International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Visualization of Correlation for Air Pollution Time Series using TAQMN Data

Harish Kumar K.S, Doreswamy

Abstract-: ***In Taiwan country Annan, Chiayi, Giran, and Puzi cities are facing a serious fine particulate matter ($PM_{2.5}$) issue. To date the impressive advance has been made toward understanding the $PM_{2.5}$ issue, counting special temporal characterization, driving variables and well-being impacted. However, notable research as has been done on the interaction of the content between the selected cities of Taiwan country for particulate matter ($PM_{2.5}$) concentration. In this paper, we purposed a visualization technique based on this principle of the visualization, cross-correlation method and also the time-series concentration with particulate matter ($PM_{2.5}$) for different cities in Taiwan. The visualization also shows that the correlation between the different meteorological factors as well as the different air pollution pollutants for particular cities in Taiwan. This visualization approach helps to determine the concentration of the air pollution levels in different cities and also determine the Pearson correlation, r values of selected cities are Annan, Puzi, Giran, and Wugu.***

Keywords**: *Air pollution, Particulate Matter ($PM_{2.5}$), Visualization, Framework, Wind Speed, Correlation, TAQMN.***

I. INTRODUCTION

Annan, Chiayi, Giran, and Puzi are the main cities of Taiwan country and these cities are facing a serious problem with particulate matter $PM_{2.5}$ as pollution issue. $PM_{2.5}$ pollution causes critical problems to health and grows rapidly in cities within the region [1]. Considerable endeavors have been made towards understanding $PM_{2.5}$ contamination in Taiwan with the spatiotemporal characterization source of units impacting meteorological components, checking and relief arrangements and well-being impacts [2].

These endeavors have enormously enhanced the information on $PM_{2.5}$ contamination and have done a momentous work in making a difference illuminate contamination relief approaches. In any case, to date small inquire about has been done to explore the energetic connections of $PM_{2.5}$ concentrations at different cities at diverse times [3].

Manuscript published on 30 September 2019
* Correspondence Author
Harishkumar K .S*, Computer Science, Mangalore University, Mangalore, Karnataka, India. Email: kskharish88@gmail.com
Doreswamy, Computer Science, Mangalore University, Mangalore, Karnataka, India. Email: doreswamyh@yahoo.com

In 2013, the Taiwan country announced the rules for decreasing the $PM_{2.5}$ concentration in the selected area by up to 25% by 2017, corresponding to the year of the 2012 level. Some of the updates for decreasing the $PM_{2.5}$ emissions are done with the regulations of the governments [4]. These efforts are required associations of $PM_{2.5}$ concentrations in various cities at various periods in a particular area. The best understanding of the relationship between the $PM_{2.5}$ and selected cities in Taiwan country is required for an action plan to help the city environment in improvement for effective measures for air pollution [5]. The aforementioned cities in the above are considered for the best understanding of the correlation between the $PM_{2.5}$ and the various cities in Taiwan country [6].

In this paper visualization techniques are deployed to describe the hidden relationships exists among the air pollutants with meteorological parameters. The interactions of air pollution in different areas of cities Annan, Chiayi, Giran, Puzi, and Wugu the effect of air pollution from meteorological conditions are visualized. The correlations of individual meteorological factors with pollutants in air pollution are described with graphical representations [7].

The remaining part of the paper is organized as follows. Section 2 describes the selected area, Section 3 introduces the phenomenon of time series analysis and visualization techniques for air pollution data analysis, Section 4 correlation analysis and last Section 5 describes the results and summary.

II. MATERIALS AND METHODS

A. Study Area

Taiwan country is located at latitude $23^0\,69'\,78''$ N and longitude $120^0 96'5''$ E as shown in Fig 1. We observe that the Taiwan Air Quality Monitoring Network (TAQMN) center is the significant premise of air quality assurance [8]. Therefore, we have collected 76 stations data and these are belonging to 60 common stations, 5 manufacturing stations, 5 traffic stations, 4 background stations, and 2 normal stations. Among these five important cities Annan, Chiayi, Giran, puzi, and Wugu are considered for study as these cities cover the larger geographical areas with a large population [9]. These cities have basic significant physical infrastructure and large population as shown in Table I.

B. Dataset Description

The data considered for the study has been observed from 1st January 2012 to 31st December 2017 that was obtained by Taiwan air quality monitoring network (TAQMN).

It has hourly data set and has been converted into a daily data set by averaging hourly data per day. We concentrate 6 major pollutants such as Nitrogen dioxide (NO_2), Ozone (O_3), Carbon monoxide (CO), Sulfur dioxide (SO_2), $PM_{2.5}$ < 2.5 micrometers in diameter, PM_{10} < 10 micrometers in diameter, and other meteorological factors such as temperature, rainfall, wind direction, wind speed, wind direction horizontally, relative humidity, and wind speed horizontally [10].

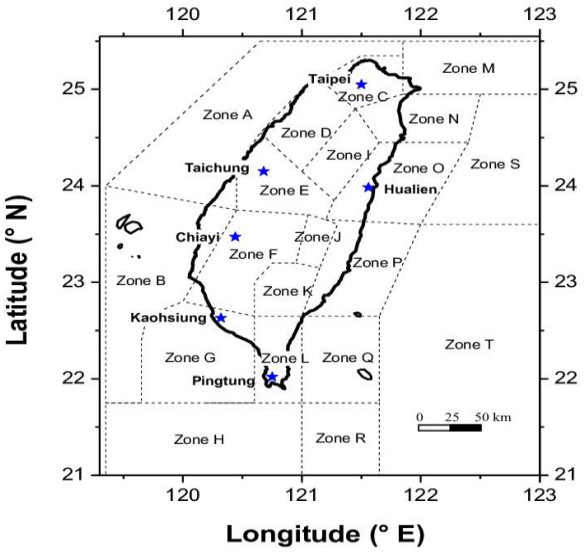

Fig. 1: Location of the study region.

TABLE I: Basic information about the 5 studied cities

City	Population (As of July 2018)*	Area of the City (Km^2)	Latitude and Longitude
Annan	1,92,887	107	23^0 47' 13" N 120^0 18' 13" E
Giran	95,905	29.87	24^0 45' 45" N 121^0 45'45" E
Wugu	82,255	34.86	25^0 4' 24" N 121^0 25'42" E
Puzi	42,176	50	23^0 46' 11" N 120^0 24' 2" E
Chiayi	2,70,254	60.03	23^0 20' 0.24" N 120^0 24' 0" E

For the quality check, the air quality monitoring data were processed to facilitate the city-based cross-correlation analysis [11]. First, for each city the concentration of the daily pollutant was calculated by averaging the hourly data from all stations within that city. Then finally, the average was calculated to better capture the general trend of the time series of averaging the daily data from all station [12].

III. VISUALIZATION FRAMEWORK

This method visualization framework to visualize continue interactions of particulate matter $PM_{2.5}$ as well as all other pollutants of air pollution and meteorological factors time series in different cities by the reference of TAQMN data. The important things in the visualization framework are the cross-correlation method, it calculates the correlations between the particulate matter $PM_{2.5}$, air pollutants and meteorological factors time series in different cities [13].

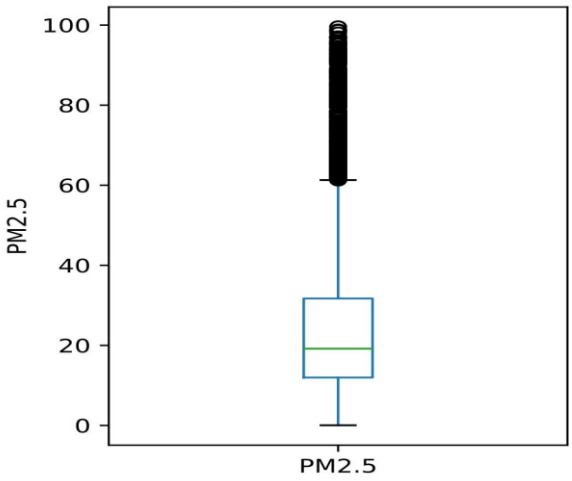

Fig. 2: Box plot of $PM_{2.5}$ data

IV. CORRELATION ANALYSIS

This section correlation analysis for the air pollution for selected cities in the Taiwan country air pollutants such as particulate matter ($PM_{2.5}$), nitrogen dioxide (NO_2), ozone (O_3), carbon monoxide (CO), sulfur dioxide (SO_2), particulate matter ($PM_{2.5}$) size as less than 2.5 micrometers in diameter, particulate matter (PM_{10}) size as less than 10 micrometers in diameter, and other meteorological factors such as temperature, rainfall, wind direction, wind speed, wind direction horizontally, relative humidity, and wind speed horizontally [14].

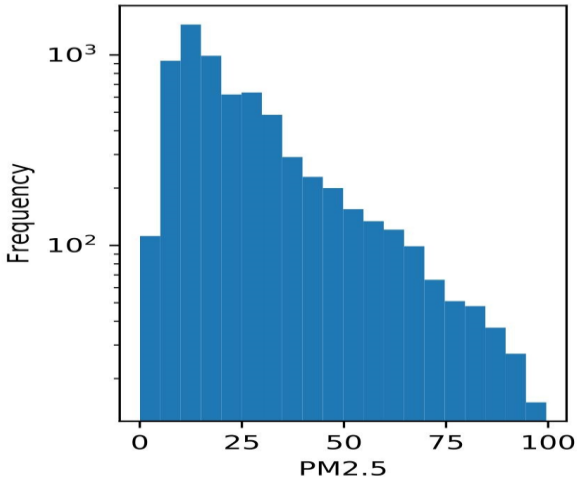

Fig. 3: Histogram of $PM_{2.5}$ data

In Fig. 2 shows the box plot of the particulate matter PM2.5 for all combined cities and Fig. 3 shows the histogram of the particulate matter $PM_{2.5}$ with related to the frequency and Fig. 4 shows the hourly particulate matter $PM_{2.5}$ with applying the probability density function that will be the determined range for 0.000 to 0.040 as a positive correlation and Fig. 5

shows the daily average particulate matter $PM_{2.5}$ from 1st January 2012 to 31st December 2017 for all selected cities combined in the Taiwan country. And Fig. 6 shows the histogram for integrated wind speed for all cities related to the frequency.

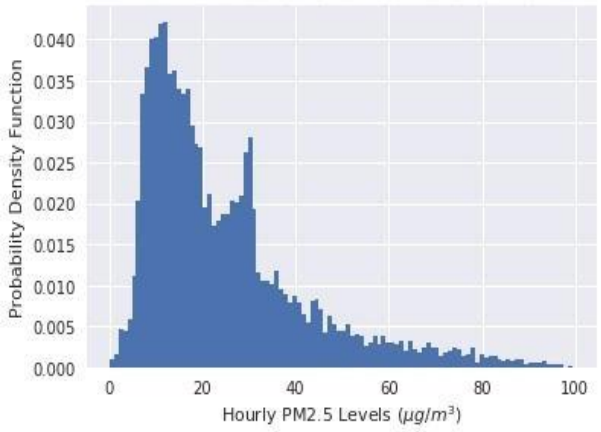

Fig. 4: Histogram for $PM_{2.5}$ in Pm clean

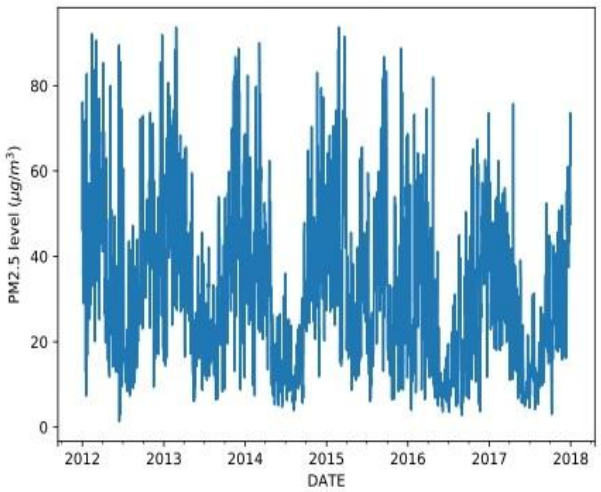

Fig. 5: Daily $PM_{2.5}$ level Chiayi station

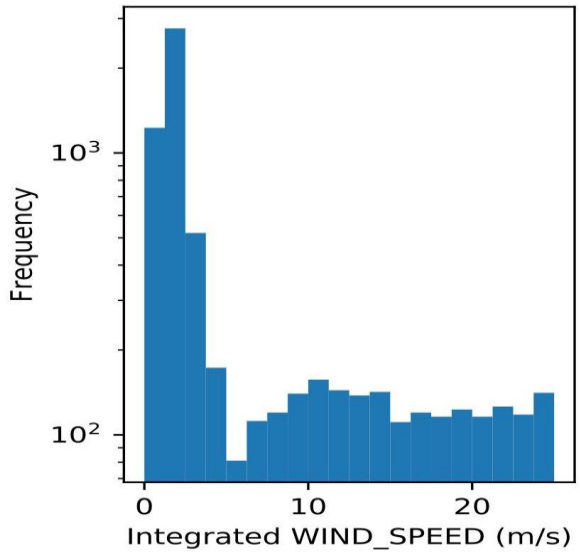

Fig. 6: Histogram of wind speed

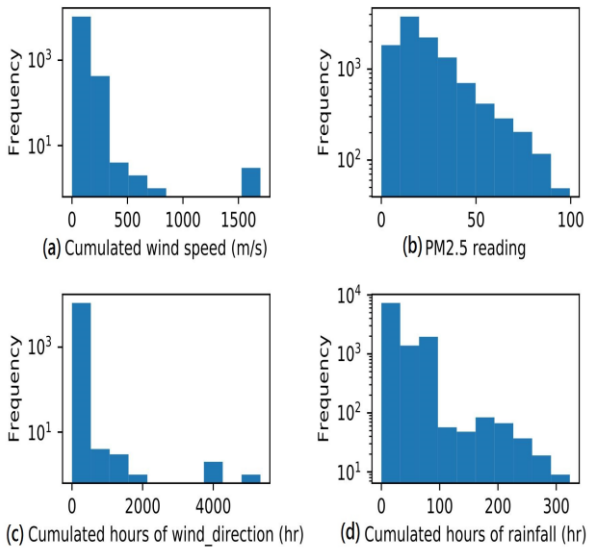

Fig. 7: Distribution of wind speed, $PM_{2.5}$ reading, wind direction, and rainfall

In Fig. 7 shows the distribution of the wind speed, particulate matter $PM_{2.5}$ reading, wind direction and rainfall for the selected cities in Taiwan country, the correlation of the visualization framework for (a) cumulated wind speed in (m/s) with frequency this as rang between the 0 to 1000 has been determined and (b) shows the particulate matter $PM_{2.5}$ reading with the same frequency of visualization, in (c) this also the cumulated hours of the wind direction with at same frequency for range 0 to 1000, and last (d) also shows the cumulated hours of rainfall with also the same frequency.

In Fig. 8(a) shows the visualization framework for the correlation matrix for Annan city in this matrix all show the air pollutants and metrological factors in this city as wind speed and wind speed horizontal as positive correlation and temperature, carbon monoxide components are the negative correlation. Fig. 8(b) Giran city temperature, carbon monoxide, nitrogen dioxide (NO_2) and wind direction are the negative correlation and other pollutants and meteorological factors are a positive correlation.

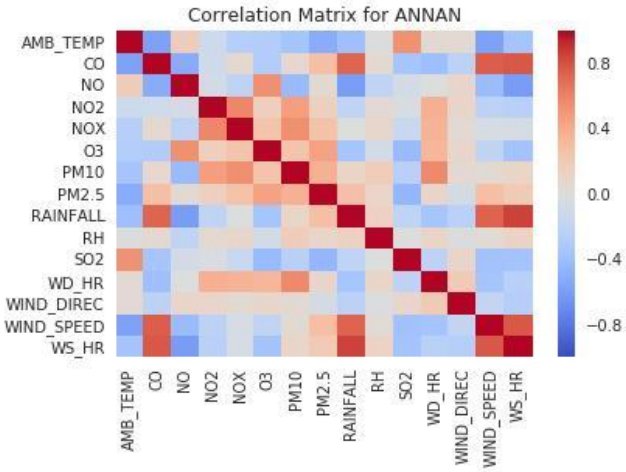

Fig 8 (a): Correlation matrix for annan station

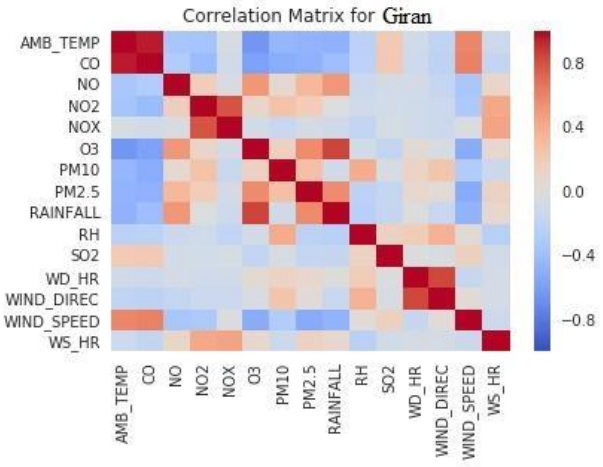

Fig 8 (b): Correlation matrix for giran station

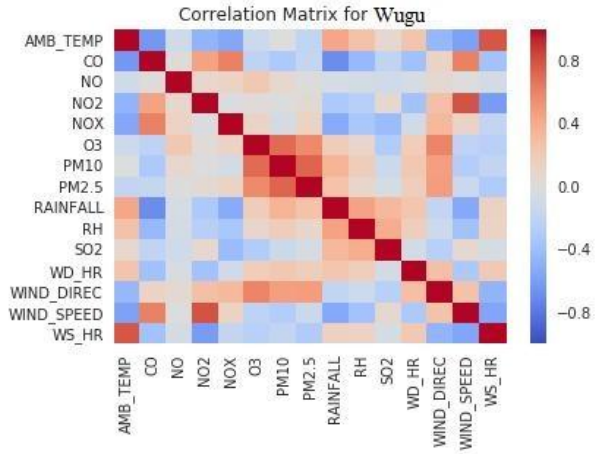

Fig 8 (c): Correlation matrix for wugu station

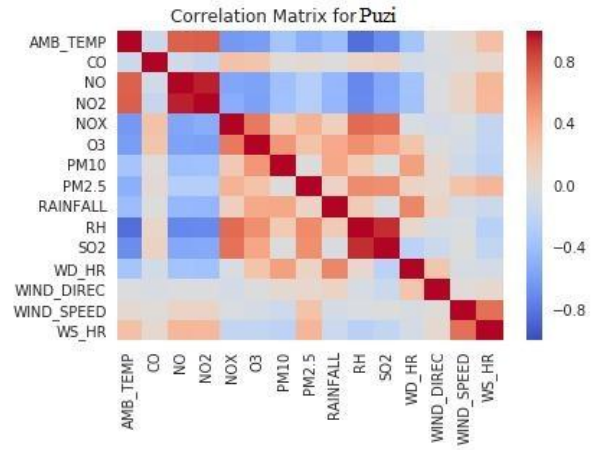

Fig 8 (d): Correlation matrix for puzi station

Fig. 8(c) Wugu city correlation matrix ozone (O_3) and particulate matter ($PM_{2.5}$ & PM_{10}) are the positive correlation and other pollutants and meteorological factors are a negative correlation. Fig. 8(d) last city as Puzi the correlation matrix nitrogen dioxide (NO_2), sulfur dioxide (SO_2) and wind speed are the positive correlation and other pollutants and meteorological factors are the negative correlation.

V. RESULT OBTAINED

$$\mu = \sum_{i=1}^{n} \frac{x_i}{n} \quad (1)$$

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \quad (2)$$

$$\rho = \frac{COV\,(X,Y)}{\sigma_x \sigma_y} \quad (3)$$

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 (y_i - \bar{y})^2}} \quad (4)$$

The mathematical definition of mean (μ), standard deviation (σ), and Pearson correlation coefficient (r) measure represented by the Equation 1, 2, 3, and 4 respectively, here in mean n is the sample size and x_i is the observed values, and in standard deviation where $x_1, \ldots, x_n$ are the observed values of the sample items, x is the mean value of these observations, and N is the number of observations in the sample and The Pearson correlation coefficient is probably the most widely used measure for linear relationships between two normally distributed variables and thus often just called "correlation coefficient" [15].

Usually, the Pearson coefficient is obtained via Least-Squares fit and a value of 1 represents a perfect positive relationship, -1 a perfect negative relationship, and 0 indicates the absence of a relationship between variables.
In Fig. 9(a, b, c, and d) are shows the visualization framework and a cross-correlation method it determines the particulate matter $PM_{2.5}$ and wind speed for selected cities Annan, Puzi, Giran, and Wugu are the Pearson r values are as shown in Table II. First two cities are a positive correlation and the other two cities are negative correlation it can be determined by using the visualization framework and a cross-correlation method [16]. Descriptive statistics for values determining (outputs) the level of pollutant emission in selected Taiwan country as shown in Table III.

TABLE II: Pearson r values for different cities

Cites	Pearson r Value	Correlation
Annan	0.29	Positive
Puzi	0.26	Positive
Giran	-0.53	Negative
Wugu	-0.13	Negative

TABLE III: Descriptive statistics for values determining the level of pollutant values

STAT	TEMP	CO	NO2	O3	PM10	PM 2.5	RF	RH	SO2	WDHR	WD	WS	WS HR
COUNT	10960	10960	10960	10960	10960	10960	10960	10960	10960	10960	10960	10960	10960
MEAN	67.67	36.47	45.17	21.62	20.89	24.84	30	8.47	21.93	73.78	41.50	33.94	9.45
STD	69.08	68.80	77.40	77.36	92.52	17.57	39.64	60.50	51.84	78.11	92.19	58.51	0.68
MIN	0	0	0	0	0.01	0.04	0	0.02	0	0	0	0	0
25%	2.13	0.64	1.51	1.19	1.95	11.95	2.06	3.69	2.17	22.41	2.27	1.93	2.54
50%	69.28	2.28	7.03	2.95	11.45	19.50	14.04	10.98	11.53	39.52	24.16	2.83	18.08
75%	84.87	24.83	26.81	29.41	27.37	31.33	49.55	34.59	24.30	88.41	61.80	45.95	37.12
MAX	319.50	311.29	327.54	4200.20	8749.12	99.54	323.29	2863.91	3778.25	2522.29	5334.87	1698.04	861.05

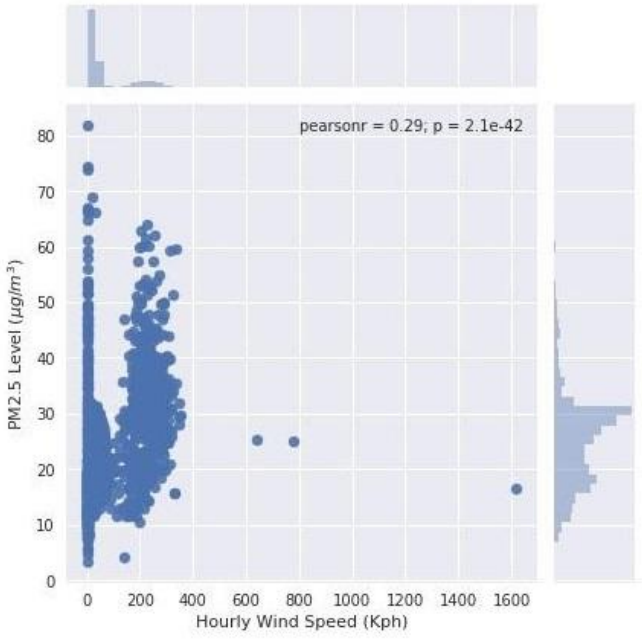

Fig 9(a): $PM_{2.5}$ versus Annan station

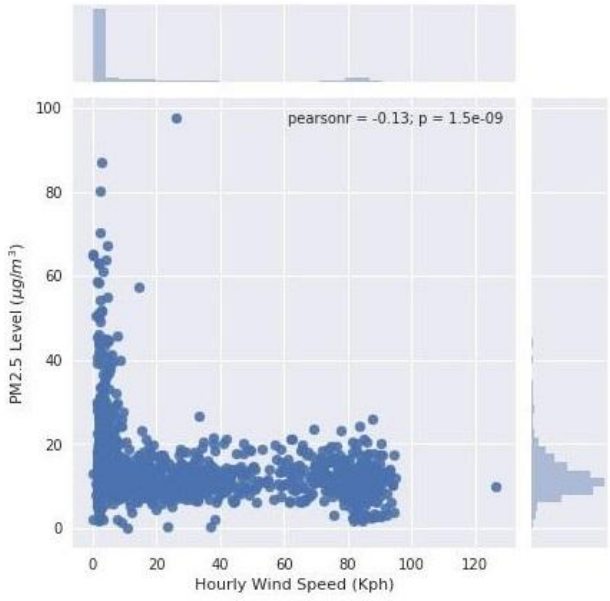

Fig 9(b): $PM_{2.5}$ versus Wugu station

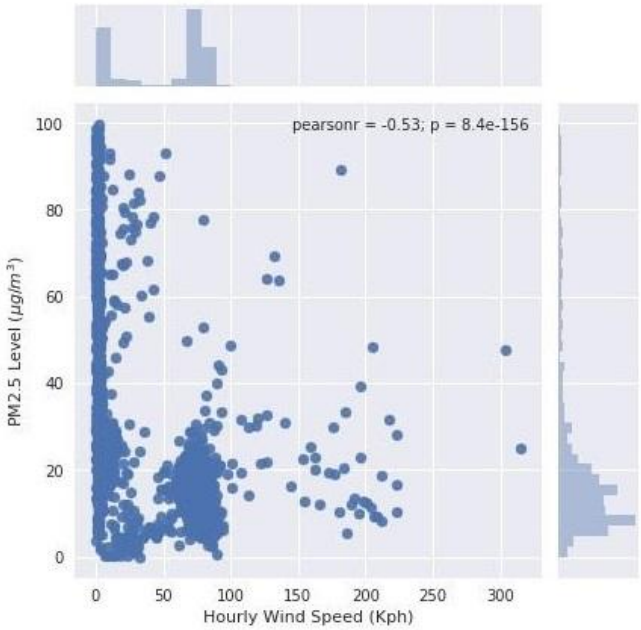

Fig 9(c): $PM_{2.5}$ versus Gairn station

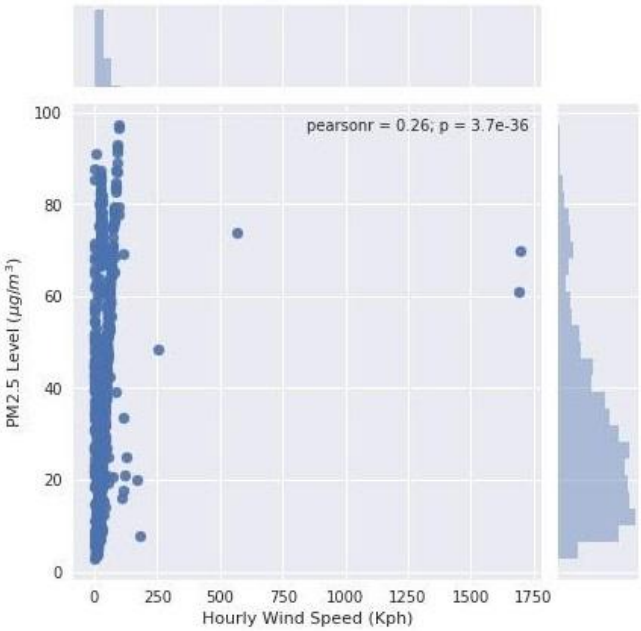

Fig 9(d): $PM_{2.5}$ versus Puzi station

VI. CONCLUSION

This research investigated the phenomenon of time series of correlation analysis of the particulate matter $PM_{2.5}$ with selected cities by using the visualization framework with the cross-correlation analysis method we analyzed the Annan, Giran, Wugu, and Puzi cities are positive and a negative correlation of the air pollution pollutants and meteorological factors. Some cities are shows the maximum positive correlation and less a negative correlation. Visualization framework and a cross-correlation method, it determines the particulate matter $PM_{2.5}$ and wind speed for selected cities Annan, Puzi, Giran, and Wugu first two cities are a positive correlation and the other two cities are a negative correlation, and also calculated the Descriptive statistics for values determining the level of pollutant values.

REFERENCES

1. Huang, Fangfang, et al. "PM2.5 spatiotemporal variations and the relationship with meteorological factors during 2013-2014 in Beijing, China." PloS one 10.11 (2015): e0141642.
2. Yang, Yong, and George Christakos. "Spatiotemporal characterization of ambient PM2.5 concentrations in Shandong province (China)." Environmental science &technology 49.22 (2015): 13431-13438.
3. Chen, Ziyue, et al. "Understanding temporal patterns and characteristics of air quality in Beijing: A local and regional perspective." Atmospheric Environment 127 (2016): 303-315.
4. Chen, Wei, Hongzhao Tang, and Haimeng Zhao. "Diurnal, weekly and monthly spatial variations of air pollutants and air quality of Beijing." Atmospheric Environment 119 (2015): 21-34.
5. Hu, Jianlin, et al. "Spatial and temporal variability of PM2. 5 and PM10 over the North China Plain and the Yangtze River Delta, China." Atmospheric Environment 95 (2014): 598-609.
6. He, Jianjun, et al. "Air pollution characteristics and their relation to meteorological conditions during 2014–2015 in major Chinese cities." Environmental pollution 223 (2017): 484-496.
7. Luo, Jieqiong, et al. "Spatiotemporal pattern of PM 2.5 concentrations in mainland China and analysis of its influencing factors using geographically weighted regression." Scientific reports 7 (2017): 40607.
8. Zhang, Renjian, et al. "Chemical characterization and source apportionment of PM 2.5 in Beijing: seasonal perspective." Atmospheric Chemistry and Physics 13.14 (2013): 7053-7074.
9. Lv, Baolei, et al. "Characterizations of PM2. 5 pollution pathways and sources analysis in four large cities in China." Aerosol Air Qual. Res 15.5 (2015): 1836-1843.
10. Chen, Ziyue, et al. "Detecting the causality influence of individual meteorological factors on local PM 2.5 concentration in the Jing-Jin-Ji region." Scientific reports 7 (2017): 40735.
11. Huang, Duruo, Wenqi Du, and Hong Zhu. "A procedure to select ground-motion time histories for deterministic seismic hazard analysis from the Next Generation Attenuation (NGA) database." (2017).
12. Pearce, John L., et al. "Quantifying the influence of local meteorology on air quality using generalized additive models." Atmospheric Environment 45.6 (2011): 1328-1336.
13. Liu, Yong, et al. "Research on the relationship between urban form and urban smog in China." Environment and Planning B: Urban Analytics and City Science 44.2 (2017): 328-342.
14. Wu, Dan, Yuan Xu, and Shiqiu Zhang. "Will joint regional air pollution control be more cost-effective? An empirical study of China's Beijing–Tianjin–Hebei region." Journal of environmental management 149 (2015): 27-36.
15. SUNG, WANG. "China Council for International Cooperation on Environment and Development (CCICED)." 生物多样性 4.Suppl. (1996): 76-88.
16. Lu, Wei-Zhen, Hong-Di He, and Li-yun Dong. "Performance assessment of air quality monitoring networks using principal component analysis and cluster analysis." Building and Environment 46.3 (2011): 577-583.

AUTHORS PROFILE

Harish Kumar K. S is currently a Research Scholar Department of Computer Science. He received B.Sc. from Gulbarga University in 2010 and M.Sc. degree in Computer Science from Vijayanagara Shri Krishnadevaraya University Ballari in 2014. After completion of his post-graduation degree. He joined as a lecturer in Vijayanagara Shri Krishnadevaraya University Ballari in the year 2015. After he joined as Research in Computer Science at Mangalore University in the year 2017 still him pursuing Ph.D. on Advanced predictive models for air pollution analysis in the smart city system. He has published more than 3 contribute research papers at National / International Journals and Conferences. His areas of research interest include Data mining, Artificial Intelligence, Machine learning, Deep learning, Big data analytics, and Air pollution analysis.

Dr. Doreswamy is currently a Professor of Computer Science in the Department of Computer Science. He received B.Sc. and M.Sc. degree in Computer Science from the University of Mysore in 1993 and 1995 respectively. After completion of his Post Graduation Degree. He joined as Associate Professor in Computer Science at Mangalore University in the year 2003. He was the Chairman of the Department of Post-Graduate Studies and Research in Computer Science during 2003-2005 and 2008-2012 and served at various capacities in Mangalore University, as a Chairman and member of DOC, DOS and UG/PG BOS and BOE in Computer Science. Started Ph.D. program in Computer Science and Technology at Mangalore University with effect from the academic year 2003-04 onwards. His areas of research interests include Data Mining and Knowledge Discovery, Artificial Intelligence, Machine learning, and Scalable Advanced Data Mining Algorithms. He has published more than 60 contributed peer-reviewed research papers at National/International Journals and Conferences. He has chaired many National and International Conferences in India. He has completed one minor research project that was sanctioned by the University Grants Commission (UGC). He completed the Major Research Project entitled "Scientific Knowledge Discovery Systems (SKDS) for advanced Engineering Materials Design Applications "from the funding Agency University Grant Commission, New Delhi, INDIA. He served as a subject expert member in many committees formed by various Universities in and outside states, Karnataka State Higher education Council. Bangalore, UGC, and KPSC. Technical advisory member too many national and international conference, editorial member too many International Journals in Computer Science and technology, Advisory members to many institutions. He has received "SHIKSHA RATTAN PURSKAR", Indian International Friendship Society, New Delhi, in the year 2009. He is guiding currently eight doctoral students in the areas of Data Mining and Knowledge Discovery; Machine learning and scalable Advanced-Data Ming Algorithms.