

Effects of Injection Wells on Seismic Activity

W. Keil & W. Cranford

October 27, 2018

Changepoint Analysis

Uses: Used to detect significance changes in the mean or variance of data.

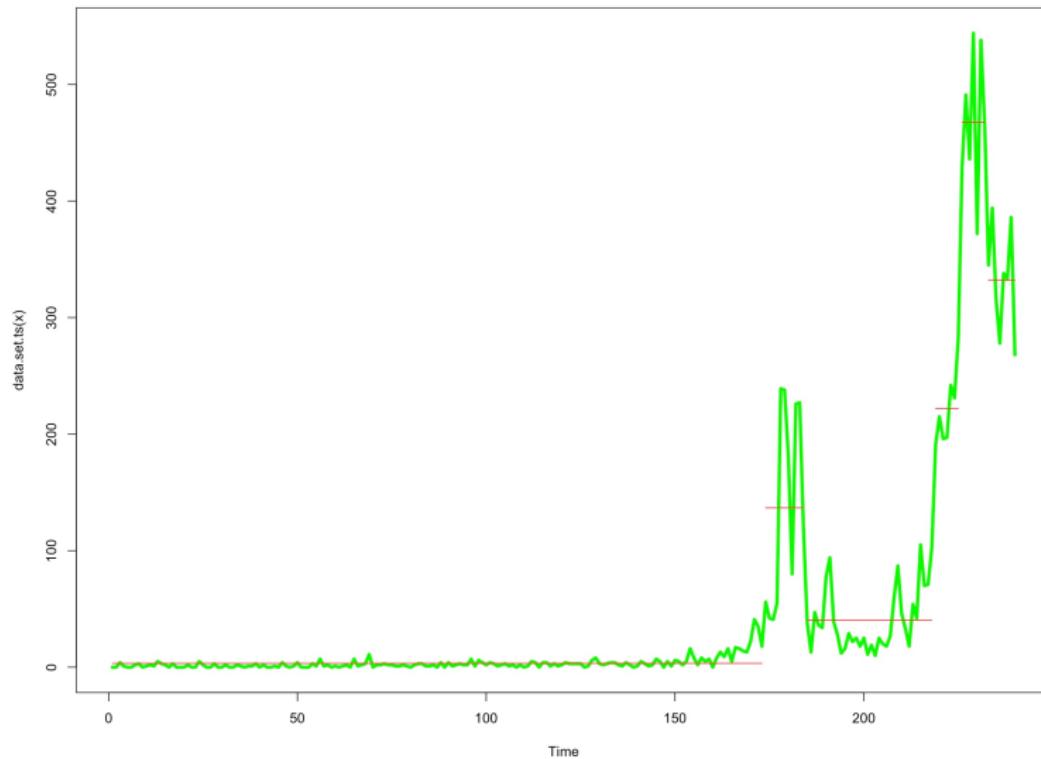
Method: Involves bootstrapping cumulative sum charts to detect change in mean or variance.

Strengths Changepoint analysis can detect significant changes in means or variances when they are not apparent or they are hard to detect by other methods.

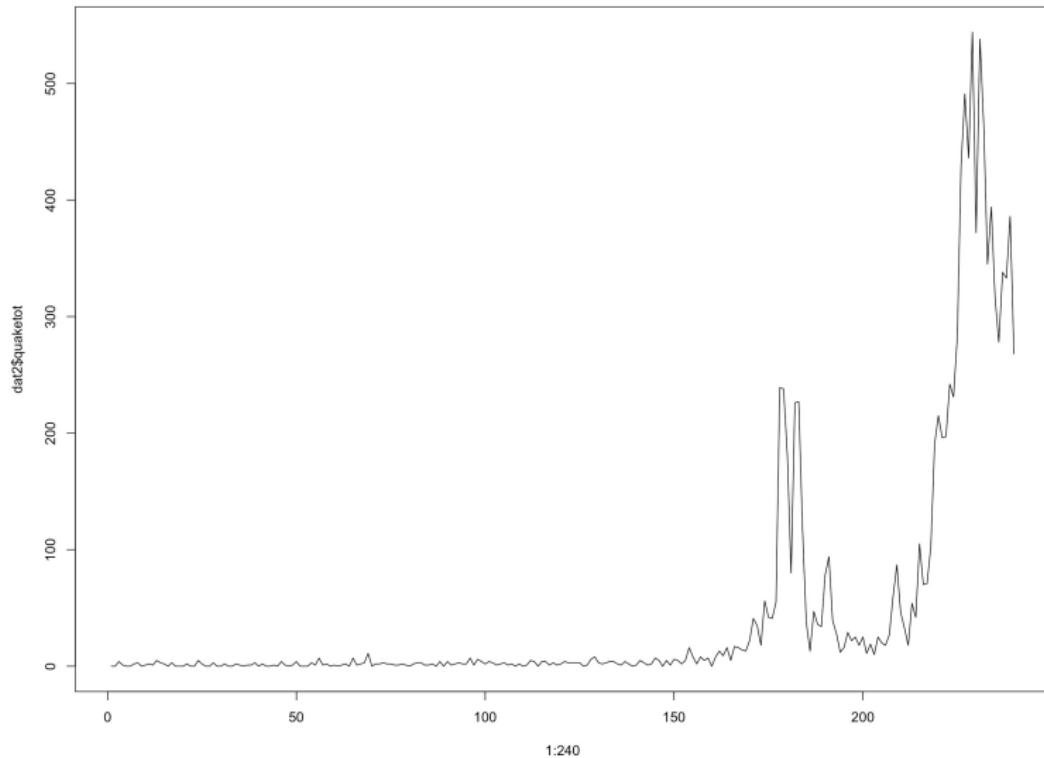
Results: Changepoint analysis detected change in the mean of the total number of quakes per month at months: 173 184 218 225 232

This translates to June 2010, May 2011, March 2014, October 2014, and May 1 2015.

Changepoint Analysis



Total Earthquakes per Month, 1996 - 2015



Arima Model

Model: Three parts: AutoRegressive(p) Differences(d) MovingAverage(q).

$$AR(p) = y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t$$

Used 8th degree auto regressive function

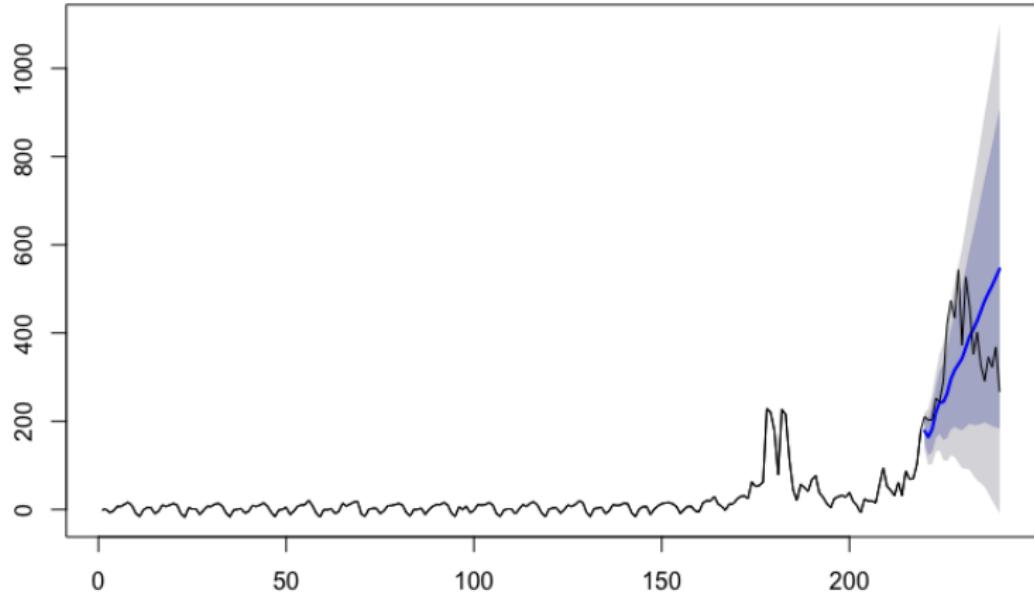
2nd degree differencing Zero order for Moving Average

$$AIC=1894.99 \quad AICc=1896.55 \quad BIC=1935.33$$

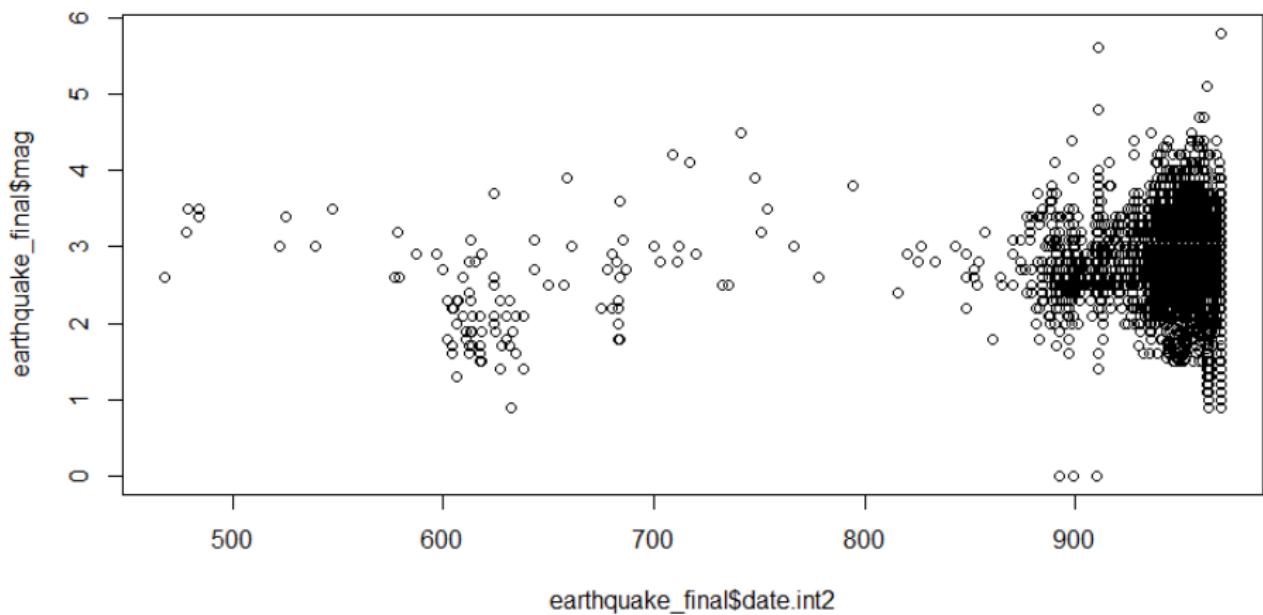
$$\phi_1 = -0.7843, \phi_2 = -0.8428, \phi_3 = -1.0816, \phi_4 = -0.4525$$

$$\phi_5 = -0.4248, \phi_6 = -0.4053, \phi_7 = -0.0802, \phi_8 = -0.0802$$

Arima Forecast Results



Number of Earthquakes Over Time



Magnitude of Earthquakes Over Time

December 1974 - January 2008:

Average of 1.96 Earthquakes > 2.5 magnitude per year.

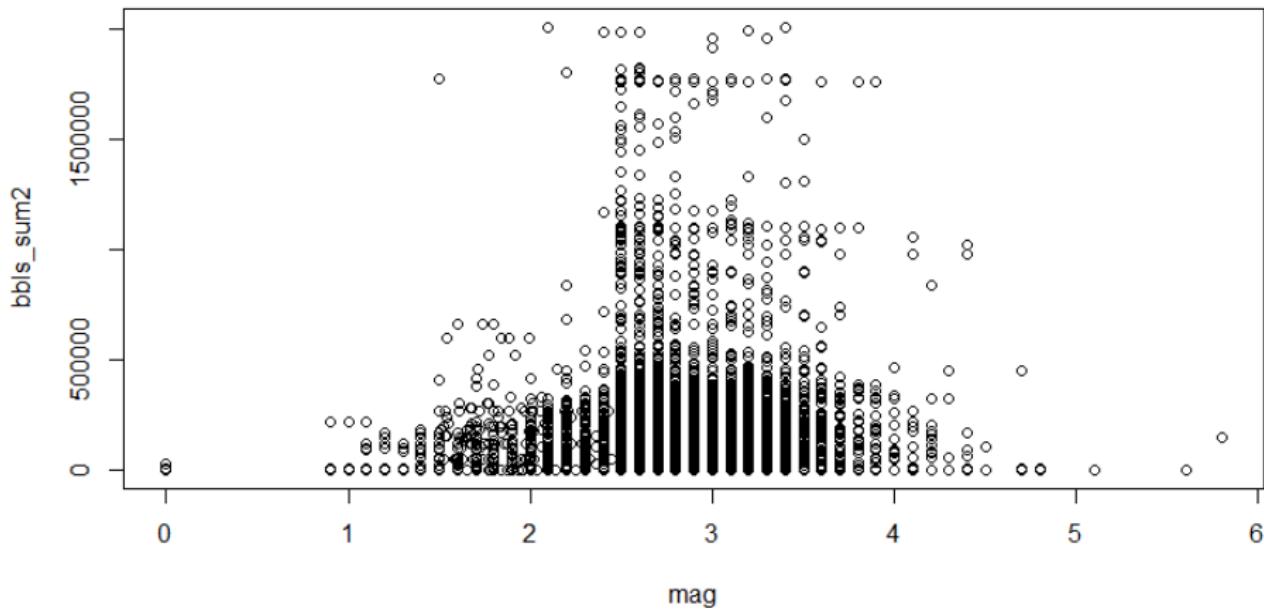
Average magnitude of all earthquakes in this period is 2.51.

February 2008 - September 2016:

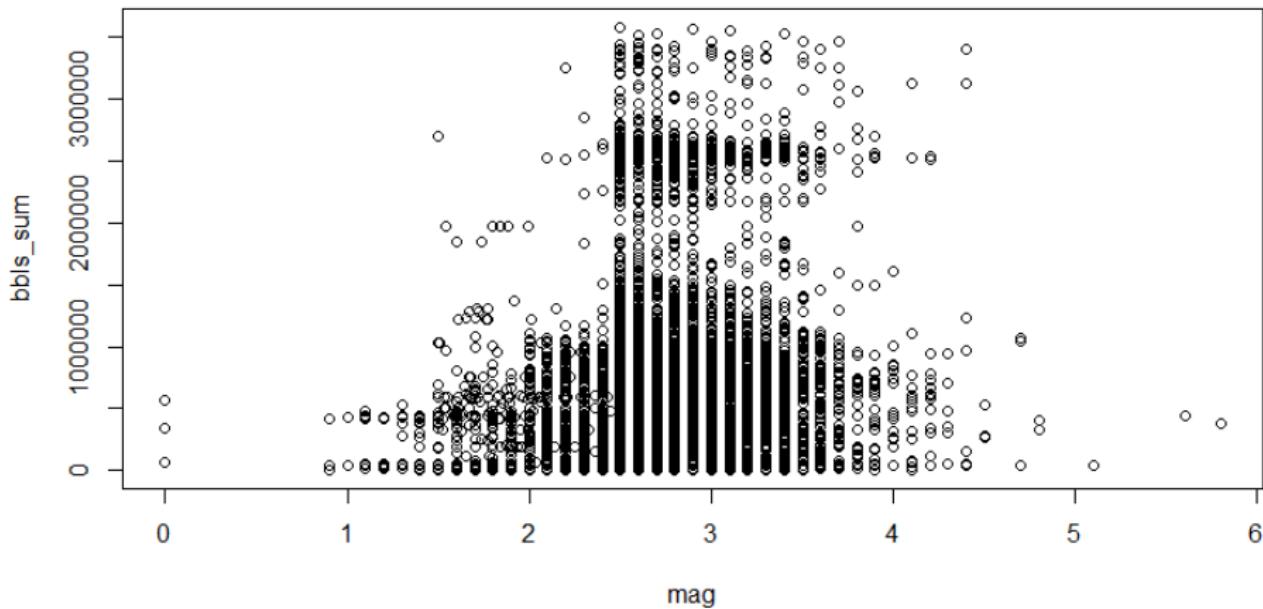
Average of 623.3 Earthquakes > 2.5 magnitude per year.

Average magnitude of all earthquakes in this period is 2.74.

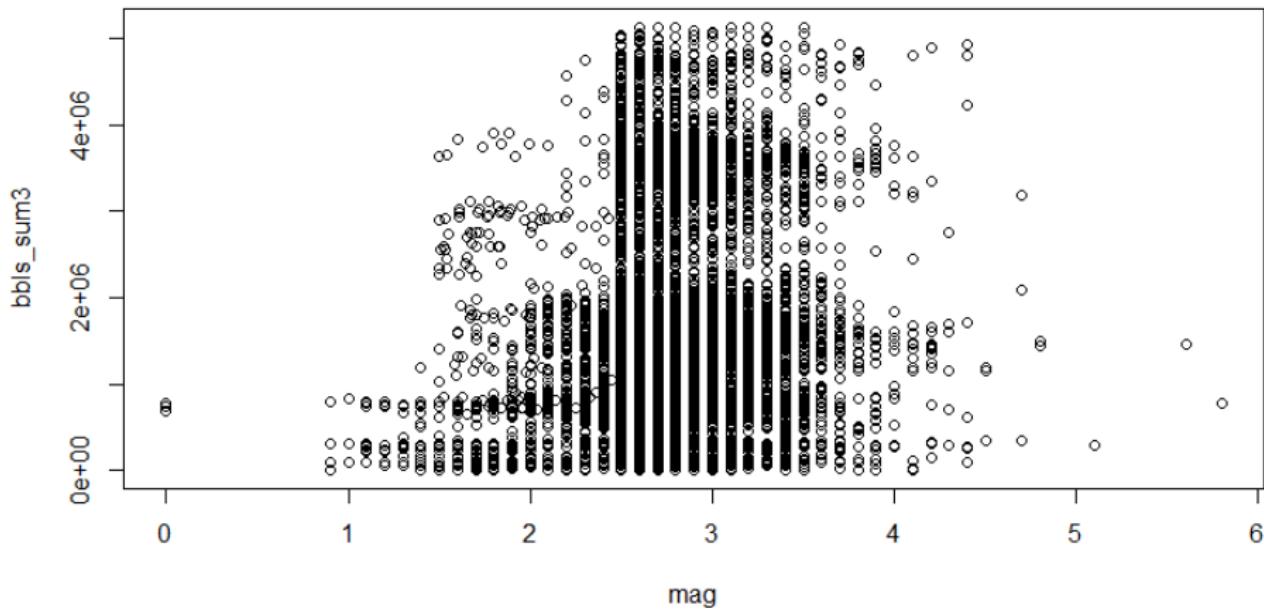
Total BBLS vs MAG for 5 Mile Radius



Total BBLS vs MAG for 10 Mile Radius



Total BBLS vs MAG for 15 Mile Radius



Model Results

Linear Model Relations

Evidence of nonzero correlation between bbls and magnitude.

Random Forest Predictions

Able to predict magnitude with good accuracy.

Significance of Results

Injection Activity Regulation

Locational Data Sets

Objective:

Try to predict the locations of earthquakes with magnitude of 2.5 or higher.

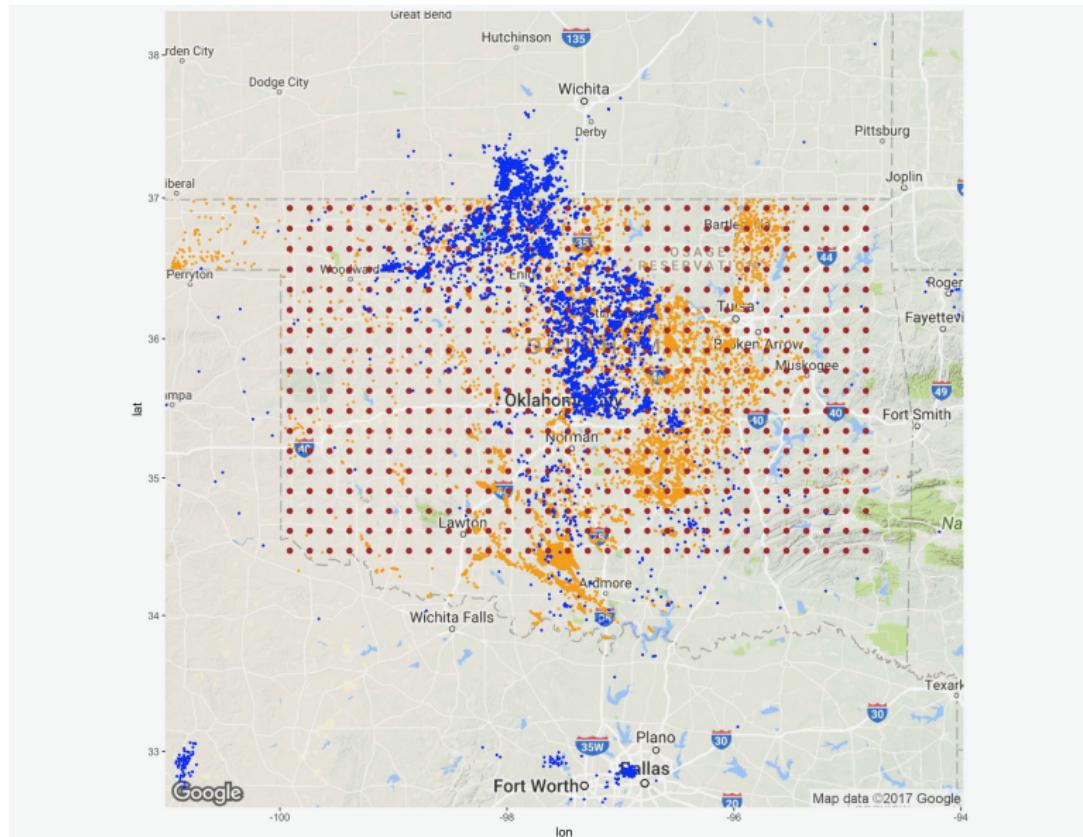
Data:

Twenty data-frames, each one representing one year of data.

The columns contain all relevant drilling and seismic information.

Each row represents a different location in Oklahoma.

Overview of Location Data Sets



Meta Ensembling

Model:

Meta Ensembling - a.k.a. Model Stacking.

A well establish technique.

Involves running multiple different models and using the results as inputs for a different model.

Motivation:

Originally ran random forest model with interesting but uninterpretable results.

Decided to take the prediction values of a random forest model and use them in inputs for a SVM model.

SVM Stacked Model

Tune (svm, catbigquake - rfpred, numwells, sumbbls)

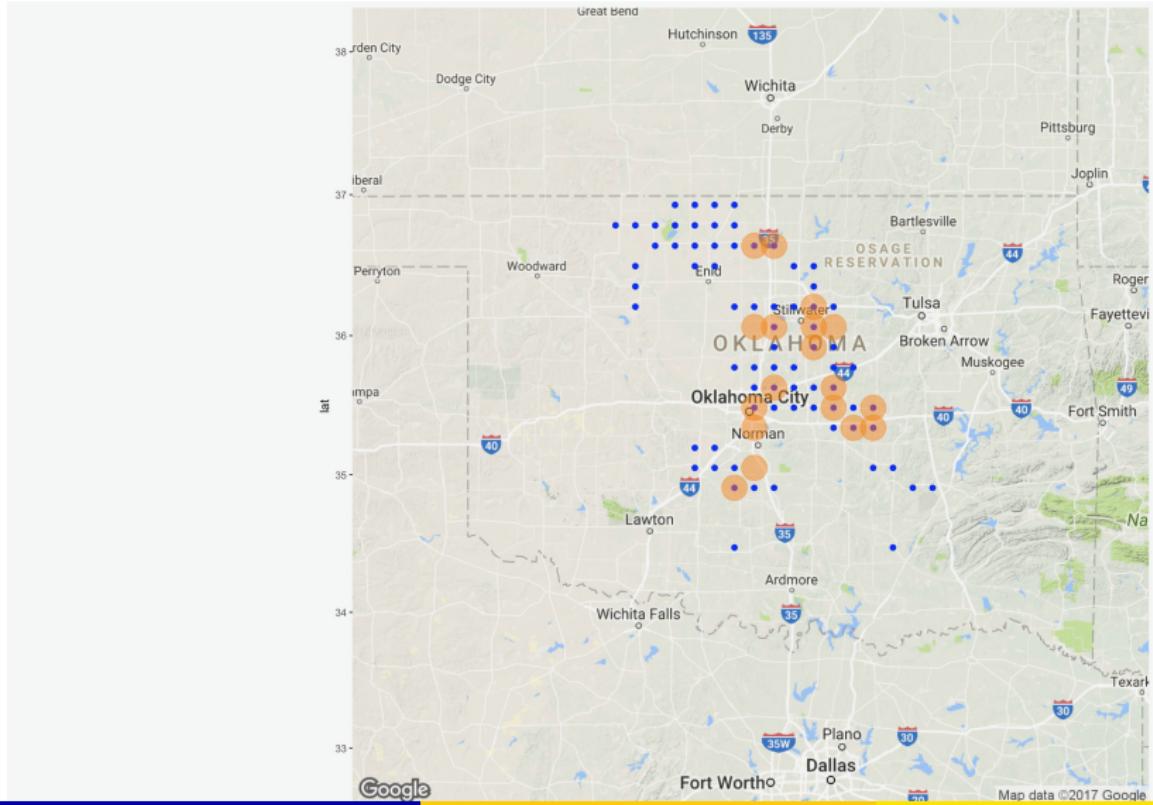
catbigquake - The categorical variable of whether a location will have at least one 2.5 magnitude earthquake in a given year.

rfpred - The prediction result of random forest model trying to predict the average magnitude of earthquakes in each location.

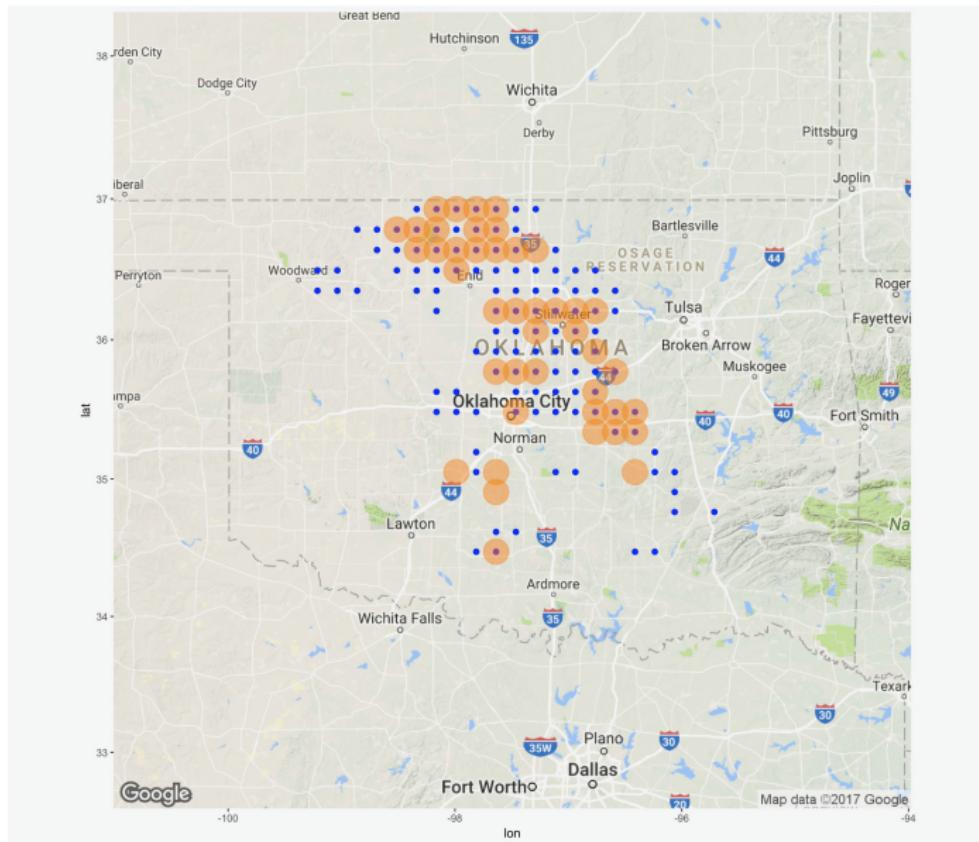
numwells - The number of wells drill per location in a given year.

sumbbls - The sum of the number of barrels of fluids pumped into the ground per year.

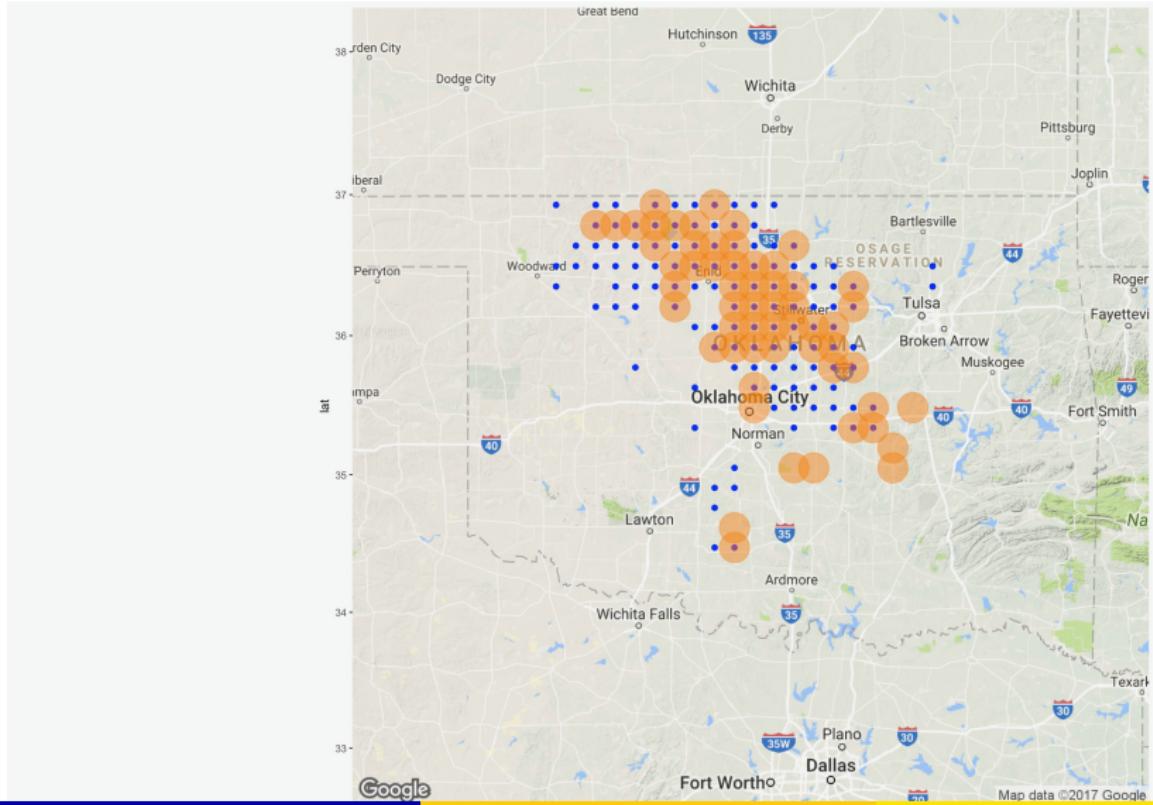
SVM 2013 - Error Rate 11.48 %



SVM 2014 - Error Rate 15.19 %



SVM 2015 - Error Rate 14.07 %



SVM Results- Confusion Matrices

2013 2014 2015

$$\begin{bmatrix} 464 & 58 \\ 4 & 14 \end{bmatrix} \begin{bmatrix} 421 & 76 \\ 6 & 37 \end{bmatrix} \begin{bmatrix} 413 & 69 \\ 7 & 51 \end{bmatrix}$$

Error Rates 11.48 % 15.19 % 14.07 %