Journal of Basic and Applied Engineering Research
p-ISSN: 2350-0077; e-ISSN: 2350-0255; Volume 6, Issue 7; July-September, 2019, pp. 396-401

http://www.krishisanskriti.org/Publication.html

Comparative Analysis of Weekly AQI Data among Selected Cities of India

Pareek N.[1], Sarmah P.[2] and Choudhury A.[3]

[1]*Research Scholar, Department of Statistics,*
Gauhati University, Guwahati, Assam, India
[2]*Professor (RTD), Department of Statistics,*
Gauhati University, Guwahati, Assam, India
[3]*Director, IDOL, Guwahati, Assam, India*
E-mail: [1]imp.info.namrata@gmail.com, [2]pranitasarma@gmail.com, [3]achoudhury@rediffmail.com

Abstract—*There has been a global concern on the ill effects of air pollution. The concern has been aggravated because even though a lot of awareness has been generated on the ill effects of air pollution yet the levels do not seem to come down even though the rate of deterioration has somewhat slackened. India and most of its cities have also managed to take the top ranks in being among the most polluted regions in the world. Air Quality Index (AQI) is a number that is used by government agencies to provide an overall measure of how polluted the air is. In this paper, we study the air pollution levels of some selected cities of India. In addition to data analysis through graphs, autocorrelation tests have been carried out for each of the selected cities. Further, a comparison has been made between the weekly AQI of each of the cities for the years 2017 and 2018 which show that the air quality has deteriorated in Chennai and Mumbai while in Kolkata, Delhi, Varanasi and Gurugram it has shown improvements in 2018 compared to 2017.*

Keywords: *Air Quality Index (AQI), weekly AQI, line graphs, autocorrelation, paired t-tests.*

1. INTRODUCTION

Air pollution can be defined, in simple terms, as the contamination of air by harmful particles that deteriorate the functioning of the environment. There is no part of the environment that has managed to save itself from the adverse affects of air pollution. Right from the smoke produced in households to the smoke emitted by factories, every activity on earth, especially by humans, is contributing to the rise of air pollution. It poses health risks among humans and animals. The World Health Organisation reports over the years have shown that India takes top ranks in being among one of the most polluted places of the world and hence is one of the reasons for concern. The government agencies have been taking certain measures to curb pollution by creating awareness among people. Yet, there is a lot on the plate that will require time and efforts. The State of Global Air 2019 report looks at how long-term exposure to outdoor and indoor air pollution has affected health. Five million cases of death from stroke, diabetes, heart attack, lung cancer, and chronic lung disease the world over were linked with air pollution in 2017, according to the report. Of these, 1.5 million were in India and China, directly related to PM 2.5 levels. [1]

An Air Quality Index is used by government agencies to communicate to the public how polluted the air is. or how polluted it is forecast to become. Public health risks increase as the AQI rises. Different countries have their own air quality indices corresponding to different national air quality standards. [2]

The Air Quality Index (AQI) was launched in India on Sept 17, 2014 under the Swachh Bharat Abhiyan. The Central Pollution Control Board along with State Pollution Control Boards has been operating National Air Monitoring Program covering 240 cities of the country having more than 342 monitoring stations.[2] The AQI ranges from 0 to 500 which is divided into six categories, namely Good (0-50), Satisfactory (51-100), Moderately polluted (101-200), Poor (201-300), Very Poor (301-400), and Severe (401-500). The proposed AQI will consider eight pollutants (PM_{10}, $PM_{2.5}$, NO_2, SO_2, CO, O_3, NH_3, and Pb) for which short-term (up to 24-hourly averaging period) National Ambient Air Quality Standards are prescribed. Based on the measured ambient concentrations, corresponding standards and likely health impact, a sub-index is calculated for each of these pollutants.[2]

In this paper, the weekly AQI data for the years 2017 and 2018 has been obtained with the help of the Central Pollution Control Board (CPCB) that publishes the AQI levels of Indian cities on a daily basis. The data that has been collected for this paper is secondary in nature.

2. OBJECTIVES

This paper has the following objectives: (1) Assessment of weekly AQI data for the years 2017 and 2018. (2) Autocorrelation analysis of the data. (3) Comparison of the AQI levels between 2017 and 2018.

3. MATERIALS AND METHODS

3.1 Line Graphs

A **line graph** is a graphical display of information that changes continuously over time. It can be used to compare different events, situations, and information. [3] This paper uses line graphs to observe the changes in the AQI levels for each of the cities over the two-year period (2017-2018).

3.2 Autocorrelation

In statistics, the autocorrelation of a random process is the Pearson correlation between values of the process at different times, as a function of the two times or of the time lag. Let X be a random process, and t be any point in time (t may be an integer for a discrete-time process or a real number for a continuous-time process). Then X_t is the value (or realization) produced by a given run of the process at time t. Suppose that the process has mean μ_t and variance σ_t^2 at time t, for each t. Then the definition of the autocorrelation between times s and t is

$$R(s,t) = \frac{\mathrm{E}[(X_t - \mu_t)(X_s - \mu_s)]}{\sigma_t \sigma_s}$$

where "E" is the expected value operator. [4]

This paper includes the calculation of autocorrelation upto lag 5 for the AQI values for each of the cities considered.

3.3 Paired t-test

The paired sample t-test, sometimes called the dependent sample t-test, is a statistical procedure used to determine whether the mean difference between two sets of observations is zero. In a paired sample t-test, each subject or entity is measured twice, resulting in pairs of observations. Common applications of the paired sample t-test include case-control studies or repeated-measures designs.

The paired sample *t*-test hypotheses are formally defined as follows: (1) The null hypothesis (H_0) assumes that the true mean difference is equal to zero. (2) The two-tailed alternative hypothesis (H_1) assumes that is not equal to zero.

The paired sample t-test has four main assumptions: (1) The dependent variable must be continuous (interval/ratio). (2) The observations are independent of one another. (3) The dependent variable should be approximately normally distributed. (4) The dependent variable should not contain any outliers.[5]

In this paper, a normality test has been performed to check if the AQI data follow the normality assumption after which paired t-tests have been performed to compare the 2017 and 2018 average air quality levels of each of the cities.

4. ANALYSIS OF THE DATA

4.1 Analysis-1: Assessment of the data

Following are the line graphs of the weekly AQI data for each of the cities for the years 2017 and 2018.

Table 4.1.1: Line graph of AQI data for Chennai

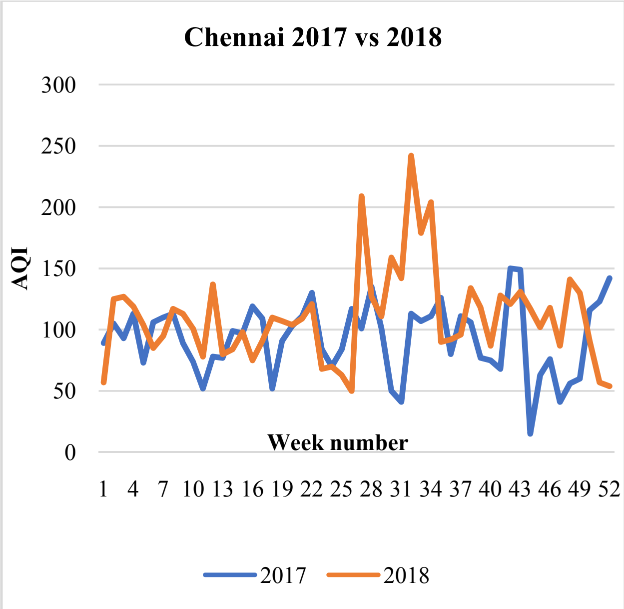


Table 4.1.2: Line graph of AQI data for Delhi

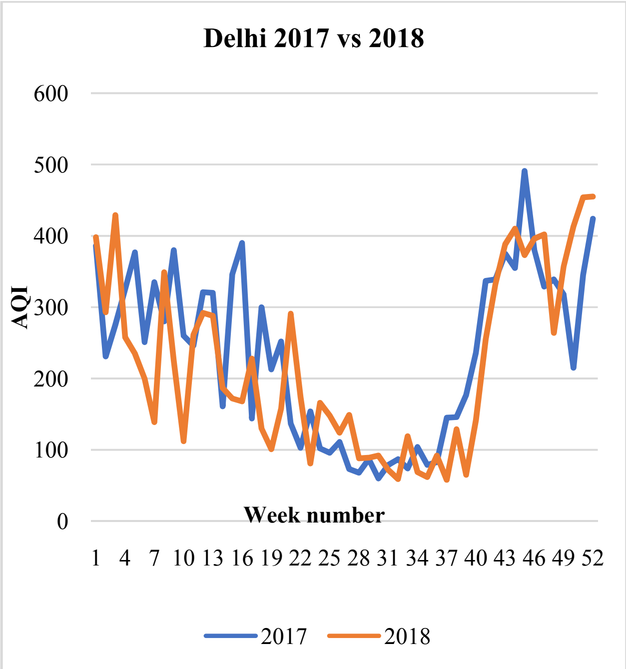


Table 4.1.3: Line graph of AQI data for Gurugram

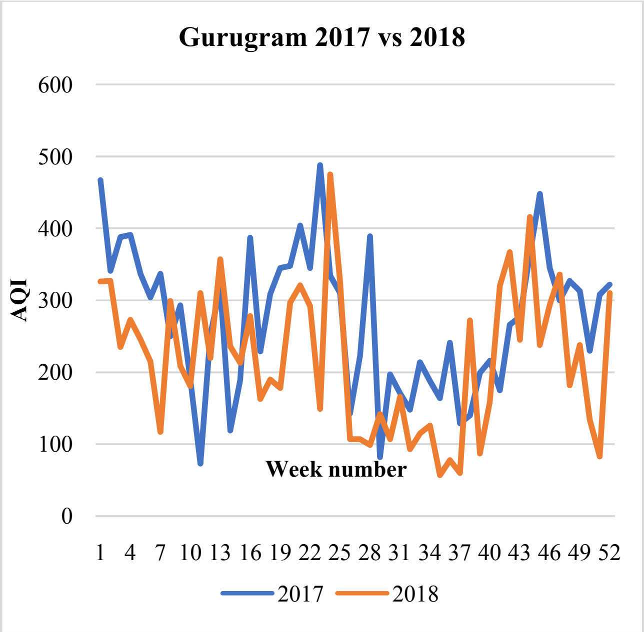


Table 4.1.5: Line graph of AQI data for Mumbai

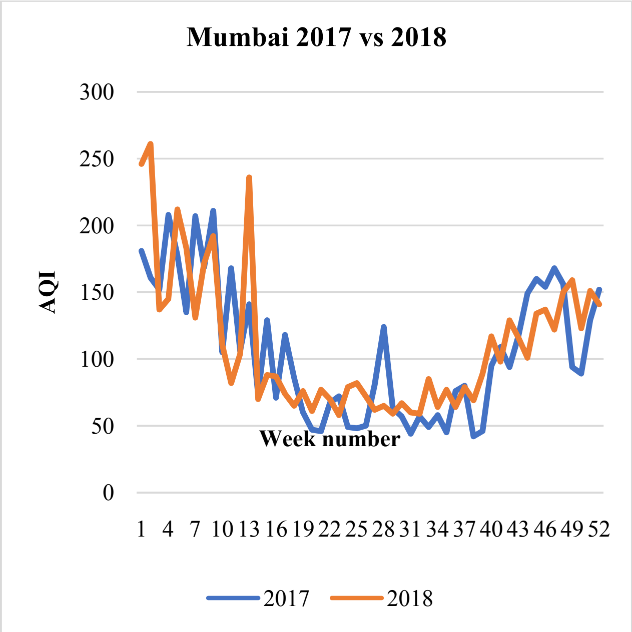


Table 4.1.4: Line graph of AQI data for Kolkata

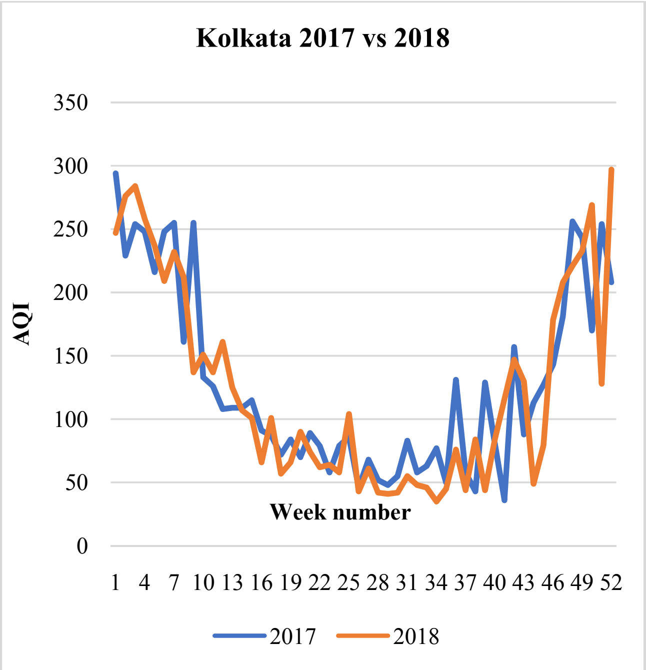
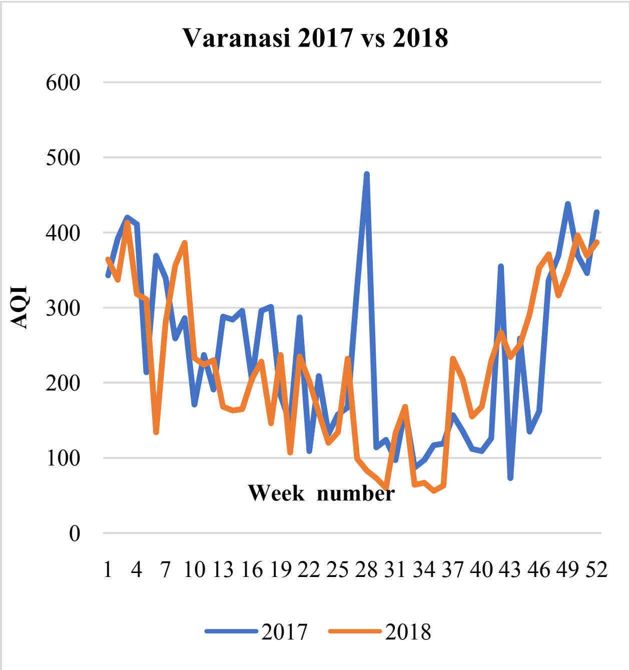


Table 4.1.6: Line graph of AQI data for Varanasi



4.2 Analysis-2: Autocorrelation tests

Table 4.2.1:Autocorrelation of Chennai AQI (2017)

Series: Chennai_17

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.248	.135	3.378	1	.066
2	-.174	.133	5.072	2	.079
3	-.188	.132	7.092	3	.069
4	-.197	.131	9.367	4	.053
5	-.065	.129	9.620	5	.087

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Table 4.2.2:Autocorrelation of Delhi AQI (2017)

Series: Delhi_17

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.714	.135	28.033	1	.000
2	.676	.133	53.714	2	.000
3	.701	.132	81.877	3	.000
4	.614	.131	103.927	4	.000
5	.445	.129	115.735	5	.000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Table 4.2.3: Autocorrelation of Gurugram AQI (2017)

Series: Gurugram_17

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.445	.135	10.926	1	.001
2	.344	.133	17.585	2	.000
3	.438	.132	28.565	3	.000
4	.198	.131	30.858	4	.000
5	.077	.129	31.217	5	.000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Table 4.2.4: Autocorrelation of Kolkata AQI (2017)

Series: Kolkata_17

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.736	.135	29.864	1	.000
2	.704	.133	57.714	2	.000
3	.684	.132	84.488	3	.000
4	.506	.131	99.481	4	.000
5	.423	.129	110.166	5	.000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Table 4.2.5: Autocorrelation of Mumbai AQI (2017)

Series: Mumbai_17

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.693	.135	26.436	1	.000
2	.690	.133	53.176	2	.000
3	.553	.132	70.677	3	.000
4	.606	.131	92.159	4	.000
5	.421	.129	102.740	5	.000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Table 4.2.6: Autocorrelation of Varanasi AQI (2017)

Series: Varanasi_17

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.477	.135	12.544	1	.000
2	.412	.133	22.096	2	.000
3	.278	.132	26.513	3	.000
4	.204	.131	28.947	4	.000
5	.204	.129	31.441	5	.000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Table 4.2.7: Autocorrelation of Chennai AQI (2018)

Series: Chennai_18

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.342	.135	6.446	1	.011
2	.215	.133	9.051	2	.011
3	.036	.132	9.126	3	.028
4	.005	.131	9.127	4	.058
5	.192	.129	11.331	5	.045

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Table 4.2.8: Autocorrelation of Delhi AQI (2018)

Series: Delhi_18

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.739	.135	30.056	1	.000
2	.605	.133	50.582	2	.000
3	.529	.132	66.635	3	.000
4	.457	.131	78.870	4	.000
5	.428	.129	89.822	5	.000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Table 4.2.9: Autocorrelation of Gurugram AQI (2018)

Series: Gurugram_18

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.330	.135	5.992	1	.014
2	.267	.133	9.990	2	.007
3	.307	.132	15.405	3	.002
4	.112	.131	16.135	4	.003
5	.045	.129	16.256	5	.006

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Table 4.2.10: Autocorrelation of Kolkata AQI (2018)

Series: Kolkata_18

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.773	.135	32.873	1	.000
2	.731	.133	62.879	2	.000
3	.604	.132	83.808	3	.000
4	.525	.131	99.917	4	.000
5	.450	.129	111.994	5	.000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Table 4.2.11: Autocorrelation of Mumbai AQI (2018)

Series: Mumbai_18

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.600	.135	19.850	1	.000
2	.412	.133	29.393	2	.000
3	.540	.132	46.134	3	.000
4	.577	.131	65.597	4	.000
5	.364	.129	73.521	5	.000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Table 4.2.12: Autocorrelation of Varanasi AQI (2018)

Series: Varanasi_18

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	.725	.135	28.950	1	.000
2	.596	.133	48.932	2	.000
3	.469	.132	61.560	3	.000
4	.489	.131	75.553	4	.000
5	.468	.129	88.652	5	.000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

4.3 Analysis-3: Paired t-tests

The null hypothesis to be tested is that there is no difference between the mean AQI values of 2017 and 2018.

Table 4.3.1: One sample Kolmogorov-Smirnov test for the normality of 2017 AQI data

One-Sample Kolmogorov-Smirnov Test

		Chennai_17	Delhi_17	Gurugram_17	Kolkata_17	Mumbai_17	Varanasi_17
N		52	52	52	52	52	52
Normal Parameters[a,b]	Mean	92.90	235.40	274.37	128.33	106.27	237.08
	Std. Deviation	29.056	118.964	98.643	73.981	50.360	112.731
Most Extreme Differences	Absolute	.091	.141	.103	.174	.115	.144
	Positive	.052	.121	.066	.174	.115	.144
	Negative	-.091	-.141	-.103	-.113	-.101	-.084
Kolmogorov-Smirnov Z		.653	1.016	.739	1.254	.830	1.040
Asymp. Sig. (2-tailed)		.788	.253	.645	.086	.495	.230

a. Test distribution is Normal.

b. Calculated from data.

Table 4.3.2: One sample Kolmogorov-Smirnov test for the normality of 2018 AQI data

One-Sample Kolmogorov-Smirnov Test

		Chennai_18	Delhi_18	Gurugram_18	Kolkata_18	Mumbai_18	Varanasi_18
N		52	52	52	52	52	52
Normal Parameters[a,b]	Mean	110.67	219.06	218.65	124.25	110.54	221.65
	Std. Deviation	38.269	123.789	99.410	79.522	51.751	103.787
Most Extreme Differences	Absolute	.125	.142	.078	.155	.181	.101
	Positive	.125	.142	.078	.155	.181	.101
	Negative	-.056	-.097	-.077	-.131	-.155	-.081
Kolmogorov-Smirnov Z		.898	1.024	.559	1.119	1.302	.730
Asymp. Sig. (2-tailed)		.395	.245	.914	.163	.067	.661

a. Test distribution is Normal.

b. Calculated from data.

Table 4.3.3: Paired sample statistics

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Chennai_17	92.90	52	29.056	4.029
	Chennai_18	110.67	52	38.269	5.307
Pair 2	Delhi_17	235.40	52	118.964	16.497
	Delhi_18	219.06	52	123.789	17.166
Pair 3	Gurugram_17	274.37	52	98.643	13.679
	Gurugram_18	218.65	52	99.410	13.786
Pair 4	Kolkata_17	128.33	52	73.981	10.259
	Kolkata_18	124.25	52	79.522	11.028
Pair 5	Mumbai_17	106.27	52	50.360	6.984
	Mumbai_18	110.54	52	51.751	7.177
Pair 6	Varanasi_17	237.08	52	112.731	15.633
	Varanasi_18	221.65	52	103.787	14.393

Table 4.3.4: Paired samples Correlations table

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Chennai_17 & Chennai_18	52	-.009	.949
Pair 2	Delhi_17 & Delhi_18	52	.708	.000
Pair 3	Gurugram_17 & Gurugram_18	52	.378	.006
Pair 4	Kolkata_17 & Kolkata_18	52	.840	.000
Pair 5	Mumbai_17 & Mumbai_18	52	.720	.000
Pair 6	Varanasi_17 & Varanasi_18	52	.542	.000

Table 4.3.5: Paired t-test for difference of means

Pair	Means	t	df	Sig. (2-tailed)
Chennai 2017 & 2018	-17.769	-2.655	51	0.11
Delhi 2017 & 2018	16.346	1.268	51	0.210
Gurugram 2017 & 2018	55.712	3.639	51	0.001

Kolkata 2017 & 2018	4.077	0.673	51	0.504
Mumbai 2017 & 2018	-4.269	-0.805	51	0.424
Varanasi 2017 & 2018	15.423	1.071	51	0.289

5. INTERPRETATION AND DISCUSSION

From analysis-1, it has been observed that there has been fluctuations in the AQI levels in Chennai, Gurugram and Varanasi while for Delhi, Mumbai and Kolkata, it has been more or less the same for the years 2017 and 2018.

Analysis-2 suggests significant autocorrelations for the AQI data of Delhi, Gurugram, Kolkata, Mumbai and Varanasi. Interestingly, there has been no indication of a significant autocorrelation among the AQI values of Chennai city.

From Analysis-3, it can be seen that the data under consideration is normal. Further, the paired t-test indicate a significant difference between the mean AQI of Gurugram in 2018 compared to 2017, while for the rest of the five cities, the test is insignificant. Also, the mean AQI of Gurugram in 2018 has decreased significantly compared to the mean AQI of 2017. These results suggest that there has been improvements in the air quality in 2018 in Gurugram, a city which otherwise is reputed to have horrendous pollution levels. So far as the other cities are concerned, Chennai and Mumbai have somehow managed to deteriorate its air quality over the one-year period. Kolkata, Delhi and Varanasi have shown slight and insignificant improvements which are not enough as per what the insignificant values of the paired t-tests indicate.

6. CONCLUSION

The average AQI levels show interesting results each year, specially when it comes to India. The analyses indicate how Chennai and Mumbai have shown degradation in the air quality and that serious actions need to be taken to improve it. Though the other four cities have shown slight improvements, it may also be due to a chance cause. These six cities are among the big cities of India which contribute to a large amount of air pollution in the country and the world as a whole.

REFERENCES

[1] World Health Day 2019: India accounts for quarter of the World's Air Pollution Deaths, 2019, accessed 29 August, 2019, <https://www.firstpost.com/tech/science/world-health-day-2019-india-accounts-for-quarter-of-the-worlds-air-pollution-deaths-6386671.html>

[2] Air Quality Index, Wikipedia, the Free Encyclopedia, accessed 29 August, 2019, <https://en.wikipedia.org/wiki/Air_quality_index>

[3] What is a line graph? Definition and examples, accessed, Chapter 1, Lesson 12, accessed 29 August, 2019, <https://study.com/academy/lesson/what-is-a-line-graph-definition-examples.html>

[4] Autocorrelation, Wikipedia, the Free Encyclopedia, accessed 29 August, 2019, <https://en.wikipedia.org/wiki/Autocorrelation>

[5] Paired Sample T-test, Statistics Solutions, accessed 29 August, 2019, <https://www.statisticssolutions.com/manova-analysis-paired-sample-t-test/>

[6] Rajamanickam R. and Nagan S. (2018) 'Assessment of Air Quality Index Data for Cities and Major Towns in Tamil Nadu, India' Journal of Civil & Environment Engineering, Vol 8, Issue 2

[7] Gulia S., et. al., (2015), Urban Air Quality Management- A review, Vol. 6, Issue 2, pp. 286-304.

[8] Edward Ming-Yang Wu, et. al., (2013), A study on the use of a Statistical Analysis Model to Monitor Air Pollution Status in an Air Quality Total Quantity Control District, Atmosphere 2013, Vol. 4

[9] Horowitz J., et al. (1979), Statistical Analysis of the maximum concentration of an air pollutant: Effects of Autocorrelation and Non- Stationarity, Vol. 3, Issue 6, pp. 811-818.