Wildfire in Australia and New Zealand

Monica Li and Stanley Nam

05/03/2020

Contents

Introduction

The wildfire in Australia is truly a tragedy. Every possible means should be taken to prevent and properly control such a painful disaster. Thanks to the cutting-edge observatory tools from NASA, we can access a detailed dataset (Paradis 2020) which provides a screenshot in the temporal development of this event. An analysis of the actual data and visualizing them can serve as a means to set up strategies for extinguishing and preventing a widespread fire like this.

This study analyses the satellite data from NASA to understand the influence of atmospheric factors on the intensity of a fire. Fire intensity, best represented by the core temperature, is a decisive factor in an efficient extinguishing plan. (Stephen J. Pyne 1996) The fire can be controlled more quickly if the core temperature is lower. Therefore, it is proper to devote limited resources to the time when the fire has a lower temperature. Among the factors that can influence the fire intensity exists atmosphere, but to the best of our knowledge, no prior research has investigated the influence of atmospheric factors in the Australian 'bushfire.' For this reason, the current research will analyze atmospheric factors represented as a categorical variable of day/night in our dataset.

Exploring the dataset

The dataset we are using (Paradis 2020) was obtained from Kaggle, uploaded by Carlos Paradis. The raw data from observation of the wildfire by two NASA Satellite Instruments were contained in the dataset. Out of the four .csv files they provided, we are limiting our scope to 'fire_archive_M6_96619.csv' only because it is the most reliable file with raw observatory data corrected by other data sources.

In this file, each row represents an observation of fire in Australia and New Zealand. Columns are for geographic information (latitude and longitude), the fire intensity (brightness, bright_t31, frp), specifications of the observation satellite instrument (scan, track, satellite, instrument, version), quality/confidence of the observation (confidence) details on the observation time (acq_date, acq_time) and other factors (daylight).

In the next subsections, we explored the dataset with plotting Correllogram and geometric diagram

Correllogram

From the correllogram below, we could see that there are some correlation between brightness and frp, scan and track. The color scheme in correllogram shows all positive correlations as blue, and all negative correlations as red.

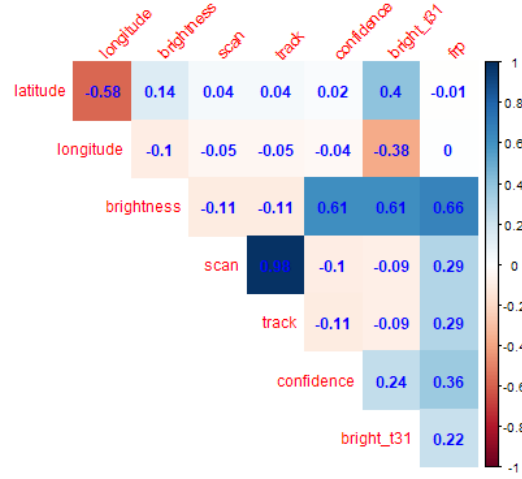**Correllogram for the numeric columns in wildfire dataset**

Figure 1: Correllogram for the wildfire data

Geometric with longitude and latitude

From the scatter plot map below, we could get a visual concept about the overall fire locations according to brightness and rediation power. The larger point on frp means higher fire rediation power, and a lighter point color, indicates a higher brightness of the fire.
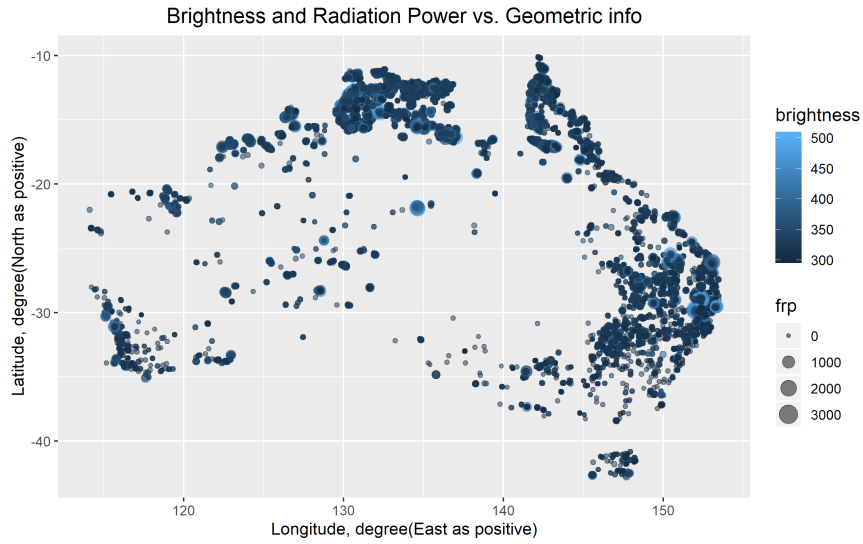

Figure 2: Geometric graph for the wildfire data

Methods

As mentioned in the introduction, the objective of our research is to understand the influence of the different atmospheric conditions between day and night on the intensity of wildfire. As for the dependent variable for the intensity of a fire, there are three in the satellite observation data: 'frp', 'brightness' and 'bright_t31.'

Among these variables, we select 'brightness' because there is a high correlation between 'frp' and 'brightness' and 'brightness' and 'bright_t31' as we see in the correlogram.

We ran a multiple linear regression analysis to understand whether the atmospheric difference between day and night (variable 'day/night') influences the intensity of wildfire (variable 'brightness'). We include two additional variables, 'scan' and 'track' which are expected to act randomly. These are variables that represent the specifications of the observation satellite instrument, so not necessarily related to the intensity. If the regression model with 'scan', 'track' and 'day/night' fits better than the model without 'day/night', namely including only two dummy variables, we can conclude that 'day/night' is a significant factor that decides the intensity of a fire.

As shown in the code chunk below, our `model` contained variable 'day/night' in addition to two dummy factors, and the null model (`model_n`) did not. These two models were compared.

```
model <- lm(brightness ~ scan + track + daynight, data = obsData)
model_n <- lm(brightness ~ scan + track, data = obsData)
signific <- anova(model_n, model, test ="Chisq")
```

Results

Significance of the variable 'day/night'

Track and daynight variables in `model` were significant as shown below.

```
tidy_model <- broom::tidy(model)
tidy_model
```

```
## # A tibble: 4 x 5
##   term        estimate std.error statistic  p.value
##   <chr>          <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept)   348.       1.46     239.   0.
## 2 scan            1.14     0.601      1.89 5.83e- 2
## 3 track         -14.1      1.97      -7.15 9.02e-13
## 4 daynightN     -18.5      0.222    -83.5  0.
```

An increase in the value of 'track' lowers the temperature by -14.0987860773409. The variable 'day/night', on the other hand disproportionally changes the temperature by -18.4910464875325, meaning that the observed fire temperature during the daytime is higher than that during night time by -18.4910464875325.

However, we cannot decide whether 'day/night' is significant or not by looking into this result alone. Therefore, we introduced a benchmark, namely another model `model_n`.

```
tidy_model_n <- broom::tidy(model_n)
tidy_model_n
```

```
## # A tibble: 3 x 5
##   term        estimate std.error statistic   p.value
##   <chr>          <dbl>     <dbl>     <dbl>      <dbl>
## 1 (Intercept)   340.       1.59     214.    0
## 2 scan            0.491    0.656      0.748 0.454
## 3 track          -9.90     2.15      -4.60  0.00000430
```

As reported above, 'scan' and 'track' were only two variables that explain 'brightness' in this model and 'track' is still significant with the p-value of 4.29569671834778e-06.

```
glance_model <- broom::glance(model) %>%
  add_column(model = "model", .before = 1)
glance_model_n <- broom::glance(model_n) %>%
  add_column(model = "benchmark model", .before = 1)
```

```
model_comp <- bind_rows(glance_model, glance_model_n)
model_comp
```

```
## # A tibble: 2 x 12
##   model r.squared adj.r.squared sigma statistic  p.value    df  logLik    AIC
##   <chr>     <dbl>         <dbl> <dbl>     <dbl>    <dbl> <int>   <dbl>  <dbl>
## 1 model    0.172         0.172  17.3     2493. 0.            4 -1.54e5 3.07e5
## 2 benc~    0.0118        0.0117 18.9      215. 1.45e-93      3 -1.57e5 3.14e5
## # ... with 3 more variables: BIC <dbl>, deviance <dbl>, df.residual <int>
```

As we can see from the comparison between two models above, the adjusted coefficient of determination (adjusted R-squared) value increases from 0.0118033144143415 to 0.171979247465454, meaning that having another predictor 'day/night' increase the fittedness of the model.

Another support for having the 'day/night' predictor comes from the anova chi-square test.

```
signific
```

```
## Analysis of Variance Table
##
## Model 1: brightness ~ scan + track
## Model 2: brightness ~ scan + track + daynight
##   Res.Df      RSS Df Sum of Sq  Pr(>Chi)
## 1  36008 12836448
## 2  36007 10755799  1   2080649 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Effect sizes

The two graphs presented below show the effect sizes of 'scan' and 'track'. What is important from them is the fact that the variables show significant gaps by daytime (in scarlet) and night time (in tortoise).
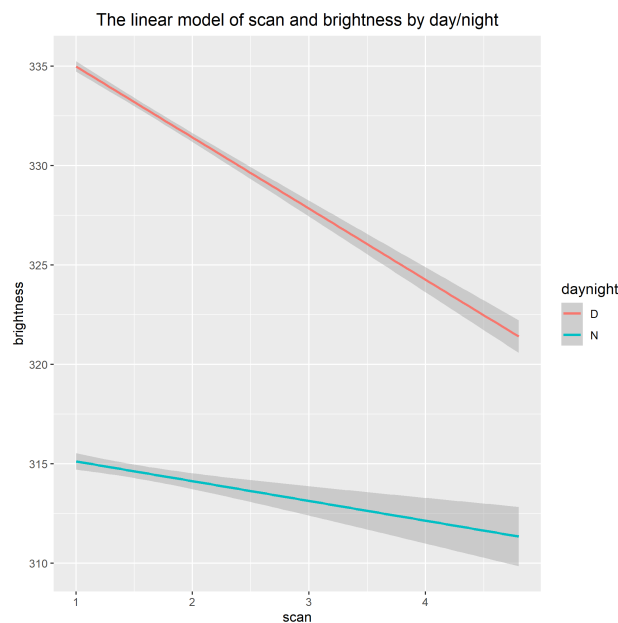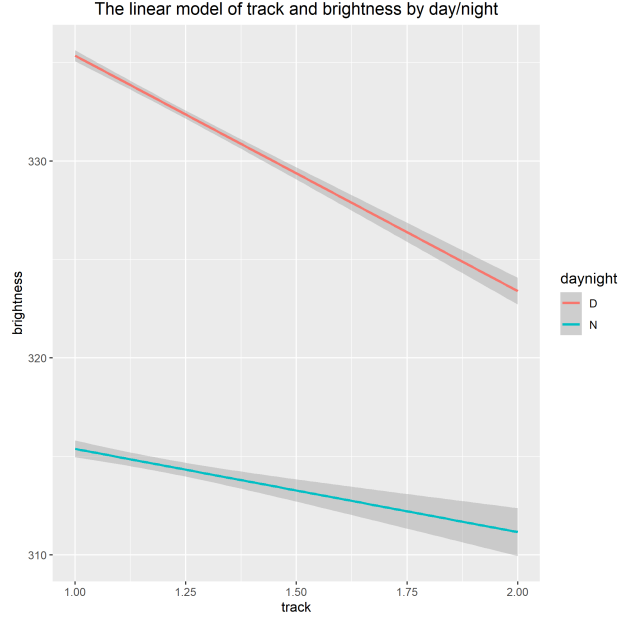

Figure 3: Scan vs. Brightness

Figure 4: Track vs. Brightness

Finally, the three subgraphs below show individual effect size. We can find that the effect of day/night is greater and clear (narrower whisker) than the other two variables.
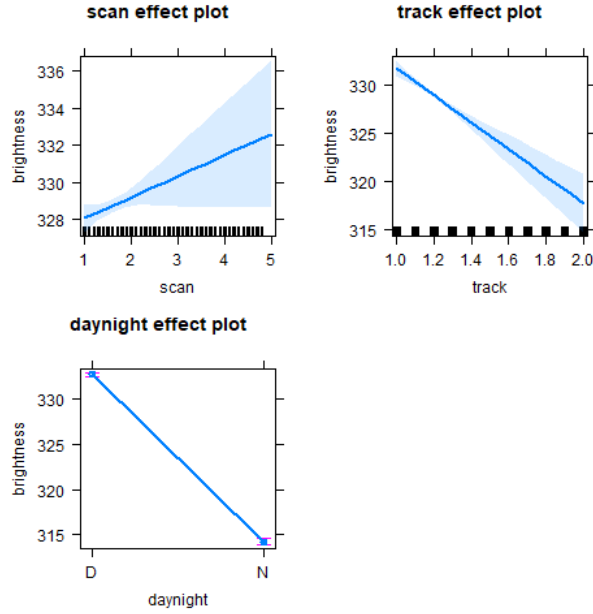

Figure 5: Individual effect size

Discussion and conclusion

According to the finding described in the Results, the atmospheric differences between daytime and night have a statistically significant effect on the intensity of a fire. Assuming that the atmospheric temperature decreases after the sunset, this may influence the core temperature of a fire as well. As widely known,

extinguishing the fire when the core temperature is lower is more effective and successful. Besides, the risk that personnel directly involved in the extinguishing tactics would strikingly decrease if the temperature is lower.

Our study suggests that in the limited resources, extinguishing strategy with a more active concentration in the night time can be more effective.

References

Paradis, Carlos. 2020. "Fires from Space: Australia Nasa Satellite Data Modisc6 and Viirs 375m from 2019-08-01 to 2020-01-11." Dataset.

Stephen J. Pyne, Richard D. Laven, Patricia L. Andrews. 1996. *Introduction to Wildland Fire*. Wiley.