Statistics as a Technology to Predict the Seasonal Variation of Air Pollution

Tanushree Bhattacharya, Tripta Narayan, Soubhik Chakraborty, Swapan Konar, Shilpi Singh

Abstract: The present study focuses on the analysis and prediction of the seasonal air quality over an industrial city of eastern India. It investigates the seasonal characteristics of three air pollutants nitrogen dioxide, PM_{10}, and sulphur dioxide (SO_2) between 2005 and 2015. The data has been obtained from the ground monitoring station of the Jharkhand State Pollution Control Board. The study concentrated on the seasons' based findings of RSPM, SO_2 and NO_X. SPSS 22 software was used to find meteorological influences on the conditions of particular matters. The study shows the strength of statistics as a technology to analyse and to make a prediction even when the available information includes only one variable.

Keywords: Seasonal variation, RSPM, Probability density function, NAAQ.

I. INTRODUCTION

Throughout the world, high levels of concentration of air pollutants have been found in most of the urban areas. It has been observed that there are differences in the air qualities in different parts of the urban areas. Also, studies shows that the local sources such as traffic and industrial emissions are among the contributing factors [1]. City air pollution is dangerous for the environment. Meteorological parameters have an influence over the concentration level of pollutants and there is sufficient literature to support this fact. A study in Kuwait showed seasonal variations in the air pollutants. The maximum concentrations of NMHC (non-methane hydrocarbon), CO, and NO_X was found in the winter whereas the minimum concentration was in the summer [2]. A similar observation has been in Japanese and Indian cities also. In Hong Kong, an increase in the concentrations of air pollutants was observed during winter season. Heavy rain and mixing layer have a great role in the existing pollution level in an area and the downfall in concentration during summer season can be seen in this context. Similarly, in Russian cities, there have been observed different concentration levels in different seasons [3]. Thus, the meteorological factor has an important aspect in the air pollution study to understand the variation and distributions.

Revised Manuscript Received on January 05, 2020.

***** Correspondence Author

Tripta Narayan *, Department of Physics, Birla Institute of Technology, Mesra, Ranchi-835215, Jharkhand, India. E-mail: tripta2609@gmail.com

Tanushree Bhattacharya, Department of Civil and Environment Engineering, Birla Institute of Technology, Mesra, Ranchi-835215, Jharkhand, India. E-mail: tbhattacharya@bitmesra.ac.in

Soubhik Chakraborty, Department of Mathematics , Birla Institute of Technology, Mesra, Ranchi-835215, Jharkhand, India.
E-mail: soubhikc@yahoo.co.in

Swapan Konar, Department of Physics, Birla Institute of Technology, Mesra, Ranchi-835215, Jharkhand, India. E-mail: skonar@bitmesra.ac.in

Shilpi Singh, School of Engineering and Technology, Amity University Patna, Patna, India. E-mail: shilpi.singh.it@gmail.com

The variation and pattern of variation of particulate matters, NO_X, and SO_2 are more often influenced by human-induced changes in land use/ land cover along with various economies, and demographics [4]. The selection of the area of study for the present research has been done keeping these aspects in mind.

In a report published by WHO in 2016, It is clearly mentioned that in India, the concentration of particulate matter are at hazardous level [5]. NO_2 is rising, but SO_2 is dropped. This is happing due to rising numbers of vehicles [7]. After Ranchi become capital of Jharkhand the increase of vehicles and rapid growth cause ambient air pollution status change [8].Moreover, Ranchi city has mining areas and industries nearby. So, it is important to study RSPM, SO_2, and NO_X.

The Jharkhand State Pollution Control Board (JSPCB) has established ground monitoring stations in four cities of the Jharkhand state, to keep an eye on the concentration levels of various air pollutants. Thus, information on the pattern of variation of each pollutant's concentration with time and the influences of the prevailing meteorological conditions are necessary for developing air pollutant control strategies.

The objectives of the study

- To find the appropriate distribution pattern for the seasonal concentration of air pollutants. With the help of continuous probability model.
- To find the influences of meteorological factors on the seasonal variation of the pollutants and
- To generate a model using multiple linear regression technique.

II. MATERIAL AND METHODS

A. Study area

Jharkhand state is the 28^{th} state of India. Its capital Ranchi is extended from latitude 22° 30' North - 23°36' and longitude 84°54' East – 85°54'. Ranchi is a hilly city with good amount of rainfall. Ranchi is situated at the average altitude of 600 meters above the mean sea level. As far as the climate is concerned, it is generally dry. The average rainfall has been observed as 1500 mm. However, the temperature in summer is recorded generally up to 40° -44°. Ranchi is surrounded by forest, which is 17.38% out of the total area. After becoming capital of Jharkhand in 2000, the population of the city is gradually increasing. The construction work increases so, vacant ground area is used, this caused change in environment. Due, to heavy vehicles load in the city, air quality also degraded and the regular rainfall condition has been observed to be changed.

Industries are also establishing in this city, these are the main causes of air pollution in the city. To monitor the air quality of the city, a monitoring station has been established at Albert Ekka Chowk. This Chowk is the most polluted and heavy traffic place.

Data collection

Data is collected from the Jharkhand State Pollution Control Board office. Furthermore, the data for the meteorological parameters such as rainfall, relative humidity temperature, and wind speed and wind direction have been collected from the India Meteorological Department office at Hinoo.

B. Method

From the JSPCB office, the data have been collected in 4 hourly, 8 hourly and 24 hourly average formats. The raw data were then arranged in the tabular form according to the requirement of the present research. With the help of Easyfit software, the probability distributions have been tried to fit the pattern. The Kolmogorov –Smirnov and Anderson –Darling tests were utilised to find how good the fit is. Finally, with the help of inverse cumulative distribution function, the probability that a particular pollutant will exceed the Air Quality Standard set for that pollutant was computed. Further, multiple linear regression technology was employed over the dependent and independent variables. SPSS software had been utilized. In SPSS, while using the "stepwise" method [9], the input data has to fulfill the following criteria:

$$Probability - of - F - to - enter \leq 0.050$$
$$and$$
$$Probability - of - F - to - remove >= 0.100$$

Mathematical expression for the probability of exceedance of a critical concentration is [10]:

$$Pr(X > x) = 1 - Pr(X \leq x) = 1 - \int_{-\infty}^{x} f(t)dt$$

III. RESULT AND ANALYSIS

During 2005-2015 RSPM, NO_X and SO_2were measured in microgram per cubic meter ($\mu g/m^3$). The frequency distribution of air pollution data is given in Table I.

Through the application of descriptive statistics, we choose the appropriate distribution for the data set. It can be seen that the available data for the RSPM concentrations (as collected) is positively skewed for winter and monsoon and negatively skewed for summer and post-monsoon whereas for NO_X and SO_2 are positively skewed for the three seasons namely Winter, Summer and Monsoon. This justifies the reason for using the accordingly skewed appropriate theoretical distributions to fit the RSPM, NO_X and SO_2 data. The graph is shown in Figures 1, 2 and 3.The parameters of distributions used in Table-I. The summary of the analysis is given in Table II.

The data of NO_X and SO_2 are less scattered as compared to RSPM. The results for the probability of exceedance of the National Ambient Air Quality Standards (NAAQS) are presented in table III.

It has been found that the RSPM concentration is crossing the NAAQ standards. Thus, for air pollution in Ranchi, RSPM is the criteria pollutant and the control strategies should focus at this pollutant more than any other pollutant. In the next step, the model has been generated keeping RSPM as dependent variable and the meteorological parameters as independent variables.

A. Multivariate Regression Model

Multivariate linear regression model (MLR) has been widely used in research work till now. It is very useful in making predictions whenever there exist linear correlations among the variables under study. In the present study, meteorological conditions and emission sources are found to be responsible for the existing air quality. So, it is very necessary to model their relationship for making effective control strategies.

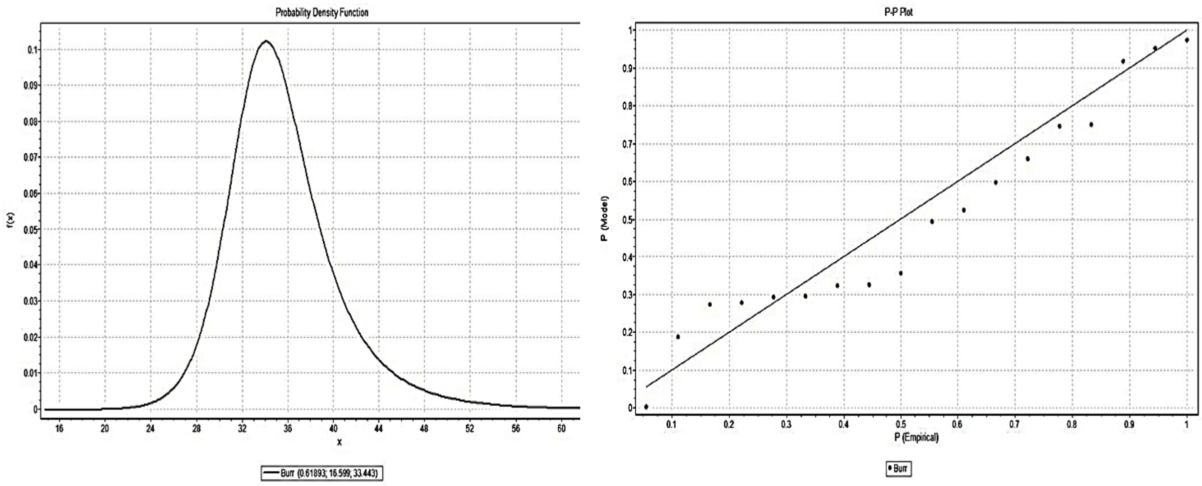

Fig. 1. NO_X (Winter)

Retrieval Number: B7599129219/2020©BEIESP
DOI: 10.35940/ijitee.B7599.019320

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

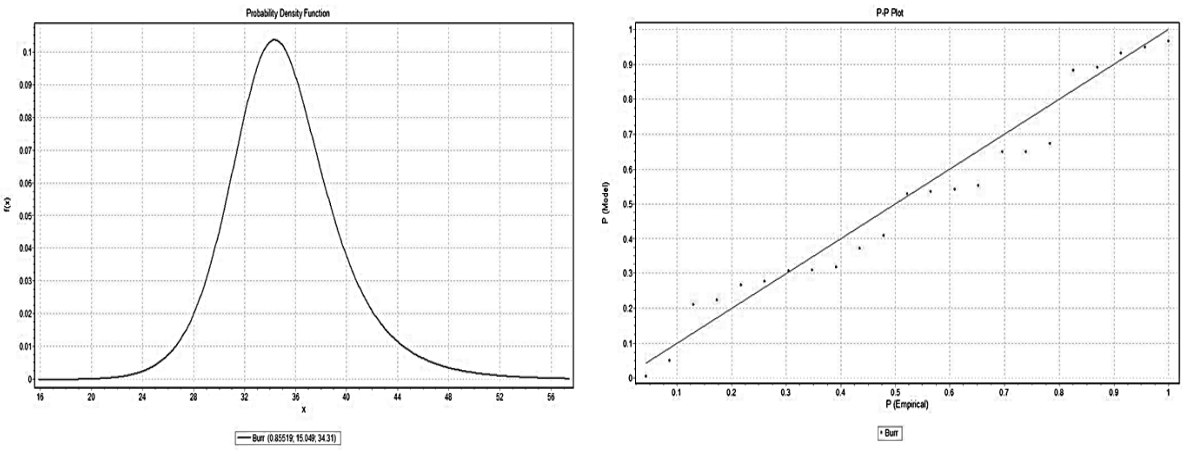

Fig. 2. NO_X (Summer)

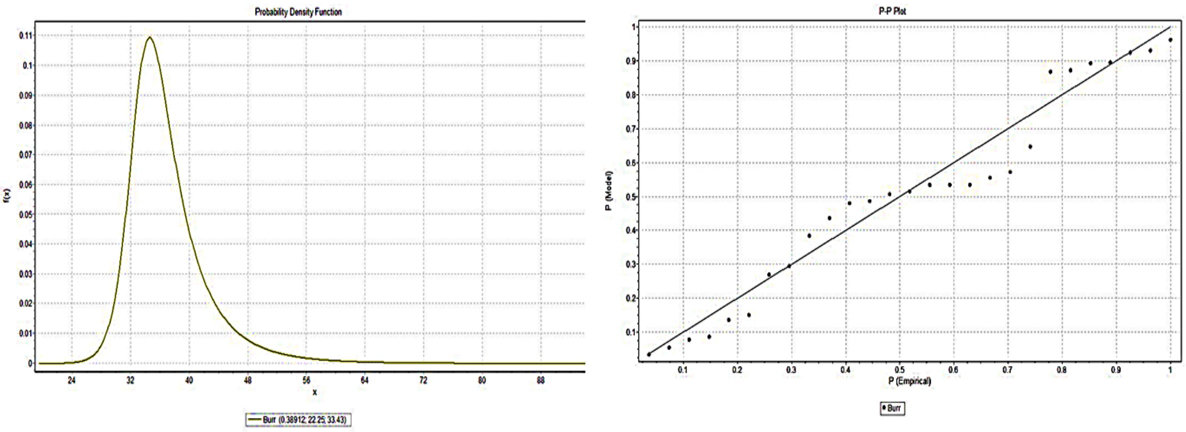

Fig. 3. NO_X (Post- monsoon)

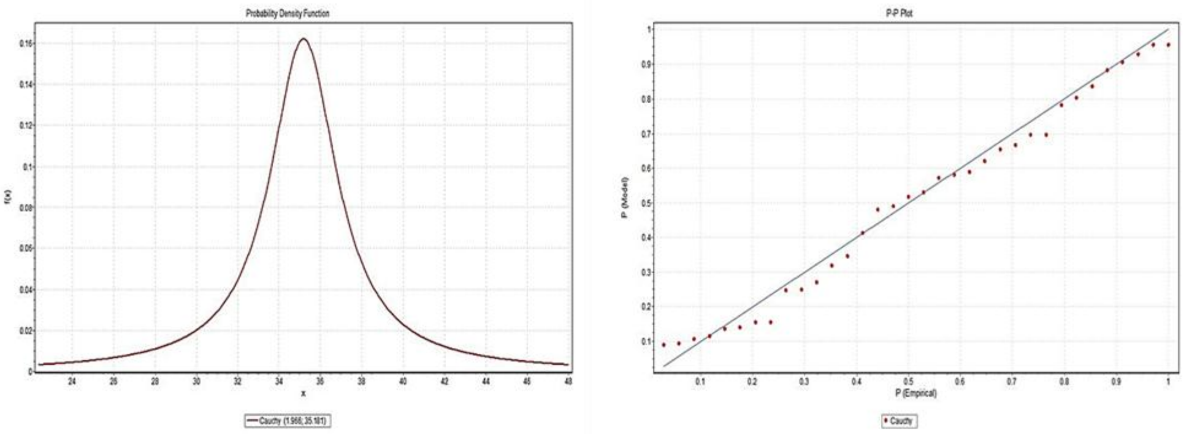

Fig. 4. NO_X (Monsoon)

Table- I: The characteristics and source of data used in the study

Parameter	Time period	Frequency	Source
RSPM	2005-2015	4-Hourly	JSPCB
SO_2	2005-2015	4-Hourly	JSPCB
NO_X	2005-2015	4-Hourly	JSPCB
Precipitation	2005-2015	Monthly	IMD
Temperature	2005-2015	Monthly	IMD
Relative Humidity	2005-2015	Monthly	IMD
Wind Speed	2005-2015	Monthly	IMD

Table- II: The goodness of fit results

	Winter		Summer	
	Distribution	*Kolmogorov Smirnov*	*Distribution*	*Kolmogorov Smirnov*
NO_X	Burr	0.16312	Burr	0.12455
SO_2	Burr	0.1256	Burr	0.14609
RSPM	Frechet (3P)	0.09468	Johnson SB	0.08317
	Monsoon		**Post-Monsoon**	
	Distribution	*Kolmogorov Smirnov*	*Distribution*	*Kolmogorov Smirnov*
NO_X	Cauchy	0.08995	Burr	0.13387
SO_2	Cauchy	0.08897	Cauchy	0.10555
RSPM	Johnson SB	0.09268	Johnson SB	0.07909

Table- III: Probability of exceedance

Sr. No.	Pollutants	Seasons	Probability of Exceedance of Indian NAAQS	Prediction (μg/m^3) with a 95% probability
1.	RSPM	Winter	1.0	134.52
		Summer	0.99	126.44
		Monsoon	0.89	96.719
		Post-Monsoon	0.95863	104.06
2.	SO_2	Winter	4.1154E-4	15.112
		Summer	7.3499E-4	16.186
		Monsoon	0.00906	12.585
		Post-Monsoon	0.01203	11.712
3.	NO_X	Winter	0.15407	28.85
		Summer	0.12798	28.516
		Monsoon	0.1233	22.768
		Post-Monsoon	0.20999	30.611

Table- IV: Descriptive Statistics

	Mean	Std. Deviation	N
RSPM	168.7296	45.98188	31
RAIN	103.6111	132.21469	31
TEMPmax	29.3945	3.95585	31
TEMPmin	17.0242	4.96021	31
RHhigh	93.5097	7.01318	31
RHlow	27.6335	14.62834	31
WS	5.7506	2.19886	31

The study of the seasonal concentration of pollutants indicates the relationship between air pollutants and meteorological parameters. The stepwise analysis was conducted to find an appropriate result. Histogram and normal probability plot have been shown in figures 4 and 5. The model is given as:

$$RSPM = 185.343 - 0.16\,Rain$$

(Resultant model is generated with the help of Table V, VI, VII and Table VIII**)**

IV. CONCLUSION

After analysing the results, it can be concluded that Statistics serves as an excellent technology to find the influences of meteorological parameters on the air pollution and also it has the potential to predict if the only distribution of the concentration over time is known.

ACKNOWLEDGEMENT

The authors would like to thank Jharkhand State pollution control board for providing data and Birla institutes of technology, for providing technical support.

Table- V: Correlations

		RSPM	RAIN	TEMPmax	TEMPmin	RHhigh	RHlow	WS
Pearson Correlation	*RSPM*	1.000	-.461	-.117	-.306	-.190	-.329	-.417
	RAIN	-.461	1.000	.226	.692	.374	.759	.654
	TEMPmax	-.117	.226	1.000	.820	-.511	-.124	.499
	TEMPmin	-.306	.692	.820	1.000	-.080	.385	.728
	RHhigh	-.190	.374	-.511	-.080	1.000	.457	.200
	RHlow	-.329	.759	-.124	.385	.457	1.000	.270
	WS	-.417	.654	.499	.728	.200	.270	1.000
Sig. (1-tailed)	*RSPM*	.	.005	.265	.047	.153	.035	.010
	RAIN	.005	.	.111	.000	.019	.000	.000
	TEMPmax	.265	.111	.	.000	.002	.253	.002
	TEMPmin	.047	.000	.000	.	.335	.016	.000
	RHhigh	.153	.019	.002	.335	.	.005	.141
	RHlow	.035	.000	.253	.016	.005	.	.071
	WS	.010	.000	.002	.000	.141	.071	.
N	*RSPM*	31	31	31	31	31	31	31
	RAIN	31	31	31	31	31	31	31
	TEMPmax	31	31	31	31	31	31	31
	TEMPmin	31	31	31	31	31	31	31
	RHhigh	31	31	31	31	31	31	31
	RHlow	31	31	31	31	31	31	31
	WS	31	31	31	31	31	31	31

Table- VI: Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.461[a]	.213	.185	41.50049	1.921

a. Predictors: (Constant), RAIN
b. Dependent Variable: RSPM

Table- VII:ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	13483.575	1	13483.575	7.829	.009[b]
	Residual	49946.418	29	1722.290		
	Total	63429.992	30			

a. Dependent Variable: RSPM
b. Predictors: (Constant), RAIN

Table- VIII: Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	185.343	9.530		19.449	.000	165.853	204.834
	RAIN	-.160	.057	-.461	-2.798	.009	-.278	-.043

a. Dependent Variable: RSPM

Retrieval Number: B7599129219/2020©BEIESP
DOI: 10.35940/ijitee.B7599.019320

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

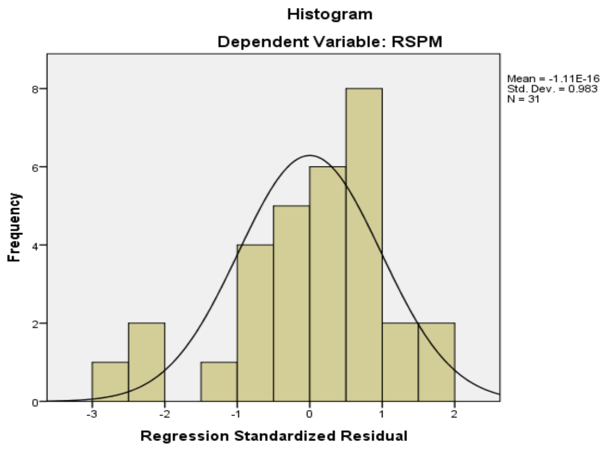

Fig. 4

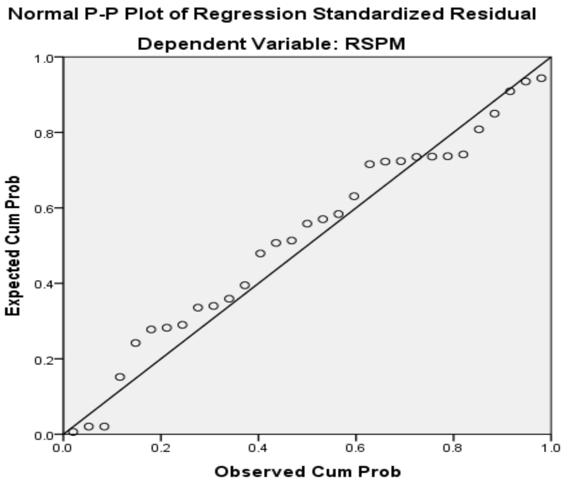

Fig. 5

REFERENCES

1. Lu, H.C.: Comparisons of statistical characteristic of air pollutants in Taiwan by frequency distribution.J. Air Waste Manag. Assoc. 53(5), 608–616 (2003)
2. Marchant, C., Leiva, V., Cavieres, M.F., Sanhueza, A.: Air contaminant statistical distributions with application to PM10 in Santiago, Chile. In: Whitacre, D.M. (ed.) Reviews of Environmental Contamination and Toxicology, vol. 223, pp. 1–31. Springer, New York (2013)
3. Mapoma, H.W.T., Tenthani, C., Tsakama, M., Kosamu, I.B.M.: Air quality assessment of carbon monoxide, nitrogen dioxide and sulfur dioxide levels in Blantyre, Malawi: a statistical approach to a stationary environmental monitoring station. Afr. J. Environ. Sci. Technol. 8(6), 330–343 (2014)
4. WHOReport: Ambient Air Pollution: A Global Assessment of Exposure and Burden of Disease. ISBN9789241511353 (2016)
5. Dubey, B., Pal, A.K., Singh, G.: Trace metal composition of airborne particulate matter in the coal mining and non-mining areas of Dhanbad region, Jharkhand, India. Atmos. Pollut. Res. 3(2), 238–246(2012)
6. Pandey, B., Agrawal, M., Singh, S.: Assessment of air pollution around coal mining area: emphasising on spatial distributions, seasonal variations and heavy metals, using cluster and principal component analysis. Atmos. Pollut. Res. 5(1), 79–86 (2014)
7. Kumar, P., „ C., Morawska, L., Norford, L., Choudhary, R., Bell, M., Leach, M.: Indoor air quality and energy management through real-time sensing in commercial buildings. Energy Build.111, 145–153 (2016)
8. Priyadarshi, N.: Environment and Geology (2016). http://nitishpriyadarshi.blogspot.in
9. Gilbert, R.O.: Statistical Methods for Environmental Pollution Monitoring. Wiley, Hoboken (1987)Mukhopadhyay, S.C.: Geomorphology of the Subarnarekha Basin: The Chota Nagpur Plateau, EasternIndia. The University of Burdwan, Bardhaman (1980)
10. Gupta, S.C., Kapoor, D.V.: Fundamentals of Mathematical Statistics: A Modern Approach. SultanChand, New Delhi (2000)

AUTHORS PROFILE

Tanushree Bhattacharya, Assistant Professor, Civil and Environment Engineering, M.Sc., M. Phil, PhD (Env. Sc.) (Jawaharlal Nehru University), NET-UGC-JRF (Env.Sc), CSIR-NET-JRF (Earth, Atmos., Ocean & Planetary Sc.)

Tripta Narayan, Research Scholar, BIT Mesra, and Assistant Professor at Amity University Patna, Bihar, India.

Soubhik Chakraborty, Professor and Head, Mathematics Ph.D., M.Sc. (Statistics), NET Qualified (UGC/CSIR) in Mathematical Sciences

Swapan Konar, Professor, Dean of Faculty Affairs and Sponsored Research, Physics M.Sc., M.Phil., M.Tech., PhD.

Dr. Shilpi Singh, currently works at the Amity School of Engineering and Technology, Amity University Patna. Shilpi does research in Software Testing and Software Engineering. Their current project is 'Design a Framework for Test Suite Optimization in Regression Testing'. Currently working on Security Testing.