Project Report

A Study of Wildfires in Australia (2005 - 2020)

Course: 19CSE304 Fundamentals of Data Science

Group Number 01

Group Members:

- Manas P P (AM.EN.U4CSE19233)
- Tarun Reddy (AM.EN.U4CSE19249)
- Sai Rakesh (AM.EN.U4CSE19240)
- SSSB Shravanth (AM.EN.U4CSE19250)

Table of Contents

- Abstract..... 03
- Introduction..... 03
- Broad Context..... 04
- Study System..... 04
- Methods..... 04
- Results..... 05
- Discussion..... 09
- Literature Cited..... 09
- Conclusion..... 09

Abstract

Data Science has a great potential to be used as a tool to deal with climate change. In this report we study the forest fires that burnt in Australia from the years 2005 to 2020. With proper collection of data, data visualisation and data science can be used to get an in depth understanding of the fires in Australia that affected 3 billion animals in 2019 and 2020 and also led to huge human costs.

Introduction

Climate crisis has been rapidly accelerated by various factors. Forest fires are becoming more devastating especially after years of forest mismanagement. Satellite data collected by NASA's NRT-MODIS uses images with each pixel estimating the fire anomalies 1 km wide. This gives us an estimate of fire area, fire brightness and fire radiative power (see Fig 1).

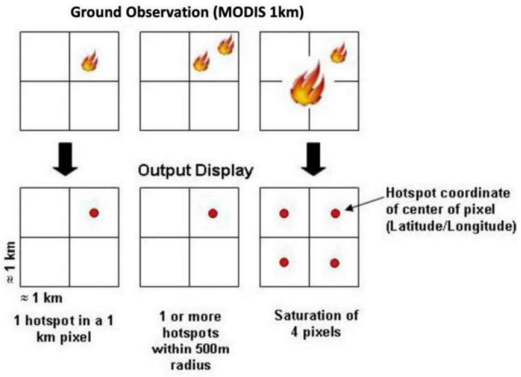

Fig 1 Collection of MODIS data

The data was averaged over the 7 states/regions in Australia from 1/1/2005. The 7 regions and their area in km² are:

1) NSW=New South Wales (801600)
2) NT=Northern Territory (1346200)
3) QL=Queensland (1727200)
4) SA=Australia (984000)
5) TA=Tasmania (67800)
6) Vl=Victoria (2525500)
7) WA=Western Australia (227600)

Fig 2 is a map of Australia showing remoteness of the 7 regions. We see that the major parts of Northern Territory, Western Australia, South Australia and Queensland are remote regions. We use information about various geographical factors like precipitation, temperature, soil water content, relative humidity, solar radiation, the land class, vegetation index and wind speed to study the effect of these parameters on wildfires.

We perform data analysis using numerical and visual techniques to study the wildfires in Australia and to discover the factors having the most influence on the wildfires across the seven regions.

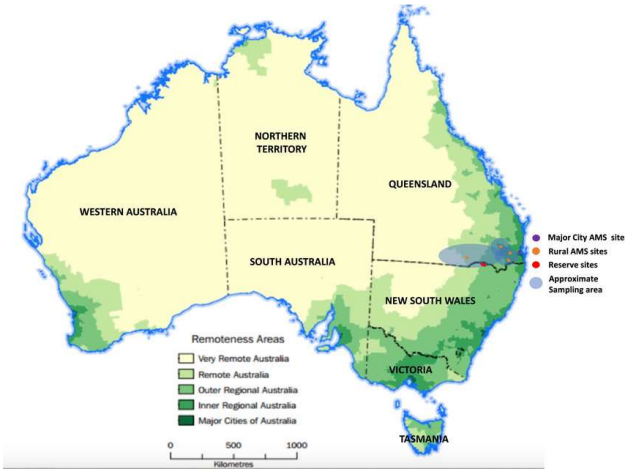

Fig 2 Remoteness of regions in Australia

Broad Context

Data Science techniques can be used to study the effect of climate change on marine biology, land use and restoration, climate change and much more. With a proper study of data it is possible for companies to measure their carbon emissions accurately, study their production process and take steps towards sustainable development. Hence, data science can play a key role in climate change mitigation. In this report we study the forest fires in Australia, the data of which was collected by NASA.

Study System

We perform an observational study on the forest fires that were recorded in Australia from the year 2005 to the year 2020. The data is split based on the Region/State of Australia, giving us seven datasets for this study. We make observations from the data about the factors most effecting the fires and we also check for seasonal patterns in the wildfires across the states.

Methods

Data Preparation: We use Python libraries NumPy and Pandas to split the data based on Regions. We also use this for feature engineering.

Data Analysis: We use data visualisation to make plots to better understand the data. We use statistical measures like mean, correlation, etc to derive insights from the data.

Results

The mean estimated fire area of each region (see Fig 3) over the entire period saw a correlation of 0.799 with the total area of that region. Fig 4 shows the percentage contribution by states to wildfire in Australia from 2005 to 2020. Fig 5 shows the annual average wildfire area of each region over the 16 year period. We see that Northern Territory, Queensland and Western Australia saw the highest area of wildfires. However, in 2019 we see that New South Wale had the highest fire areas. Fig 6 gives the correlation between the land cover area of various classes and the area of wildfire in a region.

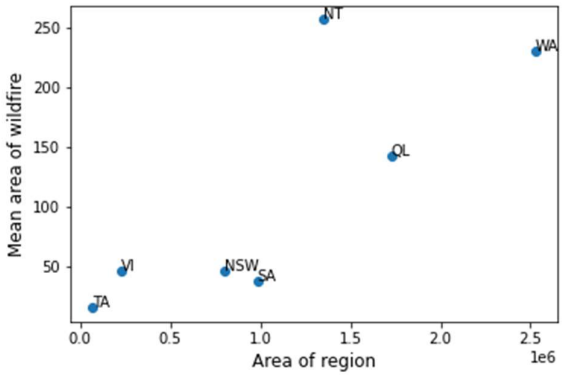

Fig 3 Mean estimated fire area of each region over the 16 year period

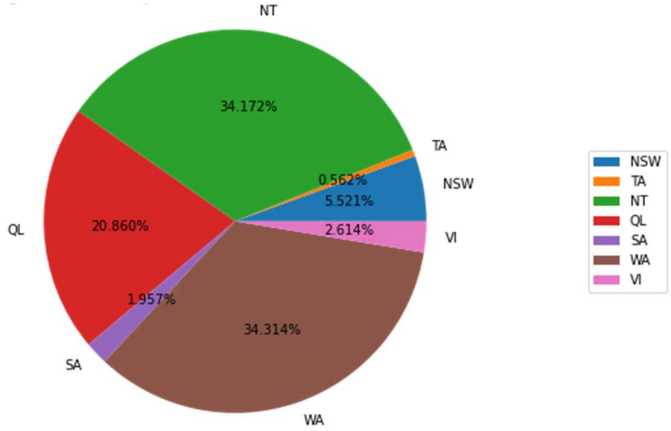

Fig 4 Percentage contribution by states to wildfire in Australia from 2005 to 2020

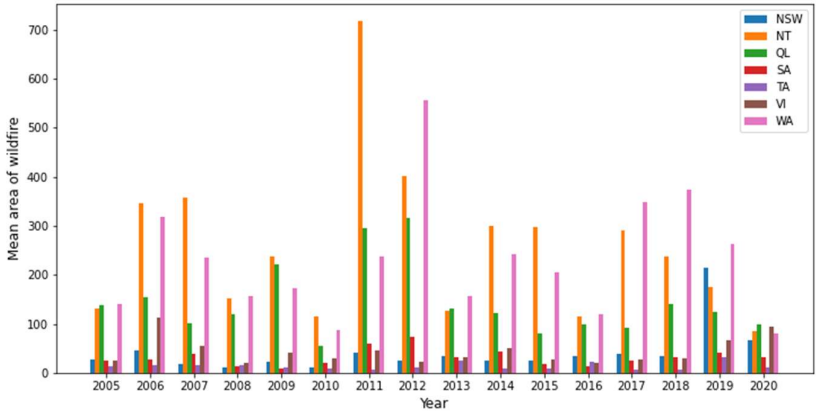

Fig 5 Annual average wildfire area of each region over the 16 year period

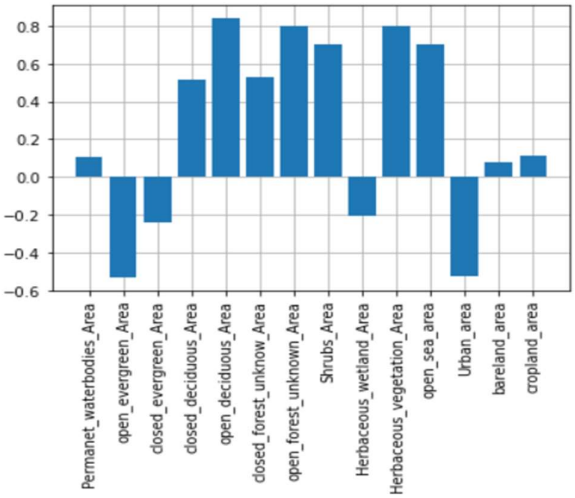

Fig 6 Correlation between land class area and wildfire area

In Fig 6, we derived the area under each land class by taking the percentage coverage in the state and multiplying it by the total area of the state and we see the most prominent correlation of the wildfire area are with open deciduous forest area, open unknown forest area, herbaceous vegetation area, open sea area and shrubs area. In Australia the seasons are defined by grouping the calendar months as follows: Spring (September, October, November), Summer (December, January, February), Autumn (March, April, May) and Winter (June, July, August). We observe that generally spring and summer recorded the higher areas of wildfires, with few regions like Western Australia seeing large fires in autumn as well. Table 1 shows the monthly average of the estimated wildfire area over the 16 year period.

Table 1 Average wildfire area for each moth across the regions

	QL	NSW	NT	SA	TA	VI	WA
Jan	96.1	60.1	81.5	61	44.1	133.7	171.4
Feb	46.4	29.6	39.4	31.2	19	57	126.3
Mar	21.5	37.3	30	16.5	13.5	44.5	76.6
Apr	19.4	63.7	121.1	31.2	16.3	58.8	182.1
May	25.2	42.9	215.6	14.9	8.6	20.8	170.2
Jun	34.6	14.5	203.1	7.2	6.3	5.3	75.5
Jul	57.5	10.2	217.1	6.7	2.5	3.6	60.3
Aug	100.2	18.9	328.4	22.9	5.6	3.9	148.5
Sep	212.6	25.8	590.6	57.7	4.8	8.2	316.8
Oct	331.9	24.9	508.44	67.9	8.9	7.9	604.3
Nov	444.5	102	308.8	51.1	7.3	11.6	568.8
Dec	276.6	130	153.6	50.2	23.2	140.5	243

The mean estimated fire brightness showed a seasonal trend with the peak each year coming in the months of summer, while winter saw lower brightness each year. Fig 7 and Fig 8 shows the plot of mean estimated fire brightness across the years in Queensland.

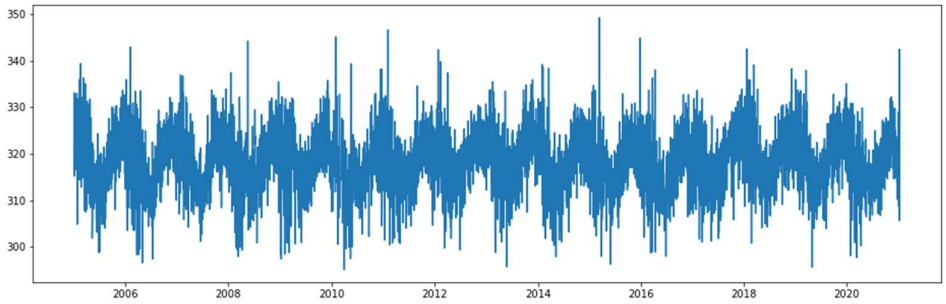

Fig 7 Plot of mean fire brightness across the years in Queensland

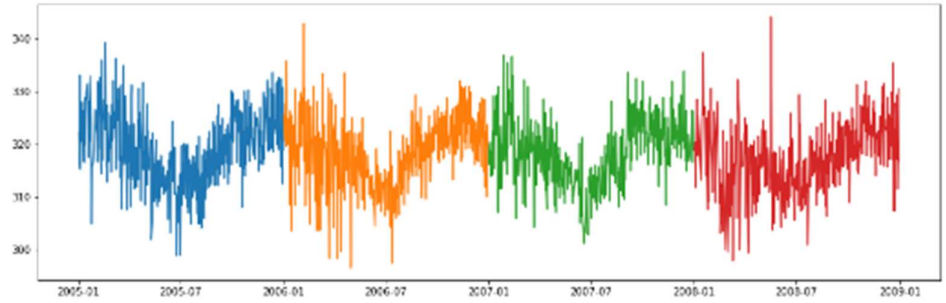

Fig 8 Plot of mean fire brightness from 2005 to 2009 in Queensland.

The fire intensity is measured using the estimated fire radiative power [3]. In [3] the fires which have radiative power less than one standard deviation of mean is labelled as Low fires, the fires with power more than one standard deviation of mean is termed as High fires, while all the fires within one standard deviation of mean in termed as medium fires. We use the same convention here. Fig 9 shows the annual mean fire radiative power of each region. Fig 10 shows the annual ratio of fires falling in each category for each of the seven regions. Fig 11 shows the percentage of the total area burnt over the years contributed by each of the three types.

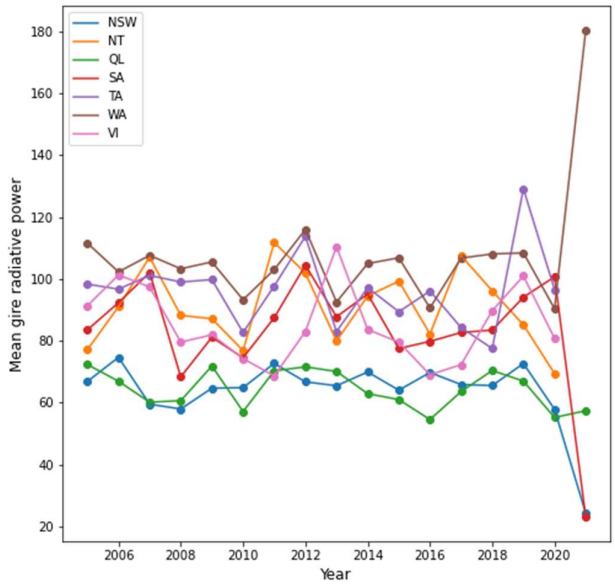

Fig 9 Annual mean fire radiative power in each region over the years

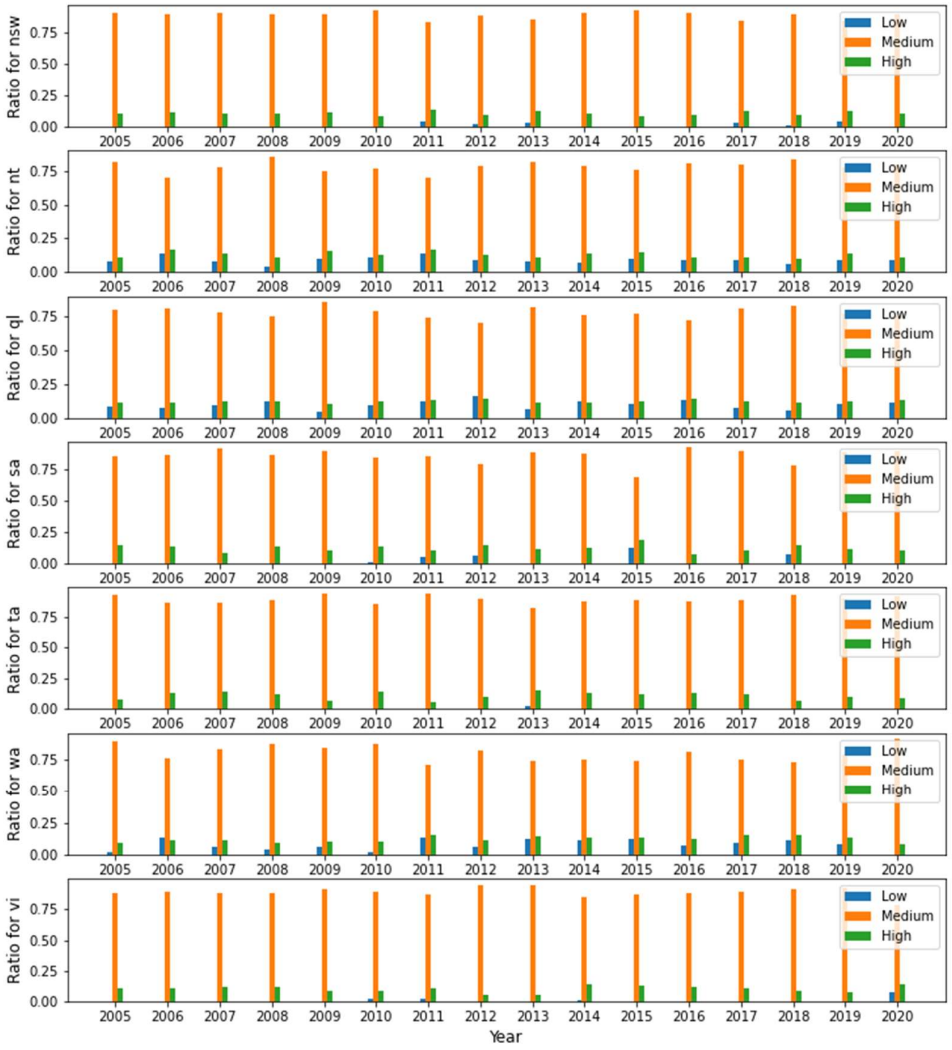

Fig 10 The annual ratio of fires belonging to low, medium and high types in each region

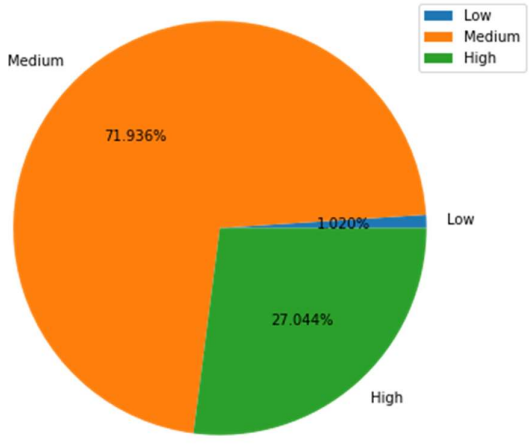

Fig 11 Percentage of forest area burnt by the fire type.

Discussion

We have seen that the area of wildfire usually increases with the area of the state/region. Northern Territory recorded the highest wildfire areas with it being the third biggest state. We saw that Northern Territory, Queensland and Western Australia made up to 89.35% of the observed wildfire areas over the entire period. 2019 and 2020 saw New South Wales and Victoria having relatively higher areas. The area of open deciduous forest, open unknown forest and herbaceous vegetation saw the highest correlations with wildfire area. While, open evergreen forest area and urban area had negative correlation below -0.5.

The fire brightness showed a seasonal trend with the higher figures coming in the summer and the lower figures in winter.

The fire intensity is measured by using the estimated fire radiative power. We classified the fires as low, medium and high fires. We see that medium fires dominated the wildfire landscape and high fires were more abundant than low fires. Western Australia saw a sharp spike in the fire radiative power in 2020, while Southern Australia and New South Wales saw a dip. Almost 72% of the observed wildfire area came from medium fires, 27% from high fires and 1% from low fires.

Literature Cited

1. GitHub - Call-for-Code Spot-Challenge-Wildfires: https://github.com/Call-for-Code/Spot-Challenge-Wildfires
2. https://community.ibm.com/community/user/datascience/blogs/susan-malaika/2020/11/10/call-for-code-spot-challenge-for-wildfires
3. Remote Sensing Data for Calibrated Assessment of Wildfire Emissions in Siberian Forests: https://www.mdpi.com/2504-3900/2/7/348

Conclusion

In this report we studied the data covering the forest fires in Australia from 2005 to 2020. We see that most of the fires were medium fires. The area of open deciduous forests, open unknown forests and herbaceous vegetation saw high correlations with the wildfire area. We saw that the fire brightness showed a seasonal trend. Northern Territory, Queensland and Western Australia saw the largest wildfires, with New South Wales and Victoria seeing an increase in the years of 2019 and 2020.