Modelling Global Temperature Anomaly: An Application of Granger Causality Analysis and Partial Least Squares Regression

Sumbal Azam, Isabelle Hegarty[1]

[1] *School of Mathematics and Statistics, University College Dublin, Ireland*

ABSTRACT

In this study, we examine a subset of variables shown to drive climate change and attempt to quantify their influence on global mean temperature rise. The variables we study are the mean mole fraction of atmospheric CO_2, number of volcanic eruptions and total solar irradiance. Using the Granger causality test, we examine the usefulness of these factors for forecasting temperature rise. We use principal component analysis in order to extract the most important factors driving variability in these data and guide our selection of predictors for regression. Finally, using partial least squares, we construct a linear predictive model for temperature. While we were only able to find Granger causality between CO_2 and temperature at the 5% level of significance, our subsequent analysis suggests that a positive correlation exists for all factors. We determine that the mean mole fraction of CO_2 is the most important variable for predicting temperature. We observe that while volcanic eruptions and total solar irradiance do have a positive relationship with temperature anomaly, their comparatively weak influence make them inefficient terms to include in our model.

1. INTRODUCTION

Global warming is the long-term rise in the average temperature of the earth surface caused by an increase in the concentration of greenhouse gases in the atmosphere (Al-Ghussain 2018). The temperature of the planet surface has been relatively constant for thousands of years, but has risen progressively over the last 150 years. It is generally believed that this rise in global temperature is due to increased levels of atmospheric carbon dioxide, methane nitrous oxide, and is caused by human activities, in particular the burning of fossil fuels.

Global temperature can fluctuate due to natural causes such as solar radiation and volcanic eruptions. Times of peak solar radiation increase the level of sunlight energy reaching the earth's surface, correlating with a rise in atmospheric temperature. In contrast volcanic eruptions initially lower global temperatures as the gases and ash released into the atmosphere create gas clouds which block sunlight from reaching the earth. However, in the long term, the gases released by volcanoes can contribute to increased global temperatures due to the greenhouse effect.

The "greenhouse effect" describes how the greenhouse gases naturally present in the atmosphere absorb thermal radiation emitted by the planet and act as a blanket over the Earth's surface, resulting in higher temperatures (Houghton 2005).

When sunlight reaches our planet, some energy is reflected back into space. The majority of the sunlight energy enters through the atmosphere and is initially absorbed by the air, land and oceans, and subsequently released back into the atmosphere in the form of heat. The greenhouse gas molecules present in the atmosphere trap some of the heat to form an insulating and protective layer surrounding the planet. For millennia, the level of greenhouse gases in the atmosphere remained relatively constant, and global temperatures likewise fluctuated very little. However, the last century, has seen increased concentrations of greenhouse gases in the atmosphere, allowing more heat to be trapped, thereby causing the global temperature of the atmosphere to rise.

According to the IPCC (Intergovernmental Panel on Climate Change), each of the last four decades has been successively warmer than any decade that preceded it in record and the approximate total human-caused global temperature increase from 1850–1900 to 2010–2019 is 1.07°C (Zhai et al. 2021). This figure is significant as it is approaching a level of irreversible damage, beyond which humanity will experience the most destructive and dangerous effects of climate change.

The Paris Agreement, an international treaty on climate change, signed by 196 parties has set its goal to limit global warming to well below 2 degrees, and preferably below 1.5 degrees. Changes to the Earth's climate caused by increased temperatures are having monumental effects including rise in sea levels, melting sea ice and increased extreme and dangerous weather conditions (Pörtner et al. 2022). The detrimental effects of these changes in global climates for people's safety, lives and livelihoods make research into the causes of climate change of utmost importance.

Existing literature almost unanimously finds that human action causes climate change. Lynas et al. (2021) analysed the content of a random 3000 papers from the 88,125 climate-related peer-reviewed papers published since 2012, and found that more than 99% of them acknowledge the role of greenhouse gas emissions in climate change. Similarly, in their annual review the IPCC conclude that "it is unequ ivocal that human influence has warmed the atmosphere, ocean and land" (Zhai et al. 2021).

Many of the analyses in the global warming literature are focused on the causes of temperature rise and whether it is naturally occurring or human caused. Research has calculated the increase in carbon dioxide in the atmosphere from 280ppm to 385ppm since the industrial revolution (Oreskes & Conway 2008), with the majority of this occurring in recent years (National Research Council 2009; National Academy of Sciences & The Royal Society 2020). Many methods have been used to measure both correlation and causation between CO_2 levels and temperature. Barnhart & Eichinger (2011), for example, examine this using empirical mode decomposition and find that CO_2 increases during the last 50 years have had a significant positive effect on the long term net radiative forcings upon Earth. This gain of energy results in warming effect. Apart from CO_2, the authors also examine the contribution of solar irradiance and find that while it does produce short term radiative forcing, the effect is much smaller than that produced by CO_2. Stips et al. (2016) additionally investigate the causality link between volcanic eruptions and global temperature using concepts based on information flow (Liang 2008). The study finds the long term contribution of solar irradiance and volcanic eruptions to be insignificant compared to that made by anthropogenic CO_2 production which is the main force driving temperature rise.

We hope this paper will contribute to the global warming literature by studying the correlation and causation between CO_2 and temperature, taking into account potential correlation occurring from solar and volcanic activity. Based on existing literature, we expect to find a positive relationship between CO_2 and temperature, but plan to ascertain how much of this temperature change in recent years could be attributed to natural causes.

This paper is organised as follows, Section 1 gives a background for our analysis. Section 2 is an explanation of the data used, its advantages and limitations. Section 3 describes the methods used in this analysis and Section 4 is composed of the results of this analysis. We conclude with Section 5.

2. OBSERVATIONS

2.1. *Sample*

Our sample consisted of observations of CO_2 concentration (ppm), global temperature anomaly (w.r.t. 1951-1980), total solar irradiance (W/m^2) and number of volcanic eruptions sampled annually from 1959 to 2021. The full data set can be found in Appendix A along with a pairs plot to illustrate the relationship between the variables.

2.2. *Temperature*

In this study, we will use the NASA global annual-mean combined land-surface air an dse-surface water temperature data (Huang et al. 2017; Menne et al. 2018). This data is an estimate of global surface temperature, comprised of mean values from a large sample of locations evenly distributed globally.

There are some limitations associated with using mean global temperature data, a single number cannot accurately represent the science behind global warming. The climate is not governed by a single statistical measure of temperature but from temperature differences which drive environmental processed like storms, sea currents and thunder (Essex et al. 2007).

2.3. *CO_2*

We use the Mauna Loa CO_2 annual mean data (Tans 2000), which measure the atmospheric CO_2 from the NOAA GMCC program at the Mauna Loa Observatory in Hawaii. This represents the number of carbon dioxide molecules in a given number of molecules of air, after removal of water vapor. We have chosen this data for our CO_2 variable for several reasons. Firstly, the Mauna Loa Observatory is at an altitude of 3400 m, making it well situated to measure air masses that are representative of a very large area. Secondly, all of the measurements are rigorously and very frequently calibrated. Finally, ongoing comparisons of independent measurements at the same site allow an estimate of the accuracy, which is generally better than 0.2 ppm.

2.4. *Total Solar Irradiance*

We use the NOAA Climate Data Record of Total Solar Irradiance to measure solar activity (Coddington et al. 2016). This represents the total, spectrally integrated energy input to the top of the Earth's atmosphere, at a standard distance of one Astronomical Unit from the Sun, measured in W per m2. There are limitations to satellite-based Total Solar Irradiance data, as it is difficult to achieve absolute accuracy in results, although differences between records are lower than 1%. Furthermore, the improvements in the scientific understanding of Solar Irradiance over time could lead to some uncertainty in time series trends in solar activity.

2.5. *Volcanic Eruptions*

We obtained records of global volcanic eruptions from the Global Volcanism Program database (Venzke 2022) maintained by the Smithsonian Institution in the National Museum of History, Washington D.C. The archive contains global records of the last 10,000 years of Earth's volcanic activity. Only confirmed eruptions were considered.

3. METHODOLOGY

3.1. *Granger Causality*

Granger causality is a statistical hypothesis test which examines causality between two variables in a time series. It measures the ability of the prior values of one variable to predict the future values of another variable. Granger (1980) classifies causality based on two principles:

1. The cause happens prior to its effect
2. The cause has unique information about the future values of its effect

The Granger causality test aims to test the hypothesis for identifying a causal effect of X on Y such that

$$P[Y(t+1) \in A|I(t)] \neq P[Y(t+1) \in A|I_{-x}(t)], \quad (1)$$

where P is probability, A is an arbitrary non-empty set, $I(t)$ is the information available at time t and I_{-x} is the information available in the modified universe where X is excluded. If this hypothesis is accepted, we say that X Granger-causes Y.

The granger causality test must be performed on stationary data, if data is non-stationary it can be transformed through differentiation to eliminate the possibility of autocorrelation. There are limitations of the granger causality test, as it does not measure true causality but rather precedence. Tests could give misguided results if both x and y are being caused by a third variable, data has a non-optimal sampling frequency, there is a nonlinear causal relationship or there are rational expectations (Maziarz 2015).

We performed the granger causality test on our data to investigate the causal relationship temperature has with CO_2, total solar irradiance and volcanic eruptions. The null hypothesis for the test is that lagged x-values do not explain the variation in temperature. The alternative hypothesis for the test is that lagged x-values explain the variation in temperature.

3.2. *Principal Component Analysis*

Principal component analysis (PCA) is a method of multivariate analysis that reduces correlated variables into a set of uncorrelated variables called principal components. These components are orthogonal and represent linear combinations of the original predictors. PCA aims to extract the maximum possible information from the original data while reducing it to a lower dimensional form. This extraction is completely unsupervised by a target variable.

In practice, this is performed by standardising the data matrix $\mathbf{X}$ so that each variable has zero mean. The covariance matrix $\mathbf{S}$ is then computed and its eigensystem is extracted. The eigenvectors are sorted in order of decreasing eigenvalue and used to form the columns of a matrix $\mathbf{Z}$. The data is transformed by projecting it in the direction of these eigenvectors. This is simply the dot product of $\mathbf{X}$ and $\mathbf{Z}$. The transformed data is therefore just a linear combinations of the original data. (Abdi & Williams 2010)

The principal components are the normalized eigenvectors of the covariance matrix $\mathbf{S}$. The first component represents the direction of greatest variance in the data. Followed by the second, third and so on. (Ringnér 2008)

We applied this approach using a data matrix whose columns consisted of observations of total solar irradiance, mean mole fraction of CO_2 and total number of volcanic eruptions for the years 1959-2021. This was implemented with the Python package `scikit-learn`

(Pedregosa et al. 2011).

3.3. *Partial Least Squares Model*

Partial least squares (PLS) regression is an extension of multiple regression analysis that uses a linear combination of predictor variables in order to model a response. The original predictor variables are transformed in such a way as to account for as much variation in the dependent variables as possible. The result is a reduced number of predictors which makes PLS a dimension reduction approach. (Carrascal et al. 2009)

The model equation reads as follows,

$$\mathbf{Y} = \mathbf{B}\mathbf{X} + \mathbf{E}, \tag{2}$$

$\mathbf{Y}$ is the vector containing the response variable, $\mathbf{B}$ is a regression coefficient matrix, $\mathbf{X}$ is the predictor variable matrix and $\mathbf{E}$ is an error term.

The matrix $\mathbf{B}$ can be represented as a product,

$$\mathbf{B} = \mathbf{W}\mathbf{Q}, \tag{3}$$

where the columns of $\mathbf{W}$ are a set of weights computed for the matrix $\mathbf{X}$ such that the covariance between $\mathbf{Y}$ and $\mathbf{XW}$ are maximised. $\mathbf{Q}$ is formed by regressing $\mathbf{Y}$ on $\mathbf{XW}$ using ordinary least squares. (Hill et al. 2006)

In the context of this work, the year, total solar irradiance, mean mole fraction of CO_2 and number of volcanic eruptions were used as predictor variables. Temperature anomaly was the response. The Python package `scikit-learn` (Pedregosa et al. 2011) was used to implement this procedure. 10-fold cross-validation was used to find the optimal number of PLS components.

4. RESULTS

Here we present the results of applying each method to our data.

4.1. *Granger Causality Testing*

In order to perform our granger causality tests, we checked for stationarity in our variables using the Augmented Dickey Fuller Test. We found that they were not stationary, so we transformed them twice through differentiation, resulting in all our variables being stationary. We chose the optimal number of lags to include in our tests by analysing the Akaike Information Criterion for each lag. We then performed our granger tests on the stationary data with a lag of 1.

Table 1. Granger causality test results to measure if each variable causes temperature

Variable	F-Statistic (p-value)
CO_2	15.2409 (0.0003)
TSI	0.0644 (0.8006)
Eruptions	0.0088 (0.9255)

These results find that CO_2 causes temperature, but do not find that TSI and Eruptions cause temperature.

The results of the granger tests are given in Table 1.

Our first granger test investigates the hypothesis that CO_2 granger causes temperature. The F-test statistic is 14.3602 and the corresponding p-value is 0.0004. Since the p-value is less than 0.05, we reject the null hypothesis of the test and conclude that knowing the level of CO_2 is useful information that helps to predict future temperature.

Our second granger test investigates the hypothesis that TSI granger causes temperature. The F test statistic turns out to be 0.0027 and the corresponding p-value is 0.9588. Since the p-value is greater than 0.05, we fail to reject the null hypothesis of the test and cannot conclude that knowing the TSI is useful for predicting the temperature.

Our final granger test investigates the hypothesis that the number of eruptions granger causes temperature. The F test statistic turns out to be 0.0365 and the corresponding p-value is 0.8493. Since the p-value is greater than 0.05, we fail to reject the null hypothesis of the test and cannot conclude that knowing the number of eruptions is useful for predicting the temperature.

4.2. *Principal Component Analysis*

We obtained the following set of eigenvectors from the data covariance matrix.

$$e_1 = \begin{pmatrix} 0.516 \\ 0.756 \\ 0.404 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0.654 \\ 0.042 \\ 0.756 \end{pmatrix} \quad e_3 = \begin{pmatrix} 0.554 \\ 0.654 \\ 0.516 \end{pmatrix}$$

These vectors correspond to the first, second and third principal component respectively. The importance of each feature (total solar irradiance, CO_2 concentration, number of eruptions) for a given component is reflected

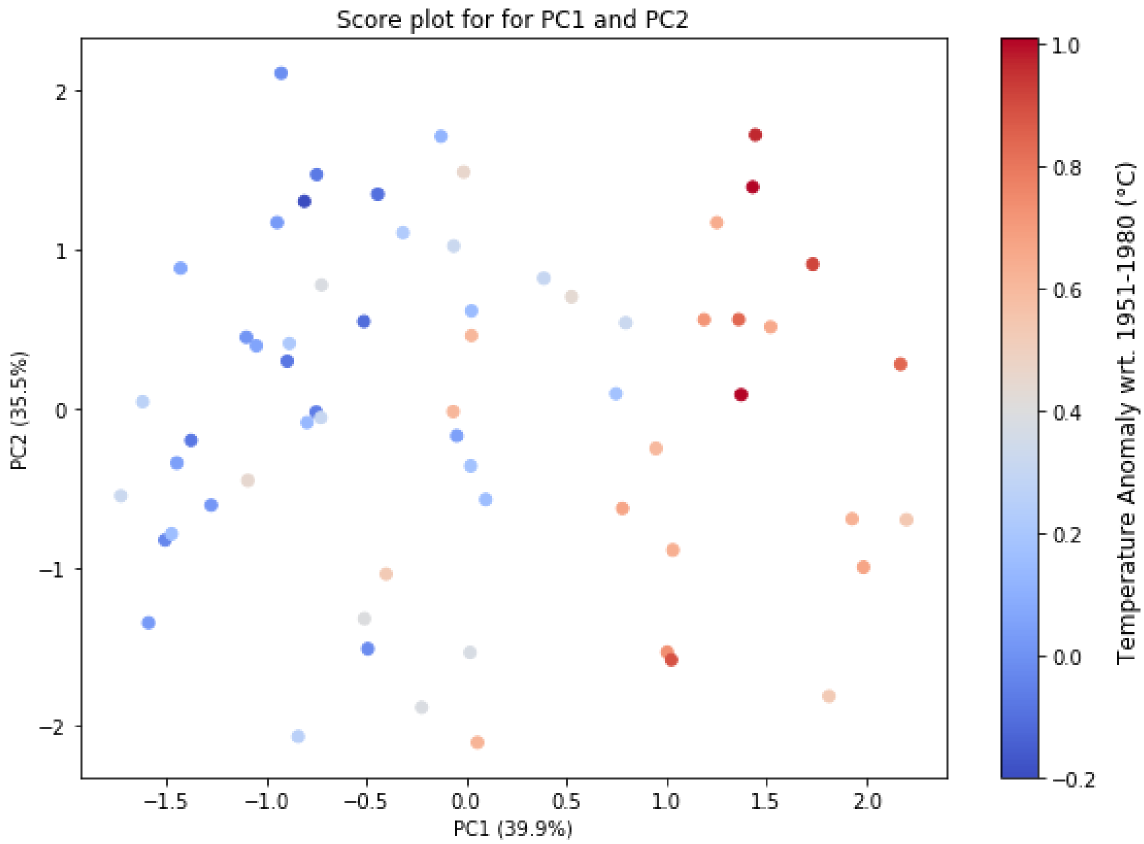

Figure 1. Score plot of PC1 and PC2, colour-coded according to temperature anomaly.

by the absolute value of corresponding entry in its eigenvector. The higher the absolute value of the score for a given feature, the higher its influence on the principal component.

We see that the first principal component (PC1) is dominated by the CO_2 concentration variable. PC2 is heavily influenced by total solar irradiance and eruptions and PC3 is mostly governed by CO_2 concentration, closely followed by total solar irradiance and eruptions.

The percentage of variance explained by each component is given in Table 2.

Table 2. Percentage of variance explained by principal components

PC	% Variance
1	39.9%
2	35.5%
3	24.6%

Note—Figures are shown rounded to one decimal place.

The first two principal components together account for just over 75% of the variance in the data. In Figure 1, we display a score plot of these two components, colour-coded by temperature anomaly.

It is seen that, the first principal component separates the data into two clusters in terms of temperature. Smaller values of PC1 are tend to occur simultaneously with smaller values of temperature anomaly while higher values of PC1 are associated with high temperature anomalies. Given that PC1 is dominated by the CO_2 concentration, this suggests a positive relationship between this variable and global temperature rise.

We do not see a similar clustering with PC2 however we observe that more extreme values of temperature anomaly tend to occur with higher values of PC2. As PC2 is governed mostly by eruptions and total solar irradiance, this indicates that high values of these variables are associated with more extreme temperature anomalies.

We conclude that all three variables provide useful information for predicting the global temperature anomaly, with CO_2 concentration being the variable of greatest influence. We expect the influence of the solar irradiance and eruption variables to be weaker but choose to include these in our partial least squares model (Section 3.3).

4.3. *Partial Least Squares Regression*

In Figure 2, we display a plot of the mean squared error obtained from 10-fold cross validation against the number of PLS components.

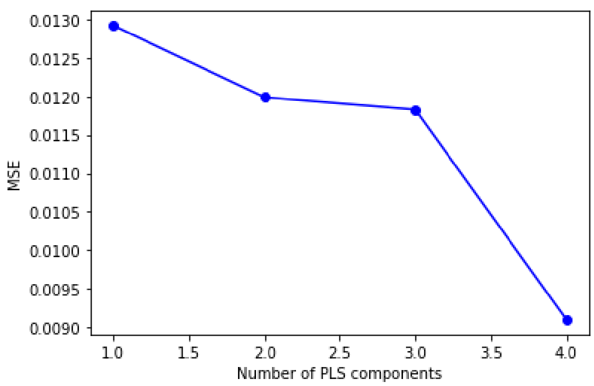

Figure 2. Mean squared error against number of PLS components.

As suggested by van der Voet (1994), we choose the smallest number of factors, above which the mean squared error does not improve significantly. This is the point where the 'statistical elbow' occurs. In Figure 2, this occurs at 2 components. Despite the reduction in the mean squared error, we do not consider 4 components as this is equivalent to performing ordinary least squares, which is unsuitable for this data due to the presence of multicollinearity.

Using 2 components, we obtained the following model equation,

$$\mathbf{Y} = \mathbf{X} \begin{pmatrix} 0.153 \\ 0.020 \\ 0.160 \\ 0.014 \end{pmatrix} + 0.350\,\mathbf{1}, \tag{4}$$

where $\mathbf{1}$ is an identity vector of length n and n is the number of rows of $\mathbf{X}$. Here, $\mathbf{X}$ is the normalised matrix of predictors and $\mathbf{Y}$ represents the temperature anomaly.

The coefficients are displayed in the following order of predictors; year, total solar irradiance, mean mole fraction of CO_2 and number of eruptions. Visibly, the the variables of greatest impact on the target are year and CO_2 concentration. The strong, positive relationship between the CO_2 variable and temperature anomaly is consistent with what we observed from our Granger analysis and principal component score plot (Figure 1).

As expected, the solar irradiance and number of eruptions do not exhibit the same magnitude of influence on the temperature anomaly. It therefore may not be cost effective to include these terms in our model as they require the collection of data which is not straightforward to obtain while having minimal influence on the target.

Furthermore, dropping these terms improves the usability of the model for forecasting future temperature anomalies. The total solar irradiance and number of eruptions for a given year are difficult variables to predict. Even with the use of existing models for forecasting volcanic eruptions (Poland & Anderson 2020) or estimating future solar irradiance values (Velasco Herrera et al. 2015), the cost of obtaining reasonable estimates for these predictors is difficult to justify given their influence on $\mathbf{Y}$.

In summary, while it is seen that solar activity and volcanic eruptions do play a role in predicting temperature, their influence is much weaker compared to the CO_2 variable which appears to be the main factor driving temperature rise.

We note that with the exclusion of the solar and volcanic terms, the model becomes even more limited in the precision it can achieve for future forecasts. The effect of a wider range of variables should be investigated if we are to build a more comprehensive model that is capable of predicting global temperature rise with minimal error.

5. CONCLUSION

This paper aimed to analyse the impact of several variables attributed to causing climate change on rising temperatures. Through Granger Causality tests we found that a significant causal relationship exists between CO_2 levels and temperature, but no causal relationship was found between TSI and temperature or number of eruptions and temperature. This implies that the level of CO_2 is a good indicator to help predict future temperature, but we have not found evidence to suggest that TSI or volcanic eruptions are useful to predict future temperatures.

Through principal component analysis we found that two principal components account for over 75% of the variance in the data. The first principal component is dominated by the CO_2 concentration and separates the data into two clusters, lower values of PC1 are associated with lower temperatures and higher values of PC1 are associated with higher temperatures, indicating a positive relationship between CO_2 concentration and temperature. The second principal component we found is dominated by TSI and eruptions and shows higher values of PC2 are associated with extreme temperatures, however we do not see the same clusters we found within the first principal component.

Finally, we modelled our data using a partial least squares regression and found a strong, positive relationship between CO_2 concentration and temperature anomalies. However, solar irradiance and number of eruptions do not show the same magnitude of influence on the temperature anomaly. Dropping these variables from our model allowed us to create a model to predict future temperature values.

Our results indicate that while the natural factors we considered, solar irradiance and volcanic eruptions do have a relationship with temperature, it is very weak compared to the strong correlation, causation and predictive power of our human activity variable, CO_2 concentration on temperature.

This agrees with much of existing literature which has commonly found that human activity influences global warming while taking into account the correlation occurring from solar and volcanic activity. We contribute to this line of research that very little of the temperature change in recent years can be attributed to natural causes, with the driving factor being CO_2 concentration. This insight will be useful for global climate change mitigation action to avoid the negative effects of climate change on people's safety, lives and livelihoods.

We believe that climate change is a complex and multi-source issue and would recommend further research to be taken to investigate the role of a multitude of other factors on global warming, including the role of other greenhouse gases, which were outside the scope of this paper.

6. SOFTWARE AND THIRD PARTY DATA REPOSITORY CITATIONS

We would like to acknowledge the use of the following additional Python packages in our analysis: `matplotlib` (Hunter 2007), `numpy` Harris et al. (2020), `pandas` (pandas development team 2020; Wes McKinney 2010), `statsmodels` (Seabold & Perktold 2010), `seaborn` (Waskom 2021) and `xarray` (Hoyer & Hamman 2017).

APPENDIX

A. DATA

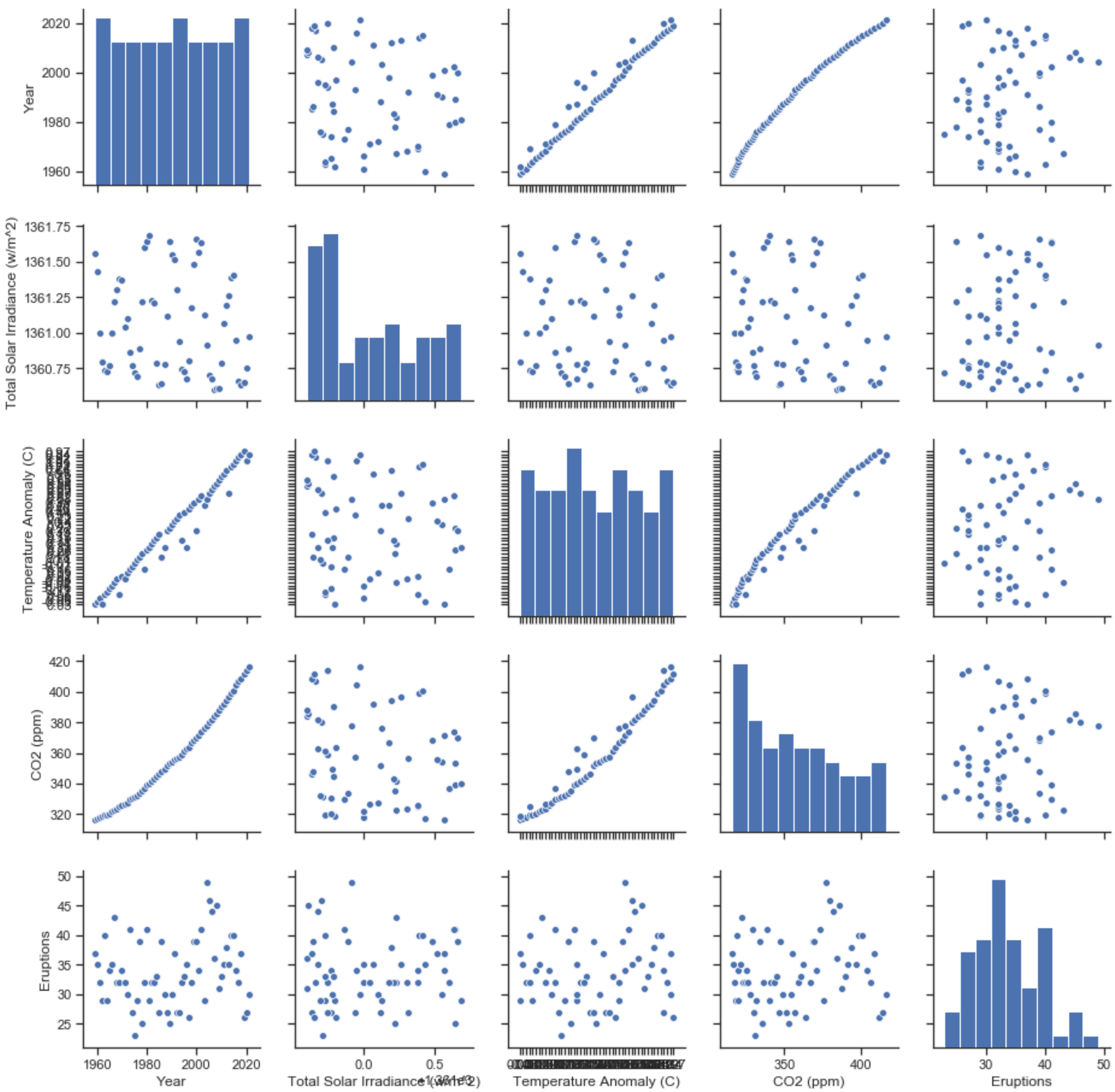

Figure A. Pairs plot of variables in Table 1.

Table 1. Data set used for analysis.

Year	TSI (W/m^2)	Temperature Anomaly (°C)	CO_2 (ppm)	No. of Eruptions
1959	1361.5613	0.03	315.98	37
1960	1361.4275	-0.03	316.91	35
1961	1360.9957	0.06	317.64	32
1962	1360.7981	0.03	318.45	29
1963	1360.7325	0.05	318.99	40
1964	1360.7277	-0.20	319.62	29
1965	1360.7740	-0.11	320.04	34
1966	1361.0021	-0.06	321.37	35
1967	1361.2230	-0.02	322.18	43
1968	1361.3037	-0.08	323.05	32
1969	1361.3771	0.05	324.62	32
1970	1361.3759	0.02	325.68	34
1971	1361.0400	-0.08	326.32	32
1972	1361.1002	0.01	327.46	30
1973	1360.8647	0.16	329.68	41
1974	1360.7710	-0.07	330.19	27
1975	1360.7155	-0.01	331.12	23
1976	1360.6946	-0.10	332.03	29
1977	1360.8905	0.18	333.84	39
1978	1361.2162	0.07	335.41	25
1979	1361.6016	0.16	336.84	32
1980	1361.6398	0.26	338.76	41
1981	1361.6833	0.32	340.12	29
1982	1361.2311	0.14	341.48	32
1983	1361.2137	0.31	343.15	32
1984	1360.7904	0.15	344.85	33
1985	1360.6329	0.12	346.35	27
1986	1360.6417	0.18	347.61	39
1987	1360.7828	0.32	349.31	30
1988	1361.1195	0.39	351.69	27
1989	1361.6433	0.27	353.20	25
1990	1361.5521	0.45	354.45	30
1991	1361.5178	0.40	355.70	37
1992	1361.3082	0.22	356.54	27
1993	1360.9427	0.23	357.21	27
1994	1360.7420	0.31	358.96	32
1995	1360.7317	0.44	360.97	33
1996	1360.6763	0.32	362.74	35
1997	1360.8073	0.46	363.88	26
1998	1361.1759	0.61	366.84	32
1999	1361.4800	0.38	368.54	39
2000	1361.6603	0.39	369.71	39
2001	1361.5686	0.53	371.32	34
2002	1361.6332	0.62	373.45	41
2003	1361.1283	0.61	375.98	29
2004	1360.9180	0.53	377.70	49
2005	1360.7004	0.67	379.98	46
2006	1360.6796	0.63	382.09	44
2007	1360.6039	0.66	384.02	36
2008	1360.6097	0.54	385.83	45
2009	1360.6066	0.65	387.64	31
2010	1360.7854	0.72	390.10	33
2011	1361.0690	0.60	391.85	35
2012	1361.1926	0.64	394.06	38
2013	1361.2631	0.67	396.74	35
2014	1361.3862	0.74	398.81	40
2015	1361.4098	0.89	401.01	40
2016	1360.9519	1.01	404.41	34
2017	1360.6597	0.92	406.76	32
2018	1360.6387	0.84	408.72	37
2019	1360.6492	0.97	411.66	26
2020	1360.7491	1.01	414.24	27
2021	1360.9749	0.84	416.45	30

REFERENCES

Abdi, H., & Williams, L. J. 2010, WIREs Computational Statistics, 2, 433, doi: https://doi.org/10.1002/wics.101

Al-Ghussain, L. 2018, Environmental Progress & Sustainable Energy, 38, 13, doi: https://doi.org/10.1002/ep.13041

Barnhart, B., & Eichinger, W. 2011, Journal of Atmospheric and Solar-Terrestrial Physics, 73, 1771, doi: https://doi.org/10.1016/j.jastp.2011.04.012

Carrascal, L. M., Galván, I., & Gordo, O. 2009, Oikos, 118, 681, doi: https:doi.org/10.1111/j.1600-0706.2008.16881.x

Coddington, O., Lean, J., Pilewskie, P., Snow, M., & Lindholm, D. 2016, Bulletin of the American Meteorological Society, 97, 1265, doi: https://doi.org/10.1175/BAMS-D-14-00265.1

Essex, C., McKitrick, R., & Andresen, B. 2007, doi: https://doi.org/10.1515/jnetdy.2007.001

Granger, C. W. 1980, Journal of Economic Dynamics and control, 2, 329, doi: https://doi.org/10.1016/0165-1889(80)90069-x

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357, doi: https://doi.org/10.1038/s41586-020-2649-2

Hill, T., Lewicki, P., & Lewicki, P. 2006, Statistics: methods and applications: a comprehensive reference for science, industry, and data mining (StatSoft, Inc.)

Houghton, J. 2005, Reports on progress in physics, 68, 1343, doi: https://doi.org/10.1088/0034-4885/68/6/r02

Hoyer, S., & Hamman, J. 2017, Journal of Open Research Software, 5, doi: https://doi.org/10.5334/jors.148

Huang, B., Thorne, P. W., Banzon, V. F., et al. 2017, Journal of Climate, 30, 8179, doi: https://doi.org/10.1175/jcli-d-16-0836.1

Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90, doi: https://doi.org/10.1109/MCSE.2007.55

Liang, X. S. 2008, Phys. Rev. E, 78, 031113, doi: 10.1103/PhysRevE.78.031113

Lynas, M., Houlton, B. Z., & Perry, S. 2021, Environmental Research Letters, 16, 114005

Maziarz, M. 2015, The journal of philosophical economics: Reflections on economic and social issues, 8, 86. https://hrcak.srce.hr/155919

Menne, M. J., Williams, C. N., Gleason, B. E., Rennie, J. J., & Lawrimore, J. H. 2018, Journal of Climate, 31, 9835, doi: https://doi.org/10.1175/jcli-d-18-0094.1

National Academy of Sciences & The Royal Society. 2020, Climate Change: Evidence and Causes (Washington, DC: The National Academies Press), doi: https://doi.org/10.17226/25733.

National Research Council. 2009, Restructuring Federal Climate Research to Meet the Challenges of Climate Change (Washington, DC: The National Academies Press), doi: https://doi.org/10.17226/12595

Oreskes, N., & Conway, E. M. 2008, Challenging knowledge: How climate science became a victim of the cold war (Contingency and Dissent in Science Project, Centre for Philosophy of Natural . . .)

pandas development team, T. 2020, pandas-dev/pandas: Pandas, latest, Zenodo, doi: https://doi.org/10.5281/zenodo.3509134

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825. http://arxiv.org/abs/1201.0490

Poland, M. P., & Anderson, K. R. 2020, Journal of Geophysical Research: Solid Earth, 125, e2018JB016974, doi: https://doi.org/10.1029/2018JB016974

Pörtner, H.-O., et al. 2022, IPCC 2022, Climate Change 2022. https://www.ipcc.ch/report/ar6/wg2/

Ringnér, M. 2008, Nature biotechnology, 26, 303, doi: https://doi.org/10.1038/nbt0308-303

Seabold, S., & Perktold, J. 2010, in 9th Python in Science Conference

Stips, A., Macias, D., Coughlan, C., Garcia-Gorriz, E., & Liang, X. S. 2016, Scientific reports, 6, 1, doi: https://doi.org/10.1038/srep21691

Tans, P. 2000, Mauna Loa CO2 Records, doi: https://doi.org/10.5860/choice.37sup-310

van der Voet, H. 1994, Chemometrics and intelligent laboratory systems, 25, 313

Velasco Herrera, V., Mendoza, B., & Velasco Herrera, G. 2015, New Astronomy, 34, 221, doi: https://doi.org/10.1016/j.newast.2014.07.009

Venzke, E. 2022, Global Volcanism Program, doi: https://doi.org/10.5479/si.GVP.VOTW4-2013

Waskom, M. L. 2021, Journal of Open Source Software, 6, 3021, doi: https://doi.org/10.21105/joss.03021

Wes McKinney. 2010, in Proceedings of the 9th Python in Science Conference, ed. Stéfan van der Walt & Jarrod Millman, 56 – 61, doi: 10.25080/Majora-92bf1922-00a

Zhai, P., et al. 2021, IPCC 2021, Climate Change 2021, doi: https://doi.org/0.1017/9781009157896