International Cartographic Association
Association Cartographique Internationale

Extracting relationship between air pollution and precipitation using spatio-temporal analysis in Tehran metropolis

Shokouh Dareshiri [a, *], Mohammadreza Sahelgozin [a], Maryam Lotfian [b], Jens Ingensand [b]

[a] *Faculty of Geodesy and Geomatics Engineering, K.N. Toosi University of Technology, Tehran, Iran, sh.dareshiri@gmail.com, m.sahelgozin@gmail.com*
[b] *University of Applied Sciences and Arts Western Switzerland, School of Business and Engineering, maryam.lotfian@heig-vd.ch, jens.ingensand@heig-vd.ch*

* Corresponding author

Abstract: Precipitation is one of the main stages of the water cycle, and it is required for the organisms to survive on the planet. In contrast, air pollution is a phenomenon that has greatly affected the human life nowadays. Population growth, development of factories and increasing number of fossil fuel vehicles are the most influencing factors on air pollution. In addition to understand nature of precipitation and air pollution, finding relationship between these two phenomena is necessary to make appropriate policies for reducing air pollution. Furthermore, studying trends of precipitation and air pollution in the past, is helpful to forecast the times and places with less precipitation and more air pollution for a better urban management. In this study, we tried to extract any probable relationship between these two parameters by investigating their monthly measured amounts in 22 municipal districts of Tehran in three epochs of time (2009, 2013 and 2017). Carbon Monoxide (CO) was considered as the indicator of air pollution. Results of the study show that the parameters have a significant relationship with each other. By using Pearson Correlation Coefficient and One-Way Variance (ANOVA) test, relationship between the data for each month and for each district of Tehran were studied separately. As the time has passed and the air pollution has increased, the correlation between the parameters in districts has decreased. In addition, during the cold months of the year, the correlations decrease since the fact that precipitation is not the only influencing factor on the air pollution due to the rise of air "Inversion". Finally, the polynomial regression model of carbon monoxide based on precipitation was extracted for each of the three years. The model suggests a degree three polynomial equation. The obtained coefficients from the regression model show that the relationship between parameters was stronger in the years with more rainfalls. This can be due to the more significant impact of other influencing factors on air pollution, such as population density, wind direction, vehicles and factories in the areas or conditions with a less rainfall.

Keywords: Precipitation, Air pollution, Pearson correlation coefficient, Polynomial regression, Modelling

1. Introduction

Continuation of life strongly depends on the water cycle. Different ecosystems on the Earth and evolution of organisms have been highly affected by precipitation during the various geological periods and human being is no exception (Valdes-Pineda et al. 2016). As a definition, every type of steam transformation to the liquid water under the high pressure which results in numerous falling water drops to the Earth surface, can be called 'Precipitation'. It might be in the shape of rain, snow, hail, dew or snowflake. Precipitation occurs whenever a part of the sky becomes much saturated by the steam, so that the steam becomes condensed and heavy enough to drop (Hasanalizadeh et al. 2015). Nowadays, precipitation is playing an important role in changing quality of life in metropolises.

As the Earth gets warmer and existence of fresh water resources is eliminated, significance of such a role increases. Large cities usually deal with air pollution and huge amount of released greenhouse gases in the air as a consequence of their dense population, large number of fossil fuel vehicles and spreading operation of factories. This is especially more critical in cold months of the year which leads to a phenomenon called 'Inversion' in meteorology. One instance of the populated and air polluted large cities around the world is Tehran, the capital of Iran. Its population has extremely raised in the recent years due to the permanent, temporary and daily migrations that are carried out with the goal of finding a better job, continuing education, medical treatments and so on. Despite of many applied policies to confront air pollution in Tehran, evidences show an increase in the air pollution while the amount of rainfall has decreased during the last few years.

Apart from understanding nature of the both precipitation and air pollution, extracting a probable relationship between these two parameters would be a beneficial key in urban management and representing solutions to reduce air pollution and to increase rainfall. In addition, research on the precipitation and air pollution trends in the past years, would be helpful for finding predictive models to forecast the times and the places with less precipitation and more

Proceedings of the International Cartographic Association, 2, 2019.
29th International Cartographic Conference (ICC 2019), 15–20 July 2019, Tokyo, Japan. This contribution underwent single-blind peer review based on submitted abstracts. https://doi.org/10.5194/ica-proc-2-23-2019

air pollution which are more critical for policy makers to concern.

In this regard, objective of the current research is to model spatio-temporal distribution of precipitation and air pollution in 22 municipal districts of Tehran in order to find a meaningful relation or correlation between precipitation and air pollution. In order to accomplish this, monthly measured amount of precipitation and daily measured amount of Carbon Monoxide (CO) in the air were considered from 8 to 10 stations in three epochs of time (2009, 2013 and 2017). Carbon Monoxide (CO), known as the invisible killer, is a colourless, odourless, tasteless, poisonous and potentially dangerous gas which is mostly produced as a result of incomplete combustion in car engines or factory equipment (Ernst and Zibrak 1998).

In the rest of this paper, related studies about precipitation and air pollution are described in section 2. Next, in section 3, the study area along with the applied data as well as their preparation process for the performed analysis were introduced. The statistical and spatial analysis used in this research to extract the relationship and correlation between precipitation and air pollution with a short discussion on the obtained results are provided in section 4, while section 5 represents conclusion of the study.

2. Related Works

Many studies including (Hasanalizadeh et al. 2015, Jahanbakhsh ASl et al. 2015, Aghamohammadi et al. 2017) have investigated precipitation and their spatio-temporal changes in different regions. Analysis and spatio-temporal models that are applied in these studies are useful tools for a more appropriate management of water resources and to find tangible and intangible impresses they might have in different regions. In addition, some studies such as (M. et al. , Carrera-Hernández and Gaskin 2007, S. et al. 2011) worked on cluster analysis or spatial interpolation to assess quality of the water resources and to find if they are suitable for drinking or farming purposes. On the other hand, a number of studies (e.g. (Khorshiddoust and Shirzad 2014)) have used multivariate statistical methods in order to identify the spatial-temporal differences of rainfall and to extract similar areas for grouping and homogenizing precipitation stations.

In contrast, there are studies that have modelled or predicted the contamination of large cities, especially metropolises, by using ground-based methods. Studies including (Brunekreef and Holgate 2002, Meshkini et al. 2017, Hadei et al. 2018) have tried to extract the relationship between air pollution and some diseases such as lung cancer and heart attack. Another category of studies (Kan et al. 2008, ATRKAR and FATHI 2017, Datta 2017), has examined the relationship between education or population and the growth of air pollution along with the seasonal changes of air pollution.

Tehran, the capital of Iran, possesses its own special economic and social situations and a large and dense population gathered together in this metropolis causing an increasing contamination in the air. Tehran's air pollution depends on various parameters such as geographic and meteorological indicators. Najafpoor et al. (2015) have studied the trend of changes in some air pollution quality indices of Tehran and their relations with meteorological data. According to their results, density of O_3 in Tehran's air has been increased in the recent years and has become more than the standard amount.

Bahari et al. (2016) classified areas of Tehran based on some different aspects including land uses, population, elevation and density of "PM2.5" (one of the contaminant indices) using meteorological parameters (such as temperature, wind speed and wind direction). They also suggested a model which is able to estimate density of PM2.5 in different locations of Tehran.

3. Data and Study Area

In this study, monthly measured amounts of rainfall in 8 precipitation stations with determined latitudes and longitudes were collected in three predefined years (2009, 2013 and 2017 with a difference of 4 years). Moreover, data of the daily measured amounts of Carbon Monoxide (CO) in the air, which is chosen to be investigated as the air pollution indicator, from 10 air pollution monitoring stations of Tehran Air Quality Control Company were collected. In the first step, monthly averages for amounts of CO were calculated for each month of the three predefined years.

Before performing any statistical analysis, distribution of the collected data was examined. Normality of data was assessed by computing their skewness and kurtosis and also by the help of Shapiro-Wilk test in Minitab17. The Shapiro-Wilk indicator is calculated by equation 1.

$$w = \frac{(\sum_{i=1}^{n} a_i X_{(i)})^2}{\sum_{i=1}^{n} (X_i - \bar{X})^2} \qquad \text{eq. (1)}$$

The Sig. value of Shapiro-Wilk test for the data was below 0.05 showing that the data were not normal. Thus, they were then normalized by the use of Logarithmic Transformation, Cox-Box and Johnson Transformation in Minitab17. After the normalization process, the data were imported as a number of tables in Microsoft Excel 2016 (a table for each year) consisting of geographical latitudes and longitudes of precipitation and air pollution stations along with monthly measures for both precipitation and air pollution of each station.

In the next step, precipitation stations as well as air pollution stations are inserted in ArcMap 10.3 software as two point layers for each month with the normalized rainfall and CO measures as their attribute tables. Then, data were interpolated by using Inverse Distance Weighting (IDW) analyses for the whole study area (22 municipal districts of Tehran) in all 12 months of the three predefined years (2009, 2013 and 2017). The result was a number of raster layers whose pixels contain values for either precipitation or CO measures.

Proceedings of the International Cartographic Association, 2, 2019.
29th International Cartographic Conference (ICC 2019), 15–20 July 2019, Tokyo, Japan. This contribution underwent
single-blind peer review based on submitted abstracts. https://doi.org/10.5194/ica-proc-2-23-2019 |

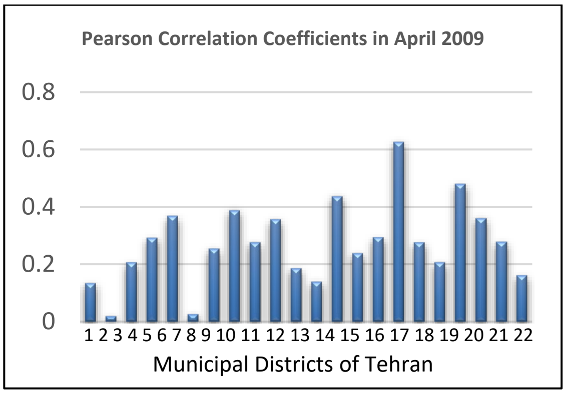
4. Relationship between Precipitation and Air Pollution

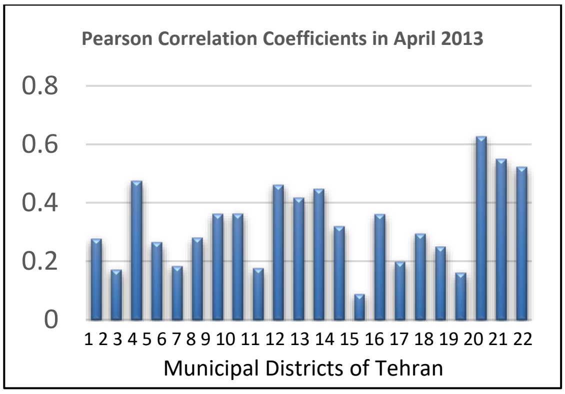
In this study, 22 municipal districts of Tehran were considered in order to have a more spatial look on the data. In this regard, means of the both precipitation and CO measures for each municipal district were calculated by utilizing Zonal Statistical as Tables analysis in ArcMap 10.3. For instance, figure (4-1) illustrates the parameters which are derived from Zonal Statistical analysis for the month April of 2009 in all 22 municipal districts of Tehran.

OID	FID_	COUNT	AREA	MIN	MAX	RANGE	MEAN	STD	SUM
0	0	4208	0.004931	0.866478	0.966672	0.100194	0.92075	0.017181	3874.514796
1	1	4265	0.004998	0.785176	1.053566	0.26839	0.861689	0.060561	3675.102354
2	2	2494	0.002923	0.868772	0.945184	0.076412	0.916802	0.019238	2286.504475
3	3	5178	0.006068	0.764171	1.219722	0.455551	0.971151	0.051553	5028.622138
4	4	4300	0.005039	0.340052	0.972474	0.632423	0.680678	0.136131	2926.91534
5	5	1822	0.002135	0.888844	1.056693	0.16785	0.966005	0.038177	1760.060368
6	6	1315	0.001541	0.907117	0.989959	0.082842	0.943633	0.013153	1240.876919
7	7	1129	0.001323	0.69012	0.952323	0.262204	0.872541	0.060508	985.098547
8	8	1668	0.001955	0.751764	1.193919	0.442155	1.001326	0.096142	1670.211248
9	9	683	0.0008	1.025869	1.170927	0.145058	1.108926	0.033047	757.396238
10	10	1010	0.001184	0.971499	1.159846	0.188347	1.100147	0.036993	1111.148525
11	11	1351	0.001583	0.926307	1.086315	0.160008	0.997327	0.034887	1347.388118
12	12	1037	0.001215	0.808694	0.94691	0.138217	0.890309	0.02901	923.25075
13	13	1216	0.001425	0.695891	0.948436	0.252545	0.865084	0.051259	1051.94225
14	14	2379	0.002788	0.470064	0.958748	0.488684	0.72219	0.147278	1718.090281
15	15	1382	0.001619	0.790215	1.061075	0.27086	0.925898	0.04894	1279.590629
16	16	728	0.000853	0.990766	1.701574	0.710809	1.292664	0.162083	941.059649
17	17	1390	0.001629	0.903791	1.654115	0.750324	1.207558	0.169619	1678.506008
18	18	932	0.001092	0.820693	1.719581	0.898889	1.281256	0.264043	1194.130232
19	19	1818	0.00213	0.791635	1.010341	0.218707	0.93826	0.034111	1705.756983
20	20	4278	0.005013	0.513488	0.885207	0.371719	0.707246	0.067504	3025.599357
21	21	6311	0.007395	0.340262	0.806205	0.465943	0.615316	0.131362	3883.256905

Figure (4-1): Derived parameters from Zonal Statistical as Table analysis for the month April of 2009

For making a better comparison, extracted means of the Zonal Statistical as Table analysis were normalized between 0 and 1. Next, the Pearson Correlation Coefficient was computed on the data derived from the analysis in IBM SPSS Statistics 22. In figure (4-2), a bar chart of the calculated Pearson Correlation Coefficients is represented in three epoch of times (April 2009, 2013 and 2017). In addition, table (4-1) contains the calculated mean Pearson Correlation Coefficients for precipitation and air pollution in each year. Average precipitation was the highest in the year 2013 when the Pearson Correlation Coefficient was also the highest. Moreover, both were the least in the year 2017. Therefore, a relationship can be figured out between the amount of precipitation and the correlation between precipitation and air pollution.





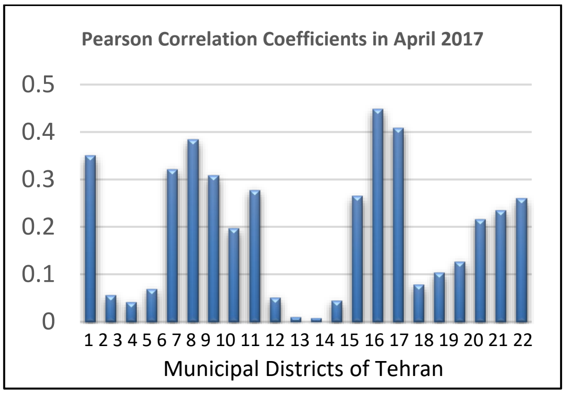


Figure (4-2): Pearson Correlation Coefficient for precipitaion and air pollution in April 2009, 2013 and 2017 for 22 districts

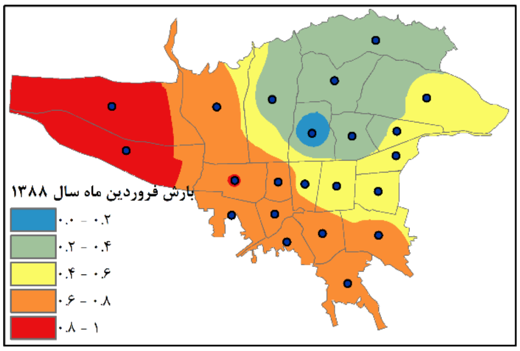
Table (4-1): Calculated mean Pearson Correlation Coefficients for precipitation and air pollution

Year	Mean correlation based on districts	Mean correlation based on months
2009	0.272837	0.21706
2013	0.329205	0.259588
2017	0.19347	0.20572

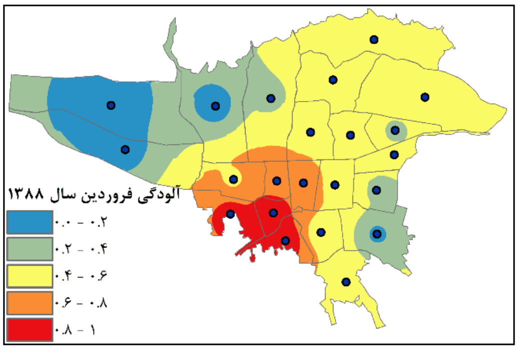
In addition, with regard to the results, the relationship between precipitation and air pollution has decreased when precipitation was less in districts. Also, air pollution and the relationship between precipitation and air pollution were varied in different districts. However, the relationship between precipitation and air pollution has been reduced when air pollution has increased or the precipitation has decreased in the recent years. Moreover, by applying the One-Way Analysis of Variance (ANOVA) test on the data in SPSS, it can be illustrated that a strong relationship exists between the amount of precipitation and air pollution. These parameters were more related while being considered in each municipal district separately rather than being considered in the whole area of Tehran which means correlation of precipitation and air pollution is more meaningful spatially than temporal.

Furthermore, the relationship between data was visualized to show trends of the data more significantly. Figures (4-3) and (4-4) are examples of the data classifications in

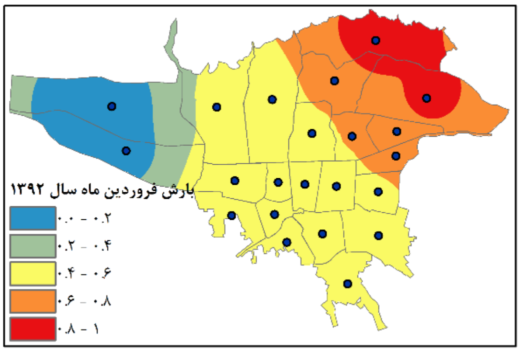
February and April of all three years. It can be seen that the areas that had more precipitation, are generally less air polluted. In addition, the data correlation was less in the cold months of the year; December, January and February were the months with the least precipitation effects on air pollution. The main reason for such a fact is a phenomenon which is called "Inversion" of the air that is known as the most important barrier of air circulation in winters causing the air pollutants to remain motionless in the air. Figures (4-3) and (4-4) show the months with the highest and the lowest correlations between precipitation and air pollution (in 2009, 2013 and 2017). It can be seen that the highest correlation was in April as a vernal month while the lowest was in February as one of the coldest months of the year.



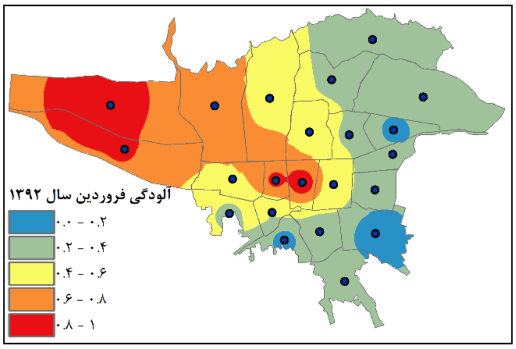
(a). Precipitation in April 2009



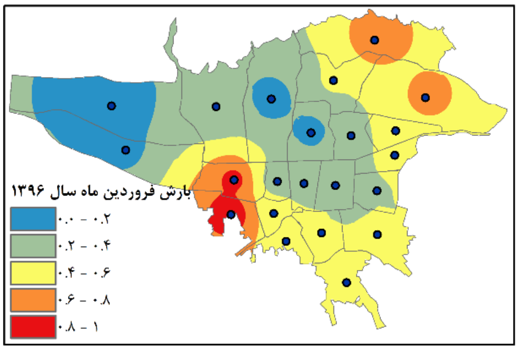
(b). Air pollution in April 2009



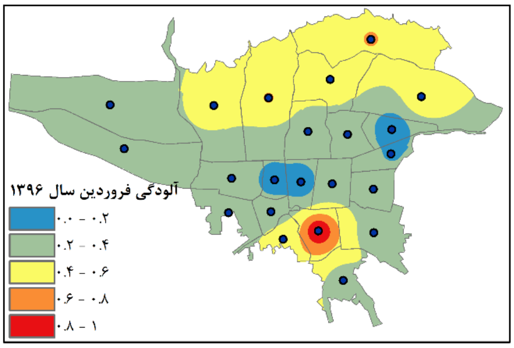
(c). Precipitation in April 2013



(d). Air pollution in April 2013

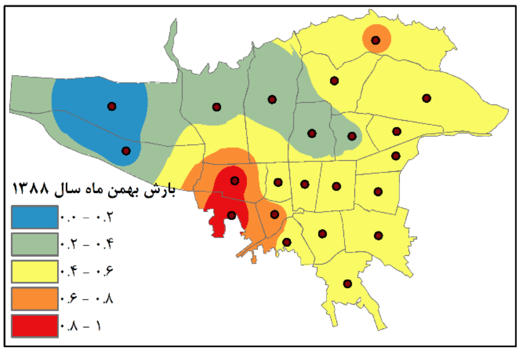


(e). Precipitation in April 2017



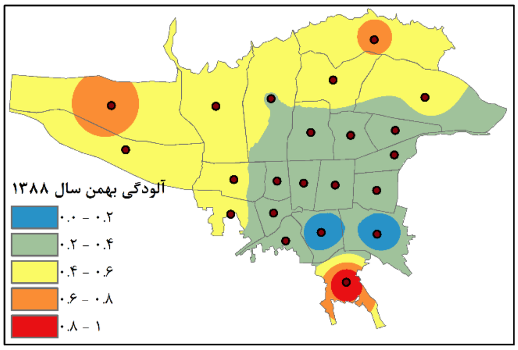
(f). Air pollution in April 2017

Figure (4-3): Classification of precipitation and air pollution data in April 2009, 2013 and 2017

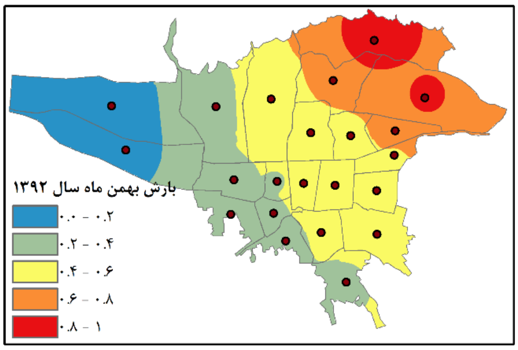


(a). Precipitation in Feburary 2009

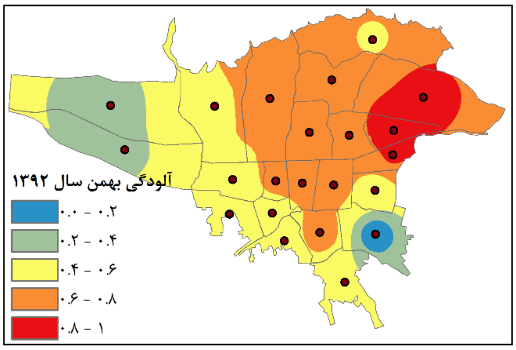
29th International Cartographic Conference (ICC 2019), 15–20 July 2019, Tokyo, Japan. This contribution underwent single-blind peer review based on submitted abstracts. https://doi.org/10.5194/ica-proc-2-23-2019 |



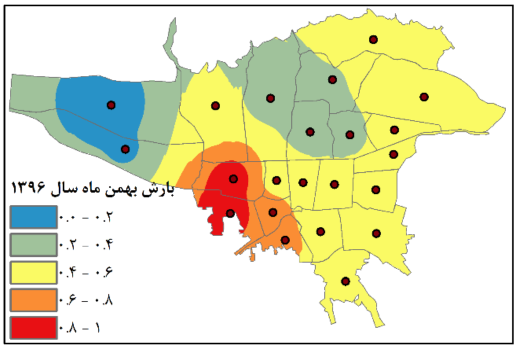
(b). Air pollution in February 2009



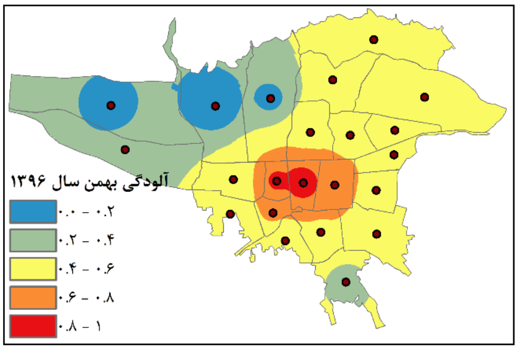
(c). Precipitation in Feburary 2013



(d). Air pollution in Feburary 2013



(e). Precipitation in Feburary 2017



(f). Air pollution in Feburary 2017

Figure (4-4): Classification of precipitation and air pollution data in Feburary 2009, 2013 and 2017

In the next step, polynomial regression models for amount of CO based on precipitation for each of the three years were extracted which show the fact that a degree 3 polynomial equation can be defined to formulize CO measures based on precipitation. The estimated coefficients of the polynomial for the year 2013 with the most rainfall is shown in Table (4-2). Coefficients for the year 2017 in this table were calculated based on coefficients of the equations for 2009 and 2013.

Table (4-2): Estimated coefficient of the polynomial regresion model for CO based on precipitation in 2013

Year	Extracted polynomial equation
2009	CO = (-2.383543) * $Precipitation^3$ + (3.416403) * $Precipitation^2$ + (-1.184290) * $Precipitation$ + (0.574204)
2013	CO = (-2.273521) * $Precipitation^3$ + (4.014026) * $Precipitation^2$ + (-1.881365) * $Precipitation$ + (0.763871)
2017	CO = (-2.3) * $Precipitation^3$ + (3.7) * $Precipitation^2$ + (-1.8) * $Precipitation$ + (0.771045)

Finally, coefficients of the polynomial for the year 2017 were also estimated based on the two previous years. The result was almost the same as the coefficients derived by the regression process. This shows that the coefficients have been almost constant during the time despite of a slight decrease which is happened in the recent years (around 2017).

5. Conclusion

In this study, we tried to analyze spatio-temporal changes of precipitation and air pollution in Tehran metropolis and to extract any probable relationship between them. In order to do this, precipitation data of meteorological stations and air pollution data of Tehran Air Quality Control Company in three predefined years (2009, 2013 and 2017) were gathered. Amounts of precipitation and air pollution were

Proceedings of the International Cartographic Association, 2, 2019.
29th International Cartographic Conference (ICC 2019), 15–20 July 2019, Tokyo, Japan. This contribution underwent single-blind peer review based on submitted abstracts. https://doi.org/10.5194/ica-proc-2-23-2019

interpolated for the whole area of Tehran for each month of the three years. Then, the correlation between precipitation and air pollution in each month was extracted using Pearson Correlation Coefficient. Results of the analysis declare that the air pollution has decreased in the months with an increased precipitation measures. This analysis was also used for the data extracted from Zonal Statistical as Table analysis within each municipal district of Tehran, which shows that the correlation has declined over the time due to the increase in effects of the other influencing factors on air pollution including population, vehicles and factories.

Moreover, results of the ANOVA test which is performed on the data indicates the correlation between precipitation and air pollution. The correlation was more meaningful while the test was performed based on the municipal districts. However, it has reduced in the months with an increased air pollution.

Also, using the polynomial regression model, the relationship of air pollution was formulated based on the precipitation data. The coefficients derived from the model have been almost constant during the three epochs of time; however, in recent years such a relationship has slightly decreased. This is probably due to the effects of other influencing factors such as population growth, wind direction and increasing the number of fossil fuel vehicles in the city.

Controlling the factors which influence air pollution is more possible by humans than precipitation. We need solutions such as managing distribution of population or making appropriate policies of daily activities, especially in the cold months of the year, to reduce traffic and air pollution. With regard to result of the current study, influence of precipitation on declining the air pollution decreases when and where the precipitation decreases. Therefore, in the areas or for the years that the precipitation is less, it is more vital to make policies to eliminate impress of the above mentioned influencing factors in order to decrease the air pollution.

Finally, as the particles with a diameter of 2.5 and 10 microns in the air that are larger than carbon monoxide remain stable and immersed in the air after precipitation more than CO, it is recommended to be considered as the air pollution indicator in the future studies.

6. References

Aghamohammadi, H., et al. (2017). "Developing an Analytical Model based on Spatial Statistics for Analyzing Rainfall in the Catchment Area of Lake Urmia." Journal of Geomatics Science and Technology **7**(2): 139-151.

ATRKAR, R. S. and Z. FATHI (2017). "The Study of Education Effects and Their Comparison on the Environment Pollution at the Different Educational Levels In Selected MENA Countries."

Bahari, R. A., et al. (2016). "Zoning of Particulate Matters (PM) Pollution Using Local Statistical Models in GIS (Case Study: Tehran Metropolisies)." Journal of Geomatics Science and Technology **5**(3): 165-174.

Brunekreef, B. and S. T. Holgate (2002). "Air pollution and health." The lancet **360**(9341): 1233-1242.

Carrera-Hernández, J. and S. Gaskin (2007). "Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico." Journal of Hydrology **336**(3-4): 231-249.

Datta, B. (2017). "Essays on Public Spending in Education, Social Protection, and Infrastructure."

Ernst, A. and J. D. Zibrak (1998). "Carbon monoxide poisoning." New England journal of medicine **339**(22): 1603-1608.

Hadei, M., et al. (2018). "DISTRIBUTION AND NUMBER OF ISCHEMIC HEART DISEASE (IHD) AND STROKE DEATHS DUE TO CHRONIC EXPOSURE TO PM2. 5 IN 10 CITIES OF IRAN (2013-2015); AN AIRQ+ MODELLING." Journal of Air Pollution and Health **2**(3): 129-136.

Hasanalizadeh, N., et al. (2015). "Modeling Spatio-temporal variation of monthly precipitation (Case Study Golestan province)." Journal of Water and Soil Conservation **22**(1): 251-269.

Jahanbakhsh ASl, S., et al. (2015). "Analysing spatio-temporal distribution of precipitation in Tabriz using cluster analysis." GEOGRAPHIC SPACE **15**(50): 59-81.

Kan, H., et al. (2008). "Season, sex, age, and education as modifiers of the effects of outdoor air pollution on daily mortality in Shanghai, China: The Public Health and Air Pollution in Asia (PAPA) Study." Environmental health perspectives **116**(9): 1183.

Khorshiddoust, A. m. and A. a. Shirzad (2014). "The Study of Precipitation in North of Iran Using Cluster and Discriminative Function Analyses." geoplanning **18**(49): 101-118.

M., L., et al. "EFFECTS OF PRECIPITATION VARIATIONS ON THE STATUS OF WATER QUALITY AND QUANTITY FOR DRINKING AND AGRICULTURE OF SEZAR RIVER." JOURNAL OF PHYSICAL GEOGRAPHY **5**: 51-62.

Meshkini, A., et al. (2017). "Spatial and Chronological Analysis of Deaths from Gastric Cancer in Tehran." geographical researches quarterly journal **31**(4): 7-23.

Najafpoor, A. A., et al. (2015). "Trend analysis of Air Quality Index criteria pollutants (CO, NO2, SO2, PM10 and O3) concentration changes in Tehran metropolis and its relationship with meteorological data, 2001-2009." Journal of Health in the Field **3**(2).

S., S., et al. (2011). "Using Cluster Analysis for Quality Evaluation of Ground water in Evan." Water engineering **4**(11): 75-85.

Valdes-Pineda, R., et al. (2016). "Spatio-temporal trends of precipitation, its aggressiveness and concentration, along the Pacific coast of South America (36–49° S)." Hydrological Sciences Journal **61**(11): 2110-2132.

Proceedings of the International Cartographic Association, 2, 2019.
29th International Cartographic Conference (ICC 2019), 15–20 July 2019, Tokyo, Japan. This contribution underwent single-blind peer review based on submitted abstracts. https://doi.org/10.5194/ica-proc-2-23-2019 |