Forecasting Fuel Ethanol Consumption

Colburn Hassman

Virginia Tech

Abstract

Accurately modeling fuel ethanol consumption is of utmost important to ethanol producers, merchants, and agricultural traders. This is especially true because biofuel consumption exhibits both trend and seasonality. This paper investigates multiple time series forecastig methods for estimating fuel ethanol consumption in the United States. The models analyzed range in complexity from mean forecasting method to Neural Network Autoregression. After specifying each model, time series cross validation techniques are used to test forecasting accuracy.

Keywords: Ethanol, Time Series Modelling, Fuel Consumption, Forecasting

Introduction

Unstable geopolitical situations culminating the 1973 US Oil Crisis prompted a reinvigoration of the biofuels sector in the United States. Dependance of foreign energy imports from unreliable or hostile trading partners became politically unconscionable. Biofuels were seen by some as an ideal solution; a completely domestic fuel which simultaneously supported rural America (Solomon, Barnes, & Halvorsen, 2007). As fossil fuel extraction technology has progressed, the United States has become a net exporter of crude oil. While US dependance on foreign oil has waned, enviromental and political activists continue to support ethanol for a myriad of reasons. A March 2019 study released by the USDA found that the life-cycle greenhouse gas emissions for corn ethanol were approximately 40% less than that of gasoline (Lewandrowski et al., 2020). In the March 2020 World Agriculture Supply and Demand Estimate (WASDE) Report, the USDA estimated that ethanol use accounted for 44% of US domestic corn consumption (NASS, 2020). A March 2018 report from the USDA attributed more than 86,000 jobs directly to the ethanol industry, along with 270,000

Correspondence concerning this article should be addressed to Colburn Hassman, Department of Agricultural & Applied Economics, Virginia Tech. E-mail: colburn7@vt.edu

indirect jobs (Jay S. Golden & True, 2018). Many of these jobs are in rural areas where ethanol remains a highly politiziced issue.

In support of the biofuels industry, the Energy Policy Act of 2005 established the Renewable Fuel Standard (RFS), mandating that fuel used for transportation contain a mininmum volume of renewable fuel. This legislation directed the EPA to administer a program to increase the amount of biofuels used for transportation from 4.0 billion gallons in 2006 to 36.0 billion gallons in 2022. Program compliance is regulated via tradable renewable identification number (RINs) credits, which are based on the energy content of the fuel (Solomon et al., 2007).

Growth in ethanol blended into gasoline has slowed in recent years. Currently the amount of ethanol blended into gasoline is around 10 percent, close to the so called "Ethanol Blendwall". Ethanol contains approximately 33% less energy per volume than pure gasoline(EIA, 2020). For this reasons and others, older vehicles cannot use fuel with more than 10% ethanol content by volume. In 2019, the EPA finalized regulatory changes to allow gasoline blended with up to 15 percent ethanol (E15) year round (EPA, 2019). This updated regulation is unlikley to translate to an immediate increase in consumption, as the EPA has previously identify the lack of ethanol infastructure as the binding constraint for ethanol consumption growth (EPA, 2017).

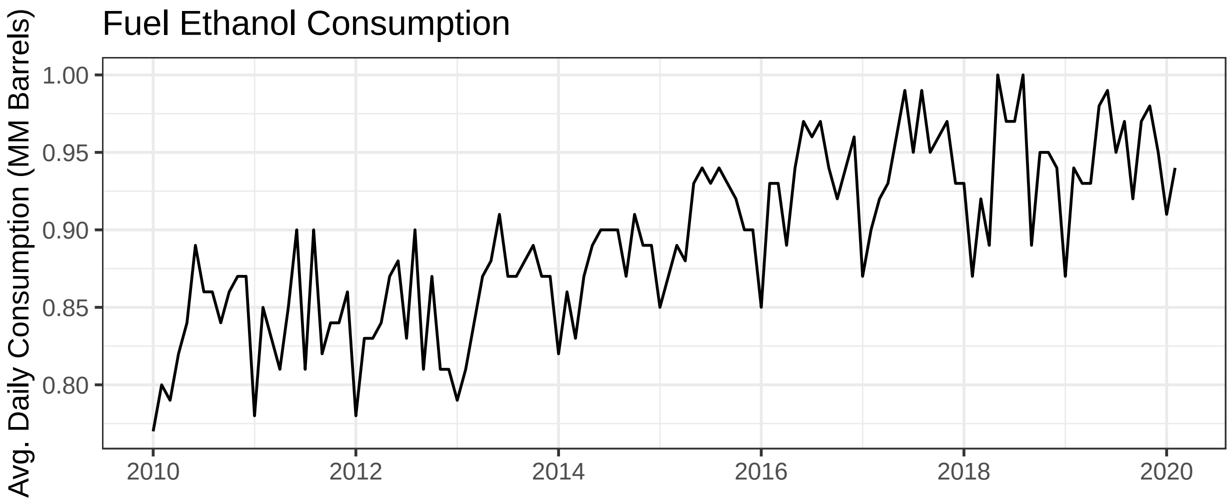
In support of the oil industry and to the frustration of biofuels, certain energy refiners are exempt from the RIN credit mandate through Small Refinery Exceptions (SREs). Reducing RIN credit demand means lowering RIN credit prices, which reduces the incentive to increase ethanol blending, which harms corn demand. In 2017, the EPA granted 1.63 billion gallons worth of exemptions. The Renewable Fuel Association estimated that this reduced corn consumption by 1.96 billion dollars (assuming 3.50 dollars per bushel) (Association, 2018).

Future growth in ethanol consumption may be brought about by increased E85 consumption which in turn may be brought about by infastructure expansion and favorable relative prices. Currently, the majority of E85 is consumed in PADD 2, which encompasses the Midwest United States, where the vast majority of ethanol is produced and where the majority of US corn is grown. Growing ethanol exports and decreasing the amount of SREs may serve as an additional source of demand.

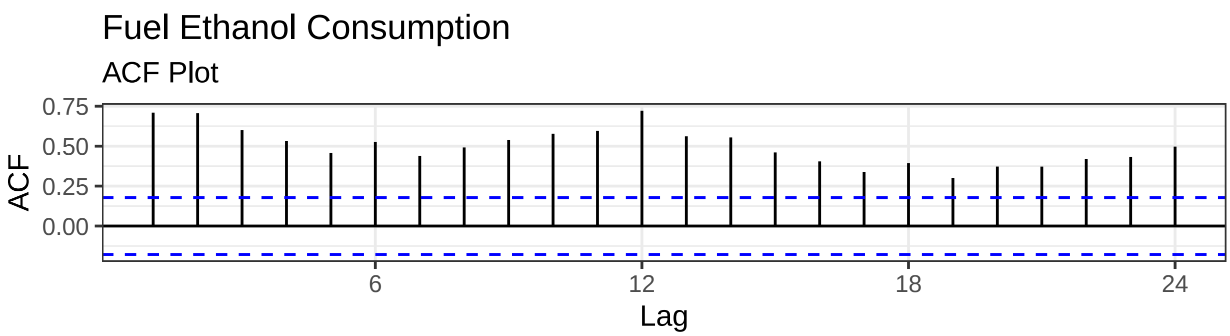
This paper examines multiple models to forecast fuel ethanol consumption in the United States with intent of comparing their accuracy for multiple forecasting horizons. Accurately forecasting ethanol consumption is an important task for ethanol producers, merchants, regulators, corn market participants, and exporters. Because the a large portion of US corn is processed into ethanol, and the United States is such a large participant in global corn markets, international participants also have incentive to accurately forecast consumption.

The following is organizes in multiple sections. The first section, Data, explores the dataset used in this analysis. The second, Model Descriptions, specifys and discusses each model used in the analysis. The Accuracy Measurement section describes the methodology and metrics for evaluating forecasting effectiveness. The Results section shares the findings of the research while the final section, c Conclusion, discusses the results suggests additional research topics.

Data

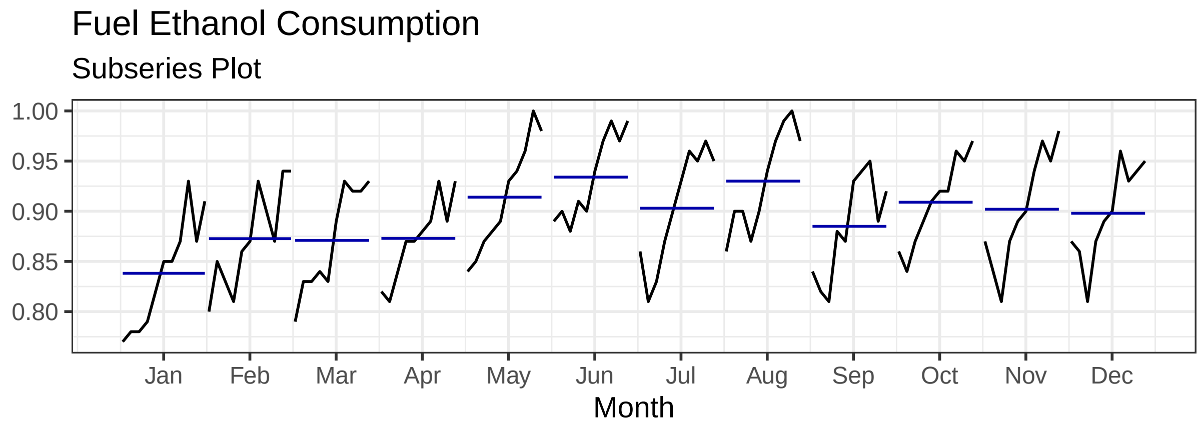


Visualization 1: Fuel Ethanol Consumption



Visualization 2: Fuel Ethanol Consumption ACF Plot

The data used is monthly fuel ethanol consumption reported in million barrels per day by the US energy Information Administration. (U. EIA, 2020) The data spanned from January 2010 to February 2020 (n=122). Based on the ACF and subseries plots, the data exhibits seasonality for which an accurate model will need to account for.



Visualization 3: Fuel Ethanol Consumption Subseries Plot

Model Descriptions

The following are analysis of the ten models used in this paper. For each model, a formal mathematical specification is offered, as well as a discussion of the merits and shortcomings of the model type.

(i) Average Method

The forecast of all future values are simply the average of the historical data. Best used for stochastic mean reverting data. This method is one of the simpliest and is often used as a benchmark to compare more complex models.This model is inappropriate for data displaying trend, and generally a poor fit for data displaying seasonality.

$$\hat{y}_{T+h|T} = \overline{y} = \frac{(y_1 + ... + y_T)}{T}$$

(ii) Naive Method

The Naive Methods uses the previous observation at the forecast. Along with the average method, it serves as the benchmark to judge all other forecasts by. The naive method is often a good forecasting method over very short forecasting horizons, but the accuracy tends to quickly diminish as the horizon is extended.

$$\hat{y}_{T+h|T} = y_T$$

(iii) Random Walk Method with Drift

A random walk is a stochastic process. Like the naive, we generally assume that the best forecast for $t = t_{t-1}$. In this case however, we account for a trend (drift) in the data. This method accounts for trend but not seasonality.

$$\hat{y}_{T+h|T} = y_T + \frac{h}{T-1}\sum_{t=2}^{T}(y_t - y_{t-1}) = y_T + h\left(\frac{y_T - y_1}{T-1}\right)$$

(iV) Seasonal Naive

The seasonal naive method is very useful for highly seasonal data. For each future season, it uses the previous season's value as the forecast. This model is appropriate for highly seasonal data with little trend. The simplicity of this model means that the data requirement is small.

$$\hat{y}_{T+h|T} = y_{T+h-m(k+1)}$$

(v) ETS

The *Error, Trend, Seasonal* (ETS) forecasting method is a flexible framework which decomposes the timeseries into distinct components.

$$y_t = T_t + S_t + R_t$$

It's flexibility comes from the ability to account for Additive and Multiplicative Errors, Trends, and Seasons, as well as dampened Trends.

ETS Model

Specifications

Trend type	No Seasonality	Additive Seasonality	Multiplicative Seasonality
N (None)	NN	NA	NM
A (Additive)	AN	AA	AM
Ad (Additive Damped)	AdN	AdA	AdM
M (Multiplicative)	MN	MA	MM
Md (Multiplicative Damped)	MdN	MdA	MdM

Table 1: ETS Model Specifications

The proper model specification can be automatically choosen using the AIC.

(vi) Holt

The Holt method is a linear trend forecasting method based on simple exponential smoothing.

$$\begin{aligned} \text{Forecast equation} \qquad & \hat{y}_{t+h|t} = \ell_t + hb_t \\ \text{Level equation} \qquad & \ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\ \text{Trend equation} \qquad & b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}, \end{aligned}$$

l_t is the estimate of the level of the series at time t and β_t is the estimate of the trend of the series. α is the smoothing parameter which is between 0 and 1. The larger the α, the more weight the model puts on recent variables. Using exponential smoothing often more accurately captures changing trends. A dampened variation of this model is used when linear forecasts are inappropriate.

(vii) Holt-Winter

The Holt-Winter model extends the original holt model to include a seasonality parameter γ over m seasons. Here is both an additive and multiplicative method depending on the type of seasonality.

$$\begin{aligned} \hat{y}_{t+h|t} &= \ell_t + hb_t + s_{t+h-m(k+1)} \\ \ell_t &= \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\ b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1} \\ s_t &= \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}, \end{aligned}$$

where $k = \frac{h-1}{m}$

γ is restricted to $0 \leq \gamma \leq 1 - \alpha$

(viii) ARIMA(p,d,q)

ARIMA models are the differenced combination of autoreggressive and moving average models.

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{1}$$

where y' is the differenced series. These types of models are defined by three paramters (p,d,q).

- $p =$ order of the autoregressive part
- $d =$ degree of first differencing involved
- $q =$ order of the moving average part

There are several cases where ARIMA specifications are identical to models previously discussed.

ARIMA Model Specifications

Special Cases

Model	Specification
White Noise	ARIMA(0,0,0)
Random Walk	ARIMA(0,1,0) with no constant
Random Walk with Drift	ARIMA(0,1,0) with a constant
Autoregression	ARIMA(p,0,0)
Moving Average	ARIMA(0,0,q)

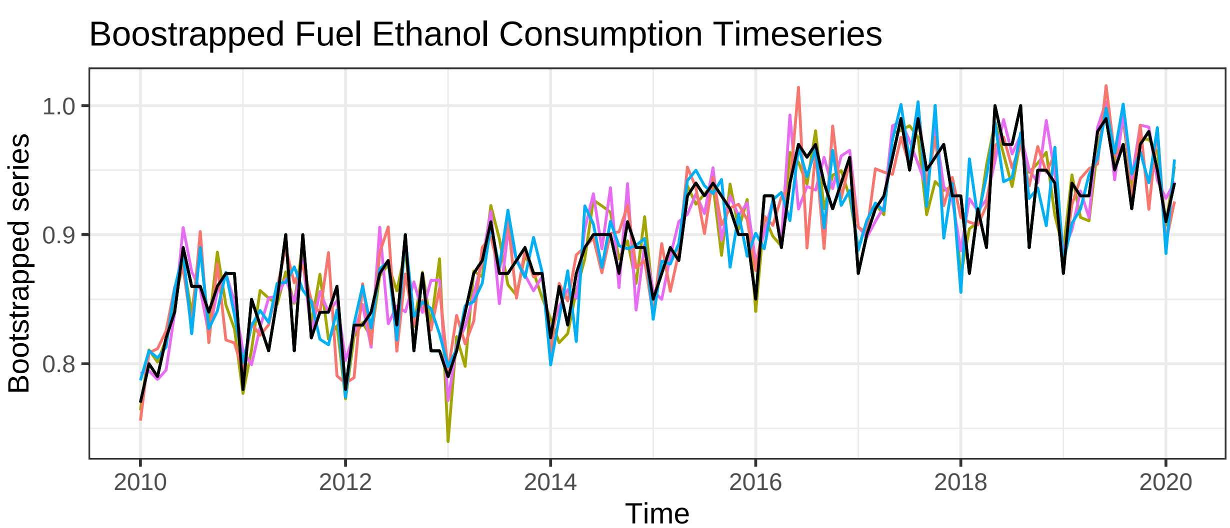
Table 2: ARIMA Model Specifications

The ARIMA framework is especially usefil because of its flexibility. The auto.arima function in the fpp2 function automatically selects the model specification to maximize AIC.

(ix) Bootstrapping

Bootstrapping time series refers to generating new time series which have similiar charactersitics as our original time series. This is often use to calculate better prediction intervals, as most prediction intervals from time series models are too narrow.Bootstrapping allows for uncertainty not only in the random error term, but also in the parameter estimates.

We can also use bootstrapping to improve forecast accuracy by producing forecasts from the bootstrapped timeseries and average together the forecasts. This is called bagging (bootstrap aggregating) and has been show to produce more accurate models.



Visualization 4: Bootstrapped Timeseries

Computational constraints limit the number of simulations run, but the method is included as an improvement of a more basic model.

(x) Neural Networks

Nueral Networks function by introducing a hidden layer of "neurons" to the model. This allows modeling of complex, nonlinear relaitonships.

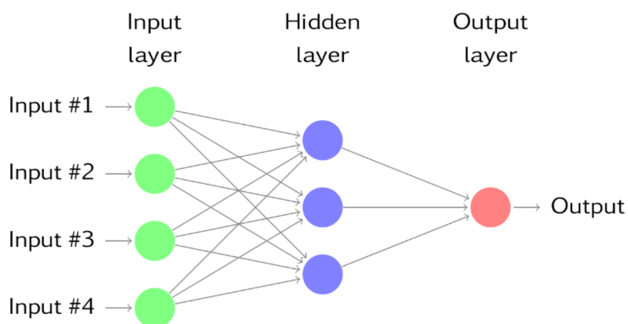


Figure 1. Neural Network Diagram

The figure above depicts a *multilayer feed-forward network* where each layer of nodes recieved inputs from the precious layer. The inputs from each node are weighted and linearly combined. The result is then modified by a nonlinear function. The values of the weightings are often restricted by a "decay parameter" which is often set to 0.1.

In timeseries data, lagged variables of the time series can be used as inputs to a neural network to create a Neural Network Autoregression (NNAR) Model, the nonlinear cousin of a linear autoregression model. Generally, the model is defined as

$$NNAR(p, P, k)m$$

where $p =$ the number of lagged inputs, P indicates seasonality, and k is the number of neurons in the hidden layer.

The function *nnetar()* will fit an NNAR model automatically, setting $P = 1$ for seasonal data, choosing the optimal number of lags (p) based on AIC, and setting $k = \frac{p+P+1}{2}$ rounded to the nearest integar. Forecasting is done iteratively.

The benefit of neural networks is that it can easily incorporate nonlinear and complex interactions which other models may be unable to include. The costs of modeling these intereactions is significant data needs and a model opacity.

Accuracy Measurement

I have choosen a forecast horizon of 12 months (h = 12). This insures that the model forecasts are full cycle of seasonality.

Rather than divide the data into a fitting and validation set to measure model accuracy, Time Series Cross Validation (TSCV) is employed. TSCV evaluates the forecast accuracy over a rolling forecasting origin, with more and more of the data being included in the training set.

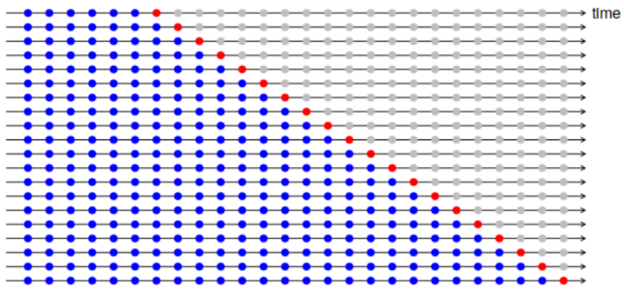


Figure 2. Illustration of Time Series Cross Validation

The tsCV function provided in the fpp2 is very useful, but it is necessary to make some minor alterations to be able to use with certain models which do not return a forecast set.

Using tscv increases our confidence in the models as it is equivalent to having the maximum amount of training and testing sets given data constraints.

Accuracy Metrics

Four metrics are used to measure forecasting accuracy; Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Percent Error (MAPE). Lower error metric values indicate a better goodness-of-fit.

Mean Absolute Error.

$$MAE = \frac{\sum_{t=1}^{T} |\hat{y}_t - y_t|}{t}$$

Mean Square Error.

$$MSE = \frac{1}{T} \sum_{t=1}^{T} (\hat{y}_t - y_t)^2$$

Root Mean Square Error.

$$RMSE = \sqrt{\sum_{t=1}^{T} \frac{(\hat{y}_t - y_t)^2}{T}}$$

Root Mean Square Error is the standard deviation of the errors.

Mean Absolute Percent Error.

$$MAPE = \frac{1}{t} \sum_{t=1}^{T} |\frac{y_t - \hat{y}_t}{y_t}|$$

The MAPE is the average absolute error as a percent, of the actual value.

Results

Goodness-of-fit Metrics

In-Sample

Models	RMSE	MAPE
Mean	0.05489	5.16695
Naive	0.04026	3.50371
Random Walk	0.04023	3.5075
Seasonal Naive	0.03106	2.86062
ETS	0.01989	1.73669
Holt	0.03392	3.02958
Holt-Winters	0.02005	1.83045
ARIMA	0.01956	1.62682
Bagged ETS	0.01914	1.70184
NNAR	0.01914	1.70184

Table 3: Goodness-of-fit Metric Comparision

Using the tscv method, the ETS, bagged ETS, Holt, Holt-Winter, and ARIMA model specifications are automatically choosen for each training set iteration to maximize the Akaike Infromation Criterion (AIC). Thus, the model specifications and parameter estimates will vary with the data used to train. Certain specifications were made beforehand. For example, the ETS models were specified to account for additive seasonality. This was choosen because the maginitude of seasonality displayed by the data does not vary significantly within the sample.

Using the entire dataset to fit a model, the ETS model includes seasonality. Thus the final model specifications are

$$\begin{aligned} \text{Forecast equation} \qquad & \hat{y}_t = l_{t-1} + b_{t-1}s_{t-m} + \epsilon_t \\ \text{Level equation} \qquad & l_t = l_{t-1} + b_{t-1} + \alpha\epsilon_t \\ \text{Trend equation} \qquad & b_t = b_{t-l} + \beta\epsilon_t \\ \text{Seasonality Equation} \qquad & s_t = s_{t-m} + \gamma\epsilon_t, \end{aligned}$$

With parameter estimates of $\alpha = 0.1248$, $\beta = 0.0001$, and $\gamma = 0.0153$, and $m = 12$.

The Holt model has smoothing parameter estimates of $\alpha = 0.2224$ and $\beta = 0.0001$. The Holt-Winters Model, which like accounts for seasonality, yields smoothing parameter estimates of $\alpha = 0.0995$, $\beta = 0.0134$, and $\gamma = 0.0004$.

The ARIMA model specification which maximizes AIC accounts for seasonality and is defined as ARIMA(0,1,1)(0,1,1) with $m = 12$.

The preferred model based on the in-sample accuracy statistics in the ARIMA model, which has the smallest errors. Models which include seasonal components have better goodness-of-fits than models which do not. The forecasting method with the poorest accuracy metrics is the mean method, which is unsurprising given both seasonality and trend.

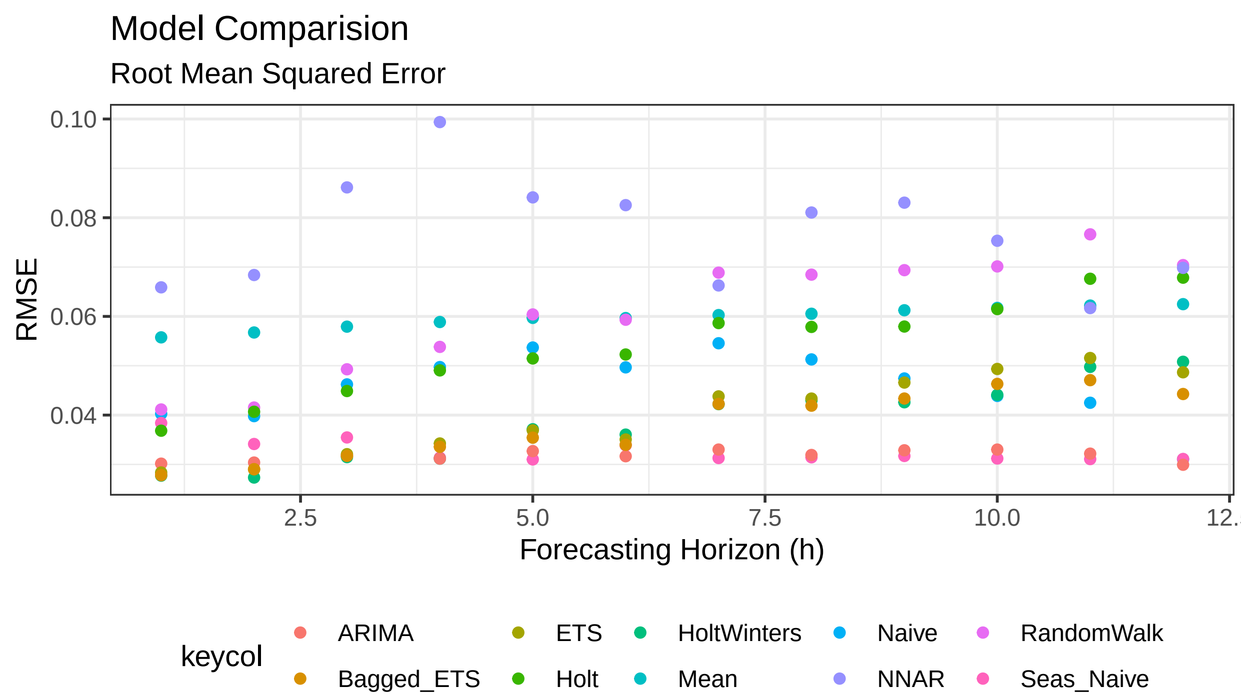
TSCV Accuracy Statistics

Model Comparision

Model	MAE	MSE	RMSE
Mean	0.05098	0.00357	0.05976
Naive	0.03737	0.00214	0.04584
Random Walk	0.04575	0.00382	0.06078
Seasonal Naive	0.02687	0.00107	0.03268
ETS	0.02748	0.00166	0.03991
Holt	0.03851	0.00299	0.05388
Holt Winters	0.0258	0.00155	0.03862
ARIMA	0.02476	0.00101	0.03173
Bagged ETS	0.02687	0.00149	0.03806
NNAR	0.06311	0.00603	0.07697

Table 4: Cross-Validation Accuracy Metrics

The above table includes averages of the goodness-of-fit metrics calculated from tscv. It is the average of the error from all forecasting horizons from all the training set iterations used. Based on this, the preferred model based on best fit is again the ARIMA model, followed by the Bagged ETS model.



Visualization 5: TSCV RMSE Forecast Horizon Plot

Seperating the errors by forecasting horizon, it is clear that there is no single best forecasting model. Instead, the most accurate model depends on the forecasting time horizon. The most accurate forecasts are those which explicitly include seasonality: ARIMA, ETS, Bagged ETS, Holt-Winters, and Seasonal Naive. For longer time horizons, the seasonal naive method, one of the most simple forecasting method, is the most accurate forecasting method. The phemonenon of random walk methods being perferred over longer time horizons has been repeatedly noted in the literature (Kilian & Taylor, 2003). Given that this method simply takes the previous value from the same season, we would expect that the accuracy would go down with stronger trend.

The bagged ETS method improves in-sample goodness of fit dispite few simulations being conducted due to computation constraints. The accuracy of this method will likely increase if more bootstrapped time series are generated.

Neural Network Autoregression (NNAR), despite being perhaps the most complex model used, does a poor job forecasting. Model accuracy would likely improve if more data was included in the training set.

The differences between the preferred in-sample model versus the preferred TSCV model illustrate the importance of validating forecasting models using training and testing sets to forecast out-of-sample. They also may indicate the need to have more data as model complexity grows. Further research may explore relationships between model complexity and improvement with additional training data. Heterogeneity of perferred methods across forecasting horizons illustrates the importance of that consideration. There is no single best method, and while more complex models may increase short term forecasting accuracy, the results support using a random walk technique (which includes seasonality) over longer forecasting horizons.

Conclusion

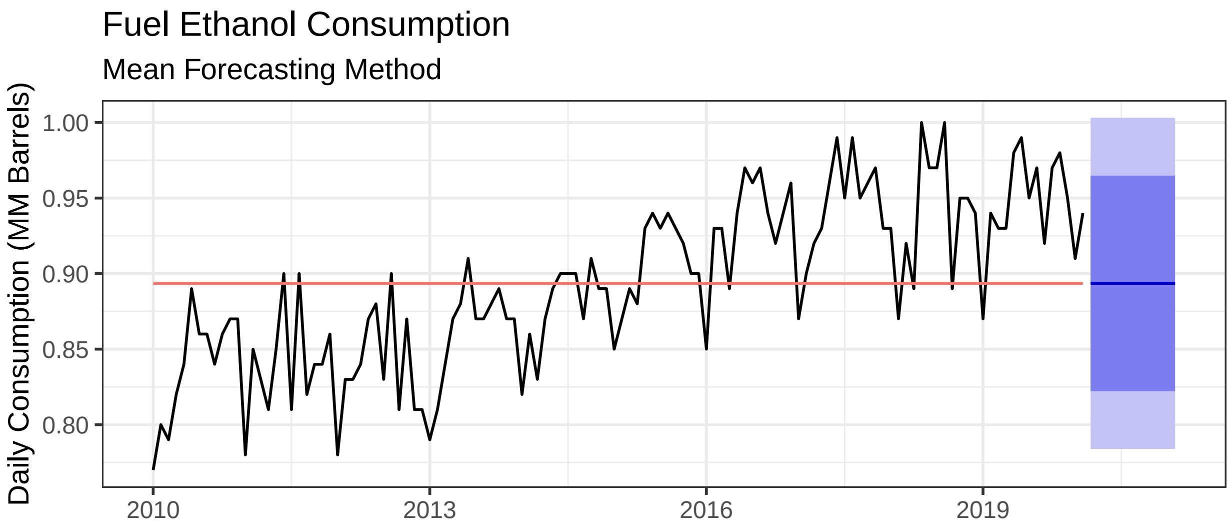
This paper analyzes and compares the ability of ten time series models in forecasting US fuel ethanol consumption. After each model was defined and discussed, forecasting ability was judged using time series cross validation.

The results support previous indications of random walk models being the preferred method of forecasting over longer time horizons. More complex models, such as the ARIMA and bagged ETS have lower errors in shorter forecasting horizons, but lose accuracy as the horizon is extended. The most complex forecasting method, Neural Network Autoregression, had very high error. Accuracy would likely increase significantly as the size of the dataset increases.

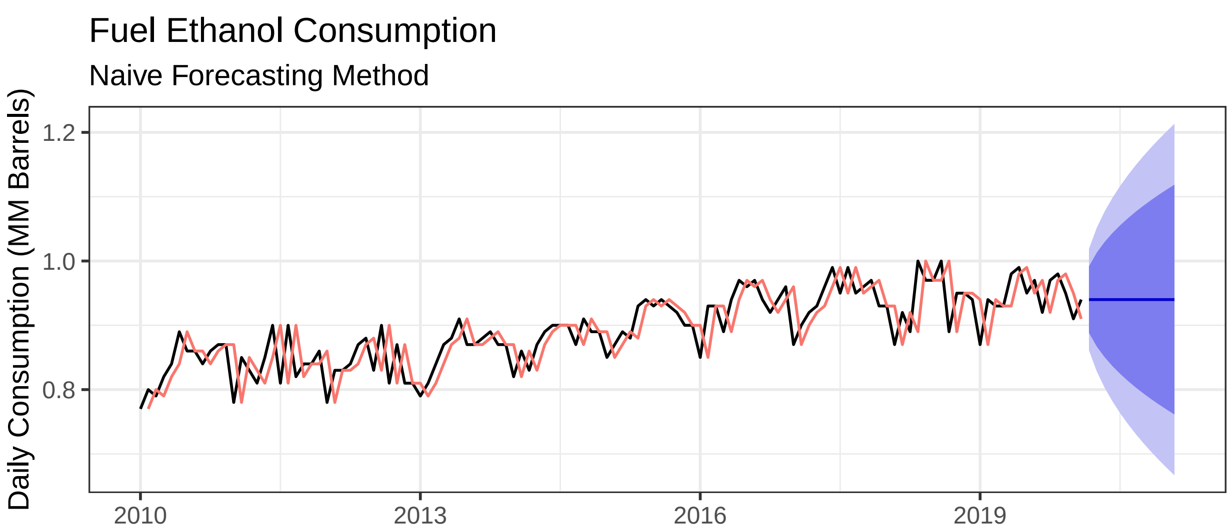
These findings are significant for many stakeholders of the ethanol industry, including producers, merchants, and traders. The results illustrate the importance of proper model choice and specification and the possibility of more accurate short-term forecasts with more complex models.

Further research may increase the data set size, explore fuel ethanol consumption in other markets, or examine other forecasting methods. More accurate forecasting models may include contemporary and lagged values of relevant data, such as disposable income which may explain the total fuel consumption, or the gasoline/ethanol price ratio, which may explain ethanol blending decisions.

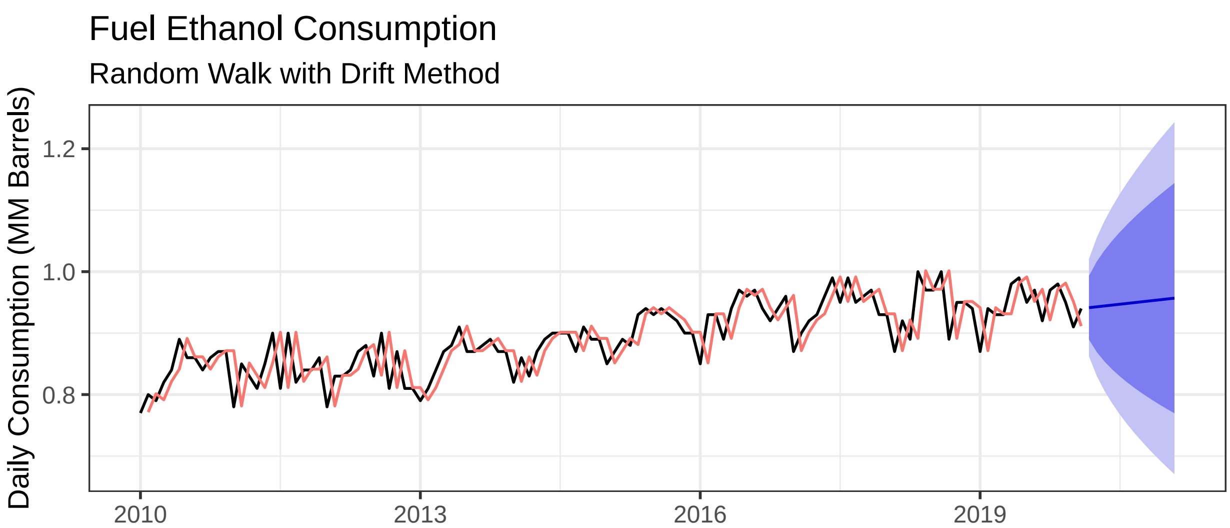
Appendix



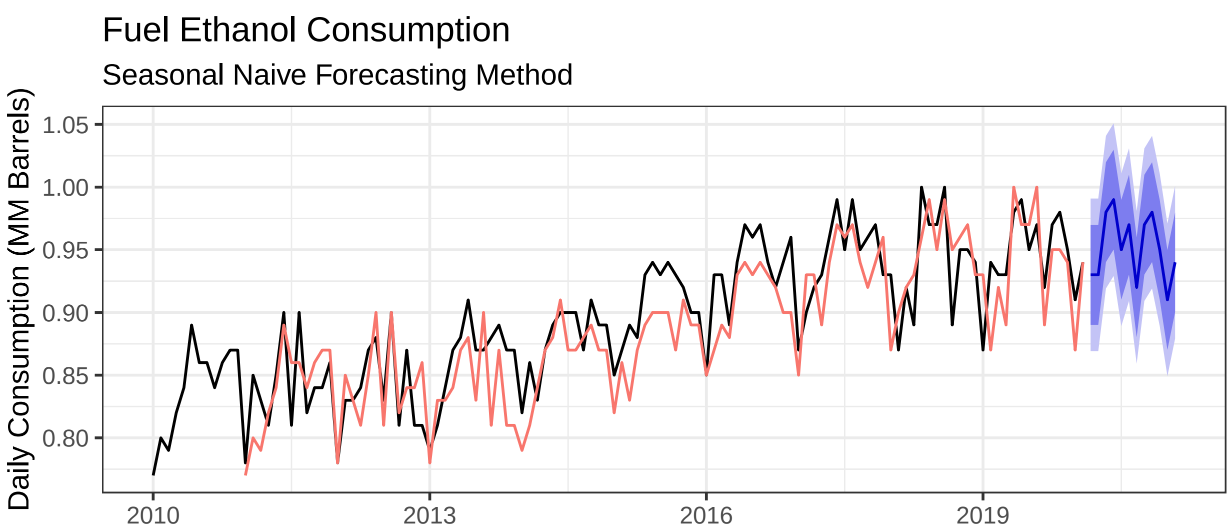

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
```

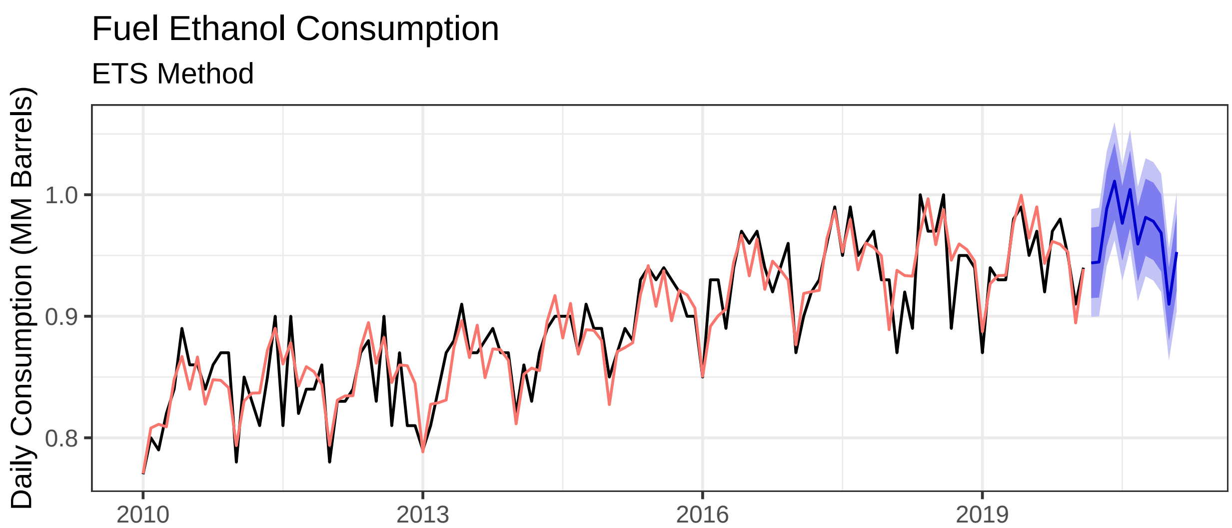


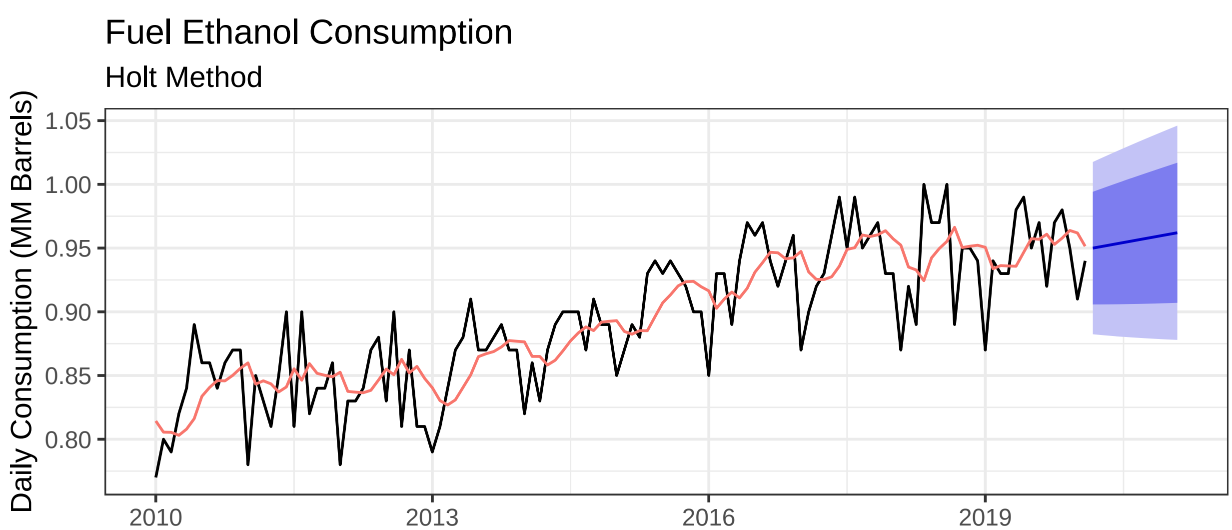

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
```

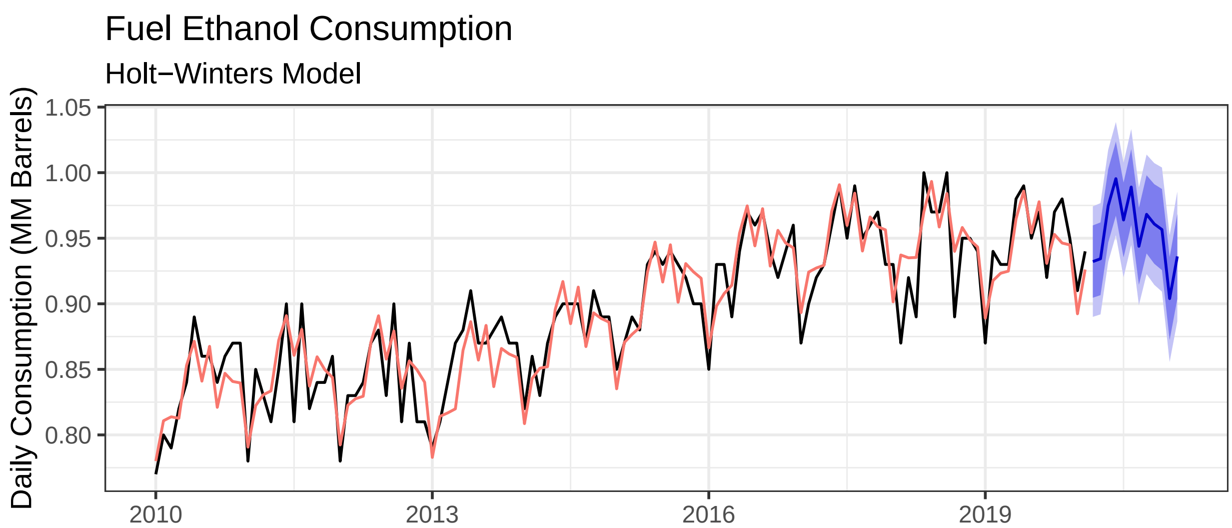


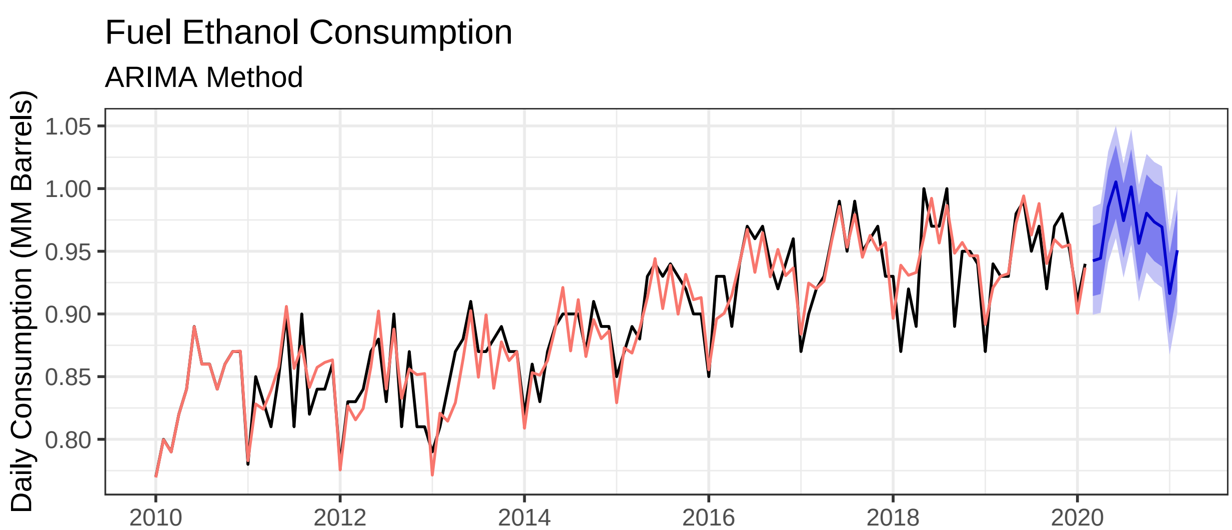

```
## Warning: Removed 12 row(s) containing missing values (geom_path).
```



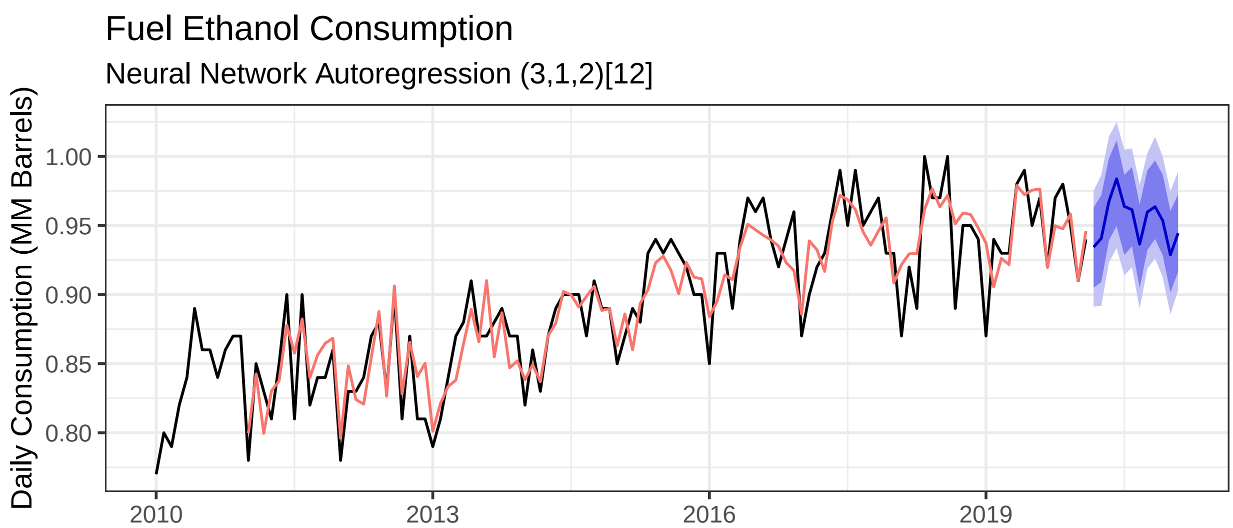

Fuel Ethanol Consumption
ETS Method
Daily Consumption (MM Barrels)
1.0
0.9
0.8
2010
2013
2016
2019

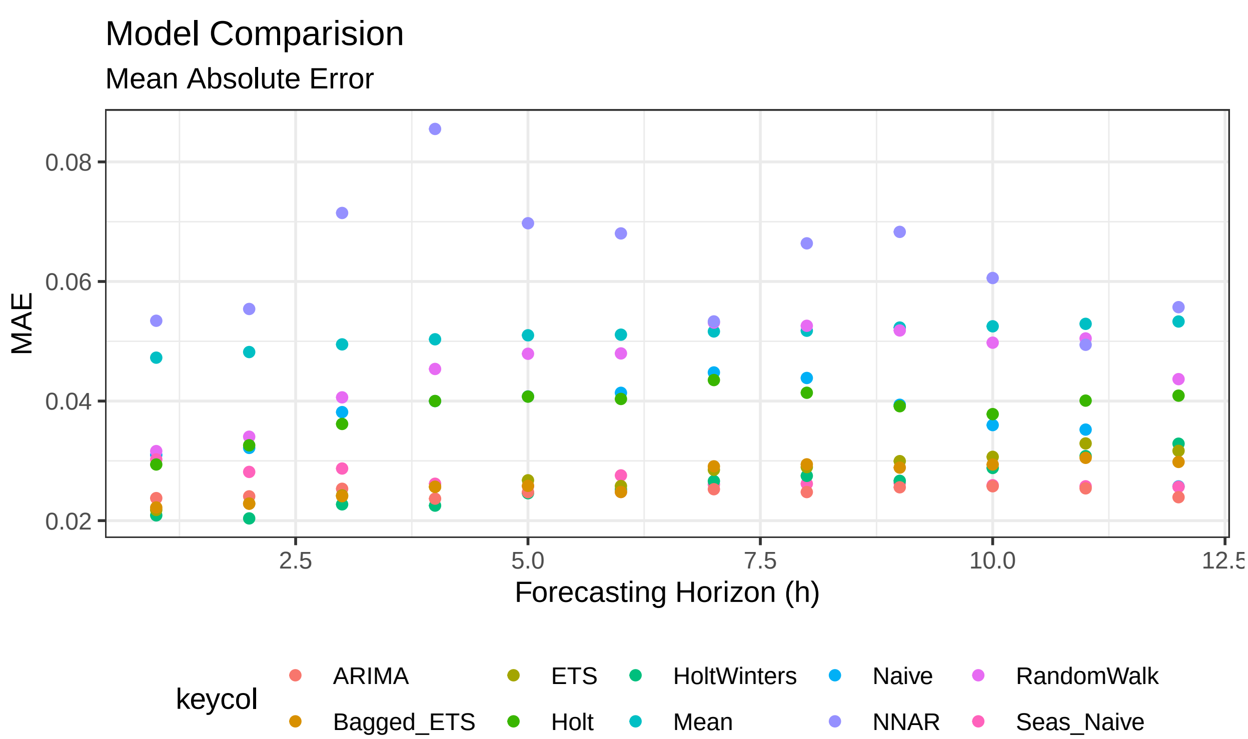

Fuel Ethanol Consumption
Holt Method
Daily Consumption (MM Barrels)
1.05
1.00
0.95
0.90
0.85
0.80
2010
2013
2016
2019

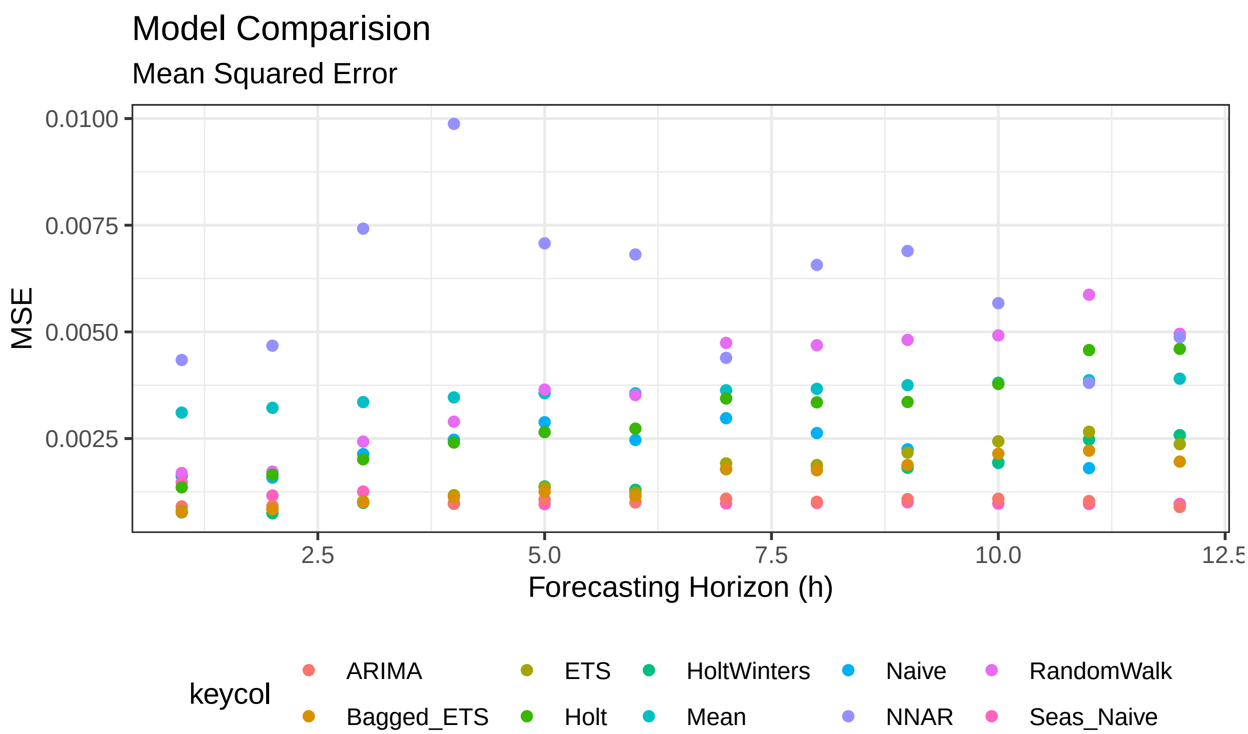





```
## Warning: Removed 12 row(s) containing missing values (geom_path).
```


Fuel Ethanol Consumption
Neural Network Autoregression (3,1,2)[12]
Daily Consumption (MM Barrels)
1.00
0.95
0.90
0.85
0.80
2010
2013
2016
2019


Model Comparision
Mean Absolute Error
MAE
0.08
0.06
0.04
0.02
2.5
5.0
7.5
10.0
12.5
Forecasting Horizon (h)
keycol
ARIMA
Bagged_ETS
ETS
Holt
HoltWinters
Mean
Naive
NNAR
RandomWalk
Seas_Naive




```
## ETS(M,A,M)
##
## Call:
##  ets(y = data_ts)
##
##   Smoothing parameters:
##     alpha = 0.1248
##     beta  = 1e-04
##     gamma = 0.0153
##
##   Initial states:
##     l = 0.8239
##     b = 0.0012
##     s = 0.9956 1.0076 1.0114 0.9906 1.0376 1.0102
##            1.0476 1.0251 0.9809 0.9816 0.978 0.9339
##
##   sigma:  0.024
##
##       AIC      AICc       BIC
## -335.1824 -329.2978 -287.5141
##
## Training set error measures:
##                         ME       RMSE        MAE         MPE     MAPE      MASE
## Training set -0.0004159376 0.01988825 0.01535063 -0.09101473 1.73669 0.5966678
```

```
##                    ACF1
## Training set -0.1000303
```

```
##
## Forecast method: Holt's method
##
## Model Information:
## Holt's method
##
## Call:
##  holt(y = data_ts)
##
##   Smoothing parameters:
##     alpha = 0.2224
##     beta  = 1e-04
##
##   Initial states:
##     l = 0.8132
##     b = 0.0011
##
##   sigma:  0.0345
##
##       AIC      AICc       BIC
## -229.5398 -229.0226 -215.5197
##
## Error measures:
##                        ME       RMSE       MAE         MPE     MAPE    MASE
## Training set 5.47503e-05 0.03392065 0.0267965 -0.1172944 3.029576 1.04156
##                   ACF1
## Training set 0.1251007
##
## Forecasts:
##          Point Forecast     Lo 80     Hi 80     Lo 95    Hi 95
## Mar 2020      0.9499781 0.9057764 0.9941798 0.8823774 1.017579
## Apr 2020      0.9510729 0.9057904 0.9963553 0.8818193 1.020326
## May 2020      0.9521676 0.9058286 0.9985066 0.8812983 1.023037
## Jun 2020      0.9532624 0.9058895 1.0006352 0.8808119 1.025713
## Jul 2020      0.9543571 0.9059716 1.0027426 0.8803579 1.028356
## Aug 2020      0.9554519 0.9060736 1.0048302 0.8799343 1.030969
## Sep 2020      0.9565467 0.9061943 1.0068990 0.8795393 1.033554
## Oct 2020      0.9576414 0.9063326 1.0089502 0.8791714 1.036111
## Nov 2020      0.9587362 0.9064876 1.0109848 0.8788289 1.038643
## Dec 2020      0.9598309 0.9066584 1.0130035 0.8785105 1.041151
```

```
##
```

```
## Forecast method: Holt-Winters' additive method
##
## Model Information:
## Holt-Winters' additive method
##
## Call:
##  hw(y = data_ts)
##
##   Smoothing parameters:
##     alpha = 0.0995
##     beta  = 0.0134
##     gamma = 1e-04
##
##   Initial states:
##     l = 0.8338
##     b = 4e-04
##     s = -9e-04 0.0041 0.0122 -0.0115 0.0345 0.01
##            0.0423 0.0223 -0.0174 -0.0188 -0.0227 -0.0542
##
##   sigma:  0.0215
##
##       AIC      AICc       BIC
## -333.8696 -327.9849 -286.2012
##
## Error measures:
##                        ME       RMSE       MAE         MPE     MAPE     MASE
## Training set 0.0001862605 0.02004706 0.0161787 -0.01484213 1.830449 0.628854
##                     ACF1
## Training set -0.03324773
##
## Forecasts:
##          Point Forecast     Lo 80     Hi 80     Lo 95     Hi 95
## Mar 2020      0.9321851 0.9046229 0.9597473 0.8900324 0.9743378
## Apr 2020      0.9343278 0.9065904 0.9620652 0.8919071 0.9767485
## May 2020      0.9746972 0.9467419 1.0026524 0.9319433 1.0174510
## Jun 2020      0.9954190 0.9671994 1.0236385 0.9522608 1.0385771
## Jul 2020      0.9638889 0.9353550 0.9924229 0.9202500 1.0075278
## Aug 2020      0.9890428 0.9601413 1.0179443 0.9448418 1.0332438
## Sep 2020      0.9438298 0.9145050 0.9731546 0.8989814 0.9886782
## Oct 2020      0.9682243 0.9384181 0.9980304 0.9226397 1.0138088
## Nov 2020      0.9608650 0.9305177 0.9912123 0.9144528 1.0072772
## Dec 2020      0.9565961 0.9256466 0.9875457 0.9092629 1.0039294
## Jan 2021      0.9039539 0.8723401 0.9355676 0.8556048 0.9523030
## Feb 2021      0.9361431 0.9038028 0.9684835 0.8866829 0.9856034
## Mar 2021      0.9408210 0.9076910 0.9739509 0.8901531 0.9914888
```

```
## Apr 2021      0.9429637 0.9089827 0.9769447 0.8909942 0.9949331
## May 2021      0.9833330 0.9484395 1.0182266 0.9299680 1.0366981
## Jun 2021      1.0040548 0.9681881 1.0399215 0.9492014 1.0589082
## Jul 2021      0.9725248 0.9356253 1.0094242 0.9160920 1.0289575
## Aug 2021      0.9976786 0.9596882 1.0356691 0.9395773 1.0557800
## Sep 2021      0.9524657 0.9133273 0.9916040 0.8926087 1.0123226
## Oct 2021      0.9768601 0.9365184 1.0172019 0.9151627 1.0385575
## Nov 2021      0.9695008 0.9279017 1.0110999 0.9058805 1.0331212
## Dec 2021      0.9652320 0.9223231 1.0081408 0.8996085 1.0308554
## Jan 2022      0.9125897 0.8683203 0.9568592 0.8448854 0.9802940
## Feb 2022      0.9447790 0.8990996 0.9904584 0.8749184 1.0146396
```

```
## Series: data_ts
## ARIMA(1,0,2)(2,1,0)[12] with drift
##
## Coefficients:
##          ar1      ma1     ma2     sar1     sar2   drift
##       0.9266  -0.9188  0.1834  -0.5793  -0.4438  0.0011
## s.e.  0.0555   0.1087  0.0974   0.0942   0.1004  0.0003
##
## sigma^2 estimated as 0.0005593:  log likelihood=255.19
## AIC=-496.37   AICc=-495.27   BIC=-477.47
##
## Training set error measures:
##                         ME       RMSE        MAE         MPE     MAPE      MASE
## Training set -0.0003350672 0.02183445 0.01641733 -0.08018675 1.837564 0.6381296
##                     ACF1
## Training set 0.003057837
```

Reference

Association, R. R. F. (2018). https://ethanolrfa.org/wp-content/uploads/2018/06/Small-Refiner-Exemptions-Economic-Impact-pdf.pdf.

EIA. (2020). How much ethanol is in gasoline, and how does if effect fuel economy? https://www.eia.gov/tools/faqs/faq.php?id=27&t=10.

EIA, U. (2020). Short-term energy outlook. https://www.eia.gov/outlooks/steo/.

EPA. (2017). Final rulemaking for modifications to fuel regulations to provide flexibility for e15 and to elements of the renewable identification number compliance system. https://www.govinfo.gov/content/pkg/FR-2017-07-21/html/2017-14632.htm.

EPA. (2019). Final rulemaking for modifications to fuel regulations to provide flexibility for e15 and to elements of the renewable identification number compliance system. https://www.eia.gov/tools/faqs/faq.php?id=27&t=10.

Jay S. Golden, J. P.-G., Robert Handfield, & True, E. (2018). Indictors of the u.s. Biobased economy. https://www.usda.gov/oce/energy/files/BIOINDICATORS.pdf.

Kilian, L., & Taylor, M. P. (2003). Why is it so difficult to beat the random walk forecast of exchange rates? *Journal of International Economics*, *60*(1), 85–107. https://doi.org/https://doi.org/10.1016/S0022-1996(02)00060-0

Lewandrowski, J., Rosenfeld, J., Pape, D., Hendrickson, T., Jaglo, K., & Moffroid, K. (2020). The greenhouse gas benefits of corn ethanol – assessing recent evidence. *Biofuels*, *11*(3), 361–375. https://doi.org/10.1080/17597269.2018.1546488

NASS. (2020). World agricultural supply and demand estimates. https://www.usda.gov/oce/commodity/wasde/.

Solomon, B., Barnes, J., & Halvorsen, K. (2007). Grain and cellulosic ethanol: History, economics, and energy policy. *Biomass and Bioenergy*, *31*, 416–425. https://doi.org/10.1016/j.biombioe.2007.01.023