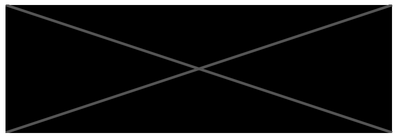
TSA Final Project



Introduction

Electric power production from fossil fuels accounts for 25 percent of the United States' greenhouse gas emissions, or over 1.6 billion metric tons of CO2 equivalent, as of 2019. Furthermore, 62 percent of all electricity comes from fossil fuels. While this has decreased by 12 percent since 1990 due to greater implementation of renewable energy generation and higher efficiency from the energy consumption side, many greater efforts in renewable energy integration are needed to reach the future goal of net-zero greenhouse gas emissions. Certain policy and economic trends have helped to encourage this growth in the past, such as the Investment Tax Credit (ITC), which promoted the installation of utility scale solar, and the Production Tax Credit (PTC), which promoted the generation of electricity from wind turbines. From 2019 to 2020, wind increased by 14 percent, utility scale solar grew by 26 percent, and small-scale solar increased by 19 percent. Meanwhile, coal declined 20 percent. Thus, as the US moves towards achieving the goal of net-zero emissions, we plan to model how renewable energy integration will look one year from today based on historical trends of renewable energy growth.

Motivation/relevance of the study

The increasing integration of renewables into the grid has important implications for energy prices, grid reliability and resiliency, and of course, carbon emissions. Forecasting and predicting future levels of renewable energy generation will be very important for planning purposes as it relates to these factors as well as economic growth, global health, energy security, and more. Projecting renewable energy integration is also important in terms of understanding whether countries are on track to meet climate goals such as achieving certain targets for renewable integration in their energy generation portfolio.

Objectives

1. In our current study, first, we will train and test the percent renewable integration in the energy mix.
2. Then, we will forecast the percent integration for 1 year (Dec 2021 to Dec 2022).
3. Understanding the correlation between renewable energy integration and electricity prices through polynomial regression.

Dataset Information

Our data set is sourced from EIA's website. It was a part of the 2022 February Review. The variables of interest in this study includes Nuclear Power Consumption (Quadrillion Btu), Hydropower Consumption (Quadrillion Btu), Geothermal Power Consumption (Quadrillion Btu), Solar Power Consumption

(Quadrillion Btu), Wind Power Consumption (Quadrillion Btu), Biomass Energy Consumption (Quadrillion Btu), Total Primary Energy Consumption (Quadrillion Btu), and Total RE Consumption (Quadrillion Btu). New variable - Total RE Consumption as a percentage of total energy consumption - was added.

Methodology/Analysis

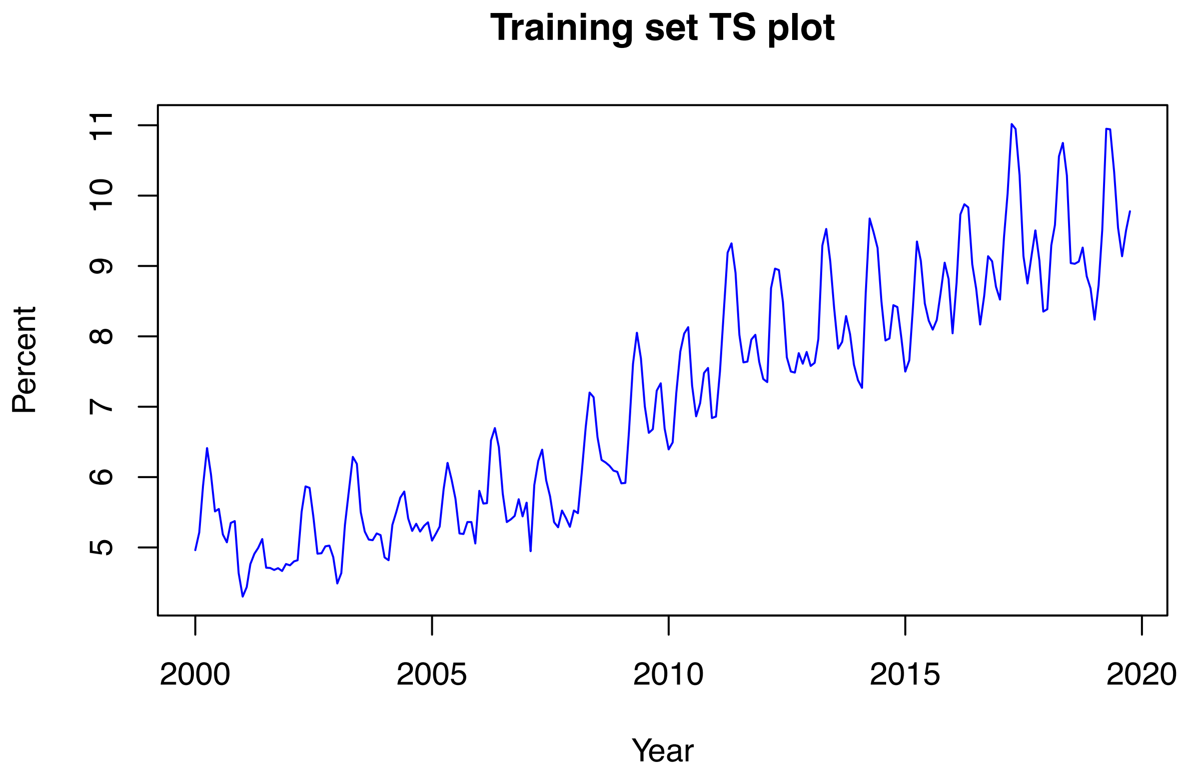
First, the data set was read and the data of interest (Date and percent renewable integration) was filtered. The variable of interest (percent renewable integration) was converted to numeric data type, and to a time series object. A training set and test set was then obtained by filtering data. Some of the models we were using required a test set of at least two full cycles and since we were using monthly data, the test sets last for 24 months. The training set begins in the year 2000 and lasts until October 2019 (two years before the end of our data set). Using head() and tail() functions, the training and test set were checked for their number of entries. The training time series data set was plotted to check for trends and seasonality. An increasing trend and a clear seasonal pattern was observed. The ACF and PACF plot were plotted. The ACF plot had a decreasing trend and signs of seasonality. The PACF plot had cut offs at lag 1 and lag 2. Using our best judgement, we predicted non-seasonal order (p,d,q) and seasonal order (P,D,Q). For the non-seasonal part, we predicted an AR model since there is a clear declining trend in ACF for the first few lags and a cut off at PACF at lag 2. For the seasonal part, we predicted an ARMA of P = 1 and Q = 1. As a check, we ran the auto.arima(). It was observed that (p,d,q) and (P,D,Q) were (1,0,2) and (1,1,2) respectively and the model should include drift. We had the drift since we did not difference the seasonal component. We decided to go with the results suggested by auto.arima(). We then tried to see if we could fit a different model that performed better than the seasonal arima model. For this, as candidate models, we decided to try the TBATS (Trigonometric seasonality, Box-Cox transformation, ARMA errors, Trend and Seasonal components) model, ETS (Error, Trend, Seasonal) model, Neural Network, and SSES (State Space Exponential Smoothing) model. A short description of the models are given below:

Description

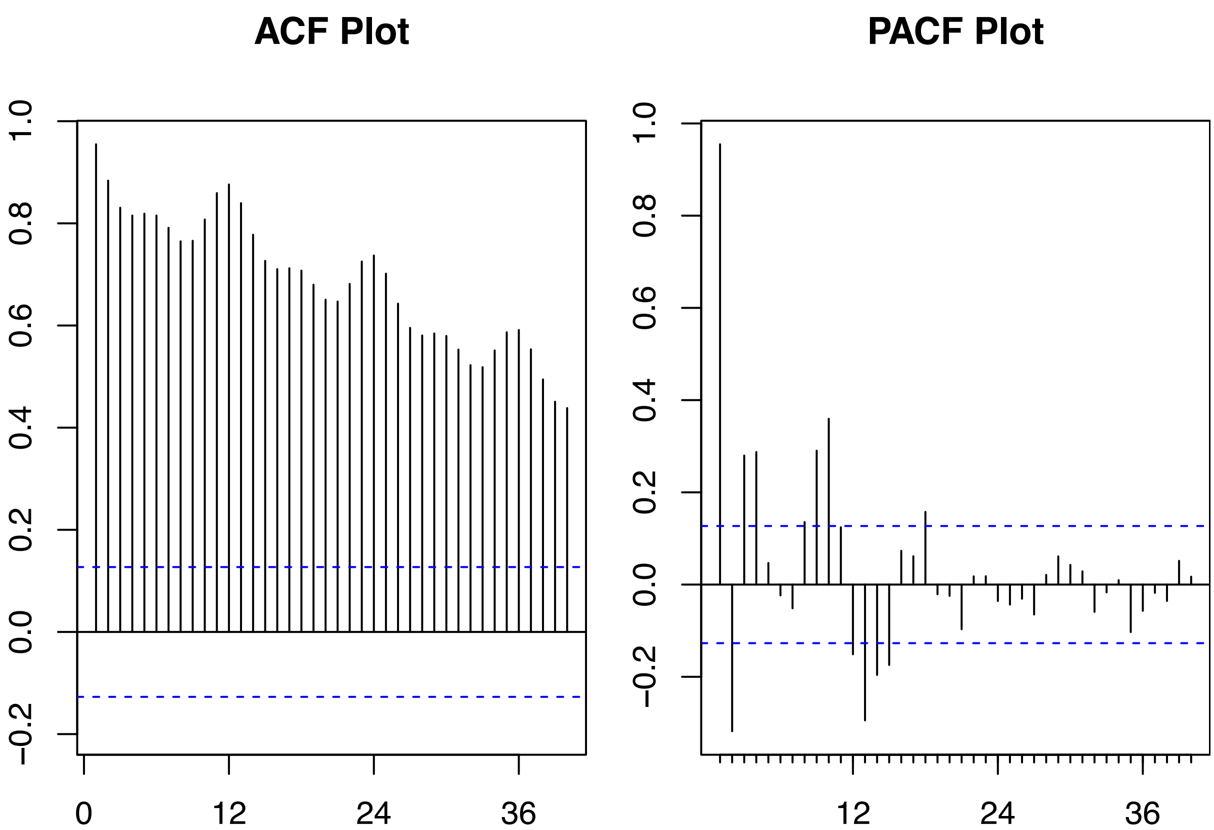
1. **SARIMA model** - SARIMA model in an extension of ARIMA that explicitly models the seasonal element in univariate data.
2. **TBATS model** - The main purpose of this is to forecast time series with complex seasonal patterns using exponential smoothing. TBATS model will consider various alternatives and fit quite a few models. It will consider models: with Box-Cox transformation and without it Box-Cox transformation, with considering Trend and without Trend, with Trend Damping and without Trend Damping, with ARIMA(p,q) and without ARMA(p,q) process used to model residuals, non-seasonal model, and various amounts of harmonics used to model seasonal effects. The final model is chosen that has the lowest AIC.
3. **ETS model** - It's a time series univariate forecasting method that use focuses on trend and seasonal components.
4. **Neural Network** - Neural Network allows complex nonlinear relationships between the response variable and its predictors.
5. **SSES model** - SSES is an extension of the exponential smoothing algorithm where a distribution assumption on the error terms to calculate the prediction interval. There are two types of error terms in the state space model: Additive and Multiplicative.

These candidate models were used to predict the test set. The plots were juxtaposed with the original values in the test set to check for accuracy. scores for all candidate models were found using the accuracy() function. The State Space Exponential Smoothing model turned out to be the best model as it had the lowest RMSE value. We then forecasted future renewable energy integration for the 2022 year using the SSES model. Using a state space model is beneficial, especially in this case, since it allows for a model with time varying parameters and models the coefficients as well as the variables. As we know, trend and seasonality with respect to renewable energy consumption are not constant and have changed over time so using a model than can analyze and update these components in real time can help improve the accuracy of our forecasts.

Time Series of the training set



ACF and PACF Plots

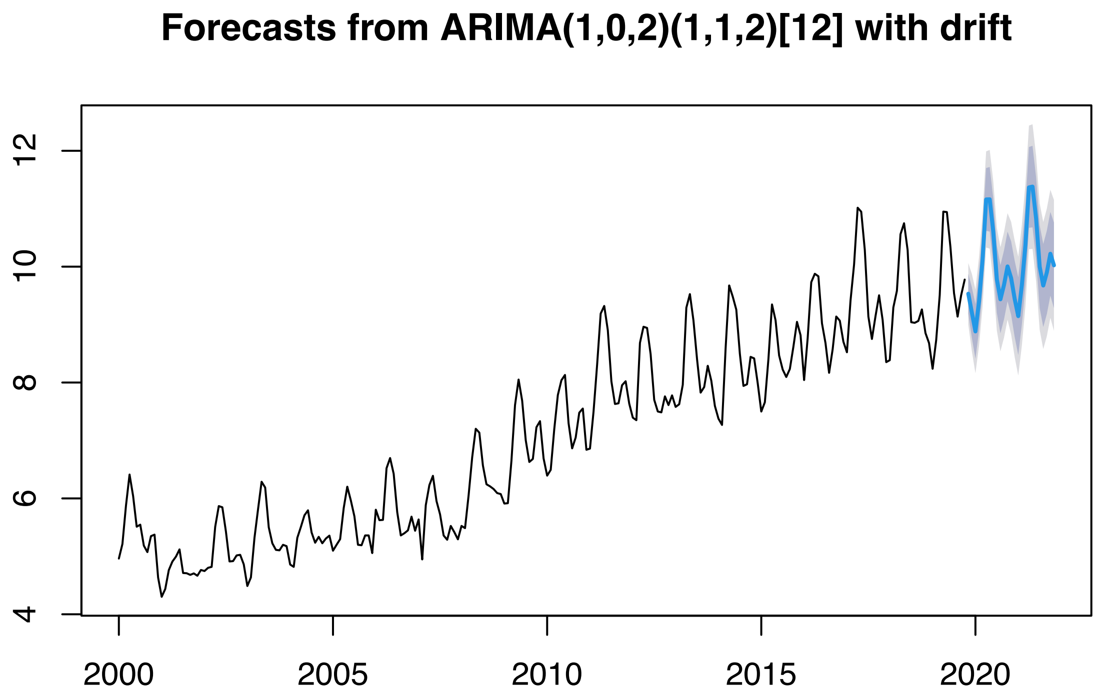


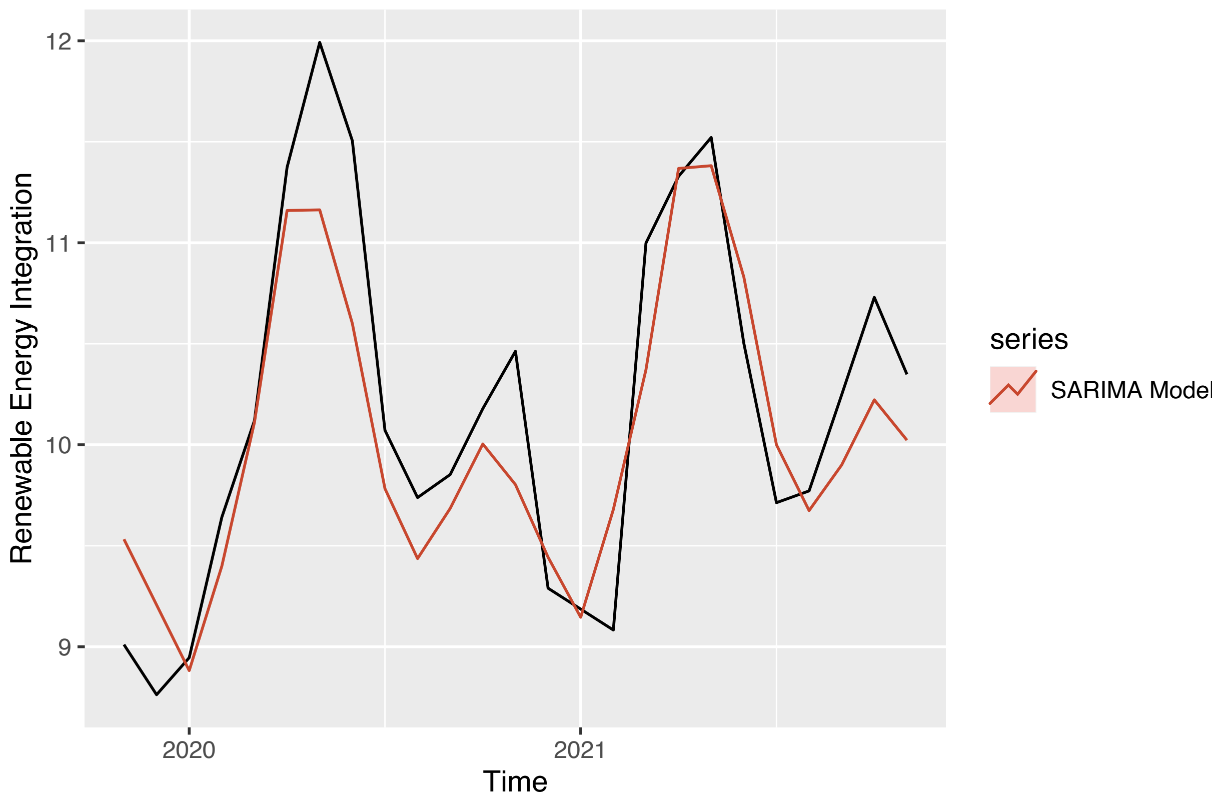
Auto Arima Results

```
## Series: TS_training_data
## ARIMA(1,0,2)(1,1,2)[12] with drift
##
## Coefficients:
##          ar1      ma1      ma2     sar1     sma1     sma2   drift
##       0.9163  -0.1505  -0.1761  -0.1699  -0.4545  -0.1963  0.0203
## s.e.  0.0424   0.0807   0.0727   0.5623   0.5550   0.3722  0.0040
##
## sigma^2 = 0.07302:  log likelihood = -25.56
## AIC=67.11   AICc=67.78   BIC=94.48
```

Let's check and see how the arima model does at forecasting the test set.

Comparing Sarima Forecast with test set


Forecasts from ARIMA(1,0,2)(1,1,2)[12] with drift
4
6
8
10
12
2000
2005
2010
2015
2020



Let's try a different model and then compare the errors.

Comparing accuracy of the models

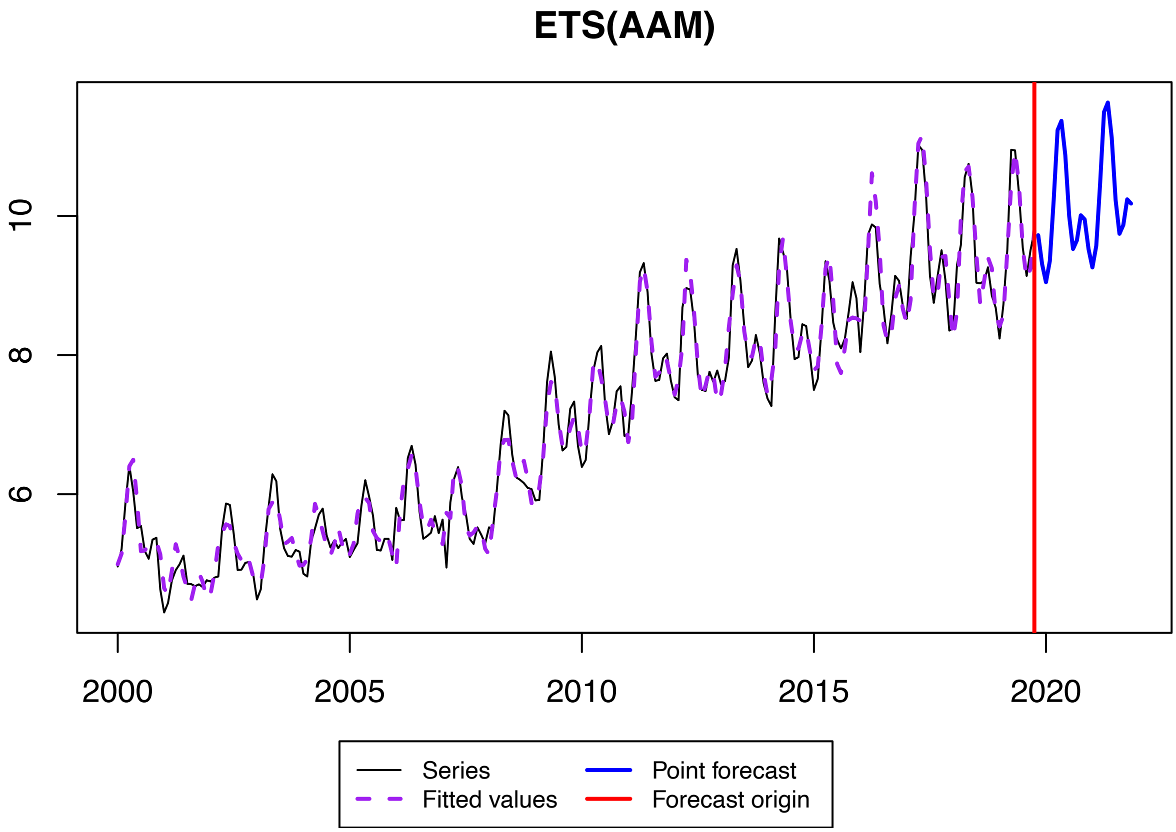
```
## The best model by RMSE is: SSES
```

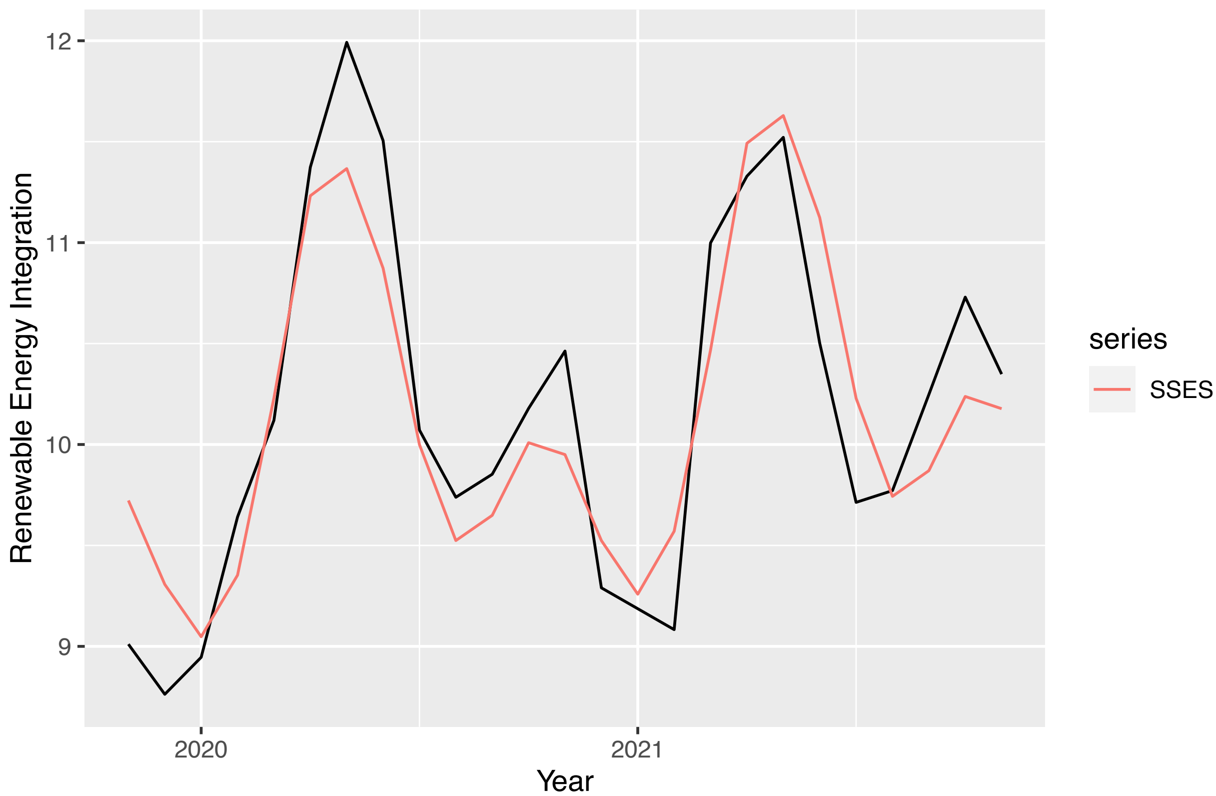
Table 1: Forecast Accuracy for Renewable Energy Integration

	ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
SARIMA	0.14284	0.41142	0.33244	1.19132	3.23145	0.26762	0.53934
TBATS	0.37306	0.57382	0.47603	3.40527	4.56051	0.40516	0.76667
STL + ETS	0.36612	0.57318	0.47968	3.34186	4.61240	0.38033	0.75932
Neural Network	0.36833	0.65228	0.51936	3.31169	4.90177	0.44841	0.90195
SSES	0.03158	0.38930	0.32527	0.11680	3.20672	0.34324	0.49739

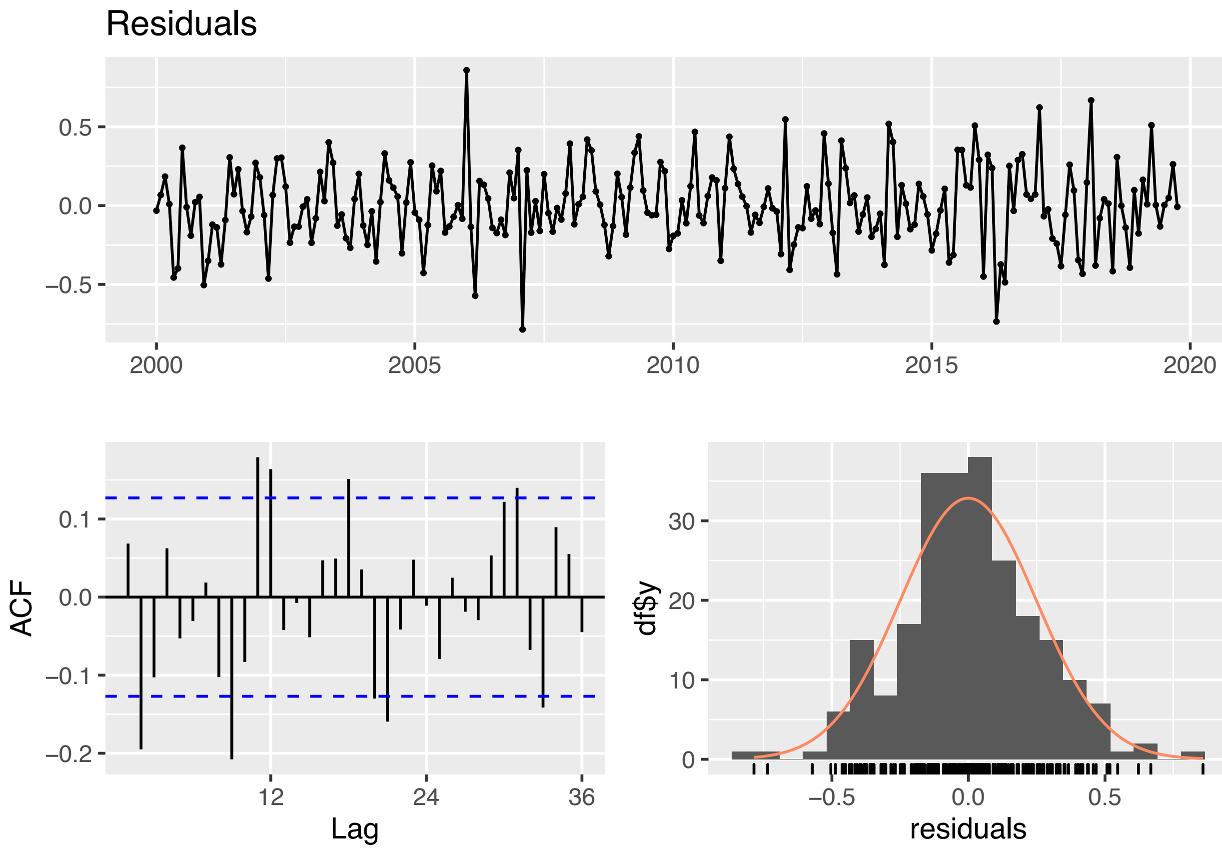
SSES model produced the most accurate results

Comparing SSES Forecast with test set


ETS(AAM)
10
8
6
2000
2005
2010
2015
2020
Series
Fitted values
Point forecast
Forecast origin

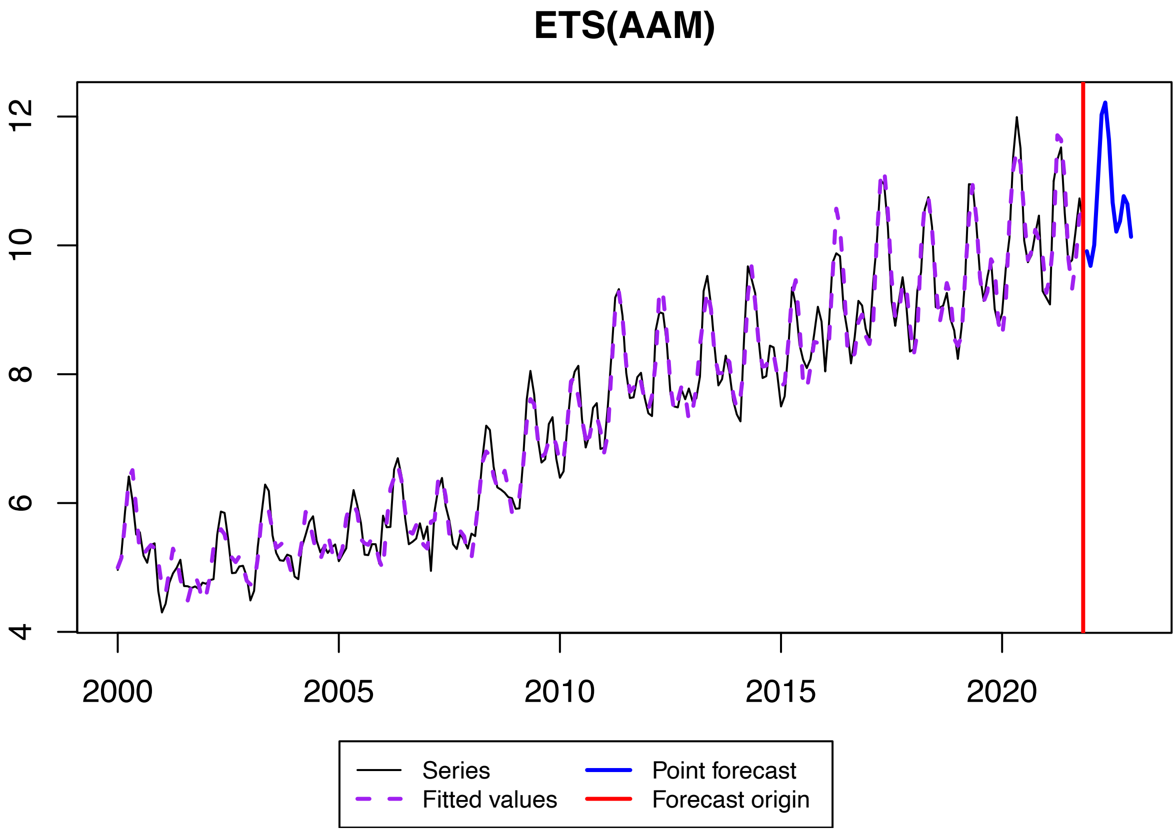


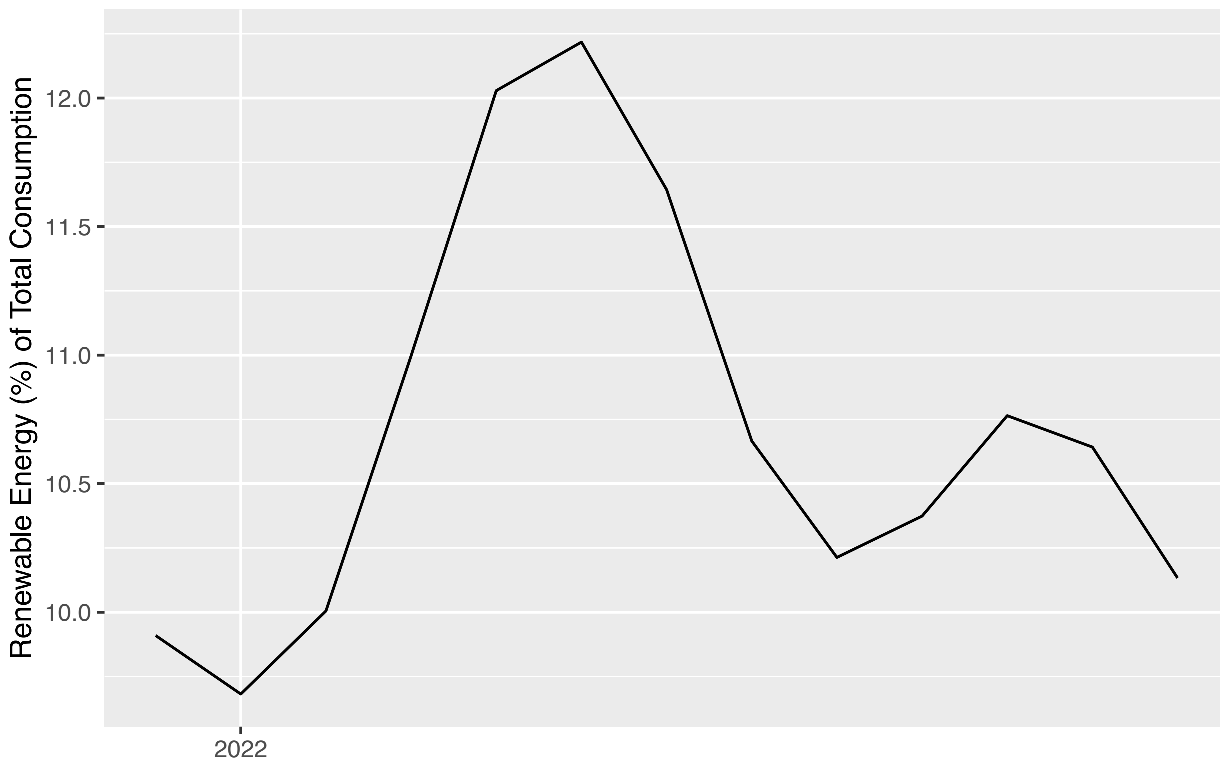
Residuals of SSES Forecast



Summary and Conclusions

Final Forecast for 2022


ETS(AAM)
12
10
8
6
4
2000
2005
2010
2015
2020
Series
Fitted values
Point forecast
Forecast origin



The data was trained from 2000 to 2020 and was tested on the year 2021. Predictions were done for 2022. From residual scores, we can infer that SSES model performed the best having an RMSE of 0.389 while Neural Network had the highest RMSE (0.6476). Therefore, SSES model was used to predict the percent of renewable integration for 2022. To better understand this metric, price was used as correlating variable. This part was done in Python. The price data for renewables was obtained from an EIA report for the years 2000 till 2020. This data was trained using a linear regression. it was inferred that the curve fitting was not accurate - low R-squared value. Hence, a polynomial regression was used to train the price data. It was observed that as percent renewable integration increased, the price also increased for the first few years and then became constant. This suggests, with an improvement in energy policies, technological advancements, and a more significant push in adopting cleaner sources of energies by many economies; the price (cents/kWh) goes down. Such results were also reported by Pregger et al. and is consistent with the literature. Then, forecasting of the prices were done for 2022 and were plotted against forecasted renewables from the SSES model. Like previous results, it was inferred that as percent renewable increased in the first few months of 2022, price increased as well and there was a drop in prices towards the end of the year. As a check, from the EIA report, it was observed that the price of renewables during Jan 2022 were 11.34 cents/kWh. From our forecasts we estimated the price for Jan 2022 as 10.41 cents/kWh (8 percent error). Our forecasted prices for forecasted renewables in 2022 could be helpful for policy makers to make more informed decisions.

Note: Plots of electricity price regression results can be seen in appendix

Benefits of using a state space model for forecast renewable energy integration

- Model with time varying parameters
- Models the coefficients as well as the variables
- Trend and seasonality are not constant and have changed over time so using a model than can analyze and update these components in real time can help improve the accuracy of our forecasts.

- Investment Tax Credit (ITC) for commercial and residential energy systems began in 2006, model can adjust for substantial level shifts.

Limitations of Study

There are, of course, a lot of limitations when it comes to modeling any TS model. However, given below are 3 main limitations.
1. Our model did not consider other exogenous variables such as percent economic growth, population growth that might affect percent renewable integration and price.
2. No extreme weather events in 2022 whose aftermath might increase the price of renewables.
3. Our model did not account for the technical, economic, environmental, and societal factors.

References

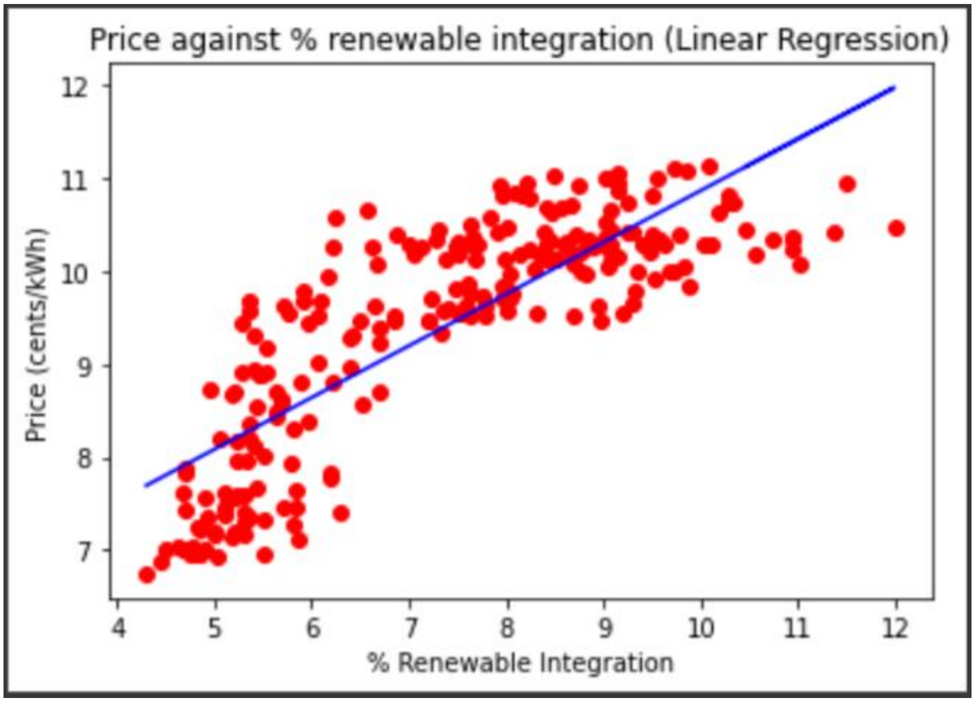
1. Sources of Greenhouse Gas Emissions. 2021. https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions

2. Renewables became the second-most prevalent U.S. electricity source in 2020. 2021. https://www.eia.gov/todayinenergy/detail.php?id=48896

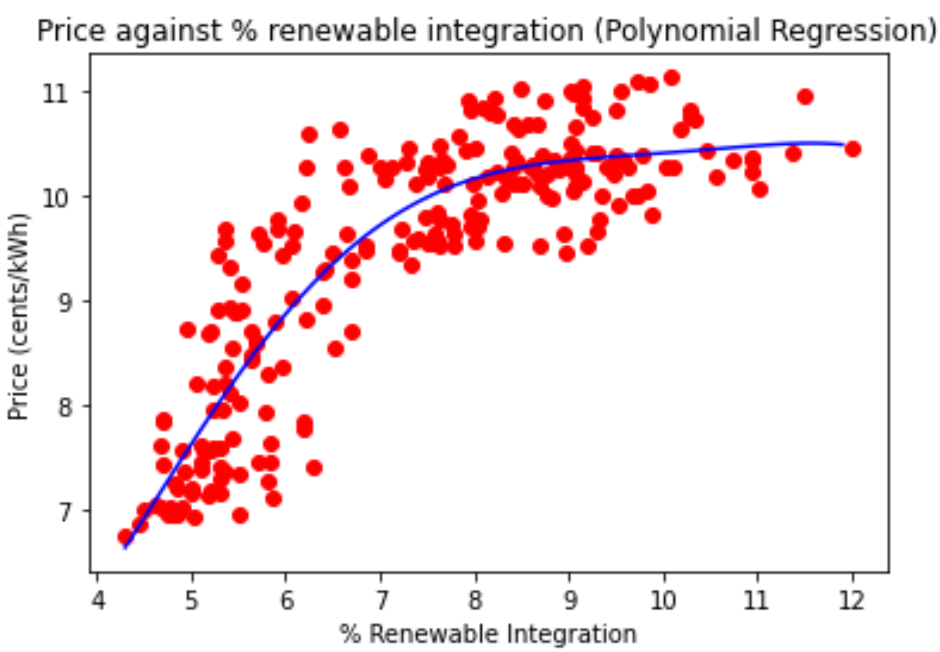
3. Pregger, T., S. Simon, T. Naegler and S. Teske (2019). Main Assumptions for Energy Pathways. Achieving the Paris Climate Agreement Goals: Global and Regional 100 percent Renewable Energy Scenarios with Non-energy GHG Pathways for +1.5 degrees C and +2 degrees C. S. Teske. Cham, Springer International Publishing: 93-130.

APPENDIX

Linear Regression plot



Polynomial Regression plot



Forecasted price against forecasted % renewable integration for 2022