Emerging Hot Spots - Earthquakes, California

Uma Maheshwari Dasari
UEP 235 : Advance GIS

1. INTRODUCTION

An earthquake is a natural occurrence caused by the movement of seismic tectonic plates beneath the surface of the Earth. Every year, the USGS (United States Geological Survey) detects around half a million earthquakes around the world. An average of 20,000 people are killed as a result of the earthquake. Predicting possible earthquake locations, particularly in earthquake-active areas or seismic hot spots, is critical for disaster preparedness and mitigation.

The west coast of US is one such earthquake-active areas because of presence of fault lines. As per USGS, California(CA) has the most earthquakes that can cause more damage than any other state in the US. There are 15700 known faults in CA, out of which 500+ are active faults, Fig.1. There is a greater than 99% probability of one or more earthquakes of magnitude 6.7 striking in CA[3].

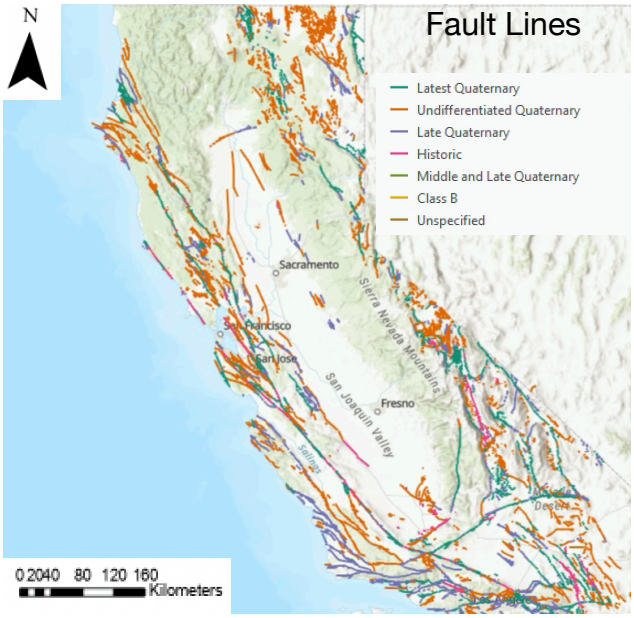

Fig.1 Faults in California region.

This paper going to be a study on an approach—emerging hot spot analysis (EHSA)—to demonstrate new, consecutive, intensifying, persistent, diminishing, sporadic, oscillating, and historical hot and cold spots in CA state.In this study, an earthquake hot spot is defined as a location with the highest earthquake risk or significant earthquake activity within an area with low or normal seismic activity. To investigate such predictions, spatial time pattern mining is often utilized. The goal of the paper is to find the variations in results of spatial-temporal analysis change with the increase in number of years in data analysis. Finally, I will also answer how far are the hotspots from the fault lines?

2. Literature Review

[1] Haque, U., Blum, P., Da Silva, P. F., Andersen, P., Pilz, J., Chalov, S. R., ... & Keellings, D. (2016). Fatal landslides in Europe. *Landslides*, *13*(6), 1545-1554.

This paper discussed a Spatio-temporal distribution of deadly landslides in 27 European countries from 1995 to 2014. The authors used the Emerging hot-spot analysis method to find the new hotspot, consequent and sporadic hot spots in Europe along with the upward trends in landslide analysis and concluded how far they are from human settlement and the risk of landslides affecting. One of the research questions includes finding trends in spatial and temporal distribution and hotspots of landslides that caused casualties.

[2] Hussain, M. S., Goswami, A. K., & Gupta, A. (2022). Predicting pedestrian crash locations in urban India: An integrated GIS-based spatiotemporal HSID technique. *Journal of Transportation Safety & Security*, 1-34.

This paper also discussed space-time cube and emerging hotspot analysis for predicting crash hotspots along urban streets in Kolkata, India.Research question on spatial analysis is trends in pedestrian crash.

3. Data Collected

3.1 Earthquake Catalog - USGS

The seismic data can be downloaded from the USGS Earthquake catalog in a CSV file. The latitude and longitude taken for CA state are from 42 to 32 and -124 to -118.5, as shown in the Fig.2. The data consists of latitude, longitude,Time and magnitude etc. parameters. Fig.3 shows the sample of dataset that is obtained. I have collected data for each year from 2012 to 2021.

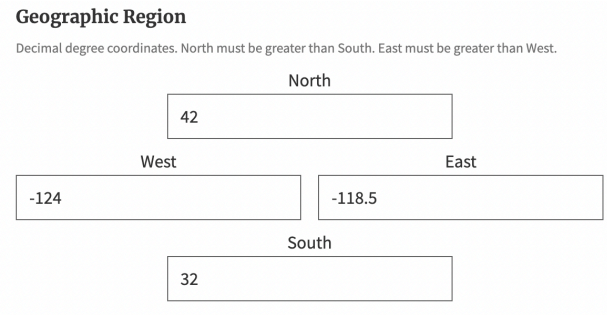

Fig. 2 snapshot of data collection region

time	latitude	longitude	depth	mag
2021-05-07T23:59:43.950Z	39.1275	-120.0276667	6.1	2.64
2021-05-07T07:27:01.830Z	36.1218333	-120.699	5.67	2.55
2021-05-07T04:37:38.190Z	39.467	-120.3195	10.8	3
2021-05-07T04:35:14.150Z	39.461	-120.3145	10.12	4.65
2021-05-07T04:26:54.470Z	39.4668333	-120.3223333	12.39	2.74
2021-05-07T04:26:31.770Z	39.4648333	-120.3168333	9.41	3.14
2021-05-06T21:52:24.570Z	38.8366667	-122.8025	2.42	2.53

Fig. 3 Snapshot of data sample

3.2 Fault lines

I have download the fault data of the USA from ArcGIS online. Link : https://services2.arcgis.com/FiaPA4ga0iQKduv3/arcgis/rest/services/Earthquake_Faults_and_Folds_in_the_USA/FeatureServer. The data download in lines vector format from the living atlas provided online by ArcGIS updated in May 2022 and owned by federal user community.

4. Methodology

4.1 *Data Merging*

The first step of this project I did is merging 1 year data into 5 years and 10 years data using simple python code as shown below,Fig. 4

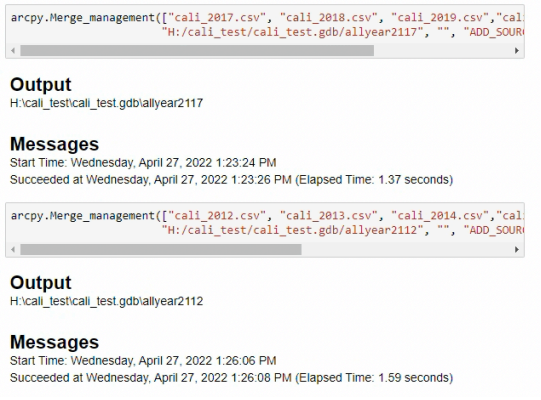

Fig.4 Snapshot of the python code

4.2 *Space-Time Pattern Mining*

Second step in the project is done using Emerging Hot spot analysis, first for 5 years (2017 to 2021) and then for 10 years (2012 to 2021). To use the Emerging hot spot analysis tool I first used 'create space time code by aggregating points' geo-processing tool and set the time step interval parameters to default to create netCDF file.Then 'visualized the space time cube in 3D' to get the time series graph of count of earthquake events vs the month. Finally used 'Emerging Hot Spot Analysis' tool, which uses Getis Ord Gi* statistics to get the 8 different patterns for hot and cold spots respectively. I set number of neighbor time step interval to 2 and my analysis resulted as in Fig. 5 and Fig. 6

Fig.5 Screen shot of Emerging Hot-spot analysis message for 5 years data

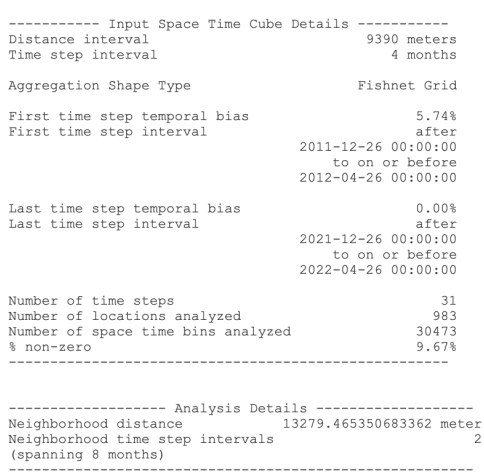

Fig.6 Screen shot of Emerging Hot-spot analysis message for 10 years data

4.3 *Summary statistics*

Final step, I have done summary statistics by using 'near' geo-processing tool. I have found the mean, median and standard deviation of the different patterns of emerging hot spot analysis with respect to the distance from the fault lines.

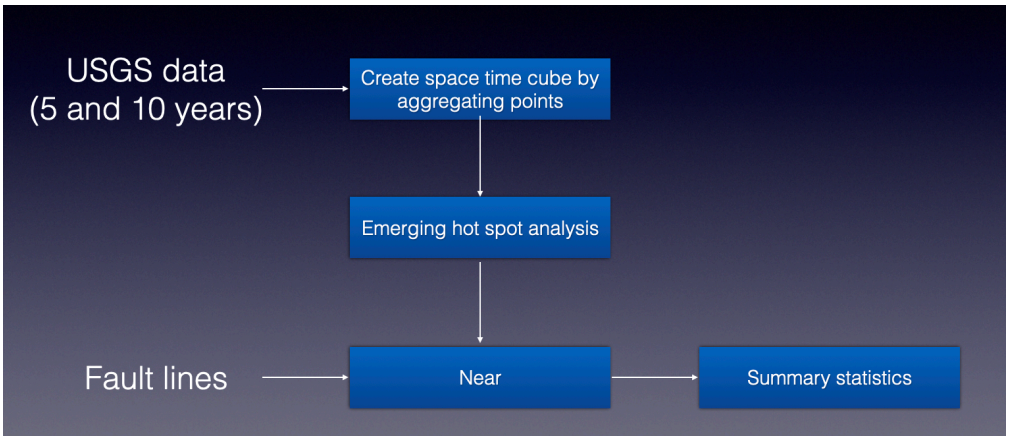

Fig 7. Flow chart of methodology

6. Errors/Warnings/Limitations

- My initial approach of this project is to do the suitability/vulnerability assessment of earthquakes but I changed my direct of project because I couldn't find the data for my project. So I realized the data we collected helps us to find the direction of project.
- I have first visualized my space time cube in 3d and then continued my emerging hot spot analysis in 3D local scene which led me to warning of excessive draw requests. I could have simply avoided it using 2d map instead of 3d local scene.
- The emerging hot spot analysis created two different new hot spot regions for 5 years and 10 years data. This is because of the highest number of event counts regions varies with the 5 year and 10 year. The new emerging hot spot regions are depending only on the counts of events hence the inconsistency in the analysis.
- Since I can't validate my hot spot regions my summary statistics also might be flawed.

7. Concluding Thoughts

EHSA has found the hot spot areas and found the new hotspots which are inconsistent and changes with the data. I can't validate my results and can't use this results for disaster preparedness and mitigation. So I think suitability analysis is better approach to find risk prone areas. Since the suitability analysis includes more geophysical parameters though it has its own limitations of inconsistency in weights arrangements.

8. References

[1] Haque, U., Blum, P., Da Silva, P. F., Andersen, P., Pilz, J., Chalov, S. R., ... & Keellings, D. (2016). Fatal landslides in Europe. *Landslides*, *13*(6), 1545-1554.

[2] Hussain, M. S., Goswami, A. K., & Gupta, A. (2022). Predicting pedestrian crash locations in urban India: An integrated GIS-based spatiotemporal HSID technique. *Journal of Transportation Safety & Security*, 1-34.

[3] https://www.earthquakeauthority.com/California-Earthquake-Risk

[4] https://storymaps.arcgis.com/stories/30df3b23b8984751986042fec53becea