Climate Change Analysis and Datasets Derived Through Computer Vision

Raymond Dueñas

B.S Candidate, Department of Mathematics and Computer Science California State University Stanislaus, 1 University Circle, Turlock, CA 95382

Received May 26, 2023; accepted June 1, 2023

Abstract

Since the 1970s, Climate Science has come to produce significant contributions towards an understanding of climate change, and humanities understanding of this world that supports it. The development of climate data is essential to the ability of Climate Science to monitor the climate system, detect contributors to climate change, measure the impacts of climate variability, and support an improved understanding of the climate system. This work aims to contribute to this body of research through the development of a novel methodology for creating climate data sets from satellite generated image data. Utilizing Computer vision on image data this work intends to produce climate data sets that can be analyzed and visualized. The resulting analysis on the developed data sets will demonstrate the effectiveness of the computer vision model to generate usable numerical data form image data.

Introduction

This interdisciplinary work will operate in both the Climate Science and Computer Science domains. The value of the incite that Climate Science provides can be seen in the $37 billion investment in relevant research by the world bank in 2022 alone. These incites include better understanding of weather phenomena like El Niño, allowing for predictions of precipitation, heat waves tidal activities and more. Climate Science has become essential to agricultural production which is charged with meeting the demand of societies' needs worldwide and is essential to the survival of humanity. Climate change effects the ability of agricultural production meet its societies demands. Changing seasonal trends affect when the crop should be planted and when they should be harvested. Changes in lengths of seasons have an impact on the ability of crops to mature. Shorter and lighter rainy seasons affect the available water supplies, and replenishment of water tables vital to the success of agricultural production. The effect of climate change on agricultural production is discussed in the work of Naveen Kumar Arora (2019). The focus of the work by is the projection of climate change over the next twenty-five years and its impact on the ability of agricultural productivity to supply a growing population worldwide. Another work by Tapan B. Pathak, et al. (2018) utilizes Climate Science to describes the effect of climate change on California ability to maintain its agricultural productivity in the future. Essential to these and other works that utilize climate science is climate data.

Computer vision is the concept of teaching a computer to gather and interpret usable information from image data. It is widely used as a tool across many disciplines. In the medical industry computer vision models are trained on large image data sets so that they may be used in the identification of abnormalities that may be missed by a fatigued human eye. Computer vision is also used in self driving cars, military and civilian drones. It also finds daily implementations every time face recognition is used to unlock a smart phone or googles image search tool is used, computer vision is at work. This work aims to use computer vision on image data in order to develop numerical datasets representing climate data.

Methods

This work utilizes satellite generated image data sets from NOAA, NASA, and Organ State University which visualize temperatures, precipitation, and snow cover as global maps examples from the datasets can be seen in Figure 1.0. The generation of a data set in a manner aligned with the goals of this work requires the development of a computer vision model that can parse information found in satellite generated image data. We first aim to develop a mode that can successfully operate on a subsection of the Western United States. After developing the ability to parse information as described on a subsection of the Western United States we need a way to take, then generate representative numerical data. To do this we measure the area in pixels that are within a targeted measurement range for each type of climate map. With respect to temperature the targeted measurement ranges include first temperatures at or below thirty-two degrees Fahrenheit, then greater than or equal to ninety degrees Fahrenheit. With respect to precipitation, we have limited observation

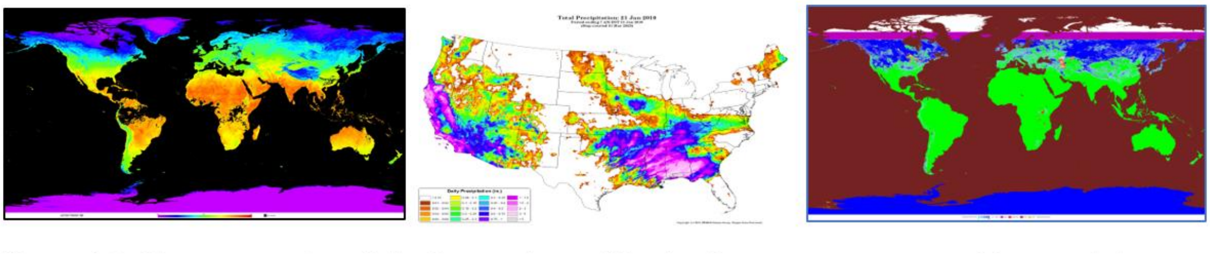

Fig. 1.0: Three Examples of the image data utilized as inputs to computer vision model.

range to greater than or equal to one half an inch to less than five inches. For snow coverage we target areas that have a measurement of one hundred percent snow coverage. Once the area in pixels within the targeted measurement range has been calculated we then use a determined pixel to mile ration to find the corresponding area in miles. This process is completed on every image provided to the model and the area in both pixels and miles is recorded into a data table. After operating the computer vision model on all three data sets, totalling more than six thousand images, we proceed to the analysis phase in which we analyse and visualize the data in order to demonstrate the effectiveness of the computer vision model at generating usable numerical data from image data.

Results

The resulting data sets from operating the developed model on over six thousand images have been calculated and visualized, some of the results from these operations are included below. In Figure 1.1 and Figure 1.2 we have graphed the generated data for temperatures, at or below thirty-two degrees Fahrenheit and at or above ninety degrees Fahrenheit, by day. We do this in order to get an initial look at the behaviour of the data produced by the model. Looking at the data in this way motivated the search for the greatest peak values on the graph of Figure 1.1. We observed that the date with the largest recorded area was found in January of 2020 with a value of 265,300 square miles. We compared this to the largest value recorded in the first year of our data 2015 which was 264,291 square mile and found that the January, 2020 reading was more than 1,000 square miles greater than the largest reading in 2015.

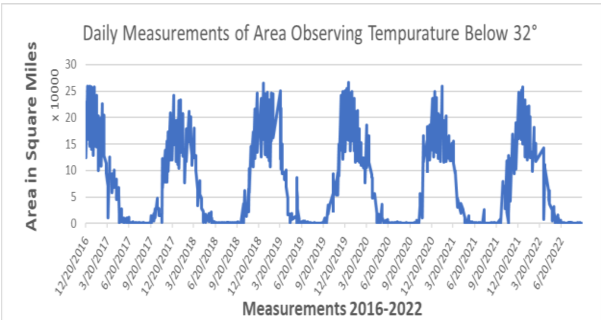

Fig. 1.1: Daily value of area where temperatures measurements were below 32°.

In Figure 1.1 and Figure 1.2 we see a repetition of peaks and valleys. Which prompted the question, how do the values in the graph compare to one

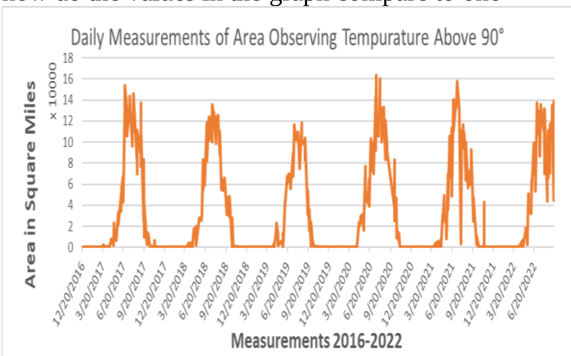

Fig. 1.2: Daily value of area where temperatures measurements were above 90°.

another in relation to their peaks and valleys. This question is answered by Figure 1.3. Figure 1.3 is a clipping from the larger graph which expands to the entirety of the observation period. Looking at the graph we see that the clusters of peaks for each considered temperature occur during the valleys of the other. This graph also brought to our attention that the amplitude of the peaks for each considered temperature are noticeably different.   With temperatures at or below 32° achieving nearly twice the amplitude in peak value. Which means that at their peak values temperatures at or below 32° covered much more area than that of temperatures above 90°.

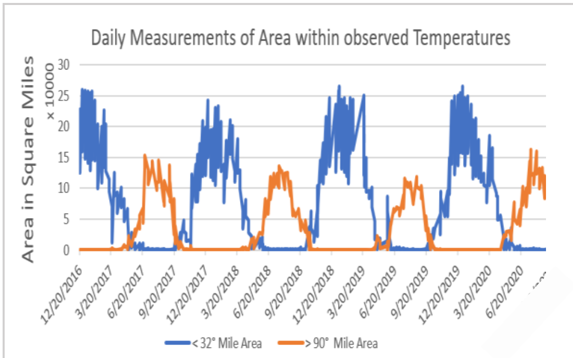

Fig. 1.3: Daily value of area where temperatures in observation range.

Year	Temperature $\leq 32^{\circ}$ F	Temperature $\geq 90^{\circ}$ F
2015	7,1936	3,0174
2016	7,2188	2,9035
2017	7,2066	3,1755
2018	6,8179	3,1957
2019	8,3385	2,4537
2020	6,8640	3,4701
2021	6,5212	3,3506
2022	6,5783	4,0459

Table. 1.0: Yearly average area with temperatures in observations range.

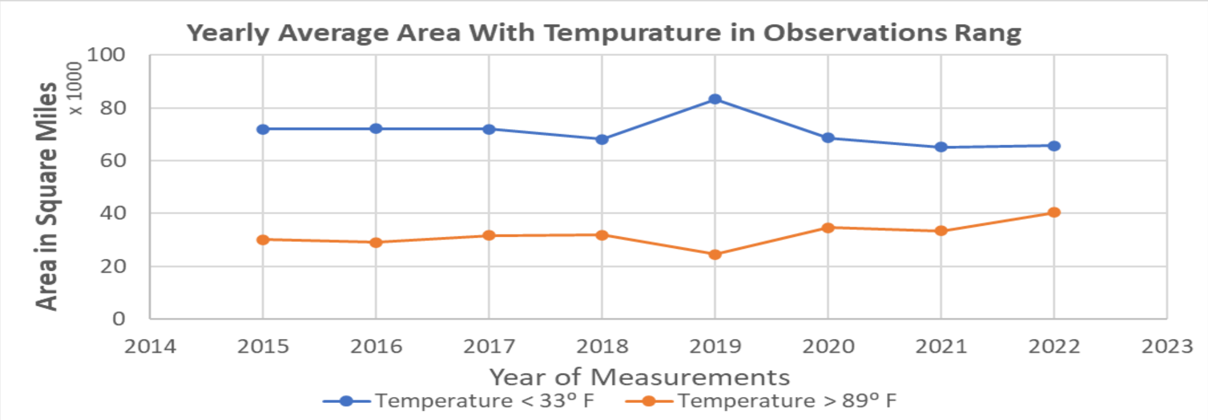

Fig. 1.4: Yearly average measurements of area with temperatures in observation range.

With a desire to see the behaviour of the data over an annual bases, the average observed measurement for each considered temperature have been calculated for each year in the data set. The results of these calculations can be seen in Table 1.0. The graph of the calculations can be found in Figure 1.4. In Figure 1.4 we can see that, during the observed period, there is a decrease in average area with the temperatures at or below 32° and an increase in average area with temperatures at or above 90°. While this is an interesting observation of the data it is important to note that the time period utilized is too small a sample size to generate results to be used as a foundation to make any claims on climate trajectory.

The analysis on the temperature data has served to demonstrate the capabilities of the model to successfully take image data and generate relevant datasets. The model was also successful in generating datasets for both snow coverage and precipitation, however because of the successful demonstration of the model work on temperature images, minimal analysis was conducted on data generated from the snow coverage and precipitation images. Graphs of the model generated data for snow coverage which can be found in Figure 1.5 and precipitation which can be seen in Figure 1.6.

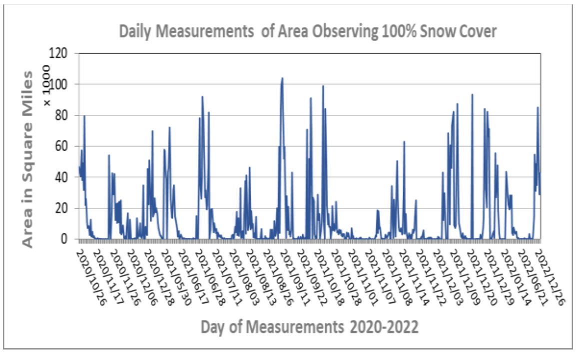

Fig. 1.5: Daily Measurements of area observing 100% snow coverage.

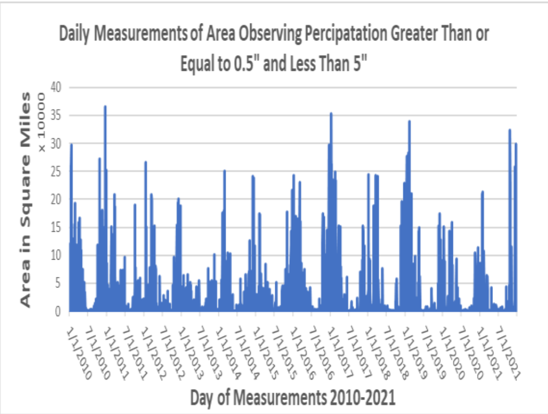

Fig. 1.6: Daily Measurements of area observing greater than or equal to 0.5” or less than 5”.

Conclusion

The goal of this work was to develop a computer vision model that could take in climate image data as an input and provide representative numerical data as an output to further support meteorologists. While this work chose to operate the developed model in the domain of climate science; however, it is important to note that the applications of this model are not limited to climate science. The model could very easily operate in other domains for example, the model could be implemented in bacteriology. Providing the model with images of bacteria, the effectiveness of different antibiotics could be compared by calculating the area of expansion or contraction of the bacteria on a petri dish over time. This would then relieve the bacteriologist from having to manually measure the bacteria on each dish, allowing for mass calculation and analysis in minutes compared to hours. Future work will include tunning the model to operate equally as successful at a global scale. Additional work could include adding a machine learning component to the model that allows the program to automatically identify an image's colour key and reliably couple pixel colour values to targeted image characteristic values. In closing as demonstrates by the analysis above this research has produced a computer vision model that can successfully utilize image data to generate numerical data.

References

L Pathak, Tapan B., et al. "Climate change trends and impacts on California agriculture: a detailed review." *Agronomy* 8.3 (2018): 25.

Copyright ©2023, PRISM Climate Group, Oregon State University, https://prism.oregonstate.edu

Map created 2023.Wan, Z., Hook, S., Hulley, G. (2015). *MOD11C2 MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 0.05Deg CMG V006* [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2023-05-04 from https://doi.org/10.5067/MODIS/MOD11C2.006

Hall, D. K. and G. A. Riggs. (2021). MODIS/Terra Snow Cover 8-Day L3 Global 0.05Deg CMG, Version 61 [Data Set]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MOD10C2.061. Date Accessed 05-04-2023.