

Assignment for Mega Project-1 (Sub-Group 9) CSIR Summer Research Training Program 2020

Aditya Sengupta *

Amity Institute of Applied Sciences, Amity University, Noida

05/08/2020

Questions -

Question 1 -

What do you mean by b-value as a seismological aspect? Write the Gutenberg-Richter Relationship and describe its importance in seismology.

Question 2 -

Briefly sketch the Geological settings of the Andaman and describe its Seismic history.

Question 3 -

Illustrate the b-value of the study area and carve out the potential hazardous region on the basis of spatial and temporal variations of b-value

Answers

Answer 1-

One of the basic seismological parameters used to describe an ensemble of earthquakes is the *b-value* in the famous Gutenberg-Richter frequency magnitude relationship relationship/law.

It characterizes the distribution of earthquakes over the observed range of magnitudes. The *b-value* is usually 1.0, but for shorter time windows it varies significantly depending on the tectonic features of the seismically active region (Prasad and SINGH 2015) (Pacheco, Scholz, and Sykes 1992). In order to understand the concept of b-value in terms of seismology, we first need to understand the Gutenberg Richter law and its significance -

Gutenberg-Richter Law ~

In seismology, the Gutenberg-Richter law (GR law) expresses the relationship between the magnitude and total number of earthquakes in any given region and time period of at least that magnitude. (Gutenberg 1956)

The corresponding Gutenberg-Richter Frequency Magnitude Relation is given by,

*adityasenguptasai99@gmail.com

$$\log_{10} N = a - bM$$

Or

$$N = 10^{a-bM}$$

Where, N is the number of events having a magnitude $\geq M$, also called the cumulative Frequency, and a and b are constants which are always same for a given value of N and M .

The significance of the GR law lies in the values of these constants. The relationship between earthquake magnitude and frequency is fairly common, but the values of a and b vary significantly from region to region giving seismologists significant amount of information regarding the seismo-tectonics/geo-tectonics of the region under study. When we say that the parameter b i.e. the *b-value* is close to the value 1.0 (in seismically active regions) we mean to say that for a given frequency say 4.0 or larger events there will be 10 times as many magnitude 3.0 or larger quakes and also 100 times as many magnitude 2.0 or larger quakes. But the *b-value* can vary significantly from 0.5 to 2.0 or even higher depending upon on the region's seismic activity. A good example of this is Earthquake swarms which can have a *b-value* as high as 2.5 which indicates a high proportion of small quakes to large ones.

The GR Law and specifically the measurement and interpretation of the spatial and temporal variations in *b-values* can provide significant information about the evolution of the seismicity of the region under study. By analysing anomalies in the spatio-temporal variation in the *b-value* of a region with respect to major quakes we can obtain spatio-temporal constraints that can be used to create, modify and/or improve earthquake forecasting models for that particular region as done by Nuannin et. al. (Nuannin, Kulhanek, and Persson 2005) for the Andaman-Sumatra region.

According to some recent work, the major anomalies that lead to these spatio-temporal variations have been:

- the stress applied to the material,
- the depth, the focal mechanism,
- the strength heterogeneity of the material and the proximity of macro-failure.

These variations in *b-value* decrease were observed prior to the failure of rock samples which were deformed in the laboratory and it leads to the suggestion that this is a precursor to major macro-failure and ultimately large magnitude earthquakes. (Smith 1981)

Now, a deviation of *b-value* from the value 1.0 can also be due to some inconsistency in the data set, i.e. it might be incomplete or it might contain errors propagated in the calculation of magnitude.

There is an apparent *b-value* decrease for smaller magnitude event ranges in all empirical catalogues of earthquakes. This effect is described as “roll-off” of the *b-value*, a description due to the plot of the logarithmic version of the GR law becoming flatter at the low magnitude end of the plot. This may in large part be caused by incompleteness of any data set due to the inability to detect and characterize small events. That is, many low-magnitude earthquakes are not catalogued because fewer stations detect and record them due to decreasing instrumental signal to noise levels. Some modern models of earthquake dynamics, however, predict a physical roll-off in the earthquake size distribution. (Bhattacharya et al. 2009)

Now, the *a-value* represents the total seismicity rate of the region. This is more easily seen when the GR law is expressed in terms of the total number of events:

$$N = N_{TOT} 10^{-bM}$$

where $N_{TOT} = 10^a$, is the total number of events. Also it is interesting to note that since 10^a is the total number of events, 10^{-bM} is the probability of those events.

Answer 2-

Geological Setting of the Andaman Region

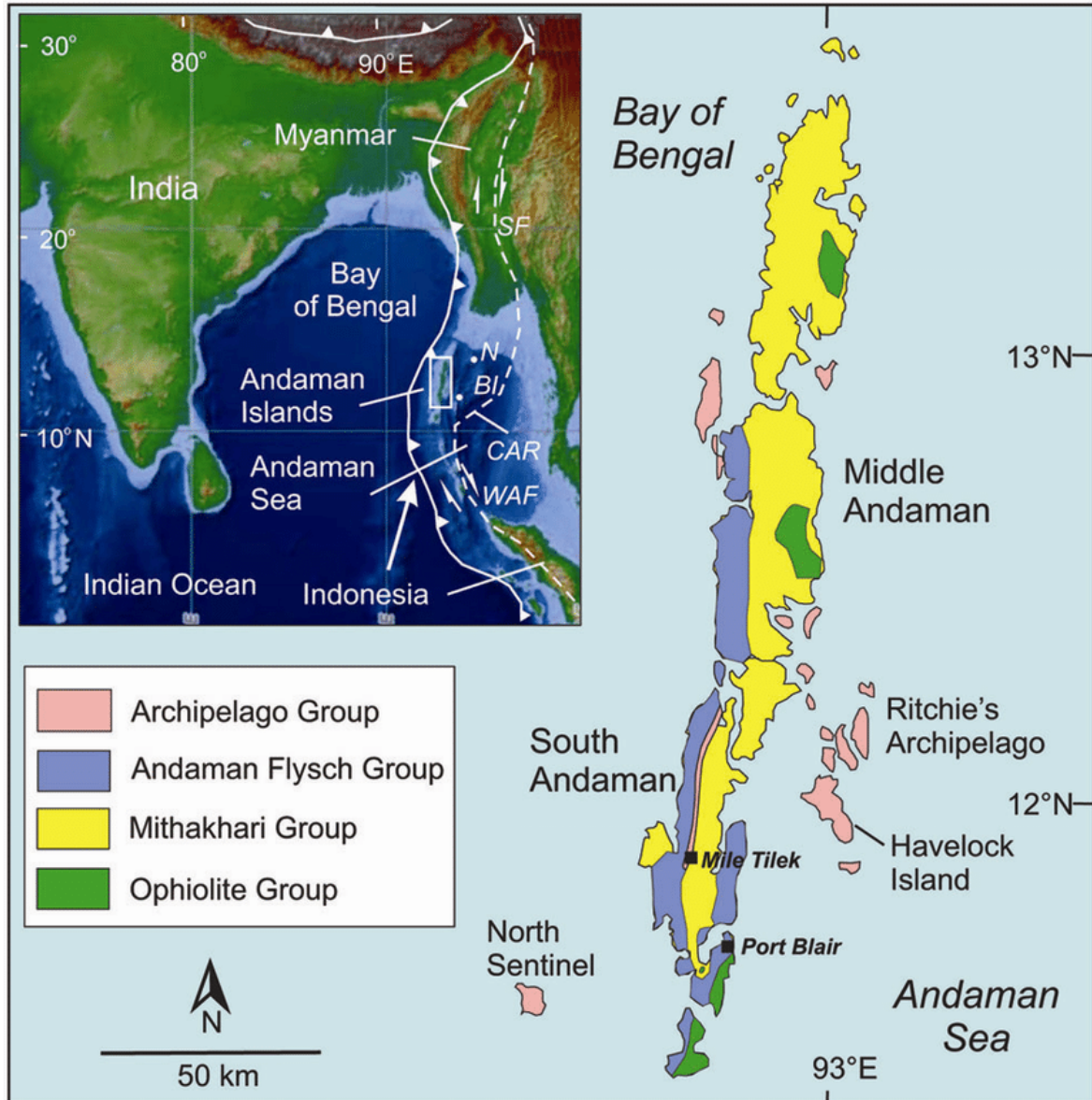


Figure 1: Geological map of the Andaman Islands

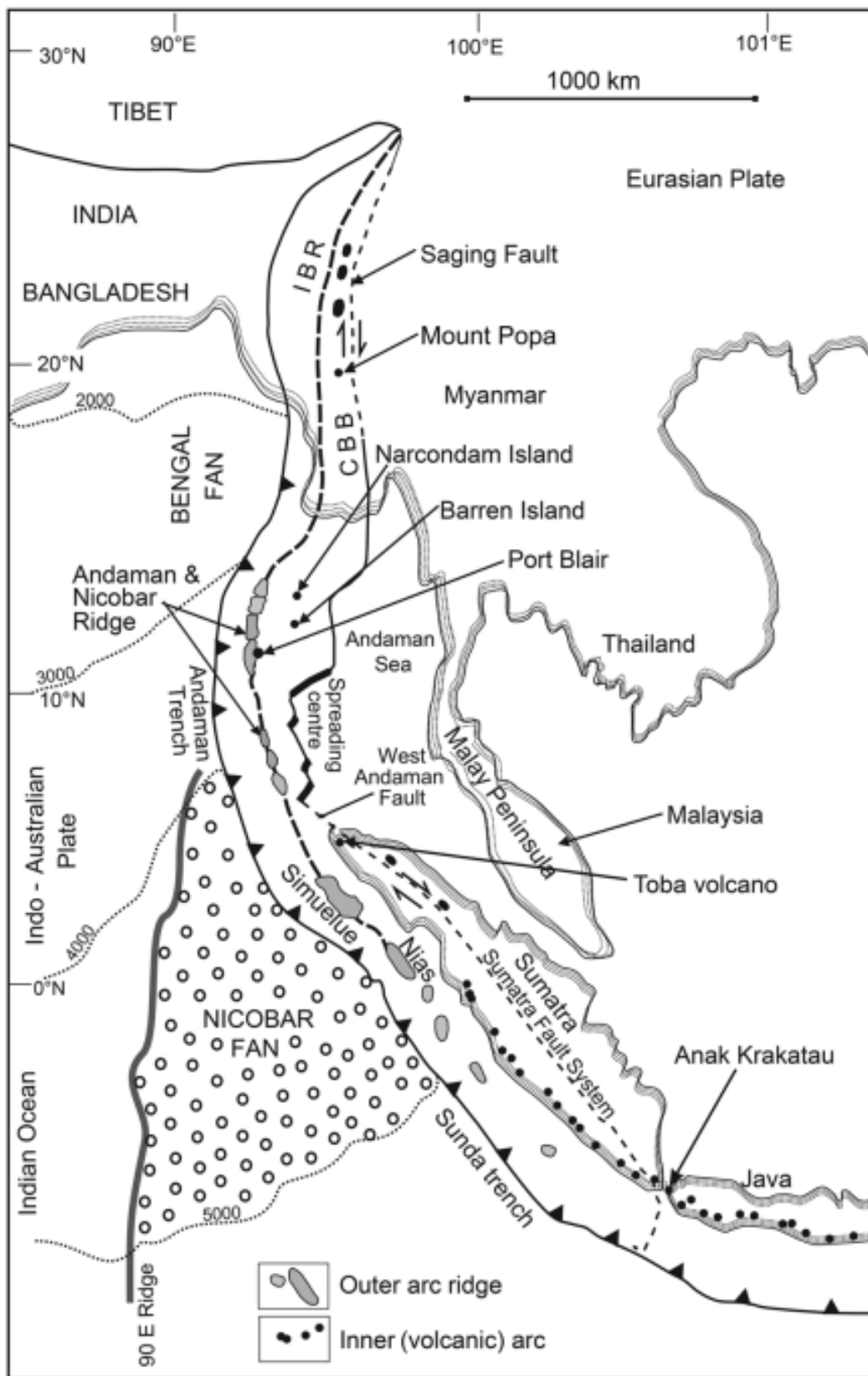


Figure 2: Tectonic map of the Western Sunda Arc

Figure 2 shows the major present-day morphotectonic elements including Andaman–Nicobar accretionary ridge(modified after Curray 2005)

The Andaman Region is an active, complex backarc basin lying above and behind the Sunda subduction zone where convergence between the overriding Southeast Asian plate and the subducting Australian plate is highly oblique. The tectonics and geological history of the Andaman Sea cannot be separated from the tectonics and geological histories of Myanmar (Burma) on the north, the Andaman and Nicobar Islands part of the accretionary prism on the western side of the Andaman Sea, and Sumatra on the south.and Nicobar Islands part of the accretionary prism on the western side of the Andaman Sea, and Sumatra on the south.(Curray 2005)

The Andaman–Nicobar archipelago that forms the western margin of the Andaman Sea is a sediment-dominated accretionary wedge (outer-arc island) associated with a convergent margin tectonic setting. The Andaman accretionary ridge consists of two stratigraphically and structurally distinct terranes, juxtaposed and telescoped into a north–south-trending high-relief fold-thrust belt formed along the obliquely subducting eastern margin of the Indo-Australian oceanic lithospheric plate. (Bandopadhyay and Carter 2017) The geology and structure of the ridge reflect the complexity of the evolving tectonics and stratigraphy of an accretionary wedge. (Curray 2005)

The Andaman– Nicobar Islands are part of a 5600 km long curvilinear belt of accretionary ridges and outer-arc islands that stretches from Sumba in Eastern Indonesia to Western Burma in the north associated with the subduction of Indo-Australian oceanic lithosphere below an overriding Eurasian (Sunda) Plate

In an accretionary setting new material is progressively accreted to the thin outer end of the accretionary wedge, while the oldest and thickest inner part of the accretionary wedge near the inner trench wall may be thickened further, increased in volume by underplating and consequently uplifted and exposed to form a chain of fold-thrust belt islands

Seismic History of the Andaman region

- Earthquake History

The map given in the Figure 3 on page 6 shows all historically recorded earthquakes in the Andaman & Nicobar Islands from the year 1980 to 2020. Two prominent red circles visible to the north-east of Port Blair, i.e. in the top half of the map, are where the only volcanoes in India are located. The northernmost of the is the Barren Island Volcano. This volcano has erupted in recent times. The other volcano is called Narcondam and is dormant. The Andaman and Nicobar Islands are located near the boundary of the Indian plate and the Burmese Microplate. The Andaman Trench marks this boundary and lies in the Bay of Bengal to the west of the archipelago. Another prominent feature is the north-south West Andaman fault which is strike-slip in nature and lies in the Andaman Sea, to the east of this island chain. The Andaman Sea, just like the Atlantic Ocean, is presently being widened by a tectonic process called “Sea Floor spreading”. This is taking place along undersea ridges on the seafloor. The Indian plate is diving beneath the Burmese Microplate along the Andaman Trench in a process known as “Subduction”. Shallow and occasional intermediate-depth earthquakes delineate the subducted slab under the Andaman-Nicobar islands joining the seismicity trend of the Indo-Burman ranges. However, it must be stated that proximity to faults does not necessarily translate into a higher hazard as compared to areas located further away, as damage from earthquakes depends on numerous factors such as subsurface geology as well as adherence to the building codes.

Figure 4 on page 7 is a scatter plot of the cumulative data set with the y-axis indicating the magnitude and the x-axis showing the index number of the corresponding magnitude. Three of the highest magnitude earthquakes having $M_b = 6.3$ and 6.2 have been marked with the respective dates on the graph.

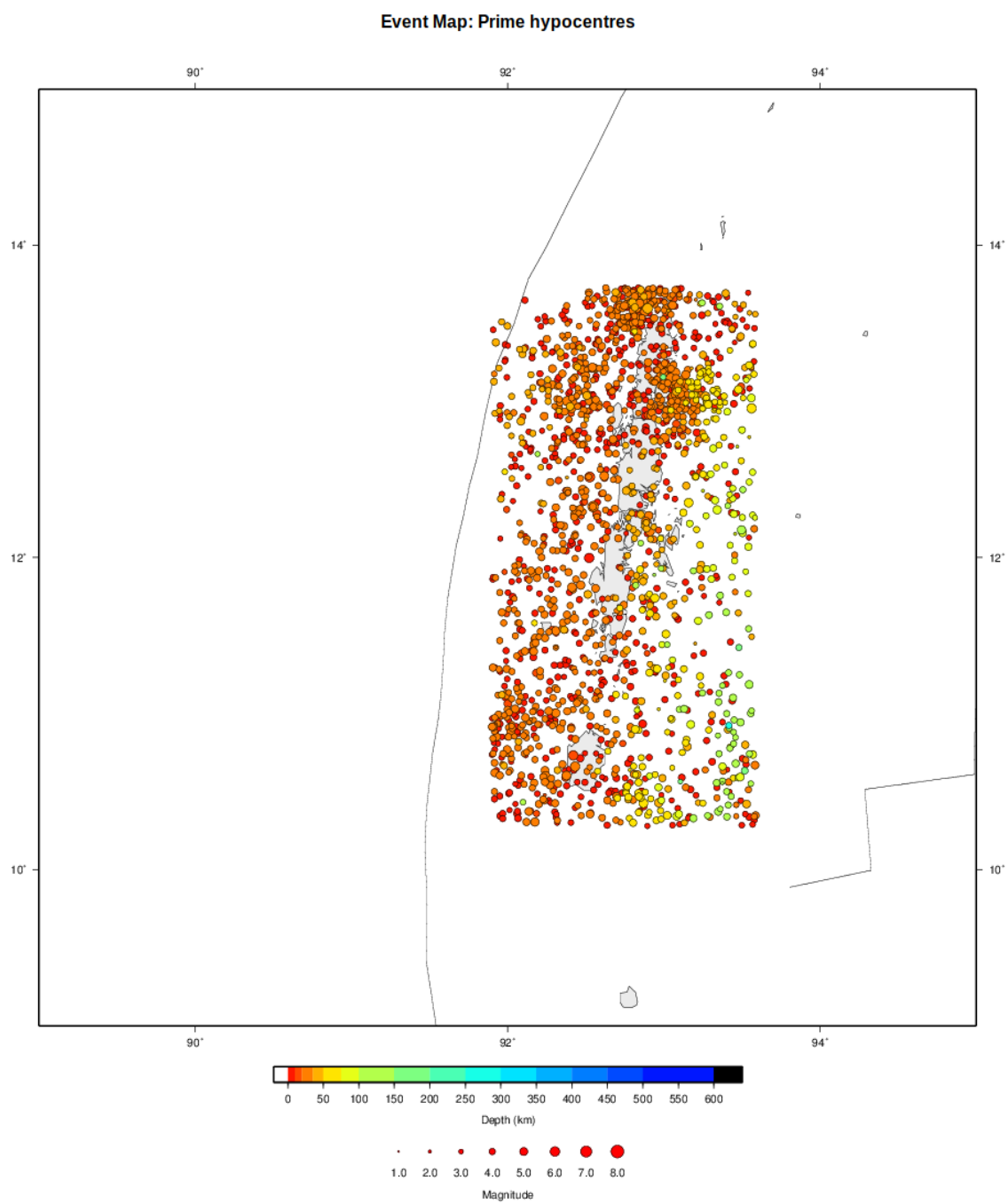


Figure 3: Andaman event map

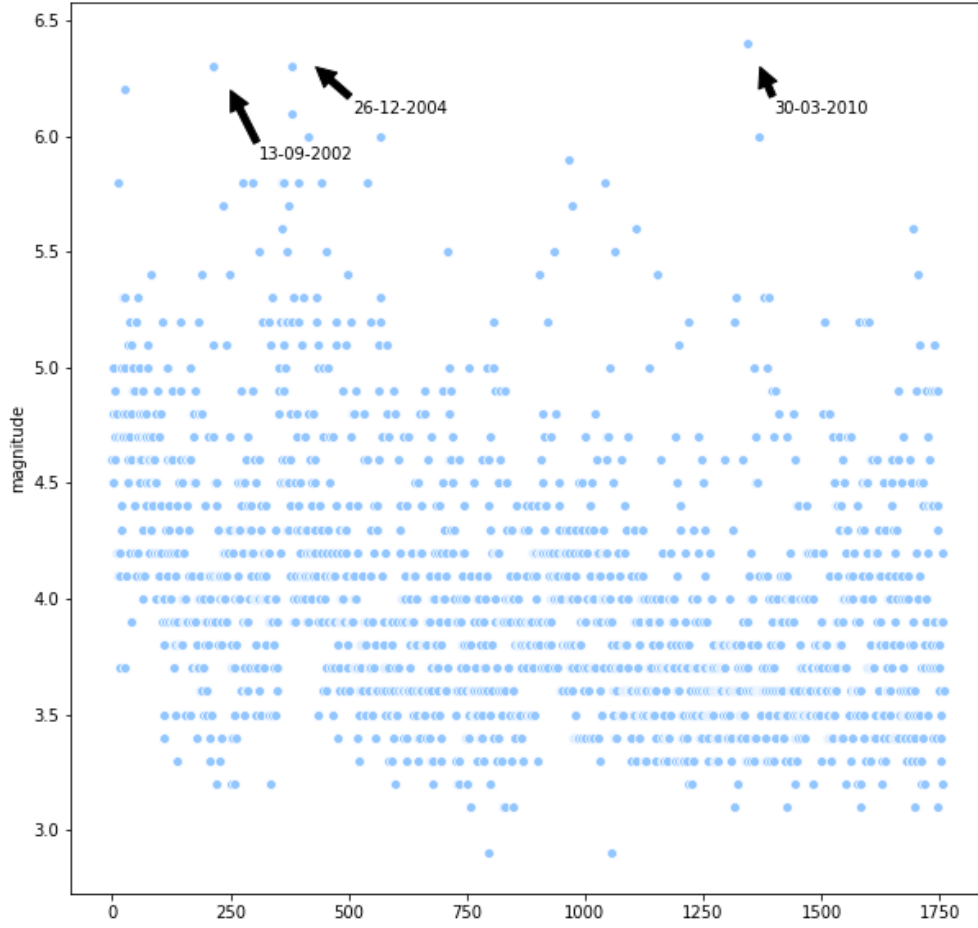


Figure 4: Andaman Overall earthquakes

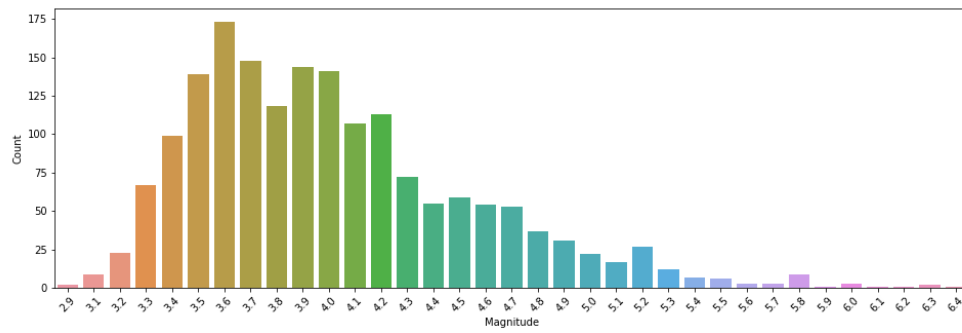


Figure 5: Magnitude Count

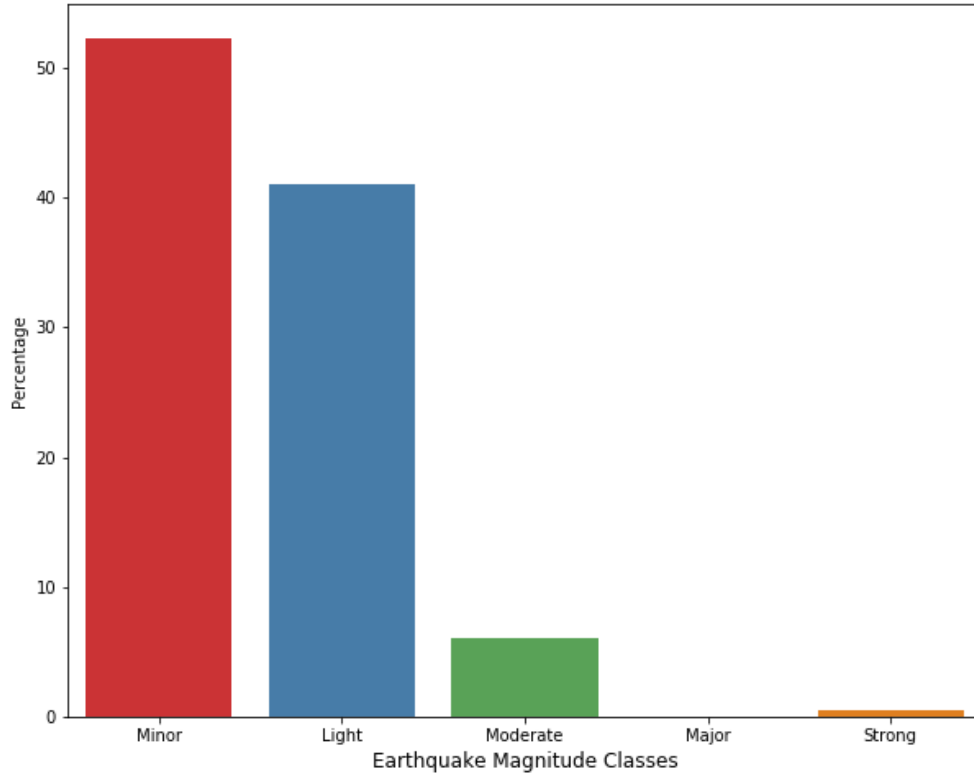


Figure 6: Magnitude Classes

Figure 5 on page 7 gives the Magnitde count of the eathquakes i.e. the number of earthquakes that occurred in the Andaman region between the years 1980 and 2020 for each recorded magnitude scale.

Figure 6 above shows the number of earthquakes that occurred for the following defined magnitude classes -

- $3.0 \leq m_b \leq 3.9 \rightarrow$ “**Minor**”;
- $4.0 \leq m_b \leq 4.9 \rightarrow$ “**Light**”;
- $5.0 \leq m_b \leq 5.9 \rightarrow$ “**Moderate**”;
- $6.0 \leq m_b \leq 6.9 \rightarrow$ “**Strong**”;
- $7.0 \leq m_b \leq 7.9 \rightarrow$ “**Major**”;
- Else \rightarrow “**Great**”.

Figure 7 on page 9 shows the Annual Frequency distribution of the earthquakes in the Andaman region. The code for the given data analysed and visualised here, is available at the end of the assignemnt.

The Andaman Region whose data has been visualised has the following coordinates-

- Top Latitude $\rightarrow 13.734^\circ$
- Bottom Latitude $\rightarrow 10.275^\circ$
- Left Longitude $\rightarrow 91.901^\circ$
- Right Longitude $\rightarrow 93.593^\circ$

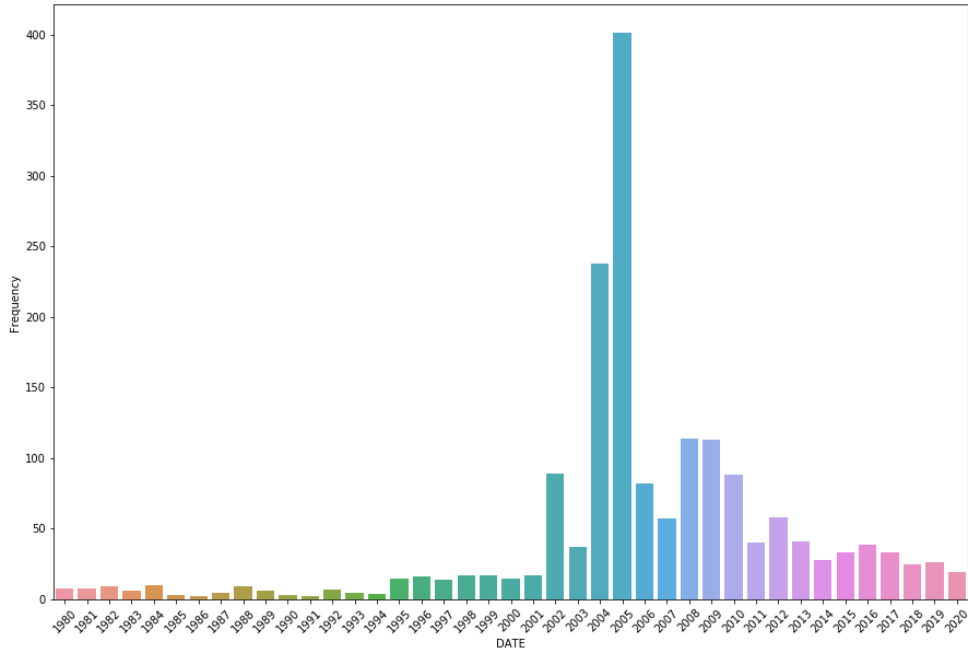


Figure 7: Distribution per year

Seismic Hazard

All of the Andaman & Nicobar Islands lie in Zone V. The entire island chain is also susceptible to tsunamis both from large local quakes and also from massive distant shocks. No warning system is presently in place for any of the islands in this chain. Since the earthquake database in India is still incomplete, especially with regards to earthquakes prior to the historical period (before 1800 A.D.), these zones offer a rough guide of the earthquake hazard in any particular region and need to be regularly updated (See also: GSHAP Hazard map for the Andaman Islands & the Nicobar Islands).

Largest Instrumented Earthquake in the Andaman & Nicobar Islands

26 December 2004 - Sumatra-Andaman Earthquake, $M_w 9.1$; $03.298N, 95.778E, D = 010.0kms, OT = 00 : 58 : 50UTC$

A “very great” earthquake struck the North Indian Ocean & the Bay of Bengal at 00:58 UTC on 26 December 2004. 2,30,210 people were estimated to have been killed in the Indian Ocean-wide tsunami generated by this earthquake. The Nicobar Islands, and to a lesser extent the Andaman Islands, were hardest hit territory in India with as many as 4,486 deaths. Extensive damage occurred in the Nicobar Islands due to the tsunami that is estimated to have attained heights of as much as 20-metres at places.

It is clear from the data analysed that majority of the earthquakes in the Andaman region have been Minor quakes or Light earthquakes with very few Strong earthquakes recorded. The major quakes occurred on the region during 2004, 2005, 2006 and 2010, some with magnitudes much higher than 6.0, and they were also followed by several minor aftershocks which is evident from the spike in the annual frequency distribution curve in years 2004 and 2005.

Answer 3-

The b-value estimation for the defined study area has been carried out using zmap7 software and MATLAB. The data was obtained from the ISC catalogue online bulletin (Centre 2001) for the Region mentioned in Answer 2. The following Map of events was obtained for the given data. -

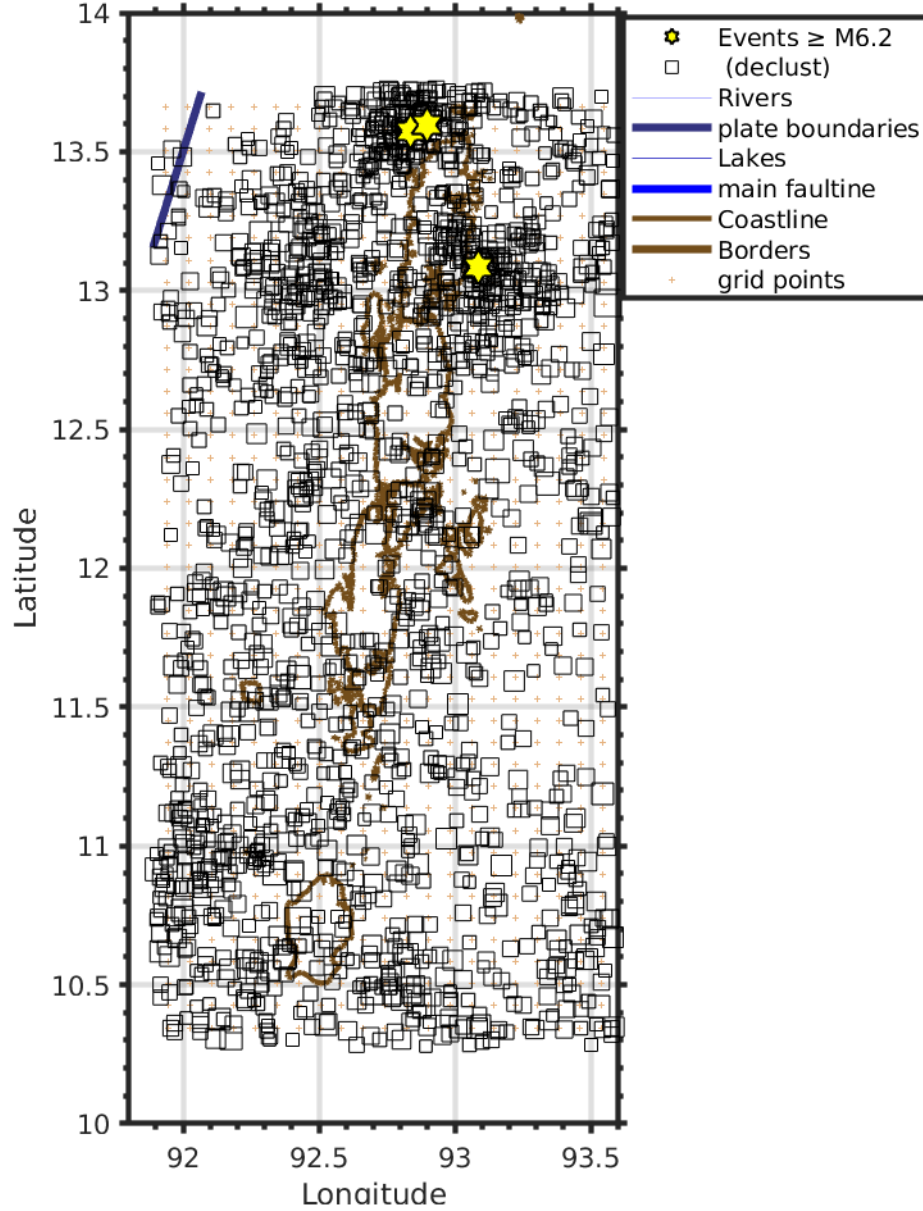


Figure 8: Seismicity map of the studied area. The stars show the epicenters of the maximum magnitude events in the catalogue

Declustering of the Earthquake catalogue

In order to analyze the expected anomalies in the spatial and temporal distribution of the seismic activity, the catalogs were declustered to remove “dependent” events. In our hazard analyses, we are

interested in modeling a process where each possible event is independent of any other. Foreshocks and aftershocks are both temporally and spatially dependent on the mainshock. When we speak of declustering, we mean removing these dependent events. Declustering an earthquake catalog results in a catalog composed of independent events. Here Reasenberg's decluster method (Reasenberg 1985) for aftershock identification was applied. In this method, events within time and distance windows are associated to form clusters and the cluster is replaced with an equivalent earthquake. Any earthquake that occurs within the interaction zone of a prior mainshock is classified as an aftershock and considered as a dependent event. Several parameters have to be chosen for the declustering procedure. A detailed description is given in Ref. (Reasenberg 1985) and in the ZMAP program manual (Wiemer 2001)

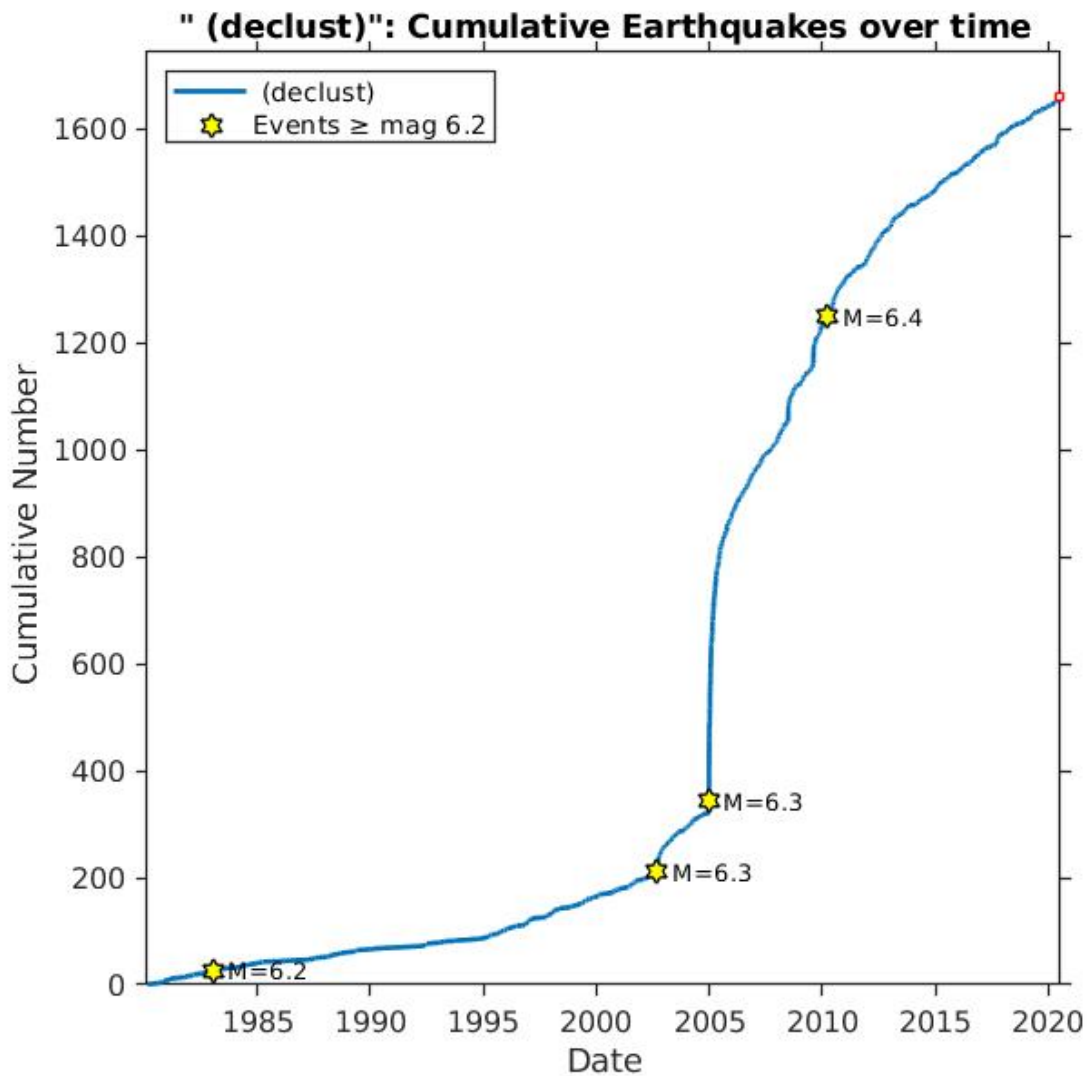


Figure 9: Cumulative Number of original and declustered data-set for ISC Catalogue of Andaman Region over time

The b -value estimation and value of the Magnitude of Completion i.e. M_c was done using the Maximum Likelihood method.

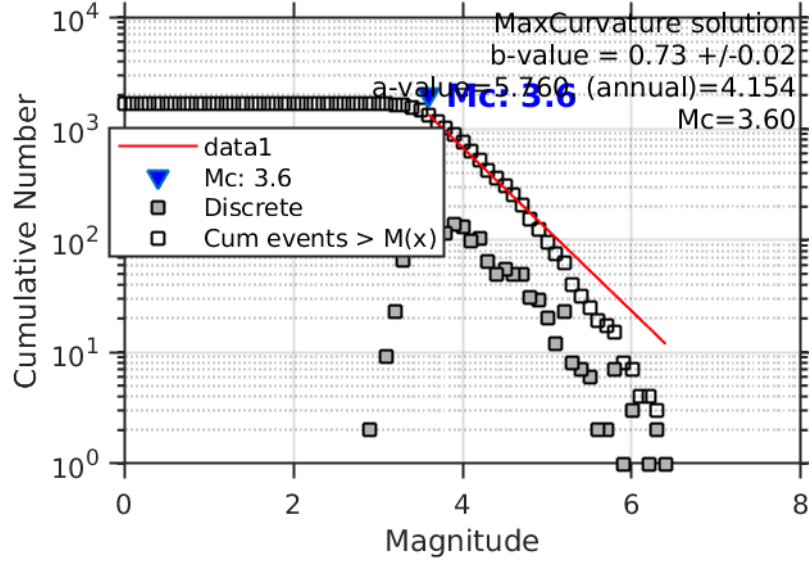


Figure 10: Frequency Magnitude Distribution

The b -value thus obtained here is 0.73 ± 0.02 and the value of M_c is 3.6 from the declusterd ISC catalogue.

An assessment of the minimum magnitude of complete recording, the so-called threshold magnitude, M_c , is an important part of data quality control. M_c is defined as the lowest magnitude for which all events in a studied space-time volume are detected. M_c varies with space and time. For most catalogs it decreases with time, because the number of monitoring seismographs continuously increases and the analysis methods improve (Wiemer and Wyss 2000).

An assessment of the minimum magnitude of complete recording, the so-called threshold magnitude, M_c , is an important part of data quality control. M_c is defined as the lowest magnitude for which all events in a studied space-time volume are detected. M_c varies with space and time. For most catalogs it decreases with time, because the number of monitoring seismographs continuously increases and the analysis methods improve (Wiemer and Wyss 2000). The magnitude of completeness of the declustered catalogs was determined by the Best Combination method in the ZMAP software. Figure 11 on page 13 shows the variation of M_c as a function of time and as expected the minimum threshold is decreasing with time as detected events has significantly increased over the years.

P.T.O.

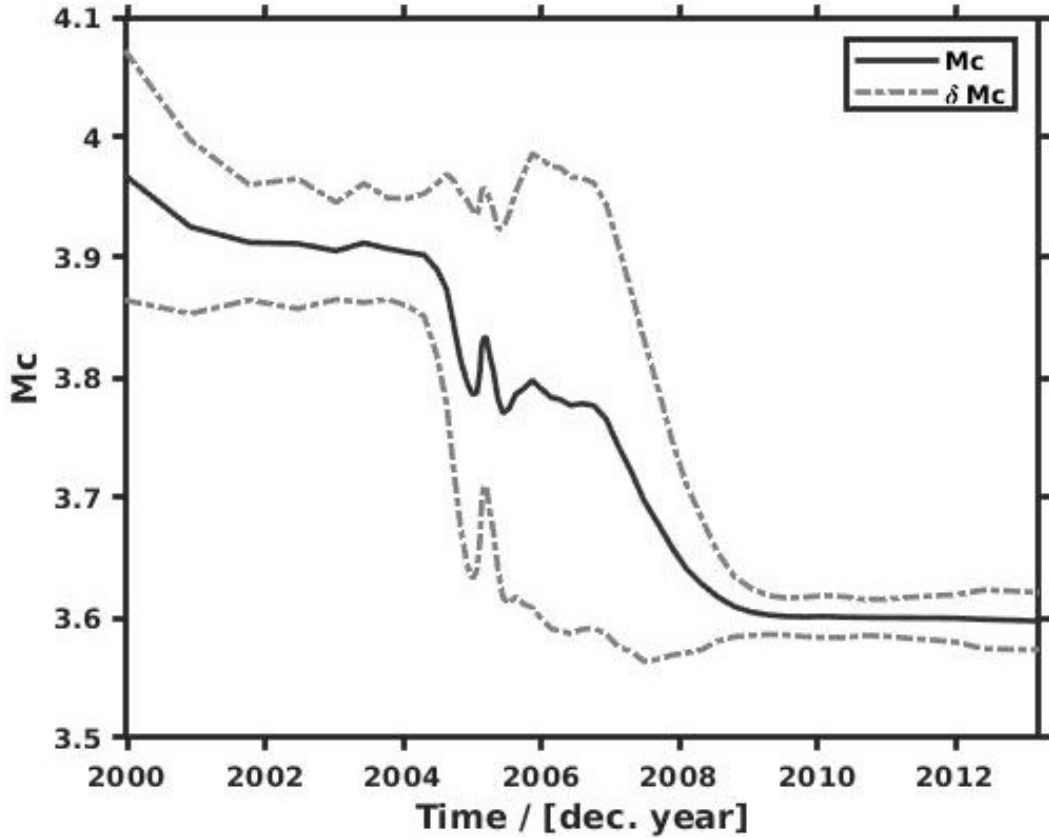


Figure 11: M_c time series

The figure 12 on page 14 shows the magnitude -frequency distribution as a function of spatial position which has been mapped by projecting earthquake epicenters on a plane. The b -value has been estimated at every node of the plane grid. using the N ($N = \text{constant}$) nearest earthquakes, or a varying number of events located within a chosen (constant) distance, R , from the node. When N is constant, it is usually in the range of 50-500 events, and the nodal separation is $0.1^\circ - 1^\circ$, depending on the density of epicenters. To visualize the variation, b -values are translated into a color code and plotted for each grid node. (Nuannin and Kulhánek 2012)

From the b val grid it is clear that Andaman-Nicobar region can be broadly categorised into two seismic zones, one in the northern andaman and the second, southern andaman region. The southern andaman region has very low b -value, whereas, the northern region which has high b -values. We know that higher b -values mean a greater proportion of low magnitude earthquakes to higher magnitude earthquakes and vice-versa for lower values of the constant b . Hence we can say that,

- High b -values \rightarrow low stressed region; small to moderate earthquakes,
- Low b -values \rightarrow high stressed region; moderate to large earthquakes.

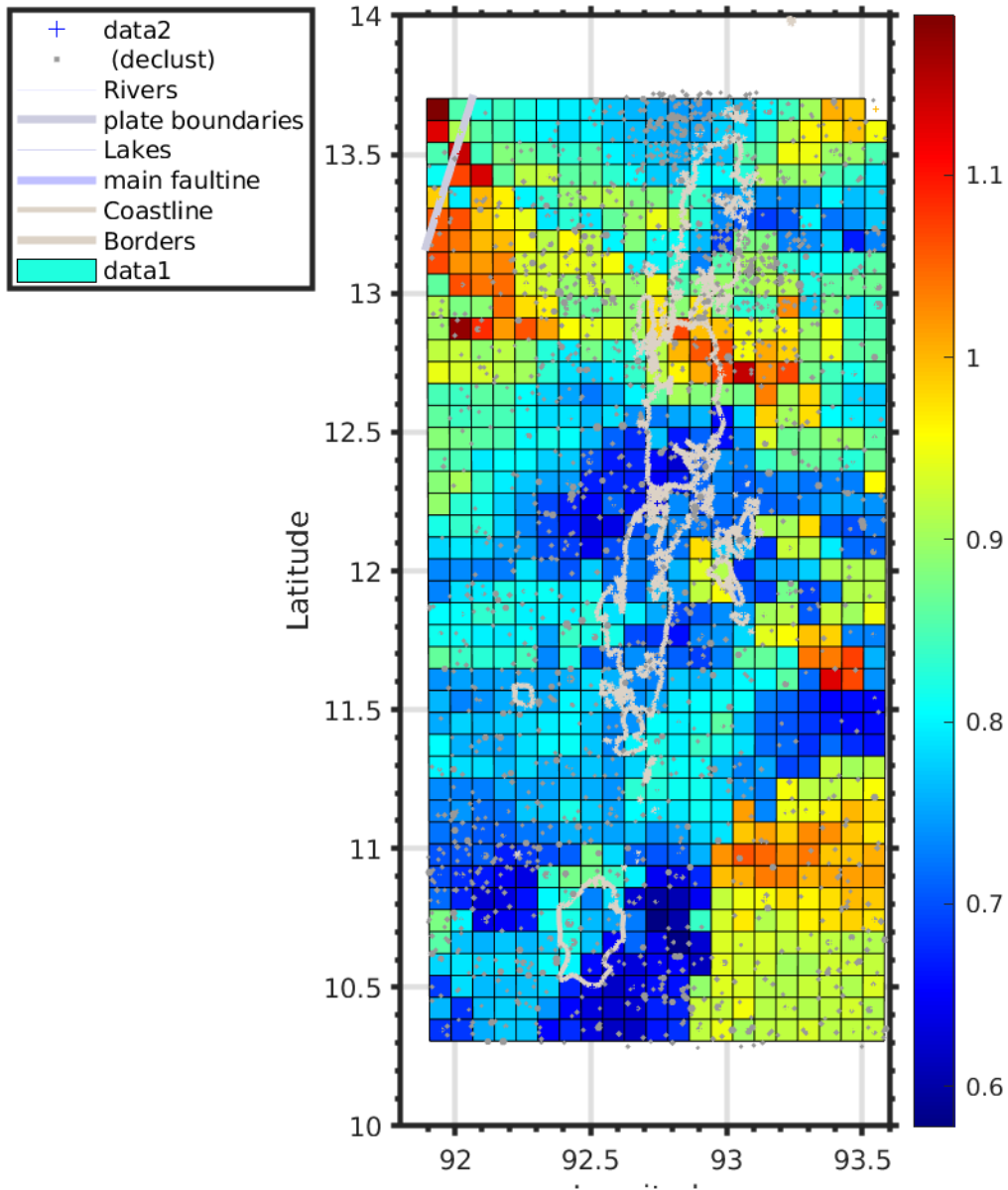


Figure 12: Spatial Variation of b-value

From figure 8 on page 10 and figure 12 above, we can easily say, that all major earthquake events of magnitude ≥ 6.20 occurred in the regions/nodes of relatively low *b-values*, clearly supporting the expectations from the Gutenberg-Richter Law. These regions of low *b-values* are the hazardous regions as these are the areas where a large number of the high magnitude earthquakes have occurred or may occur in the near future. The regions of extremely high *b-values* are essentially earthquake swarms of fairly low magnitude quakes.

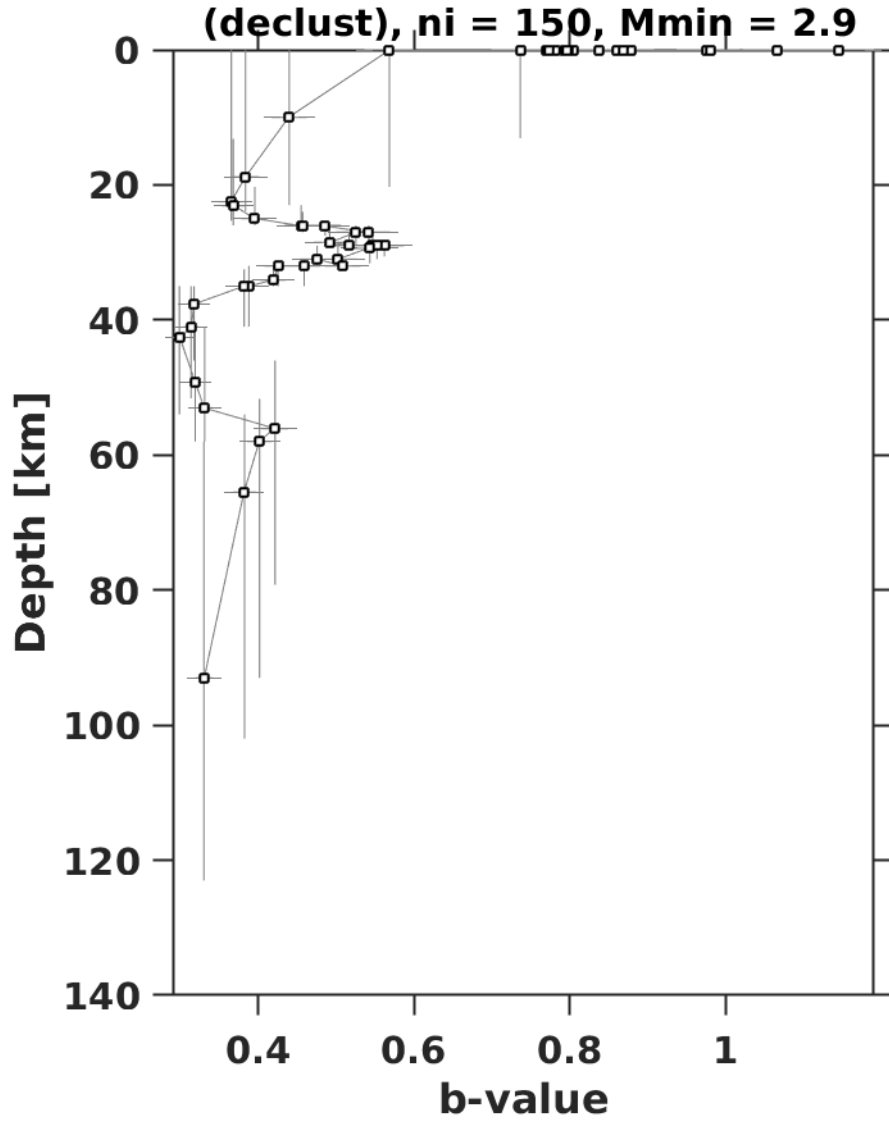


Figure 13: Variation of b-value with depth

The Gutenberg–Richter b-value is thought to reflect the stress conditions in the crust; therefore, spatial and/or temporal variations of the b-value can provide important information regarding crustal tectonics. But, the variability of b with depth is often not statistically significant and that the decrease of b with depth should be interpreted with caution as shown by many research groups including Amorese et. al. (Amorese, Grasso, and Rydelek 2010).

In this case, the variation of b -value with depth indicates a steady decline in the value of b with depth. There is a sharp increase in value of b at $\approx 30\text{km}$. and 58km . This indicates that the higher magnitude quakes are generally occurring at larger depths small quakes have their epicenter at/near the surface.

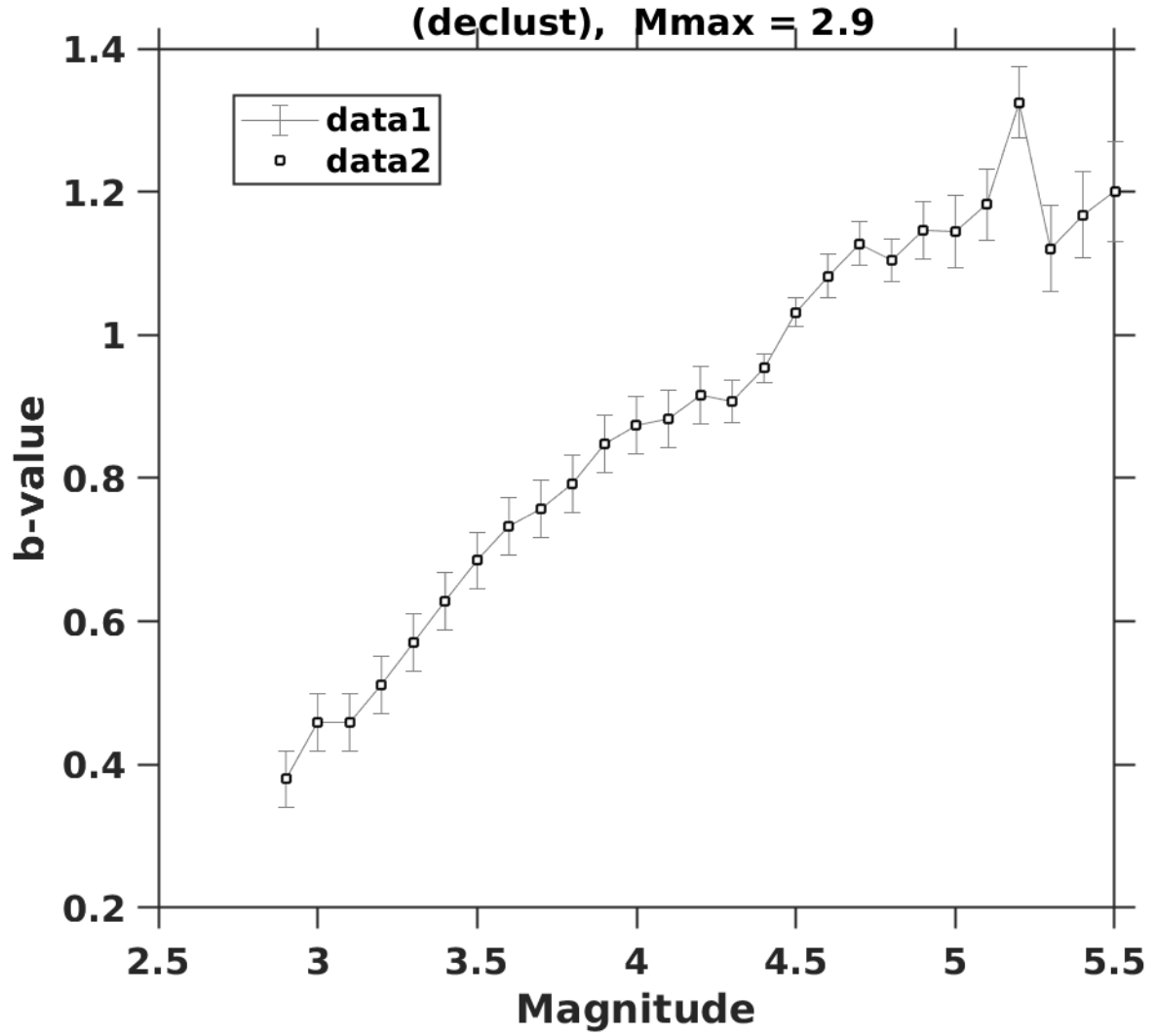


Figure 14: Variation of b-value with magnitude

The Figure 14 above shows the variation of *b-value* of the Andaman region with the magnitude of the quakes. This is in accordance with the Gutenberg-Richter law. The larger the magnitude of the quake, the larger will be the corresponding *b-value* and thus the smaller will be the corresponding cumulative frequency as seen from the following equation.

$$N = N_{TOT}10^{-bM}$$

The opposite is true for smaller magnitude quakes.

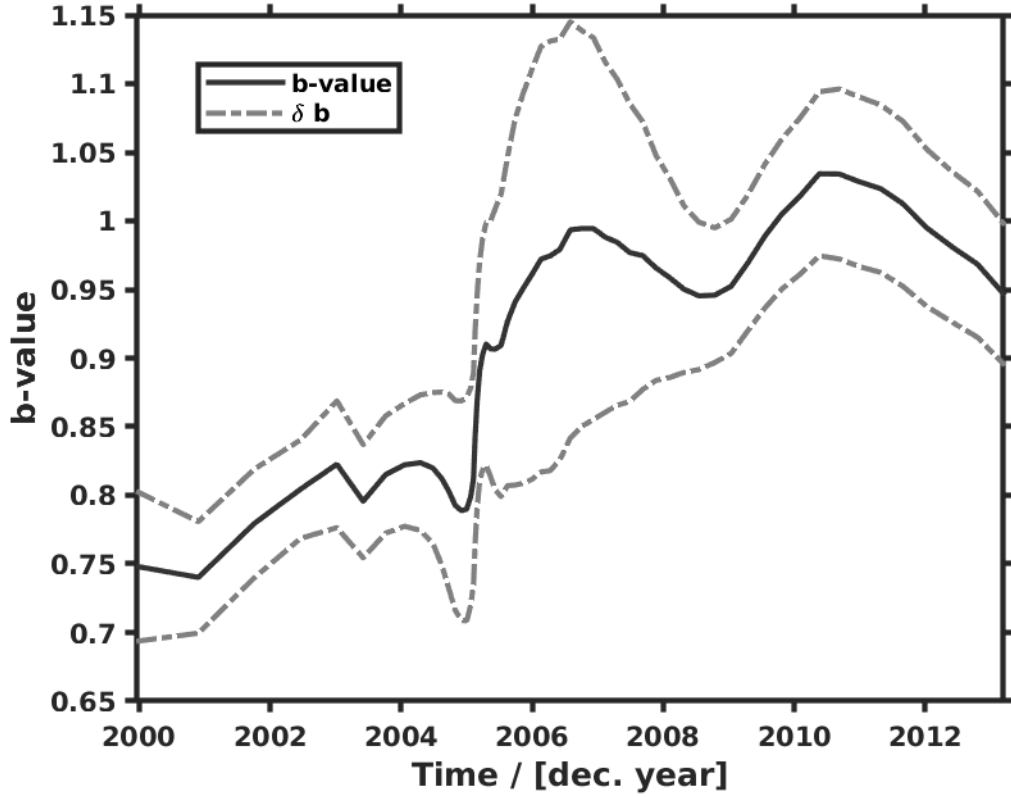


Figure 15: Temporal variation of b-value

The temporal variations of the *b-values* shown in the above figure give us a great insight into the potentially hazardous regions in the greater Andaman region. We can correlate the temporal variation with the Seismic history of the region to identify the changes in *b-value* before/after major earthquake events.

- From the figure it is clear there was a sharp decline in the *b-value* before the major quake in North Andaman in the year 2002 of moment magnitude $M_w = 6.5$ ¹
- After this, there was a steady increase in the *b-value* with time before a sharp decrease in the value of *b* right before the devastating Sumatra-Andaman quake in 2004 of $M_w = 9.1$
- Then there was another significant decline in *b-value* prior to the major quakes in 2005 and 2006 followed by many years of steady incline in the value of *b*.
- Then there was another sharp decline before the major quake in June, 2010 off the shore of Nicobar islands with a moment magnitude of $M_w = 7.5$ and then there was a steady rise before a another steady decline.

It is clear from these observations that there has been a steady decline in the b-value of the Andaman region before major quakes have occurred and this time between the sharp decline in the *b-value* in the analysed data-set and actual occurrence of the quake (also called the Precursor time) can be used to predict the time until the next major quake by using *b-values* as a precursor for predicting earthquakes of specific magnitudes depending on the decrease in *b-value* in the data gathered from the Seismology catalogues.

¹Note that the magnitudes used from the ISC catalogue for analysis are body wave magnitude i.e. M_b

References

- Amorese, D, J-R Grasso, and PA Rydelek. 2010. "On Varying B-Values with Depth: Results from Computer-Intensive Tests for Southern California." *Geophysical Journal International* 180 (1). Blackwell Publishing Ltd Oxford, UK: 347–60.
- Bandopadhyay, PC, and Andrew Carter. 2017. "Geological Framework of the Andaman–Nicobar Islands." *Geological Society, London, Memoirs* 47 (1). Geological Society of London: 75–93.
- Bhattacharya, Pathikrit, Bikas K Chakrabarti, Debashis Samanta, and others. 2009. "Fractal Models of Earthquake Dynamics." *arXiv Preprint arXiv:0906.1931*.
- Centre, International Seismological. 2001. "On-Line Bulletin." *Internatl. Seis. Cent.* Thatcham United Kingdom.
- Curray, Joseph R. 2005. "Tectonics and History of the Andaman Sea Region." *Journal of Asian Earth Sciences* 25 (1). Elsevier: 187–232.
- Gutenberg, Beno. 1956. "The Energy of Earthquakes." *Quarterly Journal of the Geological Society* 112 (1-4). Geological Society of London: 1–14.
- Nuannin, Paiboon, and Ota Kulhánek. 2012. "A Study of B-Value Precursors Applied to the Andaman-Sumatra Region." *Journal of Earth Science and Engineering* 2 (January): 166–88.
- Nuannin, Paiboon, Ota Kulhanek, and Leif Persson. 2005. "Spatial and Temporal B Value Anomalies Preceding the Devastating Off Coast of Nw Sumatra Earthquake of December 26, 2004." *Geophysical Research Letters* 32 (11). Wiley Online Library.
- Pacheco, Javier F, Christopher H Scholz, and Lynn R Sykes. 1992. "Changes in Frequency–Size Relationship from Small to Large Earthquakes." *Nature* 355 (6355). Springer: 71–73.
- Prasad, Shweta, and CHAnDRAnI SInGH. 2015. "Evolution of B-Values Before Large Earthquakes of $M_b \geq 6.0$ in the Andaman Region." *Geologica Acta: An International Earth Science Journal* 13 (3). Universitat de Barcelona: 205–10.
- Reasenber, Paul. 1985. "Second-Order Moment of Central California Seismicity, 1969–1982." *Journal of Geophysical Research: Solid Earth* 90 (B7). Wiley Online Library: 5479–95.
- Smith, Warwick D. 1981. "The B-Value as an Earthquake Precursor." *Nature* 289 (5794). Nature Publishing Group: 136–39.
- Wiemer, Stefan. 2001. "A Software Package to Analyze Seismicity: ZMAP." *Seismological Research Letters* 72 (3). Seismological Society of America: 373–82.
- Wiemer, Stefan, and Max Wyss. 2000. "Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan." *Bulletin of the Seismological Society of America* 90 (4). Seismological Society of America: 859–69.

SRTP project analysis Group 9 (Andaman)

Name - Aditya Sengupta, Institute - AIAS, Amity University, Noida

Mentor - Prof. Debasis D. Mohanty, CSIR-NEIST, Jorhat, Assam

Data obtained from the ISC catalogue at <http://www.isc.ac.uk/iscbulletin/search/catalogue/> for the Andaman Region for the years 1980 to 2020 and for min magnitude $M_b = 0.1$ to max magnitude $M_b = 10$, where M_b is the Body wave magnitude of the earthquake and for min. depth of 0 to max. depth of $6731 \approx \text{Radius of earth or } R_e$.

In [87]:

```
import pandas as pd
import requests
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import scipy.stats as stats
plt.style.use('seaborn-pastel')

import geopandas as gpd
from descartes import PolygonPatch
from matplotlib.animation import FuncAnimation
# from matplotlib.patches import Polygon
```

In [88]:

```
df = pd.read_csv('Andaman_data.csv')
```

In [89]:

```
df.head()
```

Out[89]:

	year	month	date	Time	latitude	longitude	depth	magnitude	DATE
0	1980	2	21	16:05:50	10.5276	92.9652	78.7	4.6	1980-02-21
1	1980	5	26	11:37:30	11.0789	92.8553	43.2	4.8	1980-05-26
2	1980	8	20	6:21:54	11.0277	92.9682	54.0	5.0	1980-08-20
3	1980	9	29	0:18:31	13.0346	93.0987	32.0	4.5	1980-09-29
4	1980	10	31	8:01:10	12.2750	92.8534	41.0	4.6	1980-10-31

In [90]:

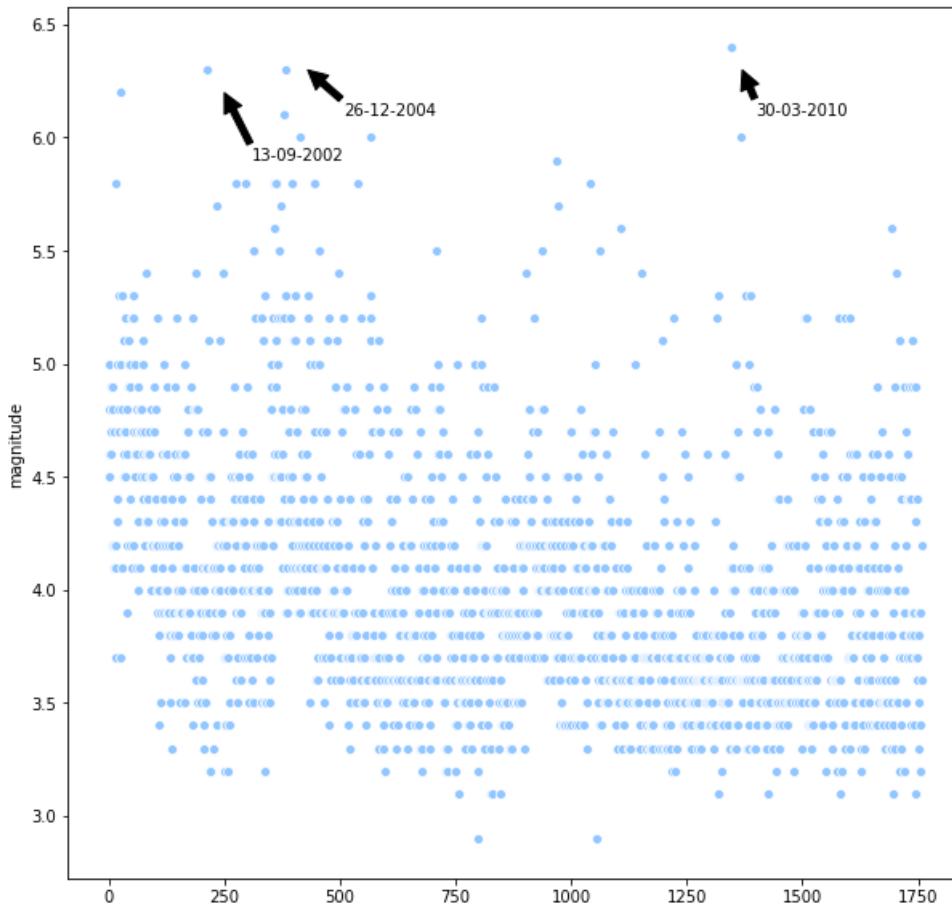
```
df.describe()
```

Out[90]:

	year	month	date	latitude	longitude	depth	magnitude
count	1759.000000	1759.000000	1759.000000	1759.000000	1759.000000	1759.000000	1759.000000
mean	2006.017624	6.469585	17.370665	12.217075	92.725736	25.825867	4.017112
std	6.697191	3.901819	9.060043	1.072680	0.445045	27.593395	0.543749
min	1980.000000	1.000000	1.000000	10.280000	91.900600	0.000000	2.900000
25%	2004.000000	3.000000	10.000000	11.174000	92.370050	0.000000	3.600000
50%	2005.000000	6.000000	17.000000	12.438400	92.749400	26.400000	3.900000
75%	2009.000000	10.000000	26.000000	13.137250	93.065350	32.450000	4.300000
max	2020.000000	12.000000	31.000000	13.727700	93.586700	291.600000	6.400000

In [91]:

```
plt.figure(figsize=(10,10))
sns.scatterplot(y="magnitude", x=df.magnitude.index , data=df)
plt.annotate('30-03-2010', xy=(1370, 6.3), xytext=(1400, 6.1),arrowprops=dict(facecolor='black', shrink=0.01),)
plt.annotate('26-12-2004', xy=(430, 6.3), xytext=(510, 6.1),arrowprops=dict(facecolor='black', shrink=0.01),)
plt.annotate('13-09-2002', xy=(250, 6.2), xytext=(310, 5.9),arrowprops=dict(facecolor='black', shrink=0.01),)
plt.savefig('overall_earthquakes_andaman.png')
```

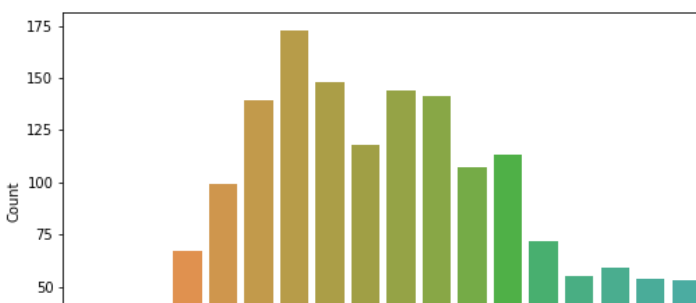


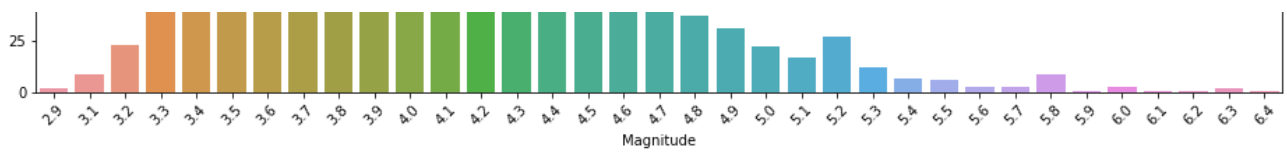
In [24]:

```
#group by magnitude counts
df_magnitude_count = pd.DataFrame(df.magnitude.groupby(df.magnitude).count())
df_magnitude_count['index'] = df_magnitude_count.index
df_magnitude_count.columns = ["Count", "Magnitude"]
```

In [92]:

```
plt.figure(figsize=(16,5))
ax = sns.barplot(x="Magnitude", y="Count", data=df_magnitude_count)
ax.set_xticklabels(ax.get_xticklabels(),rotation=45)
plt.savefig('magnitude_count_andaman.png')
```





Earthquake Magnitude classes

In [26]:

```
scale = []
for magnitude in df.magnitude:
    if magnitude >= 3.0 and magnitude <=3.9:
        scale.append("Limited Damage")
    elif magnitude >=4.0 and magnitude <=4.9:
        scale.append("Minor Damage")
    elif magnitude >=5.0 and magnitude <=5.9:
        scale.append("Slight Damage")
    elif magnitude >=6.0 and magnitude <=6.9:
        scale.append("Severe Damage")
    elif magnitude >=7.0 and magnitude <=7.9:
        scale.append("Serious Damage")
    else:
        scale.append("Great Damage")
```

In [27]:

```
df['scale'] = scale
```

In [28]:

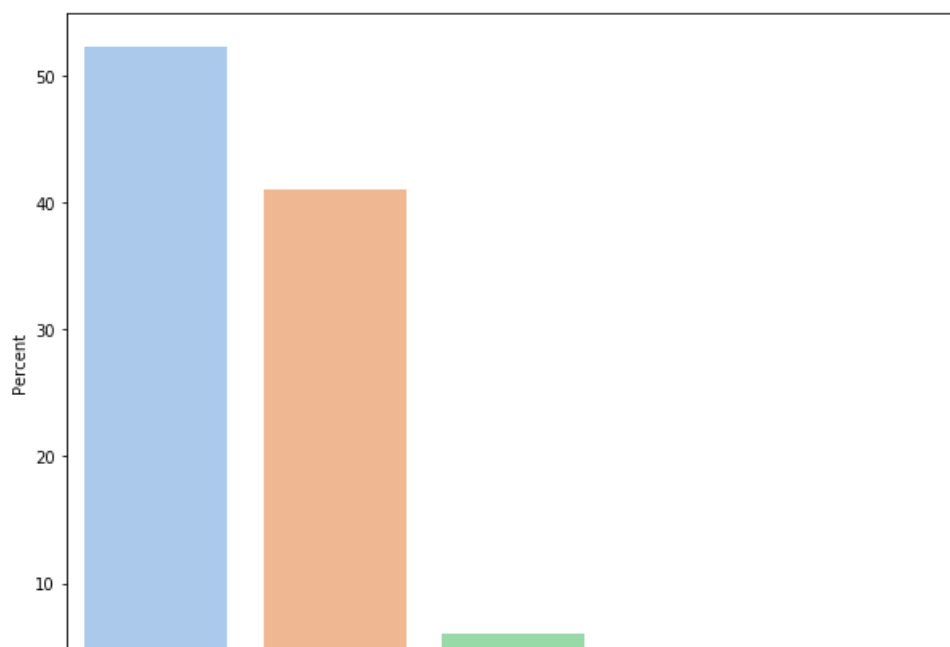
```
df_damage_count = pd.DataFrame(df.scale.groupby(df.scale).count().sort_values(ascending=False))
df_damage_count['percent'] = ((df_damage_count.scale)/sum(df_damage_count.scale))*100
```

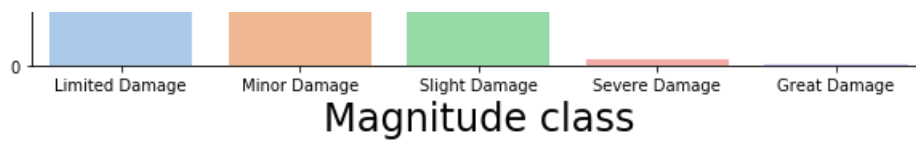
In [29]:

```
plt.figure(figsize=(10,8))
sns.barplot(df_damage_count.index,df_damage_count.percent, palette="pastel")
# plt.xticks(df_damage_count.index, rotation= 10)
plt.ylabel('Percent')
plt.xlabel('Magnitude class', fontsize=24)
```

Out[29]:

```
Text(0.5, 0, 'Magnitude class')
```





Magnitude class size: "Great", "Major", "Strong", "Moderate", "Light", "Minor", "Less than 3.9")

In [30]:

```
#array for storing the size_class
size_class = []

for magnitude in df.magnitude:
    if magnitude >= 3.0 and magnitude <=3.9:
        size_class.append("Minor")
    elif magnitude >=4.0 and magnitude <=4.9:
        size_class.append("Light")
    elif magnitude >=5.0 and magnitude <=5.9:
        size_class.append("Moderate")
    elif magnitude >=6.0 and magnitude <=6.9:
        size_class.append("Strong")
    elif magnitude >=7.0 and magnitude <=7.9:
        size_class.append("Major")
    else:
        size_class.append("Great")

#Creating a column in the dataframe called class_size
df['size_class'] = size_class
```

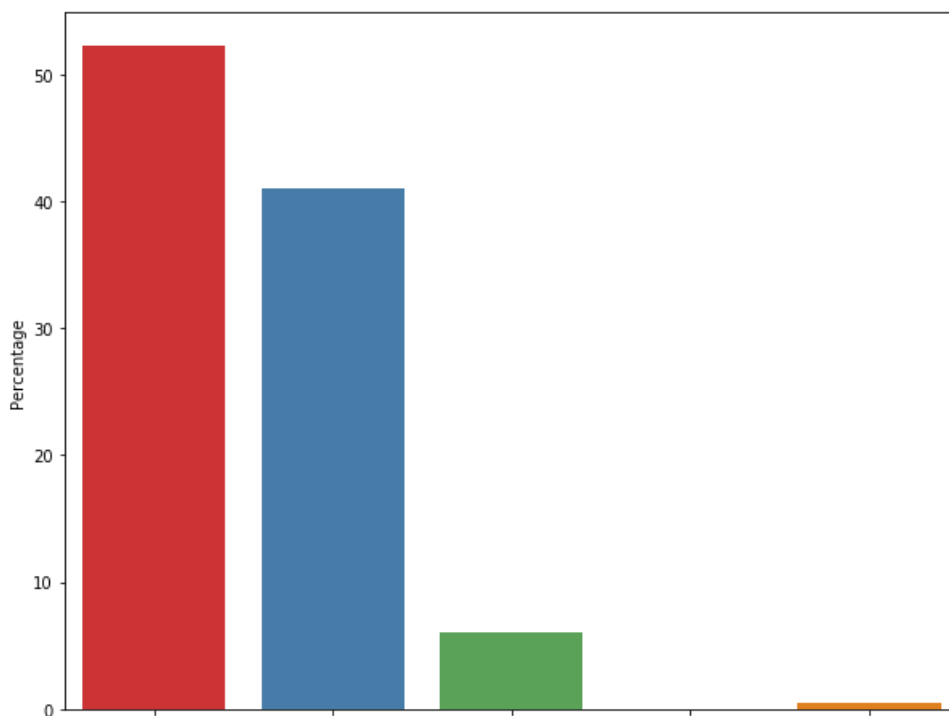
In [31]:

```
df_size_class = pd.DataFrame(df.size_class.groupby(df.size_class).count())
df_size_class['percent'] = ((df_size_class.size_class)/sum(df_size_class.size_class))*100
```

In [93]:

```
plt.figure(figsize=(10,8))
sns.barplot(df_size_class.index,df_size_class.percent, palette="Set1", order = ['Minor','Light','Moderate',
                                                                              'Major',
                                                                              'Strong'])

plt.ylabel('Percentage')
plt.xlabel('Earthquake Magnitude Classes', fontsize=12)
plt.savefig('magnitude_classes_andaman.png')
```



In [94]:

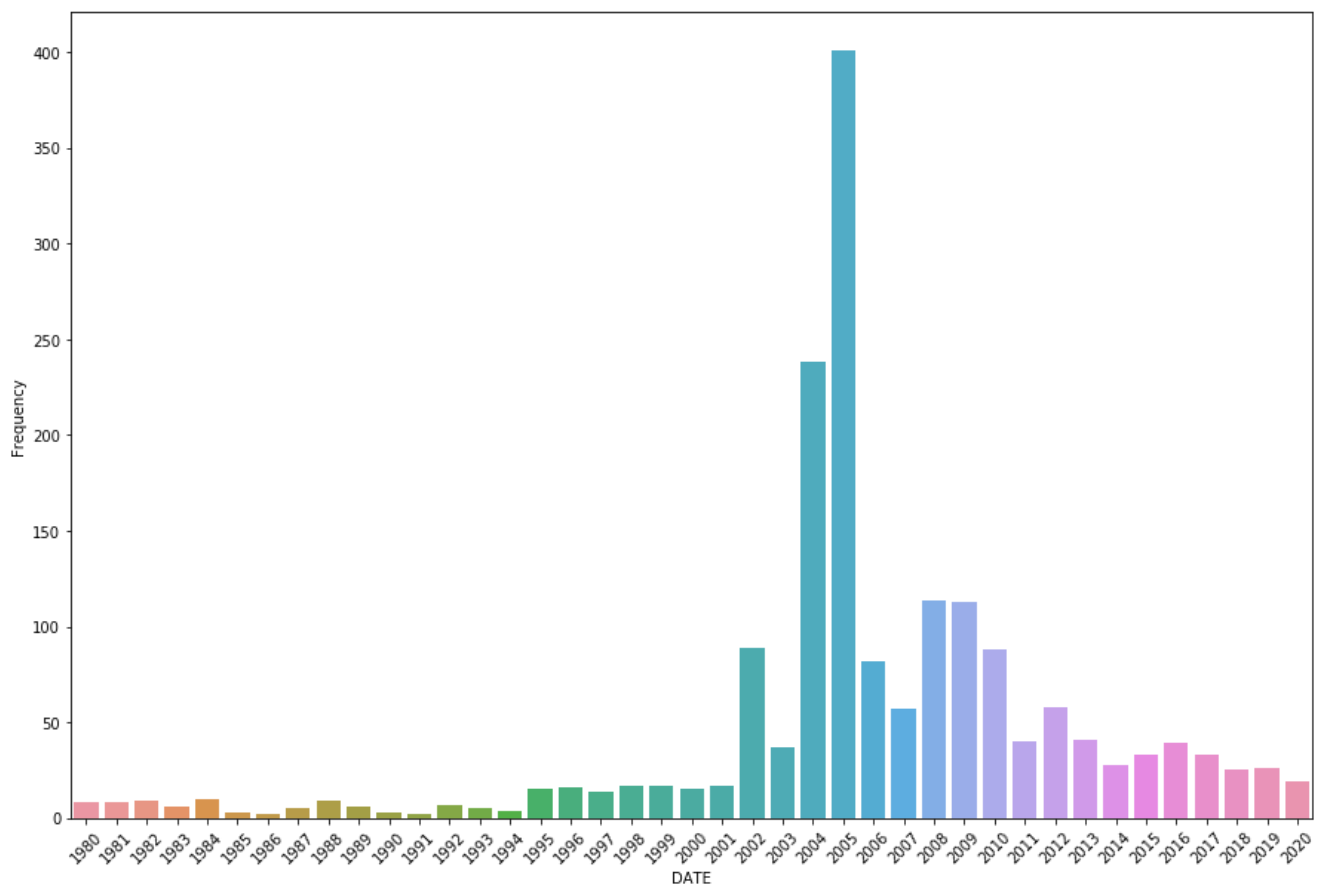
```
mag_class = []
for magnitude in df.magnitude:
    if magnitude >= 3.0 and magnitude <=3.9:
        mag_class.append("3-3.9")
    elif magnitude >=4.0 and magnitude <=4.9:
        mag_class.append("4-4.9")
    elif magnitude >=5.0 and magnitude <=5.9:
        mag_class.append("5-5.9")
    elif magnitude >=6.0 and magnitude <=6.9:
        mag_class.append("6-6.9")
    elif magnitude >=7.0 and magnitude <=7.9:
        mag_class.append("7-7.9")
    else:
        mag_class.append("Above 8")
df['mag_class'] = mag_class
```

In [95]:

```
df['DATE'] = pd.to_datetime(df['DATE'], errors='coerce')
group_by_year = pd.DataFrame(df.magnitude.groupby([df['DATE'].dt.year]).count())
group_by_year.columns = ['Frequency']
```

In [96]:

```
plt.figure(figsize=(15,10))
by_year = sns.barplot(x=group_by_year.index, y=group_by_year.Frequency)
by_year.set_xticklabels(by_year.get_xticklabels(),rotation=45)
plt.savefig('yearly_eq_andaman.png')
```



In [97]:

```
df.head()
```

Out[97]:

	year	month	date	Time	latitude	longitude	depth	magnitude	DATE	mag_class
0	1980	2	21	16:05:50	10.5276	92.9652	78.7	4.6	1980-02-21	4-4.9
1	1980	5	26	11:37:30	11.0789	92.8553	43.2	4.8	1980-05-26	4-4.9
2	1980	8	20	6:21:54	11.0277	92.9682	54.0	5.0	1980-08-20	5-5.9
3	1980	9	29	0:18:31	13.0346	93.0987	32.0	4.5	1980-09-29	4-4.9
4	1980	10	31	8:01:10	12.2750	92.8534	41.0	4.6	1980-10-31	4-4.9

In [99]:

```
df.drop(['mag_class'], axis=1, inplace=True)
df.head()
```

Out[99]:

	year	month	date	Time	latitude	longitude	depth	magnitude	DATE
0	1980	2	21	16:05:50	10.5276	92.9652	78.7	4.6	1980-02-21
1	1980	5	26	11:37:30	11.0789	92.8553	43.2	4.8	1980-05-26
2	1980	8	20	6:21:54	11.0277	92.9682	54.0	5.0	1980-08-20
3	1980	9	29	0:18:31	13.0346	93.0987	32.0	4.5	1980-09-29
4	1980	10	31	8:01:10	12.2750	92.8534	41.0	4.6	1980-10-31

In [101]:

```
df.to_csv(r'Andaman.csv', index = False)
```

In []: