The Impact of Factors Influenced by Climate Change on the Intensity of Forest Fires

Jessica Zhai, Fatimeh Hassan, Will Assad, Raghav Banka

December 14, 2020

Problem Description and Research Question

Our research question is: **How do factors influenced by climate change (temperature, humidity, wind) have an impact on the intensity of forest fires?**

Wildfires are defined as unplanned and uncontrolled fires in a naturally combustible area, such as a forest. Environmental and weather conditions can cause fires to spread quickly and last for prolonged periods of time (Berwyn 2018). Forest fires have several negative effects on the environment; not only do they destroy ecosystems and natural habitats, but they also result in extensive economical damage. The changes in air quality caused by forest fires can also greatly impact public health (Untamed Science 2020). The greenhouse gases released by forest fires can further contribute to climate change, which ultimately results in more fires. The environmental, economic, and health-related effects of forest fires can be very damaging in areas where fires are consistently occurring (Berwyn 2018).

We chose our research question because we want to understand which factors have the greatest impact on forest fires, as they have become an increasing threat in areas across the world. As we've seen in 2020, there has been a clear growth in the intensity and frequency of wildfires. Forest fires in the United States and Australia have indicated that the wildfire risk is increasing as a result of changes in the climate. Warmer areas with dry conditions are extremely prone to the threat of forest fires, and climate change has exacerbated this (Berwyn 2018). Therefore, one component of our program is to estimate the future temperature change of an area, and then predict the frequency and damage area of future forest fires of that area based on that estimation.

Climate change influences various factors such as temperature, precipitation, wind, and humidity over time. Exploring relationships between these factors and forest fires could help us better predict what causes the occurrence of forest fires. Making a connection between climate change and forest fires would also better our understanding on how to prevent and control these fires in the future.

Datasets

First Dataset: https://archive.ics.uci.edu/ml/machine-learning-databases/forest-fires/
Format of Dataset: csv file
Name of Dataset: forestfires.csv

The dataset contains information about several forest fires that occurred in Montesinho Natural Park in Portugal from 2000 to 2003. We're using this data to create a connection between forest fires and the factors that influence them.

The first and second column of the dataset are coordinates for a 9 by 9 map of the Montesinho Natural Park in Portugal. We will not be using these columns in our computations. Starting from the third column, the data displays the month, day, FFMC (Fine Fuel Moisture Code), DMC (Duff Moisture Code), DC (Drought Code), ISI (Initial Spread Index), temperature, relative humidity, wind, rain, and area of damage (which we neglected in computations).

The Fine Fuel Moisture Code (FFMC) represents fuel moisture of forest litter fuels under the shade of a forest canopy ("FW System"). It can be considered as an indicator of relative ease of ignition and flammability of fine fuel. DMC represents fuel moisture of decomposed organic material underneath the litter. It gives us insight to the live moisture stress. DC is the rating of the average moisture content of deep compact organic layers. Drought code deeply affects the intensity, severity, extent, and frequency of forest fires. ISI can be considered as the measure of how quickly a fire is expected to spread. It is based on FFMC and the wind speed.

The other factors are closely related to forest fire intensity. If an area receives no rain for a long period of time, the vegetation in that area dries out making it ideal for a forest fire. We should look at the number of millimeters of rain received depending on which we can predict which area is most prone to forest fire. Rain affects the occurrence and intensity of forest fires, as they would provide moisture and increase the humidity of the environment. Further, areas with greater temperatures are more prone to forest fires because temperature aids in the ignition of forest fuels. We also consider that higher relative humidity would decrease the chance of forest fires since it is difficult to start fires under a high humidity environment. Wind is another key factor affecting forest fires, aiding in the rate at which fires spread, hence increasing the ISI.

Second Dataset: https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data
Format of Dataset: csv file
Name of Dataset: portugaltemperatures.csv

This dataset contains information of the monthly average temperature of several big cities in Portugal from January 1753 to September 2013. We are using the data for the city Braga only, since it is the closest city to Montesinho Nation Park and will provide more accurate temperature data for that area. Starting from the first column, the data displays date-time information in format year-month-day. Then, the average temperature in Celsius, percent uncertainty, name of the city, name of the country (just Portugal), latitude, and longitude. We consider average regional temperature over time to be important data as this gives insight into what a future increase in temperature will mean for forest fires and the other variables that are temperature-dependent.

Computational Overview

A description of the modules, functions and computations in our program.

First, we created a module (entities.py) to organize our data from the data files. This module contains all of the class definitions as well the functions to read and transform the data. For example, there is a data class for the Portugal temperatures data set (PortugalTemperatureData), which represents a row of temperature data and contains the time, city, average temperature and uncertainty values which can be found in the data set as well. There is also a function that processes the csv file into a list of PortgualTemperatureData for easy access to data in computations. We also use a function in this module to process the forest fire data into a dictionary that contains a list of the values in each column mapped to the name of the column. This is more efficient than using a dataclass with the third-party module plotly. Overall, the purpose of the entities.py module is to help us contain the data in an organized way, make it more readable, and provide easy access to data in computational functions.

The models.py module contains a class to model the factors that we're exploring in various ways. We organize this using a "model" object that models data in a specific location. It includes methods to predict future changes, perform single and double regressions, graph, and animate the data. One of the methods will produce a 3D scatter plot given two independent variables and one dependent variable from the factors that influence forest fires. This is very helpful in determining relationships between the variables and displaying the data in an interactive way. Another important method plots the predicted change of DC (drought code) over the years as the temperature changes. For our analysis, it is necessary to develop a connection between DC and temperature as DC greatly affects the intensity of forest fires and increasing temperatures are a result of climate change.

We also use a method to calculate a dependent variable using the formula of a double regression trend when given two independent variables and regression parameters. Similarly, there is a method that takes two independent variables and one dependent variable from the factors that affect forest fires to model the change on the dependent variable as a result of the changes in the two independent variables. This method returns a tuple containing the constant value and the two coefficient values of the independent variables using the external library statsmodels.

Finally, our main.py module is an interactive file that calls various class methods in models.py to express the results of computations on the data visually. The main Python console will provide information and instructions on how to see the visualizations of the computations we performed. It will display an animation showing the temperature over the years in Braga, which is the city closest to the national park where the forest fires occurred. This shows a gradual increase in temperature and as an interactive feature, the user has the ability to change the year to see the average temperature of that year in Braga. A graph for the relationship between DC (drought code) and temperature will also be shown. Then the next graph will show a prediction of how DC will change over the years as the temperature does as well, to clearly develop a connection between these two factors. We used plotly in greater detail to visualize the relationships between multiple variables, for example, the double regression on ISI (how quickly a fire spreads). There is also a 3D scatter plot that can be opened and the user may manipulate the view of the graph however they would like, which is another interactive feature. After all of these visualizations are shown, the user has the ability to enter two variables they would like to see the relationship between, and a graph will be displayed to show the relationship. The user may do this as many times as they wish and the interactive portion will end when they choose to stop.

Instructions for Running the Program

1. Download the zip file containing all the Python modules, requirements and functions. We were able to upload the data sets to MarkUs so ignore steps 2 - 4 unless the data sets need to be re-downloaded.

2. Download the forest fires data set (forestfires.csv) using this link: https://archive.ics.uci.edu/ml/machine-learning-databases/forest-fires/.

3. Download the portugal temperatures data set (portugaltemperatures.csv) using this link: https://github.com/willassad/forestfires/tree/master/data.

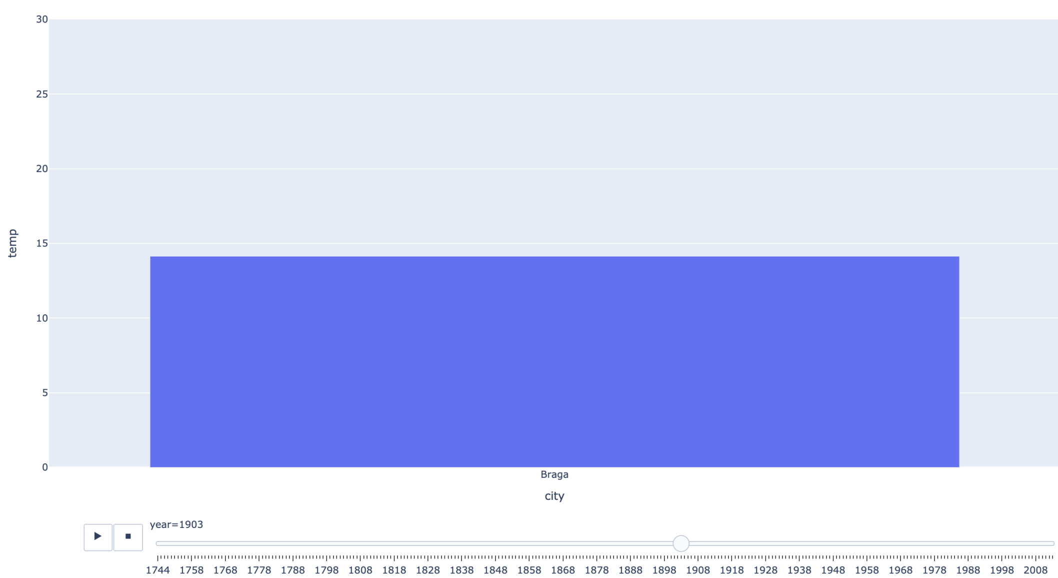
4. Save the datasets to the data folder provided in the zip file.

5. Install the required libraries listed in the requirements.txt file.

6. Run the main.py file which will provide an interactive overview of our functions and program.

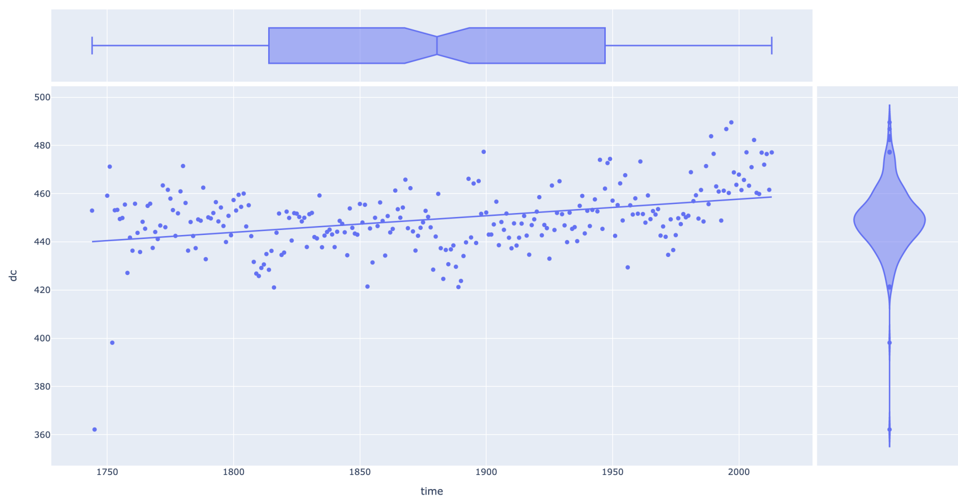
7. The instructions to see the graphs and functions can be found in the main Python console, which is interactive to the user.

8. First, an animation showing the temperature in Braga can be seen. As an interactive feature, the year can be changed using the sliding bar corresponding to years at the bottom. The following image is what it should look like.

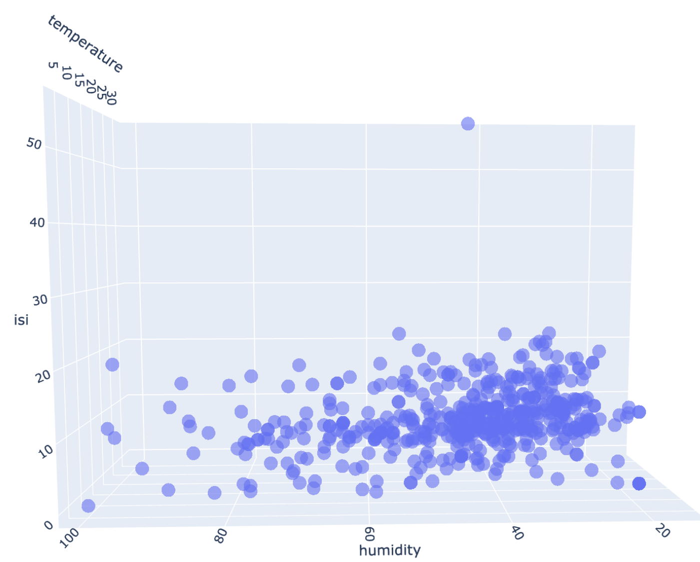


9. Next, the user may return to the interactive Python console and run the next few graphs, the first one will show the relationship between temperature and DC (drought code).

10. The next graph will show how DC will change over the years as temperature does as well. The following image is what it should look like.



11. After this, there is a double regression done on ISI (how quickly a forest fire spreads) with humidity and temperature, and a 3D scatter plot can be opened. The user can manipulate the view of the graph however they would like as an interactive feature. The following image is an example of what it could look like.



12. Finally, the user may choose two variables they would like to see the relationship between and type them in the console as prompted by the instructions. The user may do this as many times as they wish, with any of the variables, and then the interactive portion will end when they choose to stop.

Changes to Project

We're using another dataset along with the one we originally introduced in the proposal. This dataset contains information about temperature in the same location as the first dataset to further illustrate the connections between temperature and forest fires. Temperature is an important factor to explore because climate change has largely impacted it and this can help us develop a connection between the factors influenced by climate change and forest fires. In order to add more complexity to the code, we decided to include methods that perform double regressions on the data and produce scatter plots to display the results. This will also help us better understand the connections between different variables that may be related. We also added methods to animate the temperature data in order to make the changes more clear to the user. In our original proposal, we said we would use the Python library scikit, however we did not end up using it. Instead, we used the plotly library in a greater extent than we have done in class. We used the additional libraries pandas and statsmodels to do this. We also made minor changes to our research question and defined the intensity of forest fires using the factors of DC (Drought Code) and ISI (Initial Spread Index), instead of frequency and area of damage. Our general goal of the project remained the same, but our method of accomplishing the goal changed in how we decided to perform computations on the data.

Discussion

A discussion, analysis and interpretation of the results of our program and how it relates to our research question.

Our research question was originally designed to explore the factors influenced by climate change, including temperature, humidity, and wind. Our purpose was to see the kind of impact that they have on the intensity of forest fires and the connections that we could make between the factors. We defined the intensity of forest fires using two factors, DC (drought code) and ISI (initial spread index). Drought code is the average rating of the moisture content in the organic layers where a fire may occur. This factor has a large impact on the severity, extent and frequency of forest fires and thus contributes to the intensity. As for for the initial spread index, it can be defined as the measure of how quickly a fire can spread which also contributes to the intensity.

We found that DC is the factor that has the highest relation to temperature, with an R^2 value of 0.246. The other R^2 values found through linear regressions of factors in comparison to temperature are not that high as they are dependent on multiple other factors. From double regression, we can see that DMC (Duff Moisture Code) and humidity have a close relationship with temperature as well, with an R^2 value of 0.538. As the number of independent variables increases, we can get a much more accurate prediction of the dependent variable. It was necessary to be mindful of the choices we make for independent variables when performing double regressions, as many of the factors affecting forest fires are connected to each other. As a result, multicollinearity could occur and this would mean the independent variable and the resulting coefficients do not have much statistical significance.

There were other factors that we did not consider that would have an impact on forest fire intensity. Different circumstances produce different fires, and what areas of trees are involved can influence ISI. We can also consider that human-caused fires may be different than natural-occurring fires. While it is possible that an increase in human population over time may increase the number of forest fire occurrences, we assume that our definition of intensity would be dependent more on the factors we considered. Data that includes the underlying cause of each fire could further supplement this project.

We discovered that FFMC (Fine Fuel Moisture Code) is directly related to temperature, relative humidity, wind and rain. These are all factors that have been greatly influenced by climate change. ISI, which we used to help define the intensity of forest fires, is dependent on FFMC and wind. From our computations, the increase in wind speed did not show a clear change in DC or ISI so based on the data we used, we ignore wind speed as a factor that could predict forest fire intensity. This is not to suggest that wind is not an important factor, but the data we used was inconclusive. However, we learned through double regression that ISI was dependent on two variables (temperature and humidity), and these factors have a direct correlation to climate change. We can make sense of this result given conditions in which fires spread. This helps to answer our research question, suggesting there is a connection between these two factors that are influenced by climate change and ISI, which we use to define forest fire intensity. As a result of this computation, we predict that an increase in temperature will yield greater DC and ISI values, and thus more intense fires.

We used Portugal temperature data to predict future temperatures from the general trend since 1774. Through our analysis and computations, we observed that temperature had the closest relation and effect on both DC and ISI, and therefore was the biggest contributor to the intensity of forest fires. We were able to predict that higher temperatures will result in increased values of DC and ISI. This means that we can conclude the intensity of forest fires has increased over the years and will continue to increase as it is related to temperature. Climate change has caused a noticeable change in temperature over the years. Data of Portugal temperatures did not have a very evident or observable change in temperature because we averaged the monthly temperatures. However, we know that a slow and gradual increase does exist, and around the world there is a clear increase in temperature as a result of climate change. Therefore, we can reasonably conclude that the intensity of forest fires will increase over time.

Our computations and data sets were a good starting point to see the impact that factors related to climate change can have on forest fires. If we want to get an even more clear and complex understanding on this topic, it would be necessary to perform computations on much larger and varied data sets. We were able to conclude that increasing temperature has the biggest influence on an increase in forest fire intensity, answering our research question. Our computational exploration was limited by the data sets that we used. There are not many public forest fire data sets that include the factors we wanted to explore. We were limited to exploring fires in Montesinho National Park and were unable to analyze fires in other locations. For further exploration of the relationship between climate change and forest fires, it would be necessary to find data sets that contain the forest fire information of various locations to see how geographical location has an impact; our code can be reused with data from different locations. We would also suggest collecting more data on the factors that impact forest fires, such as the aforementioned issue of natural versus human caused fires.

References

Berwyn, Bob. ”How Wildfires Can Affect Climate Change (and Vice Versa).” Inside Climate News, 23 Aug. 2018, https://insideclimatenews.org/news/23082018/extreme-wildfires-climate-change-global-warming-air-pollution-fire-management-black-carbon-co2.

“Canadian Wildland Fire Information System: Canadian Forest Fire Weather Index (FWI) System.” Canadian Wildland Fire Information System — Canadian Forest Fire Weather Index (FWI) System, https://cwfis.cfs.nrcan.gc.ca/background/summary/fwi.

Cortez Paulo & Morais Anibal. ”A Data Mining Approach to Predict Forest Fires using Metereological Data.” University of Minho, http://www3.dsi.uminho.pt/pcortez/fires.pdf.

“Fire Weather Index (FWI) System.” NWCG, 15 Mar. 2019, www.nwcg.gov/publications/pms437/cffdrs/fire-weather-index-system.

“How to Extract the Regression Coefficient from Statsmodels.api?” Stack Overflow, Nov. 2017, https://stackoverflow.com/questions/47388258/how-to-extract-the-regression-coefficient-from-statsmodels-api

“Introduction — Statsmodels.” Statsmodels, https://www.statsmodels.org/stable/index.html

J.D., Hunter. “Matplotlib.” Matplotlib: A 2D Graphics Environment, IEEE COMPUTER SOC, 2007, https://matplotlib.org/.

Mne-Learning-Databases/Forest-Fires, Archive.ics.uci.edu, 2008, https://archive.ics.uci.edu/ml/machine-learning-databases/forest-fires/.

Natural Resources Canada. “Canadian Wildland Fire Information System: Canadian Forest Fire Weather Index (FWI) System.” Canadian Wildland Fire Information System — Canadian Forest Fire Weather Index (FWI) System, https://cwfis.cfs.nrcan.gc.ca/background/summary/fwi.

”Plotly Python Open Source Graphing Library.” Plotly Graphing Libraries, https://plotly.com/python/.

“The Environmental Impact of Forest Fires.” Untamed Science, 7 Feb. 2020, https://untamedscience.com/blog/the-environmental-impact-of-forest-fires/.

“Three-D Scatter Plots.” Plotly, https://plotly.com/python/3d-scatter-plots/.