2021

Investigating the effect of E30 fuel on long term vehicle performance, adaptability and economic feasibility

Adil Alsiyabi

Seth Stroh

Rajib Saha

Fuel 306 (2021) 121629

Contents lists available at ScienceDirect

Fuel

journal homepage: www.elsevier.com/locate/fuel

Full Length Article

Investigating the effect of E30 fuel on long term vehicle performance, adaptability and economic feasibility

Adil Alsiyabi, Seth Stroh, Rajib Saha *

Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, USA

ARTICLE INFO

Keywords:
Higher ethanol blend fuel
Vehicle performance
Machine learning
Economic analysis

ABSTRACT

Due to the drawbacks associated with the use of petroleum derived fuels, the use of more sustainable fuel sources has garnered increasing attention in several sectors including road transportation. However, the transition away from gasoline is often hindered by the inability of currently operating vehicles to efficiently run under alternative fuels. Therefore, the logical short-term alternative is to transition to clean fuel sources including higher-ethanol fuel blends that are compatible with current fuel systems and spark-ignition engines. In this work, the long-term adaptability and economic feasibility of non-flex vehicles to consume a 30% ethanol (E30) fuel blend was investigated. Sixteen diagnostic and operating parameters were tracked in real-time in 50 vehicles over a one-year period. The vast amount of data generated was used to train sparse regression and machine learning models to explore differences in performance and operational robustness of commercial vehicles consuming E30 blends compared to the ones consuming 15% blends. Results indicate that although modest changes can be observed in the behavior of a subset of parameters, overall performance and adaptability are not compromised by consumption of E30. It was determined that an average price difference of 2.5% is sufficient to offset the mileage loss caused by the increased ethanol concentration. Finally, we discuss the large-scale environmental impact of an incremental nation-wide shift towards E30 consumption.

1. Background

As the world transitions away from fossil fuels into more renewable energy sources, many industries have shown a growing interest in investigating and eventually transitioning to more environmentally friendly fuel sources. One such case is the transportation sector, where approximately 142 billion gallons of motor gasoline are used annually [1]. Furthermore, since the turnover of vehicles currently on the road is very slow (~15 years [2]), it will be a long time before this number decreases due to introduction of vehicles operating on alternative sources of energy (e.g. electricity or hydrogen). Therefore, the only immediate solution that can be implemented is through the use of fuel sources compatible with current engine designs, either independently or as additives. One such candidate that fits these requirements is ethanol.

Ethanol is an alcohol produced primarily from agricultural feedstocks such as corn or sugarcane making it a renewable fuel source. It has been used to power auto-engines since the 19th century [3]. Currently, all commercial vehicles can use ethanol blends up to 15% (E15), while flex-fuel vehicles can use fuel blends containing 85% (E85). Several studies were conducted to investigate how addition of higher concentrations of ethanol affects different aspects of vehicle performance and functioning. These studies looked at the fuel blends' effects on different materials that the fuel comes into contact with, the emission profile of engines operating on higher ethanol blends, and the effect that increasing ethanol concentrations has on fuel efficiency and engine performance. A number of previous studies investigated the effect of higher ethanol concentration on material compatibility in commercial vehicles. Matejovsky et al. conducted a study in 2017 that looked at the effect of increasing ethanol concentrations on the corrosivity of different materials [4]. They found that although there was a slight increase in corrosion rate with higher ethanol blends, the observed change was insignificant as compared to the lowest threshold of 0.0025 mm/year [4]. The test was repeated after increasing the moisture content, and the researchers reached the same conclusion. Another study conducted at the Oak Ridge National Lab looked at the effect that different ethanol blends had on the properties of different elastomers. The researchers found that the effect due to the increased ethanol concentration was negligible compared to variability due to the different types of

* Corresponding author.
E-mail address: rsaha2@unl.edu (R. Saha).

https://doi.org/10.1016/j.fuel.2021.121629
Received 9 April 2021; Received in revised form 28 July 2021; Accepted 31 July 2021
Available online 7 August 2021

elastomers tested [5]. A different study looked at the effect of different ethanol blends on the function of engine oil. The metrics used to determine differences between the blends were the friction coefficient and the wear scar diameter. Their results indicate that increasing the ethanol content up to 85% had minimal effect on the friction coefficient, and no effect on the wear scar diameter [6].

Furthermore, a large number of studies were conducted to study the effect of different ethanol blends on the emission profile. Table 1 summarizes the main results of these studies. The first column of the table details the different fuel blends investigated in each referenced study, and subsequent columns indicate the general pattern observed in emission profiles for each particular pollutant as the ethanol concentration increases. As can be seen from the table, both greenhouse gas emissions and other emitted pollutants such as nitrous oxide compounds and hydrocarbons decrease with higher ethanol blends. Although there are some discrepancies in the reported values due to variations in the blending process, method of fuel injection, and experimental setup, the general trend points to reduced emissions across the board.

Finally, studies investigated the effect of increasing ethanol concentrations on fuel efficiency and engine performance. Due to the lower energy density of ethanol compared to gasoline, it is assumed to be much less efficient, and that any savings resulting from the lower price of ethanol would be offset by the decrease in mileage due to the lower energy density. However, this assumption doesn't take into account other factors that offset this decrease in energy such as ethanol's higher combustion efficiency and octane rating. Studies found that there is an increase in cylinder pressure, temperature, and combustion efficiency, as well as a decrease in knocking and combustion duration (see Table S1). In turn, these effects can significantly offset the lower energy density of ethanol. Moreover, studies looking at differences in engine performance reported varying results depending on the implemented general setup, nevertheless, the general trend points to an increase in torque and brake specific fuel consumption (see Table S2).

Over the past decade, machine learning techniques have become a valuable tool in addressing questions of complex nature [14]. These techniques rely on analyzing large datasets to identify emerging patterns that are otherwise difficult to identify [15]. Different machine learning methods have been applied to obtain relationships between multiple variables, cluster data into different sets, and classify inputs into different categories [16,17]. Due to the wide array of applications that can be achieved, machine learning has found its use in almost all areas of research [14,16]. A plethora of such algorithms have been developed in the last decade, and a number of excellent reviews have detailed the strengths and weaknesses of each method [18–20]. In recent years, artificial neural networks (NNs) have gained increasing popularity due to their wide-range applicability, high performance, and their nonparametric nature, which enable modeling of complex functional forms. Within the context of energy and fuel consumption, NNs have been used to predict biofuel quality parameters such as oxidative stability using infrared data [21]. Other studies have used such techniques to investigate how different factors influence fuel consumption [22,23]. However, the main drawback of using NNs is their black-box nature, which makes it difficult to untangle how the algorithm reaches a certain prediction. To remedy this limitation, a framework called Sparse Identification of Nonlinear Dynamics (SINDy) that combines sparsity and machine learning techniques has been developed to elucidate predictive governing equations from large datasets [24]. This allows for a mechanistic understanding of how the predictions are generated and gives insights into the weighting factors that each parameter in the dataset contributes to predict the desired variable. Both NN and SINDy can be utilized to harness real-time data generated from the vehicle's on-board control module to predict how a change in one parameter (e.g. fuel ethanol content) affects other performance-related parameters. Based on the resulting analysis, it can be concluded whether or not the vehicle can adapt to the change in fuel ethanol content.

In this study, the long-term adaptability and economic feasibility of non-flex vehicles to E30 consumption was investigated. A data-driven approach was followed in which 16 diagnostic and operating parameters were collected over a one-year period from vehicles operating on both E30 (test) and E15 (control) fuel blends. A myriad of statistical and data analysis methods including sparse regression and neural networks were used to compare the behavior of these parameters temporally and between the two fuel types. Furthermore, data was collected on the average mileage per gallon for each fuel type to explore the economic feasibility of utilizing E30. It was observed that the engine control module of non-flex vehicles was capable of maintaining major operating parameters (i.e., optimum air to fuel ratio and engine temperature) at their desired set-points. Moreover, machine learning models trained to predict instantaneous volumetric efficiency (as a performance metric) displayed no difference between the two fuel types. Finally, the use of E30 fuel was found to become economically viable when price differences exceed 2.5% (i.e., when E30 is 2.5% cheaper). This demonstration paves the way for further studies investigating the long-term adaptability of high-alcohol fuel blends and is intended as a guide to legislators and consumers on the risks and benefits of consuming higher alcohol fuel blends in commercial non-flex vehicles.

Table 1
Emission profiles of pollutants under different ethanol-blends relative to gasoline. Arrows indicate the effect of increasing ethanol content on the emission of each pollutant.

Ethanol blends (%)	CO	NOx	CO2	HC	Reference
E0; E25; E50; E75; E100	↓	↓	↓	↓	[7]
E10; E20; E30; E40	–	↓	–	–	[8]
E0; E5; E10; E20; E30	↓	↑	↑	↓	[9]
E0; E5; E10; E20; E30	↓	–	↓ **or** ↑	↓	[10]
E0; E10; E20; E40; E60	↓	–	↑	↓	[11]
E5; E10; E20; E30	↓	↓	–	↓	[12]
E0; E10; E30	↓	↓	–	↓	[13]

2. Results and discussion

2.1. The effect of E30 use on fuel/air mixture and engine temperature control

To analyze the long-term adaptability of vehicles to E30 fuel, On-Board Diagnostic (OBD) parameters expected to remain constant throughout the drive cycle were investigated to determine their response to the increased ethanol content. Long-term fuel trim (LTFT) and O_2 sensor readings indicate whether the vehicle's engine control module (ECM) can adapt to the increased oxygen concentration resulting from the added ethanol. The ECM controls the air/fuel ratio (AFR) by measuring the voltage generated through the oxygen sensor which indicates the proportion of oxygen in the exhaust [25]. A voltage which is too low indicates that the engine is running lean (excess oxygen), while a high voltage indicates that the engine is running rich (excess fuel). Readings from the oxygen sensor (controlled variable) signal whether any changes are required to the fuel trim. LTFT, displayed as a percentage, is a long-term average of the adjustment made to the fuel mixture to maintain a balanced AFR. For engines running on gasoline, the optimal AFR is 14.7:1. Therefore, a new non-flex fuel vehicle consuming pure gasoline is set to have a fuel trim of 0%. Deviations from this value occur over time due to several reasons, however, as long as the values are maintained above −25% (rich fuel mixture) and below 25% (lean fuel mixture), no warnings arise and the oxygen concentration in the exhaust can still be controlled.

Due to the higher oxygen content, the optimal AFR of pure ethanol (9:1) is significantly lower than that of gasoline. Therefore, it is expected that the addition of higher ethanol blends will result in a leaner fuel mixture. For a 30% ethanol concentration, the optimal air/fuel ratio will be approximately 13:1. Subsequently, this will lead to an automatic increase in the LTFT in order to maintain a constant oxygen concentration in the exhaust (i.e. constant O_2 sensor reading). The distribution of these two parameters, recorded over an entire year, were compared between vehicles operating on E30 compared to those operating on E15. As expected, there was an average increase of approximately 3.5% (11.4% vs 7.9%) in the LTFT of vehicles operating on E30 (Fig. 1A). However, the distribution of O_2 sensor readings for the two conditions were similar (Fig. 1A). This indicates that the ECM of the tested vehicles was able to account for the increased oxygen content in the fuel. It is noted that a multitude of factors other than fuel type can affect fuel trim (e.g., air leaks, faulty fuel pumps or injectors, etc.) and combine to result in a high enough LTFT to trigger a warning [26]. However, from this analysis, it can be concluded that the average increase due to fuel type alone can be handled by the control system of non-flex fuel vehicles. Furthermore, a time-series analysis was conducted to determine whether there were any temporal changes in the amount of oxygen evolution in vehicles operating on E30, indicating any underlying mechanical deterioration or deviation from the desired set point over time. As can be seen from Fig. 1B, although the data is inherently noisy due to the scale of the measured voltage, no trend can be observed across drive cycles for either fuel type. A similar analysis was conducted to determine whether there was a significant difference in engine operating temperature between the two fuel types. As can be seen from Fig. 2, although there was a slight but significant increase in the average engine temperature of vehicles operating on E30 compared to those operating on E15 (197 °F and 195 °F, respectively), both fuel types remained within the acceptable range of 195 °F to 220 °F and significantly below engine temperatures that would trigger any warnings (240–250 °F). A comparison of the calculated slope for each vehicle in Fig. 1B and 2B verified that there was no significant difference between the change in O_2 sensor or coolant temperature readings over time between the two fuel types (p-values of 0.186 and 0.748, respectively. Furthermore, analysis of the stationarity of the time-series behavior of both parameters has been detailed in Supplementary File 1.

2.2. The effect of E30 use on vehicle performance: A comparison of volumetric efficiency

Next, more complex parameters that display transient behavior such as engine timing, throttle position, and engine speed were analyzed to look for any underlying differences that might arise from using a higher ethanol blend. To accomplish this, a performance metric had to be chosen that could be used to fit the measured parameters to. More importantly, this chosen metric would also have to be measured so that it could be used to train the fitting algorithms. Absolute load (AL) was chosen as the performance metric. In addition to being one of the parameters that could be recorded by the OBD tracker, AL is dependent on multiple other parameters and is a direct measure of volumetric efficiency. Therefore, if a model can be obtained in which the other recorded parameters can be used to predict this metric (AL), then a direct comparison can be made between models generated using data from vehicles consuming E15 and those consuming E30.

Phase I: Modeling Absolute load using SINDy

In the first phase of the analysis, a recently developed framework was implemented to determine an empirical formula relating the gathered (independent) vehicle parameters to absolute load. The framework, called sparse identification of nonlinear dynamical systems (SINDy), uses gathered data on the parameters of interest to come up with a governing equation describing the relationship between the independent and dependent variables [24]. The goal of this analysis was to (i) determine whether a relationship did in fact exist between the gathered parameters and absolute load, and (ii) to determine how coefficients of the predicted equations vary between vehicles running on E15 compared to those running on E30.

The analysis showed that 9 terms comprising 5 unique parameters were required to describe absolute load (Table 2). Since the SINDy framework relies on sparse regression, the nine identified terms comprise the minimal number of terms required to accurately predict the variable of interest (AL). The majority of the non-zero linear terms (with the exception of engine timing), which alone resulted in a correlation of approximately 0.88, were similar between the two fuel types. Furthermore, by comparing the ranges of the coefficient values, it can be seen that although most distributions are statistically different ($P < 0.01$), the magnitude of the values are often similar. The main difference to this is the effect of intake temperature on AL. As can be seen from both the linear and non-linear coefficients associated with this parameter, the effect of ambient temperature is much less pronounced on the AL of vehicles operating on E30 as compared to those using E15. In other words, all other parameters being equal, vehicles operating in the same temperature will have a higher AL when consuming E30. A mechanistic explanation of this result is not clear.

The model's performance was also evaluated using a number of conventional metrics (Table 3 and Fig. 3A). The models displayed satisfactory performance when tested on data from a different fuel blend compared to the one used for training. This indicated that if a better

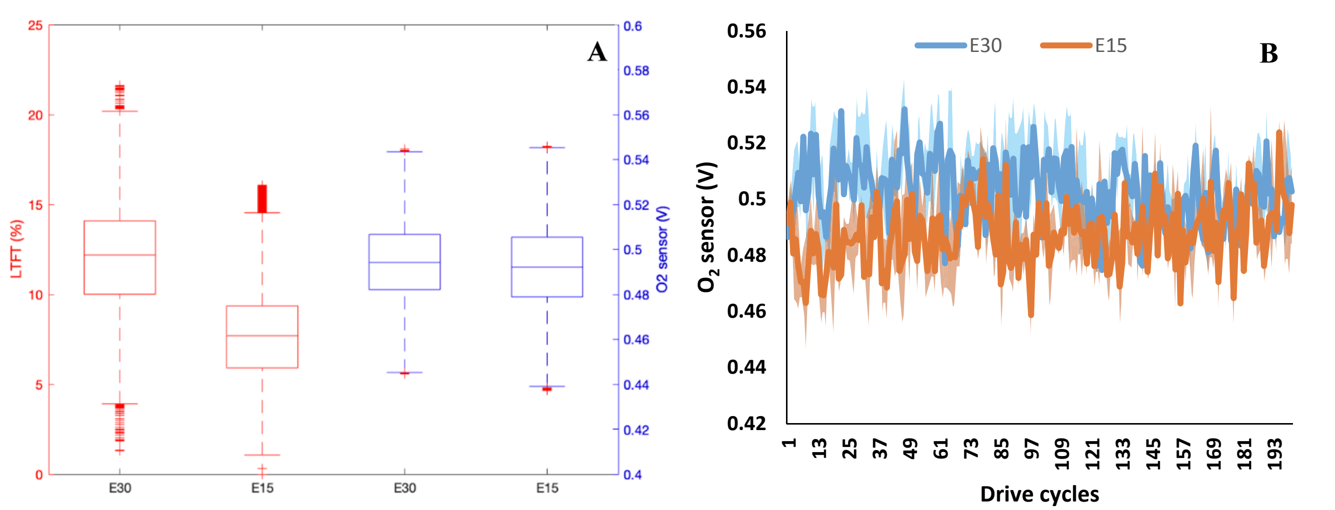

Fig. 1. Distribution and temporal analysis of parameters affecting air/fuel ratio for vehicles operating on E30 and E15. (A) Box plot of the distribution of average LTFT and O2 sensor values, and (B) O2 sensor readings across drive cycles for three randomly selected vehicles operating on each fuel type.

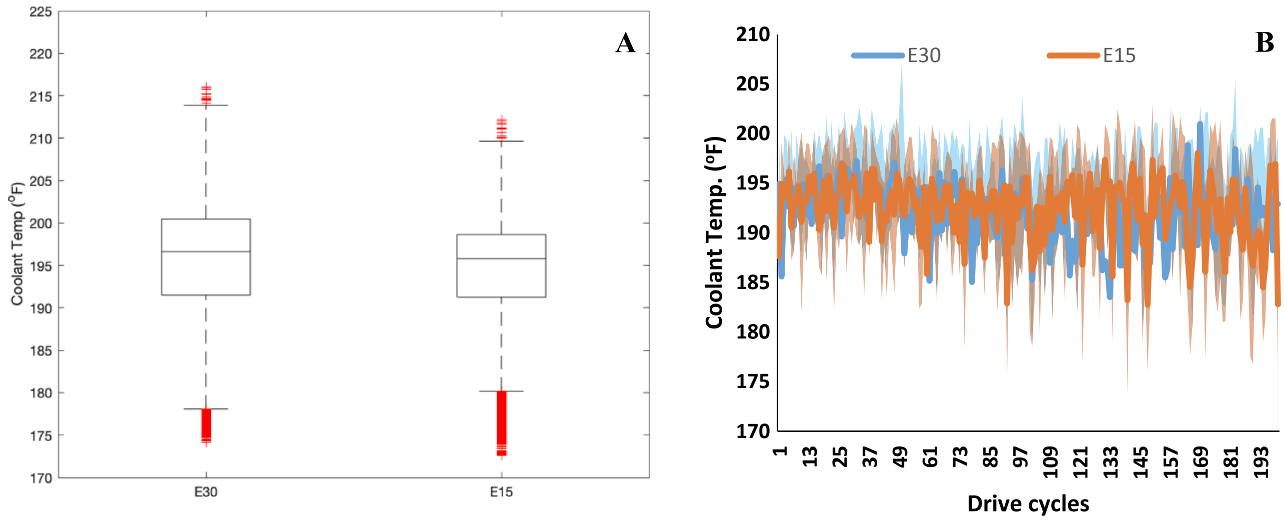

Fig. 2. Distribution and temporal analysis of engine coolant temperature for vehicles operating on E30 and E15. (A) Box plot of the distribution of average coolant temperature values, and (B) coolant readings across drive cycles for three randomly selected vehicles operating on each fuel type.

Table 2
Coefficients of the parameters predicted to affect absolute load using SINDy.

Parameter	E30	E15	*P* value
Engine Speed (RPM)	−0.2395 ± 0.0010	−0.2425 ± 0.0023	0.0513
Engine Timing (%)	0 ± 0	−0.0640 ± 0.0011	3.63E-14
Intake Air Temp (^{0}F)	−0.5848 ± 0.1210	−2.7476 ± 0.0562	8.91E-10
LTFT (%)	0.1347 ± 0.0009	0.1333 ± 0.0011	0.092
Throttle Position (%)	1.5491 ± 0.0011	1.4557 ± 0.0027	4.14E-12
(engine speed)^2	0 ± 0	−0.0671 ± 0.0011	2.41E-14
(Engine Speed)×(Throttle Position)	0.0938 ± 0.0012	0.1340 ±0.003	6.89E-09
(Intake Air Temp)^2	0.2289 ± 0.0462	0.9522 ± 0.0195	2.24E-09
(Throttle Position)^2	−0.2021 ± 0.0002	−0.1924 ± 0.0005	2.05E-10

fitting method was used, it would be possible to observe no difference between the models obtained using data from vehicles running on either fuel blend. This was the main objective behind the analysis in phase II.

Phase II: Modeling Absolute load using a deep neural network

In phase II, neural networks (NNs) were used to accomplish a similar task to that in phase I. The main differences being that the use of a NN would eliminate any biases introduced when using SINDy and would not require any pre-normalization of the data [15]. After training and validating data from each fuel type individually, the NN trained using data from vehicles operating on E15 was used to predict the absolute load of vehicles operating on E30 fuel (Fig. 3B). The rationale being that if the accuracy remains high in this case, then no observable differences exist in performance between the vehicles operating on either fuel type. As can be seen from Table 3 and Fig. 3B, the NN outperformed SINDy across all calculated metrices. More importantly, the model performed just as well when tested on a different fuel blend than the one used for training. This leads to the conclusion that based on the performance metric chosen (AL), there is no significant difference between operating on either fuel blend.

By employing the modeling frameworks above, it was possible to contextualize the vast amount of data gathered and therefore determine the effect of using E30 on engine performance. Overall, application of SINDy and NNs made it possible to predict the complex relationship between a multitude of driving parameters and to ensure that this relationship did not change with increasing ethanol content.

2.3. Analysis of the long-term economic feasibility of E30 vs. E15

The long-term economic feasibility of using E30 compared to E15 was analyzed by comparing the average cost per mile of each fuel type. To achieve this, fueling data was collected through driver logs to calculate the fuel efficiency in terms of mileage per gallon (see methods section). Variations arising due to factors other than fuel type were minimized by pairing up vehicles according to their daily driving routines and by using similar makes and models as both test and control vehicles. In addition to the large sample size, it was assumed that these measures are sufficient to eliminate any underlying factors affecting the analysis (e.g., driving routes, highway vs city miles, etc.).

Fig. 4(A) below shows the average monthly fuel efficiency of the two different ethanol blends throughout the time period of the demonstration. As can be seen, E15 is on average 0.64 MPG more efficient than E30 ($P < 0.005$). Furthermore, the fluctuations over time between the two different fuel types were found to be consistent which indicates that temporal efficiency changes due to weather likely affect both fuel blends similarly. It is also possible that the fluctuations are impacted by changes in driving patterns across different weather conditions. Furthermore, the error bars demonstrating the standard deviation for each condition further emphasize the effect of driving conditions.

Next, the collected data was used to calculate the price per mile for each fuel type. It is noted that when calculating this metric, rack prices for gasoline and ethanol were used instead of gas station prices. However, this should not affect the analysis as it is assumed that gas station

Table 3
Model performance of the methods used to predict absolute load.

Model	correlation	bias	MAE	R^2	RMSE
SINDy					
E30	0.9708	0	0.1691	0.9424	0.2399
E15	0.9713	0	0.1667	0.9434	0.2379
E30 (E15)[a]	0.9357	−0.0048	0.2429	0.8741	0.3548
E15 (E30)	0.9674	0.0034	0.1821	0.935	0.255
NN					
E30	0.9861	−0.0002	0.1112	0.9724	0.1661
E15	0.9862	−0.0002	0.1173	0.9726	0.1655
E30 (E15)	0.9826	0.0280	0.1346	0.9639	0.1901
E15 (E30)	0.9821	−0.0266	0.1418	0.9632	0.1918

[a] Model trained on E15 data and tested on E30 data.

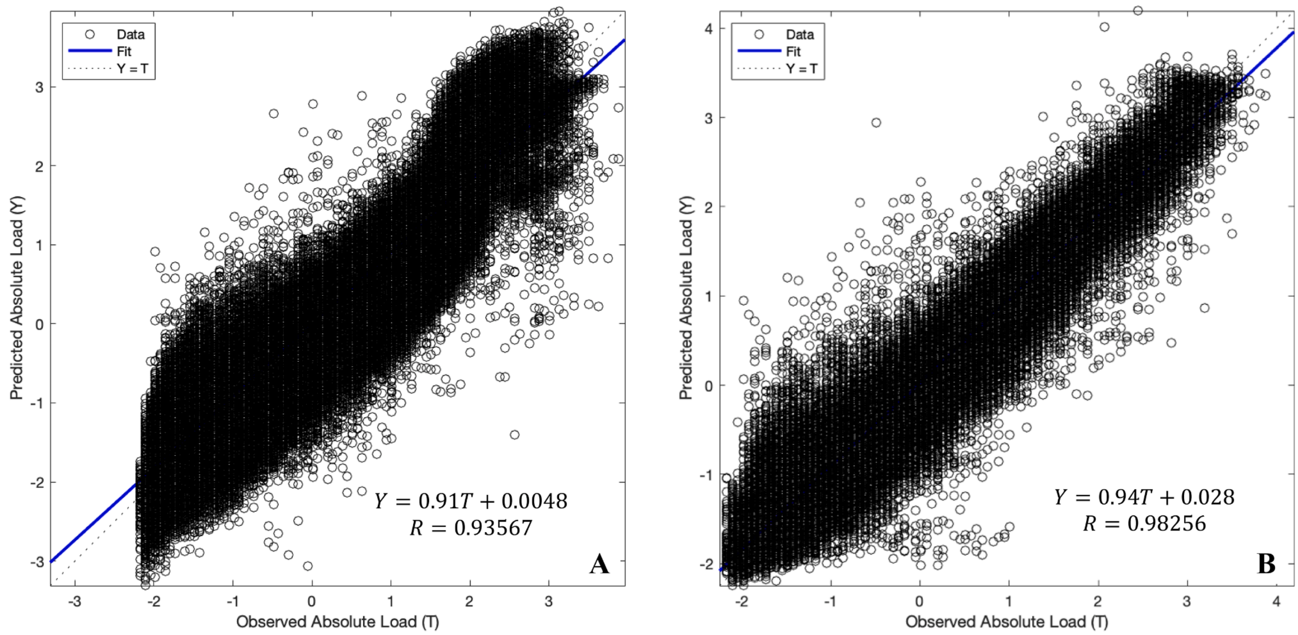

Fig. 3. Model performance of (A) the SINDy algorithm and (B) a 1 hidden layer neural network. Both models were trained using data from vehicles consuming E15 and tested on data from vehicles consuming E30.

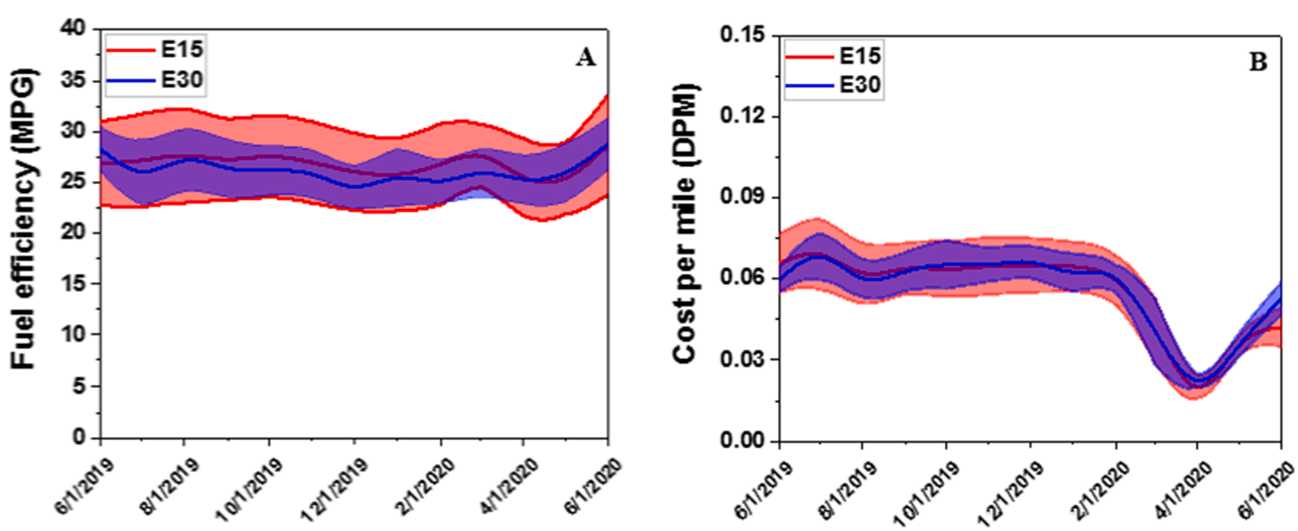

Fig. 4. Comparison of the economic feasibility of using E15 and E30. (A) Time-course fuel efficiency of the two different ethanol concentrations. (B) Time-course price per mile for the two different ethanol concentrations.

markup would be similar for both E15 and E30. Fig. 4(B) below shows the average monthly price per mile for each ethanol concentration. As can be seen, the slight differences in fuel efficiency are offset by the reduced price of E30. Furthermore, a sudden decrease in price can be observed starting from March of 2020. This decrease was due to a sudden drop in both gasoline and ethanol brought about by the COVID-19 pandemic. The effect on gasoline was exacerbated by a concurrent decline in price due to political circumstances. This explains why in April of 2020, the price of E15 fuel was cheaper than E30.

2.4. Long-term impact of E30 use on CO_2 emissions and ethanol consumption

Ethanol combustion releases 35% less carbon dioxide compared to pure gasoline. Furthermore, the combustion of ethanol made from biomass (such as corn and sugarcane) is considered atmospheric carbon neutral [27]. In fact, recent well-to-wheel life cycle analysis (LCA) has shown that the average carbon intensity for corn ethanol is approximately 46% compared to that of neat gasoline [28]. This analysis took into account nine emissions categories, including ones from upstream processes such as corn production and transport, including farming, land use change, feedstock transport, and co-product credit. It also took into account emissions resulting from fuel production and transport, addition of denaturant, and tailpipe emissions [28]. Therefore, increasing ethanol content inherently leads to reductions in overall CO_2 emission. Currently, about 95% of the gasoline sold contains 10% ethanol [29]. Moreover, a typical passenger vehicle emits 5.1–5.2 tons of CO_2 per year [30]. This value would be reduced to approximately 4.9 tons per year when 30% ethanol is used instead. When the total number of light-duty non-flex fuel vehicles is considered (~233 million vehicles [29]), this reduction becomes very significant. Fig. 5A below shows the projected cumulative reduction in CO_2 emissions resulting from only 10% of the

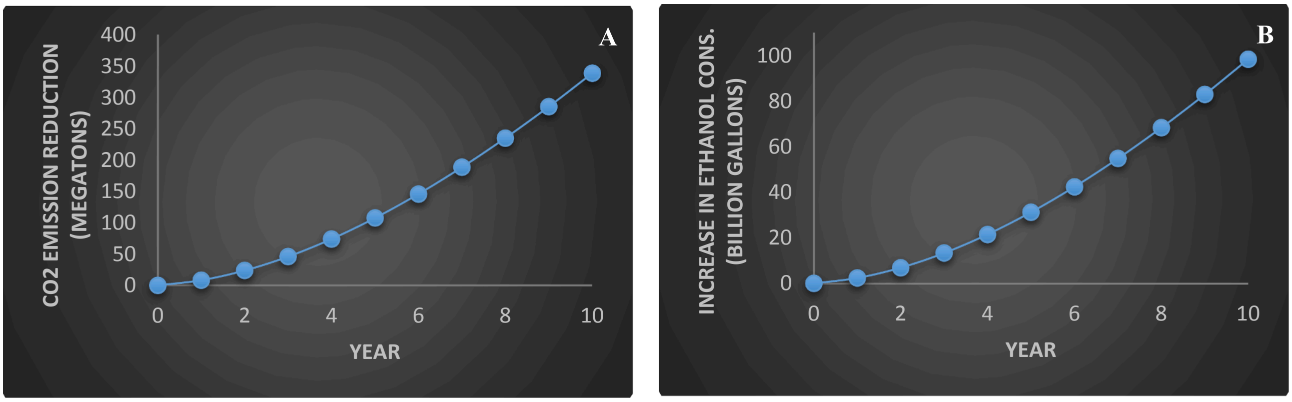

Fig. 5. Impact of using E30 on CO_2 emissions and ethanol consumption. (A) Predicted reduction in carbon dioxide emissions if 10% of US vehicles switched to E30 every year. (B) Predicted increase in ethanol consumption if 10% of US vehicles switched to E30 every year.

US's on-road fleet switching from E10 to E30 every year. Such a transition would result in an average reduction of approximately 34 megatons emitted every year. It is noted that these estimates are conservative, as they do not take into account the reduced carbon intensity of producing ethanol compared to gasoline [28]. Fig. 5B shows the corresponding increase in ethanol consumption over the same time period. Such a shift will constitute a notable milestone on the road to mitigating the damage caused by global warming.

3. Conclusion

This study investigated the long-term feasibility of using a 30% ethanol blend (E30) as compared to a 15% blend (E15) on non-flex fuel vehicles. Feasibility was defined as (i) the vehicle's ability to adapt to a fuel blend containing 30% ethanol without sudden, or gradual deviations in performance and (ii) the economic viability of using a higher ethanol blend. Performance metrics were gathered through on-board-diagnostic (OBD) trackers which continuously recorded the values of 16 parameters from 50 vehicles consuming either E15 or E30. Analysis revealed that although differences in the performance metrics were statistically significant, the magnitude of these differences were minimal. This was evident both for parameters that were expected to remain constant (fuel trim, oxygen sensor, and coolant temp), and the combinatorial behavior of parameters which were constantly changing. As expected, the difference in oxygen content between the two blends caused the long-term fuel trim of vehicles operating on E30 to be higher. However, the magnitude of this change was below the limit set by most car manufacturers (20–25%). Furthermore, the change in fuel trim allowed the vehicles' control modules to control for the amount of unburnt oxygen. This indicates that the control system of non-flex fuel vehicles is capable of handling ethanol concentrations up to 30%. In addition, analysis of transiently behaving parameters showed that according to the performance metric chosen (absolute load), no difference could be observed between the performance of vehicles operating on either fuel blend. The economic analysis showed that the use of E30 leads to an average decrease in fuel efficiency of approximately 0.64 MPG compared to E15. This minor variation in efficiency is usually offset by the lower price of E30 compared to E15. In fact, a 2.5% price difference is sufficient in counterbalancing the efficiency gap. Therefore, when the price differential is greater than 2.5%, the use of E30 becomes more economically viable. Finally, a strategy of introducing E30 fuel and gradually shifting from E15 to E30 consumption will have appreciable environmental effects due to reductions in greenhouse gas emissions. The slow turnover of vehicles currently on the road requires the proposal of immediate solutions compatible with the present infrastructure and engine design. Therefore, implementation of such a strategy can act as a buffering solution to reduce the rate of greenhouse gas emissions until improved fuel technologies and electric vehicles become more widespread.

This study constitutes a good starting point towards ensuring the safety of using 30% ethanol in non-flex fuel vehicles. Future studies can build on this demonstration in two ways: by (i) replicating this analysis on a broader range of vehicles makes and models (coarse-grained analysis), or (ii) investigating the physical effect of long-term E30 use on the fuel system under real-life driving conditions (fine-grained analysis). Studies looking into the physical effect of consuming E30 should compare the long-term accumulation of particulate matter in the fuel pump and fuel injectors between vehicles operating on E15 and those operating on E30. Both types of analysis will provide valuable information and will ensure the safety and feasibility of using higher ethanol blends in non-flex fuel vehicles.

4. Methods and experimental design

4.1. OBD tracker setup and data acquisition

Third generation kiwi OBD trackers were purchased from PLX devices. The trackers continuously log up to 16 parameters from the vehicles' diagnostic system (data is logged every second of the drive cycle). The logged data is stored in a 6 GB SD card that was retrieved at multiple time points during the demonstration to gather the data. A detailed description of the setup procedure is provided in Supplementary File 2. In short, the tracker is first connected to a mobile device before the vehicle is started. The parameters of interest and mode of data acquisition are selected, and the vehicle is turned on. A list of the selected parameters is also provided in Supplementary File 2. If the device was set up correctly, the tracker should beep at every power cycle. As noted in the supplementary information, this step can cause a number of contingencies. Therefore, after the first power cycle, the microSD cards were removed to ensure that the parameters of interest were being recorded. The kiwi tracker records each drive cycle as an independent excel file. A sample output is provided in Supplementary File 3. It is also noted that if the vehicle's battery is disconnected, the tracker can no longer accurately keep track of the date unless reconnected through the mobile device. No other parameters are affected by this disconnection.

Information regarding vehicle make/model, year, and miles driven is provided in Supplementary File 4. Out of the 50 participating vehicles, 43 accumulated sufficient data for subsequent adaptability and performance analysis. The majority of these vehicles were operated under normal driving conditions. However, nine of the vehicles were State Patrol owned muscle vehicles operated under extreme conditions. Data from these vehicles was analyzed separately, and the resulting figures are provided in Supplementary File 5. Due to the relative small dataset generated from these vehicles compared to the 41 vehicles operating under normal conditions, the data was only used to determine adaptability and economic feasibility but was not used to predict performance (Supplementary File 5).

4.2. Data pre-processing and statistical analysis

To ensure that the analyzed parameters had enough time to stabilize, only drive cycles that were longer than 10 min were analyzed. This threshold was chosen since the trackers start recording as soon as the vehicle is started, meaning that in most occasions, shorter driving times indicate that the vehicle remained idle for the entire power cycle. Furthermore, after filtering out short drive cycles, the remaining files were pooled into either test (E30) or control (E15) pools for further analysis.

Long-term fuel trim (LTFT), coolant temperature, and oxygen sensor readings constituted a set of parameters that expected to remain constant and within a certain range throughout the demonstration. Comparison of these parameters required comparing each parameter's distribution of values between the test and control case. A python code was implemented to determine whether a significant difference was present between the parameters recorded for each of the two fuel types. The boxplot was generated in MATLAB using the built-in boxplot function. Values outside of the three interquartile range were considered outliers.

4.3. Predicting volumetric efficiency using SINDy

Sparse identification of nonlinear dynamics (SINDy) is a sparse regression method used to derive (non)linear relationships between dynamic parameters [24]. This input requires the specification of a dependent variable (y) to be fit, and a set of potential variables to be used as independent variables (x). For this work, absolute load was set as the dependent variable, and the remaining parameters were input as independent variables. The output of the analysis is a sparse vector of coefficients constituting the best-fit solution to the regression problem (Table 2). The framework allows users to specify the degree of polynomial nonlinearity (n) to be tested. For this work, n was set to two to avoid overfitting. Data gathered and pre-processed from the OBD trackers of vehicles operating on each of the fuel types was randomly divided into 5 equally sized groups to determine coefficient sensitivity to the training data. A MATLAB implementation developed by the authors of the original work was used for model generation. Correlation, bias, mean absolute error (MAE), root mean square error (RMSE) and the coefficient of determination (R^2) were used to evaluate model performance.

4.4. Predicting volumetric efficiency using neural networks

The Deep Learning Toolbox was used to generate a neural network using the toolbox's fitting app. A variety of network architectures were tested (i.e. number of nodes and hidden layers) and it was found that 10 nodes and a hidden layer were sufficient. The Levenberg-Marquardt training algorithm was used with 70% of the data used for training, 15% for validation, and 15% for testing. Model performance was evaluated as described in the preceding section.

4.5. Economic analysis

Driver logs were given to each of the drivers to document fuel fill-ups (Supplementary Fig. S3). Drivers documented the date, station name and location, fuel blend, gallons filled, mileage, and any comments on performance or issues. Drivers were instructed to fill-up the tank at each filling stop. In cases where the tested ethanol blend (E15 or E30) was not available, the drivers were instructed to use E10. Such cases were taken into account when calculating the cost per mile. 48 out of the 50 participating vehicles accumulated sufficient data for the economic analysis. Data from the logs was entered into an Excel spreadsheet ("Compiled log data.xlsx") for subsequent analysis. This dataset can be found in the GitHub page associated with study (https://github.com/aalsiyabi/E30). Miles traveled on each tank was calculated by taking the difference in mileage between gas fills. In cases where the mileage entered was inconsistent, or when other important parameters such as the number of gallons filled were missing, the entire row of data was labelled problematic and omitted from subsequent analysis. Furthermore, cases where an incorrect fuel blend was used were also discarded. Out of a total of 2873 data points, each constituting a gas fill-up, 4.1% of the data was manually discarded for the reasons stated above. Average fuel efficiency (i.e. MPG) was then calculated by dividing the distance traveled in miles by the volume of fuel added. This data was consolidated by removing log data from vehicles with substantial deviations in their calculated MPG. Outliers, defined as values below the first interquartile or above the third interquartile range, were omitted. These values usually arise due to errors in logging the mileage or the amount of gasoline added. Finally, the monthly average and standard deviation was calculated for vehicles using each of the two fuel blends.

A three-terminal average of gasoline and ethanol rack prices from Doniphan, North Platte, and Omaha was used to calculate the price per mile based on the percentage of ethanol used. The daily price for these fuel types was averaged for each month unless the standard deviation was greater than 10 cents. For March 2020, the standard deviation for both fuel types were approximately 40 cents, so daily prices were used instead. For April and May 2020, a 10-day average was sufficient to obtain a standard deviation of <10 cents. Using the calculated miles per gallon values, a monthly average was used to generate the price per mile values.

4.6. Predicting long-term effect of using E30 on CO_2 emission and ethanol consumption

Detailed calculations of how the values obtained for Fig. 5 are provided in Supplementary File 6. Briefly, typical fuel efficiency (in miles per gallon) and miles driven by a typical light-duty vehicle per year obtained from US Department of Transportation reports [31] were used to determine the average gallons consumed per year. These values were also calculated for E30 fuel using the 2.5% difference in efficiency determined in this study. CO_2 emission reduction was subsequently calculated based on 8.9 kg/gal and 5.77 kg/gal CO_2 emitted for gasoline and ethanol respectively.

Declaration of Competing Interest

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

Acknowledgments

We gratefully acknowledge funding support from Nebraska Ethanol Board (22-1106-0023).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fuel.2021.121629.

References

[1] U.S. Energy Information Administration. Annual Energy Outlook 2020 2019.

[2] Keith D, Houston S, Naumov S. Vehicle fleet turnover and the future of fuel economy. Environ Res Lett Feb. 2019;14:21001. https://doi.org/10.1088/1748-9326/aaf4d2.

[3] Al-Hasan M. Effect of ethanol-unleaded gasoline blends on engine performance and exhaust emission. Energy Convers Manag 2003;44(9):1547–61. https://doi.org/10.1016/S0196-8904(02)00166-8.

[4] Matějovský L, Macák J, Pospíšil M, Baroš P, Staš M, Krausová A. Study of Corrosion of Metallic Materials in Ethanol-Gasoline Blends: Application of Electrochemical

Methods. Energy Fuels Oct. 2017;31(10):10880–9. https://doi.org/10.1021/acs.energyfuels.7b01682.
[5] K. Knoll et al., Effects of Intermediate Ethanol Blends on Legacy Vehicles and Small Non-Road Engines, Report 1 - Updated, 2008.
[6] Khuong LS, Masjuki HH, Zulkifli NWM, Mohamad ENiza, Kalam MA, Alabdulkarem Abdullah, et al. Effect of gasoline–bioethanol blends on the properties and lubrication characteristics of commercial engine oil. RSC Adv 2017;7(25):15005–19. https://doi.org/10.1039/C7RA00357A.
[7] Çelik MB, Colak A. The use of pure ethanol as alternative fuel in a spark ignition engine. J Fac Eng Archit Gazi Univ Sep. 2008;23:619–26.
[8] Schifter I, Diaz L, Rodriguez R, Gómez JP, Gonzalez U. Combustion and emissions behavior for ethanol–gasoline blends in a single cylinder engine. Fuel 2011;90(12):3586–92. https://doi.org/10.1016/j.fuel.2011.01.034.
[9] Hsieh WD, Chen RH, Wu TL, Lin TH. Engine performance and pollutant emission of an SI engine using ethanol-gasoline blended fuels. Atmos Environ 2002;36(3):403–10. https://doi.org/10.1016/S1352-2310(01)00508-8.
[10] Wu Chan-Wei, Chen Rong-Horng, Pu Jen-Yung, Lin Ta-Hui. The influence of air-fuel ratio on engine performance and pollutant emission of an SI engine using ethanol-gasoline-blended fuels. Atmos Environ 2004;38(40):7093–100. https://doi.org/10.1016/j.atmosenv.2004.01.058.
[11] T. Topgül, H. Yücesu, C. ÇINAR, A. Koca, The effects of ethanol-unleaded gasoline blends and ignition timing on engine performance and exhaust emissions, 31 (Jan. 2006) 2534–2542.
[12] Chen R-H, Chiang L-B, Chen C-N, Lin T-H. Cold-start emissions of an SI engine using ethanol–gasoline blended fuel. Appl Therm Eng 2011;31(8):1463–7. https://doi.org/10.1016/j.applthermaleng.2011.01.021.
[13] He Bang-Quan, Jian-Xin Wang, Hao Ji-Ming, Yan Xiao-Guang, Xiao Jian-Hua. A study on emission characteristics of an EFI engine with ethanol blended gasoline fuels. Atmos Environ 2003;37(7):949–57. https://doi.org/10.1016/S1352-2310(02)00973-1.
[14] Razavian N, Knoll F, Geras KJ. Artificial Intelligence Explained for Nonexperts. Semin Musculoskelet Radiol Feb. 2020;24(1):3–11. https://doi.org/10.1055/s-0039-3401041.
[15] Warner Brad, Misra Manavendra. Understanding Neural Networks as Statistical Tools. Am Stat 1996;50(4):284–93. https://doi.org/10.1080/00031305.1996.10473554.
[16] Abiodun Oludare Isaac, Jantan Aman, Omolara Abiodun Esther, Dada Kemi Victoria, Mohamed Nachaat AbdElatif, Arshad Humaira. State-of-the-art in artificial neural network applications: A survey. Heliyon 2018;4(11):e00938. https://doi.org/10.1016/j.heliyon.2018.e00938.
[17] K.0G. Liakos, P. Busato, D. Moshou, S. Pearson, D. Bochtis, Machine Learning in Agriculture: A Review, Sensors 18(8) (2018). doi: 10.3390/s18082674.
[18] S. Ray, A Quick Review of Machine Learning Algorithms, in: 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), 2019, pp. 35–39, doi: 10.1109/COMITCon.2019.8862451.
[19] B. Mahesh, Machine Learning Algorithms -A Review, 2019.
[20] S.0S. Dash, S.K. Nayak, D. Mishra, A Review on Machine Learning Algorithms BT - Intelligent and Cloud Computing, 2021, pp. 495–507.
[21] S.S. Dash, S.K. Nayak, D. Mishra, A Review on Machine Learning Algorithms BT - Intelligent and Cloud Computing, 2021, pp. 495–507.
[22] S. Wickramanayake, D. Bandara, Fuel consumption prediction of fleet vehicles using Machine Learning: A comparative study. 2016.
[23] Schoen Alexander, Byerly Andy, Hendrix Brent, Bagwe Rishikesh Mahesh, Santos Euzeli Cipriano dos, Miled Zina Ben. A Machine Learning Model for Average Fuel Consumption in Heavy Vehicles. IEEE Trans Veh Technol 2019;68(7):6343–51. https://doi.org/10.1109/TVT.2510.1109/TVT.2019.2916299.
[24] S.L. Brunton, J.L. Proctor, J.N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proc Natl Acad Sci. Apr. 2016;113(15):3932 LP – 3937. doi: 10.1073/pnas.1517384113.
[25] Dragos T, Dumitru I, Oţăt O, Racila L, Geonea I, Rotea C. Oxygen Sensor Testing for Automotive Applications. Appl Mech Mater Feb. 2020;896:249–54. https://doi.org/10.4028/www.scientific.net/AMM.896.249.
[26] DuMont RJ, Evans JA, Feist DP, Studzinski WM, Cushing TJ. Test and Control of Fuel Injector Deposits in Direct Injected Spark Ignition Vehicles. SAE Int 2009. https://doi.org/10.4271/2009-01-2641.
[27] National Research Council, Renewable fuel standard: Potential economic and environmental effects of U.S. biofuel policy, Washington, DC, 2011. doi: 10.17226/13105.
[28] Scully MJ, Norris GA, Alarcon Falconi TM, MacIntosh DL. Carbon intensity of corn ethanol in the United States: state of the science. Environ Res Lett 2021;16(4):43001. https://doi.org/10.1088/1748-9326/abde08.
[29] EIA. Almost all U.S. gasoline is blended with 10% ethanol. 2016. https://www.eia.gov/todayinenergy/detail.php?id=26092 (accessed Jul. 04, 2021).
[30] EPA. Greenhouse Gas Emissions from a Typical Passenger Vehicle. 2018. https://www.epa.gov/greenvehicles/greenhouse-gas-emissions-typical-passenger-vehicle (accessed Jul. 04, 2021).
[31] USDOT. Average Annual Miles per Driver by Age Group. 2018. https://www.fhwa.dot.gov/ohim/onh00/bar8.htm (accessed Jul. 04, 2021).