Predicting Temperature with Monthly Global Average Methane Concentrations

Jingze Dai, Rachael Stephan, & Zhuocheng Zhang

April 25, 2025

Introduction

Greenhouse gases (GHG) are natural, atmospheric gases that can absorb infrared radiation[1]. GHGs will reabsorb heat energy emitted by Earth and retain it within the atmosphere, re-radiating heat back towards the surface[1]. This phenomena is known as the greenhouse effect. Therefore, GHGs are important determinants of Earth's energy budget and resulting climate[1]. Concentrations of greenhouse gases will naturally oscillate and there are other external climate drivers as well (e.g., Milankovich cycles)[1]. However, the anthropogenic consumption of fossil fuels since the Industrial Revolution have resulted in elevated concentrations of GHGs[1]. The most influential greenhouse gases with regards to contribution to the greenhouse effect are carbon dioxide, methane, and water[1].

Methane is commonly regarded as the second most important GHG compared to carbon dioxide - especially considering methane has lower atmospheric concentrations, but carbon dioxide and methane are difficult to compare due to their residence time and radiative forcing capabilities[2]. Methane has a more potent radiative forcing (i.e., higher heat retention) but a residence time on the order of decades compared to carbon dioxide's centuries[1]. Assuming an equal amount of each gas is released, this will result in methane having more power over climate warming than carbon dioxide at small time scales[2]. Over 20 years in this scenario, methane will be 80 times as powerful as carbon dioxide[2]. Over the next 100 years in this scenario, methane will be 28 times as powerful as carbon dioxide[2]. Therefore, methane is an important consideration for near-future climate considerations, such as within our lifetime.

Two of the major contributors to methane emissions are agricultural practices and fossil fuels[3]. The fermentation during the digestion process in cattle is the primary source of agricultural methane. It is estimated that up to 264 pounds of methane worldwide are emitted from cattle annually[4]. Methane emitted through fossil fuels are caused by extraction leaks, crude oil processing, and transportation of the fossil fuels. This results in an estimated global annual emission of 120 million metric tons[5].

Monitoring atmospheric GHG concentrations is important for climate monitoring and tracking the effectiveness of climate change mitigation – including in the green energy transition. This report will investigate the temporal trends in methane concentrations considering the exogenous factors of beef production and socioeconomic factors (as a proxy for fossil fuel consumption). We will use our methane projections to extrapolate trends in climate, represented by temperature anomalies based on an averaged climate from 1901 to 2000.

[1] Mann, M.E. (2025, April 11). greenhouse gas. Encyclopedia Britannica. https://www.britannica.com/science/greenhouse-gas

[2] MIT Climate (2024, January 4). Why do we compare methane to carbon dioxide over a 100-year timeframe? Are we underrating the importance of methane emissions? MIT Climate Portal. tps://climate.mit.edu/ask-mit/why-do-we-compare-methane-carbon-dioxide-over-100-year-timeframe-are-we-underrating

[3] Raymond, P., & Hamburg, S. (2024, November 18). Yale experts explain methane emissions. Yale Sustainability. https://sustainability.yale.edu/explainers/yale-experts-explain-methane-emissions

[4] U.S. Environmental Protection Agency. (2020, October). Agriculture and aquaculture: Food for thought. https://www.epa.gov/snep/agriculture-and-aquaculture-food-thought

[5] International Energy Agency. (2024). Global Methane Tracker 2024: Key findings. IEA. https://www.iea.org/reports/global-methane-tracker-2024/key-findings

Data Sources

The following data sources were used in this analysis and their respective wrangling is described below.

1. *Globally averaged methane concentration data*

- Frequency: monthly
- Units: ppb
- Format: csv file
- Source: NOAA Global Monitoring Laboratory

The initial methane data set was an excel `.csv` file with columns of the year, month, decimal year, average methane concentration, average uncertainty, trend value, and trend uncertainty on a single sheet. It was processed as follows:

- The header and data were read in as separate data frames because of the excel file formatting.
- The header information was set as the column names for the data.
- A date was created using the month and year columns using lubridate's `make_date` function. The day was assumed to be the first of the month.
- Data set was verified for missing data - none were present.
- Data frames were saved as the full data, training period data, and test period data. The training and testing periods are expanded upon in the methods section.

2. *Globally averaged temperature anomaly data (based on a 1901-2000 average)*

- Frequency: monthly
- Units: degrees Celsius
- Format: csv file
- Source: NOAA National Center for Environmental Information

The initial temperature anomaly data set was an excel `.csv` file with columns of the date and temperature anomaly on a single sheet. It was processed as follows:

- Only the data was read into a data frame from the excel file.
- The column names for the data frame were manually set.
- The date column was converting into date format using lubridate's `ym` function.
- Data set was verified for missing data - none were present.
- Data frames were saved as the full data, training period data, and test period data. The training and testing periods are expanded upon in the methods section.

3. *Aggregated global beef production*

- Frequency: yearly
- Units: metric tons
- Format: csv file
- Source: U.S. Department of Agriculture

4. *Socioeconomic factors data set including consumer price index (CPI), industrial production, merchandise exports, and merchandise imports.*

- Frequency: monthly
- Units: index (2005=100)
- Format: csv file
- Source Federal Reserve Bank of Dallas

Data Summary

The data frames were not combined for analysis, but the relevant factors have been merged into a data frame for summary purposes below. All of the monthly data is shown in a single data frame and yearly data (i.e., beef production) is shown in another data frame)

Table 1: First 10 cases of global methane (ppb), temperature anomaly (degrees celsius), and socioeconomic indicators

Date	Methane	Temperature.Anomaly	CPI	Exports	Imports	Industrial.Production
1983-07-01	1625.96	0.23	4.35	14.42	30.64	43.65
1983-08-01	1628.05	0.36	4.41	14.72	31.78	44.29
1983-09-01	1638.43	0.42	4.47	15.05	31.32	44.77
1983-10-01	1644.80	0.22	4.53	15.00	33.81	44.78
1983-11-01	1642.61	0.35	4.59	15.23	32.02	45.15
1983-12-01	1639.54	0.24	4.65	15.28	31.74	45.61
1984-01-01	1638.73	0.31	4.71	15.80	35.95	45.97
1984-02-01	1638.82	0.20	4.77	16.00	36.07	46.24
1984-03-01	1640.87	0.31	4.84	16.03	36.68	46.20
1984-04-01	1643.96	0.12	4.90	15.66	38.07	46.17

Table 2: First 10 cases of beef production (metric tons)

Year	Beef.Production
1983	49034412
1984	50439336
1985	51351004
1986	53107880
1987	53075240
1988	53568584
1989	53795532
1990	55351170
1991	55351156
1992	54537624

Table 3: Summary statistics for all of the variables used in the analysis from July 1, 1983 to November 1, 2024, Beef Production is provided until 2023.

Variable	Count	Mean	SD	Min	Max	Pct25	Pct75
Methane	497	1781.26718309859	70.5946016831821	1625.96	1941.81	1739.23	1820.87
Temperature.Anomaly	497	0.606700201207243	0.290952233671071	0.02	1.44	0.39	0.8
CPI	497	93.2523138832998	57.3428981287527	4.35	208.43	45.95	139.02
Exports	497	105.180482897384	71.2689141432184	14.42	259.8	38.7	173.05
Imports	497	82.8749094567404	31.0533848481298	30.64	140.91	47.9	112.91
Industrial.Production	497	99.844245472837	38.7991876406803	43.65	177.6	62.6	132.51
Beef Production	41	61971509.6585366	8101405.92508226	49034412	76560810	54537624	69255790

Analysis

Training and Test Periods

The training period was designated as July 1, 1983 to December 1, 2021. The test period was designated as January 1, 2022 to November 1, 2024. These periods were used to filter the data to create test and training data subsets.

Methane Concentration Predictions

GENERAL DESCRIPTION OF EACH METHANE DATA SET WITH GENERAL GRAPH

The methane and temperature time series were decomposed into the component parts with the `decompose` function. The ACF and PACF were also examined. **EXPLAIN ACF AND PACF PLOT**

ACF AND PACF PLOT

The following models were used to predict methane concentrations without the use of exogenous variables.

1. ARIMA using `auto.arima` and decomposed seasonality
2. SARIMA using `auto.arima` with Fourier terms
3. ARIMA with Fourier terms with `auto.arima` function
4. ETS Model with the `stlf` function
5. Neural Network Model with the `nnetar` function
6. TBATS Model with the `tbats` function
7. State Space Model - Exponential Smoothing with the `es` function
8. State Space Model - BSM with the `StructTS` function

The following models were used to predict methane concentrations with the use of exogenous variables (i.e., using the `xreg` term).

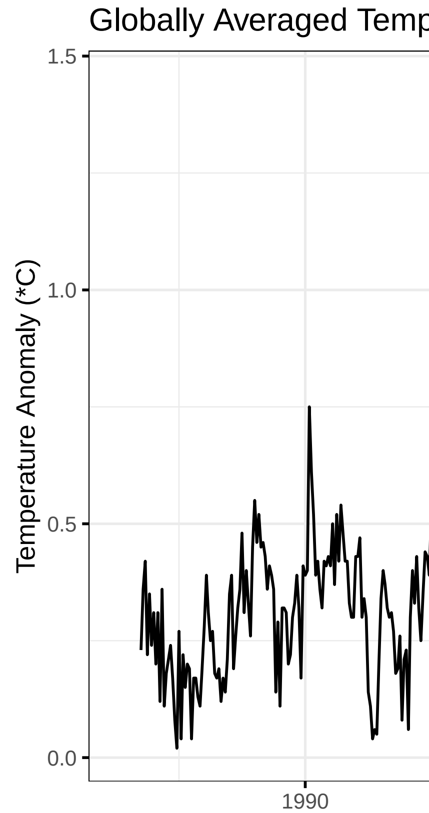
1.

All of the above models were averaged together in every possible combination.

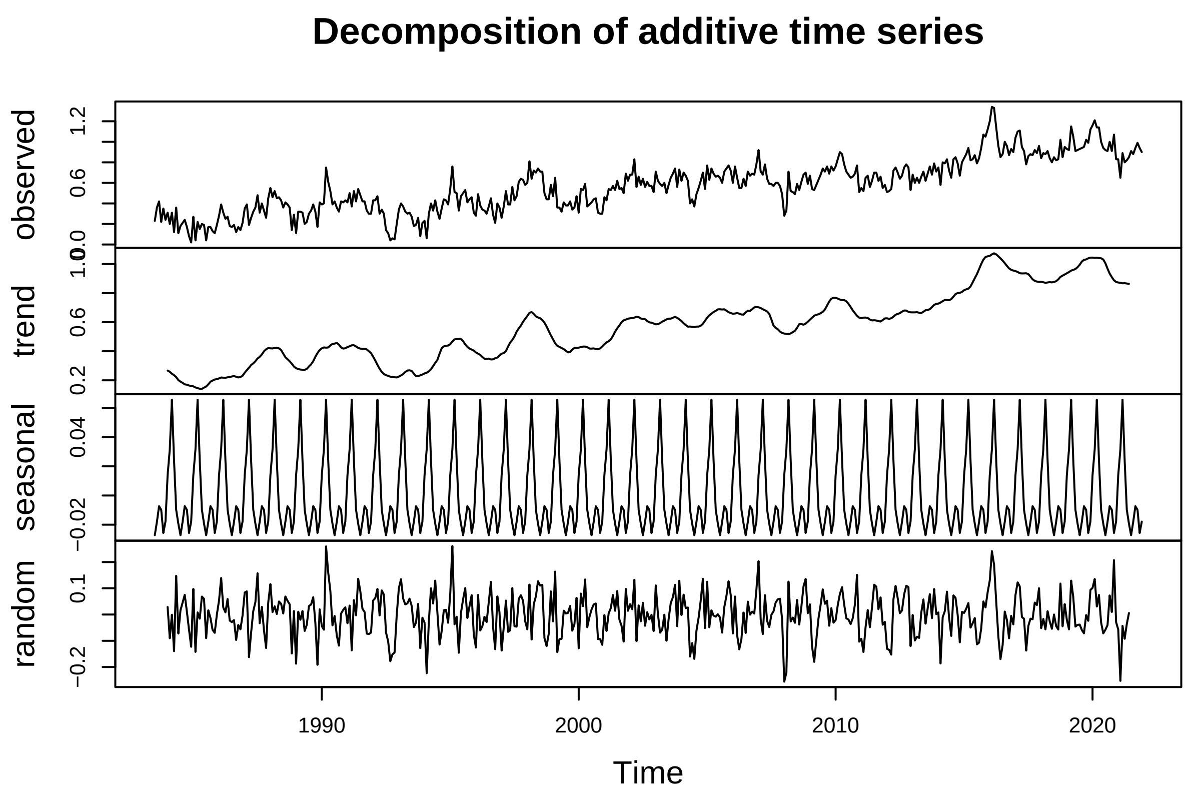
The performance of all of the models were examined with the `accuracy` function and the methane test data subset. The performance metrics for the top 10 models are shown below.

TALK ABOUT TOP 3 MODELS AND LOOK AT THEIR RESIDUALS

Temperature Anomaly Predictions



The temperature anomaly data has a general increasing trend. There is a seasonal trend.



Conclusions

Delete this title: Footnotes