

Assessment of Variation of Air Pollutant Concentration within Monitoring Stations in Ernakulam

Shilpa Anil¹, Sindhu P.²

Post Graduate Scholar, Department of Civil Engineering, College of Engineering, Trivandrum,
Thiruvananthapuram, India¹

Assistant Professor, Department of Civil Engineering, College of Engineering, Trivandrum,
Thiruvananthapuram, India²

ABSTRACT: This paper presents a study on the estimation of variation of air pollutant concentration in different areas based on the type of activities carried out in a region and ascertaining the need of separate monitoring stations in these areas. There are seven National Ambient Air Quality Monitoring Stations (NAAQMS) in Ernakulam which monitors Respirable Suspended Particulate Matter (RSPM), Suspended Particulate Matter (SPM), SO₂ and NO₂. Statistical approach consisting of Analysis of Variance (ANOVA) was applied to determine whether there is any significant difference in pollutant emissions between various monitoring stations. The temporal variation in the same period was also evaluated. ANOVA test helped in ascertaining the variation of pollutant concentration among the stations. Then, Tukey Kramer test was conducted to identify the stations which show significant difference in pollutant concentration for all the four pollutants. It was observed that in some stations there was no significant difference in pollutant concentration. Hence separate monitoring stations in such areas can be reduced. This data can be used as a basis for selecting monitoring stations in an area.

KEYWORDS: Tukey Kramer, ANOVA, null hypothesis, test statistics.

I. INTRODUCTION

Air pollution refers to the contamination of air which causes difficulty for plants, animals and humans to survive. In India, air pollution is quite a serious issue^[1]. Hence, monitoring stations are being set up under the National Air Quality Monitoring Programme (NAMP) across all the states so as to monitor the air quality by Central Pollution Control Board (CPCB). Many stations are being set up in different districts in Kerala as a part of this programme to measure the quality of air. Initially, the stations were categorized as located in industrial, residential and commercial zones. Later the air quality standards have been revised by the Central Pollution Control Board on April 11, 1994 and the classification of stations was reduced into two zones such as industrial, residential, rural and other areas and ecologically sensitive area^[2].

The main objective of the study was to assess whether there exists any significant difference in pollutant concentration between these monitoring stations. The temporal variation of air pollutants in each of these stations was carried out for a period of four years. Tukey Kramer test was then conducted to identify the pair of stations which show significant difference in pollutant concentration. The necessity of seven number of monitoring stations in Ernakulam was thus investigated.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

II. RELATED WORK

Rob et al., (2013) developed Land Use Regression (LUR) models in thirty-six study areas in Europe and explained the spatial variation of air pollutant concentration in each area. Centrally and locally available Geographic Information System (GIS) variables were used as predictors. Finally, a leave-one out cross-validation procedure was used to evaluate the model performance^[3]. Allaa et al., (2013) evaluated air quality in Baghdad city by analysis of heavy metals (Ni, Cu and Pb) using ANOVA two-way test. Baghdad's air pollution levels and the health and economic impact of those air pollution levels were assessed. ANOVA two-ways test with respect to space and time was used for the study. From the study, it was found that heavy metals Ni, Cu and Pb were one of the main reasons of air pollution in Baghdad city^[4]. Nur et al., (2015) carried out a preliminary assessment for pollutant sources identification using a statistical approach consisting of Analysis of Variance (ANOVA) and Principal Component Analysis (PCA). By ANOVA test, the significant difference in pollutant concentration between monitoring sites was assessed. Then, by PCA the various pollution sources in each monitoring site was identified based on the activities carried out in the area^[5]. Quinet al., (2015) explored the relationship between neighborhood definition and exposure assessment using twelve exposure assessment techniques. The pollutant concentration of each area was calculated based on pollution sources, topography, urban morphology, meteorology, etc. Exposure estimates were taken at different buffer zones around buildings and the estimates were found to be different for each buffer zone. This variation was then related to the definition of each neighborhood^[6]. Kamalet al., (2015) conducted a cluster analysis to reduce the air quality monitoring station network density in Delhi, India with a minimum loss of air quality data. The existing air quality monitoring stations have been grouped using fuzzy similarity measures. Non-linear mapping method and fuzzy c-mean (FCM) clustering method was used for further classification of monitoring stations. Finally, decision was taken based on the results of both fuzzy similarity measure and FCM^[7].

III. METHODOLOGY

The study area selected was Ernakulam, the industrial capital of Kerala. There are seven air quality monitoring stations in Ernakulam which are so located to cover industrial, residential and sensitive areas as shown in Table 1.

Table 1. NAAQMS in Ernakulam

Sl No.	Station name	Index	Category
1	Ernakulam South	S ₁	Residential
2	M. G. Road	S ₂	Residential
3	Vytilla	S ₃	Residential
4	Irumpanam	S ₄	Industrial
5	Kalamassery	S ₅	Industrial
6	Eloor 1	S ₆	Residential
7	Eloor 2	S ₇	Industrial

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

The monitoring stations were located at different spatial levels as shown in Fig. 1.

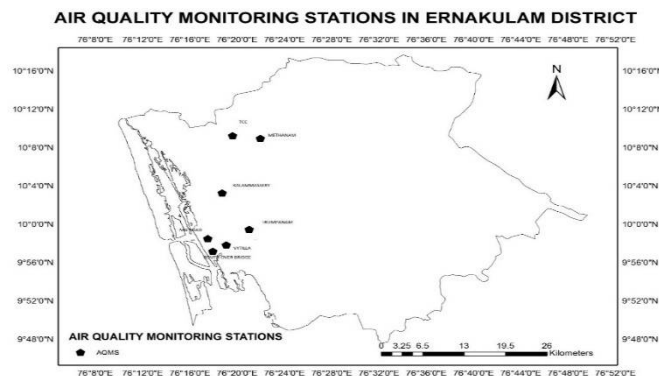


Fig. 1. Map showing Air Quality Monitoring Stations

Statistical analysis of data using Analysis of Variance was conducted to determine whether there is any variation in pollutant emission among stations. The stations where there was significant difference in pollutant concentration was identified through Tukey Kramer test. The temporal variation of air pollutant concentration was analysed over a period of four years. The need of seven air quality monitoring stations in Ernakulam was also ascertained. For the seven air quality monitoring stations, four-year daily data from 2011- 2014 of SPM, RSPM, SO₂ and NO₂ was collected from Kerala State Pollution Control Board (KSPCB), Ernakulam.

For conducting ANOVA test, first a null hypothesis was framed that the mean value of pollutant emissions is the same. The test statistics f and the critical value of f which was determined from ANOVA tables as a function of degrees of freedom were compared. If $f > f_{critical}$, then the null hypothesis is rejected. The p value, which is the probability of a value of f greater than or equal to the observed value and the fifth percentile value α were compared. α value is the probability of rejecting the null hypothesis given that it is true. It is usually set at or below 5%. If $p < \alpha$, the null hypothesis is rejected and we accept the alternate hypothesis that the mean value of pollutant emissions was different. The test statistics f is obtained from the mean sum.

The mean sum of squares is computed by dividing the sum of squares by the degree of freedom.

$$\text{Mean sum} = \frac{\text{Sum of Squares}}{\text{d.o.f}} \quad (1)$$

where,

$$\text{d.o.f} = \text{degree of freedom} = n - 1 \quad (2)$$

Three sum of squares were calculated- total sum of squares, between sum of squares and within sum of squares. Total Sum of Squares is the sum of the squares of the deviations of all the observations from their mean and it relates to the total variance of the observations. Total sum of squares can be partitioned into between sum of squares and within sum of squares.

Total Sum of Squares,

$$\text{TSS} = \text{Sum of square of each observation} - C \quad (3)$$

Where,

C = Correction factor

Correction factor was calculated by dividing the grand total square by number of observations.

$$\text{Correction factor, } C = \frac{(\text{sum of observations})^2}{\text{number of observations}} \quad (4)$$

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

Sum of squares between groups examines the differences among the group means by calculating the variation of each mean around the grand mean.

Between Sum of Squares,

$$BSS = \frac{(\text{Sum of observation in each group})^2}{\text{Number of observations in each group}} - C \quad (5)$$

Sum of squares within groups examines variation of individual scores around each group mean.

Within Sum of Squares,

$$\text{Within Sum of Squares} = TSS - BSS \quad (6)$$

Tukey Kramer test was conducted to identify the stations which show significant difference in pollutant concentration. The test was conducted by using the software Graph Pad Prism. If the test statistics q is greater than the critical value of q i.e., $q > q_{\text{critical}}$, then means are said to be significantly different and vice versa. First step was to find the absolute difference between average of each pair of group to be compared. Then the test statistics q was determined.

Test statistics,

$$q = \frac{Y_A - Y_B}{SE} \quad (7)$$

Where,

Y_A = larger of the two means being compared

Y_B = smaller of the two means being compared

$$SE = \text{Standard Error} = s \sqrt{\frac{2}{n}} \quad (8)$$

Where, s = variance

n = number of observations

IV. EXPERIMENTAL RESULTS

ANOVA test was conducted to identify the variation in pollutant concentration between different monitoring stations using Microsoft Excel and it yielded the following results as shown in Table 2.

Table 2. Variation of RSPM Concentration Between Stations In 2014

Anova: Single Factor						
SUMMARY						
Groups	Count	Sum	Average	Variance		
S ₁	12	836.416	69.701	395.107		
S ₂	12	854.929	71.244	326.423		
S ₃	12	817.972	68.164	749.252		
S ₄	12	1279.615	106.635	2886.666		
S ₅	12	563.118	46.926	200.269		
S ₆	12	630.957	52.580	120.185		
S ₇	12	655.285	54.607	179.647		

ANOVA						
Source of Variation	SS	Df	MS	f	P-value	f _{crit}
Between Groups	28343.213	6.000	4723.869	6.807	8.1×10^{-06}	2.219
Within Groups	53433.036	77.000	693.936			
Total	81776.249	83.000				

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

From the above results, it was observed that the two conditions $f > f_{critical}$ and $p < \alpha$ were satisfied. Therefore, we reject the null hypothesis and accept the alternate hypothesis. Thus, it was concluded that there exists significant difference in RSPM concentration between stations in 2014. Similarly, ANOVA test was conducted for RSPM for the subsequent years and for the remaining pollutants for the four-year time period. From the results of ANOVA, it was observed that except for SO₂ concentration in the years 2012 and 2013, in all other cases there exists significant difference in pollutant concentration.

The temporal variation of each pollutant in each monitoring station was determined in the same manner as that of the variation of pollutants among stations and was as shown in Table 3.

Table 3. Temporal Variation of RSPM Concentration in M.G. Road

**Anova:
Single
Factor**

SUMMARY

Groups	Count	Sum	Average	Variance
2011	12.000	479.544	39.962	254.316
2012	12.000	798.604	66.550	1199.877
2013	12.000	893.947	74.496	590.823
2014	12.000	854.929	71.244	326.423

ANOVA

Source of Variation	SS	Df	MS	f	P-value	f _{crit}
Between Groups	8921.398	3.000	2973.799	5.016	0.004	2.816
Within Groups	26085.830	44.000	592.860			
Total	35007.229	47.000				

In this case also, it was observed that the conditions $f > f_{critical}$ and $p < \alpha$ were satisfied. Thus, the null hypothesis was rejected and alternate hypothesis was accepted. Thus, it was concluded that there exists significant difference in RSPM concentration over the years in M.G. Road. ANOVA test was conducted for RSPM for the remaining stations and for the remaining pollutants over the years and it was observed that almost all the stations show variation of pollutant concentration over the years. But, RSPM concentration in Kalamassery and SPM concentration in Ernakulam South were found to be free from temporal variation.

Tukey Kramer test was conducted to identify the pair of stations which show significant difference in pollutant concentration. A sample of the results obtained from the software is shown in Table 4.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

Table 4. Tukey Kramer Result of Variation in RSPM Concentration between Stations in 2014

Tukey's multiple comparisons test	Mean Diff.	95% CI of diff.	Significant?	Summary	Adjusted P Value
S ₁ vs. S ₂	-1.54	-34.1 to 31.0	No	Ns	> 0.9999
S ₁ vs. S ₃	1.54	-31.0 to 34.1	No	Ns	> 0.9999
S ₁ vs. S ₄	-36.9	-69.5 to -4.37	Yes	*	0.0160
S ₁ vs. S ₅	22.8	-9.79 to 55.3	No	Ns	0.3529
S ₁ vs. S ₆	17.1	-15.4 to 49.7	No	Ns	0.6876
S ₁ vs. S ₇	15.1	-17.5 to 47.7	No	Ns	0.7983
S ₂ vs. S ₃	3.08	-29.5 to 35.6	No	Ns	> 0.9999
S ₂ vs. S ₄	-35.4	-68.0 to -2.83	Yes	*	0.0243
S ₂ vs. S ₅	24.3	-8.24 to 56.9	No	Ns	0.2763
S ₂ vs. S ₆	18.7	-13.9 to 51.2	No	Ns	0.5948
S ₂ vs. S ₇	16.6	-15.9 to 49.2	No	Ns	0.7157
S ₃ vs. S ₄	-38.5	-71.0 to -5.91	Yes	*	0.0104
S ₃ vs. S ₅	21.2	-11.3 to 53.8	No	Ns	0.4388
S ₃ vs. S ₆	15.6	-17.0 to 48.1	No	Ns	0.7732
S ₃ vs. S ₇	13.6	-19.0 to 46.1	No	Ns	0.8674
S ₄ vs. S ₅	59.7	27.1 to 92.3	Yes	****	< 0.0001
S ₄ vs. S ₆	54.1	21.5 to 86.6	Yes	****	< 0.0001
S ₄ vs. S ₇	52.0	19.5 to 84.6	Yes	***	0.0001
S ₅ vs. S ₆	-5.65	-38.2 to 26.9	No	Ns	0.9984
S ₅ vs. S ₇	-7.68	-40.2 to 24.9	No	Ns	0.9914
S ₆ vs. S ₈	-2.03	-34.6 to 30.5	No	Ns	> 0.9999

Test details	Mean Mean Mean SE of							
	1	2	Diff.	diff.	n ₁	n ₂	q	DF
S ₁ vs. S ₂	69.7	71.2	-1.54	10.8	12	12	0.203	77
S ₁ vs. S ₃	69.7	68.2	1.54	10.8	12	12	0.202	77
S ₁ vs. S ₄	69.7	107	-36.9	10.8	12	12	4.86	77
S ₁ vs. S ₅	69.7	46.9	22.8	10.8	12	12	2.99	77

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

S ₁ vs. S ₆	69.7	52.6	17.1	10.8	12	12	2.25	77
S ₁ vs. S ₇	69.7	54.6	15.1	10.8	12	12	1.99	77
S ₂ vs. S ₃	71.2	68.2	3.08	10.8	12	12	0.405	77
S ₂ vs. S ₄	71.2	107	-35.4	10.8	12	12	4.65	77
S ₂ vs. S ₅	71.2	46.9	24.3	10.8	12	12	3.20	77
S ₂ vs. S ₆	71.2	52.6	18.7	10.8	12	12	2.45	77
S ₂ vs. S ₇	71.2	54.6	16.6	10.8	12	12	2.19	77
S ₃ vs. S ₄	68.2	107	-38.5	10.8	12	12	5.06	77
S ₃ vs. S ₅	68.2	46.9	21.2	10.8	12	12	2.79	77
S ₃ vs. S ₆	68.2	52.6	15.6	10.8	12	12	2.05	77
S ₃ vs. S ₇	68.2	54.6	13.6	10.8	12	12	1.78	77
S ₄ vs. S ₅	107	46.9	59.7	10.8	12	12	7.85	77
S ₄ vs. S ₆	107	52.6	54.1	10.8	12	12	7.11	77
S ₄ vs. S ₇	107	54.6	52.0	10.8	12	12	6.84	77
S ₅ vs. S ₆	46.9	52.6	-5.65	10.8	12	12	0.743	77
S ₅ vs. S ₇	46.9	54.6	-7.68	10.8	12	12	1.01	77
S ₆ vs. S ₈	52.6	54.6	-2.03	10.8	12	12	0.267	77

Here, significant difference in pollutant concentration was observed between the pairs of stations S₁ vs. S₄, S₂ vs. S₄, S₃ vs. S₄, S₄ vs. S₅, S₄ vs. S₆ and S₄ vs. S₇. This was because, the q value was greater than the q_{critical} value and also the p value was less than α . Among the pair of stations, S₄ vs. S₅ and S₄ vs. S₆ have shown the greatest value of a four-star significant difference. Similarly, for the all the pollutants for the four year period the test was conducted and the stations with significant difference were identified.

The Tukey Kramer results of variation of RSPM concentration during the four years was as shown in Table 5.

Table 5. Tukey Kramer Results of Variation of RSPM Concentration

Sl no.	Stations Considered	Significant difference among stations? (Yes/ No)			
		2011	2012	2013	2014
1	S ₁ vs. S ₂	No	No	No	No
2	S ₁ vs. S ₃	No	No	Yes	No
3	S ₁ vs S ₄	No	No	No	Yes

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

4	S ₁ vs. S ₅	No	No	No	No
5	S ₁ vs. S ₆	Yes	No	No	No
6	S ₁ vs. S ₇	Yes	No	No	No
7	S ₂ vs. S ₃	No	No	Yes	No
8	S ₂ vs. S ₄	No	No	Yes	Yes
9	S ₂ vs. S ₅	Yes	No	No	No
10	S ₂ vs. S ₆	Yes	No	No	No
11	S ₂ vs. S ₇	Yes	No	No	No
12	S ₃ vs. S ₄	No	No	Yes	Yes
13	S ₃ vs. S ₅	No	No	No	No
14	S ₃ vs. S ₆	Yes	No	Yes	No
15	S ₃ vs. S ₇	Yes	No	No	No
16	S ₄ vs. S ₅	Yes	No	Yes	Yes
17	S ₄ vs. S ₆	Yes	Yes	Yes	Yes
18	S ₄ vs. S ₇	Yes	No	Yes	Yes
19	S ₅ vs. S ₆	Yes	No	No	No
20	S ₅ vs. S ₇	Yes	No	No	No
21	S ₆ vs. S ₇	No	No	No	No

For each pollutant in each year the variation was assessed. In the case of RSPM, the pair of stations which have shown no significant difference in pollutant concentration during the four year period were S₁ vs. S₂, S₁ vs. S₅, S₃ vs. S₅, S₆ vs. S₇. Similarly, the test was conducted for the remaining pollutants and the stations where there is no significant difference in pollutant concentration was identified.

The results of Tukey Kramer test were summed up and the stations where there is no variation in the concentration of any of the pollutants were identified and as shown in Table6.

Table 6. Stations showing no Significant Variation in Pollutant Concentration

Pollutants	Monitoring stations showing no variation
RSPM	S ₁ - S ₂ , S ₅ S ₃ - S ₅ S ₆ - S ₇
SPM	S ₁ - S ₂ , S ₅ S ₂ - S ₅ , S ₆ , S ₇ S ₃ - S ₇ S ₆ - S ₇
NO ₂	S ₁ - S ₂ , S ₃ , S ₅ S ₃ - S ₅ S ₆ vs. S ₇
SO ₂	S ₁ - S ₂ , S ₃ , S ₄ , S ₅ S ₂ - S ₃ , S ₅ S ₃ - S ₄ , S ₅ S ₄ - S ₅ , S ₇ S ₆ - S ₇

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

It was observed that in the case of each of the four pollutants the stations Ernakulam South, M.G. Road and Kalamassery did not show significant variation in concentration. The stations Eloor 1 and Eloor 2 also did not show any significant variation in pollutant concentration. Thus, the number of monitoring stations in these areas can be reduced.

V. CONCLUSION

The variation of air pollutant concentration within the monitoring stations in Ernakulam was assessed. By Analysis of Variance and Tukey Kramer test the stations which have shown significant difference in pollutant concentration were identified. The temporal variation of pollutant concentration for RSPM, SPM SO₂ and NO₂ was also determined. Thus, it was observed that seven monitoring stations in Ernakulam is not required as there is no significant difference in pollutant concentration. It was found that among the three monitoring stations Ernakulam South, M.G. Road and Kalamassery only one station is required and among Eloor 1 and Eloor 2 only one station is required.

ACKNOWLEDGEMENT

The authors thank the Department of Civil Engineering of College of Engineering Trivandrum for providing the necessary facilities for carrying out the study. The authors are grateful to the Kerala State Pollution Control Board for providing pollutant concentration data used in the study.

REFERENCES

- [1] Congbo Songa, Lin Wua, Yaochen Xieb, Jianjun Hec, Xi Chend, Ting Wanga, Yingchao Lina, Taosheng Jina, Anxu Wang, Yan Liua, Qili Daia, Baoshuang Liua, Ya-nan Wanga, and Hongjun Mao, "Air Pollution in China: Status and spatiotemporal variations", *Environmental Pollution*, Vol.227, pp.334–347, 2017.
- [2] www.cpcb.nic.in.
- [3] Rob Beelen, Gerard Hoek, Danielle Vienneau, and Marloes Eeftens, "Development of NO₂ and NO_x land use regression models for estimating air pollution exposure in 36 study areas in Europe The ESCAPE project", *Atmospheric Environment*, Vol.72, pp.10, 2013.
- [4] Allaa, M., Aenab, Singh, S.K., and Ali Jabir Lafta, "Critical assessment of air pollution by ANOVA test and human health effects", *Atmospheric Environment*, Vol.71, pp.84-91, 2013.
- [5] Nur Diyana Mohamada, Zulfa Hanan Ash, and Melawani Othman, "Preliminary assessment of air pollutant sources identification at selected monitoring stations in Klang Valley, Malaysia", *Procedia Environmental Sciences*, Vol.30, pp.121 – 126, 2015.
- [6] Quentin, M. T., Frederic Mauny, Daniel Joly, Stephane Francois, and Nadine Bernard, "Air pollution in moderately polluted urban areas: How does the definition of "neighborhood" impact exposure assessment?", *Environmental Pollution*, Vol.206, pp.437-448, 2015.
- [7] Kamal Jyoti Maji, Anil Kumar Dikshit, and Satish Salunkhe, "Adequacy Assessment of Air Quality Monitoring Stations in Delhi, India using Fuzzy Similarity and Fuzzy c-Mean Clustering", *International Journal of Scientific & Engineering Research*, Vol.6, pp.576-581, 2015.