weekly motor gasoline supply and non-pandemic gasoline demands are provided for comparison¹⁸. **b**, State-level mobility projections. The inputs for projection are from Apple and Google. The mobility level at 0% in Apple mobility refers to the baseline on 13 January 2020²⁰. The mobility level at 0% for Google mobility refers to the baseline of the median value for the corresponding days of the weeks from 3 January to 6 February 2020¹⁹.

of the demand under the non-pandemic scenario on 8 May 2020, and it is about 82% on 5 June 2020¹⁸. Therefore, the motor gasoline demand in the United States 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 2020, 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 2020 continues to rapidly grow until October 2020. 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 US 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 2020 declines to about 6.1 million BPD before it starts to recover by mid-August 2020, but the demand will not be worse than it was in April 2020. 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 10 June 2020 are presented in this work, the PODA model is updated regularly with the evolution of the pandemic (<https://covid19-mobility.com> and <https://teem.ornl.gov/poda.shtml>).

This study is also able to project the motor gasoline demands with pandemic inputs from other pandemic projection models. Given the forecast period and the availability of model source codes, we chose

the pandemic projections from the MIT model as an example^{9,24}. The purple curve in Fig. 5 shows the projected gasoline demands in the United States 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 (<https://covid19-projections.com/about>)²⁴, as shown in Supplementary Fig. 5.

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 clearly increased, even though the pandemic was far from over. Some researchers have warned that too early a reopening in some states could result in more infections and deaths²⁵, but a continuation of shutdown measures, or postponing reopening, will mean a continued reduction in mobility. To further explore the potential impact of reopening and no-reopening policies on motor gasoline demand, we created a hypothetical scenario (the 'no-reopening' scenario) in which the reopening policy is postponed by four weeks (that is, the reopening is not implemented until late May and early June 2020). Here, the travel mobility and motor gasoline demand from 24 April to 21 May 2020 (four weeks) was assumed to be postponed by four weeks to 22 May to 18 June 2020. Travel mobility and motor gasoline demand from 24 April to 21 May 2020 were assumed to be linearly extended from 23 April 2020. In addition, the study by Fowler et al. shows that, in the United States, a strict stay-at-home order can help to give 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 weeks²⁶. These reduction rates were built into the no-reopening scenario in comparison with the scenarios with the YYG model under the facts with reopening (<https://covid19-projections.com/about>). In the no-reopening scenarios, the reciprocal effect of mobility on the evolution of the pandemic was not considered.

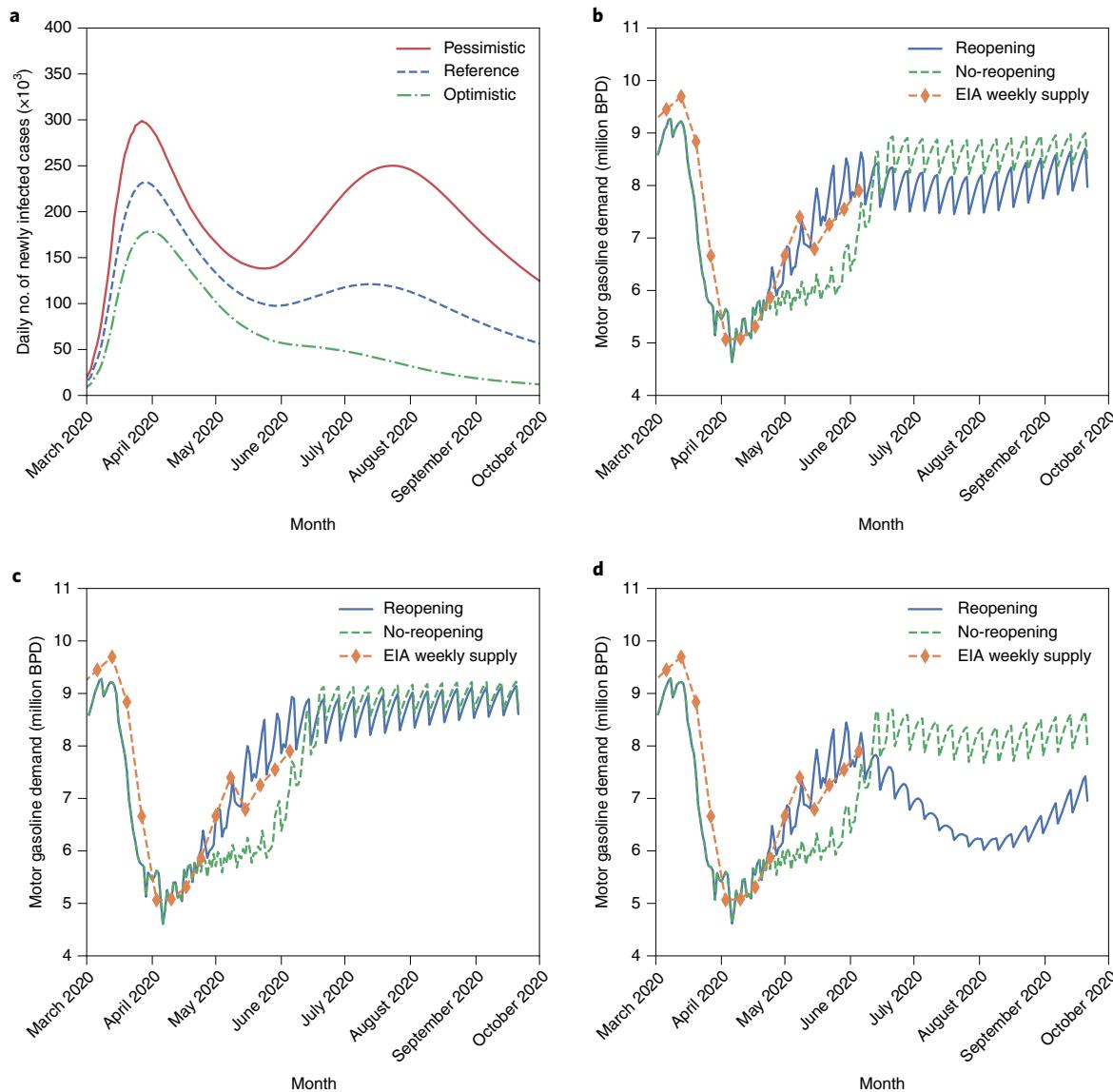


Fig. 4 | Projections of motor gasoline demands under reopening and no-reopening policies. **a**, Three scenarios of daily number of newly infected cases projected by the YYG model (<https://covid19-projections.com/about>). **b**, Gasoline demand projection that compares reopening and no-reopening policies under the reference YYG scenario. **c**, Gasoline demand projection that compares reopening and no-reopening policies under the optimistic YYG scenario; **d**, Gasoline demand projection that compares reopening and no-reopening policies under the pessimistic scenario.

Figure 4 shows the differences in motor gasoline demand that result from the 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 more slowly than it would with a reopening policy during this period. However, after the reopening in late May and early June 2020, the rebound in gasoline demand is more prominent than that in late April and early May 2020. The cumulative motor gasoline demands during the period from 23 April to 21 September 2020 under both reopening and no-reopening policies were calculated in these three pairs of scenarios, as shown in Fig. 4. 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 that results from the no-reopening policy is about 40.0 million barrels less than that from the reopening policy. Under the pessimistic scenario, the cumulative motor gasoline demand that

results from the no-reopening policy is about 109.5 million barrels more than that 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.

Conclusions

A machine-learning model, PODA, was developed to predict travel mobility in combination with motor gasoline demand. This model allows a review of the mobility trends from Google and Apple, compares these with the non-pandemic period, analyses 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

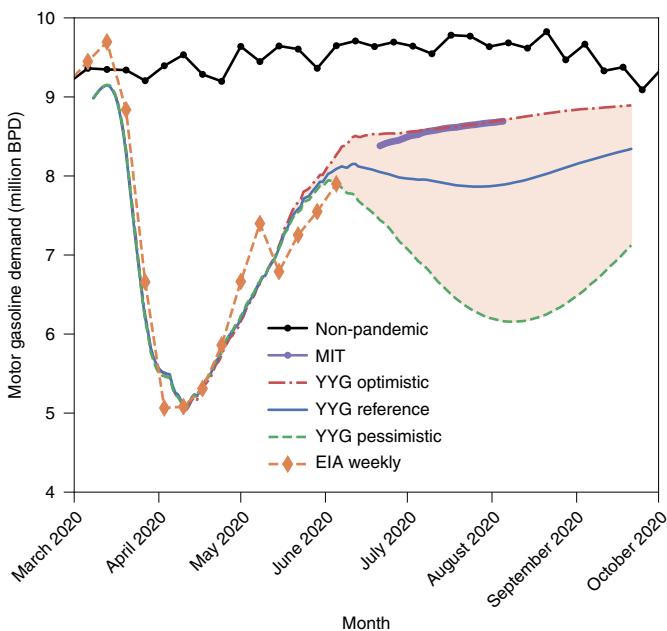


Fig. 5 | Gasoline demand under different scenarios (projected by 10 June 2020). Gasoline demand projections are based on the three scenarios of the YYG model and that of the MIT model (<https://covid19-projections.com/about>)²⁴. EIA weekly motor gasoline supply and non-pandemic gasoline demands are provided for comparison¹⁸.

combined with the pandemic models referenced by the US Centers for Disease Control and Prevention, 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 2020. It is unlikely that motor gasoline demand in the United States 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 the 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 recovered by then. Under the pessimistic infection scenario, the second wave of infections in mid-June to August 2020 could substantially lower the gasoline demand again, but it would not be worse than it was in April 2020.

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 more quickly. However, under the optimistic scenarios, a reopening policy recovers the motor gasoline demand sooner.

The contribution of this study is the creation of a framework to investigate and project 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 the pandemic on the energy industry and economy, on outlook macroeconomics and on social activities. However, the model has some limitations that should be addressed. First, it assumes the gasoline demands from non-light-duty vehicles and other sectors, which account for 8% of the total gasoline

consumption in the United States, are constant during the pandemic^{23,27}. 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 publicly available websites (<https://covid19-mobility.com> and <https://teem.ornl.gov/poda.shtml>). 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 developed a method with a 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 a correlation analysis between the pandemic and mobility data^{19,20}, which are presented in Supplementary Fig. 1. The neural network model has five 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 centred 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 Figs. 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 PyTorch²⁸. The network is optimized using Adam²⁹, with a default learning rate of 1×10^{-3} and a batch size of 64. To avoid potential overfitting, the L₂ regularization is applied with a weight decay factor of 1×10^{-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 was chosen to maximize the accuracy of both training and test datasets, and the performance of different numbers of layers and nodes is compared in Supplementary Table 4. Normally, the network reaches a good performance after 5,000 epochs. The results of the neural network model validation are provided in Supplementary Figs. 10 and 11, and described in Supplementary Note 5.

The future evolution of the COVID-19 pandemic is needed to project the future gasoline fuel demand. In this study, we adopted the pandemic projection of the YYG model and used the projection from the MIT model for comparison (<https://covid19-projections.com/about>)^{9,24}. 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 US daily confirmed cases as an input rather than the daily newly infected cases, as the confirmed cases have more impact 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 many 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:

$$N_{\text{infected},d} = \eta \times N_{\text{confirmed},d+16} \quad (1)$$

where η is calculated so that it best fits the total US confirmed cases with the historically projected total infected cases from the YYG model for the reference, lower and upper scenarios. The term $d+16$ means the confirmed cases value has a time delay of 16 days compared with the newly infected cases. Supplementary Fig. 5 shows the calculated daily number of confirmed cases based on the newly infected cases projected by YYG on 10 June 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 2017 National Household Travel Survey (NHTS) are used as a benchmark³⁰. Estimated miles of travel by trip activities from the NHTS are presented in Supplementary Table 2, and the calculation method is

described in Supplementary Note 6. Normally, the gasoline demand from personal travel mobility is comparatively inelastic⁷. However, the motor gasoline demand decreased 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 $\mathbf{Q}_{n,d} \in R^K$:

$$\mathbf{Q}_{n,d} = \begin{bmatrix} q_{n,d}^1 \\ q_{n,d}^2 \\ \dots \\ q_{n,d}^K \end{bmatrix}, n \in \{1, 2, \dots, N\} \quad (2)$$

where N denotes the total number of states in the United States. 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 the case where the historical mobility data are from Apple, then $K=1$, and the dynamic index of mobility is 'driving'. Owing to the tracking errors of mobility (for example, Google uses the locational and duration data instead of actual route distance data to quantify the mobility trend) and the partial mismatching in definitions of trip activities between Google and NHTS, mobility adjustment factors must be used for a better calibration and projection. 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 ($q_{n,d}^3$) and parks ($q_{n,d}^4$). The changes of mobility by trip activity, based on NHTS trip activities and used to project the motor gasoline demand, include the activities work ($m_{n,d}^1$), school, day care and religious activity ($m_{n,d}^2$), medical and dental services ($m_{n,d}^3$), shopping and errands ($m_{n,d}^4$), recreational and social ($m_{n,d}^5$), transport someone ($m_{n,d}^6$), meals ($m_{n,d}^7$) and something else ($m_{n,d}^8$). The indices of dynamic mobility from Google, $q_{n,d}^k$, need adjustments with $\mathbf{X} = [x^1, x^2, \dots, x^9]$ before they are converted to $m_{n,d}^p$, 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 equation (3):

$$\mathbf{M}_{n,d} = \begin{bmatrix} m_{n,d}^1 \\ m_{n,d}^2 \\ \dots \\ m_{n,d}^P \end{bmatrix}, n \in \{1, 2, \dots, N\} \quad (3)$$

Through the mobility adjustment factor $\mathbf{X} = [x^1, x^2, \dots, x^9]$, we can connect the changes of mobility ($\mathbf{M}_{n,d}$) by trip activity and the indices of dynamic mobility ($\mathbf{Q}_{n,d}$) based on their definitions of trip activities, respectively^{19,30}. The conversion equations are shown in equation (4):

$$\left\{ \begin{array}{l} m_{n,d}^1 = q_{n,d}^1 \times x^1 \\ m_{n,d}^2 = q_{n,d}^1 \times x^2 \\ m_{n,d}^3 = q_{n,d}^2 \times x^3 \\ m_{n,d}^4 = q_{n,d}^2 \times x^4 \\ m_{n,d}^5 = [x^9 \times q_{n,d}^3 + (1 - x^9) \times q_{n,d}^4] \times x^5 \\ m_{n,d}^6 = q_{n,d}^3 \times x^6 \\ m_{n,d}^7 = q_{n,d}^3 \times x^7 \\ m_{n,d}^8 = q_{n,d}^3 \times x^8 \end{array} \right. \quad (4)$$

Using the changes in mobility for trip activity (p) in state (n) on date (d), $m_{n,d}^p$, we can estimate the miles travelled for different trip activities listed in Supplementary Table 2. They are calculated by equation (5):

$$t_{n,d}^p = S_n^p \times (1 + m_{n,d}^p), p \in \{1, 2, \dots, P\} \quad (5)$$

where P denotes the total number of NHTS trip activities, $t_{n,d}^p$ is the estimated miles travelled for an average household for trip activity (p) in state (n) on date (d) and S_n^p is the average miles travelled 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 five-week period from 3 January to 6 February 2020¹⁹.

Lastly, the estimated total miles travelled $T_{n,d}$ of people in state (n) on date (d) is shown in equation (6):

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

where $\sum_{p=1}^P t_{n,d}^p$ denotes the total miles travelled 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 travelled by people in the same state on the same date, as shown in equation (7):

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

where G_d denotes the nationwide motor gasoline demand in the United States on date (d), G_B is the medium value of motor gasoline demand in the United States during the five-week period from 3 January to 6 February 2020 and T denotes the total miles travelled by people on normal days. This period is used as a benchmark for mobility change by Google¹⁹. Finally, y_n denotes the on-road demand share in state (n) in the United States, which is shown in Supplementary Table 3.

We solve equation (8) to calibrate the values of the adjustment factor vector \mathbf{X} and constant C :

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

where equation (8) is used to minimize the sum of the squared absolute percentage error. $G_{EIA,W}$ denotes the weekly motor gasoline demand in the United States, which is published by the US EIA every week. $\overline{G_d}$ is the average weekly value of the daily estimated motor gasoline demand (G_d). The term s.t. denotes 'subject to'.

EIA provides values of weekly finished motor gasoline supplies, which are used to estimate gasoline demand in the United States³¹. For this study, we used this data to calibrate the Motor Gasoline Demand Estimation Module. A quadratic equation was adopted to fit the seasonality of weekly motor gasoline demands in 2016–2019, as discussed in Supplementary Note 4, and it is 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 were fairly stable in 2017–2018³², and therefore the model assumes that the state shares of motor gasoline consumption before the pandemic remain the same in 2020. Moreover, the transportation sector accounts for 94–98% of motor gasoline consumption³². 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.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data used for model development are either provided in Supplementary Information or publicly available from EIA weekly gasoline demand, https://www.eia.gov/dnav/pet/pet_cons_wsup_k_c_w.htm; US pandemic data, <https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/>; YIG pandemic projection model, https://github.com/youyanggu/covid19_projections; MIT pandemic projection model, <https://github.com/COVIDAnalytics/DELPHI>; Apple mobility, <https://www.apple.com/covid19/mobility>; Google mobility, <https://www.google.com/covid19/mobility/>; government policy, <https://github.com/COVID19StatePolicy/SocialDistancing/blob/master/data/USStatesCov19distancingpolicy.csv>; NHTS 2017, <https://nhts.ornl.gov/> and 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 the figures in the main manuscript and Supplementary Information are provided as a Source Data file. With the evolution of the pandemic, the US motor gasoline demand and mobility level projection is updated periodically and posted at <https://covid19-mobility.com> and <https://teem.ornl.gov/poda.shtml>. Source data are provided with this paper.

Code availability

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

Received: 27 May 2020; Accepted: 29 June 2020;

Published online: 17 July 2020

References

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* <http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic> (2020).
3. COVID-19 Dashboard by the Center for Systems Science and Engineering at Johns Hopkins University (Johns Hopkins University, accessed 23 June 2020); <https://coronavirus.jhu.edu/map.html>

4. Sood, N. et al. Seroprevalence of SARS-CoV-2-specific antibodies among adults in Los Angeles County, California, on April 10–11. *JAMA* **323**, 2425–2427 (2020).
5. Egan, M. Oil prices turned negative. Hundreds of US oil companies could go bankrupt *CNN Business* (21 April 2020); <https://www.cnn.com/2020/04/20/business/oil-price-crash-bankruptcy/index.html>
6. Metz, T. *Coronavirus and Energy, a Sector Challenged by Geographical Concentrations* Energy and environment Working Paper 1 (Groupe d'études Géopolitiques, 2020).
7. Kilian, L. The economic effects of energy price shocks. *J. Econ. Lit.* **46**, 871–909 (2008).
8. *World Economic Outlook Update, June 2020: A Crisis Like No Other, An Uncertain Recovery* (International Monetary Fund, 2020); <https://www.imf.org/en/Publications/WEO/Issues/2020/06/24/WEOUpdateJune2020>
9. Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860–868 (2020).
10. Pasco, R. et al. *COVID-19 Healthcare Demand Projections: Austin, Texas* (The University of Texas at Austin, 2020); https://sites.cns.utexas.edu/sites/default/files/cid/files/covid-19_analysis_for_austin_march2020.pdf
11. 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).
12. 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).
13. 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).
14. 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).
15. *Short-Term Energy Outlook* (US Energy Information Administration, accessed 12 July 2020); <https://www.eia.gov/outlooks/steo/>
16. Swales, V., Lu, D. & Mervosh, S. See which states and cities have told residents to stay at home. *New York Times* (accessed 13 May 2020); <https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html>
17. *March 2020 Traffic Volume Trends* (US Department of Transportation Federal Highway Administration, accessed 12 June 2020); https://www.fhwa.dot.gov/policyinformation/travel_monitoring/20martvt/
18. *Weekly Petroleum Status Report* (US Energy Information Administration, accessed 12 June 2020); <https://www.eia.gov/petroleum/supply/weekly/>
19. *COVID-19 Community Mobility Reports* (Google, accessed 12 May 2020); <https://www.google.com/covid19/mobility/>
20. *Mobility Trends Reports* (Apple, accessed 12 May 2020); <https://www.apple.com/covid19/mobility>
21. *Daily Confirmed COVID-19 Deaths per Million, Rolling 7-Day Average* (OurWorldData, accessed 13 May 2020); <https://ourworldindata.org/grapher/daily-covid-deaths-per-million-7-day-average>
22. *Healthcare Facilities: Preparing for Community Transmission* (US Centers for Disease Control and Prevention (accessed 13 May 2020); <https://www.cdc.gov/coronavirus/2019-ncov/hcp/guidance-hcf.html>
23. *Gasoline is the Main US Transportation Fuel* (US Energy Information Administration, accessed 12 June 2020); <https://www.eia.gov/energyexplained/gasoline/use-of-gasoline.php>
24. Bertsimas, D. *DELPHI Epidemiological Case Predictions* (MIT, accessed 15 June 2020); <https://www.covidanalytics.io/projections>
25. 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. Preprint at <http://www.medRxiv.org/2020.05.04.20090670> (2020).
26. 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. Preprint at <http://www.medRxiv.org/2020.04.13.20063628> (2020).
27. *Annual Energy Outlook 2020* (US Energy Information Administration, 2020).
28. Paszke, A. et al. in *Advances in Neural Information Processing Systems* Vol. 32 (eds Wallach, H. et al.) 8026–8037 (Curran Associates, Inc., 2019).
29. Kingma, D. P. & Ba, J. Adam: method for stochastic optimization. Preprint at <http://arXiv.org/1412.6980> (2014).
30. *2017 National Household Travel Survey* (US Department of Transportation Federal Highway Administration, 2018); <http://nhts.ornl.gov>
31. *Petroleum & Other Liquids—Weekly Supply Estimates* (US Energy Information Administration, accessed 13 May 2020); https://www.eia.gov/dnav/pet/pet_sum_sndw_a_epm0f_vpp_mbblpd_w.htm
32. *State Energy Data System* (US Department of Energy, Energy Information Administration (accessed 14 May 2020); <https://www.eia.gov/state/seds>.

Acknowledgements

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 US Department of Energy. The authors thank the developer of the YYG model, Y. 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

The 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 were 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.

Competing interests

X.H., L.S., S.P. and J.B. are employees of Aramco Services Company, which is a US subsidiary of Saudi Aramco. Saudi Aramco is a fully integrated energy company whose business interests includes the sale of crude oil and refined products, which include gasoline.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41560-020-0662-1>.

Correspondence and requests for materials should be addressed to X.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

- All data used in this study are publicly available and free to download.
- (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 projection: https://github.com/youyanggu/covid19_projections.
(4) Apple Mobility: <https://www.apple.com/covid19/mobility>
(5) Google Mobility: <https://www.google.com/covid19/mobility/>
(6) Government policy: <https://github.com/COVID19StatePolicy/SocialDistancing/blob/master/data/USstatesCov19distancingpolicy.csv>
(7) National Household Travel Survey: <https://nhts.ornl.gov/>
(8) EIA State Energy Data System: <https://www.eia.gov/state/seds/>

Data analysis

- We used Python 3.7 to analyze the data and model development. The following python packages are used: pandas, numpy, shutil, torch, pytorch, seaborn, sklearn, tqdm, and matplotlib.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data used for model development are either provided in the Supplementary Material 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 projection: https://github.com/youyanggu/covid19_projections.
- (4) Apple Mobility: <https://www.apple.com/covid19/mobility>
- (5) Google Mobility: <https://www.google.com/covid19/mobility/>
- (6) Government policy: <https://github.com/COVID19StatePolicy/SocialDistancing/blob/master/data/USstatesCov19distancingpolicy.csv>
- (7) National Household Travel Survey: <https://nhts.ornl.gov/>
- (8) EIA State Energy Data System: <https://www.eia.gov/state/seds/>

The results support the findings of this study are provided in the main text and Supplementary Material. With the evolution of the pandemic, the U.S. motor gasoline demand projection will be updated weekly and posted at <https://covid19-mobility.com/>.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

In this paper, we developed a machine-learning based model to quantify and analyze the impact of the COVID-19 pandemic on people's travel/trip activities. These trips are connected to motor gasoline demand based on the probabilities of miles taken for different travel activities. By adopting the pandemic projections from other studies, we used the model to project the short-term fuel demand as well as studying the impact of government intervention.

Research sample

We used the following data in this study, which are publicly available and free to download.

- (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 projection: https://github.com/youyanggu/covid19_projections.
- (4) Apple Mobility: <https://www.apple.com/covid19/mobility>
- (5) Google Mobility: <https://www.google.com/covid19/mobility/>
- (6) Government policy: <https://github.com/COVID19StatePolicy/SocialDistancing/blob/master/data/USstatesCov19distancingpolicy.csv>
- (7) National Household Travel Survey: <https://nhts.ornl.gov/>
- (8) EIA State Energy Data System: <https://www.eia.gov/state/seds/>

Sampling strategy

NA

Data collection

We did not collect the data ourself, but using the data published by others.

Timing

We used daily US National and state level pandemic and mobility data, weekly EIA US gasoline data from 2/1/2020 to current for this study. We also use the monthly fuel demand form 2016-2019 as the no-pandemic fuel demand baseline.

Data exclusions

No data were excluded.

Non-participation

NA

Randomization

NA

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies
<input checked="" type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology and archaeology
<input checked="" type="checkbox"/>	Animals and other organisms
<input checked="" type="checkbox"/>	Human research participants
<input checked="" type="checkbox"/>	Clinical data
<input checked="" type="checkbox"/>	Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging