

Capstone Two – Final Project Report

Problem Statement:

In recent years wildfire occurrence within the continuous US and associated destruction have become increasingly concerning. The increased rate and duration of recent wildfires requires properly distributed resources to effectively combat them. The goal of this capstone project is to develop a predictive model of wildfire occurrence and destruction based on recent trends. Generating a predictive model of wildfire occurrence could assist in fighting or preventing future wildfires.

Data Wrangling:

The data set for this project has been collected from the Fire Information for Resource Management System (FIRMS). The data covers 10 years of recording from January 2010 to January 2020. The data set is requested from the FIRMS system for a specified range of time and provided in CSV format. The data source also contains information for wildfires around the world enabling expansion of the dataset.

Exploratory Data Analysis:

The data set contained multiple parameters that could be used as the focus of a timeseries analysis. The columns provided are as follows:

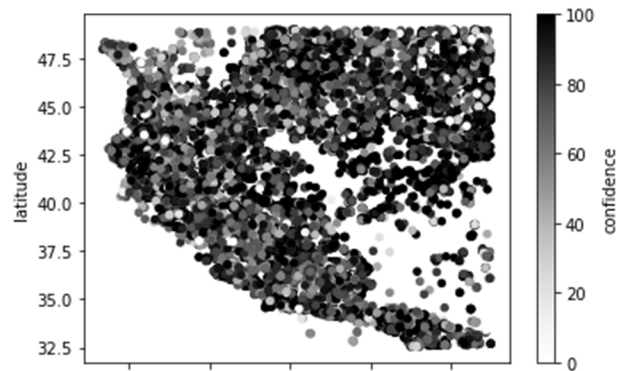
- Longitude
- Latitude
- Brightness – temperature K
- Scan
- Track
- Acq_Date
- Acq_Time
- Satellite – Aqua and Terra
- Confidence
- Version –software version when collected
- Bright_T31 – temperature K
- FPR – Fire Radiative Power
- DayNight

Through exploration of the data set Brightness was identified as a key variable. Here the brightness of a fire refers to its strongest point and for the purposes of modeling a single brightness point will denote a single fire. A secondary complication of the brightness value is present in the data set due to the collection method. Each point is recorded from an image of a single point in time. Because of this it is possible that multiple bright points are recorded with the same timestamp.

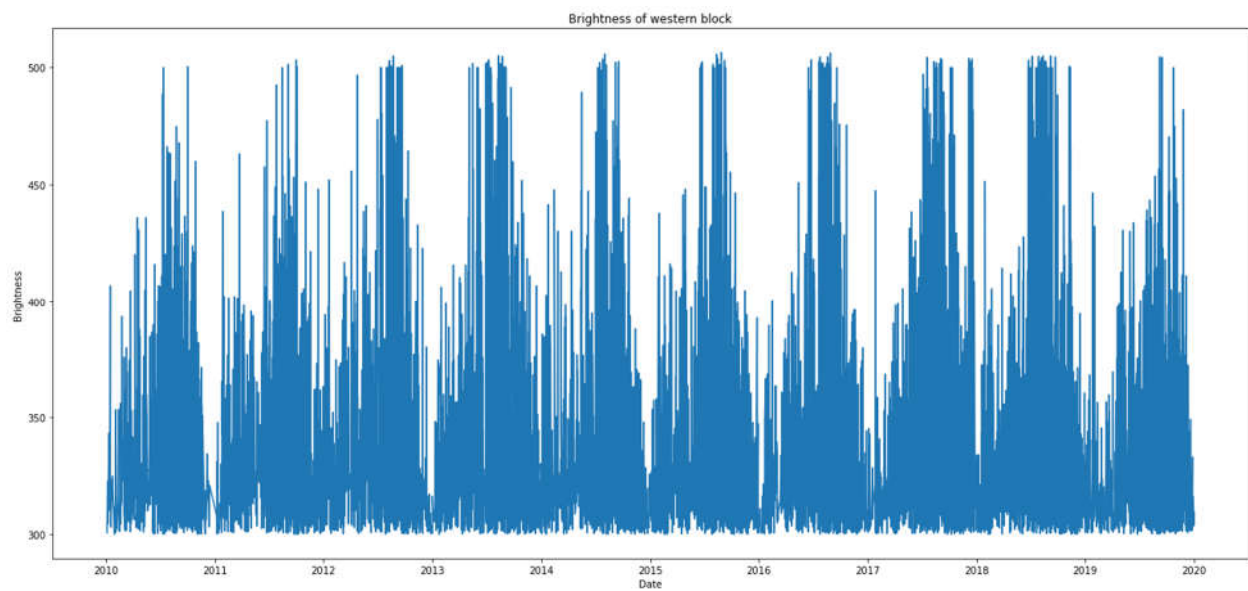
As a result of multiple brightness timestamps the dataset was down sampled to the highest brightness value for a single timestamp. A future improvement to the dataset is restructuring the dataset to follow individual fires.

Analysis:

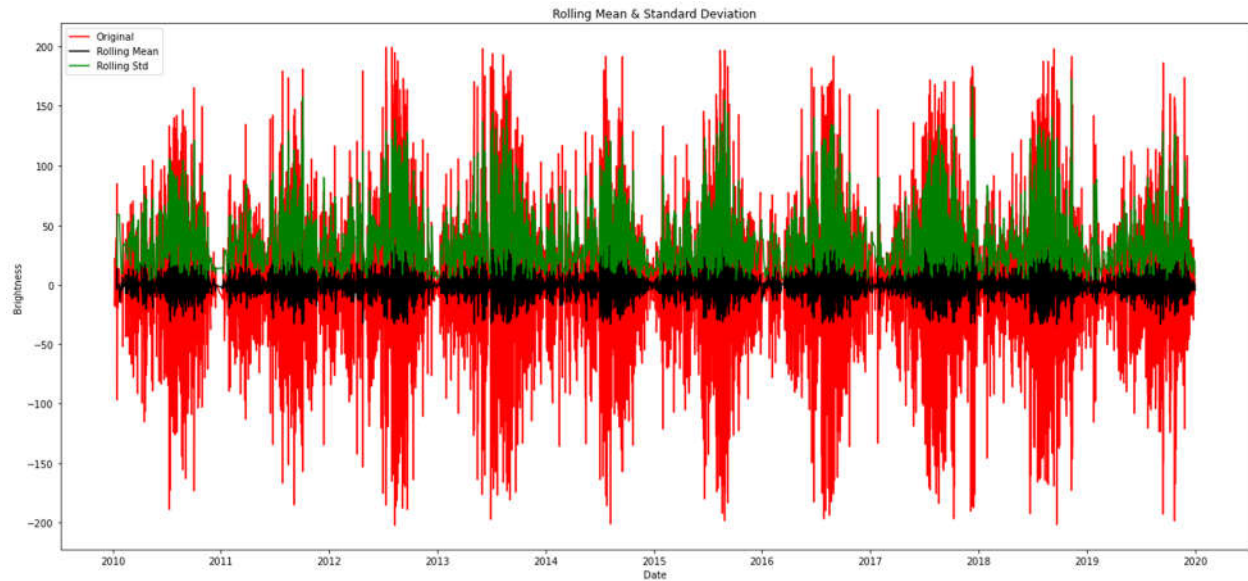
Down sampling of the initial data set was done in multiple steps. First a block of the western United States was filtered from the data set.



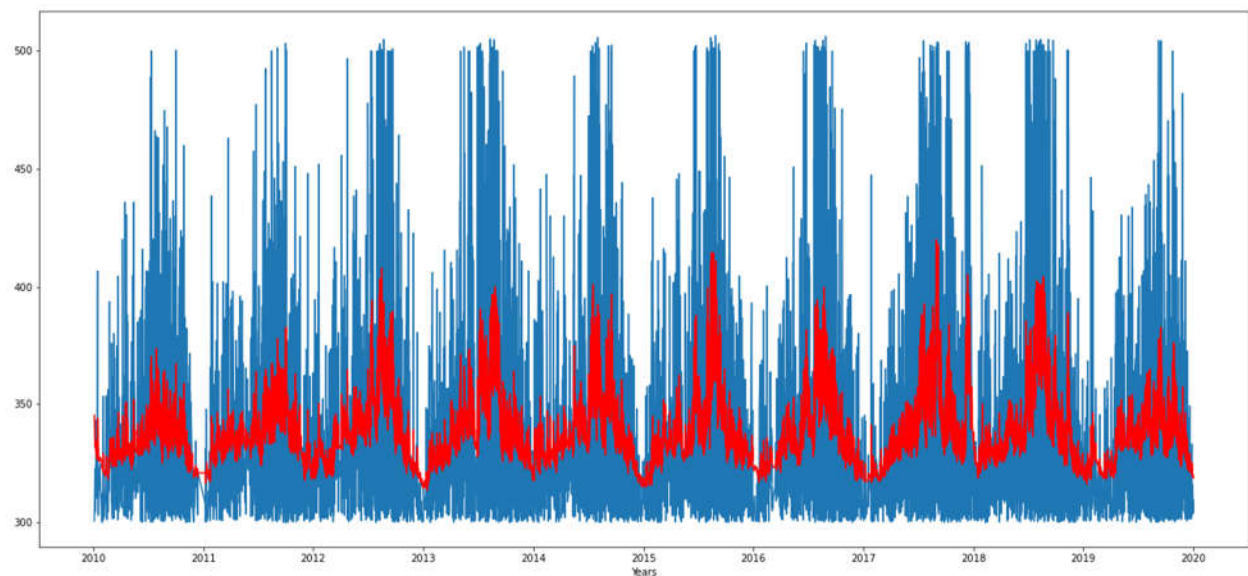
Checking the down sampled data set of the western block did not develop any substantial gaps in the time series.



Following selection of the data set it was determined that the data was not stationary. Correcting the dataset so that it was stationary was completed through using the difference.



Next the model output was examined in comparison to the recorded brightness values.



Results:

From the model comparison peak values are not accurately predicted. While the model follows the apparent trend brightness values additional work is required. Therefore, current conclusions indicate that while the model does not fit the dataset it acts as a clear starting point.

Future Work:

To improve the current model additional work in simplifying the timeseries is necessary. Additionally, the individual parameters provided, and combinations of parameters should be examined for use as variables in the timeseries.

A more accurate method of geographic selection should be determined with the future goal of selecting individual states, counties, or parks.

Working through this dataset provided numerous exciting diversions. While complications in tracking individual fires was beyond the scope of the current project it provides a clear continuation of the research. Additional research in fire spread can assist in connecting various brightness values as a single fire and would allow for the forecasting of individual fires.