**Climate Change Solutions : Analysis and Insight**

Patrick Magat, Tohidul Islam, Alice Maharjan, Antonio Jordan

Professor Erik Grimmelmann

CSC 59867 Senior Design 2

December 21, 2021

**Abstract**

Climate change refers to significant changes in global temperature and weather patterns, primarily due to human-induced emissions of greenhouse gasses. The effects of global climate change include increased wildfires, sea-level rise, droughts, reduced agricultural yields, and much more. As these effects have continuously worsened within the 21st century, we’ve made it our goal to showcase how our climate is affected with and without implementing several climate change solutions. As the largest contributor to global warming, atmospheric carbon dioxide levels will serve as our input variable for an ARIMA model where we forecast future values for the next 20 years. Afterward, we explore and analyze four common climate change solutions: renewable resources, carbon capture and sequestration, electric vehicles, and energy-efficient buildings. In doing so, we overcome modeling obstacles we came across, dive into multivariate time series forecasting using the VARS model, and review our findings in-depth. Finally, our conclusion compares our results to the original forecasts and discusses whether our solutions impact future carbon dioxide levels.

**Introduction**

Climate change is natural. Long before humans, Earth naturally went through cooling and warming phases caused by volcanic eruptions, changing concentrations in greenhouse gasses, and even solar variations. More recently, however, human activities have been the main driver in the unnatural shifts in our climate, most significantly the burning of fossil fuels. The burning of coal, gas, and oil has severely increased greenhouse gasses’ concentration within our atmosphere. Due to their ability to trap and radiate heat, this event has led to significant global warming. If we continue to ignore our contribution to this problem, the Intergovernmental Panel on Climate Change (IPCC) forecasts a temperature increase of 2.5-10 °F over the next century. This will lead to already occurring disastrous effects, such as increased droughts, loss of species, and a warmer, rising ocean. As important as this topic is, our group decided to explore solutions that could help in our fight against climate change.

The four solutions discussed below are renewable resources, carbon capture and sequestration, electric vehicles, and energy-efficient buildings. Each plays its own role in mitigating or adapting to the rise in carbon emissions, our primary indicator of global warming. Carbon dioxide is the most important of the greenhouse gasses for several reasons. It’s more abundant, stays in the atmosphere longer, and absorbs less heat per molecule. As a result, it’s responsible for two-thirds of the energy imbalance on Earth. In order to see how our solutions would affect future climate, we initially forecasted future carbon emissions using Python’s ARIMA model.

**ARIMA Model**

The ARIMA model, short for AutoRegressive Integrated Moving Average, provides a means of forecasting univariate time series data. The model is characterized by the parameters P which represents the autoregressive portion of the model, D which represents the integrated portion, and Q which represents the moving average. The formula for the ARIMA model, shown in the middle of *Figure 1*, combines the AR and MA predictors to obtain the series Yt which is then differenced a number of times given by parameter D. The autoregressive portion of the model serves to predict future values of the dependent variable based on a linear combination of its past values. The AR model, given by the leftmost formula in *Figure 1*, obtains the value for the next variable Yt by taking the sum of the past P \*Y-values along with some constant and “white noise”1 (Hyndman). The moving average portion of the model aims to predict future values by utilizing past forecast errors. The MR model, given by the rightmost formula in *Figure 1*, obtains the value for the next variable Yt by taking the past Q error values t-1…t-q summed up with some constant and “white noise”t . A major objective in the modeling of solutions to climate change was to determine the most optimal values of p, d, and q for which an accurate forecast would be made.

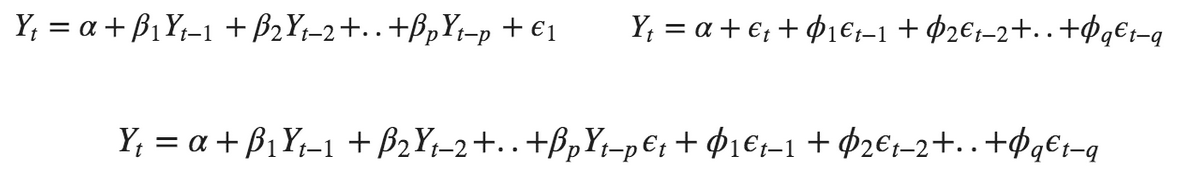


Figure 1: AR Model (left), MR Model (right), and ARIMA Model (center)

## 

## **Renewable Resources**

## The emergence of renewable energy has given us the best solution for mitigating and stabilizing the levels of greenhouse gasses in our atmosphere. The burning of fossil fuels for electricity, heat, and transport is the direct source of these heat-trapping gasses, especially carbon dioxide. According to an International Energy Agency (IEA) report released in 2018, the Buildings Operations sector accounts for 28% of global CO2 emissions, whereas transportation takes up 23%.Outstanding statistics such as these prove that the utilization of fossil fuels for our daily needs must come to an end as we begin to transition to more natural sources. In fact, renewable energy is now the fastest-growing energy source globally and within the United States, with renewables making up 29% of electricity generation by the end of 2020 worldwide. Hydroelectric power takes the lead with 16.8%, while wind energy follows up with a little over 6%. These are just a few examples of why changing our primary energy source to renewable resources such as these is the best way to stop using fossil fuels and reduce climate change.

**Data**

Our dataset comes straight from British multinational oil and gas company, Bp. Their Statistical Review of World Energy 2021 looks at the decline of fossil fuels and carbon emissions amidst COVID and the rise of renewable energy. Carbon dioxide emissions are measured in million tonnes, with it dating back to 1965. As our primary indicator of global warming, this served as our time series to feed into the ARIMA model. Luckily, we didn’t have to look far for data on renewable resources, as the same dataset contained information on various types. That list includes nuclear, hydroelectric, solar, wind, geothermal, and biomass energy. While each renewable dated back to 1965 as well, the dataset separated them into energy generation and energy consumption. In our case, we chose to work with energy generation, which is measured in terawatt-hours. Consequently, the addition of these new variables led us to adopt a new model in order to see how renewables would influence future carbon emissions. This new model is known as Vector AutoRegression.

**VARS Model**

Unlike the ARIMA model, Vector AutoRegression or VARS uses multivariate time series forecasting. As a result, we can use features or variables other than carbon emissions to forecast it. Because our goal is to see the effect of renewable energy on the future forecasts of carbon emissions, a VARS model serves its purpose perfectly, as all the variables influence each other. This fact subsequently creates a system of equations with one equation per variable. *Figure 2* showcases a system of 2 equations, an example of a VAR model with two time series. Including carbon emissions, we started with six time series in total. However, soon to be discussed, modeling obstacles and data cleaning decisions forced our hands to remove two of them for the benefit of the project.

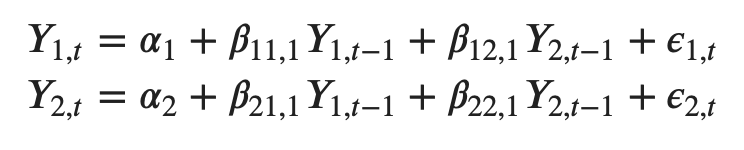


Figure 2: Var(1) Model with two time series

**Data Cleaning/Stationary tests**

After importing the dataset, our first problem and probably the most significant issue was making the data stationary. Stationary data implies that the statistical properties, such as mean and standard deviation, of a system do not change over time. That is especially important for time series analysis because a majority of forecasting models assume each point is independent of one another. We need the data’s overall behavior to be constant to perceive and predict future values accurately. In order to do so, our data should show no trend or seasonality. Trend is a long-term increase or decrease in the data, while seasonality describes patterns that repeat regularly over time.

Not only did our six time series show a trend, as seen in *Figure A1*, but the results of a particular statistical test proved our data to be non-stationary. This test is known as the Augmented Dickey-Fuller test (ADF) and is common for testing stationarity by making the presence of a “unit root” the null hypothesis. Accepting the null hypothesis means the time series has a unit root and therefore, is non-stationary. These results meant we would have to perform differencing or logarithmic transformations to make the data stationary. Differencing helps to remove trend and seasonality by stabilizing the mean. When differencing is not enough, log converts the data in a logarithmic scale before differencing to help.

Our issue arose when no amount of differencing was working, nor was taking the log due to division by zero errors. Solar and wind generation, in particular, had a lot of zero values up until the 1990s that was causing the problem. In an attempt to correct the issue, we tried dropping missing values or filling them in using Panda’s backward fill (*bfill*) method. Backward fill solved the error and made our data stationary but produced complications when modeling. Ultimately, we decided to remove both solar and wind generation from our dataset. In turn, this decision helped us achieve stationarity in three differences, which can be seen in *Figures A2, A3, and 3*.

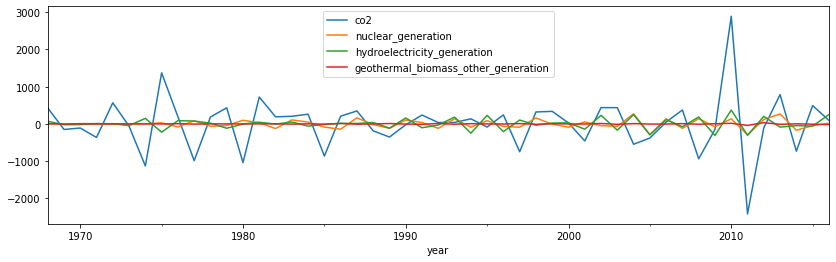


Figure 3: Plot of Third Difference

**Selecting Order P**

Following our data transformation, we ran into our next problem, which was selecting an order P by looking at the AIC. Typically, one wants to see the AIC drop the lowest and then increase again. However, in our case, with a maximum of 8 lags, the AIC value just kept decreasing, which made it difficult to choose a good number for p. As a result, we trained the model using each lag order as our p to see what performed the best. Our performance metric was the serial correlation for each series. In the end, we chose to use a lag order of 4.

**Inverting the Transformation**

Another difference from ARIMA is the inversion of our transformations to retrieve the actual forecast. Initially, the forecasts are generated on a scale of the training data. Therefore, to bring it back to the original scale, one needs to “de-difference” it as many times as the data is differenced. The trickiest part of this process was identifying how to roll back the third difference. To roll back a difference, one must take the most recent values of the original series’ training data and add them to a cumulative sum. This part consisted of a lot of trial and error because choosing the wrong values resulted in inaccurate plots afterward.

**Results**

When overcoming all these problems and successfully training the VAR model, our Forecast vs. Actual plot seen in *Figure 4* came as a surprise. At first glance, the forecasts don’t look close to our observed values at all, especially carbon emissions. However, our evaluation metrics say otherwise. The mean absolute percentage error for carbon emissions was only 6%. Mean absolute percentage error (MAPE) measures a forecast system’s accuracy. Higher percentages are not what we want since it’s a measure of error. Our carbon emissions forecast showed a greater decline in emissions and wasn’t as far off as we thought. Unfortunately, our MAPE for our other time series was much higher, with nuclear at 15% and hydroelectricity at 23%. It’s hard to say how well renewables influence carbon emissions from our VARS model, but we believe further adjustments could fix that. Much like how renewable energy will help reduce atmospheric carbon dioxide, carbon capturing and sequestration aims to do the same.

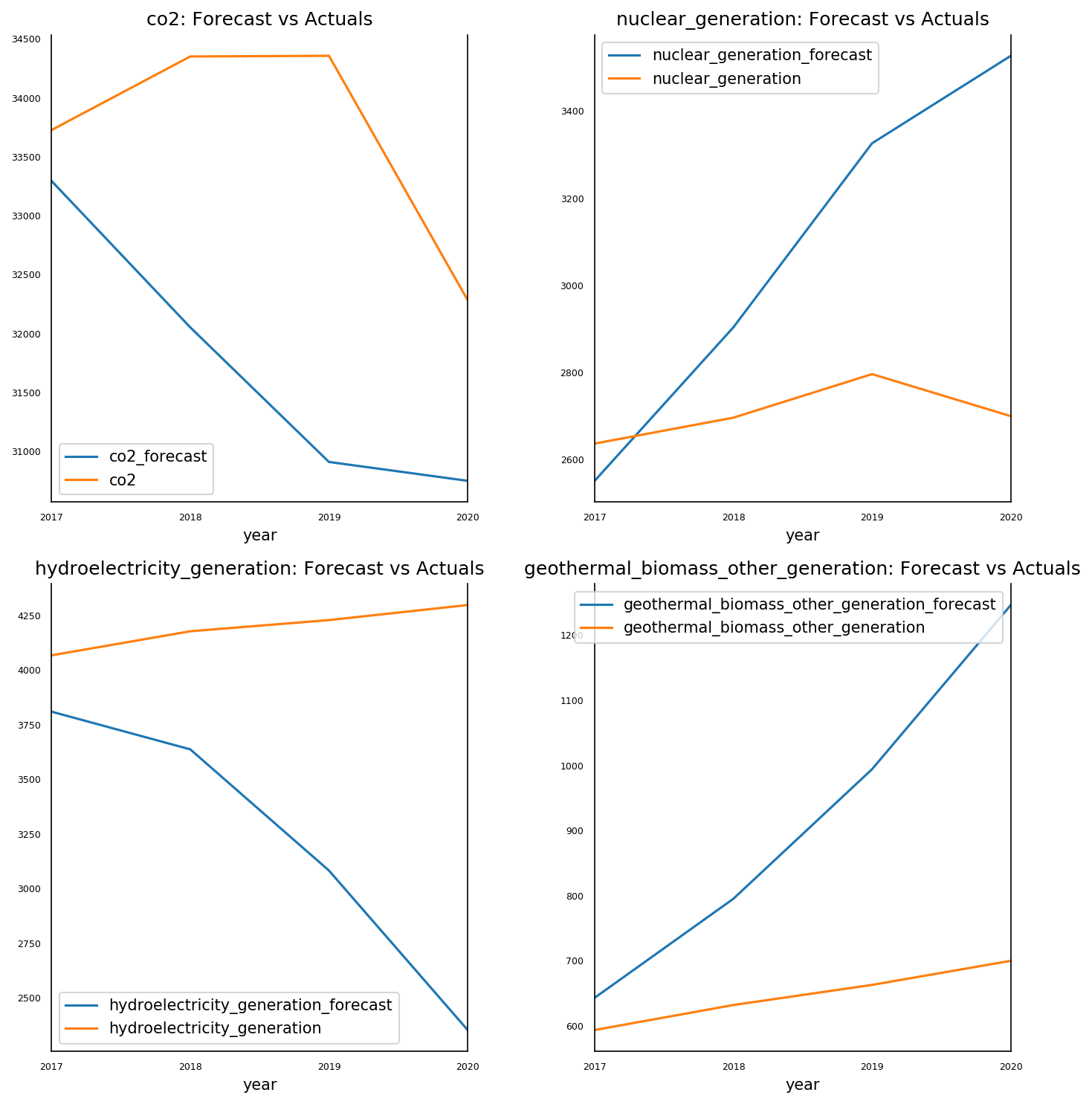


Figure 4: Forecast Vs Actual Plot

**Carbon Capture and Sequestration (CCS)**

While many attempts to reduce climate change have focused on the shift to more sustainable sources of energy, there is still a significant percentage of the world that is reliant on the burning of fossil fuels and biomass; much of which are not slated to make a transition to cleaner forms of energy in the near future. Specifically, stationary sources of CO2 emissions such as factories and processing plants are the largest contributors to the annual global greenhouse gas emissions. According to a 2016 breakdown of emissions by sector, nearly 49.4 billion tons of CO2 emissions were produced that year, of which 43 percent, or approximately 21 billion tons, of CO2 were attributed to the energy producing and industry sectors (Ritchie and Roser). In order to tackle the issue of climate change at its roots, a solution that targets the present-day major sources of CO2 emissions is desired. One such solution that can reduce the CO2 emitted from major sources is the use of carbon capturing and sequestration. Carbon capturing and sequestration is the process of capturing CO2 emissions “at the point of combustion” (Berend) in stationary facilities, and then processing, selling off, or storing the CO2 deep in geological formations or onsite. By analyzing CCS technology and its growth, we can forecast its effectiveness in reducing global CO2 emissions in the future.

**CCS Data Collection**

## In order to produce a forecast of the effectiveness of carbon capturing and sequestration in relation to global CO2 emissions, we obtain a dataset of all operational, suspended, and in construction carbon capturing and sequestration facilities since 1972 up to 2028 from the 2021 status report by the Global CCS Institute (“Global Status Report”). Using the given max CO2 storage capacity of each facility from this dataset, we produced the cumulative sum of annual global storage capacities. The capacity of facilities with operations suspended is excluded from the total capacity in the years after the date of suspension. This is done in order to maintain the accuracy of the total global CCS storage capacity at that point in time. The resulting dataset shows an almost exponential growth in the total global CO2 storage capacity over the last 40 years, visualized in *Figure 5*, with a notable change in the past decade. This observation is supported by the Paris Agreement which was an international treaty adopted in 2015 to reduce the effects of climate change to under two degrees celsius (“Paris Agreement”). The surge in CCS facilities in recent years can be attributed to the global push for reducing climate change incentivized by this treaty.

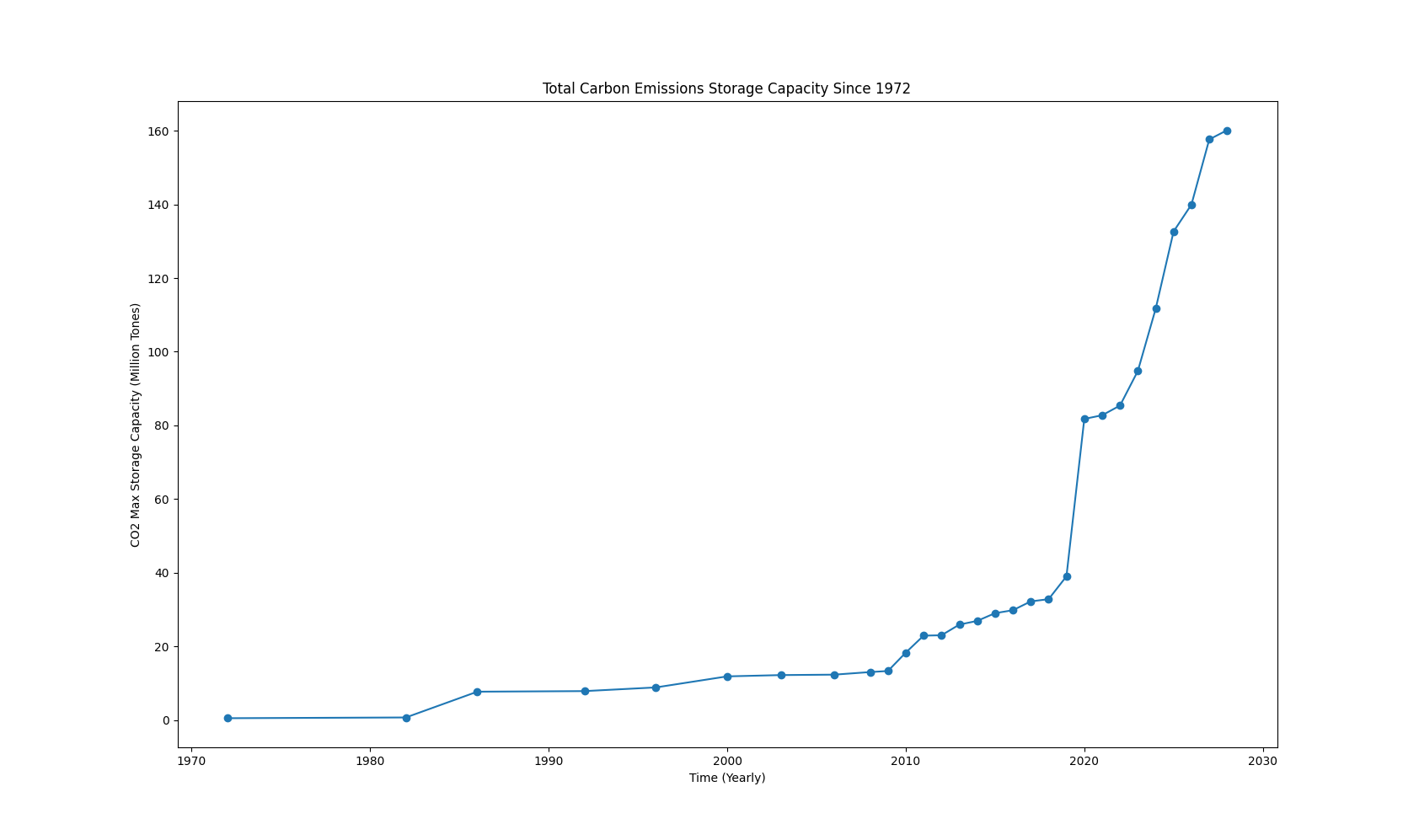


Figure 5: Total Global CO2 storage capacity

## The data for global CO2 emissions is once again obtained from the British multinational oil and gas company, Bp. To then obtain the corresponding dataset representing the emissions produced by the energy producing and industry sectors we obtain the dataset of annual breakdowns of contribution to CO2 emissions by economic sectors. The sectors pertaining to energy production and industrial processes are then extracted and summed up to determine the average contribution made by both sectors annually. The resulting value was 56 percent of global greenhouse gas emissions being attributed to these sectors, and a 43 percent contribution of global CO2 emissions. An approximate dataset for annual global CO2 emissions from the energy producing and industry sectors is then produced from the annual global CO2 emissions dataset as shown in *Figure 6*.

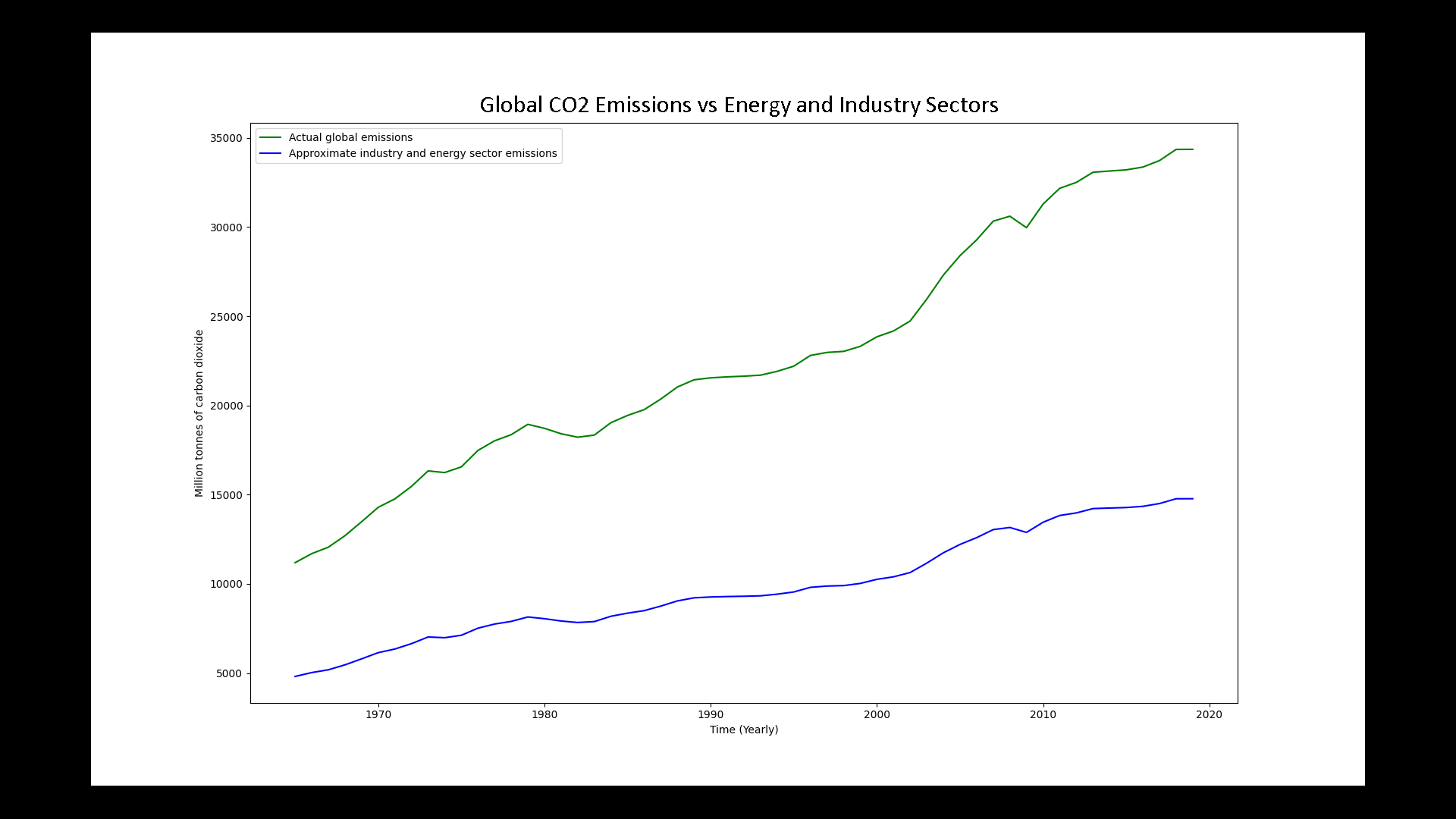
****

Figure 6: Energy and Industry Sectors’ CO2 emissions

**CCS Model and Choosing Optimal Parameters**

## A significant part in creating our ARIMA model of CCS CO2 capacity was determining the optimal values for the number of lags in the dependent variable (p), the number of times the data is differenced to make it stationary (d), and the number of lags in the forecast error (q). To check the data for stationarity we perform the Dickey-Fuller test on the CCS dataset; the resulting test statistic is significantly larger than the critical values and thus the null hypothesis was rejected. This indicated that our dataset for the global CO2 capacity was almost guaranteed to be non-stationary, thus differencing is required in order for forecasts to be accurate. Knowing that the data is non-stationary and that a differencing would be needed, we begin testing for the most optimal p, d, q values using a brute force approach. An ARIMA model is created for each unique combination of p, d, q with values ranging from 1 to some value *x*. The akaike information criterion is then extracted and compared with all previous formulations of the ARIMA model, the parameters producing the model with the lowest AIC value is then labeled to be the most optimal. However, this brute force approach to determining the model of the best fit grows significantly in computation time as the value of *x* is increased. In general, the number of ARIMA models constructed is given by x3; thus, at just a value of x = 10 there are 1000 distinct ARIMA models that need to be formulated and compared to determine the lowest AIC value, the resulting computation time goes well over 30 minutes. As a result, we opted for x = 8 as the upper bound on the values for p, d, and q which produced an optimal set of p = 2, d = 3, q = 7.

## In order to verify that the given values produced a model that would accurately represent the data, we produced the diagnostics plot shown in *Figure 7*. The plot of the standardized residual indicates that the produced model stays relatively close to the expected data points with a jump around 2020. Based on the normal Q-Q plot, the model data has roughly the same distribution as the expected normal distribution. In addition, the resulting autocorrelation plot shows that the autocorrelation in the data is relatively low which suggests that the model is a good fit for further data analysis. Based on these observations we concluded that the values of p = 2, d = 3, and q = 7 would provide a suitable model for forecasting the growth of global CO storage capacity based on observed CCS facility data.

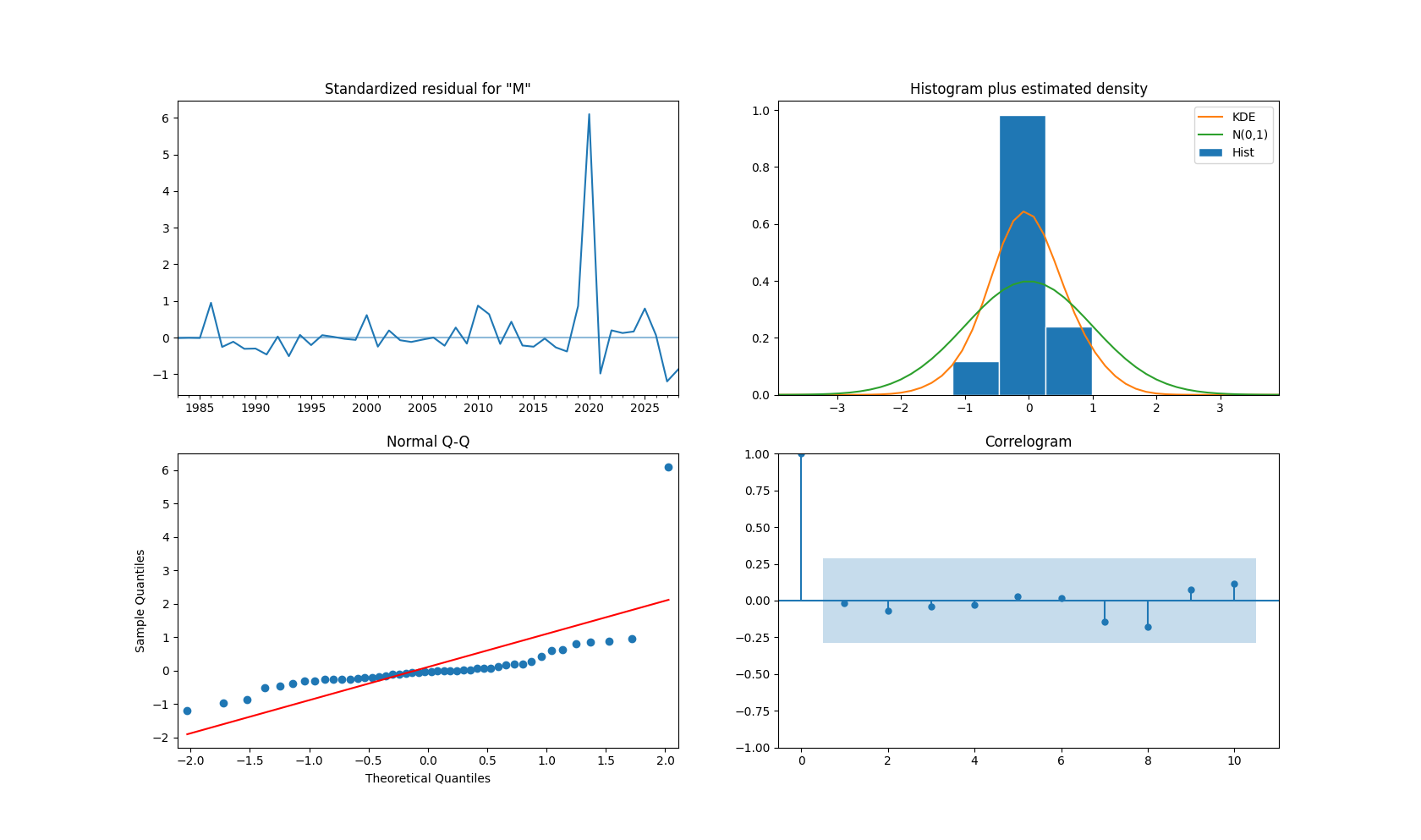


Figure 7: Diagnostics Plot

**Net Zero**

## A major goal of carbon capturing and sequestration technology is to reduce global carbon dioxide emissions from applicable sources to what is known as net zero by the year 2050. At this point, the CCS facilities worldwide would capture and store an amount of CO2 equivalent to the amount produced by stationary sources. In order to determine the plausibility of this goal, we construct an ARIMA model to forecast the growth of CO2 storage capacity over the next two-hundred years; the forecast for global CO2 emissions from the energy producing and industry sectors is then added in using the dataset derived earlier with the p, d, q parameters of p = 1, d = 1, and q = 3 selected in the same manner as before. The resulting plot, illustrated in *Figure 8* along with the confidence intervals of both models, indicates that with the current rate of growth in the number of CCS facilities, achieving net zero will not be possible until the year 2216 where the total annual CO2 being produced and stored is approximately 16,588,000,000 tons per annum. In the year 2050, global CO2 emissions from the energy and industry sectors is projected to rise to approximately 16,192,000,000 tons per annum with a global CCS storage capacity of only 692,000,000, leaving 15.5 billion tons of CO2 to be released into the atmosphere. With an average annual storage capacity of 1.335 million tons per facility, based on the observed dataset, we would need an additional 11.6 billion facilities built by 2050 if net zero is to be achieved using only current carbon capture and sequestration technology. Such a task can be deemed infeasible due to constraints imposed by funding and time.

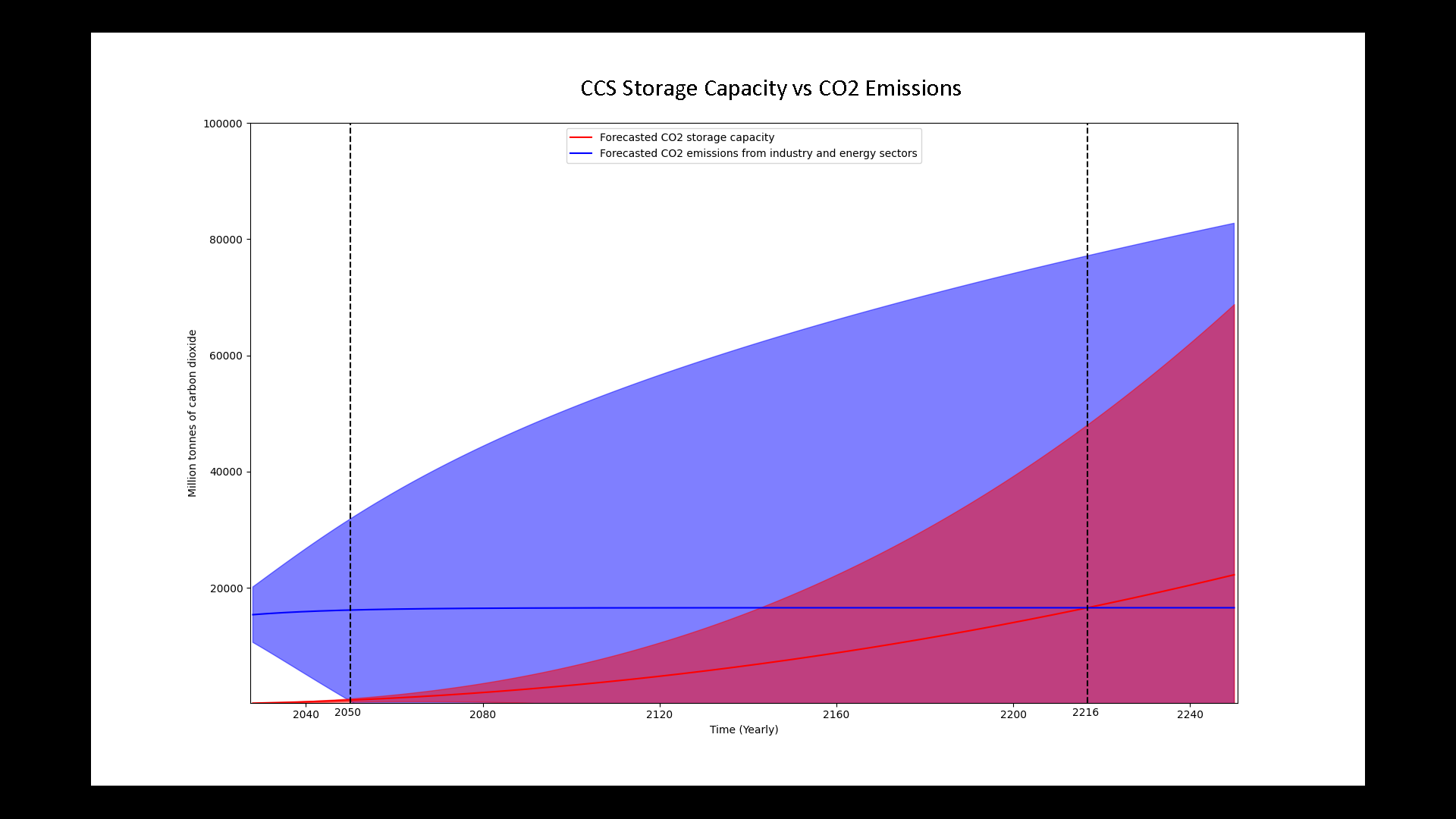


Figure 8: CO2 storage capacity vs emissions, forecast of Net Zero

**Limitations and Assumptions**

## When considering the datasets for the formulation of the CCS solution, we were faced with the problem of very sparse data. The use of carbon capturing and sequestration technology is relatively new and rarely used before the 21st century due to the additional costs that come with it. In most cases implementation of CCS technology results in 50 to 80 percent increase in the cost of electricity (Rhode). This additional overhead, along with the lack of incentives to encourage the reduction in emissions prior to the 21st century, results in the number of CCS facilities being few and far between in the beginning of the dataset. In addition, the models derived in this section depend on the assumption that the annual percentage of CO2 emissions contributed by the energy and industry sectors remain roughly close to the average of 43 percent seen in the last decade. It is entirely possible for an event to take place in the future that alters how much of global CO2 emissions is attributed to these sectors. In such a case, the onset of net zero would shift to an earlier date given that the percent contribution decreases, or shift even further from present day otherwise. Furthermore, as seen from the model in *Figure 8*, the confidence interval for the forecasts of global CO2 emissions and storage capacity grow rapidly as they move away from the observed data. This creates some uncertainty on the reliability of the prediction for net zero.

## **Electric Vehicles**

Transportation is one of the highest contributors to the United States’ carbon emissions, roughly taking around 30%. Half of the emissions from the transportation sector is due to privately owned light vehicles. A huge cause of this is that out of all the major automotive markets, the United States has the lowest adoption rates of Electric Vehicles (EV) by an extremely significant margin. There are plenty of reasons thrown around as to why this is the case but at the end of the day, this is why the transportation sector is nearly tied with the energy sector for carbon emissions in the United States.

In order to combat this, the federal government has decided to set a goal of 50% EV sales by 2030. As this is not a legislature, they are not placing an expectation on the people to meet this quota. Instead, it sets the tone for the federal government, state governments, and manufacturers to push the market towards EVs. Such examples include tax rebates, regulations where manufacturers need to have and sell a certain number of EVs per year, and so on.

The aim of this model is to see the effects of this model should that goal be met. By extension, the model would be able to tell if the current EV trends, should things remain the same, can meet this set goal. As will be demonstrated in a short while, the forecasted sales by 2030 are way below the target number of sales.

**Model Structure**

The idea is to use the ARIMA model to perform simple forecasts and perform analysis based on the results. The forecasts would be overall United States’ emissions, vehicle sales, vehicle registrations, EV sales, and EV registrations. Further models will be made based on these forecasts such as the extent at which increased EV sales would improve yearly carbon emissions. Lastly, extra hypothetical scenarios will be drafted up if the forecasted sales of EVs do not reach 50% of vehicle sales by 2030.

For the data used in this model, it is difficult to find a comprehensive set of data that came from one source. This is mainly due to the issue of EVs being relatively new technology. Data is not as matured and there is a lot of variance in what one can find. Therefore the possibility exists that the forecasts may have some level of inaccuracy, however the findings seem to be in line with what is expected. In the end, the data used for the EV sales and EV registrations is taken from IEA.org titled “Trends and developments in electric vehicle markets”, overall vehicle sales from statista.com titled “Light vehicle retail sales in the United States from 1976 to 2020”, vehicle registration from the Hedges Company “US Vehicle Registration Statistics”, and the dataset from BP used in the introduction filtered to only count US.

Before modeling is performed, some assumptions and limitations about the model need to be addressed. One mainly has to deal with the fact that the world is not a static environment. As such, the model assumes this is the case. No other external factors, whether it be regulation, market trends, advancements in technology, power grid issues, a global pandemic, and so on, will take effect during the forecasting period of 2021 till 2030. What this means is that any improvement towards the forecasted carbon emissions in the model would mostly come from more EVs on the road. Furthermore, only battery operated EVs will be considered as they effectively have 0 carbon emissions when it comes to operating them all the time. Including plug-in hybrid EVs is not feasible as there is no reasonable way of knowing their exact contribution towards improving carbon emissions as that is on an owner-by-owner basis. Lastly, the EVs are assumed to never be in accidents and be reliable enough to last 10 years. Therefore, each EV sale counts as an EV registration. This assumption can be made because of the positive correlation.

Before moving on with modeling, one also needs to consider the limitations of the ARIMA model. While there are several, one needs to be addressed, the fact that EV data is not mature enough. This introduces a lot of variance when it comes to changing parameters on top of not having a lot of options for said parameters to begin with. As shown in the next figure, this is the most reasonable set of parameters (3, 0, 0) for electric vehicle sales given those assumptions.

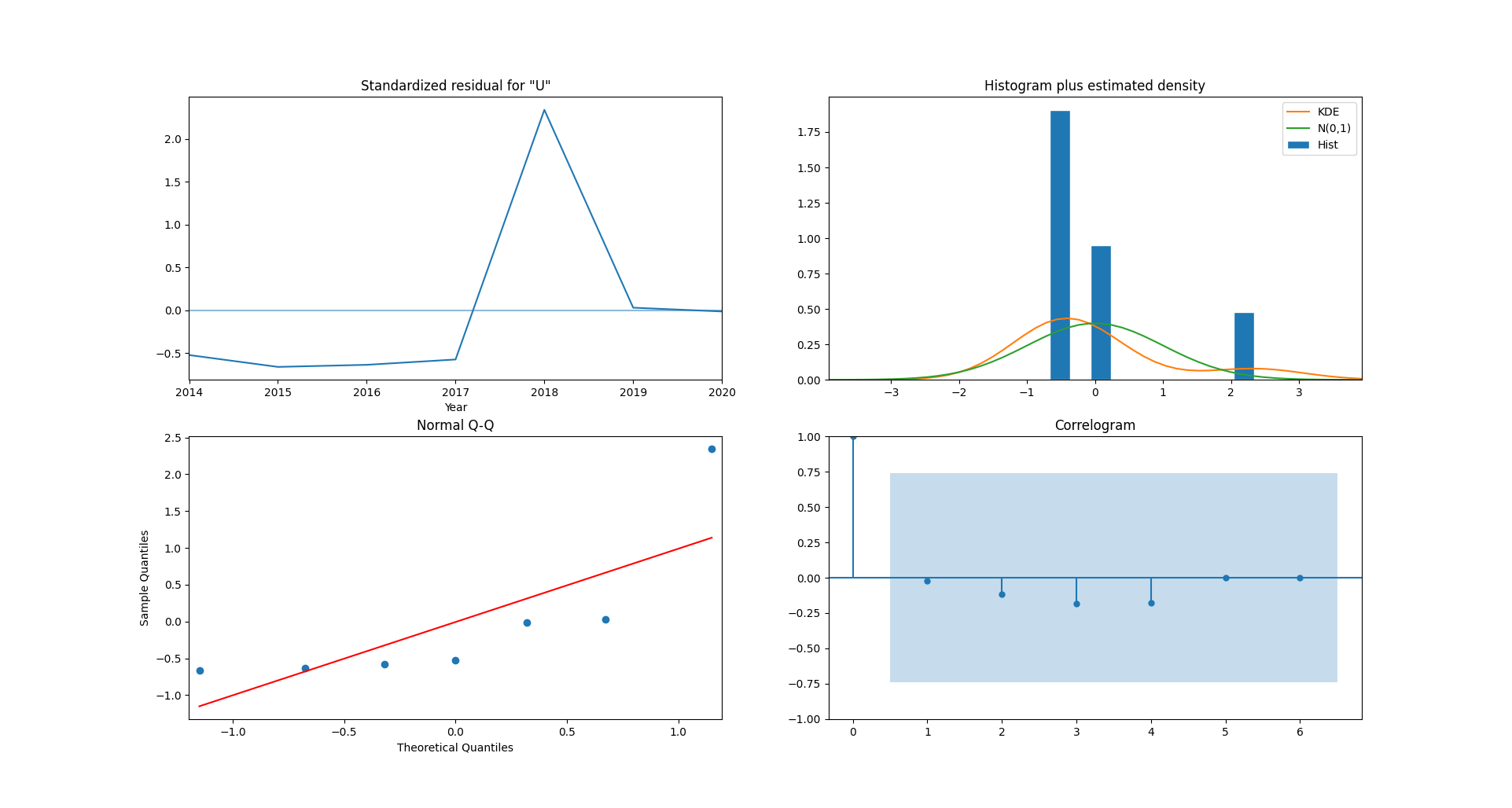


Figure 9: Diagnostics

**Modeling Results**

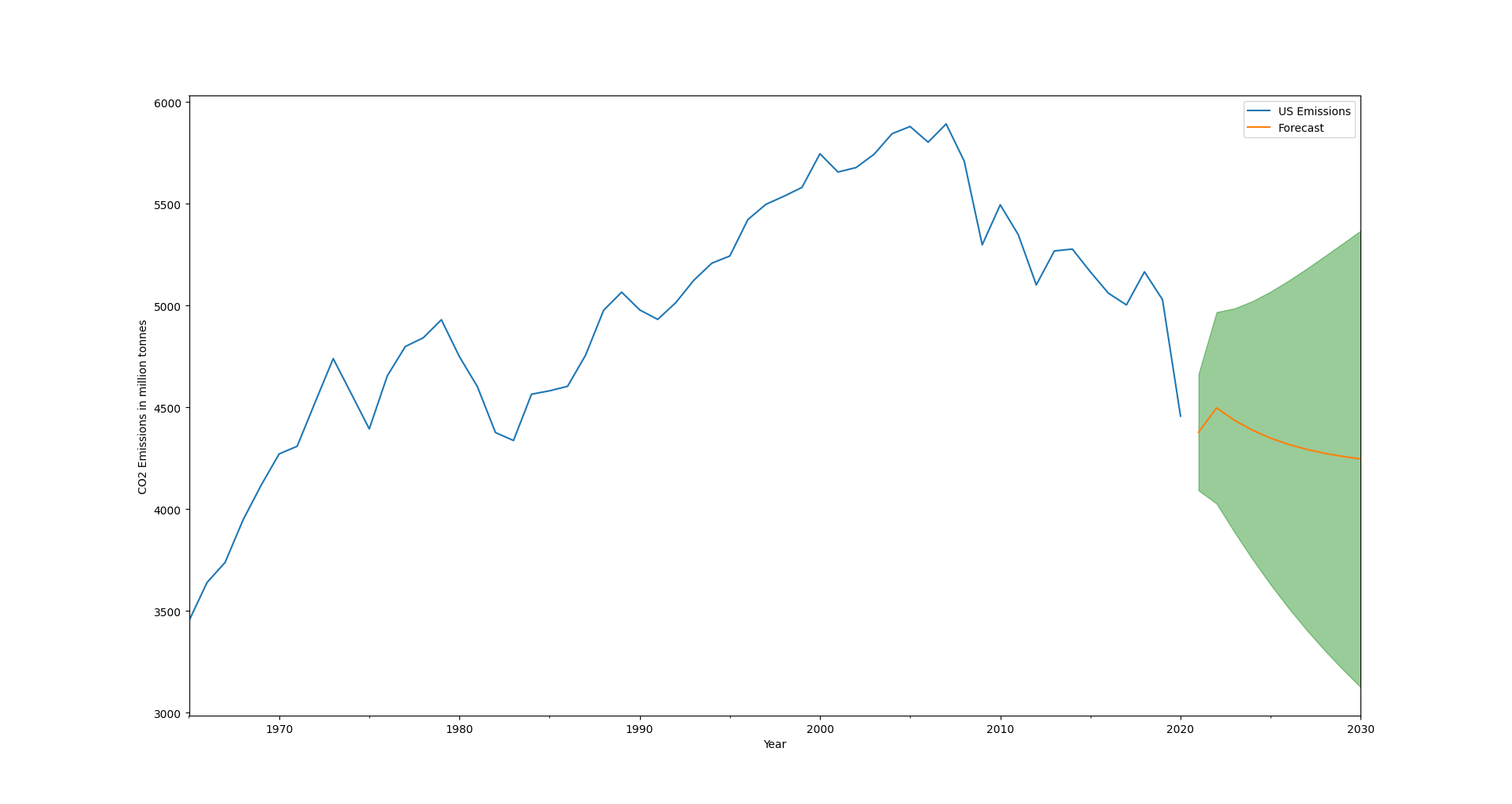
****

Figure 10: U.S. Emissions Forecast

Here, one can see that the forecast expects carbon emissions to continue decreasing naturally. Part of the sharp decrease in the past year is naturally due to the pandemic but the model does not know that. Regardless, emissions have been on a downward trend prior so this result remains reasonable.

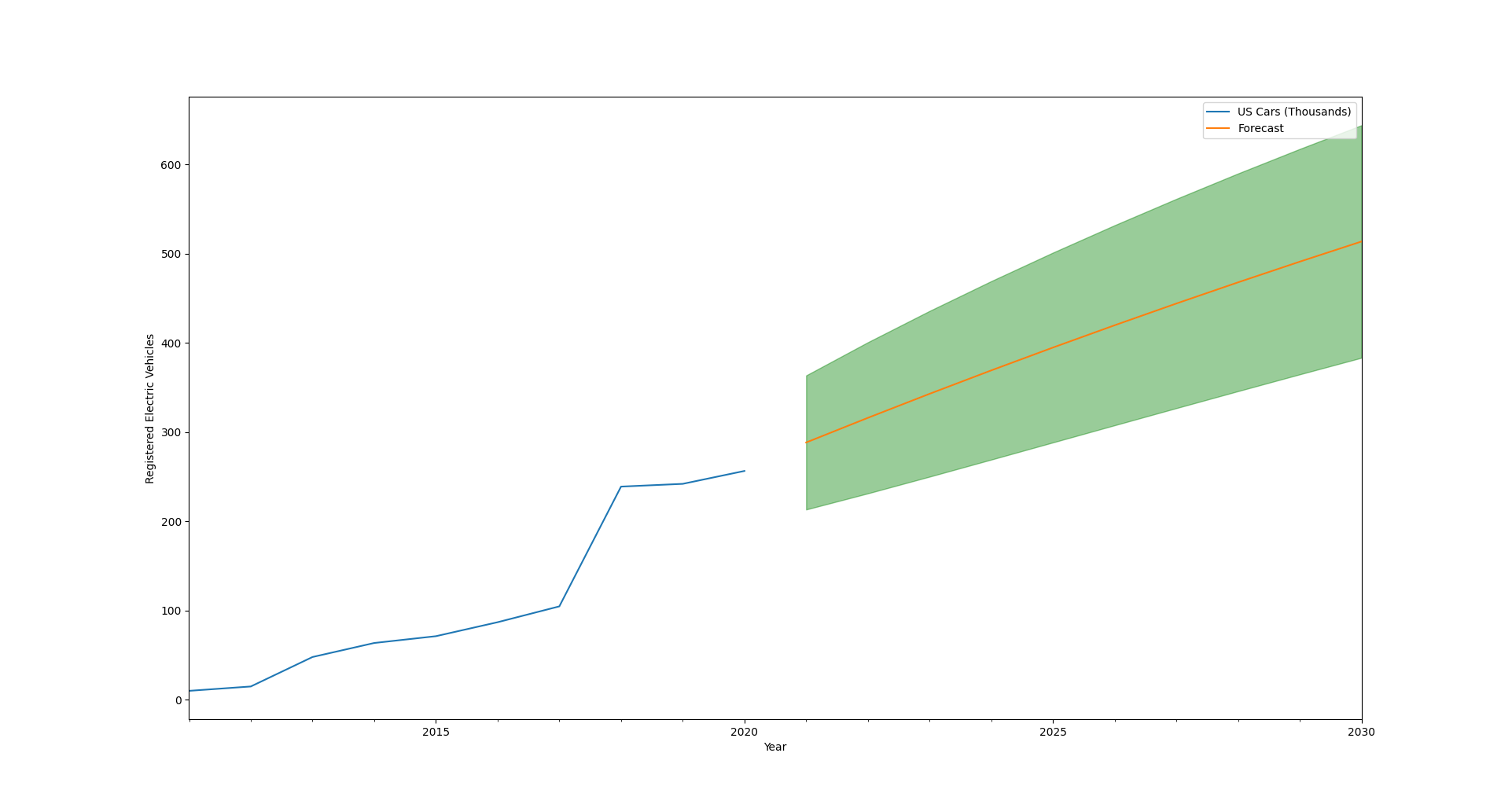
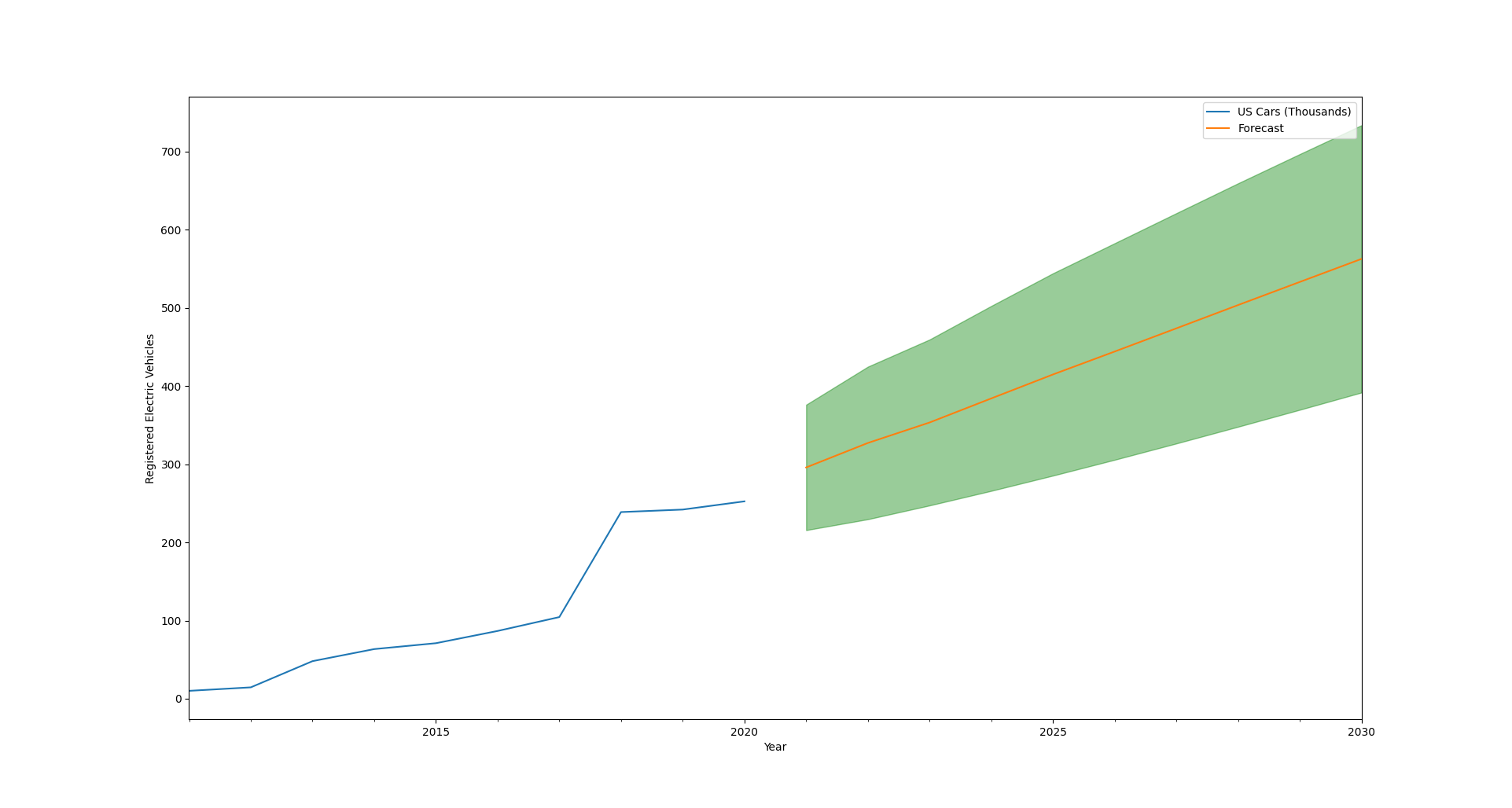


Figure 11: EV Sales Forecast Figure 12: EV Registrations Forecast

Both forecasts are for EV sales and EV registrations respectively. As one can see, they look very similar to one another both in actual values and forecasts. Therefore, if they have a strong positive correlation, one dataset could be used for both. Results show that they have a near perfect positive correlation at 0.9999. Therefore, EV sales will be used for both moving forward. For full sized versions of these images, please refer to the appendix.

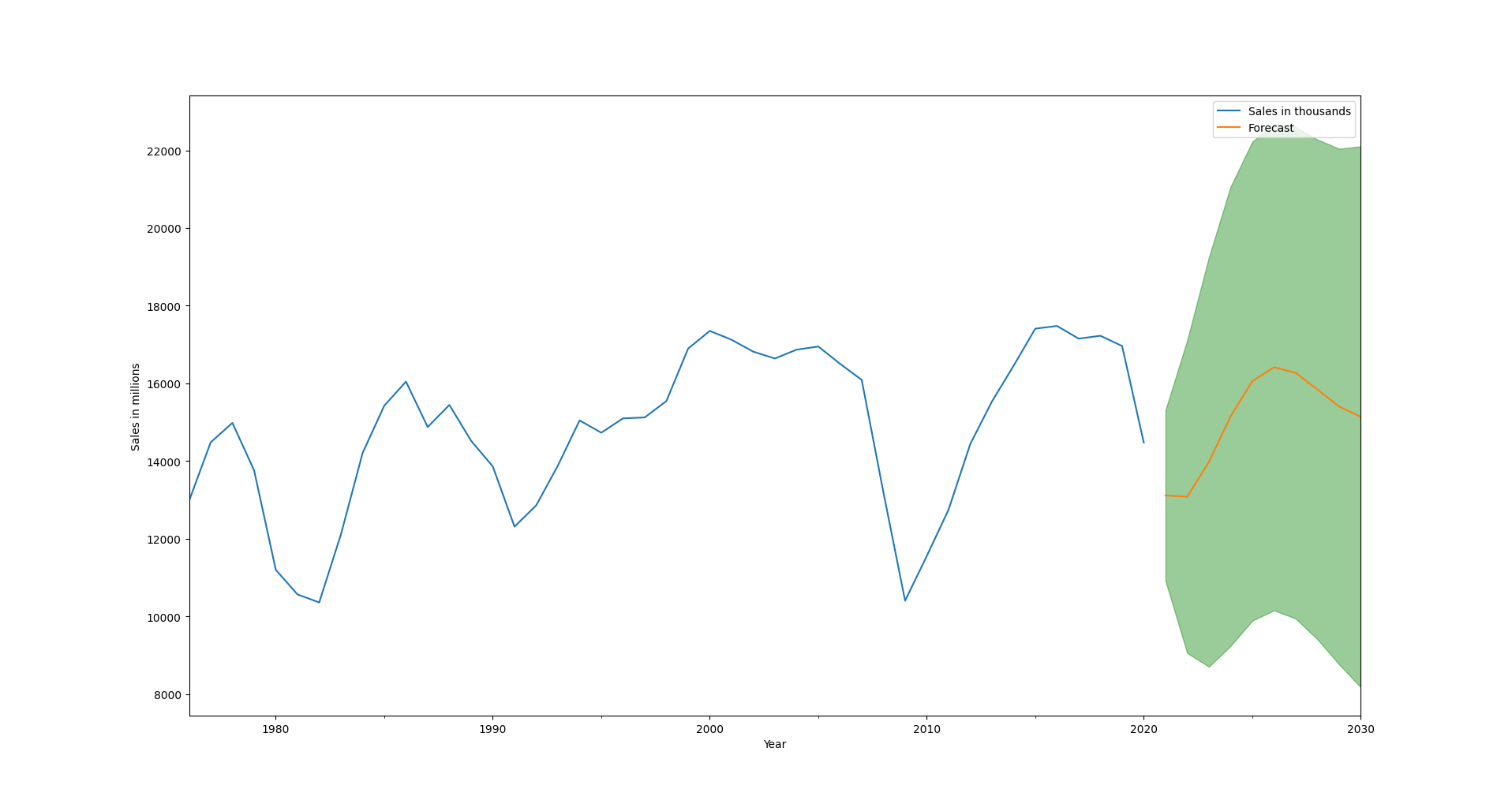


Figure 13: EV Sales

Meanwhile, vehicle sales continue to fluctuate and the model is able to forecast a similar behavior happening in the future. The confidence interval is somewhat wide but that is to be expected given said fluctuating nature as the possibility of it breaking that pattern exists.

Just from a glance, it is clear that EV sales would not reach 50% of light vehicle sales by 2030. This does illustrate that, out of all major automotive markets, the United States has a very low adoption rate of EVs. Therefore, hypothetical situations need to be drafted up in order to model the solution. For this model the following growths are used: linear, exponential, and S-curve. S-curves can be observed in technology where there is a huge spike in early adoption then slows down creating an S like curve hence the name.

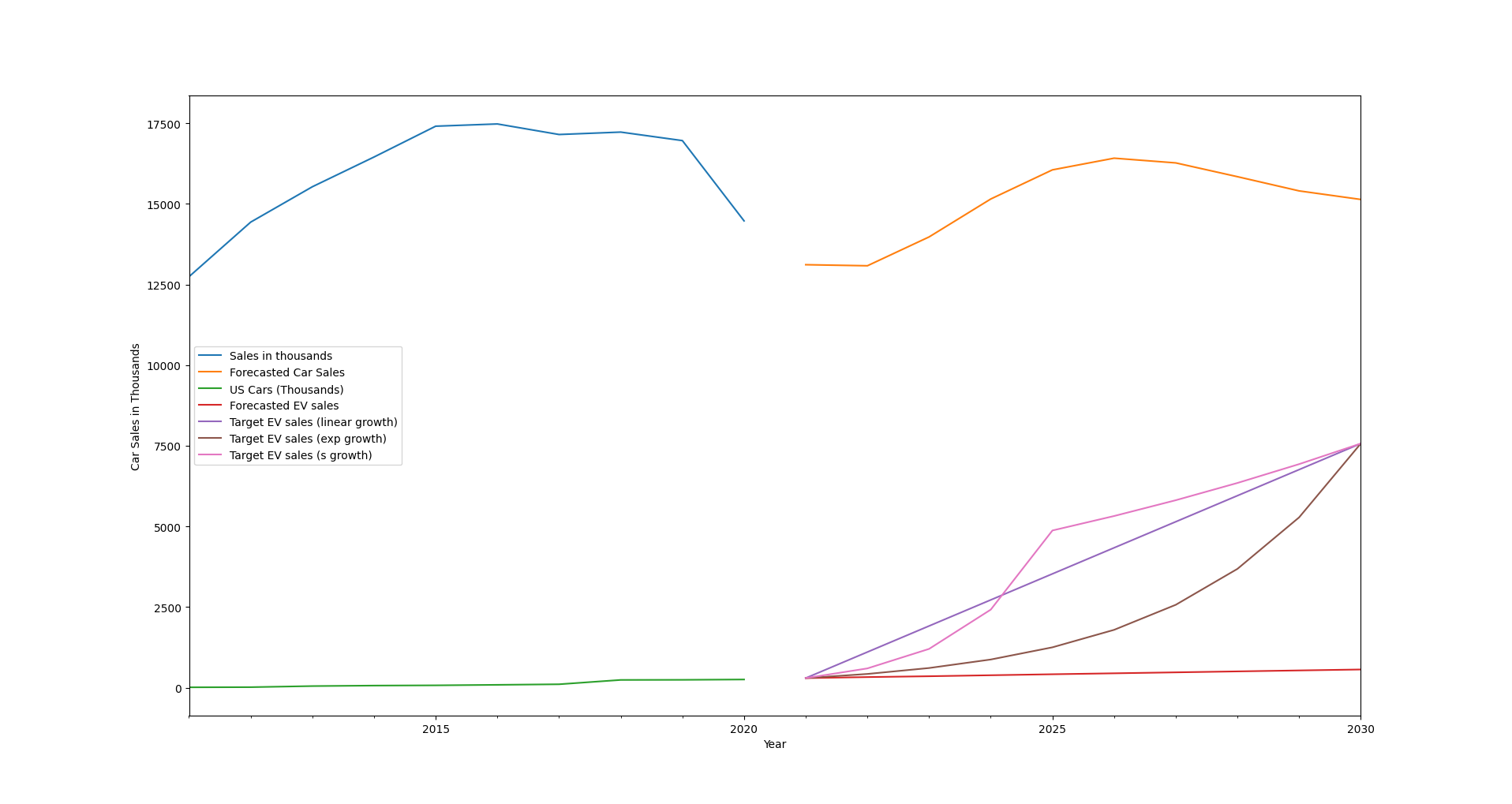


Figure 14: Sales Comparison

Now that the scenarios have been set up. It is time to determine the effects of having these many EVs on the road. According to EPA.gov, the average light passenger vehicle emits 4.6 metric tons of carbon dioxide every year. By subtracting that amount multiplied by the number of electric vehicles sold that year; and subsequently registered as implied by the correlation, on top of subtracting the EVs that have lasted 10 years, the result is as follows.

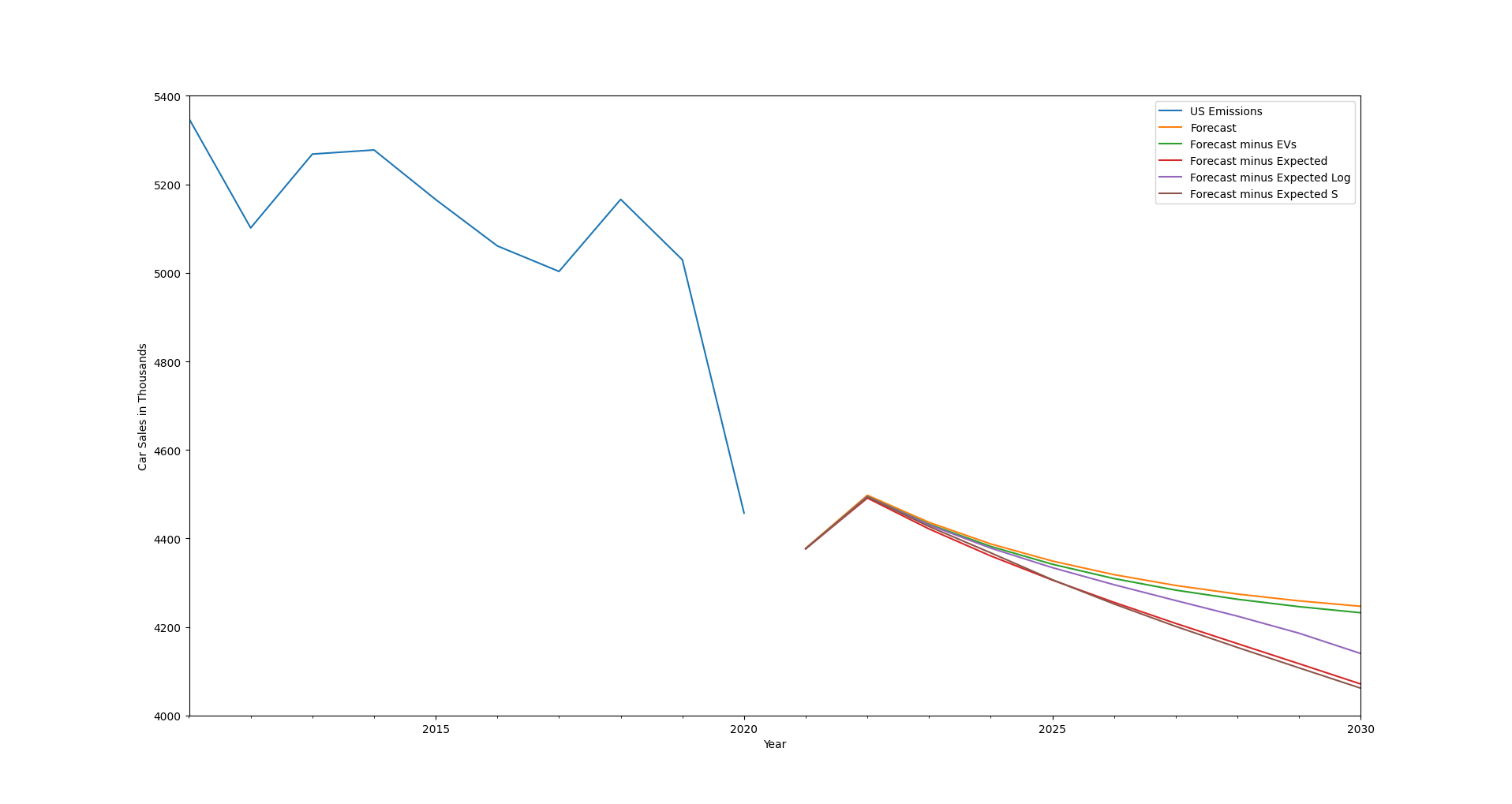


Figure 15: Effect of EV Sales on Emissions

The thing of note here is that despite three of the curves reaching 50% sales by 2030, their effect on reducing carbon emissions is different. This adds up as the earlier EVs get sold, the more internal combustion engine vehicles get replaced over time therefore decreasing the amount of emissions per year. For a more visual representation, these are the number of registered EVs per year in these scenarios.

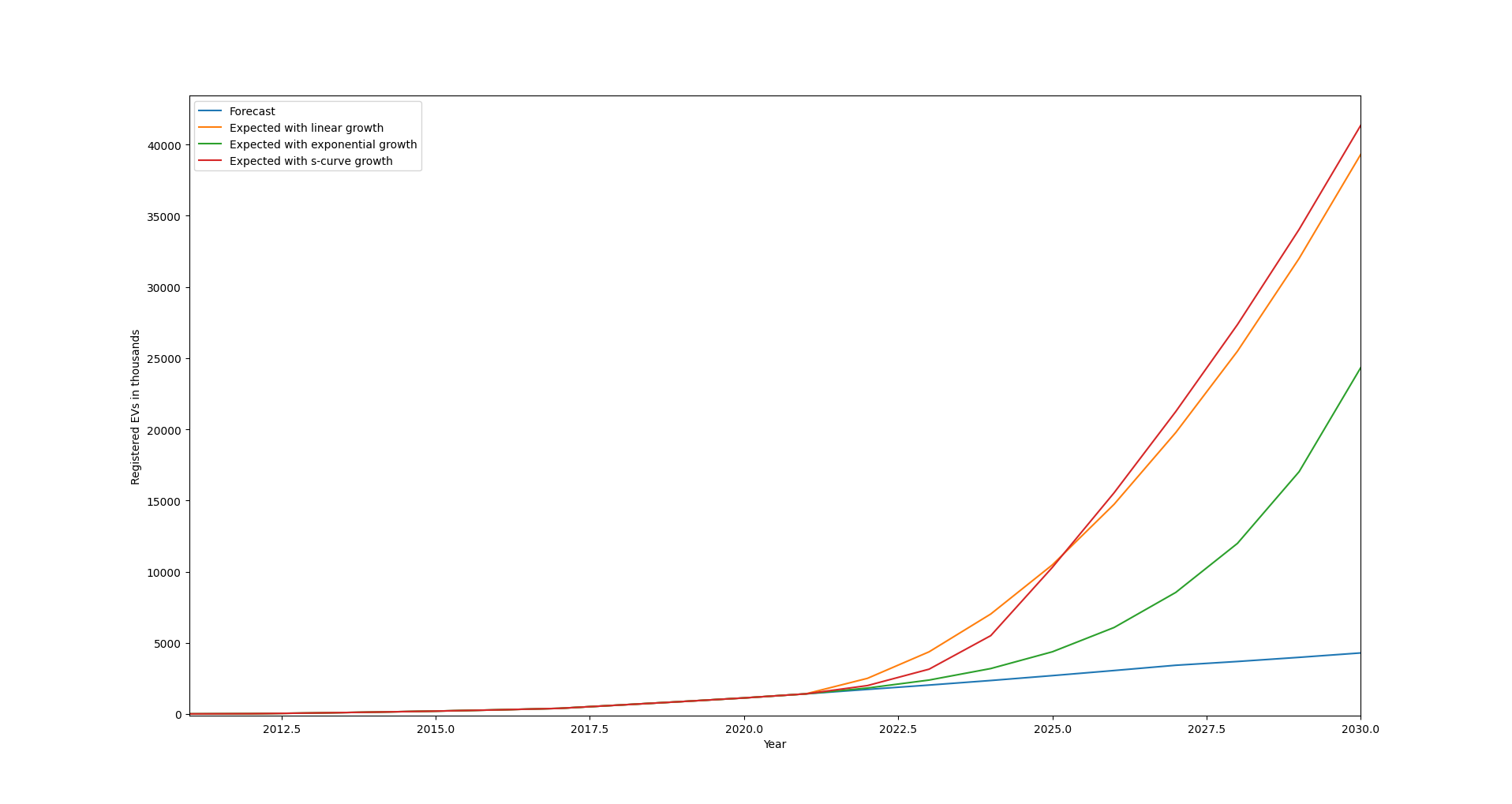


Figure 16: Cumulative Registered EVs

**Verification and Analysis**

Estimated number of registered vehicles by 2030: 309,309,894

Percentage of light passenger vehicles: 35%

Number of light passenger vehicles: 108,258,463

Estimated US emissions by 2030: 4246.67 million tonnes

Percentage contributed by light passenger vehicles: 15%

Emissions caused by light passenger vehicles: 637 million tonnes

Market Share = Registered EVs / Registered Light Vehicles

Theoretical Decrease = Market Share \* 637 million tonnes

Actual Decrease = Emissions w/o solution - Emissions w/ solution

By using these formulas, with the data collected during modeling, the results are as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | % of EVs on the road | Emissions by 2030 | Decrease from forecast | Theoretical decrease |
| Predicted | 3.97% | 4232.13 | 14.54 | 25.26 |
| 50% Linear Growth | 36.32% | 4071.00 | 175.67 | 231.37 |
| 50% Exp. Growth | 22.48% | 4139.91 | 106.76 | 143.22 |
| 50% S Growth | 38.21% | 4061.62 | 185.05 | 243.37 |

Table 1: Growth in EVs vs CO2 emissions

As is evident, the modeled decrease is always lower than the theoretical decrease. There are many reasons as to why this is the case. Most of which is based on the fact that the theoretical decrease is calculated naively. There are multiple factors which can show that the decrease in emissions is going to end up lower than the theoretical value. Two give two examples, the average American household owns 1.85 cars. While, given the nature of averages, this does affect the average emissions of vehicles per year, it does not take into account which vehicle from households that own multiple get used more. Of course this also can be used as an argument to claim that the actual value can be higher than the theoretical one if most of the families use EVs as their main car or if all the cars they own are EVs. However, an issue of EVs at the moment is range. While it has been improving over time, range and, by extension, charger availability remains a legitimate concern. Especially considering that the United States has the highest yearly mileage out of every country by a decent margin. Given that, it makes the scenario that multiple car owning households use their internal combustion engine vehicles for longer trips, at the very least, much more likely. Therefore, one can reasonably expect the modeled values to be lower than the theoretical values. What matters is that the results are consistently lower without huge fluctuations in how much there is a decrease.

Looking at the results, it is clear why there is a push for EVs. Once most of the issues have been solved and transitioning from combustion engine vehicles become painless, it is a straightforward way to decrease the emissions on one of the United States’ largest sectors when it comes to emissions. However, economical and logistic reasons make them a hard sell for most.

## **Energy Efficient Buildings**

## Among the few solutions for the carbon reduction method we went over, one of the solutions we will discuss is the use of energy-efficient buildings. Buildings generate nearly 35% of the annual global CO2 emissions. While one might question why the unreasonable amount of CO2 has been gathered because of buildings alone, the primary reason is that emissions are separated into two different sections, operation and building. Twenty-two percent of those emissions are because of building operations, whereas the remaining 13% is for the manufacture of raw materials to construct the buildings. These two numbers add up to the hefty amount of CO2 emissions. Even among those emissions, only about 5% of large buildings emit an equal amount to 95% of the small buildings. Hence, if the remedies to reduce these reductions are put in place, we could definitely see a positive curve in global CO2 emissions within our environment.

## **Data Collection**

We got our initial data for the world CO2 emission from oil and gas company Bp. Luckily, there were many resources for CO2 data on the internet, as this is one of the most concerning problems at hand. So, getting the raw data was not as challenging as expected. Architecture 2030 is a company that is researching extensively to reduce carbon emissions in the environment, with their initial target to make buildings zero carbon emission by the year 2030. Considering their data, we modified or transformed our initial dataset from Bp to predict specific scenarios such as an increase in buildings. Then, the data was fed into the ARIMA model to find the prediction in each of these scenarios.

## **Selecting the parameters**

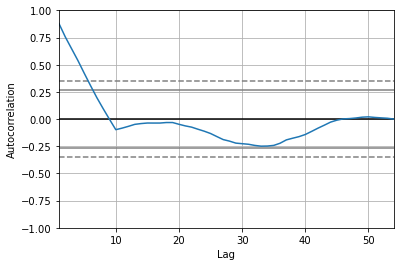
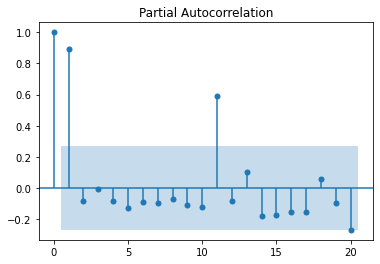


Figure 17: Plot of partial autocorrelation Figure 18: Plot of autocorrelation

We select the p and q values from analyzing the partial auto-correction and autocorrelation graphs. Here, the partial auto-correlation has three points out of the bound, which gives us 3 for the value for q. For the value of p, we look at the curve of autocorrelation, specifically where the graph intersects the bound of the graph. In this case, the p-value for the ARIMA model will be 6 or 7 based on the lag values of the autocorrelation. We repeat this step for all six data scenarios fed into the ARIMA.

## 

## 

Figure 19: Diagnostics plot for ARIMA model

By selecting the p and q values, we get the following graph for the ARIMA model. The residual test showcases that the relative prediction remains pretty stationary throughout except in the year 1970, which is because of the sudden spike in the raw data for that time. The p and q values were calculated from the previous two graphs. The acquired ARIMA data is used for the prediction. Again, we repeat the same process for all the scenarios.

## **Scenarios for the Modeling and Prediction**

We go through each scenario and run the data through the ARIMA model. Afterward, we retrieve the one-step-ahead forecast prediction from the ARIMA model.

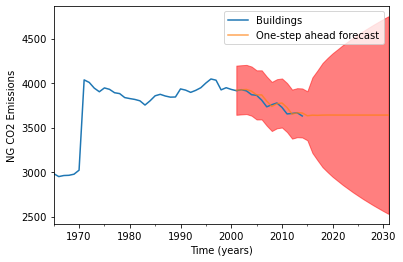


Figure 20: Emission forecast until 2030

The above graph is derived from the raw data where we can see the prediction of CO2 till year 2030.

**1st SCENARIO:**

In an ideal environment with our current innovations for infrastructures, the study by Architecture 2030 shows that the carbon reduction can reach up to 40% by the year 2050 if we pursue reduction extensively. Our first prediction is as follows:

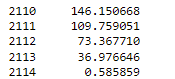
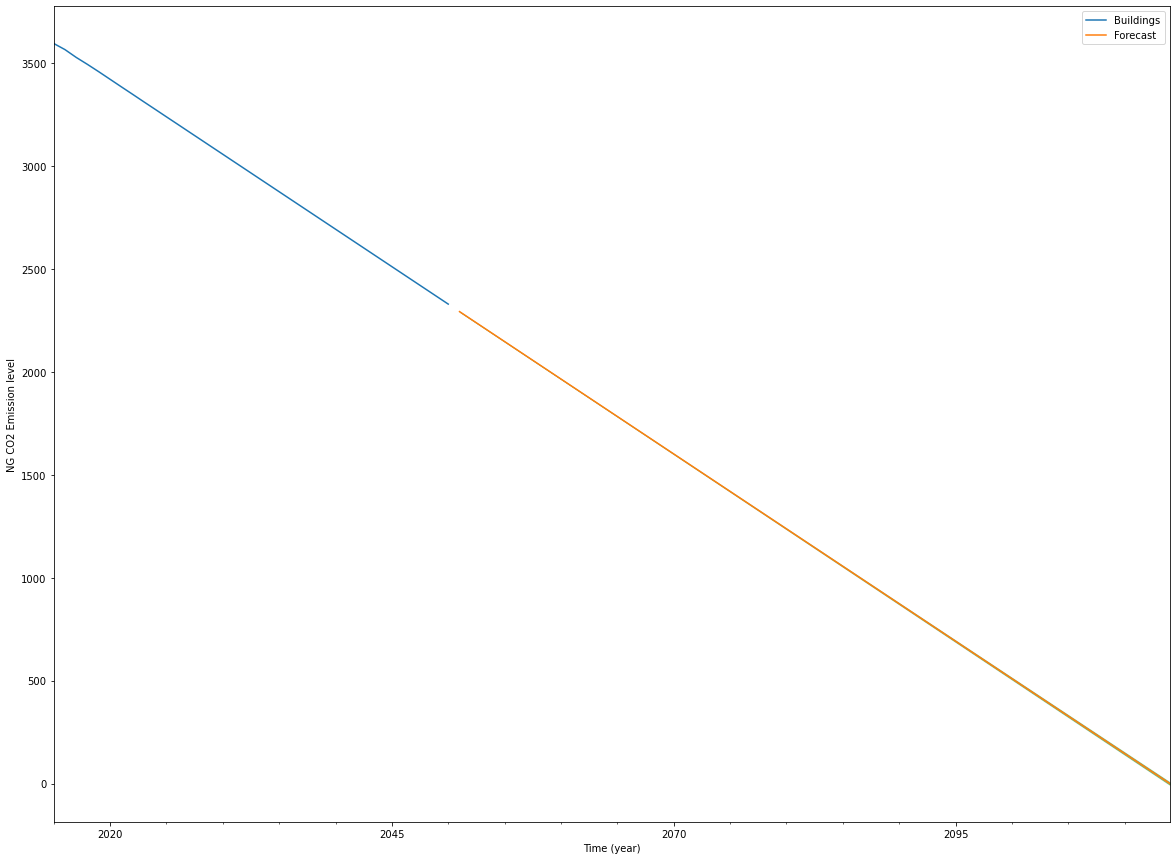


Figure 21: New emissions prediction with current solution for CO2 reduction

The graph predicts that it would take almost another century, i.e till 2114 with current rate till we reach zero carbon emissions for buildings keeping every other variable constant, hence the linear graph.

We then move on to the other scenarios, introducing the carbon reduction methods on the predicted data we received.

The modes of reduction can be broadly explained in three terms.

●Reuse – reuse of existing materials and designing for deconstructions

●Reduce – Include material optimization and the specification of low to zero carbon materials

●Sequester – including the design of carbon sequester sites ad the use of carbon sequestering materials

**2nd SCENARIO:**

Considering the increase in buildings, the growth rate of the buildings is very high and is expected to almost double by the year 2060. This keeps the increased building rate in the and CO2 reduction rate the same in the equation. We get the following graph which demonstrates the hyped CO2 emissions over the years.

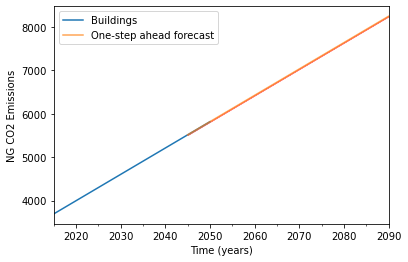


Figure 22: Emissions prediction with increasing number of buildings (doubled by 2060)

**3rd and 4th SCENARIO:**

In the following scenarios, we include CO2 reduction methods on only new-formed buildings for the left graph. Afterward, we apply CO2 reduction methods on old buildings. As a result, we got two distinctly different graphs with a vast improvement in the emission of CO2.

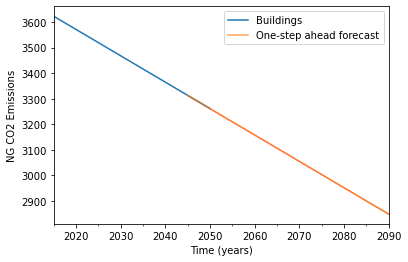
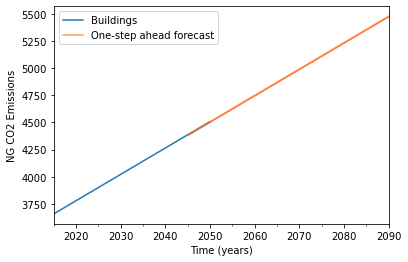


Figure 23: CO2 emissions prediction with new renovated buildings (left) and new renovated/reconstruction of old buildings (right)

**5th Scenario and Analysis:**

All the predictions were made using ideal situations, ignoring all other variables. However, realistically, there are always uncertainties and non-ideal situations. We have to consider the following variables and solution approaches below.

* By the year 2060, building numbers are likely to double. That means CO2 emission sources are likely to double as well.

Steps to reduce those emissions

* Existing buildings will need to undergo energy upgrades involving a combination of improvements in EEB, shift to electric or district heating systems powered by carbon-free resources compared to the gas powered, become more reliant on renewable energy, and electrification of buildings
* New buildings will be required to implement all those resources. Additionally, reuse of materials is a must because production of construction materials are responsible for a lot CO2 emissions as well

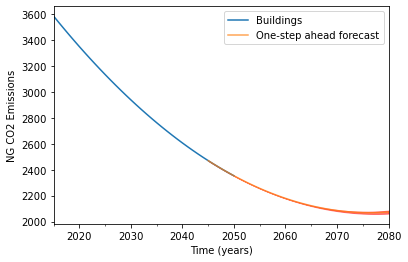


Figure 24: CO2 emission prediction with all variables from scenario 1 through 4

We get the curved predictions when we include all the variables, which is understandable because the number of buildings will continuously increase. With current technology, we can not obtain zero carbon emissions from buildings if we practically put it in action, which can be seen from the graph's curve. It might even increase as the years go on, but CO2 emissions are greatly reduced in contrast to the initial prediction. Moreover, zero carbon emissions are not possible in our case. The best we can do is decrease it.

## **Conclusion**

The effects of climate change have existed for millions of years and are evident now more than ever. While it is the process of how the planet evolves and is not something humans can completely stop, we can reduce the adverse effects caused by human actions to prevent accelerated and unnatural changes to the climate. Our project puts forward different means of controlling carbon emissions so that the people and the ecosystem could be well preserved for the future generations to come. We put forward the data gathered and ran it through the ARIMA model and acquired the respective conclusion for each of the approaches. For our first approach, renewable resources; the original and predicted values of a majority of our time series didn’t show a similar pattern. Mean absolute percentage error was only 6% for CO2 emissions, indicating how much closer the forecast is than it actually looked. An analysis of carbon capturing and sequestration resulted in the conclusion that with the current rate of growth in CCS facilities, net zero carbon dioxide emissions will be achievable by the year 2216 based on the forecast obtained with our ARIMA model. In order to achieve the Paris Agreement’s goal of net zero by 2050 dramatic increase is necessary to the number of CCS facilities produced in the next few years; however, due to the magnitude of growth needed and limitations imposed by costs, land use, and time, achieving net zero by 2050 using only CCS technology may be infeasible. Electric Vehicles results were consistent with what is expected. Less cars that emit CO2 mean less emissions. However, the model does not measure emissions caused by producing EVs nor whether the power grid can handle it. Finally, use of EEB can significantly reduce the CO2 emission from building sectors if planned and implemented efficiently and have the potential to even have lesser CO2 emission than today while doubling the number of buildings.

References

“Accelerating to Zero by 2040!” Architecture *2030*, architecture2030.org/accelerating-to-zero-by-2040/#:~:text=For%20embodied%20carbon%20%E2%80%93%20a%2040,entire%20built%20environment%20by%202040.

Berend Smit, et al. *Introduction To Carbon Capture And Sequestration*. Imperial College Press, 2014. *EBSCOhost*, search-ebscohost-com.ccny-proxy1.libr.ccny.cuny.edu/login.aspx?direct=true&db=e000xna&AN=711885&site=ehost-live.

Carlier, Mathilde. “Light vehicle retail sales in the United States from 1976 to 2020.” *Statista*, Statista, 4 Aug 2021, https://www.statista.com/statistics/199983/us-vehicle-sales-since-1951/

“CO₂ Emissions: Energy Economics.” *Bp Global*, www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy/co2-emissions.html.

“Global Status Report.” *Global CCS Institute*, www.globalccsinstitute.com/resources/global-status-report/.

“Greenhouse Gas Emissions from a Typical Passenger Vehicle.” *EPA*, March 2018, https://www.epa.gov/greenvehicles/greenhouse-gas-emissions-typical-passenger-vehicle

Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2.

“Paris Agreement.” *Climate Action*, ec.europa.eu/clima/eu-action/international-action-climate-change/climate-negotiations/paris-agreement\_en.

Prabhakaran, Selva. “Arima Model - Complete Guide to Time Series Forecasting in Python: ML+.” *Machine Learning Plus*, 22 Aug. 2021, https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/.

Prabhakaran, Selva. “Vector Autoregression (VAR) - Comprehensive Guide with Examples in Python.” *Machine Learning Plus*, 22 Nov. 2021, https://www.machinelearningplus.com/time-series/vector-autoregression-examples-python/.

Rasheed, Rayhaan. “Why Does Stationarity Matter in Time Series Analysis?” *Medium*, Towards Data Science, 12 July 2020, https://towardsdatascience.com/why-does-stationarity-matter-in-time-series-analysis-e2fb7be74454.

“Renewable Energy.” *Center for Climate and Energy Solutions*, 10 Nov. 2021, https://www.c2es.org/content/renewable-energy/.

Rhode, Emily. “Carbon Capture and Storage (CCS) Pros and Cons.” *Treehugger*, Treehugger, 13 Aug. 2021, www.treehugger.com/carbon-capture-and-storage-ccs-pros-and-cons-5120005.

Ritchie, Hannah, and Max Roser. “Emissions by Sector.” *Our World in Data*, 11 May 2020, ourworldindata.org/emissions-by-sector.

Shaftel, Holly. “Climate Change Adaptation and Mitigation.” *NASA*, NASA, 13 Dec. 2021, https://climate.nasa.gov/solutions/adaptation-mitigation/.

Shinn, Lora. “Renewable Energy: The Clean Facts.” *NRDC*, 15 June 2018, https://www.nrdc.org/stories/renewable-energy-clean-facts.

Smiljanic, Stasha. “How Many Americans Own Cars?” *Policy Advice*, Policy Advice, 15 September 2021, https://policyadvice.net/insurance/insights/how-many-americans-own-cars/

“Statistical Review of World Energy: Energy Economics: Home.” *Bp Global*, https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html.

“Trends and developments in electric vehicles markets.” *IEA*, 2021, https://www.iea.org/reports/global-ev-outlook-2021/trends-and-developments-in-electric-vehicle-markets

“US Vehicle Registration Statistics.” *Hedges Company*, 2021, https://hedgescompany.com/automotive-market-research-statistics/auto-mailing-lists-and-marketing/

Appendix

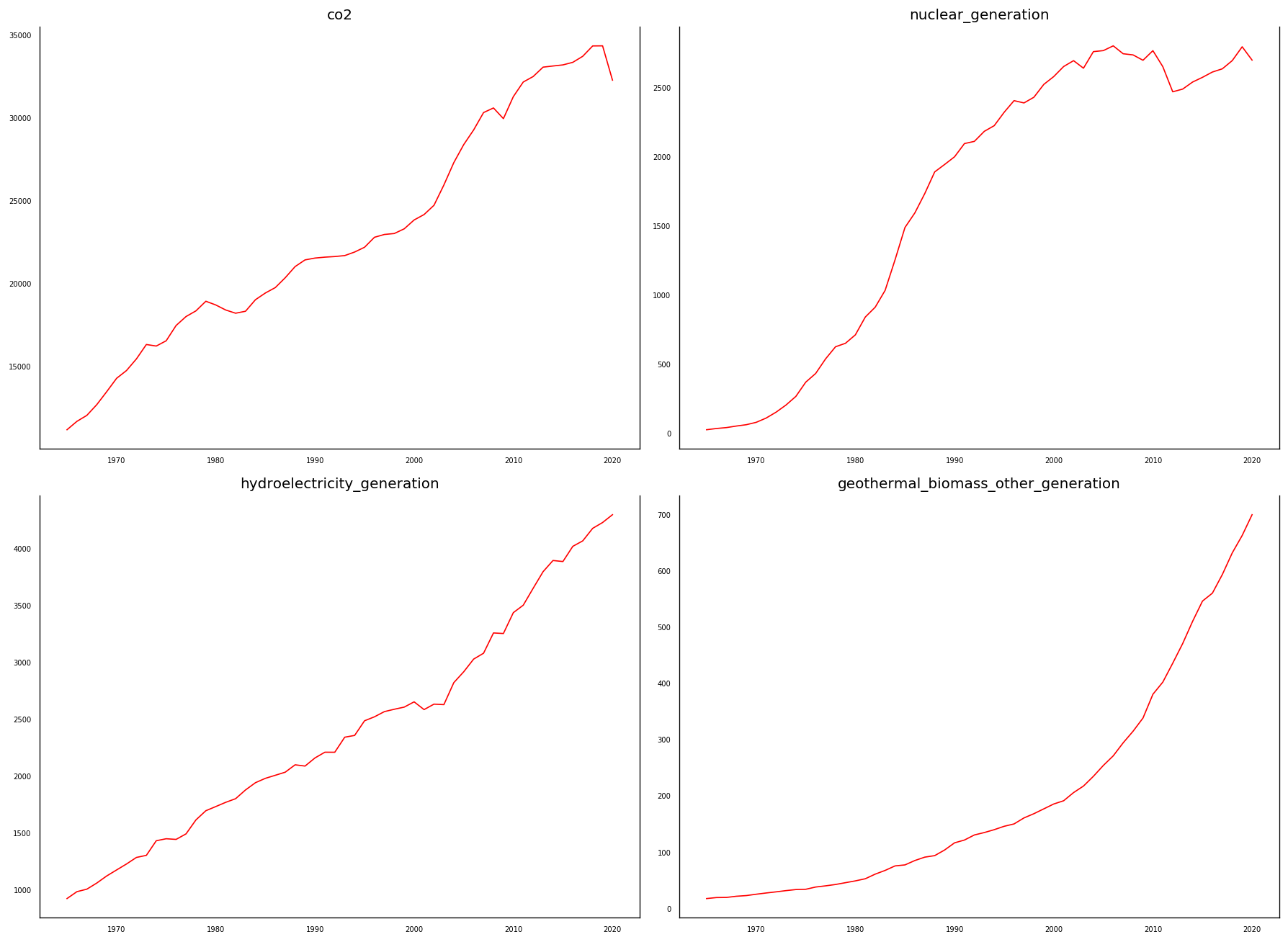


Figure A1: Time Series Visualization

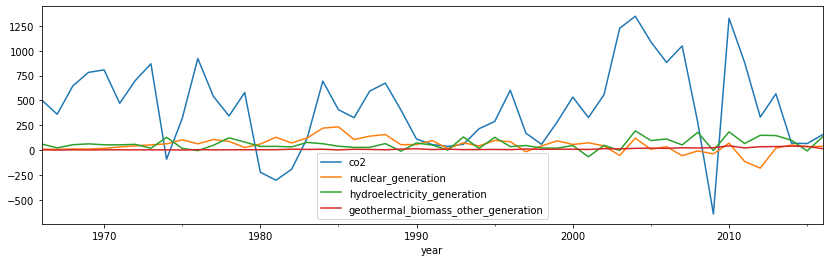


Figure A2: First Difference Plot

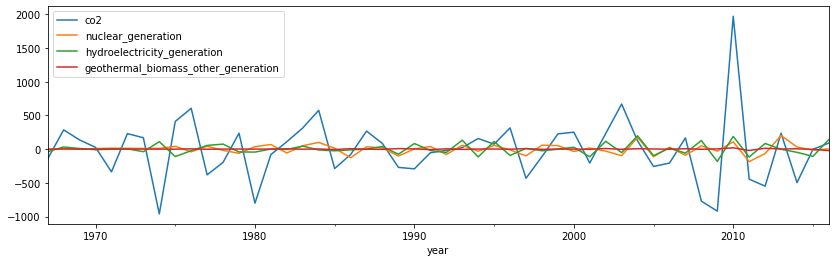


Figure A3: Second Difference Plot

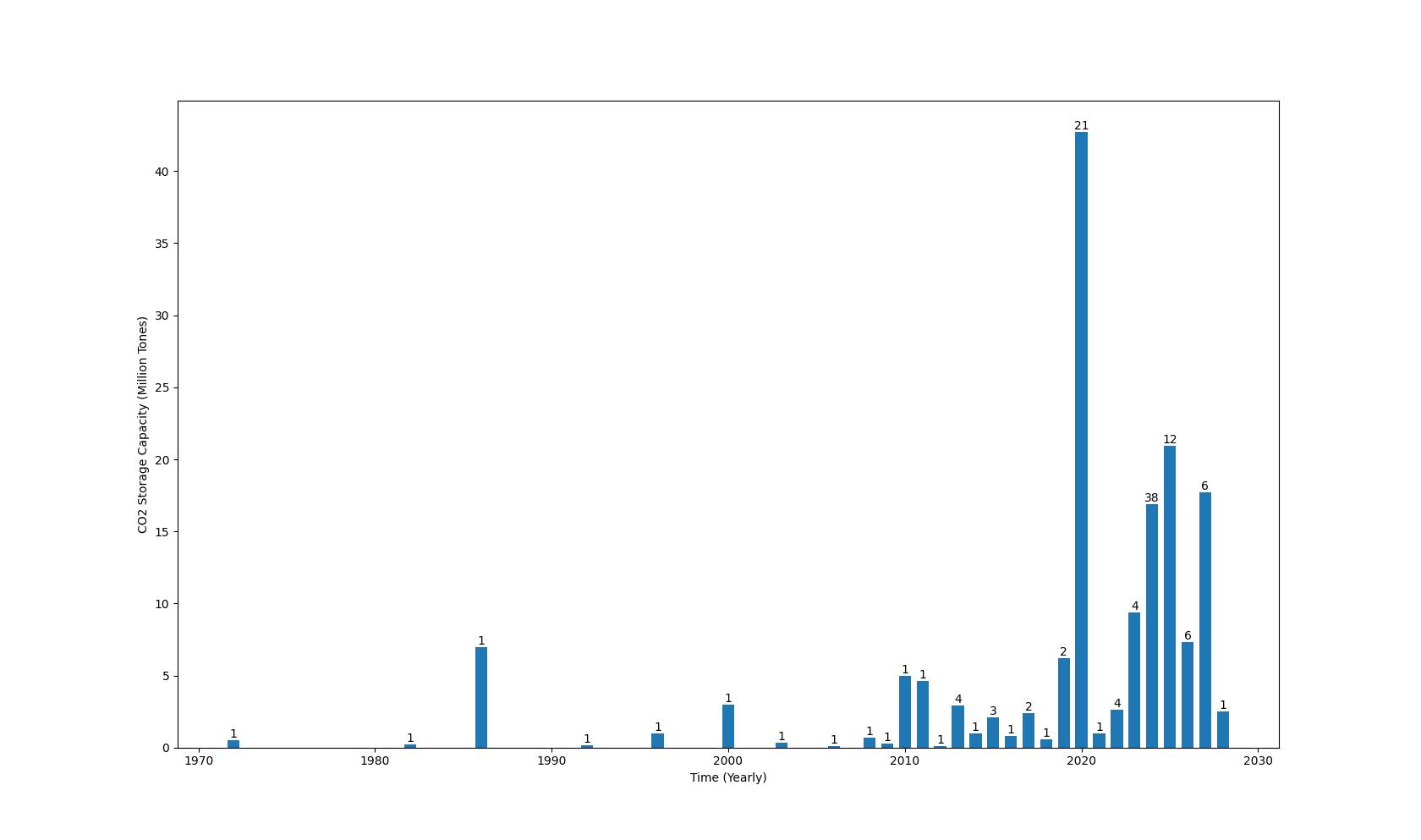


Figure A4: CCS Facilities Opened Since 1972 and their CO2 Storage Capacities

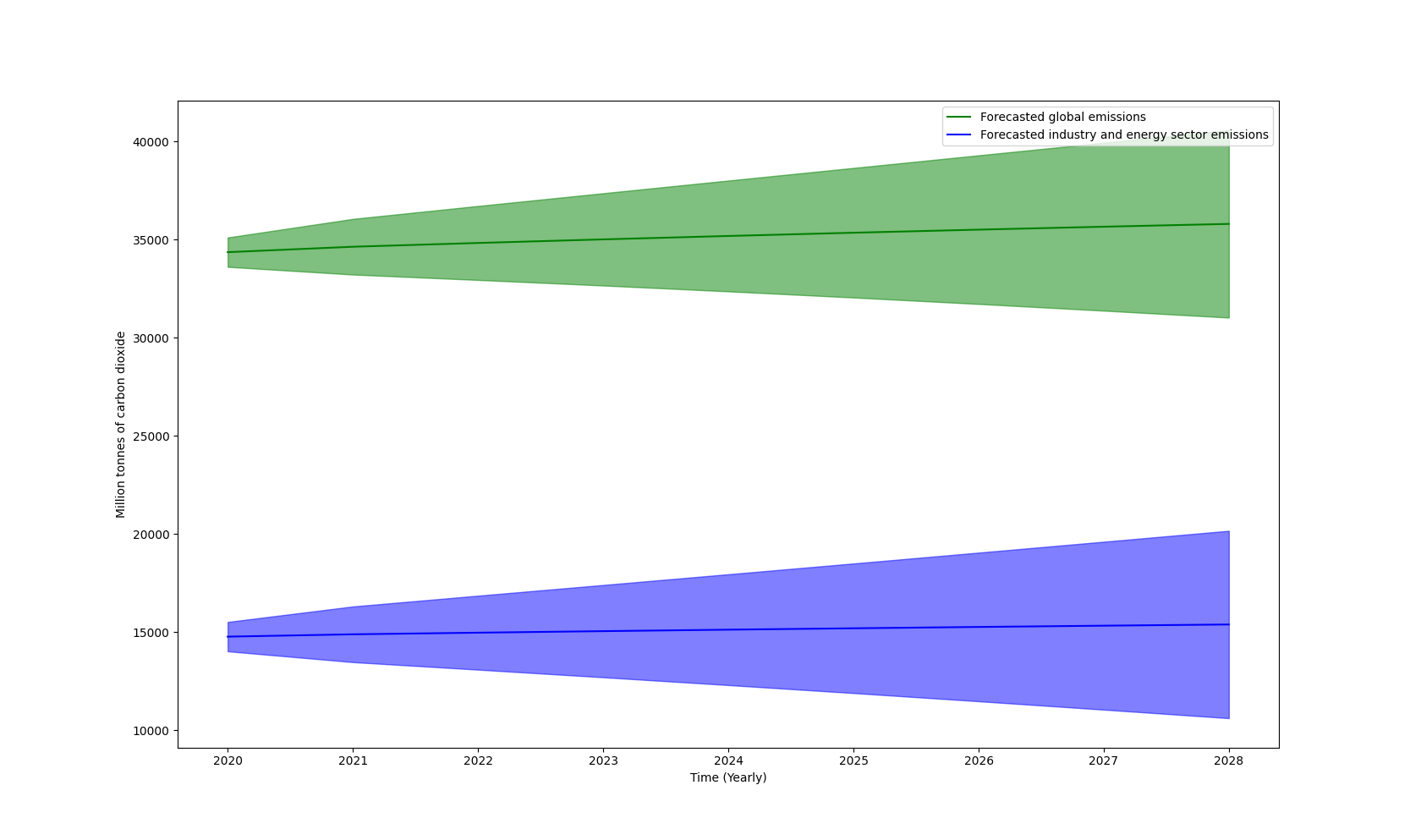


Figure A5: Forecast of Global CO2 Emissions Compared to Energy and Industry Sectors

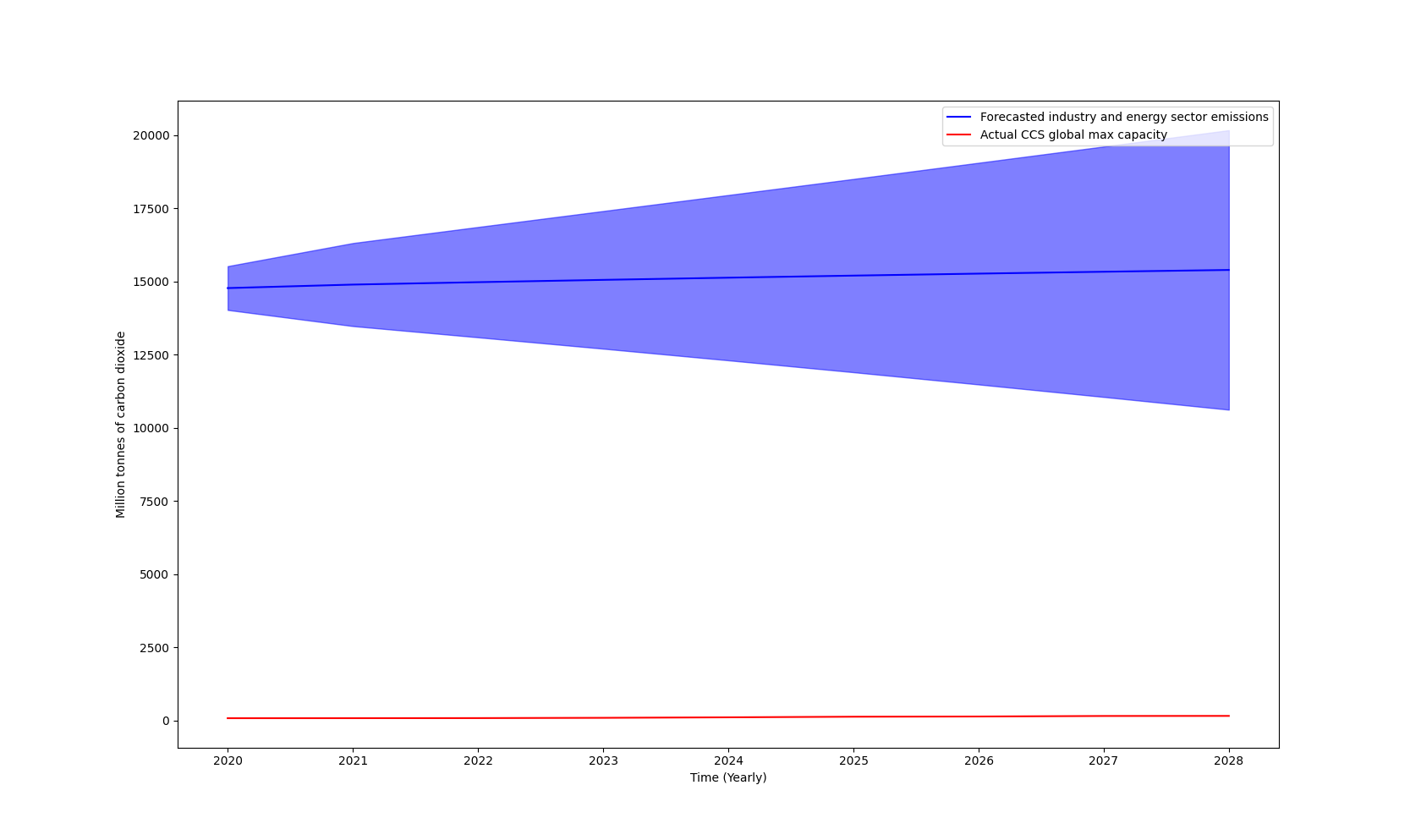


Figure A6: Forecast of Energy and Industry Sector Emissions Compared to Observed CO2 Storage Capacities

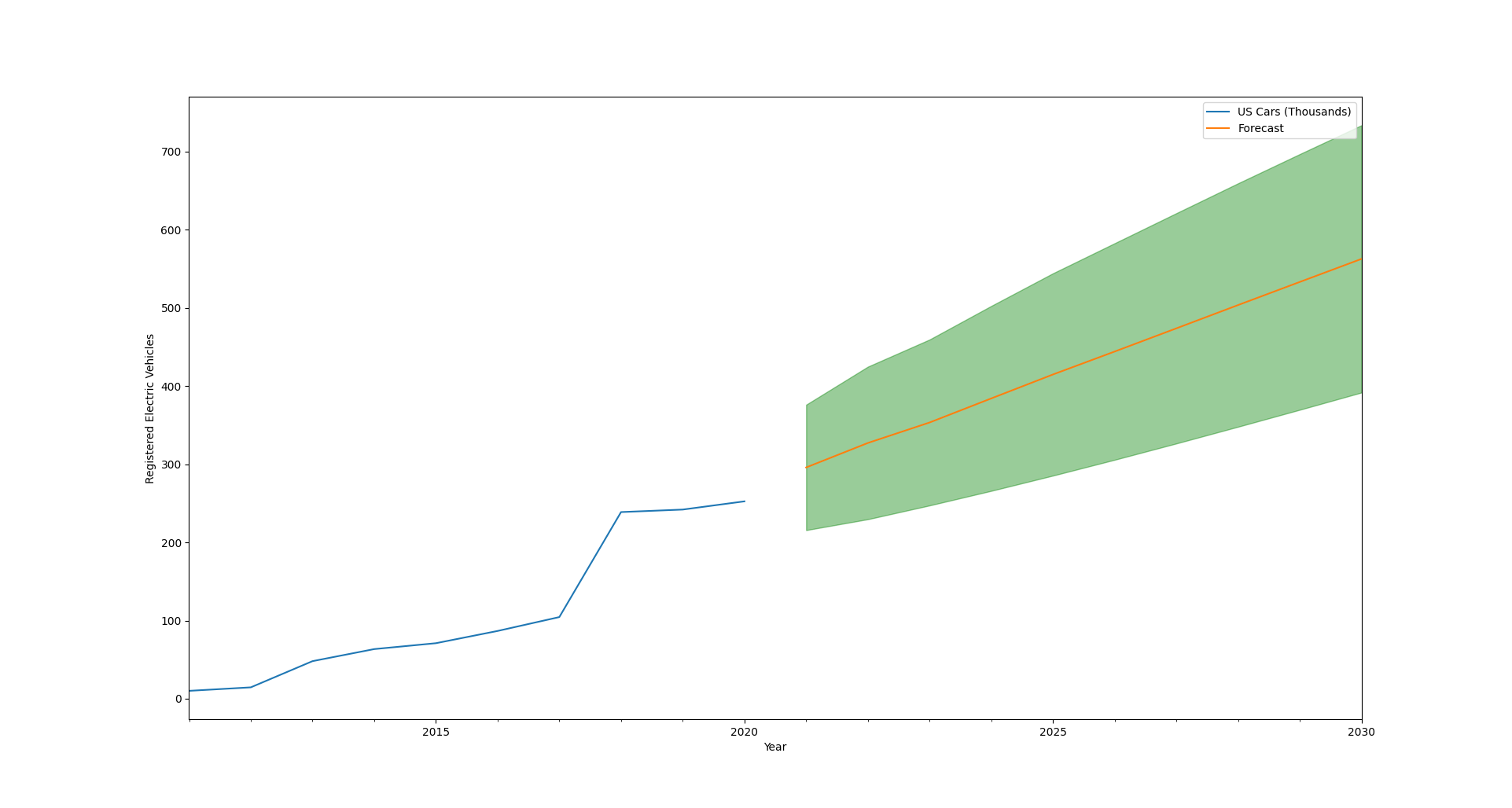


Figure A7: Forecasted Electric Vehicle Sales until 2030

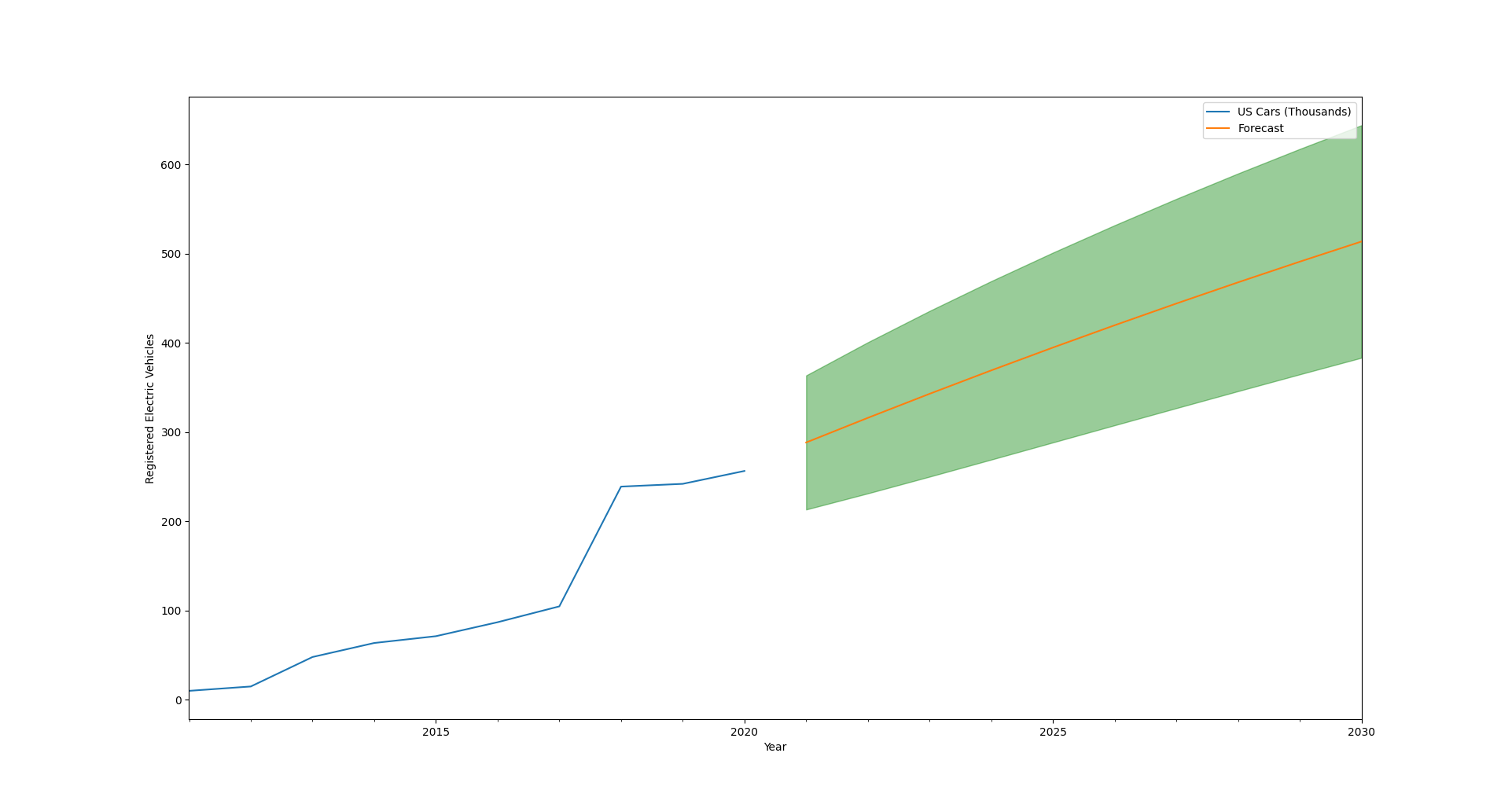


Figure A8: Forecasted Electric Vehicle Registrations until 2030