2019 Australian Bushfires and Weather Trend Analysis

Olivia Alexander
olivia.alexander@sjsu.edu
014545986

Jeehee Choi
jeehee.choi@sjsu.edu
012565696

Sukriti Mishra
sukriti.mishra@sjsu.edu
014580696

Ziwei Song
ziwei.song@sjsu.edu
014557231

1. Introduction

Australia is one of the world's greatest diversities. But today, Australia is burning, after a span of 8 horrendous months. Wildfires are an uncontrolled fire caused by natural vegetation, coined “bushfires” Australia and “brush fires” in North America. There are both natural and human factors that can lead to bushfires. In the case of the 2019 Australian bushfires, both factors likely played a role in the start, severity, and longevity of these events. A few key weather variables that are shown to be responsible for the bushfires are excessive heat, high wind speed, and drought conditions resulting in dry vegetation. When this combination of factors collide, a fire can get quickly out of control. Since August 2019, at least 27 million acres of Australia have burned, more than 14 times the area that burned in California in 2018. An estimated 1.25 billion animals have lost their lives as a result of these horrendous bushfires. Australia’s bushfires have released 400 megatons of carbon dioxide into the atmosphere which can cause human respiratory problems, greenhouse effects, and ultimately global warming. According to a BBC report, temperatures have already risen by more than one degree Celsius since 1920. 2019 was the hottest year on record, as shown in Figure 2a. In December 2019, the temperature was more than 100°F and there seems to be a trend of Australia getting warmer over time. There are three purposes of this project: (1) Study the natural causes of the fire in the form of historical weather readings. Search for any correlations or trends over time for weather variations and subsequent fires as a sign of climate change. (2) Explore the states, regions of Australia that were most adversely affected by these events. With the analysis of the bushfires, patterns or trends of the fires can be detected in order to make predictions for future reference. (3) Examine the timeline of fires in Australia as well as the timeline of weather patterns during the bushfires as they were most severely occurring in late 2019.

2. Related Work

The increase in frequency of Australian bushfires over the span of few years has suggested a major climatic change in coming years. There is a crucial requirement of scientific research and understanding for climatic change due to Australian bushfires to support Australian society. The bushfire occurrences from various climate data contain spatial and temporal information as well as climatic aspects of bushfires. There is an aim to derive the contextual information of a model to calculate accurate prediction for future bushfire hotspots. From weekly climatic surfaces, researchers developed an ensemble method on the basis of a two-layered machine learning model to establish a relationship between fire incidence and climatic data to provide highly accurate bushfire incidence hot spot estimation with a global accuracy of 91% [8]. Their analysis also showed that Australian weekly bushfire frequencies increased by 40% over the last 5 years, especially during the late summer months, indicating a serious climatic shift. Data from two NASA satellites, the Terra and Aqua satellites, detect the emission of infrared radiation by fires. Figure 2b shows that 2019 was the most active fire year in the last 20 years, and that the regions of the southeastern states of New South Wales and Victoria have been adversely affected. According to wildfire researchers, the consolidation of extremely dry and extremely hot conditions are additional factors responsible for the more powerful fires in Australia [9].

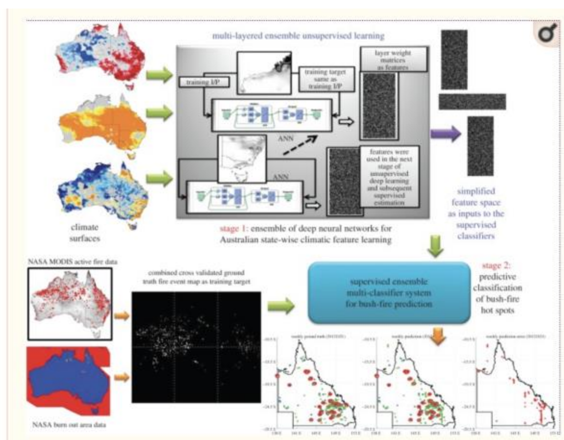

Figure 1: Deep learning mechanism for weekly bushfire frequency estimation and hot spot estimation

K.Hennessy et al. [7] explored how the fire-weather risk is related to climate change and proposed improved methods on assessments of fire risk in the paper. Fire-weather risk is affected by a number of weather variables, including temperature, humidity, wind speed and precipitation. In order to measure the fire-weather risk in a numerical way, two indexes were introduced in the work. One was the Forest Fire Danger Index (FFDI) and the other was Grassland Fire Danger Index (GFDI). FFDI is commonly used by weather forecasters and fire services agencies in Australia as an indicator on risk and hazard of forest fires.

This work provided an improved measurement on fire risk assessment in two ways. The first approach was re-evaluating the raw weather data, and selecting the most representative daily weather data. Another approach was adjusting the data variance, and increasing weights in extreme values.

This work also recalled some previous assessment on calculation of FFDI. Beer and Williams (1995) found how the amount of carbon dioxide increased the FFDI. They used simulated daily weather data from a climate change model and applied the dataset into a fire risk model. This paper examined the disadvantages of this methodology that the errors in the simulated baseline climate will cause bias in the result. Cary (2002) explored how climate change by 2070 would affect FFDI by two steps. The first step was using monthly average simulated climate change data from the CSIRO regional climate model (DARLAM), combined with observed daily weather data. The second step was applying the above data into the ANU FIRESCAPE model and concluded the result of 5%-20% increase in FFDI.

In the end, this work analyzed FFDI results compared with climate change simulation results. In our paper, we particularly concern three major Australia cities, Coffs Harbour, Sydney and Newcastle. Hennessy et al. mentioned two of them. The report provided two sets of comparison. One was the average annual accumulated FFDI from 1970 to 2003 and prediction on increase of FFDI by 2020 and 2050. The other was the average number of days where FFDI was extremely high and prediction on the average number of days where FFDI could be extremely high by 2020 and 2050. For Coffs Harbour, the average annual accumulated FFDI was 2002, and could increase 2%-6% by 2020, 5%-15% by 2050; there were 4.4 days when the FFDI was very high and this could increase to 4.7-5.6 days by 2020 and 5.1-7.6 days by 2050 [7].For Sydney, the average annual accumulated FFDI was 2158, and could increase 2%-7% by 2020, 5%-19% by 2050; there were 8.7 days when the FFDI was very high and this could increase to 9.2-11.1 days by 2020 and 9.8-15.2 days by 2050. From the above statistics, we can tell that the risk of fire weather for the two major cities are potentially high in the future. Our paper will then focus on weather variables such as temperature, relative humidity and wind speed.

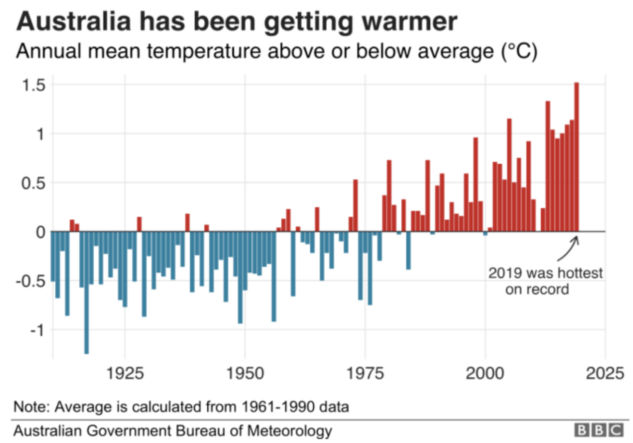

Figure 2a: Annual Mean temperature of Australia over the past century

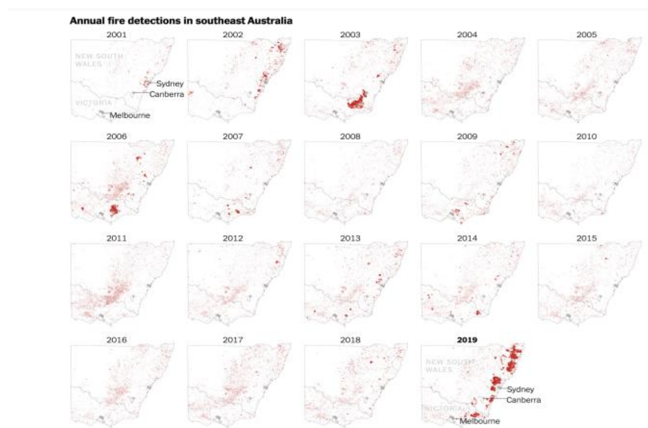

Figure 2b: Total fire detections in southeast Australia each year from 2001 to 2019.

3. Methods

Using open-source datasets through the Australian Government Bureau of Meteorology as well as Kaggle, weather and fire data from Australia were collected. The aim of this analysis was to gain insights about the weather patterns in Australia, leading up to and during the deadly bushfires starting in late 2019. The Australia bushfire dataset used in this study was recorded by MODIS, a NASA satellite instrument, to identify active fires and thermal anomalies. The fire dataset acquired included over 350,000 records of fires by pixel, where each pixel is a 1 kilometer (0.62 miles) area. Feature selection was performed on the initial 14-column dataset in order to condense the dataset into the most important variables: latitude, longitude, and date. The dataset included the time period of August 1st 2019 to January 11th, 2020, almost 6 months of daily fire data in Australia. This dataset was preprocessed using Python in order to count the number of fires/areas burned by the number of records per date. This calculation allows for the summarization of fires over the

6 month period to be visualized in a time series. This visualization can be used to explore the timeline of fire occurrence in Australia for the bushfire occurrences from August 2019-January 2020 (Figure 3). The fire dataset was also spatially joined with the Australian state administration and New South Wales city administration boundary shapefiles using ArcGIS to obtain city and state administration for each fire event [6]. Another technique applied to the fire dataset was to visualize the distribution of fires across Australia using the latitude and longitude features from the dataset, plotting the records individually (Figure 4a). This action also allowed for the identification of dense fire regions for further analysis in relation to the weather datasets.

It was reported by The Verge that the state of New South Wales had been affected by the Australian bushfires most severely, with the coast specifically being extremely damaged [5]. Therefore, the distribution of fires in New South Wales was visualized cumulatively by month in Figure 4b for further investigation. It was then hypothesized that a suitable visual representation of the data shown in Figure 4b could be shown as a series of maps where monthly fire occurrences are displayed to show the timeline of fires as well as the location of occurrences (Figure 5a-e). Additionally, the overall fire distribution in New South Wales was examined to identify the top 10 cities with the most fire occurrences, which were all along the eastern coast (Figure 6a-b). An interactive Tableau dashboard, Figure 7, was then developed, to showcase three major visualizations to explore the 2019 Australian bushfire occurrences: the continental distribution in Figure 4a with the additional affordance to filter by state, the New South Wales Distribution in Figure 4b with an interactive menu to show individual monthly values as shown in Figure 5, and the top 10 cities in New South Wales for fire frequency shown in Figure 6.

To narrow the scope of the weather pattern investigations and to evaluate these dense fire regions more thoroughly, three major cities along the coast of New South Wales were extracted from the dataset and analyzed for weather trend patterns. Sydney was selected as a major city near Blue Mountains National Park, the area with the most frequent fires in New South Wales. Similarly, the town of Coffs Harbour was selected as a second case study, which is a coastal city near Widden, the fourth most affected city in New South Wales, as well as Newcastle, located near Newton Boyd, also in the top 10 New South Wales bushfire list. These three New South Wales coastal cities, all dispersed across the eastern coast near the most affected regions of Australia for the 2019 bushfires, serve as a macro-level indicator of the lingering effects of the bushfire occurrences.

The Australia weather datasets collected contained daily weather observations from various weather stations across the continent over a decade, from December 1st, 2008 to December 31st, 2019, and included 24 variables. The daily weather data was provided for approximately 50 general city locations across Australia. The weather variables selected for this study were temperature variables such as minimum and maximum values, rainfall, humidity, and wind speed as it has been reported that these features can be strong indicators of fire disturbances [2]. There were two readings each for humidity and wind speed, one at 09:00 and the other at 15:00. One of the case study locations, Newcastle, was missing a significant number of readings, with the entirety of 2015-2018 as NAs, for all four readings of humidity and wind speed. Therefore, this location was not included in these two weather plots as it was believed this would show a biased result. Additionally, 2017-2019 was not included for the humidity and wind speed visualizations because there were several missing values for the other two locations, Sydney and Coffs Harbour, as well.

Before analysis, the weather dataset was wrangled to remove NA values as well as any other error values. Each city in the dataset had approximately 3,000 daily entries, which would make the trendline of this visualization quite noisy and less effective to the audience. Therefore, for the 10-year historical weather trends, the average annual value per feature was calculated in order to smooth the data, resulting in a straightforward story for the viewer. Comparing these sets of time series trends for various variables, the aim was to identify patterns in the weather data in years leading up to the Australia bushfires that would explain why some areas burned so severely (Figure 8a-e). Additionally, the monthly weather values for the three case study sites for August 2019-December 2019 were visualized in order to examine how the weather was changing during the deadly bushfire season of 2019, and which features were reacting most severely (Figure 9a-e).

4. Results

4.1 Application of Principles

The visualizations in this study are organized following the Gestalt laws of grouping: proximity, similarity, good continuation, and good form. The bushfire event data in the same state are colored-coded in Figure 4a. Proximity shows small assembled fire data in a composition. Furthermore, the principle of similarity is shown in the fire data gathering within the Australian state borderlines in Figure 4a. Within Figure 8 and Figure 9, there is a clear representation of the principle of good continuation, as each plot within the figures does not change form for any unnecessary reason, and remains consistent. Additionally, the decade-interval of historical weather data (Figure 8) and the weather data during the fires (Figure 9) show similar patterns of plot line colors in each chart to represent corresponding variables. The last principle is good form. The four red color saturation and intensity of color allow grouping together similar colors or shapes in Figure 4b.

Another important aspect of visualizations is formed: length, width, orientation, size, shape, enclosure. For dashboards and storyboards, each visualization is recommended to be in similar length, width, and orientation. The size of the chart can be adjusted to fit the size of the board, respectively. Figure 7a-d displays different sizes of chart mainly because of the long length of a variable, 'Blue Mountains National Park,' in the top-right of the chart. The visualization must enclose each chart containing all the necessary data. For example, a dashboard should be on one page with no scroll and few drop-down options. Big bold title and left-aligned texts in legends and titles in a visualization increase readability. Another way to improve readability for viewers is by using Miller's magic number of 7 bits of information (plus or minus 2 bits). Almost all of the visualizations have between 5 to 9 variables of information, with the exception of Figure 6a showing 10 cities/pieces of information because the 10th city, Newton Boyd is crucial to represent the Coffs Harbour data. Furthermore, texts in Figure 6a were originally in all upper case letters but changed to lower case letters to enhance readability.

Additionally, removing data-ink increases effectiveness and expressiveness. No outline color exists in Figure 4a-b to 7a-d. This can say that the fire event data overlap each other to form a big merged dataset. Different color-coding is used in Figure 7a-d to avoid confusion within the dashboard. The Australia map consists of blueish colors, while only New South Wales state is in red (Figure 4a). The zoomed-in version of this state in Figure 4b uses red saturation color to represent bushfire events by different months in 2019. The bar graph and heat map in Figure 6 uses rainbow coloring to represent bushfire events by different city administrations. These different color-coded graphs with openstreetmap backgrounds helps readers find locations of major cities in New South Wales in relation to the bushfire events. Another way to improve graph quality is by removing redundancy. Figure 6a shares the same color-code with Figure 6b, so its legend is removed to avoid redundant information.

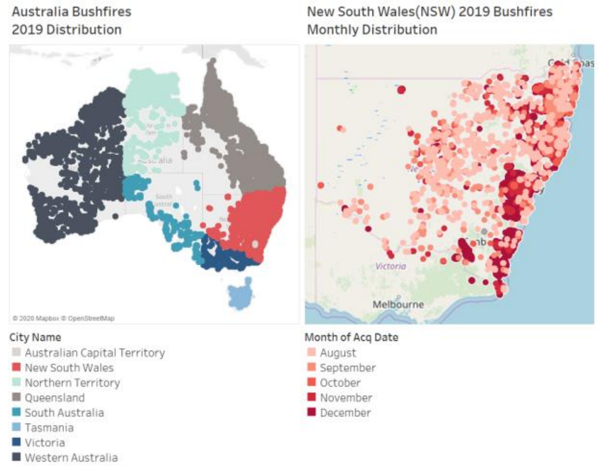

Figure 4a-b: 2019 Bushfire Distributions- Australia & New South Wales

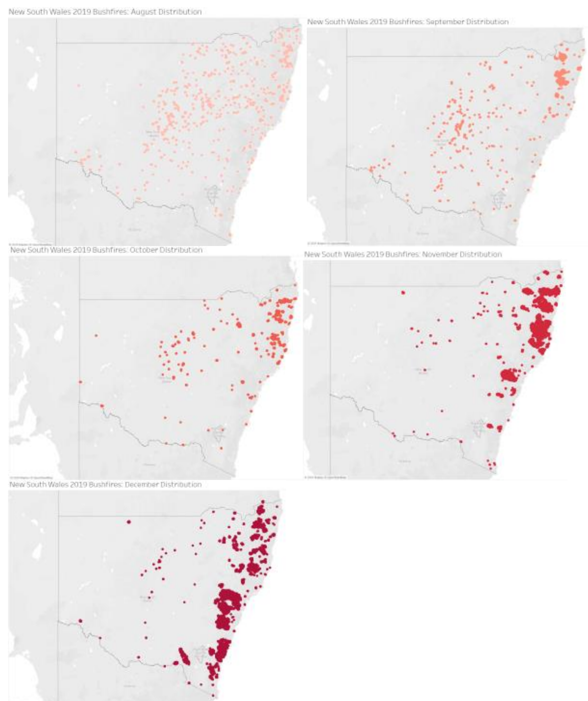

Figure 5a-e: New South Wales Monthly Fire Distributions for 2019 Bushfires- August to December

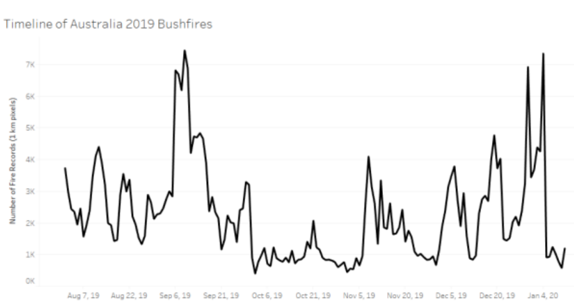

Figure 3: Timeline of 2019 Australia Bushfires from August 2019-January 2020

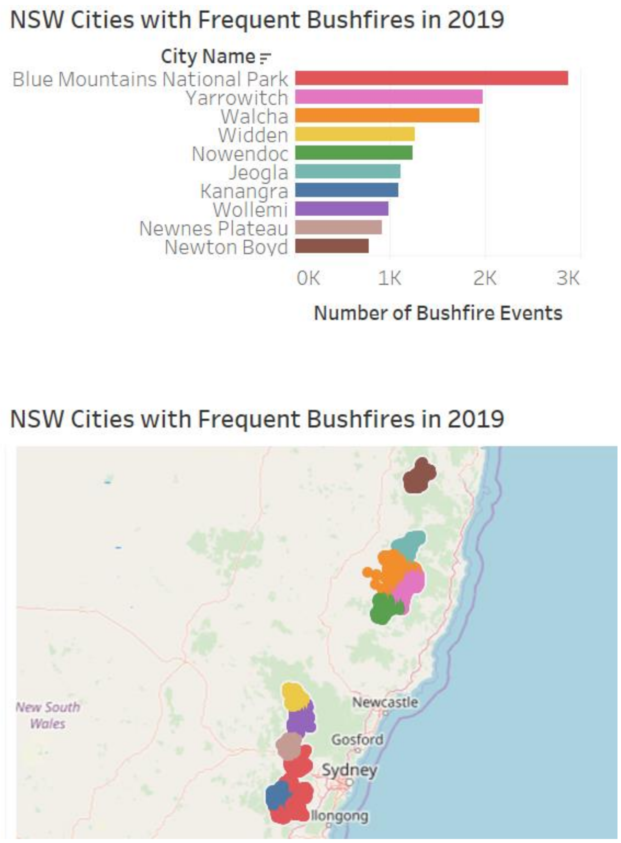

Figure 6a-b: Top 10 Cities in New South Wales for Fire Frequency During 2019 Bushfires

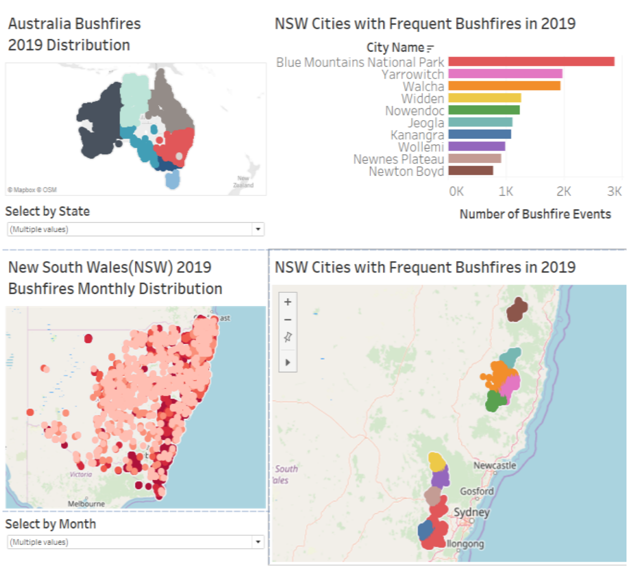

Figure 7a-d: 2019 Australia Bushfire Dashboard

4.2 Visualization Techniques

The hierarchy of data uses the top-down approach to illustrate the most important information with the order of top left, bottom left, top right, then bottom right. Visualizations in a dashboard are organized with hierarchy while a storyboard is ordered with time.

A dashboard has the following infographic formats: static, zooming, clickable, and interactive. Figure 7a-d illustrates static function with text, charts, and maps. Zooming functionality allows zooming in and out of complex visualizations such as maps. Clicking affordances have a drop-down menu to select one or multiple filters then produces different graphs accordingly. In presentation mode, storyboards demonstrate changes in infographics by each chart, similar to animations.

There are three main types of charts used in this study: bar graphs, heat maps, and line charts (time series plots). Below discusses the merits of each type of visualization used in this report. Bar graphs are effective in comparing multiple features and values that fall within the same datasets. In Figure 6a, the city name is nominal data, and the number of bushfire events is discrete data as an integer. Bar graphs are easily displayed and interpretable with the city values on the y-axis and the number of bushfire events on the x-axis. Another frequently used visualization is the heat map. In this project, heat maps with longitude, latitude, and other quantitative values are useful to represent geographical data with color and density. Heat maps are used in this study to visualize bushfire events by region and time, summarized in Figure 7. Line charts are effective in displaying the trends of change through time. Figure 8a-e and Figure 9a-e illustrate fluctuations and patterns of annual and monthly data associated with different cities' weather datasets.

4.3 Evaluation

On a continental level, Australia suffered severe bushfires during August to December 2019. The fire events were mostly distributed along the coast, and New South Wales had fire events with high density. Considering the Blue Mountain National Park, located close to Sydney, with huge vegetation in New South Wales, the damage was substantial. The main method to analyze the bushfires in this research was to examine the historical weather data leading up to these extreme bushfires as well as the weather data during the bushfires. Using the three selected case study locations as major coastal cities near some of New South Wales' most affected areas, five major weather factors that can affect the chance of fires were analyzed.

Using a decade-long trend from 2009-2019 for most of the features, the Australian climate preceding the 2019 bushfires were captured and visualized for evaluation. As the eastern coast of Australia was affected most severely, these three sites are visualized together in each subfigure for Figure 8 and 9, grouped by weather variable. This technique is used to demonstrate the potential consistent trends of weather variables within the three locations during the past decade that could have contributed to a higher probability of deadly fires in the area, examined in Figure 8, as well as to demonstrate how the weather variables changed and reacted during the fires in late 2019, examined in Figure 9.

A combination of fluctuations between a number of weather variables can lead to a higher probability of fire occurrences. Therefore, the trends of each of these pertinent variables are examined within Figure 8 with respect to annual averages for maximum temperature, minimum temperature, total rainfall, percent humidity, and wind speed. Considering the fact that August through December is fall to winter in Australia, a small amount of temperature change by month is natural. Any unusual phenomenon needs to be diagnosed.

The following are evaluations from the 10-year historical climate data (Figure 8) for the three case study locations, as well as observations from the 5-month interval of August 2019-December 2019 during the 2019 Australia bushfires for the three case study location of Sydney, Coffs Harbour, and Newcastle:

- Maximum Temperature (Figure 8a, 9a): The 10 years of historical weather data shows cyclical increases and decreases with a range of 22 to 25 C' while the August to December 2019 data shows steady growth of around 17 to 27 C'.

- Minimum Temperature (Figure 8b, 9b): The 10 years of historical weather data shows cyclical increases and decreases with a range of 12 to 16 C', with a slight overall increase, while the August to December 2019 data shows 8 to 19 C'.

An observation of an increase in both maximum and minimum temperature during the bushfires, seen in Figure 9, suggests an overall rising shift in the range of temperature values. This shows that during the recent bushfire season, whereas it would normally be expected for the maximum and minimum temperatures to drop during these late months of the year, they both continuously rose, suspectedly as a consequence of the severe bushfires surrounding these cities.

- Rainfall (Figure 8c, 9c): The 10 years of historical weather data for total rainfall shows a consistent decrease. In particular, the city of Coffs Harbour shows dramatic decrease; 6mm in 2009 and 2.4mm in 2019. The August to December 2019 data shows an overall decrease in rainfall for two of the three case study locations.

The rainfall data for Sydney and Newcastle during the bushfires in Figure 9c show similar trends of increased dryness through the bushfire occurrences by starting between 1 and 3.5 mm of average monthly rainfall in August and ending the year with almost 0 mm of rainfall in November to December. It can be assumed that drought would have been a major issue due to lack of rainfall. Additionally, Figure 10 shows that the bushfires were most prevalent in November and December near Sydney and Newcastle.

- Humidity (Figure 8d, 9d): The 10 years of historical weather data shows cyclical increases and decreases with a range of 64% to 72% from 2009 to 2016. The August to December 2019 data shows a range of 58% to 77%, with an increasing trend. The city of Newcastle shows the highest humidity rate with 77% in December.

- Wind Speed (Figure 8e, 9e): historical data shows that most wind speed ranges from 14 to 16 km/hr between 2009 and 2016. No further data is available. Newcastle city's lowest wind speed 12.7km/hr in 2014. The recent data shows Coffs Harbour had the lowest wind speed at 13 km/hr in August while Sydney and Newcastle had around 23 km/hr. The wind speed ranged from 13 to 23km/hr.

As there was no available data from 2017-2018 for humidity and wind speed readings, these years are not included in the 10-year historical climate trend analysis. This discrepancy does not allow for a full picture of the weather leading up to the 2019 bushfires for the purposes of this analysis. However, from what can be observed up through 2016, there was an overall decrease in humidity for the case study sites. Using humidity and wind speed data for 2019 to observe the weather changes during the bushfires, Figure 9d and 9e show the most recent trends for these two weather variables for the case studies. There was an overall increase in humidity during this time period as well as a cyclical increase and decrease in wind speed for two of the sites, Sydney and Newcastle, as 2019 came to a close. For Coffs Harbour, however, the wind speed dramatically increased from 13 km/hr in August, when the bushfires began to spread rapidly, gradually up to 19 km/hr in November, then to 17 km/hr in December.

Figure 8a-e: Pertinent Historic Weather Reading Time Series for Australia Bushfire Case Studies

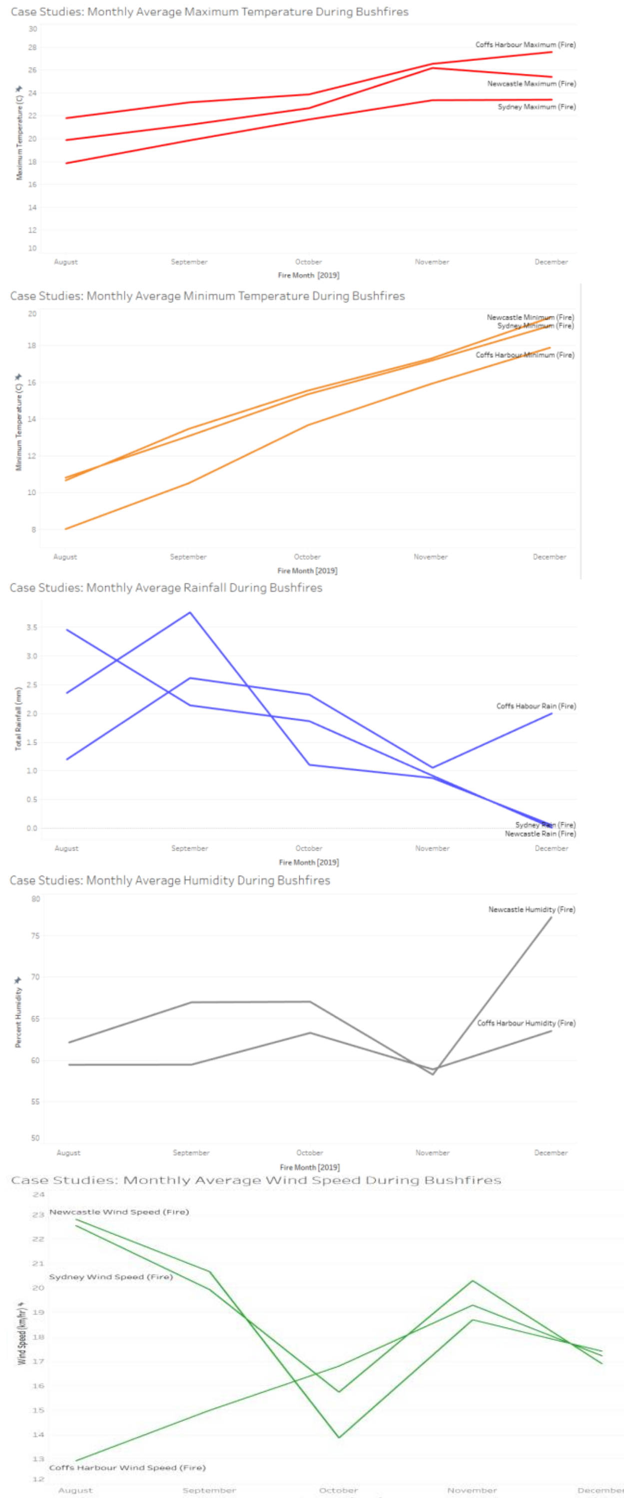

Figure 9a-e: Pertinent Weather Reading Time Series during Australia Bushfires Case Studies

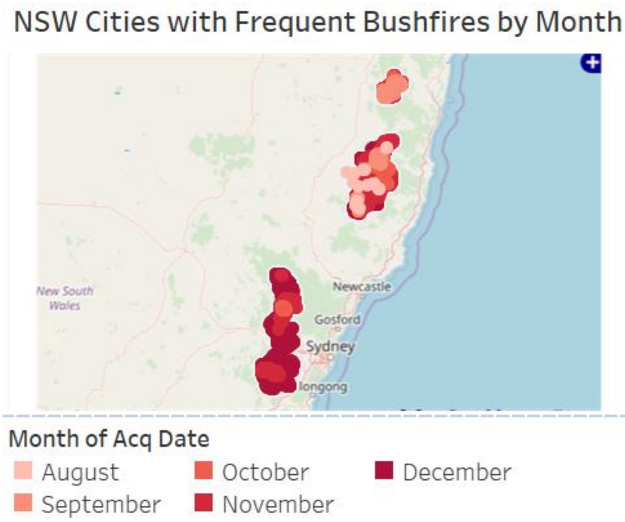

Figure 10: Monthly Fire Frequency Trend during 2019 Bushfires

5. Discussion

The visualization products in this report examine the recent Australia bushfire events in two forms. The first method is the dashboard to visualize the distribution of the fire events, summarized in Figure 7. The time series plot in Figure 3 from August 2019 to January 2020 shows the audience how the events developed during this 6-month period. From the plot, two peak events are observed on September 6, 2019, and January 4, 2020. Then we discovered the events with respect to the regions, locations, and cities. The 2019 Australia Bushfire distribution is displayed by an interactive map with a selection of state as an affordance. The next step was to illustrate our case studies in New South Wales, as the most affected state in Australia by these bushfires, in multiple ways. In order to keep consistent with the approach of exploring the continent-wide bushfire events, we built an interactive map to track monthly fire events along with top citywide fire distributions in 2019.

The second form to examine this crisis in Australia was to produce and display time-series plots for several crucial weather variables related to fire occurrences such as temperature, humidity, rainfall, and wind speed. Comparison between 2009-2016, or 2009-2019 depending on the variable, historical weather data with the same variable data in 2019 provides observations that the New South Wales region had a warm and dry period during the bushfires. Thus, the weather at the time had certain impacts on the severity of the fire events, as well as on the fire behavior and the ease of suppression.

As mentioned in the related work section, a combination of weather variables, such as in the Forest Fire Danger Index, is used as an indicator of fire weather risk. The trend lines we created help to identify how each variable is related and affected the fire events and could also be used as a supplementary interpretation of FFDI. There are still a number of knowledge gaps remaining: (1) The historical data collected in this study ranged from 2009 to 2016 for two of the five weather variables being examined, humidity. Daily weather readings from 2017 to 2018 for a complete historical weather data analysis for these two variables were not obtainable and thus caused a discrepancy when trying to accurately convey the entire weather trends leading up to these bushfire events. (2) Daily weather readings were removed if any variable was missing or contained an error value and outlier. For more accurate results, the quality of the daily weather data can be improved by either collecting data from a more reliable source or filling missing values with predictive values.

6. Future Work

There are several challenges and scopes of improvements in this project which can be considered for future research. For future work, the implementation of machine learning algorithms on the full Australian wildfires dataset for the prediction analysis of wildfires in the future would be a great approach. Through exploratory data analysis, experts can make informed decisions and cost-effective decision making. Additionally, interactive visualization with more advanced performance can be supplemented to the existing visualizations. By adding more functionalities, we could attract domain experts, data scientists, and climate change sectors.

As it was suspected that climate change had an impact on the 2019 Australia bushfire events, regional or local predictions of fire weather can be derived by fitting historical weather data into many existing weather prediction models to get climate change scenarios. Although fire events are not only caused by climate factors, the predictions on fire risk weather can nonetheless be used as a reference for local fire stations to gain awareness for the times with high risks of fire.

Furthermore, we can extend this analysis to other regions in Australia, e.g. Queensland, Western Australia. Some weather variables may display different patterns for different regions. A comparison of the trends of the same weather features in different regions is useful to generate a complete climate change report through key years.

7. Lessons Learned

The main takeaways from this project are setting a goal and parameters, preprocessing datasets, creating meaningful graphs, and finding important insights among graphs.

In this project, we wanted to see the correlation between historic and near-current weather data and the 2019 bushfires in Australia. We have learned that we can collect data by acquiring open-source data or requesting data from a government website. We also learned that the format of data needs to be consistent when modifying and merging with other datasets. Additionally, Tableau and ArcGIS need to be packaged with a dataset when sharing with other people, for ease of accessibility. We set the weather data range for 10 years as this time period is usually the minimum length to be considered as climate. Also, the fire data ranges from August 2019 to December 2019. The three cities of Sydney, Coffs Harbour, and Newcastle were selected since they are major cities on the coast of New South Wales with an abundance of weather data and are in close proximity to the major bushfire events in New South Wales in 2019.

The quality of data used represents the quality of data analysis results: garbage in, garbage out. The accuracy and performance of any model is influenced by the algorithms and the quality of the dataset. The preprocessing techniques are very important because it prepares the dataset in a form that the algorithm can handle. Therefore, data preprocessing takes the majority of the time in data analysis. With the data available from the web, data cleaning, wrangling, joining, and formatting were carefully performed to preprocess the data. Data transformation can include a range of tasks such as converting data types, cleaning data by removing null or duplicate data, and performing aggregations. Missing or incorrect values were also removed to clean the noise. We learned that it is better to drop a variable of a dataset if a majority of the values are missing, as plugging in average/similar values could mislead the overall trend of data or introduce unintentional bias.

Learning how to visualize the dataset by using data visualization tools such as Tableau and ArcGIS was challenging but rewarding. We learned how each button functions; change graph types; how to input map background for heatmaps; change widths and lengths of graphs; demonstrate graphs in storyboards and dashboards; change colors and shapes of icons; change the orientation of texts in graphs because overlapped texts were hard to distinguish; remove redundant legends or variables to increase readability; add drop-down options in dashboards; change headline font style and size; change range of values for a graph axis; change allis names for plot labeling purposes; filter data within Tableau; etc. Creating meaningful visualizations is paramount yet difficult considering the correlation with other variables and graphs and then to visualize this information in an effective and efficient way. We tried to illustrate graphs in a simple and uncomplicated way so that readers can immediately understand. Considering the data type, whether it's numerical, ordinal, nominal, discrete, continuous, interval, or ratio data, and the story of the visualization, the visualization types and color codes were determined.

Time series plots are a great method to visualize long ordered data. The time series plots in this study were separately created for the 10 years of historic weather data, 5 months of recent weather data, and 5 months of recent fire events data. The heat maps demonstrated the macro and micro view of bushfire events to help the viewer understand the timeline and distribution of events across the continent as well as areas of interest. The weather data visualizations show the long-term and short-term trends of each weather variable in consideration. We were able to gain insights from the monthly fire event areas with respect to the weather datasets. From the background study, we learned that the regions of the southeastern states of New South Wales and Victoria had been adversely affected during the bushfires events in 2019. We learned about the history of bushfire events in Australia and a few of the factors potentially responsible for the bushfires such as low rainfall, dry weather conditions, dry vegetation and an increase in overall temperature.

Since the purpose of analyzing the weather data was to gain insights on climate trends or patterns, those data are more valuable when they are visualized. Here, we also learned how to identify and handle the outliers in the time series plots. Charts and graphs make trends patterns easier to see, as well as allowing the interpretation of large scales of data possible. In addition to static images, we learned that interactive visualizations help gain attention from the audience by showing the dynamic data. With Tableau, we learned that creating a dashboard with multiple graphs was effective to tell a story and demonstrate our correlation analysis. Additionally, we experienced the application of a storyboard by showing various weather variables in our study as different segments in a story, both for the 10-year historical trend data as well as the 5-month interval during the 2019 Australia bushfires.

References

[1] Bom.gov.au. 2020. Climate Data Online. [online] Available at: <http://www.bom.gov.au/climate/data/> [Accessed 11 March 2020].

[2] CRC, T., 2020. Fire Fundamentals: Fire Weather. [online] Learnline.cdu.edu.au. Available at: <http://learnline.cdu.edu.au/units/env207/fundamentals/weather.html> [Accessed 11 March 2020].

[3] Kaggle.com. 2020. Fires From Space: Australia. [online] Available at: <https://www.kaggle.com/carlosparadis/fires-from-space-australia-and-new-zeland> [Accessed 11 March 2020].

[4] Kaggle.com. 2020. Rain In Australia. [online] Available at: <https://www.kaggle.com/jsphyg/weather-dataset-rattle-package> [Accessed 11 March 2020].

[5] The Verge. 2020. What You Need To Know About The Australia Bushfires. [online] Available at: <https://www.theverge.com/2020/1/3/21048891/australia-wildfires-koalas-climate-change-bushfires-deaths-animals-damage> [Accessed 10 March 2020].

[6] “NSW Suburb/Locality Boundaries - PSMA Administrative Boundaries,” 17-Feb-2020. [Online]. Available: https://data.gov.au/dataset/ds-dga-91e70237-d9d1-4719-a82f-e71b811154c6/details. [Accessed: 11-Mar-2020]

[7] K. Hennessy, C. Lucas* N. Nicholls* J. Bathols, R. Suppiah and J. Ricketts, “Climate change impacts on fire-weather in southeast Australia”, CSIRO and Bureau of Meteorology, 2005

[8] R. Dutta, A. Das and J. Aryal, "Big data integration shows Australian bushfire frequency is increasing significantly," Royal Society Open Science, vol. 3, no. 2, p. 150241, 2016.

[9] N. Popovich, D. Lu and B. Migliozzi, "Here’s Where Australia’s Destructive Wildfires Are Burning," The New York Times, 13 January 2020. [Online]. Available: https://www.nytimes.com/interactive/2020/01/02/climate/australia-fires-map.html. [Accessed March 2020].

Appendix

1. AUS_Bushfire_2019.ppkx : A packaged ArcGIS
2. AUS_WildFire5_Ginny.twbx : Fire Maps
3. NWS_3City_WeatherData.zip : Raw data of 2017 to 2019 Weather data of Sydney, Coffs Harbour, Newcastle
4. City3_WeatherData_2017_2019.xlsx : Cleaned data of NWS_3City_WeatherData.zip file in one excel.
5. fire_archive_V1_96617.csv : August - September fires
6. fire_nrt_M6_96619.csv : October - January fires
7. weatherAUS.csv : Raw data for Australia weather 2008 - 2017
8. timeline.twbx : time series plot of fires by date
9. weather_plots_FINAL_Ginny.twbx : Time series plots for weather variables for case studies