Climate change since 1750

Proposal Paper

Kaixin Chen
Department of Computer Science
University of Colorado Boulder
1111 Engineering Dr
Boulder, Colorado 80303
changshawuyanzu@gmail.com

Xinyi Zhang
Department of Computer Science
University of Colorado Boulder
1111 Engineering Dr
Boulder, Colorado 80303
xizh3366@colorado.edu

Yonghao Zhang
Department of Computer Science
University of Colorado Boulder
1111 Engineering Dr
Boulder, Colorado 80303
yozh2004@colorado.edu

ACM Reference format:
Kaixin Chen, Xinyi Zhang, and Yonghao Zhang. 2016. Climate change since 1750. In *Proceedings of ACM Conference, Washington, DC, USA, July 2017 (Conference'17),* 3 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

1 PROBLEM STATEMENT

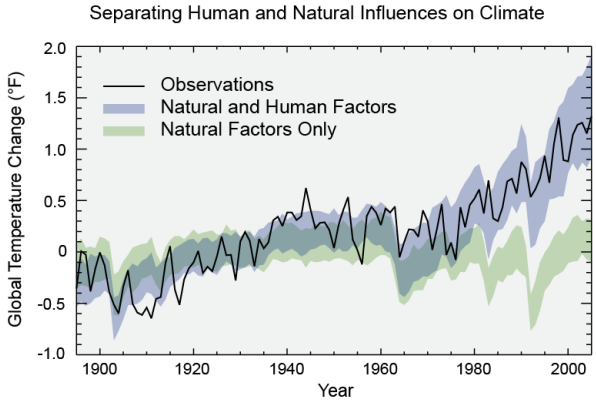
In this paper, we will examine the trend of climate change since 1750 in different area and country all over the world. We map the climate data with average temperature, country, latitude, and longitude while also concerning the uncertainties of each numerical data. In the analyze part, we will explore the relationship between temperature and latitude and discover the land average high and low trends of the climate. We would also concern the effect of melting glaciers and the development of industry to have a more conclusive ideal about the climate change. The results from this analysis could help those climate studies discovering the phenomenon of global warming and helping those people that are not familiar with the climate change get a better and straightforward understanding about the changes in climate these years.

2 LITERATURE SURVEY

There are many existing studies analyzing how global climate has changed, which includes analysis on temperature, human effects, and some natural factors. These studies also concluded how to solve or slow down the process of the climate change and what consequences the climate change might have.

2.1 EPA: United States Environmental Protection Agency

The report of EPA claimed that the primary cause of climate change is imbalance between energy entering and leaving the planet's system. As for we know, when incoming energy from the sun is absorbed by the Earth system, Earth warms and vice versa. The imbalance of energy is caused by the greenhouse effect, natural changes including changes within the sun and changes in Earth's orbit and changes in reflectivity. Greenhouse effect that causes the atmosphere to retain heat. Natural changes affect can affect the intensity of the sunlight that reaches Earth's surface that affects how much energy reaches Earth's system. Finally, changes in reflectivity also affect how much energy enters Earth's system because the amount of sunlight that reaches Earth can be reflected or absorbed that depends on Earth's surface and atmosphere. [1]



2.2 NASA: Climate Change and Global Warming

The NASA research about the climate change also concludes the global warming. The research conclusion of NASA could be regarded an authentic Wikipedia; in other words, the research includes evidence, causes, effects, and solutions. We might focus our attention on the Climate Resource Center, which including the analysis of the climate change through many aspects. These results are useful for us to prove the credibility of our project and could also help us conclude many other aspects we might need in the final conclusion part. [2]

3 PROPOSED WORK

We found a great data set from Kaggle as the raw data we tend to analysis and we divided the pre-processing work into three parts that containing data extraction, data cleaning, and data integration. [3] As this data set is kind of big for our test calculation at the beginning, we will extraction the first 1000 data points and store them as a temporary data set to make the coding part quicker and easier. And we also noticed that the land temperature data from 1750 to 1850 contains average temperature only instead of maximum temperature and minimum temperature, we will probably

Conference'17, Washington, DC, USA

DOI: 10.1145/nnnnnnn.nnnnnnn

get rid of them while performing some analysis. As the time when we do the actual calculation and plot the graphs, we will make a throughout check to find the errors in the whole data set. Data integration is one of the most important parts in our study. For some of our interesting questions, we need to compare the data between different latitude, different area, and different time period. We aim to answer the following questions in the course of the project:

(1) Is Global Warming a fact or fiction?
(2) Is temperature rising a global phenomenon or it only happens in the certain area?
(3) What is the relationship between the climate change and altitude?
(4) What is the relationship between the seasonality and climate trend?
(5) What is the relationship between the highs and lows trend and climate trend?

The most noticeable difference between our study and what we have found previously in the literature survey is our project based on the database we choose to use. We not only compare the numerical data but also combine what we find through the data with other natural and human effects that have not been concluded in the dataset we will use. Also, we will only compare our conclusion with the already known analysis instead of using them as the basic theory of our study.

4 DATA SET

We obtained the five global temperatures since 1750 datas from the Kaggle.com. [3] The raw data comes from the Berkeley Earth data page, which is affiliated with Lawrence Berkeley National Laboratory. These five data sets repackaged 1.6 billion temperature reports from 16 pre-existing archives. In this dataset, we have include several files:

- Date: starts in 1750 for average land temperature and 1850 for max and min land temperatures and global ocean and land temperatures
- LandAverageTemperature: global average land temperature in celsius
- LandAverageTemperatureUncertainty: the 95% confidence interval around the average
- LandMaxTemperature: global average maximum land temperature in celsius
- LandMaxTemperatureUncertainty: the 95% confidence interval around the maximum land temperature
- LandMinTemperature: global average minimum land temperature in celsius
- LandMinTemperatureUncertainty: the 95% confidence interval around the minimum land temperature
- LandAndOceanAverageTemperature: global average land and ocean temperature in celsius
- LandAndOceanAverageTemperatureUncertainty: the 95% confidence interval around the global average land and ocean temperature

Other files include:

- Global Average Land Temperature by Country (GlobalLandTemperaturesByCountry.csv)
- Global Average Land Temperature by State (GlobalLandTemperaturesByState.csv)
- Global Land Temperatures By Major City (GlobalLandTemperaturesByMajorCity.csv)
- Global Land Temperatures By City (GlobalLandTemperaturesByCity.csv)

5 EVALUATION METHODS

One of the most evaluation methods for our project is to check the credibility through comparing our results with the existing studies or conclusions. As the climate change is a relatively old and well-known topic that has been studied for many years, it would be not hard for us to find plenty of related studies. We will not perform our project base on these already-completed results absolutely; however, it is necessary for our team to compare the results as the data set we decided to use contains only temperature issue while the climate change is a complicated process. Moreover, many existing studies are pretty authentic as many famous results have been accepted world-widely for many years. Furthermore, checking whether our data is useful to examine the question we are interested in and could our data provide the answer to these questions are necessary evaluation methods. Even if we obtain many results from the data set, we still need to make sure we do the correct and proper analysis that could provide the enough information for us to draw the conclusions.

6 TOOLS

- **SQL**
  Helping host the preprocessed database.
- **Matlab**
  Tool we might use at beginning to analysis the data and plot the graph. Might switch to Python if the complete data is too large for Matlab to compute efficiently.
- **Weka**
  Containing tools for data pre-processing, classification, regression, clustering, association rules, and visualization.
- **Python**
  Another tool to process the data efficiently.
- **Excel**
  The application that stores the original data.
- **Latex**
  Using LaTeXto form the report.

7 MILESTONES

- **Data collection and extraction**
  First thing to do is extract data from an original source and store them.
- **Data pre-processing**
  Find the errors in data set and clean them. And then for some of our interesting question, we have to compare data between a different area in the certain area.
- **Process for derived data**
  We will only use the data points after 1850 because only data after 1850 has the minimum and maximum temperature for each country and each city.

- **Evaluate data and summary** Evaluate the credibility of the result and make conclusions.

8 SUMMARY OF PEER REVIEW SESSION

In this session, we put our focus especially on the project topic and what results we want to obtain from the project. We searched various topic and decided to choose climate as the final target as the database of climate is abundant and it is valuable topic deserved to be discussed. Furthermore, we divided the work and set which tool to get start with. Few concerns that we thought were important for this project were: As the climate change is an old topic, we need to make sure the problems we decided to solve are valuable and could be solved through analyzing the data. Also, more factors, including the human effects, should be included in our final conclusion.

REFERENCES

[1] United States Environmental Protection Agency. Climate Change Science. (????).
[2] NASA's Jet Propulsion Laboratory. 2017. Global Climate Change: Vital Signs of the Planet. (2017).
[3] Berkeley Earth Data Website. 2016. Climate Change: Earth Surface Temperature Data. (2016).

A THE BODY OF THE PAPER

A.1 Problem Statement

A.2 Literature survey

A.2.1 EPA: United States Environmental Protection Agency.

A.2.2 NASA: Climate Change and Global Warming.

A.3 Proposed Work

A.4 Data set

A.5 Evaluation Methods

A.6 Tools

A.7 Summary of peer review session

A.8 References