Time Series Data Analysis (DSA301)

AY 2023-2024 Term 2

Group Project

Name	Student ID
Rayner Yeoh Wei-En	01366079
Chua Yu En, Cassandra	01423143
Janias Tan Jia Xuan	01416788
Syvester Song Shi Jun	01407528

Executive Summary

Our team has decided to work on option 1 of the project options provided - ‘Climate Change, Human Cause or not?’. Our project aims to investigate the global effects of climate change and human activity. Our research will involve reviews of literature on climate change and human activities.

We will focus on two parts to investigate if climate change is a product of anthropology: (i) using climate data, such as greenhouse gas levels, global average temperatures, sea levels, and other key variables to measure forecast future temperatures; and (ii) investigating the relationship between climate data and human activity. We will also include a robustness test, to account for the differences between empirical climate research and our statistical tests.

Literature Review: Climate Change

The consensus among scientists is that climate change represents one of the most significant ecological and social challenges of the 21st century (Dietz et al., 2020), along with efforts to show how human activities are the primary drivers of climate change (IPCC, 2014).

Relationship between Temperature and CO2 Emissions

The increased production of Carbon Dioxide (CO2) has significantly accelerated the atmospheric temperature and will remain the driving factor of climate change (Sarmiento & Gruber, 2002). A study has also shown that both temperature and CO2, an important greenhouse gas, have seen similar patterns due to the heating or cooling of the Earth's surface will lead to the concentrations of greenhouse gases in the atmosphere, which would cause further warming or cooling of the atmosphere (United States Environmental Protection Agency, n.d.).

Relationship between Temperature and Sea Levels

Research has also shown that temperatures affect sea levels greatly. The temperature of the Earth is primarily due to the accumulation of heat-trapping greenhouse gases, and more than 90% of this heat is absorbed by water bodies, causing water levels to rise as water expands (NASA, n.d.) The relationship appears to be unidirectional, with temperature changes causing sea level changes, and the relationship is affected by a combination of several factors (Ezer & Updyke, 2024).

Relationship between Human Activity and Temperature

Human activities have been the main contributor to climate change, according to a study of over 88000 climate-related papers published since 2012 (Lynas et al., 2021). The Intergovernmental Panel on Climate Change (IPCC), also confirms that the unprecedented rise in global temperatures since the 19th century has been predominantly caused by human activities rather than natural factors like solar or volcanic activity (IPCC). Excessive human activities have resulted in a higher temperature than pre-industrial levels, with warming ranging from 0.8 to 1.2 degrees (IPCC, 2018).

Relationship between Human Activity and CO2 Emissions

Human activities have also increased global greenhouse gases by 30% compared to pre-industrial levels (Sarmiento & Gruber, 2002). Studies have shown that the main production of CO_2 comes from human activity (CO2 Human Emissions, 2017). From here, it makes sense to conclude that if modelling CO_2 results in increased temperature, then human activities have indeed caused climate change.

The interaction between climate and our community has a complex historical context, where events such as the Little Ice Age suggest an intertwined relationship between human activity and climate variability (Dietz et al., 2020). This historical point of view emphasises the importance of our project's investigation, where we will look into more studies reinforcing that human activity plays a significant role, and provide empirical evidence and analysis from various perspectives.

Approach

To analyse if climate change is a product of anthropology, we wanted to investigate if global temperatures after 1970 truly deviated from the pre-existing trend. This means that for all our modelling processes, we will use training data till 1970 and forecasting post-1970s temperatures. We chose 1970 due to our domain knowledge, and machine learning considerations. For domain knowledge, our research on climate change has elucidated that human activity has increased sharply since the 1970s due to increasing industrialization and agriculture modernisation. Therefore, it makes sense to forecast pre-1970s and compare it to actual post-1970s temperatures. From a machine learning perspective, we attain more valuable conclusions by forecasting approximately 50 years of data than by having short forecasts. Hence, 1970 to 2024 will provide us with this range of forecasts. This will be done for the univariate temperature model and 3 variable climate models: CO2, sea levels, and temperature.

However, we will be building a new model, on top of the previous ones which forecasted post-1970s. This model will consist of more recent human data, such as GDP and agricultural production. The aim is to provide a more detailed analysis on human activity, as opposed to using the classic and traditional 3 variable climate data. However, the limitation of using this recent human data is that it narrows the time frame of our analysis, as most of this data starts from the 1960s and is recorded annually. As such, we decided to forecast post-2000s, using the 1960s to 2000s as the set that we would train and evaluate our models.

Lastly, we will choose the best model based on each of the different approaches mentioned, and compare the actual temperature with our model's forecast to determine if there is truly a statistically significant change in global temperature at a 95% confidence level.

Methodology

1. Dataset

Variable (column name)	Duration (annual)
Temperature (temp)	1743 - 2024
Co2 Emissions per global capita (co2_pc)	1850 - 2023
Change in Sea Levels (sea_level)	1880 - 2013
Methane per global capita (methanepc)	1961 - 2021
GDP per global capita (gdppc)	1960 - 2021
Energy Consumption per global capita (global_energypc)	1965 - 2022
Agriculture Production per global capita (foodpc)	1965 - 2020

Table 1: Corresponding duration for data used

1.1 Justifying chosen Indicators for Human Activity Data

Based on our research, we found numerous studies that show a connection between human activities and greenhouse gas emissions. Human activities like industrialisation and agricultural modernisation have a positive correlation with CO2 emission due to factors such as the increase in industrial output, fuel, and electricity consumption. Each of these activities contributed to the release of greenhouse gases, such as CO2 and methane, into the atmosphere, leading to changes in Earth's climate system (Shijie Li et al., 2019). These impacts encompassed changes in temperature, sea level rise, and extreme weather events (Chovancová & Tej, 2020) (Lapinskienė et al., 2017).

There were 2 limitations in the data collection process:

1. Historical climate data usually starts from the 1800s, but they are recorded annually. This means that our scope of analysis is confined to annual data, instead of monthly or quarterly data, especially when it involves these variables.

2. Human activity data typically starts from the 1960s, which means that the duration of our analysis is rather short. This is further pertinent when we attempt to perform a train-test split, in which we may have insufficient training data to work with.

Additional considerations from our data processing:

1. Normalised data - per capita

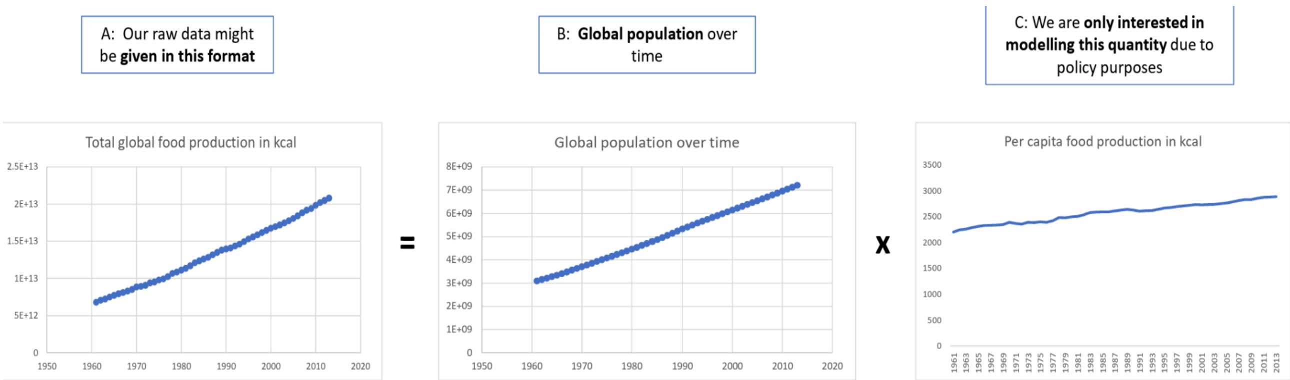


Image 1: From C, we are more interested in modelling per capita data

We work on per capita data for our variables instead of absolute global values to normalise the different variables. Therefore, we divided our variables over the global population on a year-by-year basis. Our main intuition behind this is to scale variables appropriately, such that variables with exceptionally large magnitudes (eg. C02 emissions) will not dominate over other variables especially those with low magnitudes (change in sea level, temperature) in our modelling processes. Additionally, per capita data can account for population growth.

Intuitively, as population increases, global carbon emissions increase. However, each individual might not be polluting more, and they might even be polluting less. Yet, absolute value data will not be able to give us these insights. By using per capita data, we could better evaluate the trends of emissions relative to population growth, to provide greater insights into our modelling process.

2. Quarterly and annual data

<u>Quarterly temperature data</u>

For the temperature dataset, we were able to attain quarterly data, hence univariate modelling of temperature could be done quarterly.

<u>Annual data for other variables</u>

However, after our extensive research for datasets, we found that other variables, such as sea levels and emissions, are only retrievable in annual form. As such, our multivariate analysis will be conducted based on annual forecasts.

With these in mind, we decided to have 3 different approaches for our modelling:

(i) modelling univariate quarterly temperature time series from the 1880s onwards, **to forecast post-1970s**

(ii) modelling historical climate data (green data) from the 1880s onwards, **to forecast post-1970s**

(iii) modelling climate and human activity data (green and yellow data) from the 1960s onwards, **to forecast post-2000s**

2. Exploratory Data Analysis (EDA)

Before model building, we conducted EDA for both green and yellow data to see how each respective time series behaves over time.

2.1 Visualisation

2.1.1 Visualisation of temperature data (Quarterly)

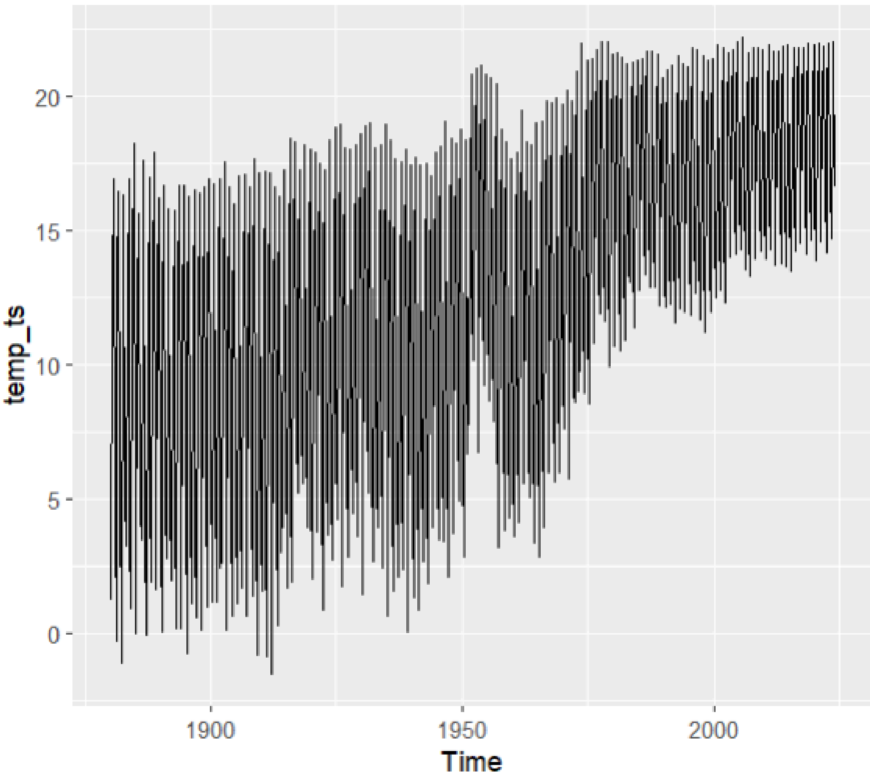


Image 2: Quarterly temperature data from 1880 to 2024 shows an upward trend

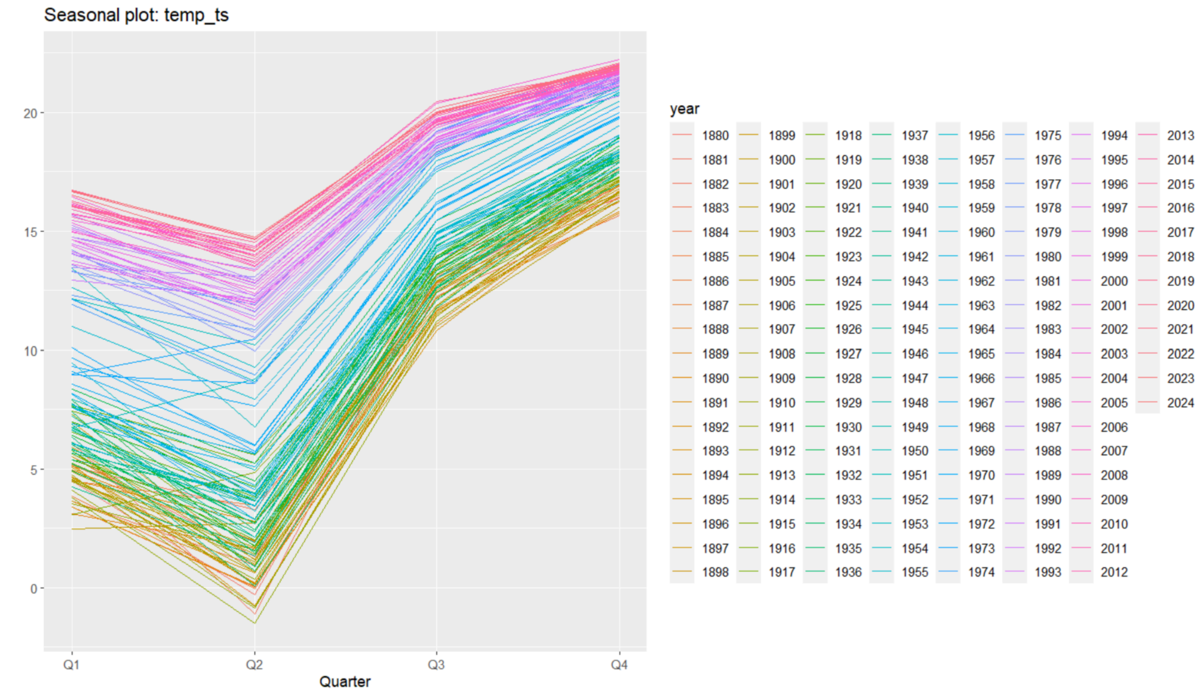


Image 3: Seasonality in each year

From our quarterly temperature time series, we see an upward trend, increasing even sharply after the 1970s. Zooming into each year, the seasonal trend is that, assuming we start at quarter 1, then quarter 2 will see a sharp decrease in temperatures, then increase in quarter 3 and even more in quarter 4.

2.1.2 Visualisation of climate data (Annual)

We observe that all variables in climate data (green data) have an increasing trend. In particular, C02 emissions per capita have been increasing at a very high rate, as compared to sea levels and average global temperatures.

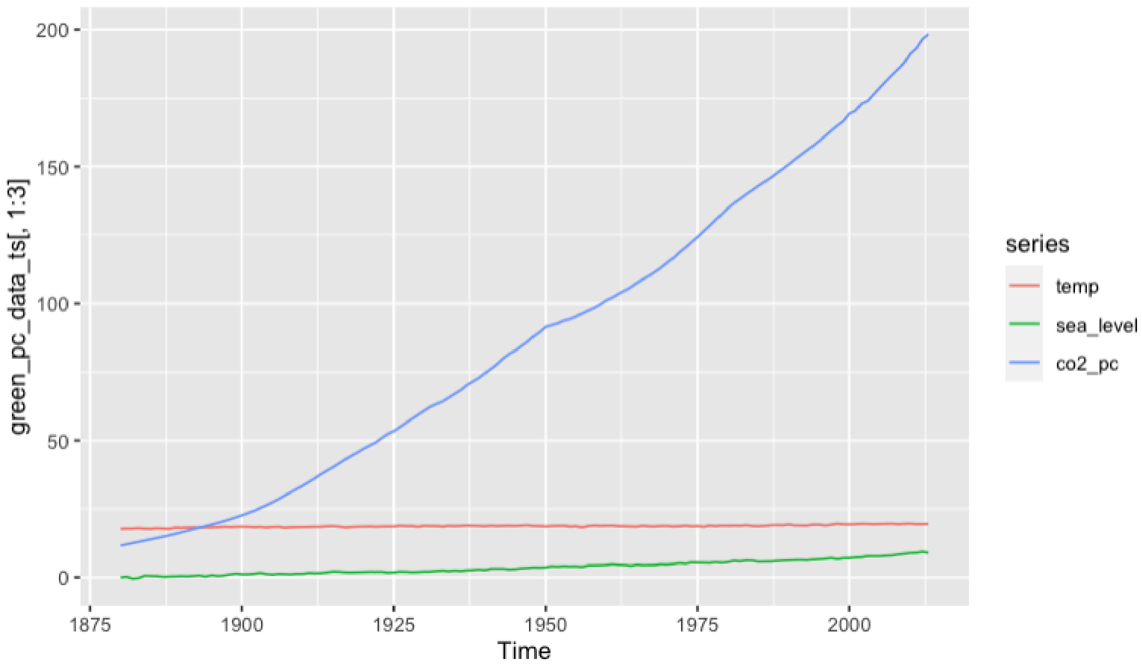


Image 4: Plot of all climate variables

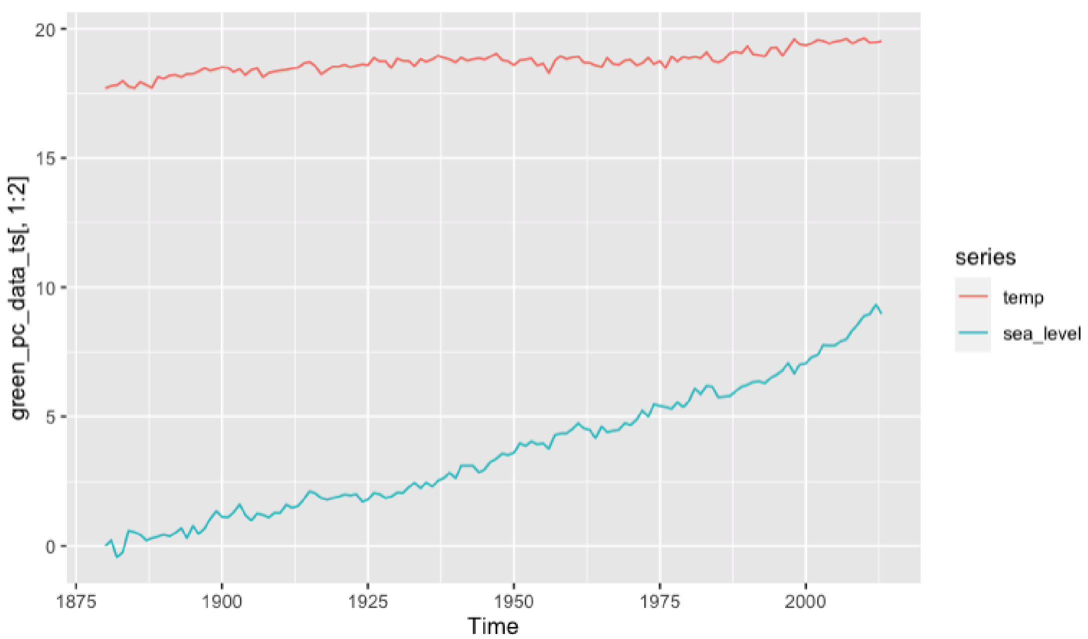


Image 5: Plot of sea level and temperature

However, zooming in, we see that changes in sea levels have also been increasing at a very drastic rate. From our literature review, carbon emissions are caused by human activity, such as industrialization. This suggests that, at least from this figure alone, human activities might be influencing climate change.

2.1.3 Visualisation of human activity data (Annual)

For human activity data (yellow data), we noticed that all of the human data depicts an upward increasing trend except for methane per capita, where it possesses a decreasing trend.

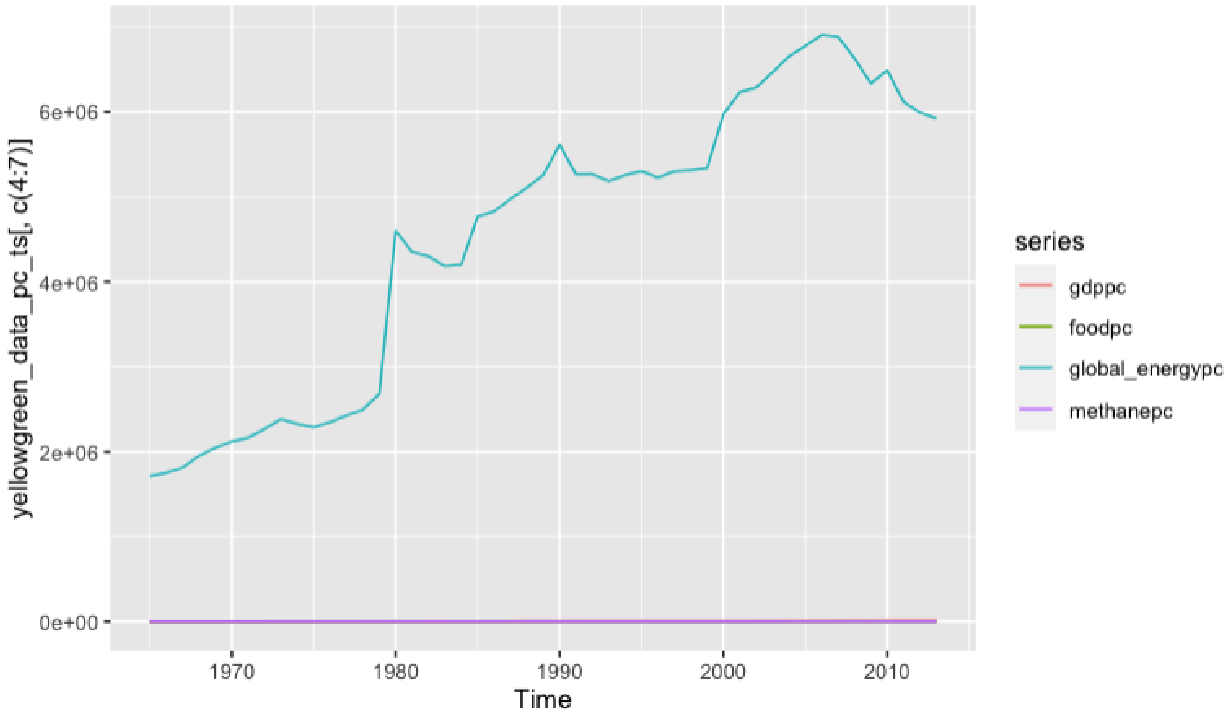


Image 6: Plot of human activity variables

We observed that the global energy consumption per capita is larger in magnitude than the other human activity data, and this is attributed to the units of energy consumption. Therefore, to understand plots of the remaining variables we remove some of the human activity variables to have a look at the relationships.

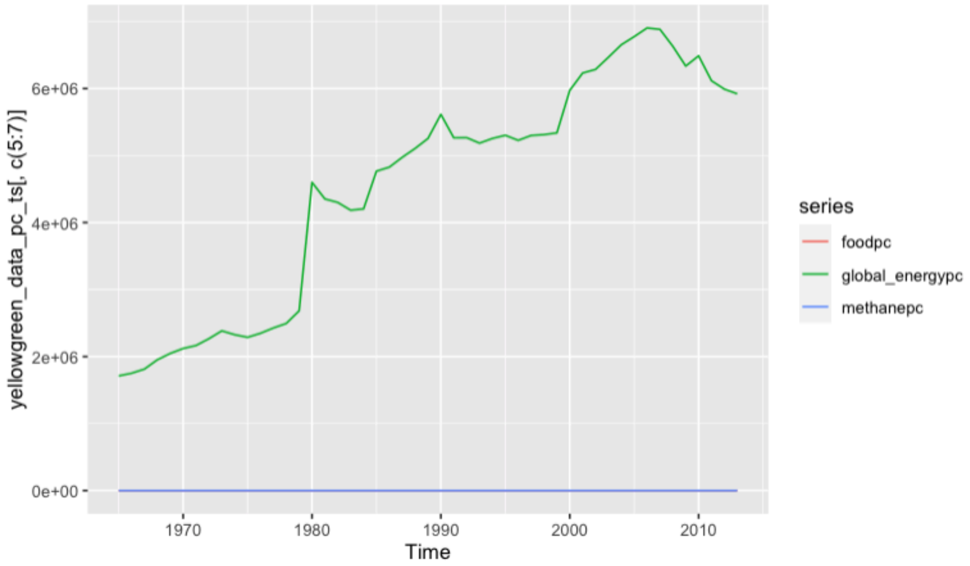


Image 7: Plot of agriculture, energy consumption, and methane in per capita

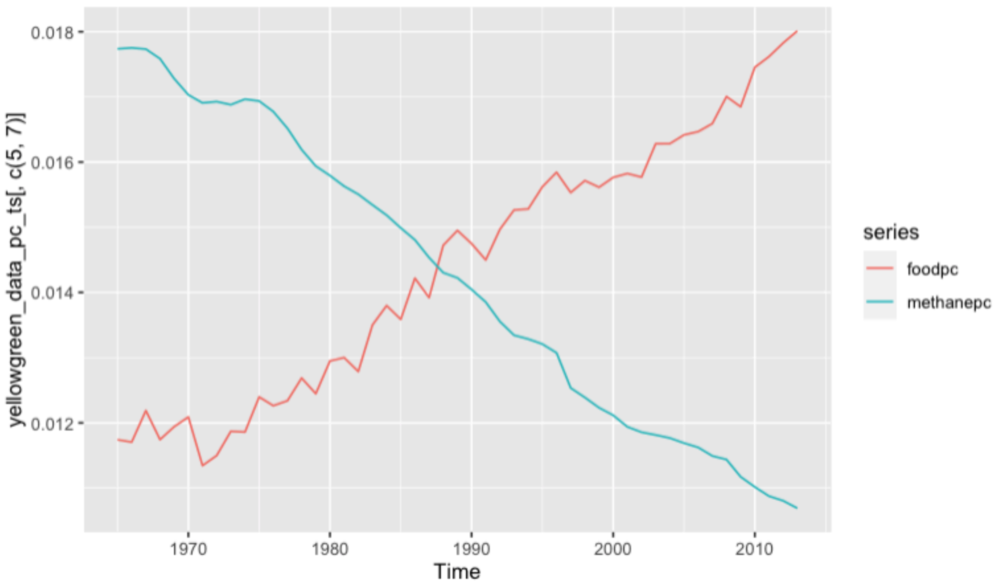


Image 8: Plot of methane per capita and agriculture per capita

From the plots above, agriculture per capita, GDP per capita, and energy consumption per capita increase over time, but methane per capita decreases over time. The increasing agriculture per capita and energy consumption per capita could be attributed to increasing globalisation and economic activity (Tabash et al., 2024). The decreasing levels of methane per capita is due to a reduction in emissions from landfills, coal mines, and natural gas systems which offsets the increasing emission of emissions from livestock (Environmental Protection Agency, 2023).

3. Model Building

Firstly, for the model-building workflow, our approach is as follows:

1. We would train our models on 1880 to 1952 data as a training set and forecast from 1953 to 1970 as our test set.
2. We would then choose and select an optimal model with the best accuracy on the test set.
3. We would then use this model to train on the entire pre-1970s data and forecast into the future.
4. We then compare forecasted model results with the actual temperature observed in our world. This would then be the comparison between 1970s climate projection vs actual climate projection. Should there be significant deviations from our model's predictions, then we conclude that increased human activities post-1970s did cause temperatures to increase.

One key consideration for our model-building process is the concern of overfitting. It is crucial to keep in mind the number of variables that our models will contain. This is to prevent overfitting, especially as yearly data can lead to a very limited number of records. This problem is further exacerbated by the fact that VAR and VECM models scale in the number of coefficients very quickly as their order increases (see VAR equation below), which further necessitates us to exercise careful model selection. As a general rule of thumb, a machine-learning model should have at least 10 observations per coefficient.

$$Y_{1,t} = A_1 + \beta_{11,1}Y_{1,t-1} + \beta_{12,1}Y_{2,t-1} + u_{1,t}$$

$$Y_{2,t} = A_2 + \beta_{21,1}Y_{1,t-1} + \beta_{22,1}Y_{2,t-1} + u_{2,t}$$

Each equation estimates 1 constant, plus K coefficients for each lag. Hence, the total across K equations is $K(pK+1) = K + pK^2$; where p = number of lags and K = number of variables.

Next, as mentioned earlier, the univariate model will be quarterly while the multivariate model will be annual.

3.1 Univariate models

For univariate modelling, we will be working with quarterly temperature data from the 1880s to 2024.

3.1.2 Additive Decomposition: Modelling trend cycle and seasonal components

We create a time series object ranging from 1880 to 1970s, to model pre-1952 and forecast post-1953. Firstly, we perform a time series decomposition to help us understand the trend, seasonality, and remainder components. This is done using the decompose() function which performs classical additive decomposition.

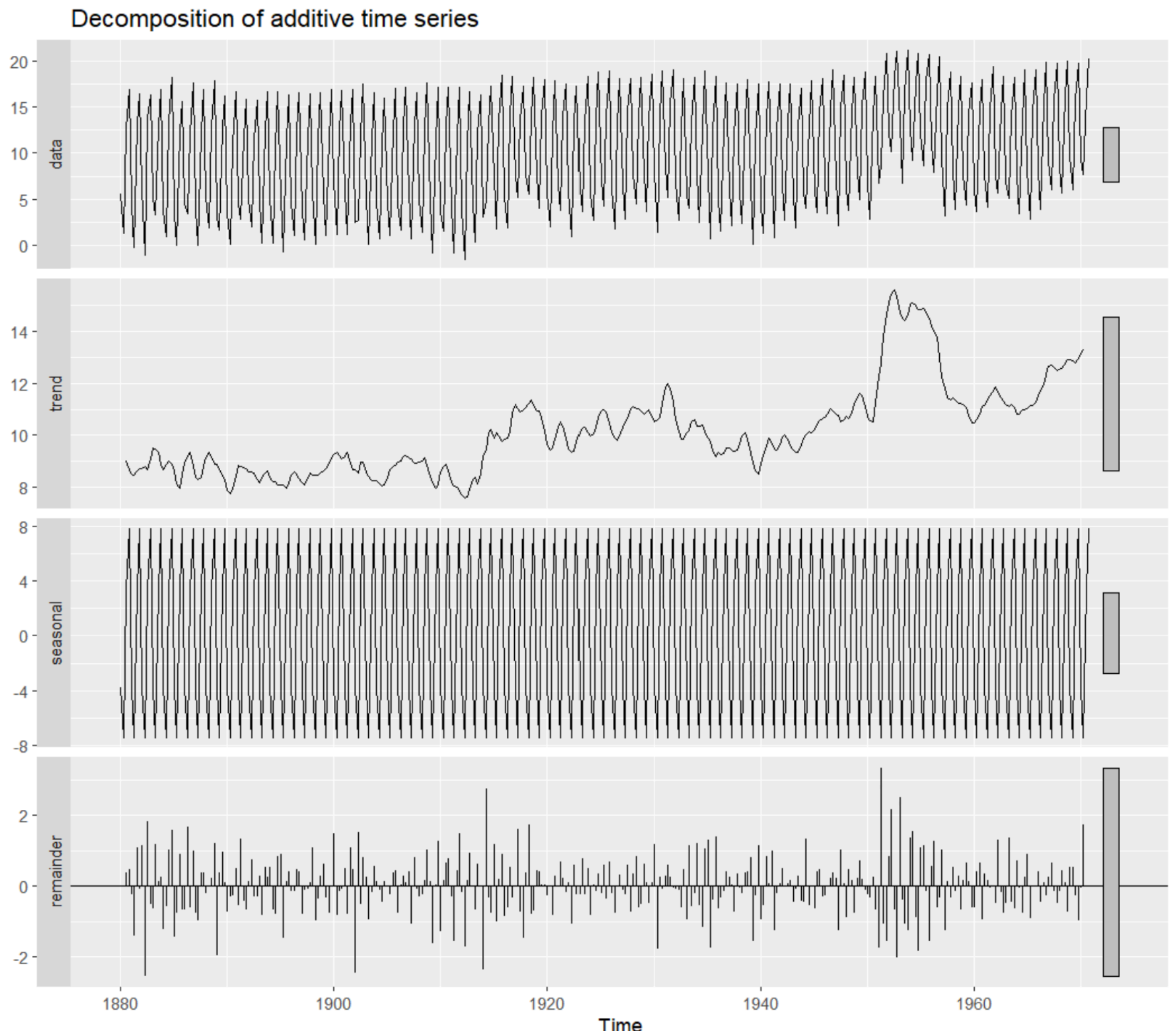


Image 9: Classical decomposition

As evident, it is clear that there is an upward trend, especially a sharp increase towards the end of the time series from about the 1950s onwards. There is also clear seasonality in our data. Hence, we conclude that the data is not stationary.

Next, we conduct a KPSS unit root test, to statistically prove that the time series is not stationary.

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: mu with 5 lags.

Value of test-statistic is: 2.3534

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

Image 10: Test for stationarity

Since the test statistic of 2.3534 lies outside the 5% critical value of 0.463, we reject the null hypothesis and conclude that the series is not stationary. This leads us to believe that data transformation is needed to make the series stationary before modelling, such as differencing, box-cox transformation, or logistic transformation.

Running nsdiffs(), we see that just one seasonal differencing is required. This is the result of a seasonal differencing:

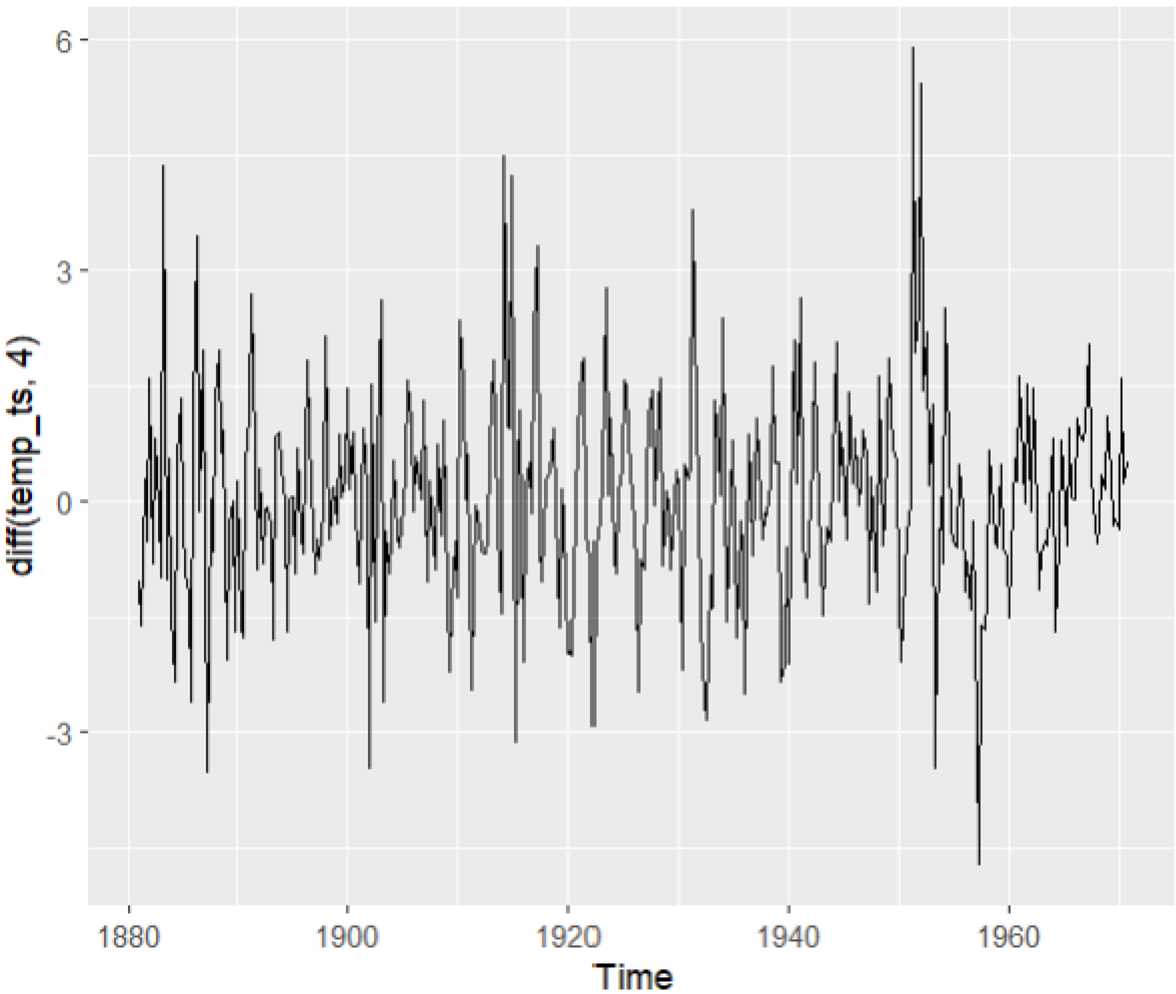


Image 11: Series autoplot after 1 seasonal difference

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: mu with 5 lags.

Value of test-statistic is: 0.0483

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

Image 12: Test for stationarity

The autoplot shows that the series is relatively stationary, and this is supported by the KPSS test which shows that the test statistic (0.0483) now lies within the 5% critical value (0.463). Thus, the series is stationary.

Next, we experiment with BoxCox. We had to shift the plot up by about 4.8 units as the stationary time series object had a minimum value of -4.6, and we added 4.8 as BoxCox only works on positive data. Based on the visual inspection, where the red line is the original stationary time series object and the black line is its BoxCox transformed version, we see that there is not much significant difference. Therefore, based on visual inspection only, we conclude that BoxCox is not necessary. Furthermore, the original series is already stationary and a large number of transformations would make it harder to interpret our forecasts.

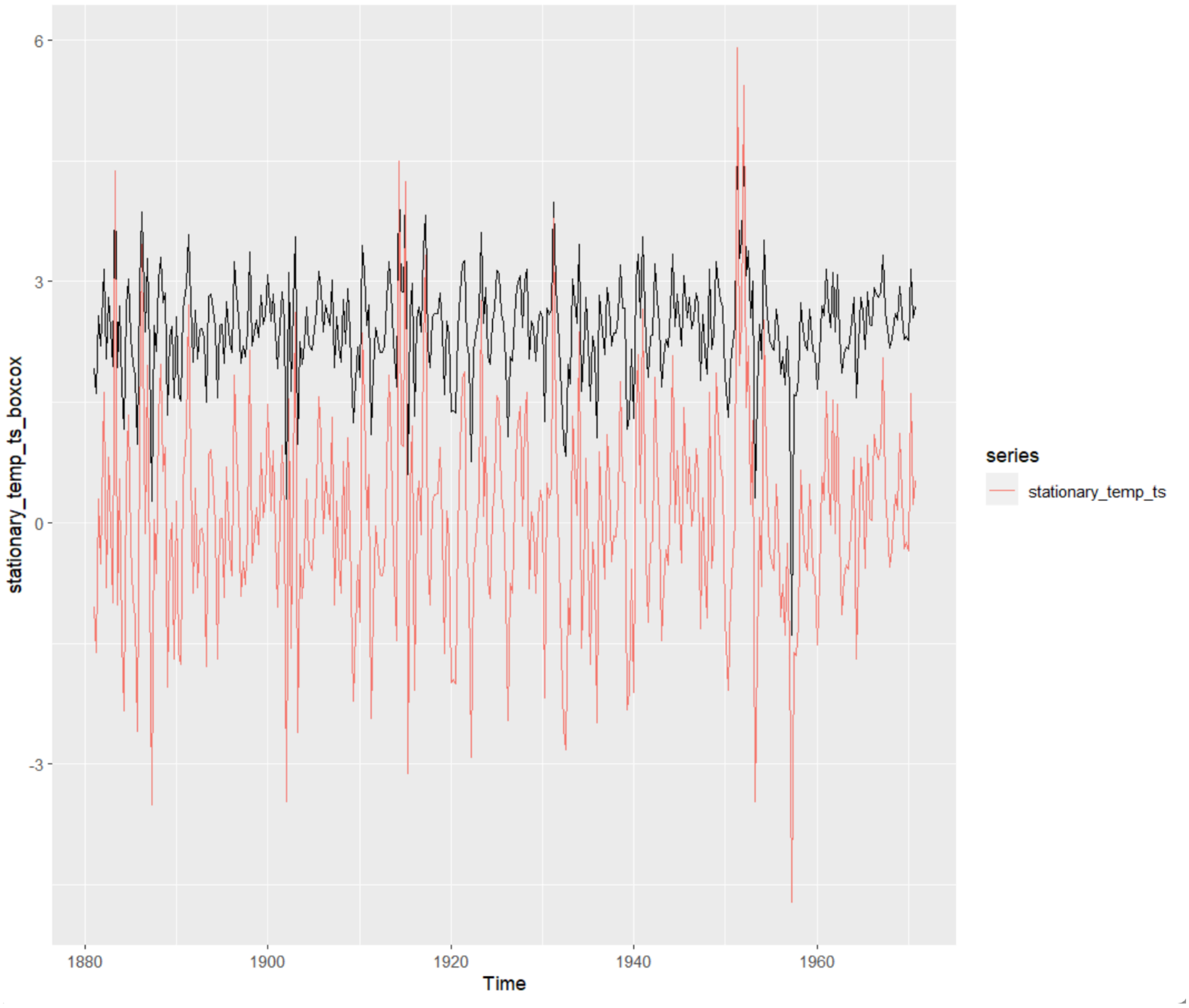


Image 13: Boxcox vs first seasonal diff (no significant difference)

Thereafter, we will split the stationary series into the training and test set, and model different benchmark methods on the trendcycle and seasonal components. We also tried stl decomposition, on top of the standard classical decomposition. Stl decomposition might be more effective than classical decomposition. This is because classical decomposition assumes that seasonal patterns are constant over time, and are not robust to outliers. Whereas stlf uses a rolling window consisting of the recent seasonal data, hence potentially providing more accurate seasonality. This is because it adapts to changes in seasonal patterns over time, making it more robust to shifts or irregularities in the data.

The following table shows our models, along with the decomposition, seasonal and trend-cycle methods:

Model	Decomposition	Seasonal Component	Trendcycle component
additiveforecast1	classical	snaive	naive
additiveforecast2	classical	snaive	mean
additiveforecast3	classical	snaive	drift
additiveforecast4	stl	snaive	naive
additiveforecast5	stl	snaive	drift

Table 2: The models we built and their loadouts

The following is the layered plot of the different forecasts.

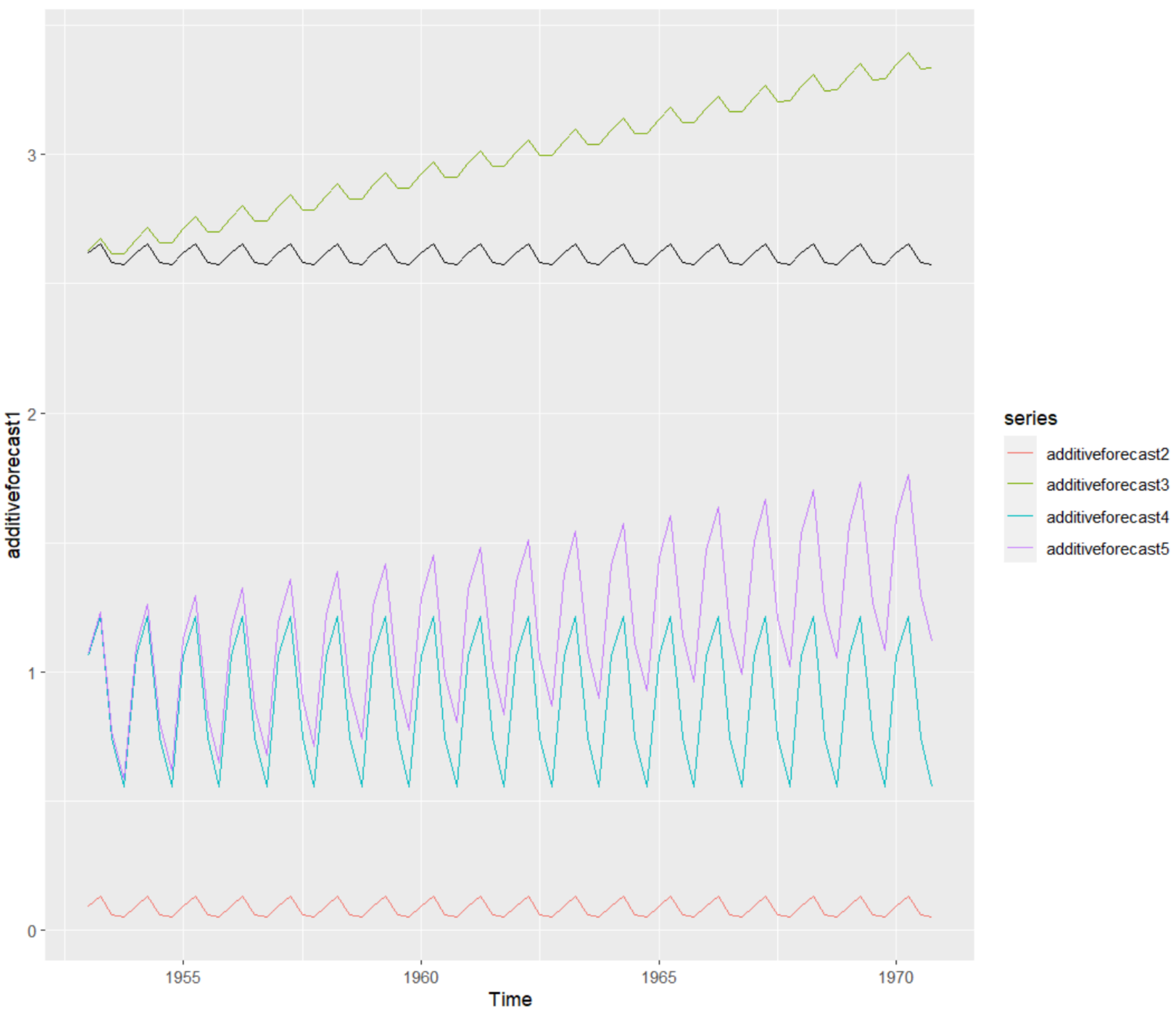


Image 14: Forecasts from classical modelling

Below is the same plot with the actual series. It appears that the forecasting methods are not very great, especially due to the high amplitude/variance of the test series, including a sudden sharp decrease in values during the 1957 period.

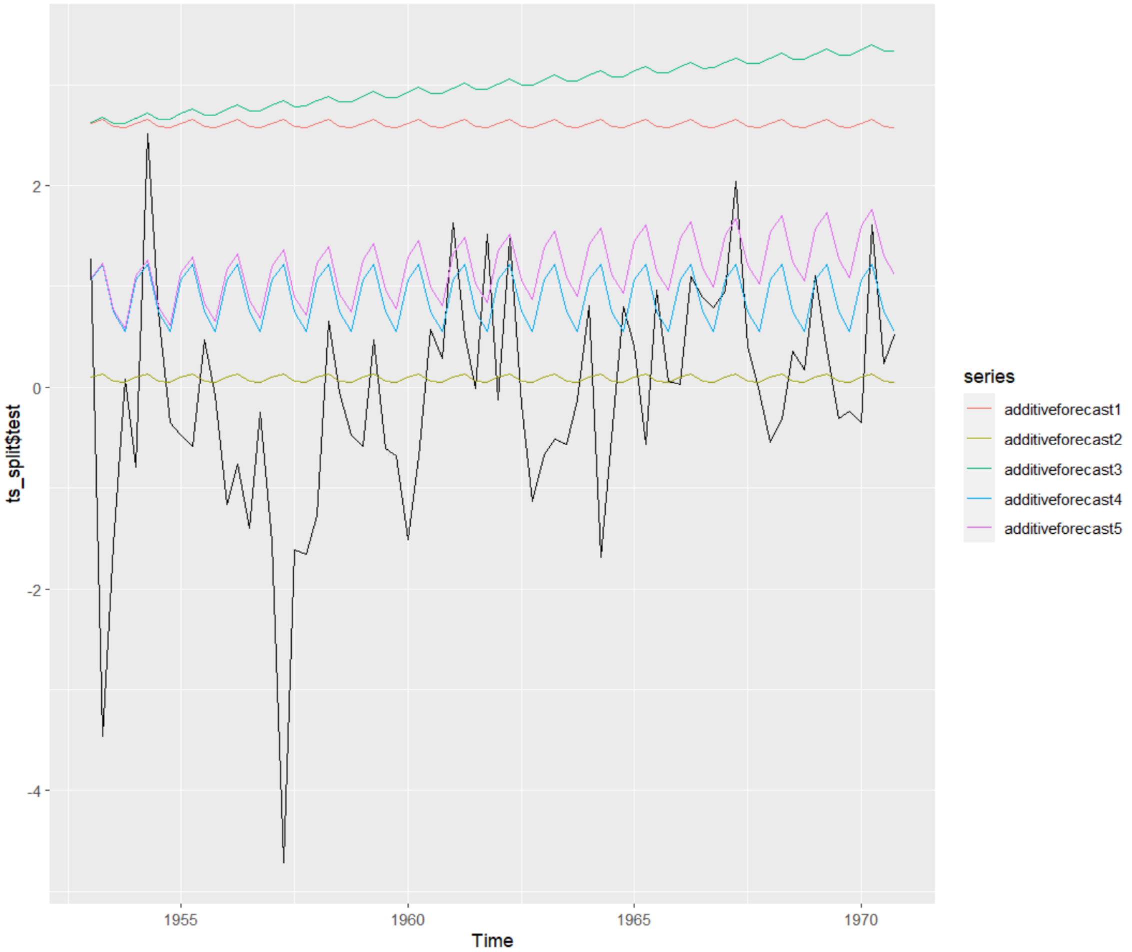


Image 15: Comparison of forecasts with actual series

Lastly, we gather out-of-sample accuracy. This is to help us decide which model is optimal.

Model	RMSE	MAPE
additiveforecast1	2.952171	1069.596
additiveforecast2	1.15434	110.8086
additiveforecast3	3.292869	1239.456
additiveforecast4	1.549341	341.4477
additiveforecast5	1.719449	461.2836

Table 3: Model performances

Our best model is model 2, based on the numbers alone. However, we have to conduct a LJung Box test on the full data to investigate the model's effectiveness.

Unfortunately, all the models fail the Ljung Box test. Thus, we conclude that in this case, the classical methods are not very effective, as residuals still contain time series information. This leads us to consider other forms of modelling such as SARIMA.

3.1.3 SARIMA

Out of Sample Accuracy: RMSE: 5.234782, MAPE: 30.51816

Using the same training data from 1880 to 1952 as above, we run auto.arima() to build an optimised SARIMA model for us. This gave us a SARIMA(2,0,1)(0,1,1)[4] model. This means that the model consists of 1 order of seasonal differencing, 1 order of seasonal moving average, 2 orders of autoregression, and 1 order of moving average.

Thereafter, we must ensure that the model passes the Ljung-Box test, such that we know that there is no time series information (autocorrelation) present in the residuals. If there was time series information still captured in our residuals, the model is invalid and we would need to derive another SARIMA model.

```
        Ljung-Box test

data:  Residuals from ARIMA(2,0,1)(0,1,1)[4] with drift
Q* = 2.7615, df = 4, p-value = 0.5985

Model df: 4.   Total lags used: 8
```

Image 16: Ljung Box test on SARIMA(2,0,1)(0,1,1)

Fortunately, our model passes the Ljung box test with a p-value of 0.5985 > 0.05, hence we fail to reject the null hypothesis that there is no time series autocorrelation present.

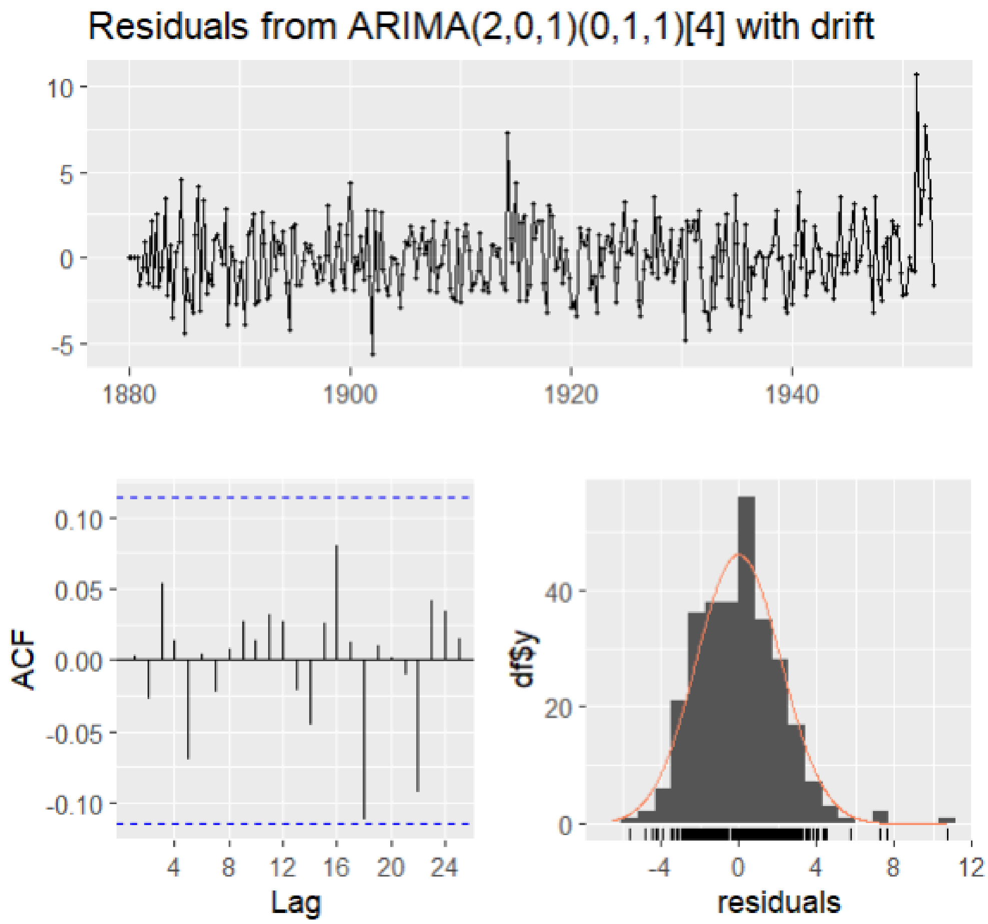


Image 17: Plot of the Ljung box test

This is consistent with the graphical output of the Ljung Box test, which shows the residuals having constant mean and variance on the topmost graph. Additionally, ACF shows no significant autocorrelation in the lags.

Our final step is to attain out-of-sample accuracy scores. Here, we look at our test set RMSE and MAPE. This gives us the performance of the model using 1952 to 1970s as the test data. We derived a test RMSE of 2.788031 and a test MAPE of 33.85267.

Overall, we then train this optimal SARIMA(2,0,1)(0,1,1) model on the pre-1970s data and forecast it post-1970s. Our final test scores are an RMSE of 5.234782 and a MAPE of 30.51816.

3.1.4 STLF ARIMA

Another method of applying ARIMA can be through the stlf function in R. This is a method that combines ARIMA, BoxCox, and time series decomposition. It performs the following steps:

1. Transforms data using BoxCox with “auto” lambda
2. Performs STL decomposition using MSTL to automatically calibrate seasonal windows
3. Automatically builds ARIMA models on seasonally adjusted components (trend-cycle)
4. Uses snaive on the seasonal component.
5. Recombines (3) and (4) and reverse BoxCox to forecast.

While we attained accuracy scores for the automated ARIMA built, STL + ARIMA(0,1,1), the automated stlf Arima model fails the Ljung box test, which will not be useful for our modelling process. This is because residuals still contain time series information.

```
        Ljung-Box test

data:  Residuals from STL +  ARIMA(0,1,1) with drift
Q* = 23.796, df = 7, p-value = 0.001237

Model df: 1.   Total lags used: 8
```

Image 18: Ljung Box test on stlf arima model

The p-value of 0.001237 means that we reject the null hypothesis and that there is time series autocorrelation present in our model.

Hence, through our univariate modelling process, we conclude that SARIMA(2,0,1)(0,1,1) is the most appropriate model and the model we have chosen for forecasting. This is the only model that passes the Ljung box test, while the others all fail.

3.2 Climate Data Models (Green)

3.2.1 Sampling

Since the data for the climate activity model ranges from 1880 to 2013 of yearly data, it gives us 134 years (observations). We will be splitting our data set into 3 different sections as mentioned previously (1880-1952, 1953-1970, post 1970s). Hence, we will have 74 observations as the training set, 16 observations as a test set, and the remaining to allow us to assess if there is truly climate change. We note that the training and test set will be split into an 80-20 ratio.

3.2.2 VAR

3.2.2.1 Temperature, sea levels, Carbon dioxide

Out of sample accuracy: RSME = 0.3549815, MAPE = 1.651139

When constructing our VAR, the first step is to apply VARselect() on our time series object. This is akin to auto.arima(), which will help us determine the orders needed for our optimal VAR model. The following is an example of the output:

```
$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
     2      2      1      2
```

Image 19: SC selection for VAR model

In this case, we look at the SC(n) value, as this represents the order that provides us with the optimal BIC score as SIC, SC, and BIC are the same measurements. The reason why we are interested in BIC over AIC is that VAR coefficients scale very quickly as order increases, as such we resort to using a larger penalty factor such as BIC when selecting our VAR models, to prevent overfitting.

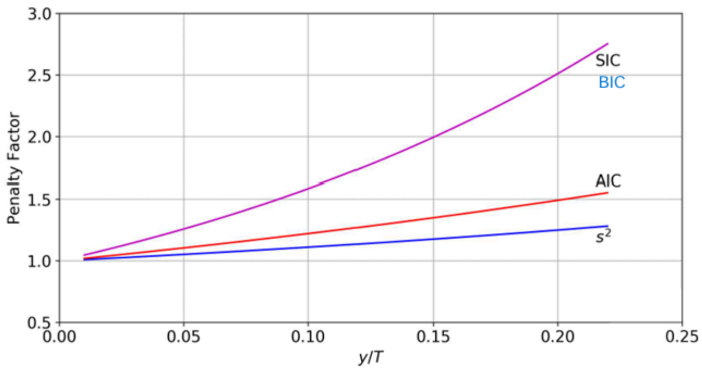
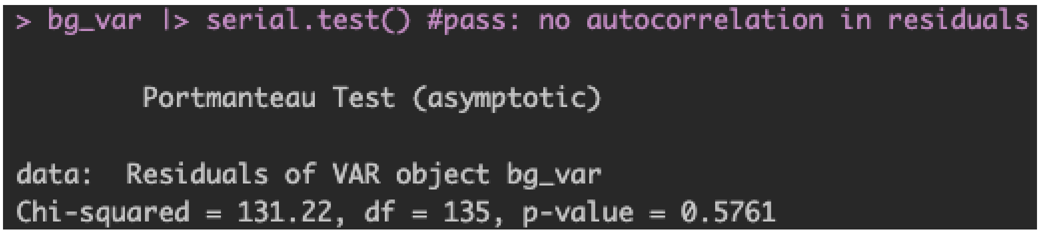


Image 20: Graph comparing different penalty factors

Then, for any VAR model that we have selected to build, we will apply the Portmanteau test to test for autocorrelation in residuals.



```
> bg_var |> serial.test() #pass: no autocorrelation in residuals

	Portmanteau Test (asymptotic)

data:  Residuals of VAR object bg_var
Chi-squared = 131.22, df = 135, p-value = 0.5761
```

Image 21: Portmanteau test of autocorrelation in residuals

From this output, the p-value of 0.5761 means that we fail to reject the null hypothesis and that there is no significant time series information, to conclude that our VAR model passes tests.

$$temp_{1,t} = A_2 + \beta_{11} temp_{t-1} + \beta_{12} sea\_levels_{t-1} + \beta_{13} co2\_pc_{t-1} + u_{1,t}$$

$$sea\_levels_{1,t} = A_2 + \beta_{21} temp_{t-1} + \beta_{22} sea\_levels_{t-1} + \beta_{23} co2\_pc_{t-1} + u_{2,t}$$

$$co2\_pc_{1,t} = A_3 + \beta_{31} temp_{t-1} + \beta_{32} sea\_levels_{t-1} + \beta_{33} co2\_pc_{t-1} + u_{3,t}$$

Since this model is a VAR(1) K=3, p=3, we have (1+1*3)*3 = 12 parameters, there is severe overfitting for the model. Therefore, we might have to consider this when deciding on the best green model.

3.2.3 VAR-X

3.2.3.1 Temperature and Carbon Dioxide

Out of sample accuracy: RSME = 0.2314613, MAPE = 1.054244

```
> # temp and co2pc
> # parameters = (1+1 + 2*2)*2 = 12 (overfitting)
> green_data_pc_split$train[,c(1,3)] |> VARselect(type = "const") #lag = 2
$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
    10      2      2     10

$criteria
                  1            2            3            4            5            6
AIC(n) -6.564711379 -6.903043130 -6.835436751 -6.782326418 -6.674931745 -6.614120912
HQ(n)  -6.484977651 -6.770153583 -6.649391386 -6.543125234 -6.382574743 -6.268608091
SC(n)  -6.362316090 -6.565717648 -6.363181077 -6.175140551 -5.932815685 -5.737074659
FPE(n)  0.001409424  0.001005365  0.001076888  0.001137896  0.001270888  0.001356752
                  7            8            9           10
AIC(n) -6.773508545 -6.662790640 -6.807857824 -6.972656187
HQ(n)  -6.374839906 -6.210966182 -6.302877548 -6.414520092
SC(n)  -5.761532100 -5.515884002 -5.526020994 -5.555889164
FPE(n)  0.001164162  0.001311364  0.001146546  0.000985606

> bg_varx_13 <- VAR(green_data_pc_split$train[,c(1,3)] , p=2, exogen =green_data_pc_split$train[,2], type = 'const')
#lag = 2, parameters = (1+1 + 2*2)*2 = 12 (overfitting)
Warning: No column names supplied in exogen, using: exo1 , instead.
> bg_varx_13  |> serial.test() #pass test: no autocorrelation in residuals

        Portmanteau Test (asymptotic)

data:  Residuals of VAR object bg_varx_13
Chi-squared = 45.954, df = 56, p-value = 0.8286

> fc_bg_varx_13  <- bg_varx_13 |> predict(dumvar = matrix(green_data_pc_split$test[,2]), n.ahead = out_of_sample_year
s)
> accuracy(green_data_pc_split$test[,1], fc_bg_varx_13$fcst$temp[,1]) #RSME = 0.3657443, MAPE =1.67904
                ME      RMSE       MAE     MPE    MAPE
Test set 0.3196359 0.3657443 0.3196359 1.67904 1.67904
```

Image 22: VAR-X for temperature and CO2pc, with sea levels as exogenous

As we take the same steps for building VAR-X for temperature and CO2pc, with sea levels as exogenous, we realise that the model is overfitted.

3.2.3.2 Temperature and Sea levels

Out of sample accuracy: RSME = 0.3922482, MAPE = 1.845636

```
> # temp and sea level
> # parameters = (1+1 + 2*1)*2 = 8 (overfitting)
> green_data_pc_split$train[,1:2] |> VARselect(type = "const") #lag=1
$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
     2      1      1      2

$criteria
                   1             2            3             4             5             6
AIC(n) -7.3070525254 -7.3378476857 -7.278002727 -7.2603727973 -7.1782980019 -7.0656933559
HQ(n)  -7.2273187975 -7.2049581393 -7.091957362 -7.0211716137 -6.8859409997 -6.7201805351
SC(n)  -7.1046572363 -7.0005222039 -6.805747052 -6.6531869301 -6.4361819420 -6.1886471033
FPE(n)  0.0006708835  0.0006508642  0.000691778  0.0007054884  0.0007682421  0.0008637443
                   7            8            9           10
AIC(n) -6.9580380742 -6.87326169 -6.764504066 -6.787677009
HQ(n)  -6.5593694348 -6.42143723 -6.259523789 -6.229540914
SC(n)  -5.9460616288 -5.72635505 -5.482667235 -5.370909986
FPE(n)  0.0009679956  0.00106247  0.001197346  0.001185875

> bg_varx_12 <- VAR(green_data_pc_split$train[,1:2] , p=1, exogen = green_data_pc_split$train[,3], type = 'const') #l
ag = 1, (1-1+2*2)*2 = 8
Warning: No column names supplied in exogen, using: exo1 , instead.
> bg_varx_12 |> serial.test() #pass test

	Portmanteau Test (asymptotic)

data:  Residuals of VAR object bg_varx_12
Chi-squared = 36.191, df = 60, p-value = 0.9936

> fc_bg_varx_12 <- bg_varx_12 |> predict(dumvar = matrix(green_data_pc_split$test[,3]), n.ahead = out_of_sample_year
s)
> accuracy(green_data_pc_split$test[,1],fc_bg_varx_12$fcst$temp[,1]) #RSME = 0.3922482, MAPE = 1.845636
                ME      RMSE       MAE      MPE     MAPE
Test set 0.3518478 0.3922482 0.3518478 1.845636 1.845636
```

Image 23: VAR-X for temperature and sea levels, with CO2pc as exogenous

Similar to the above steps, we build a VAR-X for temperature and sea levels, with CO2pc as exogenous, and we realise that the model is overfitted.

3.2.4 VECM-X

3.2.4.1 Temperature and Sea Levels

Out of sample accuracy: RMSE = 0.4518534, MAPE = 2.104086

VECM or VECM-X is used when time series are cointegrated; they "move together" over time. This means that the time series themselves may not be stationary, but a linear combination of them is stationary (the difference between them is constant). The time series must require the same number of differencing for VECM to work. The drawbacks of our models (VAR, ARIMA) so far are that they require time series to be stationary before these methods can be applied, which ends up removing levels of information, which comes at a cost. There may be information embedded in the levels. Therefore, VECM can model data with unit roots directly, thus addressing the previous problem, to obtain generally superior results, if variables are cointegrated. Before this, we would need to conduct an Engle-Granger Cointegration test - consisting of checking for orders of differentiation on time series and the Augmented Dickey-Fuller (ADF) test on residuals. If there are variables that are cointegrated with each other, we can use the Johansen Test to find the number of cointegrated relationships. Following this, we can use the Ljung box test to check for no autocorrelation in the residual. The difference between VECM and VECM-X is that VECM-X includes exogenous variables in the modelling.

```
#bg model: VECM-X with exogeneous variable as temp
ndiffs(green_data_pc_split$train[,1]) # temp requires 1differencing
ndiffs(green_data_pc_split$train[,2]) # sea level requires 1 differencing
ndiffs(green_data_pc_split$train[,3]) #co2pc requires 2 differencing
```

```
[1] 1
[1] 1
[1] 2
```

Image 24: number of differencing required for climate data

Since the number of differencing required for temperature and sea level is the same, but carbon dioxide emissions per capita are different, we have to use VECM-X since carbon dioxide cannot have a cointegration relationship with sea level and temperature.

```
######################
# Johansen-Procedure #
######################

Test type: maximal eigenvalue statistic (lambda max) , with linear trend

Eigenvalues (lambda):
[1] 0.118959137 0.003070402

Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 1 |  0.25  6.50  8.18 11.65
r = 0  | 10.39 12.91 14.90 19.19

Eigenvectors, normalised to first column:
(These are the cointegration relations)

                temp.l2 sea_level.l2
temp.l2       1.0000000    1.0000000
sea_level.l2 -0.1498368   -0.5730806

Weights W:
(This is the loading matrix)

               temp.l2 sea_level.l2
temp.d      -0.2155995  0.009483675
sea_level.d  0.1618327  0.022568322
```

Image 25: Johansen Procedure to identify the number of cointegrating relationships

From the image above, we notice that the critical value for cointegrating relationship (r) <= 1 is 0.25 < 6.50 at a 10% significance level. Therefore, we can conclusively say that there is 1 cointegrating relationship between temperature and sea levels.

```
###############################################################
# Augmented Dickey-Fuller Test Unit Root / Cointegration Test #
###############################################################

The value of the test statistic is: -7.018


###############################################################
# Augmented Dickey-Fuller Test Unit Root / Cointegration Test #
###############################################################

The value of the test statistic is: -5.9453
```

Image 26: ADF Unit root test for stationarity on residuals

The absolute value of test statistics for the ADF unit root test for stationarity is > |1.96|Therefore, we can conclude that the residuals are stationary.

```
        Ljung-Box test

data:  Residuals
Q* = 10.448, df = 10, p-value = 0.4021

Model df: 0.   Total lags used: 10


        Ljung-Box test

data:  Residuals
Q* = 7.8217, df = 10, p-value = 0.6462

Model df: 0.   Total lags used: 10
```

Image 27: Ljung-box test to test for no autocorrelation in residuals.

Likewise, Ljung box test for residuals also confirms that there is no autocorrelation in the residuals as p-value > 0.1 at a 5% level of significance.

Since this model is a VECM-X, K=2, p=2, r=1, we have ((1+1+1)+1*2)*2 = 14 parameters, there is severe overfitting for the model. Therefore, we might have to consider this when deciding on the best green model.

3.2.5 ARIMA-X

3.2.5.1 ARIMA-X on Temperature with Sea Levels and Carbon Dioxide Emissions Per Capita as Exogenous Variables

Out of sample accuracy: RSME = 0.3567035, MAPE = 1.6316

3.2.6 Statistical tests for significance for all multivariate models built

3.2.6.1 VAR Granger causality test results

Model specification	**Hypothesis Tested**	**Granger Causality (5% significance)**
VAR(1) for temperature, co2pc, sea level	H0: temp do not Granger-cause sea_level co2_pc H1: temp does Granger-cause sea_level co2_pc	p-value = 0.0001952 < 0.05 - Reject H0 - Temperature granger causes sea level and CO2pc
	H0: sea_level do not	p-value = 0.7659 > 0.05

	Granger-cause temp co2_pc H1: sea_level Granger-cause temp co2_pc	- Cannot Reject H0 - Sea level does not granger cause temperature and CO2pc
	H0: co2_pc do not Granger-cause temp sea_level H1: co2_pc Granger-cause temp sea_level	p-value = 0.0007024 < 0.05 - Reject H0 - CO2pc granger causes temperature and sea level.
VAR-X(2) for temperature, co2pc, with sea level as exogenous	H0: temp do not Granger-cause co2_pc H1: temp does Granger-cause co2_pc	p-value = 0.02636 < 0.05 - Reject H0 - Temperature granger causes CO2pc.
	H0: co2_pc do not Granger-cause temp H1: co2_pc Granger-cause temp	p-value = 0.4399 < 0.05 - Reject H0 - CO2pc granger causes temperature.
VAR-X(1) for temperature, sea level with co2pc as exogenous	H0: temp do not Granger-cause sea level H1: temp Granger-cause sea level	p-value = 0.09229 > 0.05 - Cannot reject H0 - Temperature does not granger cause sea level
	H0: sea level does not Granger-cause temp H1: sea level Granger-cause temp	p-value = 0.7554 > 0.05 - Cannot reject H0 - Sea level does not granger cause temperature

Table 4: VAR granger causality test results

Based on our VAR statistical tests, we see that the relationships between the 3 variables are ambiguous, and not very clear-cut. This is because some variables granger cause one another, but in no VAR specification does all variables granger cause one another.

3.2.6.2 ARIMA-X t-test results

Model Specification	Hypothesis Tested	T-Statistics (5% significance)
ARIMA-X on Temperature with Sea Levels and CO2pc as Exogenous	H0: Sea level does not explain changes in temperature H1: Sea level explains changes in temperature	Critical Value = 0.6367288 < 1.96 - Cannot reject H0 - Sea level does not explain the changes in temperature

	H0: CO2pc does not explain changes in temperature H1: CO2pc explains changes in temperature	Critical Value = 2.288505 > 1.96 - Reject H0 - CO2pc explains the changes in temperature

Table 5: ARIMA-X t-test results

The ARIMA-X test result is backed up by a study that showed that CO2 emissions and global temperature are related by using a long memory model, ARFIMA, assuming that CO2 emissions are weakly exogenous, the result produced a significant positive effect in the temperatures. Without the assumption, the result indicates that the two variables differ in their orders of integration, which might create challenges for analyzing their relationship because their long-term behaviors may not be related in a meaningful way, which can invalidate the analysis of cointegration between them (Gil-Alana & Monge, 2020).

3.3 Human Activity Models (Yellow)

In this section, we will only be considering ARIMA-X. From a climate perspective, as previously mentioned in our literature review, human activity causes temperature in a one-directional relationship. Since the relationship is not bi-directional, we decided to reject VAR modelling. Based on the equation of VAR, VAR modelling requires variables to be simultaneously related to each other, such a variable can be forecasted based on the lagged values of another.

3.3.1 Sampling

Since the data for the climate activity model ranges from 1965 to 2013 consisting of yearly data, it gives us 49 years (observations). Since we will be splitting our data set into 3 different sections as mentioned previously, we will have 28 observations for the training set, 7 observations as the test set, and the remaining 14 years will allow us to assess if there is truly climate change.

3.3.2 VAR(X), VECM(X)

Since our training set is limited to 28 observations, we are unable to build a VAR model that is not severely overfitted due to the exponential increase in parameters at higher lags. Therefore, we have decided to drop VAR, VAR-X, VECM, and VECM-X even if there might be a bidirectional relationship between human activity variables and temperature.

3.3.3 ARIMA-X

3.3.3.1 ARIMA-X with Temperature and other human activity variables as the exogenous variable.

Out of sample accuracy: RSME = 0.245348, MAPE = 1.084634

However, it is noted that our ARIMA-X model is also limited because of the number of observations available in the training set. We also confirmed that the ARIMA-X(0,0,0) passes the Ljung box test, with a p-value < 0.05. Therefore, at a 5% level of significance, we cannot reject the null hypothesis and hence, there is no autocorrelation in the residuals.

3.3.4 Statistical Tests for human activity model

3.3.4.1 ARIMA-X t test results

Model Specification	Hypothesis Tested	T-Statistics (5% significance)
ARIMA-X on Temperature with Human Activity variables (GDPpc, methane pc, energy consumption pc, agriculture pc)	H0: GDPpc does not explain changes in temperature H1:GDPpc explains changes in temperature	Critical Value = 0.7639196 < 1.96 - Cannot reject H0 - GDPpc does not explain the changes in temperature
	H0: Agriculture pc does not explain changes in temperature H1: Agriculture pc explains	Critical Value = 0.266043 < 1.96 - Cannot Reject H0 - Agriculture pc does not explain the changes in

	changes in temperature	temperature
	H0: Energy pc does not explain changes in temperature H1: Energy pc explains changes in temperature	Critical Value = 0.0001420065 < 1.96 - Cannot reject H0 - Energy pc does not explain the changes in temperature
	H0: Methane pc does not explain changes in temperature H1: Methane pc explains changes in temperature	Critical Value = 0.6396764 < 1.96 - Cannot reject H0 - Methane pc does not explain the changes in temperature

Table 6: ARIMA-x t test results

4. Is there really Climate Change?

4.1 Temperature Forecast Model (Quarterly)

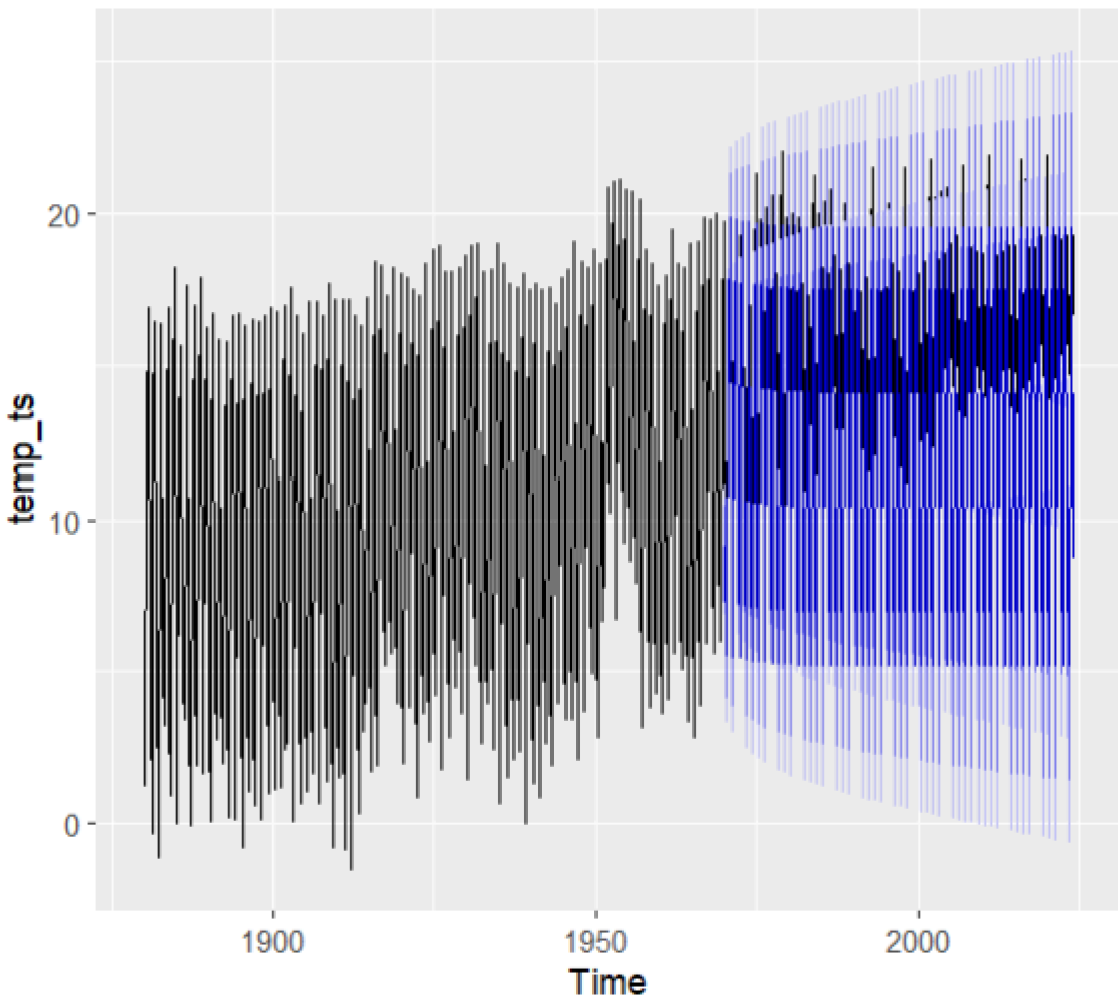


Image 28: Comparison plot of univariate SARIMA to actual temperature data post-1970s

Based on our SARIMA(2,0,1)(0,1,1)[4] model forecasts (in blue), we see that the actual temperature post-1970s lies within the 95% confidence interval of our forecast. From our forecasts, there is not much significant change in temperature pre-1970s and post-1970s, at least from a SARIMA forecast point of contention.

4.2 Climate Data Forecast Model (Annual)

4.2.1 ARIMA-X (1,0,1)

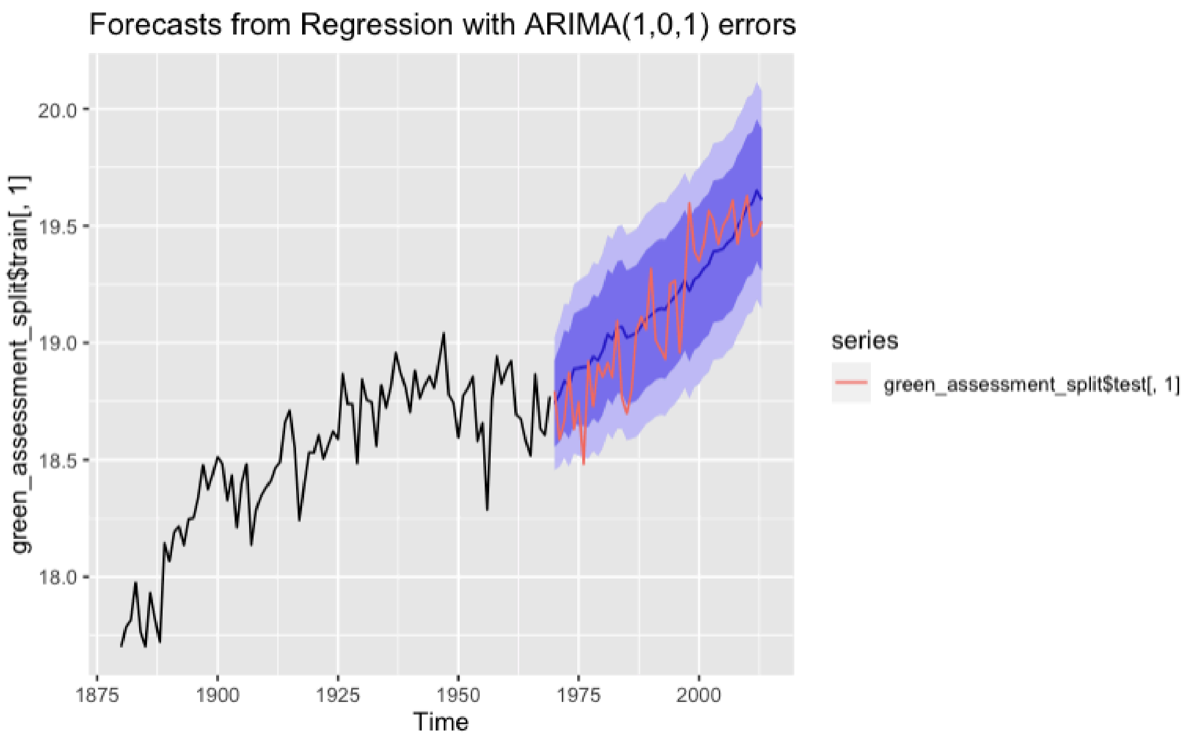


Image 29: Comparison plot of climate model forecast to actual temperature data post-1970

Based on our ARIMA-X (1,0,1) model forecasts (blue line), we see that the actual temperature post-1970s lies within the 95% confidence interval of our forecast (darker blue shaded area). From our forecasts, we conclude that there is not much significant change in temperature pre-1970s and post-1970s, at least from an ARIMA-X forecast point of contention.

4.3 Human Activity Data Forecast Model (Annual)

4.3.1 ARIMA-X (0,0,0)

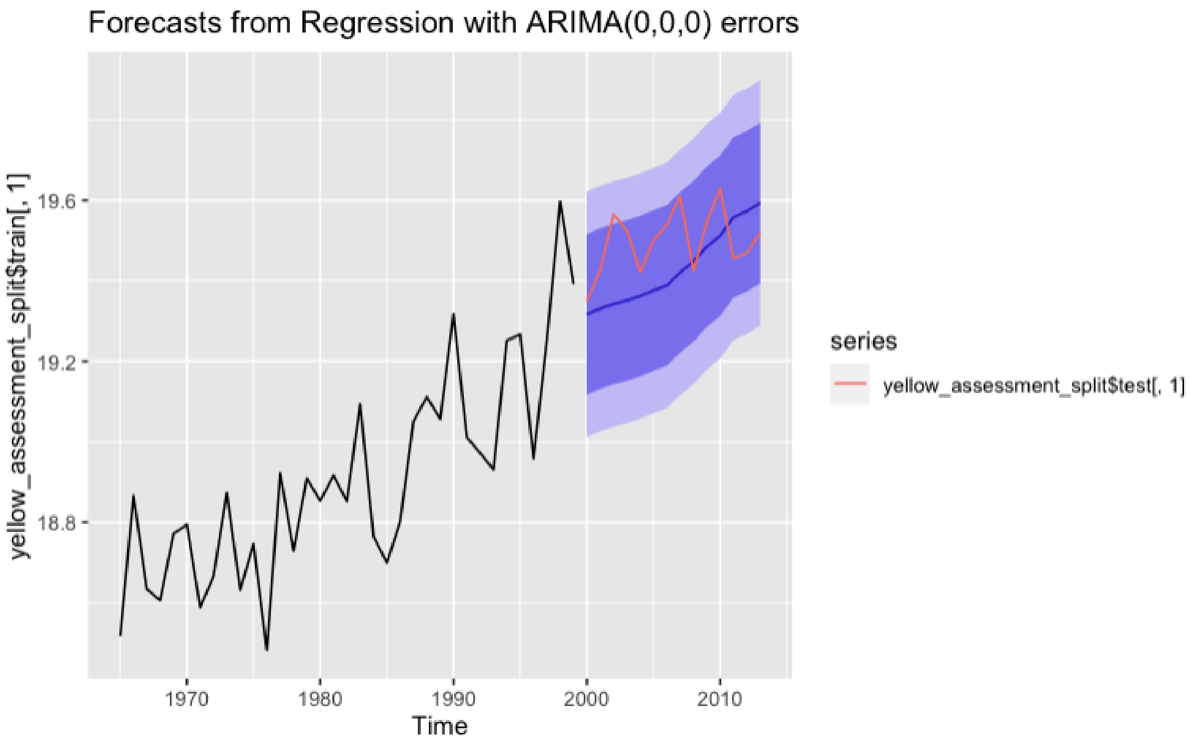


Image 30: Comparison plot of human activity model forecast to actual temperature data post-2000s

Based on our ARIMA-X (0,0,0) model forecasts (blue line), we see that the actual temperature post-2000s lies within the 95% confidence interval of our forecast (darker blue shaded area). From our forecasts, we conclude that there is not much significant change in temperature pre-2000s and post-2000s, at least from an ARIMA-X forecast point of contention.

5. Robustness

5.1 Summary of relationships between variables

Variables	Our Statistical tests	Empirical Studies
Temperature and Co2pc	Bidirectional	Bidirectional
Temperature and Sea levels	No relationship	Unidirectional: Temperature ⟶ Sea levels
Temperature and Methane pc	No relationship	Unidirectional: Temperature ⟶ Methane
Temperature and GDPpc	No relationship	Bidirectional
Temperature and Agriculture pc	No relationship	Bidirectional
Temperature and Energy consumption pc	No relationship	Bidirectional

Table 7: Summary of relationships for all variables

Our findings from empirical studies indicate that the relationship between temperature and methane is unidirectional, with temperature being the primary driver of changes in methane (Cui et al., 2015). Furthermore, we found that the relationship between temperature and GDP, agricultural, and energy consumption, appears to be bidirectional, with temperature affecting economic output, agricultural productivity, and energy consumption. The increases in economic output, agricultural productivity, and energy consumption also cause the temperature to rise (Yao, 2021) (Kahn et al., 2019) (Mutual Effects of Climate Change and Agriculture, n.d.).

5.2 Reasons why the climate data in our model is different from empirical studies

5.2.1 Additional Factors / Processes not accounted for

Our literature review separately studies the relationship between two variables. For example, processes such as the thermal expansion of seawater are incorporated into the research studies, however, our model does not take into account specific processes and factors that our research study does. Another possible reason is that our model might fail to capture some of the results due to assumptions made that differ from our literature review studies.

5.2.2 Sample Size of Climate Data

Although we have temperature data starting from 1743 to 2024, upon combining it with CO2pc and sea levels, the duration of our climate data shrank to 134 observations, ranging from 1880 to 2013. Given that we require the number of observations to be at least 10 times the number of parameters in our model, we are limited by the types of models we can build since some of the models might be overfitted. Therefore, the statistical conclusions that arise from the models might not be a good representation of the entire picture.

5.3 Reasons why human activity data in our model is different from empirical studies

5.3.1 Sample Size of Human Activity Data

We acknowledge that our sample size for human activity data might be limited as there are only 49 observations in our entire dataset which limits us from building models such as VAR(X) and VECM(X) as those require more observations. Therefore, we are unable to test for Granger causality since there are no suitable VAR models that can be built without severely overfitting our model. This limits our analysis in our statistical tests as we did not take into account the possibility of bidirectional causality in human activity variables and temperature.

5.3.2 Categorisation of Training and Test Sets

We acknowledge that our sample size for human activity data might be limited as there are only 49 observations in our entire dataset, causing the allocation of data into our training, test, and assessment sets to be limited by a small number of observations. This is further an issue when we consider that we had only 28 observations as the training set and 7 as the test set, to forecast 14 years of data. As such, our model might not have learnt the relationships over time well enough, and given a larger time frame of analysis, our statistical results might be more accurate.

5.4 Limitation in Dataset

Besides having a small sample size, other limitations that our data might be subjected to are inconsistency, and inaccuracy in datasets. We retrieved our data from a wide variety of sources which might result in potential differences in data collection. Differences in data collection methods may result in varying conditions that we might not be able to account for systematic errors of measurement. Since climate data is also influenced by external factors such as human activities and natural phenomena (e.g. solar and volcanic activities), it makes the separation of the effects of climate change from other influences challenging and potentially introduces biases into the analysis.

5.5 Limitations in ARIMA-X models

For all the ARIMA-X models, we assume that we have full knowledge of the exogenous variables, and hence, we use the actual values of exogenous variables without forecasting it (i.e. we used exogenous values from the test set). This enables us to vary temperature while keeping everything else constant (ceteris paribus) as the error will only stem from the model's ability to forecast temperature since the error that stems from forecasting exogenous variables will not be considered. However, we do note that if we were to forecast a longer duration into the future, we would then have to predict the exogenous variables too.

Conclusion

Although all of our models conclude that there is no climate change and that it is not caused by humans, we believe that our model is limited in forecasting the true actual trajectory of temperature post-1970s since our models possess some level of error. Therefore, our conclusion is based on the premise that our forecast for temperatures using the model is highly accurate such that it allows us to fully determine whether actual temperatures are statistically different from actual temperature. Furthermore, this is substantiated by the fact that the literature reviews point to the prevalence of climate change since the 19th century with human activities exacerbating the impacts, causing global temperature to rise about 0.8 to 1.2 degrees (IPCC, 2018).

References

Azam, M., Khan, A. Q., Abdullah, H., & Qureshi, M. E. (2015). The impact of CO2 emissions on economic growth: evidence from selected higher CO2 emissions economies. *Environmental Science and Pollution Research*, *23*(7), 6376–6389. https://doi.org/10.1007/s11356-015-5817-4

Chovancová, J., & Tej, J. (2020). Decoupling economic growth from greenhouse gas emissions: the case of the energy sector in V4 countries. *Equilibrium (Toruń)*, *15*(2), 235–251. https://doi.org/10.24136/eq.2020.011

Climate change indicators: U.S. greenhouse gas emissions | US EPA. (2023, December 13). US EPA. https://www.epa.gov/climate-indicators/climate-change-indicators-us-greenhouse-gas-emissions#:~:text=Methane%20emissions%20decreased%20by%2016,activities%20such%20as%20livestock%20production

Cui, M., Ma, A., Qi, H., Zhuang, X., Zhuang, G., & Guo-Hui, Z. (2015). Warmer temperature accelerates methane emissions from the Zoige wetland on the Tibetan Plateau without changing methanogenic community composition. *Scientific Reports*, *5*(1). https://doi.org/10.1038/srep11616

Ezer, T., & Updyke, T. (2024). On the links between sea level and temperature variations in the Chesapeake Bay and the Atlantic Meridional Overturning Circulation (AMOC). *Ocean Dynamics*. https://doi.org/10.1007/s10236-024-01605-y

Gil-Alana, L. A., & Monge, M. (2020). Global CO2 emissions and global temperatures: Are they related. *International Journal of Climatology*, *40*(15), 6603–6611. https://doi.org/10.1002/joc.6601

Lapinskienė, G., Peleckis, K., & Slavinskaitė, N. (2017). ENERGY CONSUMPTION, ECONOMIC GROWTH AND GREENHOUSE GAS EMISSIONS IN THE EUROPEAN UNION COUNTRIES. *Journal of Business Economics and Management (Online)*, *18*(6), 1082–1097. https://doi.org/10.3846/16111699.2017.1393457

Lynas, M., Houlton, B. Z., & Perry, S. (2021). Greater than 99% consensus on human caused climate change in the peer-reviewed scientific literature. *Environmental Research Letters*, *16*(11), 114005. https://doi.org/10.1088/1748-9326/ac2966

Sarmiento, J. L., & Gruber, N. (2002). Sinks for anthropogenic carbon. *Physics Today*, *55*(8), 30–36. https://doi.org/10.1063/1.1510279

Tabash, M. I., Elsantil, Y., Hamadi, A., & Drachal, K. (2024). Globalization and Income Inequality in Developing Economies: A Comprehensive analysis. *Economies*, *12*(1), 23. https://doi.org/10.3390/economies12010023

The Intergovernmental Panel on Climate Change. (IPCC, 2014). Climate Change 2014 Synthesis Report Summary for Policymakers. https://www.ipcc.ch/site/assets/uploads/2018/02/AR5_SYR_FINAL_SPM.pdf

The Intergovernmental Panel on Climate Change. (IPCC, 2018). Global Warming of 1.5 degree.

CO2 Human Emissions. (2017, December 13). Main sources of carbon dioxide emissions: CO2 Human Emissions. https://www.che-project.eu/news/main-sources-carbon-dioxide-emissions#:~:text=Since%20the%20Industrial%20Revolution%2C%20human,dioxide%20concentrations%20in%20the%20atmosphere

Thomas Dietz, Rachael L. Shwom, Cameron T. Whitley (2020, July). Climate Change and Society. https://doi.org/10.1146/annurev-soc-121919-054614

Shijie Li, Chunshan Zhou, Shaojian Wang (2019, May). Does modernisation affect carbon dioxide emissions? A panel data analysis. https://doi.org/10.1016/j.scitotenv.2019.01.373

United States Environmental Protection Agency (n.d.). Getting to the core: the link between temperature and carbon dioxide. https://archive.epa.gov/climatechange/kids/documents/temp-and-co2.pdf

NASA (n.d.). Understanding Sea Level, Thermal Expansion. https://sealevel.nasa.gov/understanding-sea-level/global-sea-level/thermal-expansion/#:~:text=The%20warming%20of%20Earth%20is,increase%20in%20global%20sea%20level.

Yao, J. (2021). Electricity Consumption and Temperature: Evidence from Satellite Data. *Social Science Research Network*. https://doi.org/10.2139/ssrn.3799614

Kahn, M. E., Mohaddes, K., Ng, R. N. C., Pesaran, M. H., Raissi, M., & Jin, Y. (2019). *Long-Term Macroeconomic Effects of Climate Change: A Cross-Country Analysis*. https://doi.org/10.3386/w26167

Mutual effects of climate change and agriculture. (n.d.). https://www.eurasian-research.org/publication/mutual-effects-of-climate-change-and-agriculture/