Correlation Analysis between Crop Yield and Greenhouse Gases

Tianxiang Cheng[1, †], Zhenhai Tu[2, †], Junjie Zhao[3, *, †], Yixuan Zhu[4, †]

[1] Suzhou foreign language school Jiangsu, China

[2] Long Island Lutheran high school Guangdong, China

[3] Beijing Royal School Beijing, China

[4]Beijing Eighty School Beijing, China

* Corresponding Author Email: zhaojunjie@st.brs.edu.cn

[†]These authors contributed equally.

Abstract. The relationship between the production of crops and emission of the greenhouse gases is rarely studied in the previous works, which deserves more attention. In this study, four types of greenhouse gases (i.e. carbon dioxide, carbon monoxide, ozone, and sulfur dioxide) and four types of crops (i.e. rice, wheat, maize, and soybean) are employed to find the correlation between them. To be more specific, this paper used least squares approximation to estimate the trend of greenhouse emission and crop production yield starting from 2000 to 2020 during the experiments. Finally, the experimental result indicated that carbon dioxide emissions increased by 100% when wheat production decreased by 130.774%. In addition, the emission of ozone also has a certain impact on the production of crops. It can be seen that the relationship between the production of crops and emission of the greenhouse gases can be established well and may benefit future agricultural production.

Keywords: Greenhouse gases; least squares approximation.

1. Introduction

Nowadays, with the progress of technology, humans are more focused on the scientific things, but ignore the nature. They belittle the importance of the gift the Earth gave us. They forget the rivers, the forests, and the gases. When it comes to the greenhouse gases, many people are not familiar with them. The only impression is the greenhouse gases warm the world, but not knowing how it influence our surroundings. People have to face the influences it brings. With the expansion of the country's infrastructure and the increase in population, the environment has been greatly polluted, and air quality has dropped severely. Also, it affects the crops harvest and the global temperature. The amount of greenhouse gases increases 1 per cent per year.

Although the emission of methane and nitrous oxide, decreased a bit, the emission of fluorinated gases increased by 3.8 per cent, and this trend would be continuous. The unstopped growth of the emission of greenhouse gases is worrying, but there are plausible reasons that the source of greenhouse gases cannot be removed. The sources are electricity, heat, transportation, fuel combustion, agriculture, and so on. Every element cannot be removed or reduced easily, and the simple removal can cause a huge negative impact of people’s daily life. According to the IPCC's Scientific Assessment of Climate Change, greenhouse gas emissions from the atmosphere have increased significantly, and it is expected that by the end of the 21st century [1], the concentration of carbon dioxide in the atmosphere will double from the current 350 PPM to 700 PPM, and the temperature will rise by 1.5-4.5 °C [2]. This greenhouse effect will have a serious impact on society, economy and other aspects, especially on agriculture. It is well known that experiments have shown that in environments with high carbon dioxide concentrations, plants grow faster and taller. But the result of "global warming" can affect atmospheric circulation, which in turn alters the global distribution of rainfall and the moisture content of soils on the surface of continents. The impact of "global warming" on the regional climate is not clearly understood, and the transformation of plant

ecology cannot be determined. At all, it is hard to determine whether the greenhouse gases have positive impact on agriculture or not.

During the past studies, greenhouse gases and agriculture seem two popular topics. There are many researches' studies about greenhouse gases and agriculture, but it is failed to find one research which combines the two parts. For example, Cole et al. reviewed the technologies that can reduce at the emission of the greenhouse e. g. carbon dioxide and nitrous oxide [3]. Duxbury et al. analyze the Impacts of agricultural land development and agricultural production practices on greenhouse gas emissions [4].

However, no research has ever made an experiment which control the variables and record the data. So, this study has to find the data in the same period in same regions, and combine them together. Also, there is no available greenhouses which can easily change the content of greenhouses gases, so it can be accepted why the experiment is never completed by the researchers. The experiment of this control variable is obviously unrealistic and impossible to complete at the present stage, because this experiment requires a greenhouse that can control the gas content, but this kind of greenhouse has not been studied at the present stage. It is worth mentioning that developing this kind of greenhouse is also a goal of our project.

Table I. Sample dataset-1.

Gas Type / Year	2000	2005	2010	2020
CO2 (million tons)	5 729.820	5 703.154	5 352.050	4 285.892
CO (thousand tons)	3.791667	2.336111	1.627778	1.797222
O3 (ppm)	0.082396	0.080487	0.073612	0.067516
SO2 (ppb)	78.572917	74.4635416	40.850000	10.340625

Table II. Sample dataset-2.

Crops Type / Year	2000	2005	2010	2020
Maize (Tons/hectare)	8.59	9.29	9.60	10.79
Rice (Tons/hectare)	4.898	5.185	5.239	5.477
Wheat (Tons/hectare)	2.82	2.83	3.12	3.35
Soybean (Tons/hectare)	2.562	2.861	2.923	3.379

To be more specific, this study mainly focusses on finding the relationship between the type of greenhouse e. g. carbon dioxide and ozone and the production yield of different types of crops e. g. rice and wheat. Least squares approximation as a typical method was employed in experiments to fit the collected dataset. Our project will be contributed in this field. First of all, through this experiment, this study can appeal to people to protect the environment, they can understand from the impact of greenhouse gas on agriculture. The environment affects every aspect of our daily lives. The greenhouse gases exhaled, the gases emitted by cars or the gases caused by combustion are all around us, affecting our daily life to a greater or lesser extent. They may reduce crop yields, may contribute to global warming, may contribute to sea level rise. This project can at least make people reflect on their daily energy conservation, and thinking about their production of greenhouse gases. Second, a lot of new research can also be carried out, such as smart planting, a greenhouse that can regulate the gas content can be designed, which will undoubtedly increase the yield of crops. Different plants have different demands for different gas concentrations, so it is possible to adjust the content of different gases in the greenhouse to increase the output of crops. For example, some plants need carbon dioxide, and more carbon dioxide will increase their productivity. Photosynthetic rate increases, transpiration coefficient decreases, dry matter accumulates, and larger and more fruits are produced.

2. Methods

2.1. Data description

In the project, we focus our research primarily on the United States. We choose the dataset in four types of greenhouse gases ranging from 2000 to 2020, including carbon dioxide, carbon monoxide, ozone, and sulfur dioxide, and four types of crops ranging from 2000 to 2020, including rice, wheat, maize, and soybean, in our research scope. The carbon dioxide data refers to the gross direct emissions from fuel combustion in the United States, which is measured in million tons, and the data is provided by the International Energy Agency. The carbon monoxide data measures the concentration of carbon monoxide throughout the United States, which is counted in thousand tons. This data is used by EPA, state, tribal and local agencies to ensure that carbon monoxide remains at levels that protect public health and the environment. Ozone data indicates the concentration of ozone throughout the United State, which is measured in ppm, part per million. Sulfur dioxide data shows the concentrations of sulfur dioxide throughout the United States, which is counted in ppb, part per billion. The data of rice, wheat, maize, and soybean are provided by OECD, Organization for Economic Co-operation and Development. All these four types of data indicate the Crop production in the United States, which are measured in tons per hectare. The sample data of our dataset can be found in Table I and Table II.

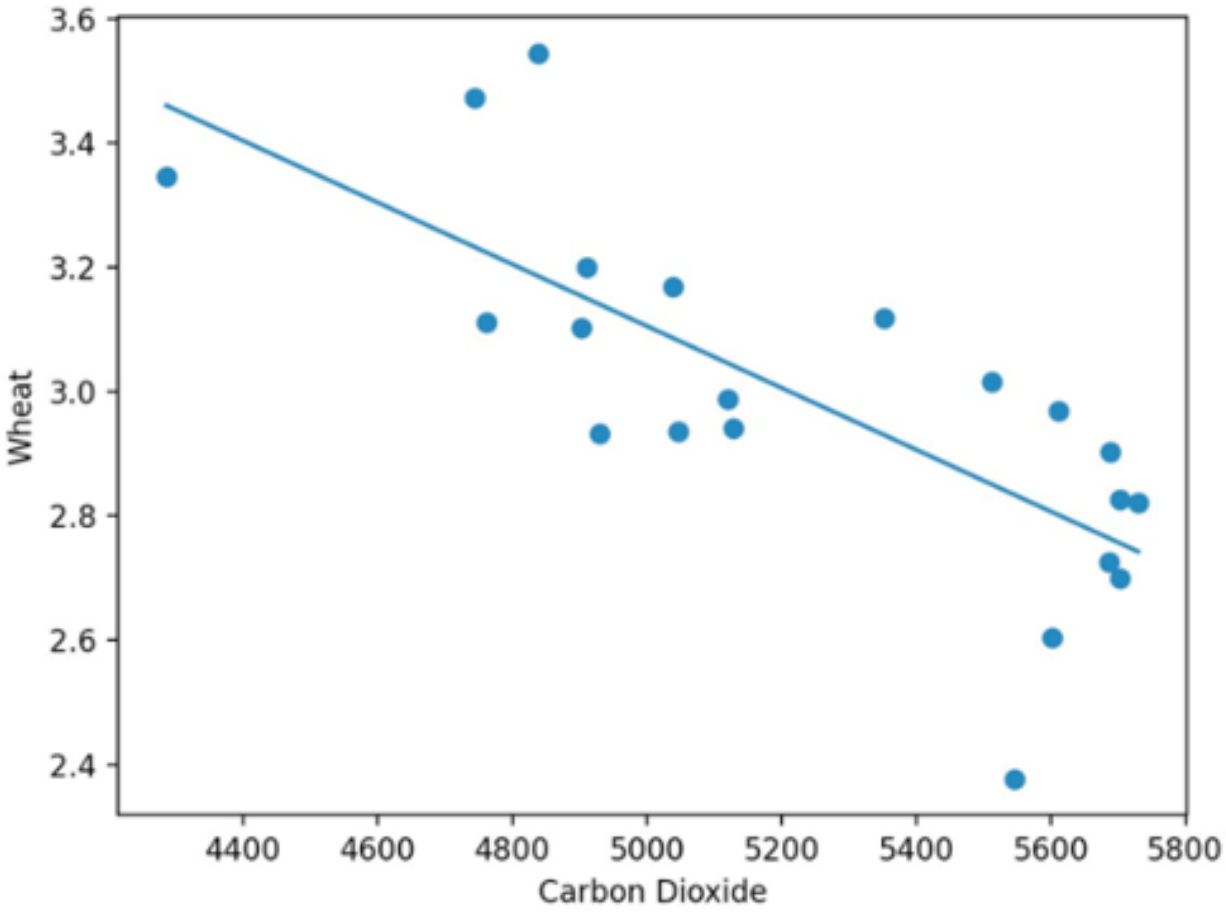

Fig 1. The relationship between the carbon dioxide and the wheat.

```
Model slope:  -0.000497119343718726
Model intercept: 5.591336169133342
The effect of 100% increase in Carbon Dioxide emission or concentration on Wheat yield: -130.74407619976952 %
```

Fig 2. The percentage of products that go down based on dioxide emission.

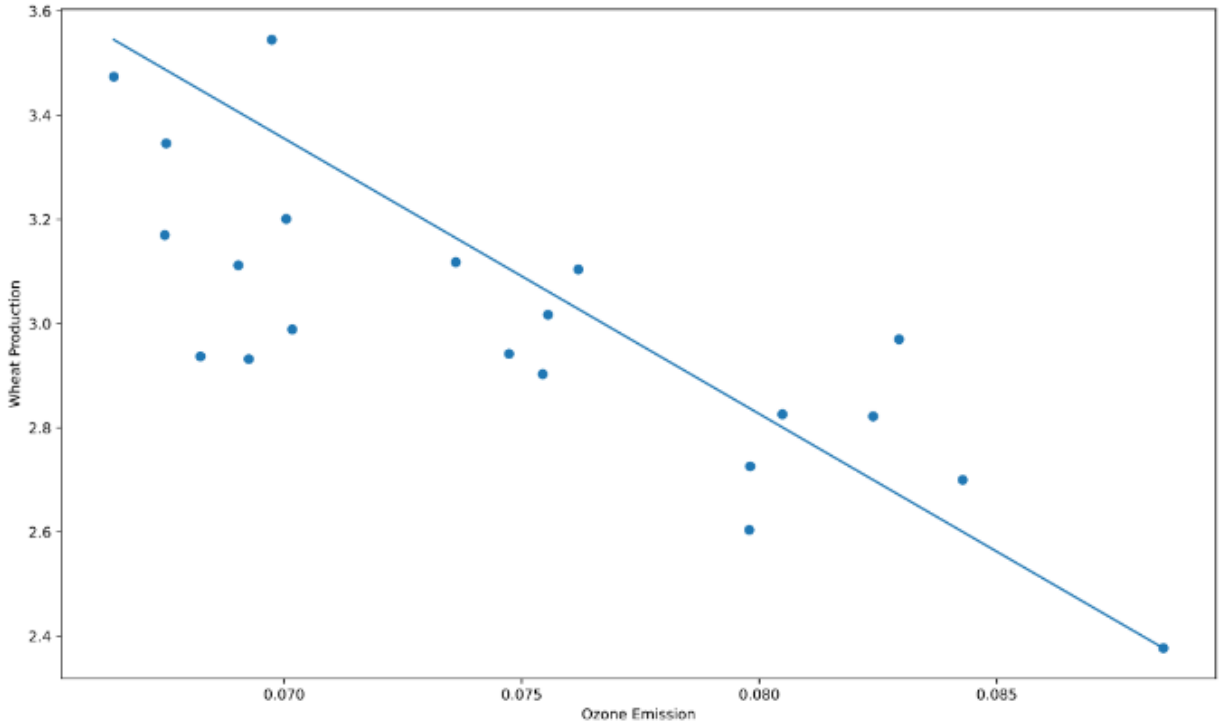

Fig 3. The relationship between the ozone emission and the wheat.

```
Model slope:  -33.97816199476392
Model intercept: 5.534113365822748
The effect of 100% increase in Ozone emission or concentration on Wheat yield: -132.80545971323902 %
```

Fig 4. The percentage of products that go down based on ozone emission.

2.2. Analysis methods

In the data analysis, we use least squares approximation to estimate the trend of greenhouse emission and crop production yield starting from 2000 to 2020 [5-7]. The method of least squares is a standard method in regression analysis to approximate the solution of systems by minimizing the sum of the squares of the residuals (the difference between an observed value, and the fitted value provided by a model) made in the results of each individual equation. In order to directly observe the relationship between the quantity of gas emission and the number of years, we use linear regression model to approximate a linear function with the specific slope and intercept.

First, we start with the carbon dioxide emission. With the help of Python [8, 9] and its packages, such as Matplotlib [10, 11], we could present our data on the scatter plot. However, in this stage, it is hard to find any trend between data points. Second, by importing scikit-learn [12, 13], a python package, we create a linear regression model and input the carbon dioxide data to train the model. Third, after we derive the estimated coefficient and intercept, we use Matplotlib to draw a smooth linear function on the plot to represent the trend of the carbon dioxide emission along with the year. Then, we repeat this process by using the data of other three types of greenhouse gases, carbon monoxide, ozone, and sulfur dioxide. Also, we process the crop production data among rice, wheat, maize, and soybean in the same way.

After finding the trend between each type of greenhouse gas and the number of year and the trend between the production of each type of crop and the number of the year, we try to find the relationship between the change in the quantity of greenhouse gases and the change in the quantity of crop production yield. First, we calculate the total change of each crop production yield from 2000 to 2020 by the crop production quantity in 2020 minus the crop production in 2000. Second, we use the same way to calculate the total change of each greenhouse gas emission from 2000 to 2020. Third, we use

the total change of crop production for one type of crop divide the total change of emission for each type of greenhouse gases to see which type of greenhouse gases will have a bigger impact on the production yield for each specific crop.

3. Result and discussion

Our study, with the help of Python, predicted the effects of different greenhouse gases on crop yields, and through the integration and analysis of the data, we found that the two gases that had the greatest impact on crop yields were carbon dioxide and ozone. Taking wheat as an example (because wheat data is the most intuitive), when carbon dioxide emissions increased by 100%, wheat production decreased by 130.774% (Fig. 1); When ozone emissions increased by 100%, wheat production decreased by 132.005 per cent (Fig. 2). We converted annual data for different greenhouse gases and data for different pancakes into tables (Fig. 3, Fig. 4). From our table, we can intuitively see that from 2000 to 2020, carbon dioxide and ozone in the United States are declining. In 2000, the United States had 57,298,200 tons of carbon dioxide, but by 2020, it had become 4,285,892 tons, a decrease of 14,439,280 tons. In 2000, the ozone in the United States was 0.082396 ppm, and by 2020, the ozone will become 0.067516 ppm, a decrease of 0.01488 ppm. Total U.S. wheat production has been increasing from 2000 to 2020: from 2.82 Tons/hectare at the beginning to 3.35 Tons/hectare. Therefore, we conclude that when carbon dioxide and ozone in the atmosphere decrease, crop yields increase.

Because there is not enough data (most of our data comes from government websites, and each greenhouse gas unit is different), and we ignore a lot of external factors that affect crop yields, such as rainfall, temperature, etc. So, our group's data analysis is relatively crude. Therefore, our study needs some scientists to do some relevant experiments to come up with more accurate and valuable data. Taking the experimental carbon dioxide concentration on crop yield as an example: the experimenter wants to grow the same crop in a greenhouse containing different concentrations of carbon dioxide, and observe the growth and yield changes while ensuring the health of the crop and controlling the constants (sunlight, humidity and temperature, etc.), and then record the specific data. There are three data to be recorded by the experimenters: 1, the carbon dioxide concentration of different greenhouses, 2, the growth time of the crop, and 3, the yield. Using this kind of experiment can make the data more accurate, resulting in better data and estimates.

4. Conclusion

This paper mainly talks about the impact of greenhouse gases on cereals' yields. The content includes the introduction of why this study want to do such researches, the information of the datasets that this paper used in the research, method and workflow throughout the data analysis, and the result of our research. The dataset contains the values of different types of greenhouse gases, including carbon dioxide, carbon monoxide, ozone, sulfur dioxide, and so, in the United States from 2000 to 2020. The dataset also includes the yield record of four types of cereals, wheat, rice, soybean, and maize, in the United States from 2000 to 2020. In the part of data analysis, this study uses the least square method and the linear regression to find the relationship between the amount of greenhouse gases in the atmosphere and the cereals production yields. As a result, the result finds that two types of greenhouse gases, carbon dioxide and ozone, have stronger effects on cereals production yield. When the centration of carbon dioxides or ozone in the atmosphere doubles, the cereals production yields will decrease more than twice. This outcome demonstrate that human cannot ignore the impact of greenhouse gases on agriculture and people should protect our ecosystem and reduce the air pollution as much as possible.

References

[1] The intergovernmental panel climate change. https://www.ipcc.ch/, 2022.

[2] China Meteorological Network - China Meteorological Administration government portal. (n.d.), April 27, 2022, http://www.cma.gov.cn/

[3] C. Cole, et al. "Global estimates of potential mitigation of greenhouse gas emissions by agriculture." Nutrient cycling in Agroecosystems, vol. 49, 1997, pp. 221-228.

[4] J. Duxbury, “The significance of agricultural sources of greenhouse gases.” Fertilizer research, vol. 38, 1994, pp. 151-163.

[5] P. Drineas, et al. “Faster least squares approximation. Numerische mathematik,” vol. 117, pp. 219-249, 2011.

[6] K. Gabriel “Least squares approximation of matrices by additive and multiplicative models.” Journal of the Royal Statistical Society: Series B (Methodological), vol. 40, pp. 186-196, 1978.

[7] K. Gabriel. "Least squares approximation of matrices by additive and multiplicative models." Journal of the Royal Statistical Society: Series B (Methodological), vol. 40, pp. 186-196, 1978.

[8] G. vanRossum, Guido. "Python reference manual." Department of Computer Science [CS] R 9525, 1995.

[9] R. Vallat, "Pingouin: statistics in Python." J. Open Source Softw. 3.31, 2018.

[10] D. Hunter. "Matplotlib: A 2D graphics environment." Computing in science & engineering 9.03, 2007, pp. 90-95.

[11] E. Bisong, "Matplotlib and Seaborn." Building machine learning and deep learning models on google cloud platform. Apress, Berkeley, CA, 2019, pp. 151-165.

[12] F. Pedregosa, et al. "Scikit-learn: Machine learning in Python." the Journal of machine Learning research, vol. 12, 2011, pp. 2825-2830.

[13] O. Kramer, "Scikit-learn." Machine learning for evolution strategies. Springer, Cham, 2016, pp. 45-53.