

# Analysis of Global Warming Using Machine Learning

Harvey Zheng

Phillips Academy, Andover, MA, USA

Email: zheng.harvey5@gmail.com

**How to cite this paper:** Zheng, H. (2018) Analysis of Global Warming Using Machine Learning. *Computational Water, Energy, and Environmental Engineering*, 7, 127-141.

<https://doi.org/10.4236/cweee.2018.73009>

**Received:** March 16, 2018

**Accepted:** July 28, 2018

**Published:** July 31, 2018

Copyright © 2018 by author and  
Scientific Research Publishing Inc.

This work is licensed under the Creative

Commons Attribution International

License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Climate change is a controversial topic of debate, especially in the US, where many do not believe in anthropogenic climate change. Because its consequences are predicted to be dire, such as a mass ocean extinction and frequent extreme weather events, it is important to learn what causes the warming in order to better combat it. In this study, the first challenge dwells on how to construct reliable statistical models based on massive climate data of 800,000 years and accurately capture the relationship between temperature and potential factors such as concentrations of carbon dioxide (CO<sub>2</sub>), nitrous oxide (N<sub>2</sub>O), and methane (CH<sub>4</sub>). We compared the performance several mainstream machine learning algorithms on our data, which includes linear regression, lasso, support vector regression and random forest, to build the state of the art model to verify the warming of the earth and identifying factors contributing the global warming. We found that random forest outperforms other algorithms to create accurate climate models which use features including concentrations of different greenhouse gases to precisely forecast global atmosphere. The other challenges in identifying factor importance can be met by the feature of ensemble tree-based random forest algorithm. It was found that CO<sub>2</sub> is the largest contributor to temperature change, followed by CH<sub>4</sub>, then by N<sub>2</sub>O. They all had some sorts of impact, though, meaning their release into the atmosphere should all be controlled to help restrain temperature increase, and help prevent climate change's potential ramifications.

## Keywords

Global Warming, Machine Learning, Prediction

## 1. Introduction

The general scientific consensus is that the Earth is warming. Over the past century, the temperature has already climbed 0.5°C [1]. The “warming of the cli-

mate is unequivocal”, with the last decade being the warmest decade since 1850 [2]. However, there is still debate over if global warming is actually occurring, and if it is, then if it is anthropogenic. A majority, about 51%, of people in the US do not believe in anthropogenic climate change, with 31% of these people saying the warming is natural, and 20% saying the warming is not occurring [3]. With the US being the only country out of 196 in the UN to not sign the Paris Agreement, a commitment to combat climate change, it is clear that there are still many people in the world that deny climate change is caused by humans or is even occurring at all.

An argument against anthropogenic global warming of the Earth is due to the increased solar activity in the past few years. This point is moot, since the Sun goes through an eleven-year cycle of solar activity, and the Earth has been continuously warming for the past decade, which does not make sense. As others have said, solar activity has no correlation with global temperature [2]. Of course, other conjectures against climate change exist.

According to the Intergovernmental Panel on Climate Change’s (IPCC) latest report, the main driving force of global warming is the increase in concentration of carbon dioxide in the atmosphere [2]. The vast majority of carbon dioxide recently added to the air is from burning fossil fuels, or because of humans [2]. Increased CO<sub>2</sub> concentration in the air causes increased temperature on the earth, which is known as the greenhouse effect, meaning the atmosphere will trap heat that is released by the sun [4].

So why we should worry about increased CO<sub>2</sub> concentration and global warming? The consequences of global warming can be catastrophic. The increased CO<sub>2</sub> concentration in the air will also lead to an increase in CO<sub>2</sub> absorbed by the ocean, which means the ocean will become more acidic. The pH of the ocean has already decreased 0.1 [2]. If the ocean becomes too acidic, the results will be dire, as many organisms will be unable to adapt to acidity, resulting in significant loss of coral reefs and other underwater organisms [5].

In addition, many studies outline dire consequences involved with the global warming effects. According to IPCC report, the number of hurricanes, as well as the intensity of hurricanes, will increase due to the warming ocean water, putting coastal states at risk [2]. In general, extreme weather events, such as droughts and floods, will occur more often with global warming. In addition, a warmer temperature means less ice in glaciers and the polar ice caps, and will result in a significant rise in sea level [2]. The current projection is 50 cm to 100 cm by 2100, leaving many cities underwater [2]. Climate change will also cause mass migrations, both within countries and across borders, since more people will lose their homes to extreme weather [2]. Food security will also become an issue for many countries [2]. Although increased CO<sub>2</sub> can help with crop production, many other factors will mitigate the benefits, such as lack of water due to droughts and changing temperature, causing a decrease in output [2]. Increased CO<sub>2</sub> concentration in the atmosphere can also decrease the nutritional

value of crops [2].

However, as more studies point out, there are many important factors that contribute to the global warming besides the concentration of  $\text{CO}_2$ . Gases including  $\text{CH}_4$  have much stronger global warming effects than  $\text{CO}_2$  [4]. However, there is no research on studying the importance of these factors that contribute to global warming. By learning which factors contribute more to global warming, society can work on mitigating the effects of the factors.

In order to determine the next step to help mitigate climate change, the main factors that drive climate change should be investigated to know how significant each factor is. This study will focus on these factors, as well as many other factors that have potential to cause differences in global temperature. A previous study of temperature over the past 1000 years was conducted, including solar activity, volcanic activity, and greenhouse gas (GHG) concentration [6]. In the study, it was found that greenhouse gases predicted the temperature closer than the other two factors did [6].

Meanwhile, machine learning has been applied more and more widely on environmental protection problems and achieves promising results. Chen *et al.* [7] explored the application of double parallel feed forward neural network on estimating the suspended sediment loads to assist water resources management. Olyaie *et al.* [8] compared performance of different neural networks on suspended sediment load of river system. Artificial neural network was also studied for evaluating energy consumption and environmental life cycle for incineration and landfill system in [9]. Taormina [10] combined neural network with base flow separation and binary-coded swarm optimization to forecast river quantities.

Theory of variable fuzzy sets and fuzzy binary comparison method have been investigated on assessing water quality in [11]. Those works demonstrate the applicability of machine learning techniques on environmental issues.

In this paper, our first aim is to validate global warming based on the collected public data. After, machine learning algorithms are employed to investigate the effects different factors have on the global temperature. Then, we will analyze the plots generated from the algorithms, as well as draw conclusions from the plot.

The paper proceeds as follows: Section 2 is about the dataset we have. Section 3 is about how the data was used in conjunction with different machine learning algorithms and what the algorithms are. Section 4 is about the results from the machine learning analysis of the data. Section 5 summarizes the results and includes how the findings from this paper can be used in future projects.

## 2. Data

### 2.1. Data Collection

Data from the past 800,000 years will be compiled from a variety of public databases, such as the National Oceanic and Atmospheric Administration and the United States Environmental Protection Agency. The data used will include:  $\text{CO}_2$

in parts per million (PPM) [12], N<sub>2</sub>O in parts per billion (PPB) [13], CH<sub>4</sub> in PPB [14], the year, and temperature difference between the average temperature of the last 100 years [15]. More accurate temperature data over the past 100 years is obtained from Lawrence Berkeley National Lab. NO<sub>2</sub>, CH<sub>4</sub>, and CO<sub>2</sub> are used because they are all greenhouse gases that help cause climate change [16] [17] [18].

## 2.2. Data Preprocessing

The data collected over the 800,000 years are not aligned with each other. For example, there may be CO<sub>2</sub> and a corresponding temperature in year 1900, but may lack the corresponding N<sub>2</sub>O and CH<sub>4</sub> concentration at that time. To prepare the data for machine learning, we use linear interpolation to align the data, since machine learning algorithms cannot handle missing data points effectively.

## 3. Methodology

### 3.1. Temperature Increase Analysis

The global temperature change over the past 100 years will be visualized based on the public data provided by Lawrence Berkeley National Lab. The trend of global warming can be observed in the plotted average global temperature over the past 70 years. The coefficient of determination ( $R^2$ ) between global temperature and time is also computed, which can further validate statistically the increase of global temperatures along with time.

### 3.2. Factor Analysis

To investigate the possible factors that contribute to the global temperature increase, we need to conduct factor analysis on potential factors such as CO<sub>2</sub> concentration. Many research works have been conducted to show there is a strong relationship between temperature and CO<sub>2</sub>. The common technique to analyze potential factors includes visual check and statistical correlation computation. In this work, we first visualize the variations of temperature and CO<sub>2</sub>, and we also compute the  $R^2$  to validate the correlation observed statistically.

### 3.3. Applying Machine Learning Algorithms

Machine learning is a collection of statistical methods to analyze trends, find relationships, and develop models to predict things based on data sets. The machine learning algorithms we explore for this global warming study are random forest, support vector regression (SVR), lasso, and linear regression.

#### 3.3.1. Random Forest

Random forest is an algorithm that uses trees as building blocks to construct more powerful prediction models. The algorithm takes an ensemble of a certain number of trees. When building these decision trees, the splits will be based off a random number of predictors, less than the number in the full set. By restricting

the number of predictors in each tree, the strong predictors do not drown out weaker predictors, and the final result (the average of the results of each decision tree) of many uncorrelated trees will reduce variance of the predictions. The averaged final result will also be more accurate than if all predictors were used, as a strong predictor won't always be used, decorrelating the trees from certain predictors, and making the average less variable and thus more reliable.

### 3.3.2. Support Vector Regression

Support vector machines, or SVM, are algorithms that use hyperplanes (a line in more than 3 dimension) to create regressions. Essentially, the algorithm tries to separate the different types of data using a hyperplane that has the largest margin between the groups in a multi-dimensional space. If there is a point of data outside the margin, then there will be a penalty that will affect if the hyperplane really is the optimal choice. SVM can use different kernels, or different ways of finding the hyperplane in a high dimensional space. Support vector regression (SVR) is an extension of this, creating a regression from the principles of SVM. SVR, like in other regressions, also has a loss function, but it is only increased when the residuals are greater than a certain constant.

### 3.3.3. Lasso

Lasso, or least absolute shrinkage and selection operator, is an algorithm that uses shrinkage, or when data is shrunk toward a certain point like the mean. The algorithm uses L1 regularization, which adds penalty based on the sum of the absolute value of coefficients, and will shrink some coefficients to zero if they play no role. This prevents the model from over fitting and creating a more general model. At the same time, lasso tries to minimize the sum of squares of the data.

### 3.3.4. Results and Analysis

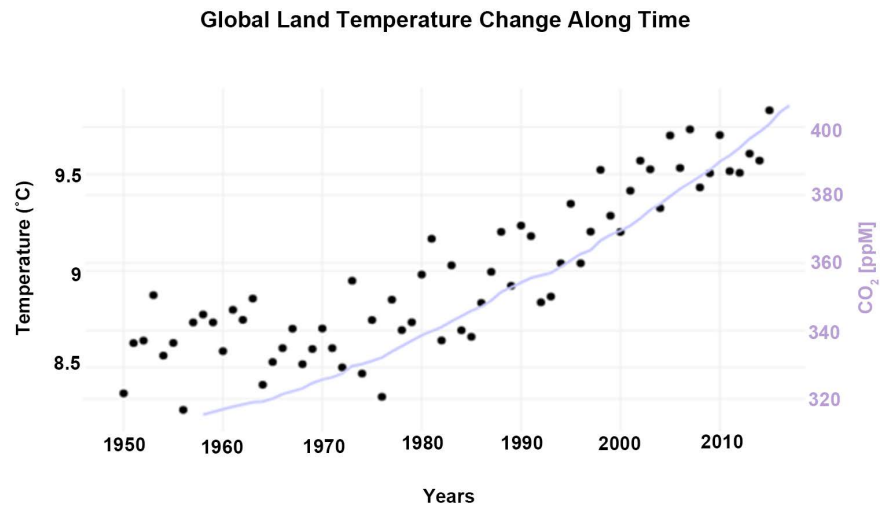
With the results, many conclusions can be drawn, since random forests output feature correlations and such using numbers. This will be conducted multiple times and averaged to get as accurate of a result as possible.

## 4. Results

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your journal for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper.

### 4.1. Temperature Change Over Time

Data about the temperature and the CO<sub>2</sub> concentration over the past 70 years were plotted on a graph (**Figure 1**). From the plot, the trend that the temperature has warmed over the past few decades is present. In addition, the graph of



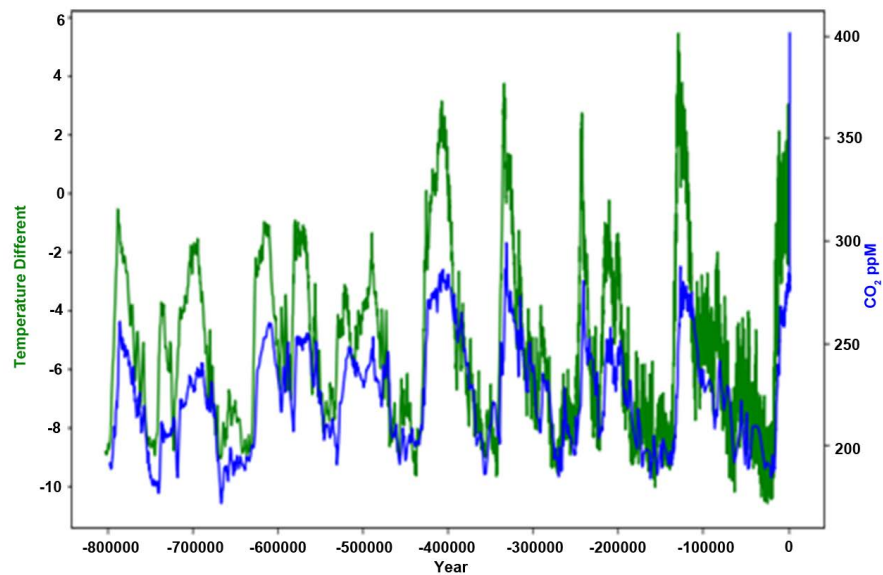
**Figure 1.** Plot of CO<sub>2</sub> ppm and average temperature since 1950.

the concentration of carbon dioxide also correlates with the temperature graph, suggesting that they are related and that it may be a large cause of the warming of the Earth.

To further verify the relationship between carbon dioxide and temperature, as shown in **Figure 2**, other data with the CO<sub>2</sub> concentration over the past 800,000 years were plotted with the difference in temperature when compared to the average of the past 100 years. Through inspection of the new plot, it can be seen that they are heavily related, and that the concentration of CO<sub>2</sub> heavily influences the temperature of the Earth. Whenever the concentration in CO<sub>2</sub> rises, the temperature rises, and vice versa. Since the increase in CO<sub>2</sub> concentration has been attributed to humans, and CO<sub>2</sub> PPM and temperature seem to be related, it can be inferred that humans caused the rise in temperature through an increase in CO<sub>2</sub>.

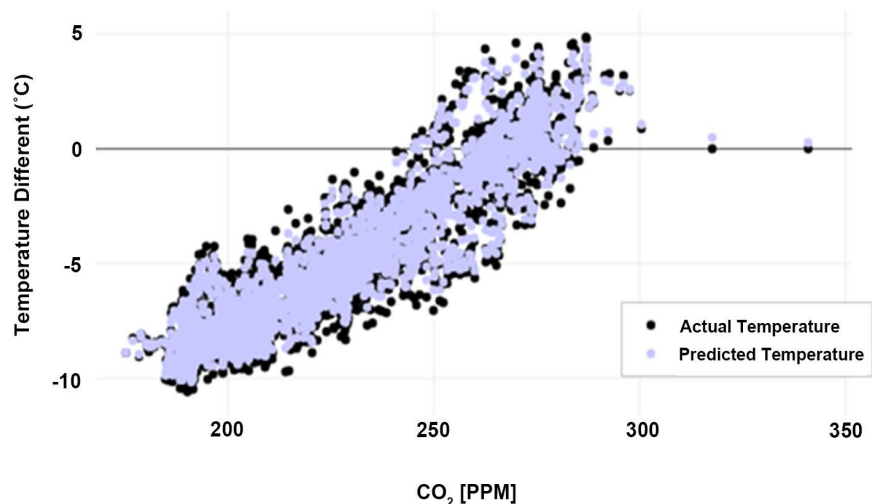
#### 4.2. Applying Machine Learning Algorithms

The data collected over the past 800,000 years was randomly split into two even samples, one for training and one for testing. We further employed 8-fold cross validation during training process to search for suitable hyperparameters and prevent models from overfitting during training. Then, three different machine learning algorithms were compared: random forest, lasso, and support vector regression. With each algorithm, the parameters were tuned to fit the data and generate accurate training results. The visual results are shown in **Figures 3-14**. Here, we provide the key hyperparameters we used for different machine learning algorithms here. The hyperparameters here are selected in hyperparameter ranges we provided by using the 8-fold cross validation. The selected hyperparameters for random forest are 300 for number of trees used, 2 for max number of features, 1 for minimum number of samples required to be at a leaf node. For SVR, we use 2.0 for penalty C of the error term and radial basis function kernel



**Figure 2.** Plot of CO<sub>2</sub> ppm and temperature difference from the average of the last 100 years over the previous 800,000 years.

#### Predicted vs. Actual Temperature Change (Random Forest, Training)



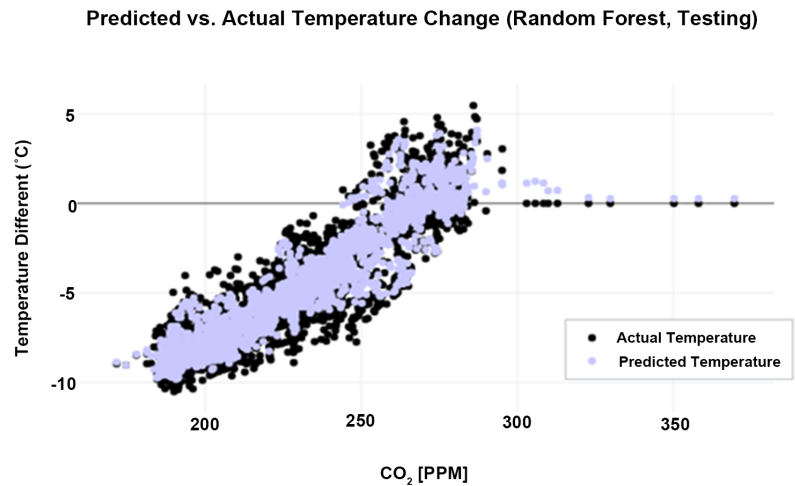
**Figure 3.** Plot of predicted temperature using random forest and the actual temperature from the training set vs. the concentration of CO<sub>2</sub>.

based on cross-validation results. We use 1.0 for regularization term coefficient in Lasso algorithm.

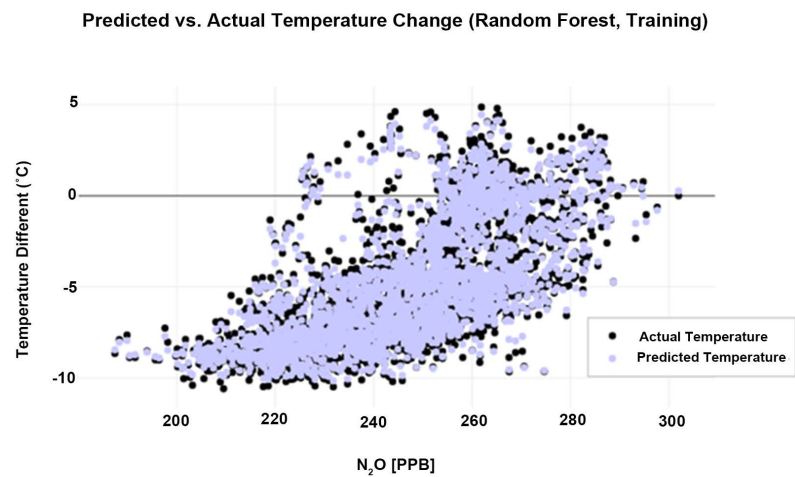
The resulting predictions were then graphed against the values from the data set. The plots for the two most accurate algorithms are shown below.

We aim to use the trained model to predict temperature given different potential factor values, therefore our problem is a regression problem. Mean squared error (MSE) measures the average of the squares of errors between our model predictions and real data and is suitable for regression problems. Other

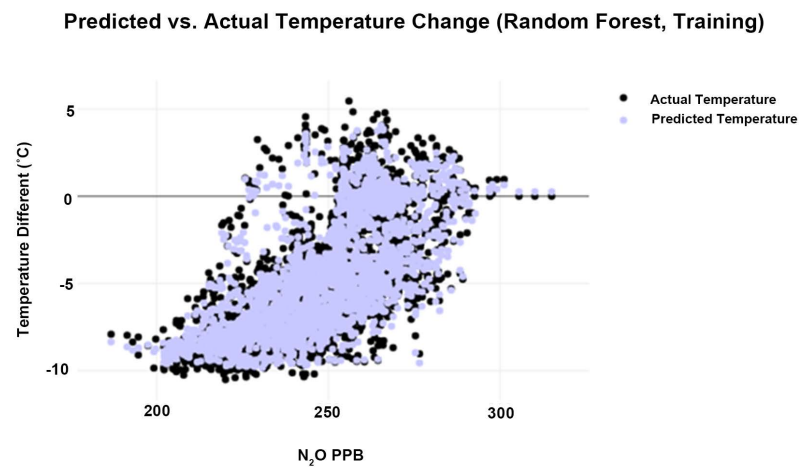




**Figure 4.** Plot of predicted temperature using random forest and the actual temperature from the testing set vs. the concentration of CO<sub>2</sub>.

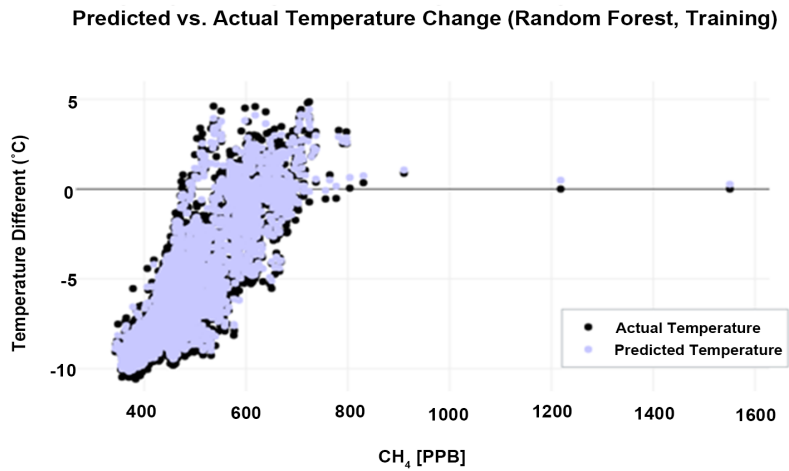


**Figure 5.** Plot of predicted temperature using random forest and the actual temperature from the training set vs. the concentration of N<sub>2</sub>O.

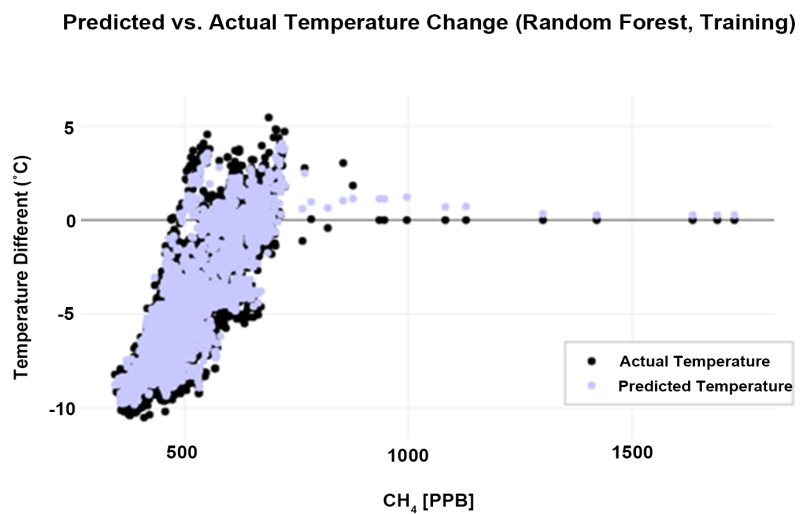


**Figure 6.** Plot of predicted temperature using random forest and the actual temperature from the testing set vs. the concentration of N<sub>2</sub>O.

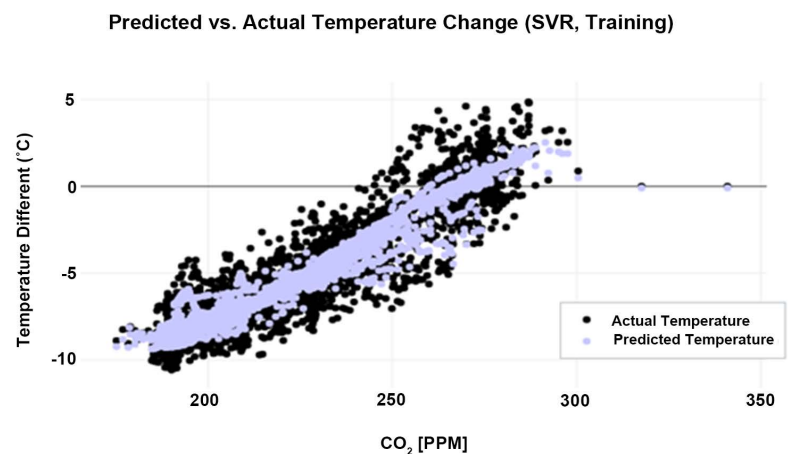




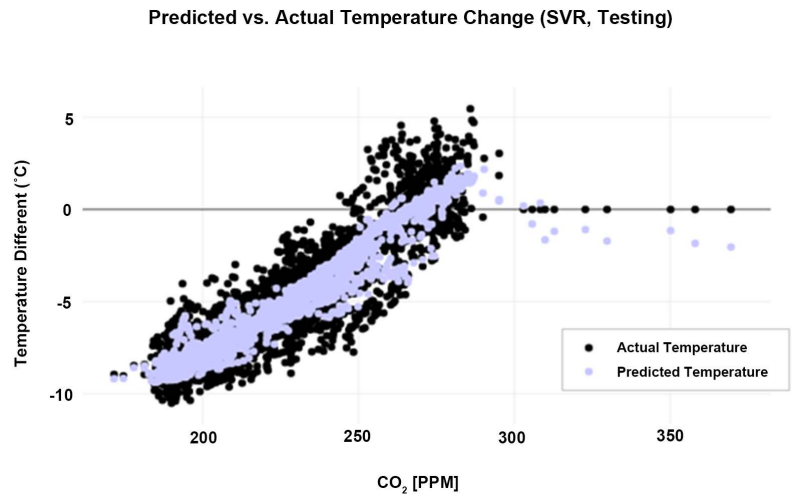
**Figure 7.** Plot of predicted temperature using random forest and the actual temperature from the training set vs. the concentration of  $\text{CH}_4$ .



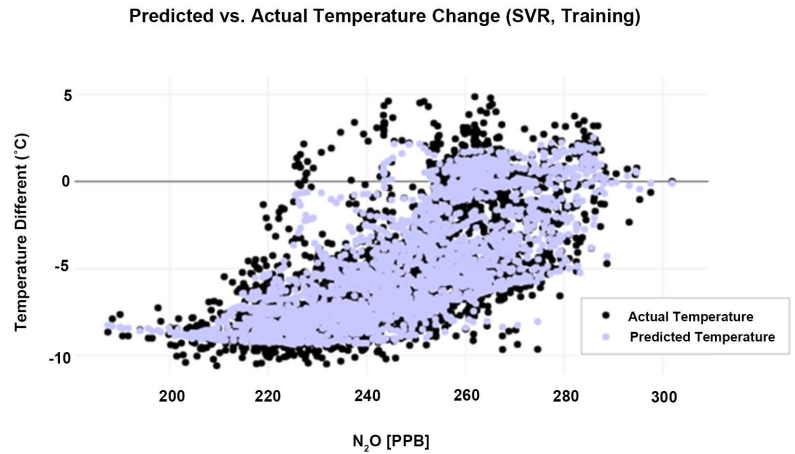
**Figure 8.** Plot of predicted temperature using random forest and the actual temperature from the testing set vs. the concentration of  $\text{CH}_4$ .



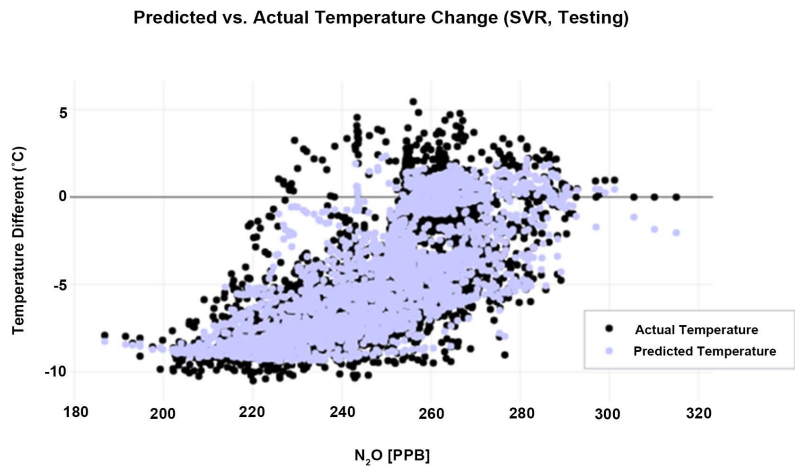
**Figure 9.** Plot of predicted temperature using SVR and the actual temperature from the training set vs. the concentration of  $\text{CO}_2$ .



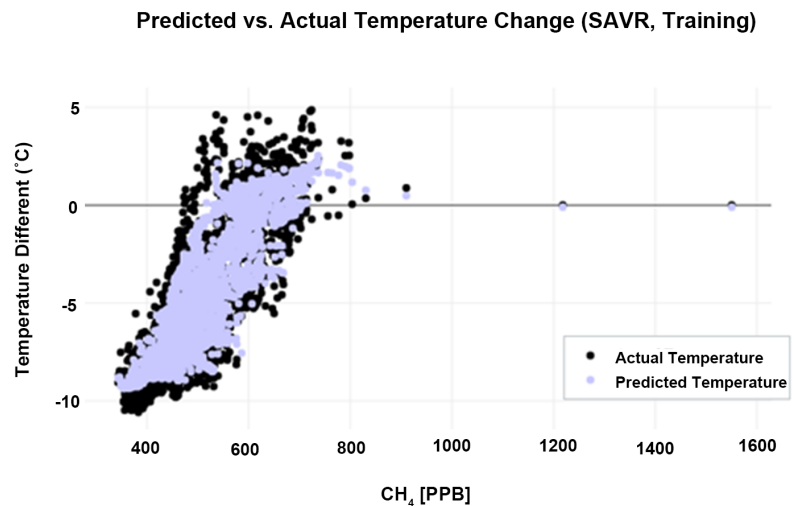
**Figure 10.** Plot of predicted temperature using SVR and the actual temperature from the testing set vs. the concentration of  $\text{CO}_2$ .



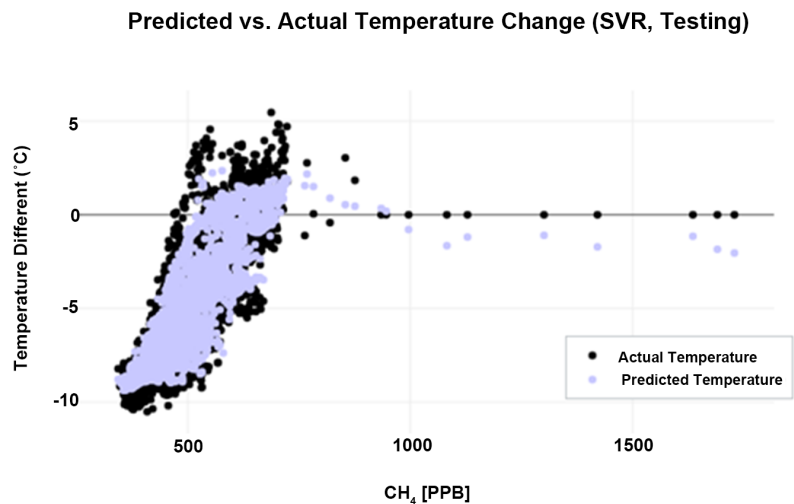
**Figure 11.** Plot of predicted temperature using SVR and the actual temperature from the training set vs. the concentration of  $\text{N}_2\text{O}$ .



**Figure 12.** Plot of predicted temperature using SVR and the actual temperature from the testing set vs. the concentration of  $\text{N}_2\text{O}$ .



**Figure 13.** Plot of predicted temperature using SVR and the actual temperature from the training set vs. the concentration of  $\text{CH}_4$ .



**Figure 14.** Plot of predicted temperature using SVR and the actual temperature from the testing set vs. the concentration of  $\text{CH}_4$ .

score criteria such as mean absolute error can also be used, but they provide no better fitting models for our problem, so we use MSE to quantify the accuracy of the model employed. The training and testing MSE results for compared algorithms are shown in [Table 1](#). It is clear that random forest creates the most accurate models. From the inspection of the plots, random forest is visually more accurate, and creates the most accurate model for predicting the temperature differences based on the concentrations of  $\text{N}_2\text{O}$ ,  $\text{CO}_2$ , and  $\text{CH}_4$ . We see that random forest runs efficiently on this dataset and has an effective method to estimate missing data. Thus, to build a more accurate model to predict the temperature with a larger set of features, random forest would be the best option out of these four algorithms. The accuracy of the algorithm also allows for an accurate feature importance chart in [Table 2](#).

**Table 1.** Mean squared error of each model within the training or testing data.

Algorithm	Training MSE	Testing MSE
Random Forest	0.1289	0.9557
Lasso	1.8819	2.2088
SVR	1.3740	1.5267
Linear Regression	1.8796	2.2689

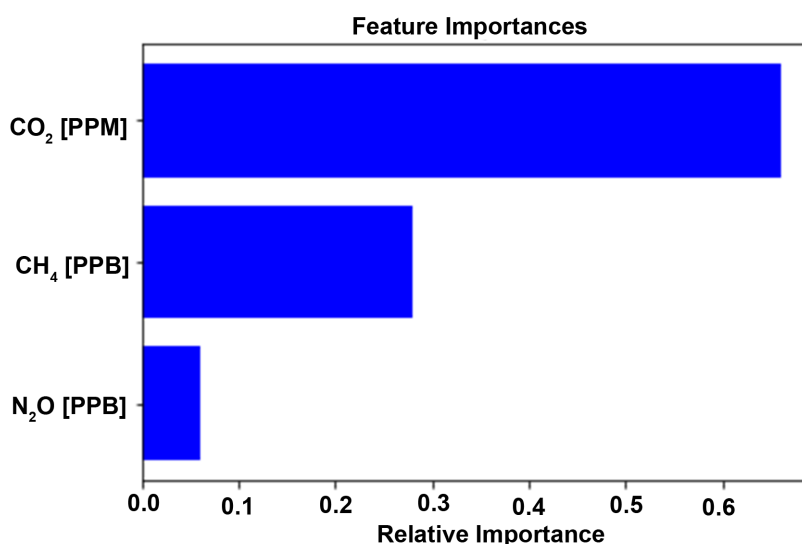
**Table 2.** Importance of each feature, as determined by Random Forest.

Feature	Relative Importance
CO <sub>2</sub> PPM	0.6598
CH <sub>4</sub> PPB	0.2795
N <sub>2</sub> O PPB	0.0607

As visible from the feature importance chart in **Figure 15**, CO<sub>2</sub> is the most significant feature in temperature change, at a factor of 0.6598, followed by methane, which has a factor of 0.2795, then the least significant would be N<sub>2</sub>O, at 0.0607. Through machine learning, the claim set forth by the IPCC and other studies, that CO<sub>2</sub> is the biggest contributor to temperature change, is confirmed. The chart also shows that the effects of CH<sub>4</sub> and N<sub>2</sub>O are also prevalent, and still affect the temperature of the earth.

Carbon dioxide is a very big factor in determining the temperature of the air. This means that the amount of carbon dioxide that humans (and not nature) are putting into the air is contributing a large amount to the changes in temperature [14]. Both methane and N<sub>2</sub>O have considerable impacts as well. In fact, there is actually little methane in the air when compared to CO<sub>2</sub> (about 1.82 PPM for CH<sub>4</sub> vs. about 399 PPM for CO<sub>2</sub>) [19], yet the effect of methane is still massive and should never be underestimated because a unit methane has much greater greenhouse effect than a unit CO<sub>2</sub>. Even if there isn't much of a gas in the air, it can still change the temperature. Thus, attention should be paid to all three of these gases.

We proved these three factors contribute to global warming when they are increased in concentration. As the IPCC noted, the effects of global warming can be catastrophic [2]. Further, we can use our constructed model and combine the greenhouse releases prediction data to forecast the temperature in the future, which can contribute to the control and prediction of global warming. Now that we have verified the effects of greenhouse gases on the global temperature, the next step is to figure out how to limit the concentrations of these gases in the atmosphere, in order to slow the global temperature rise. In addition, with more data about the concentration of different greenhouse gases going back thousands of years, the models can be strengthened and become more accurate, and can also determine if the extent to which other gases affect the temperature.



**Figure 15.** Importance of each feature, as determined by Random Forest.

## 5. Conclusions

As evident from the first part of the results, there is an upward trend in temperature, which correlates with the upward trend in CO<sub>2</sub> concentration. From the correlation analysis between the concentration of CO<sub>2</sub> and the temperature, we further show that increase in CO<sub>2</sub> concentration causing the temperature rise.

Afterward, we compared different machine learning algorithms in predicting the temperature using the concentrations of three gases: CO<sub>2</sub>, CH<sub>4</sub>, and N<sub>2</sub>O. It is apparent that random forest is by far the most accurate algorithm of the three tested. By adding more features and more data to train it, it will become even more accurate, and become a useful model for temperature change. This means by predicting the future outputs of CO<sub>2</sub>, CH<sub>4</sub>, N<sub>2</sub>O, and any other features that the algorithm is trained with, random forest will accurately predict the temperature.

The feature importance data gathered from random forest also tells an important story. In our study, we show that CO<sub>2</sub> dominates the global temperature changes, but it is important to note that the unit of CH<sub>4</sub> and N<sub>2</sub>O is ppb while the unit of CO<sub>2</sub> is ppm, which indicates that the effect of CH<sub>4</sub> and N<sub>2</sub>O should never be underestimated.

In current work, only three factors are considered as contributing factors to temperature change can be further considered such as atmospheric circulation, currents, and biodiversity. We compared four machine learning algorithms which have been proven to provide satisfactory performance in many cases. However, other machine learning algorithms, especially ensemble-based algorithms such as xgboost, as well as neural network can also be investigated for seeking better models in future work.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] NASA (n.d.) How Is Today's Warming Different from the Past? <https://earthobservatory.nasa.gov/Features/GlobalWarming/page3.php>
- [2] Pachauri, R.K. and Meyer, L. (2015) Climate Change 2014: Synthesis Report. [https://www.ipcc.ch/pdf/assessment-report/ar5/syr/SYR\\_AR5\\_FINAL\\_full\\_wcover.pdf](https://www.ipcc.ch/pdf/assessment-report/ar5/syr/SYR_AR5_FINAL_full_wcover.pdf)
- [3] Funk, C. and Kenedy, B. (2016) Americans' Views on Climate Change and Climate Scientists. <http://www.pewinternet.org/2016/10/04/public-views-on-climate-change-and-climate-scientists/>
- [4] Environmental Protection Agency (n.d.) Overview of Greenhouse Gases. <https://www.epa.gov/ghgemissions/overview-greenhouse-gases>
- [5] NASA (n.d.) The Consequences of Climate Change. <https://climate.nasa.gov/effects/>
- [6] Crowley, T.J. (2000) Causes of Climate Change over the Past 1000 Years. *Science*, **289**, 270-277. <https://doi.org/10.1126/science.289.5477.270>
- [7] Chen, X.Y. and Chau, K.W. (2016) A Hybrid Double Feedforward Neural Network for Suspended Sediment Load Estimation. *Water Resources Management*, **30**, 2179-2194. <https://doi.org/10.1007/s11269-016-1281-2>
- [8] Olyae, E., *et al.* (2015) A Comparison of Various Artificial Intelligence Approaches Performance for Estimating Suspended Sediment Load of River Systems: A Case Study in United States. *Environmental Monitoring and Assessment*, **187**, 189. <https://doi.org/10.1007/s10661-015-4381-1>
- [9] Ashkan, N.-P., *et al.* (2017) Modeling of Energy Consumption and Environmental Life Cycle Assessment for Incineration and Landfill Systems of Municipal Solid Waste Management-A Case Study in Tehran Metropolis of Iran. *Journal of Cleaner Production*, **148**, 427-440. <https://doi.org/10.1016/j.jclepro.2017.01.172>
- [10] Taormina, R., Chau, K.-W. and Sivakumar, B. (2015) Neural Network River Forecasting through Baseflow Separation and Binary-Coded Swarm Optimization. *Journal of Hydrology*, **529**, 1788-1797. <https://doi.org/10.1016/j.jhydrol.2015.08.008>
- [11] Wang, W.-C., *et al.* (2014) Assessment of River Water Quality Based on Theory of Variable Fuzzy Sets and Fuzzy Binary Comparison Method. *Water Resources Management*, **28**, 4183-4200. <https://doi.org/10.1007/s11269-014-0738-4>
- [12] United States Environmental Protection Agency (2016) Global Atmospheric Concentrations of Carbon Dioxide over Time. [https://www.epa.gov/sites/production/files/2016-08/ghg-concentrations\\_fig-1.csv](https://www.epa.gov/sites/production/files/2016-08/ghg-concentrations_fig-1.csv)
- [13] United States Environmental Protection Agency (2016) Global Atmospheric Concentrations of Methane over Time. [https://www.epa.gov/sites/production/files/2016-08/ghg-concentrations\\_fig-2.csv](https://www.epa.gov/sites/production/files/2016-08/ghg-concentrations_fig-2.csv)
- [14] United States Environmental Protection Agency (2016) Global Atmospheric Concentrations of Nitrous Oxide over Time. [https://www.epa.gov/sites/production/files/2016-08/ghg-concentrations\\_fig-3.csv](https://www.epa.gov/sites/production/files/2016-08/ghg-concentrations_fig-3.csv)
- [15] National Oceanic and Atmospheric Administration (2007) Epica Dome C Ice Core 800kyr Deuterium Data and Temperature Estimates. [ftp://ftp.ncdc.noaa.gov/pub/data/paleo/icecore/antarctica/epica\\_domec/edc3deuttemp2007.txt](ftp://ftp.ncdc.noaa.gov/pub/data/paleo/icecore/antarctica/epica_domec/edc3deuttemp2007.txt)
- [16] Wuebbles, D.J. (2009) Nitrous Oxide: No Laughing Matter. *Science*, **326**, 56-57. <https://doi.org/10.1126/science.1179571>

- 
- [17] Dickens, G.R. (2003) Review: A Methane Trigger for Rapid Warming? *Science*, **299**, 1017. <https://doi.org/10.1126/science.1080789>
  - [18] Hansen, J., Sato, M., Reudy, R., Lo, K., Lea, D.W. and Medina-Elizade, M. (2006) Global Temperature Change. *Proceedings of the National Academy of Sciences of the United States*, **103**, 14288-14293. <https://doi.org/10.1073/pnas.0606291103>
  - [19] Steig, E. (2004) How Do We Know that Recent CO<sub>2</sub> Increases Are Due to Human Activities? RealClimate.  
<http://www.realclimate.org/index.php/archives/2004/12/how-do-we-know-that-recent-cosub2sub-increases-are-due-to-human-activities-updated/>