International Journal of Geo-Information

MDPI

Article

An Assessment of Spatial Pattern Characterization of Air Pollution: A Case Study of CO and $PM_{2.5}$ in Tehran, Iran

Roya Habibi, Ali Asghar Alesheikh *, Ali Mohammadinia and Mohammad Sharif

Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology, Tehran 19967 15433, Iran; rhabibi@mail.kntu.ac.ir (R.H.); ali.mohamadinia@email.kntu.ac.ir (A.M.); msharif@mail.kntu.ac.ir (M.S.)

* Correspondence: alesheikh@kntu.ac.ir; Tel.: +98-21-8878-6212

Academic Editors: Jamal Jokar Arsanjani and Wolfgang Kainz

Received: 6 May 2017; Accepted: 27 August 2017; Published: 31 August 2017

Abstract: Statistically clustering air pollution can provide evidence of underlying spatial processes responsible for intensifying the concentration of contaminants. It may also lead to the identification of hotspots. The patterns can then be targeted to manage the concentration level of pollutants. In this regard, employing spatial autocorrelation indices as important tools is inevitable. In this study, general and local indices of Moran's I and Getis-Ord statistics were assessed in their representation of the structural characteristics of carbon monoxide (CO) and fine particulate matter ($PM_{2.5}$) polluted areas in Tehran, Iran, which is one of the most polluted cities in the world. For this purpose, a grid (200 m $\times$ 200 m) was applied across the city, and the inverse distance weighted (IDW) interpolation method was used to allocate a value to each pixel. To compare the methods of detecting clusters meaningfully and quantitatively, the pollution cleanliness index (PCI) was established. The results ascertained a high clustering level of the pollutants in the study area (with 99% confidence level). $PM_{2.5}$ clusters separated the city into northern and southern parts, as most of the cold spots were situated in the north half and the hotspots were in the south. However, the CO hotspots also covered an area from the northeast to southwest of the city and the cold spots were spread over the rest of the city. The Getis-Ord's PCI suggested a more polluted air quality than the Moran's I PCI. The study provides a feasible methodology for urban planners and decision makers to effectively investigate and govern contaminated sites with the aim of reducing the harmful effects of air pollution on public health and the environment.

Keywords: spatial autocorrelation; spatial clusters; Moran's I; Getis-Ord; air pollution; Tehran

1. Introduction

Nowadays, air pollution is one of the most important challenges in urban life due to its impacts on the environment, climate and public health [1]. The most important air pollutants that have adverse effects on human health include carbon monoxide (CO), nitrogen dioxide ($NO_{2)}$, sulfur dioxide (SO_2), ozone (O_3) and fine particulate matter ($PM_{2.5}$) [2]. These are considered in calculations of air quality index (AQI) and monitored by air quality stations in cities. Effects of sustained exposure to high air pollution levels is devastating to human health, and is a cause of various types of disorders such as cardiovascular and respiratory disease, eye irritation, diabetes, damage to the body's immune, neurological, and reproductive systems, and a decrease in life expectancy [3–7].

Population growth, industrial activities, traffic patterns, and urbanization play important roles in increasing concentrations of urban air pollutants [8]. Tehran has been experiencing a severe air pollution problem. Significant efforts to reduce and control the concentration level of pollutants and poor urban air quality have been made by major urban authorities. In the colder seasons, air pollution

intensifies due to inversion phenomena, which cause pollutants to become trapped close to the Earth's surface [9]. In recent years, air pollution during the winter months has altered the air quality status to critical and caused the closure of many schools and government agencies, with serious damage to the economy as well as public health [10]. To reduce damage and protect health in urban areas, detecting spatial clusters of air pollutants and determining hotspots based on their concentration levels throughout the city is crucial. Analysis of these patterns plays a significant role in determining the quality of life, site selection for establishing sensitive buildings such as schools or hospitals, and planning traffic routes that minimize the inhalation of pollutants. Patterns also provide valuable information for environmental health impact assessments [11]. At a region of interest, it is assumed that the amounts of pollutants were obtained from observations of the closest monitoring stations. However, observed values contain previous concentrations of pollutants at that location and do not express small changes in emissions [12]. Therefore, the presence of spatial clusters of pollutants in these areas indicates likely locations for the accumulation of pollutants, rather than representing sources of emissions, for which factors such as geographical and climatological conditions play a role.

Spatial autocorrelation is one of the main analytical tools in geographic information science (GIS). Spatial autocorrelation can authenticate that spatial data do not follow a random distribution and one or more underlying processes have led to this pattern [13]. Among spatial autocorrelation approaches for identifying spatial patterns and hotspots, Moran's I index is the most popular method capable of detecting explicit outliers [14], and has been utilized in many disciplines such as disease occurrence [15], urban planning and management [16,17], social networks studies [18], and soil pollution [19,20]. For air pollution, Fang et al. measured the spatial autocorrelation of AQI at the city level using global and local Moran's I, and estimated the comprehensive impacts and spatial variations of China's urbanization process on air quality [21]. The study investigated emission inventory data gathered from 289 cities in China and found a significant spatial dependence in the AQI values. Based on 2015 AQI data that were gathered from 161 major cities in China, Pu et al. also evaluated spatial distribution and seasonal spatial variation of AQI using Moran's I indices. In addition, the most influential urban indicators of AQI were estimated quantitatively [22]. Another study examined 16 polycyclic aromatic hydrocarbons (PAHs) pollutants at an industrially contaminated site [23]. The investigation used global Moran's I for detecting autocorrelations significance and local Moran's I for identifying hotspots. Getis-Ord statistics (Getis) is another method of spatial pattern characterization. In the literature, some studies applied this approach to identifying levels of spatial clusters. The mapping of hotspots of underground water quality in three regions of Jordan was conducted using local Getis [24]. The study revealed spatial and temporal changes in pH, nitrate and conductivity levels in 2004 and 2010 using data from 57 wells. Regarding air pollution, Réquia et al. estimated vehicle emissions for main traffic routes and evaluated their spatial pattern by applying global Moran's I and Getis tests [25]. Nevertheless, there is a lack of comparative research into spatial indices in previous studies. Therefore, this study contributes to the literature with spatial indices comparison for representation of the structural characteristics of contaminated areas. Although recognition of the factors that lead to these spatial patterns is important, this issue is not investigated in this paper and will be left for future studies. The objectives of the present study are: (1) investigation of the concentration of CO and $PM_{2.5}$ contaminations in Tehran city using hotspots detection and examining their significance statistically; and (2) comparing the spatial autocorrelation indices. We believe that the results will provide important information for science-based policy making in Tehran city to improve public health and to develop proper strategies for controlling the air quality.

2. Materials and Methods

Due to limited number of air pollution monitoring stations and their inappropriate distributions in the study area, a grid is applied to the entire surface of the study area. Concentration levels of CO and $PM_{2.5}$ from the 38 points were interpolated using the IDW method to generate a continuous surface and then concentration values for each pixel of the grid. The pixel size of the grid is 200 m $\times$ 200 m.

Different pixel sizes (as spatial resolution) will lead to different results using various data, and the study areas will experience ecological fallacy [13]. However, our examination and implementation of the model indicate that 200 m is the best size for our purpose. IDW is used to estimate CO and $PM_{2.5}$ concentration in the non-sampled locations. IDW is a deterministic interpolation method, with its estimation based on weighted average [26]. Its model formula is given as:

$$Z(\mathrm{x}) = \frac{\sum_{i=1}^{n} w_i(\mathrm{x})\mathrm{z}_i}{\sum_{j=1}^{n} w_j(\mathrm{x})}, w_i(\mathrm{x}) = \frac{1}{d(x, x_i)^p} \quad (1)$$

where $Z(x)$ is the predicted value at interpolated point x, x_i is the definite point, d is the distance between predicted and definite points, n is the total number of definite points used, w_i is the weight assigned to point i and p is a positive real number as power (usually considered as 2), which is the main factor affecting the accuracy of the model. Points closer to the interpolated point obtain greater weighting values. As the distance increases, assigned weights to known points decline and as the power parameter rises, this effect decreases [27]. In this paper, the power value was considered as 2. This parameter achieved by minimizing the prediction error calculated from the cross-validation approach. Cross-validation is an iterative process in which one sampled point is left out and the interpolation model is performed by the rest of the sampling points. Then, an estimated value is obtained for that observed point. This process goes on until a value is estimated for all the observed data points. The prediction error will be calculated using a comparison of the observed and estimated values. Both the Moran's I and Getis indices were calculated based on a distance band of 400 m (inverse Euclidean distance). All statistics and analyses were performed by employing ArcGIS 10.2.

2.1. Study Area

This study was conducted in the city of Tehran, the capital of Iran. Tehran is divided into 22 urban districts. According to the 2016 census, the Tehran metropolitan area contains almost 13.3 million urban residents [28], and 10 million related to diurnal migration [12]. Geographically, the city spreads from latitude 35°35′ N to 35°48′ N and longitude 51°17′ E to 51°33′ E. It is located at the foot of Alborz Mountains in the north, Jajroud valleys in the east, Karaj valleys in the west and the southwestern margin of central desert from the south. The average elevation of Tehran is 1200 m above sea level and there is 700 m of altitude difference between the highest and lowest points. This means that weather conditions vary across the city.

2.2. Data Sources

Hourly CO and $PM_{2.5}$ concentrations in 2015 were obtained from 38 air quality monitoring stations. Seventeen stations belonged to the Department of Environment and 21 were administered by the Air Quality Control Company of Tehran (http://air.tehran.ir/). Monthly and seasonal average measurements were calculated using this hourly basis. Figure 1 shows the locations of the monitoring stations and urban districts. Table 1 shows several statistical parameters, including minimum, maximum, average and standard deviation for concentrations of CO and $PM_{2.5}$ in districts of Tehran.

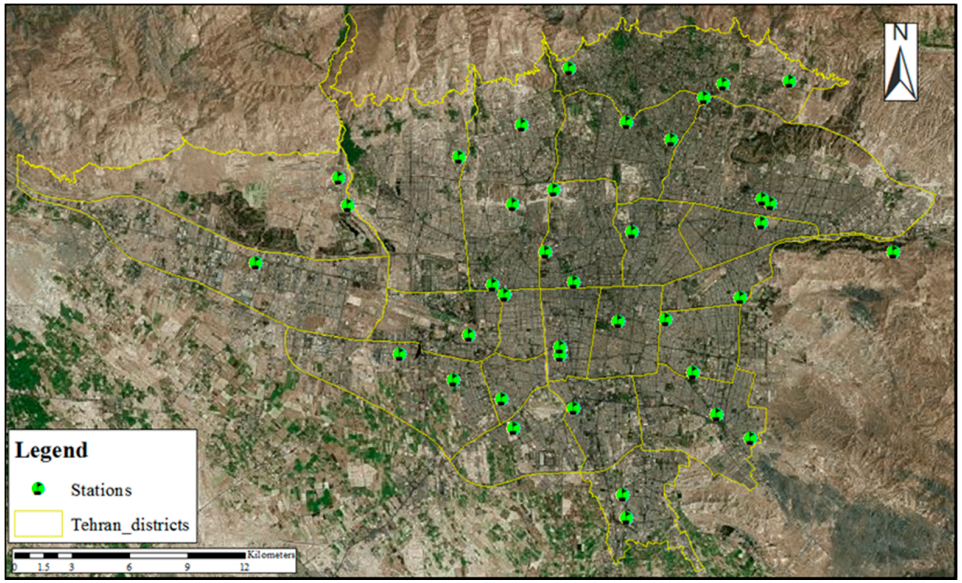

Figure 1. Tehran districts and air quality monitoring stations.

Table 1. Statistical parameters for carbon monoxide (CO) and fine particulate matter ($PM_{2.5}$) in districts of Tehran.

District	CO (ppb)				PM2.5 (μg/m^3)				District	CO (ppb)				PM2.5 (μg/m^3)			
	Min.	Max.	Ave.	Std.	Min.	Max.	Ave.	Std.		Min.	Max.	Ave.	Std.	Min.	Max.	Ave.	Std.
1	2.87	3.69	3.26	0.25	17.33	40.19	28.45	6.00	**12**	2.58	3.27	2.85	0.42	16.70	41.61	32.63	6.70
2	1.70	3.40	2.39	0.50	15.78	38.54	28.01	5.78	**13**	1.73	3.10	2.29	0.41	13.51	35.66	23.02	5.98
3	2.23	3.43	2.66	0.34	18.23	54.35	32.43	8.74	**14**	2.29	4.19	2.69	0.49	15.44	45.65	31.76	7.97
4	2.35	3.42	2.68	0.36	19.32	39.19	31.64	5.28	**15**	1.78	4.08	2.91	0.74	15.39	52.80	31.10	10.10
5	2.05	3.51	2.54	0.39	17.24	39.55	28.80	5.74	**16**	2.66	6.16	3.25	0.91	17.13	59.01	36.89	11.41
6	2.36	4.25	2.88	0.51	15.20	41.53	27.28	6.64	**17**	2.55	3.88	2.98	0.38	19.81	55.79	38.60	9.91
7	2.23	3.84	2.88	0.50	18.83	41.94	32.32	6.43	**18**	1.76	2.92	2.10	0.30	21.43	51.56	35.56	8.29
8	2.21	3.24	2.53	0.35	17.35	36.83	29.15	5.03	**19**	2.30	5.50	3.41	0.98	18.00	52.34	35.52	8.20
9	1.74	3.78	2.58	0.55	20.80	50.85	35.33	7.71	**20**	1.89	2.99	2.30	0.35	23.79	55.70	41.47	9.80
10	2.48	4.50	3.10	0.55	18.17	45.26	33.68	7.00	**21**	1.71	4.17	2.37	0.61	19.25	47.92	31.95	7.36
11	2.72	5.17	3.18	0.63	19.25	44.71	34.74	7.16	**22**	2.03	3.15	2.41	0.33	18.83	43.37	31.16	6.55

2.3. Spatial Autocorrelation Methods

In terms of spatial datasets, *self-correlation* or *autocorrelation* is a property referring to the characteristic distance or length, or lags, by which the dataset is correlated with itself [13]. Most of the statistical tests are performed by defining a null hypothesis. Global Moran's I and Getis-Ord general G statistics are inferential and the null hypothesis for them is an independent random process (IRP) or complete spatial randomness (CSR). IRP/CSR refers to a stochastic process in which there is an equal probability of a point taking place in any location and this location is independent from the others [29]. In this paper, the null hypothesis was defined as the lack of clustering pattern in CO and $PM_{2.5}$ (randomness of air pollution pattern) throughout the city. Spatial correlation of Tehran's overall pollution concentrations was measured using global Moran's I and Getis-Ord general G indices.

2.3.1. Global Spatial Correlation of Contaminants

The average amount of clustering throughout the geography is indicated using global clustering [30]. Spatial correlation of Tehran's overall pollution concentrations was measured using global Moran's I and Getis-Ord general G indices. In a given period t, global Moran's I (I_t) is calculated using the following equation [31]:

$$I_t = \frac{N}{S_0} \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij}(X_i - X)(X_j - X)}{\sum_{i=1}^{N} (X_i - \overline{X})}$$
$$S_0 = \sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij} \quad (2)$$

where X_i and X_j, respectively, are concentrations of pollutants in the i^{th} and j^{th} pixel, $\overline{X}$ is the mean of concentrations, W_{ij} is the spatial weight between the pixel i and j, N is the total number of pixels, and S_0 is the aggregate of all spatial weights.

G statistics is used by Getis for measuring hot and cold spots, with positive and negative Z-scores, respectively [31]:

$$G = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij} X_i X_j}{\sum_{i=1}^{N} \sum_{j=1}^{N} X_i X_j}, \forall j \neq 1 \quad (3)$$

The variables are the same as used in Moran's I. The values of both indices change in the range of $[-1, 1]$. The values greater than 0 shows an overall positive spatial correlation, while the values less than 0 indicate a negative correlation and 0 values express insignificance.

2.3.2. Local Spatial Autocorrelation of Pollutants

Global indices are suitable for studying spatial distribution; however, they do not provide information about the locations of clusters. Therefore, local spatial autocorrelation indices indicate the similarities or differences of one location compared with its neighbors. Hence, local Moran's I [32] and Getis-Ord G_i [33] indices are used to detect the locations of clusters in this study. The equation for local Moran's I is:

$$I_i = Z_i \sum_{i=1}^{N} W_{ij} Z_j$$
$$Z_i = (y_i - \overline{y})/S \quad (4)$$

where W_{ij} is the weight between i^{th} and j^{th} points, Z_i and Z_j are the Z-scores in point i and j, y_i is the pollutant value at point i, $\overline{y}$ is the mean of its values and S is the sum of all weights. When facing clusters with similar values in a definite region of study area, the local G_i statistic is suitable for

detecting non-stationary in data. However, an area may have high value, but its neighbors have low values, which is not considered statistically significant. The equation for local G_i is indicated as follows:

$$G_i = \frac{\sum_{j \neq i} W_{ij} y_i}{\sum_{i=1}^{N} y_i} \quad (5)$$

where the notations are the same as local Moran's I. In comparison to local Moran's I, in local G_i sums, rather than covariance, are computed.

The values of indices depend on the assumptions of the spatial weights matrix (w_{ij}). We must construct a matrix that precisely reflects our assumptions about air pollution spatial phenomenon. We can also allocate different weights to neighbors depending on our definition of neighbor. Moreover, an alternative is to use a distance decay function. However, selection of a spatial weights matrix should be consistent with the spatial pattern, data and study area; an IDW matrix has been evaluated as the best method for running models.

3. Results and Discussion

Table 2 shows the *Z*-score values of global Moran's I and Getis-Ord general G methods that are appropriate indices for investigating the intensity of clustering. All *p*-values for both methods are so small to be considered as 0, so these are not taken into account. Therefore, the null hypothesis is rejected with a 99% confidence level and 1% error, which indicate significant clustering for both CO and $PM_{2.5}$. Statistical tests conducted on collected data show that the maximum *Z*-score value and Moran's I for CO occurred in December (102.85 and 0.86, respectively), and the minimum values occurred in March (57.52 and 0.48, respectively). Also, for $PM_{2.5}$, maximum values are 74.61 and 0.63 in July, and minimum values are 52.02 and 0.44 in May, respectively. Considering seasonal results, the maximum and minimum *Z*-score and Moran's I take place in fall (292.24 and 1.004, respectively) and winter (289.75 and 0.996, respectively), for CO. These critical values also happen in fall (289.25 and 0.995, respectively) and winter (285.54 and 0.982, respectively) for $PM_{2.5}$.

Table 2. Monthly and seasonal results of global Moran's I and Getis-Ord General G indices on carbon monoxide (CO) and fine particulate matter ($PM_{2.5}$).

Time Interval		PM2.5			CO		
		Getis	Moran's I		Getis	Moran's I	
		Z-Score	Index Value	Z-Score	Z-Score	Index Value	Z-Score
Season	Spring	84.38	0.994	288.87	104.37	1.000	290.96
	Summer	98.08	0.985	286.31	107.62	0.998	290.28
	Fall	91.06	0.995	289.25	146.52	1.004	292.24
	Winter	72.00	0.982	285.54	82.70	0.996	289.75
Month	Jan.	44.76	0.445	52.7	46.92	0.491	58.11
	Feb.	45.47	0.453	53.64	47.34	0.496	58.69
	Mar.	46.48	0.467	55.26	46.78	0.486	57.52
	Apr.	46.03	0.458	54.22	54.5	0.571	67.55
	May	44.18	0.44	52.02	55.34	0.582	68.81
	June	57.63	0.589	69.63	51.79	0.589	69.63
	July	61.79	0.631	74.61	57.31	0.603	71.28
	Aug.	50.85	0.516	61.04	64.34	0.666	78.81
	Sep.	46.34	0.467	55.29	54.06	0.554	65.52
	Oct.	48.80	0.497	58.84	51.90	0.497	58.84
	Nov.	52.7	0.536	63.38	52.96	0.550	65.12
	Dec.	53.66	0.547	64.77	91.33	0.868	102.85

On the other hand, the Getis' results show that the observed general G values for both pollutants in all months are 2×10^{-6} and in all seasons are 3×10^{-6}. Table 2 shows the minimum and maximum Z-scores of Getis for CO are 46.78 and 91.33 in March and December, respectively. For $PM_{2.5}$, the minimum and maximum values were recorded in May (44.18) and July (61.79), respectively. The Z-score specifies the intensity of clustering, and results in Table 2 indicate that the Z-score values of Moran's I are higher than for Getis. Larger clusters of Moran's I in comparison with Getis are presented. Our results confirm previous research that concluded that December, when fall becomes winter, is a critical month in air pollution analysis [34,35]. The reopening of schools by the end of summer also causes an increase traffic and maximizes air pollution in fall, more attention should also be paid to this risky season [12]. Air pollution is reduced in winter when there are atmospheric precipitations, and wind is also an important factor in reduction and transmission of air pollution [36,37].

A low negative Z-score in Moran's I method indicates a statistically significant spatial data outlier. The results indicate an absence of CO or $PM_{2.5}$ outliers throughout the year. High Z-score values throughout the study period show that CO and $PM_{2.5}$ are highly clustered in Tehran. The *p*-value explains the randomness probability of the pattern, which is less than 1% for both pollutants (close to 0). The indices values exhibit positive spatial correlation at a 99% confidence level. It means that districts with relatively high level of pollutants are close together, and the same is true for those with low levels.

Figure 2 displays monthly and seasonal Z-score values for CO and $PM_{2.5}$. Similar monthly trends and results are concluded for both pollutants utilizing Moran's I and Getis approaches, but the Z-score values of Moran's I are always more than Getis (with the maximum value in December). Seasonally, pollutant trends are not similar. Here, Moran's Z-score values are still more than Getis-Ord, however, the gap between them is increased especially in fall for CO. This means that Moran's I is more sensitive in cluster detection than Getis. Alternatively, Getis is more selective in introducing regions as clusters throughout the study area.

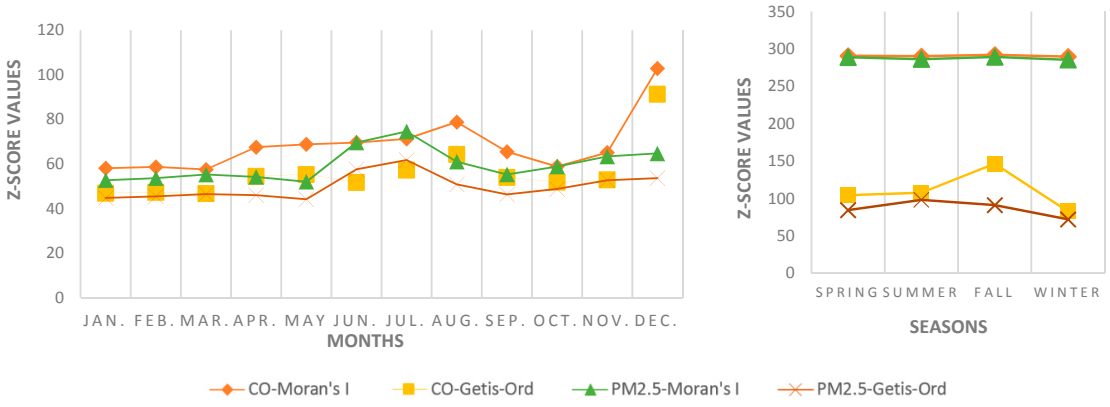

Figure 2. Monthly and seasonal Z-score values for carbon monoxide (CO) and fine particulate matter ($PM_{2.5}$).

In this article, monthly and seasonal air pollution clustering was investigated; however, due to the number of maps, only seasonal clustering is presented in Figure 3. Local Moran's I exhibits statistically significant clusters and outliers for a 95% confidence level by dividing the study area into three categories: high-high (HH), low-low (LL), and not statistically significant parts, presenting high Z-score value areas surrounded by low value ones (HL), and *vice versa* (LH). HH values imply that pollutant site with high value of Z-score is surrounded by other highly polluted locations. The inverse is true for the LL category. Spatial distribution of HH and LL values suggest the true structure of hotspots and cold spot of the pollutants. As mentioned above, no outliers were found in Moran's I Z-scores; therefore, there are not any clusters with high-low values characteristics in the city. Local G_i indicates statistically significant clusters for three confidence levels including 99, 95, and 90%. Generally, the

Getis maps are split into seven classes: hotspots and cold spots at three confidence levels and the not statistically significant parts.

Cold spots were more common in the western part of city, which is consistent with previous studies, as wind direction in Tehran is from west to east [38,39]. Hotspots were situated in a strip from northeast to south in all seasons—this strip is characterized by a high population density and shopping centers. A few cold spots can also be found in eastern districts—these areas seem to have a low population density and less traffic. Moreover, there is no cold spot at the 95 or 99% confidence level in fall, which means that the air quality is poorer in this season, and people are exposed to higher pollutant concentrations. In contrast to CO, PM2.5 hotspots and cold spots areas divide the city into two northern and southern regions. Cold spots are situated in the mid-north and north, while hotspots are mostly located in the south. Therefore, we have clean areas in north half of Tehran and polluted areas are gathered in south half. These findings are consistent with the fact that as we travel from north of Tehran to south, the population and price of houses decreases.

Although local G_i statistics detect clusters even at the 90% confidence level, Getis clusters are located inside Moran's I clusters. This means that clusters detected by Moran's I cover larger areas than Getis-Ord ones; therefore, the Getis method is more inflexible in considering areas as clusters. Moran's I also discovered some clusters with lower Z-score values in western regions, while no signs of these exist on the Getis maps.

Taking CO hotspot distributions into consideration, it seems that districts 6, 11, 16, 17, 19, intersection of districts 1, 3, 4 and the eastern half of 18, present potentially problematic pollutant concentrations. Additionally, clean areas are generally distributed over the north and western half of the study area including districts 2, 9, 18, 22 and western parts of 1. There are also some cold spots in the eastern borders of Tehran encompassing districts 4, 13, 15, and the north of 14.

$PM_{2.5}$ clusters are found in the southern parts of the city as well as districts 16, 20, the eastern half of 18, southern parts of 11, intersection of 12, 13 and 14, and the eastern part of district 3. Districts 2, 6, the northern parts of 8 and 11, the western parts of 5, 8, 13 and 14, and the eastern part of 11 are free of $PM_{2.5}$. Therefore, detection clusters show which districts need more attention and lead to appropriate allocation of resources by decision makers in Tehran. Generally, hotspots of both CO and $PM_{2.5}$ are in the southern half of the city. This is primarily because of the topography of Tehran, which slopes from north to south, and its surrounding mountains. These lead to poor wind circulation with low wind speed, and to western winds being trapped so the pollutants remain in the city. Another reason is temperature inversion, especially in these regions [40].

Figure 4 shows the monthly and seasonal numbers of CO and $PM_{2.5}$ clusters obtained by each method. To present reasonably comparable results, only those clusters of Getis with 95 and 99% confidence levels are considered. Generally, the number of CO clusters is more than those of $PM_{2.5}$. Prevention plans should be considered in high risk areas. Getis-Ord found more clusters in the first half of the year and Moran's I found more clusters in the second half. Hence, temporal analysis would be useful. However, seasonal results do not follow this trend. Overall, spring and winter have more clusters than summer and fall.

With respect to CO, the maximum number of clusters were detected in March by Moran's I and in July by Getis-Ord. Regarding $PM_{2.5}$, both methods have identified the maximum number of clusters in May and the minimum number of clusters in June (at the beginning of the school holidays). There are few clusters detected in March using Moran's I and in July by Getis (the middle of summer holidays). On the other hand, there is a bigger difference between the results of two methods regarding CO. The number of CO clusters discovered in March and June are the maximum and minimum values by Moran's I and in June and December by Getis statistic.

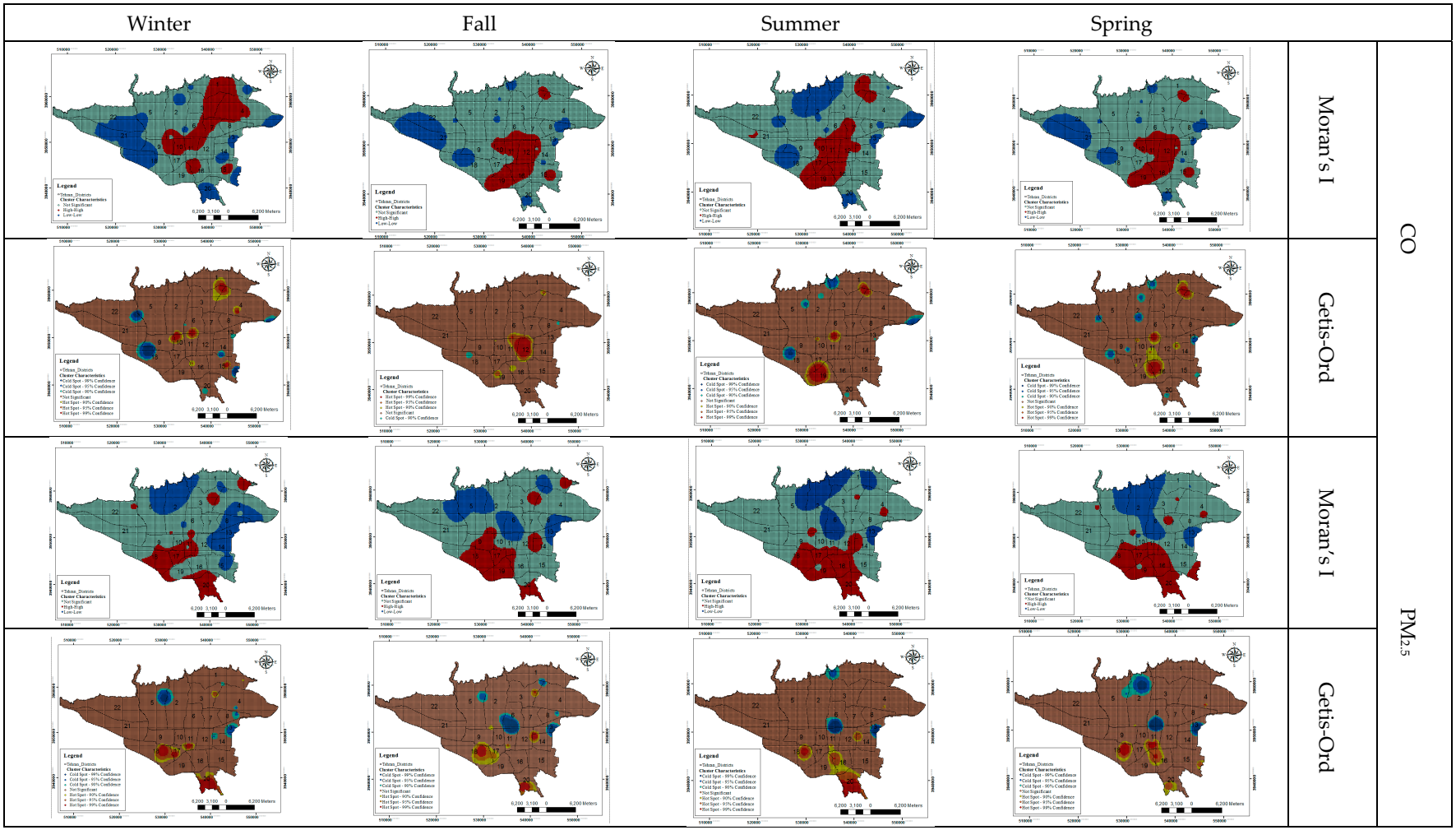

Figure 3. Seasonal clusters detected using local Moran's I and Getis-Ord G_i.

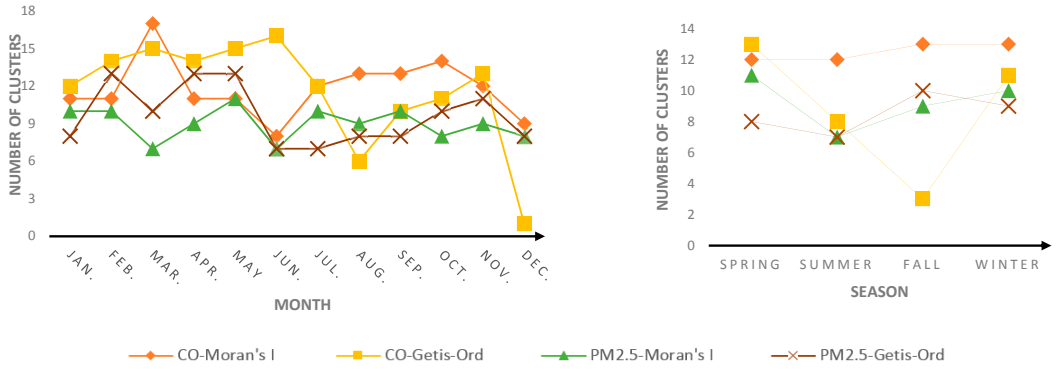

Figure 4. Monthly (**left**) and seasonal (**right**) number of carbon monoxide (CO) and fine particulate matter ($PM_{2.5}$) clusters identified by local Moran's I and Getis-Ord G_i.

As mentioned before, Moran's I considers larger areas for each cluster (see Figure 3). This results in the aggregation of clusters in some cases, and, consequently, a reduction in overall number of cluster numbers. This method identified clusters in some places which were not detected as clusters in Getis. Figure 5 displays the percentage of detected areas as hotspots or cold spots of CO and $PM_{2.5}$ throughout Tehran. In both kinds of clusters, CO and $PM_{2.5}$ clusters of local Moran's I have the highest average and those of Getis get the lowest.

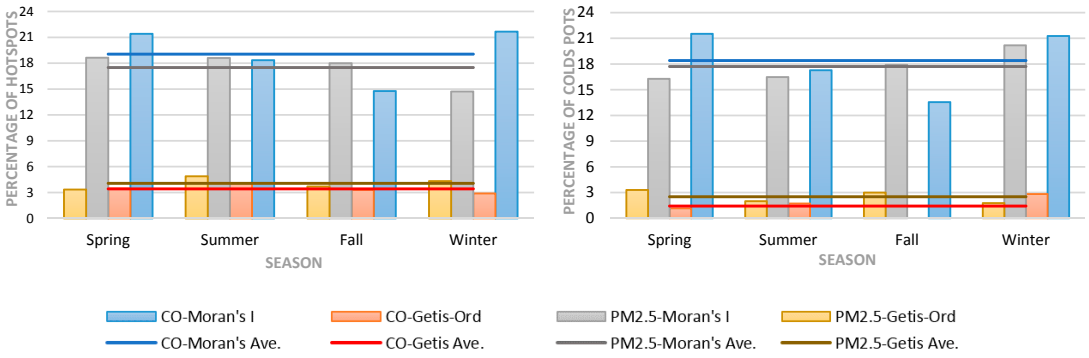

Figure 5. The percentage of hotspots (**left**) and cold spots (**right**) of carbon monoxide (CO) and fine particulate matter ($PM_{2.5}$) throughout the study area.

Detected spatial patterns of contaminants were compared by *Z*-score values, the number of detected clusters, and their location. Considering both methods detect similar places due to spatial autocorrelation, it is expected that spatial patterns will also have similarities. This issue is depicted in Figure 3. In fact, the detected clusters are found in areas where air pollution has the same behavior, due to location properties. Although, *Z*-score values have mostly similar trends for both methods, statistically significant clustered areas had differences which lead to dissimilarities in the number of detected clusters. To assess the performance of two methods meaningfully, an indicator termed the *pollution cleanliness index (PCI)* was proposed. This indicator quantitatively measures cleanliness or pollution in the study area, which is detected by two clustering methods as follows:

$$\mathrm{PCI} = \frac{\text{percentage of hotspots}}{\text{percentage of cold spots}} \tag{6}$$

The introduced method compensates for the effect of differences in detected cluster areas for both methods and reports on the general air quality of the city. As the hotspots suggest polluted sites and the cold spots indicate clean areas, PCI values that are bigger than 1 imply that the city has been

polluted and values that are smaller than 1 show the study area is generally clean. As PCI values approach 1, the spatial distribution of hotspots and cold spots are identified as more homogeneous. This means that there were as many polluted areas as there were clean areas in the city. Table 3 displays the values of this index. The overall result shows that the Getis method has a higher index in comparison with local Moran's I. This means that the Getis method states the city is located at higher level of pollution than found by Moran's I. PCI values that are closer to 1 imply more homogeneity in the spatial structure of contamination. This happened more often in PCI values of Moran's I.

Table 3. Developed pollution cleanliness index (PCI) values.

Time Interval		$PM_{2.5}$		CO	
		Getis-Ord	**Moran's I**	**Getis-Ord**	**Moran's I**
Seasons	Spring	1.012	1.145	2.961	0.995
	Summer	2.452	1.129	2.358	1.062
	Fall	1.231	1.006	-	1.09
	Winter	2.422	0.73	1.02	1.019
Months	Jan.	2.033	0.823	0.949	0.906
	Feb.	2.237	0.693	0.908	1.043
	Mar.	1.1	0.797	1.15	1.137
	Apr.	1.816	0.639	3.288	0.888
	May	0.76	0.933	2.99	1.239
	June	1.934	1.217	1.554	1.028
	July	4.459	1.102	7.182	0.916
	Aug.	1.062	0.958	11.177	1.099
	Sep.	0.935	1.067	2.722	0.887
	Oct.	1.117	1.071	2.419	0.955
	Nov.	1.548	1.122	2.406	0.975
	Dec.	0.942	1.156	-	2.039

If no cold spot is identified, t**he** returning value of the index will be infinite. This issue occurred in December and fall for Getis. Additionally, the only times that the city was specified as clean by both methods are January for CO and May for $PM_{2.5}$ (middle of winter in Tehran). This might be because of atmospheric precipitation or the wind. The air conditions of city were reported polluted in March, May, January, August, and December by two methods for CO, and clean in June, July, October, and November for $PM_{2.5}$ (summer and fall in Tehran). Moreover, considering CO, there were big differences between results of the two methods in July and August. A similar result was found in fall, where no cold spots were identified. These differences imply that Getis identified more heterogeneity in air quality than Moran's I all over the city.

4. Conclusions

In the current study, Moran's I and Getis-Ord spatial autocorrelation indices were employed for hotspot and cold spot detection of CO and $PM_{2.5}$ contaminants in Tehran city, and comparisons drawn between them. This investigation was conducted monthly and seasonally in 2014. Detected spatial patterns of contaminants were compared by *Z*-score values, the number of detected clusters, and their location, and the authors introduced the pollution cleanliness index (PCI). The results of the global spatial autocorrelation analyses found high clustering levels for both CO and $PM_{2.5}$ in the study area, which are significantly different from random at the 99% confidence level. CO has a higher level of clustering than $PM_{2.5}$ throughout the period, as found by both spatial autocorrelation methods. In the monthly results, the critical values for both pollutants occurred in the same months and, as time went on, a remarkable growing trend was detected during one-year period. However, there was no sensible similarity in the seasonal results. As stated by local indices, the distribution of CO and $PM_{2.5}$ hotspots and cold spots were in the opposite direction. $PM_{2.5}$ clusters separated the city into two northern and

southern parts, as most of cold spots were situated in the north half and hotspots were in the south. Nevertheless, CO hotspots covered an area from northeast to southwest Tehran and cold spots spread over the rest of city. Taking the clustering results into consideration, local Moran's I considered larger areas for each cluster. To compare the detected clusters of two methods meaningfully, the PCI indicator was introduced. PCI compares the spatial pattern homogeneity considering hotspots and cold spots. This index showed that the Getis-Ord statistic reports more polluted air conditions than Moran's I. Clusters, therefore may not be considered as pollution resources, but there are several factors involved in this issue that should be investigated in detail in future works.

In addition to the reasons mentioned in this paper, some activities have been performed in Tehran to prevent pollution. There are restrictions for heavy goods vehicles that prohibit them being driven on city roads or only allow them to be driven at night. Moreover, days of week have been divided into odds and evens according to right number of cars' plates. Cars require technical diagnosis cards, issued when it has passed all the technical considerations. National holidays can have a huge impact on Tehran's air pollution. We hope our study will be useful for planners and decision makers for detecting risky pollution areas so they can allocate resources appropriately and work to reduce Tehran's air pollution levels.

Acknowledgments: The authors would like to express sincere thanks to personnel of the Department of Environment and Air Quality Control Company of Tehran who provided the essential information used in this study.

Author Contributions: Roya Habibi, Ali Asghar Alesheikh and Mohammad Sharif conceived and designed the experiments; Ali Mohammadinia performed the experiments and statistical analysis; Roya Habibi analyzed the data and wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, L.; Qian, J.; Ou, C.-Q.; Zhou, Y.-X.; Guo, C.; Guo, Y. Spatial and temporal analysis of air pollution index and its timescale-dependent relationship with meteorological factors in guangzhou, china, 2001–2011. *Environ. Pollut.* **2014**, *190*, 75–81. [CrossRef] [PubMed]
2. Alesheikh, A.A.; Oskouei, A.; Atabi, F.; Helali, H. Providing interoperability for air quality in-situ sensors observations using gml technology. *Int. J. Environ. Sci. Technol.* **2005**, *2*, 133–140. [CrossRef]
3. Dix-Cooper, L.; Eskenazi, B.; Romero, C.; Balmes, J.; Smith, K.R. Neurodevelopmental performance among school age children in rural guatemala is associated with prenatal and postnatal exposure to carbon monoxide, a marker for exposure to woodsmoke. *Neurotoxicology* **2012**, *33*, 246–254. [CrossRef] [PubMed]
4. Correia, A.W.; Pope, C.A., III; Dockery, D.W.; Wang, Y.; Ezzati, M.; Dominici, F. The effect of air pollution control on life expectancy in the united states: An analysis of 545 us counties for the period 2000 to 2007. *Epidemiology* **2013**, *24*, 23. [CrossRef] [PubMed]
5. Weichenthal, S.; Villeneuve, P.J.; Burnett, R.T.; van Donkelaar, A.; Martin, R.V.; Jones, R.R.; DellaValle, C.T.; Sandler, D.P.; Ward, M.H.; Hoppin, J.A. Long-term exposure to fine particulate matter: Association with nonaccidental and cardiovascular mortality in the agricultural health study cohort. *Environ. Health Perspect.* **2014**, *122*, 609. [CrossRef] [PubMed]
6. Kim, K.-H.; Kabir, E.; Kabir, S. A review on the human health impact of airborne particulate matter. *Environ. Int.* **2015**, *74*, 136–143. [CrossRef] [PubMed]
7. Wong, C.-S.; Lin, Y.-C.; Hong, L.-Y.; Chen, T.-T.; Ma, H.-P.; Hsu, Y.-H.; Tsai, S.-H.; Lin, Y.-F.; Wu, M.-Y. Increased long-term risk of dementia in patients with carbon monoxide poisoning: A population-based study. *Medicine* **2016**, *95*, e2549. [CrossRef] [PubMed]
8. Ma, T.; Zhou, C.; Pei, T.; Haynie, S.; Fan, J. Quantitative estimation of urbanization dynamics using time series of dmsp/ols nighttime light data: A comparative case study from china's cities. *Remote Sens. Environ.* **2012**, *124*, 99–107. [CrossRef]
9. Mohammadi, H.; Cohen, D.; Babazadeh, M.; Rokni, L. The effects of atmospheric processes on tehran smog forming. *Iran. J. Public Health* **2012**, *41*, 1. [PubMed]

10. Vafa-Arani, H.; Jahani, S.; Dashti, H.; Heydari, J.; Moazen, S. A system dynamics modeling for urban air pollution: A case study of tehran, iran. *Transp. Res. Part D* **2014**, *31*, 21–36. [CrossRef]
11. Tashayo, B.; Alimohammadi, A.; Sharif, M. A hybrid fuzzy inference system based on dispersion model for quantitative environmental health impact assessment of urban transportation planning. *Sustainability* **2017**, *9*, 134. [CrossRef]
12. Amini, H.; Taghavi-Shahri, S.M.; Henderson, S.B.; Naddafi, K.; Nabizadeh, R.; Yunesian, M. Land use regression models to estimate the annual and seasonal spatial variability of sulfur dioxide and particulate matter in tehran, iran. *Sci. Total Environ.* **2014**, *488*, 343–353. [CrossRef] [PubMed]
13. O'Sullivan, D.; Unwin, D. *Geographic Information Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2014.
14. Li, W.; Xu, B.; Song, Q.; Liu, X.; Xu, J.; Brookes, P.C. The identification of 'hotspots' of heavy metal pollution in soil–rice systems at a regional scale in eastern china. *Sci. Total Environ.* **2014**, *472*, 407–420. [CrossRef] [PubMed]
15. Mohammadinia, A.; Alimohammadi, A.; Habibi, R.; Shirzadi, M.R. Spatial and statistical analysis of leptospirosis in Guilan province, Iran. *Int. Arch. Photogramm., Remote Sens. Spa. Inf. Sci.* **2015**, *40*, 497–502. [CrossRef]
16. Blazquez, C.A.; Celis, M.S. A spatial and temporal analysis of child pedestrian crashes in santiago, chile. *Accid. Anal. Prev.* **2013**, *50*, 304–311. [CrossRef] [PubMed]
17. Fan, C.; Myint, S. A comparison of spatial autocorrelation indices and landscape metrics in measuring urban landscape fragmentation. *Landsc. Urban Plan.* **2014**, *121*, 117–128. [CrossRef]
18. Lee, Y.; Kwon, P.; Yu, K.; Park, W. Method for determining appropriate clustering criteria of location-sensing data. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 151. [CrossRef]
19. Zhao, K.; Fu, W.; Liu, X.; Huang, D.; Zhang, C.; Ye, Z.; Xu, J. Spatial variations of concentrations of copper and its speciation in the soil-rice system in wenling of southeastern china. *Environ. Sci. Pollut. Res.* **2014**, *21*, 7165–7176. [CrossRef] [PubMed]
20. Zhang, C.; Luo, L.; Xu, W.; Ledwith, V. Use of local moran's i and gis to identify pollution hotspots of pb in urban soils of galway, ireland. *Sci. Total Environ.* **2008**, *398*, 212–221. [CrossRef] [PubMed]
21. Fang, C.; Liu, H.; Li, G.; Sun, D.; Miao, Z. Estimating the impact of urbanization on air quality in china using spatial regression models. *Sustainability* **2015**, *7*, 15570–15592. [CrossRef]
22. Pu, H.; Luo, K.; Wang, P.; Wang, S.; Kang, S. Spatial variation of air quality index and urban driving factors linkages: Evidence from chinese cities. *Environ. Sci. Pollut. Res.* **2017**, *24*, 4457–4468. [CrossRef] [PubMed]
23. Liu, G.; Bi, R.; Wang, S.; Li, F.; Guo, G. The use of spatial autocorrelation analysis to identify pahs pollution hotspots at an industrially contaminated site. *Environ. Monit. Assess.* **2013**, *185*, 9549–9558. [CrossRef] [PubMed]
24. Alqadi, K.A.; Kumar, L.; Khormi, H.M. Mapping hotspots of underground water quality based on the variation of chemical concentration in amman, zarqa and balqa regions, jordan. *Environ. Earth Sci.* **2014**, *71*, 2309–2317. [CrossRef]
25. Réquia, W.J.; Koutrakis, P.; Roig, H.L. Spatial distribution of vehicle emission inventories in the federal district, brazil. *Atmos. Environ.* **2015**, *112*, 32–39. [CrossRef]
26. Shepard, D. A two-dimensional interpolation function for irregularly-spaced data. In Proceedings of the 23rd ACM National Conference, Las Vegas, NV, USA, 27–29 August 1968; pp. 517–524.
27. Atabi, F.; Moattar, F.; Mansouri, N.; Alesheikh, A.; Mirzahosseini, S. Assessment of variations in benzene concentration produced from vehicles and gas stations in tehran using gis. *Int. J. Environ. Sci. Technol.* **2013**, *10*, 283–294. [CrossRef]
28. SCI. Census Results of 2016. Available online: www.amar.org.ir (accessed on 25 April 2017).
29. Shaffer, M.J.; Bishop, J.A. Predicting and preventing elephant poaching incidents through statistical analysis, gis-based risk analysis, and aerial surveillance flight path modeling. *Trop. Conserv. Sci.* **2016**, *9*, 525–548. [CrossRef]
30. Baller, R.D.; Richardson, K.K. Social integration, imitation, and the geographic patterning of suicide. *Am. Soc. Rev.* **2002**, *67*, 873–888. [CrossRef]
31. Mitchell, A. *The Esri Guide to Gis Analysis, Volume 2: Spatial Measurements and Statistics*; Esri Press: Redlands, CA, USA, 2005.
32. Anselin, L. Local indicators of spatial association—Lisa. *Geogr. Anal.* **1995**, *27*, 93–115. [CrossRef]
33. Getis, A.; Ord, J.K. *Local Spatial Statistics: An Overview*; John Wiley and Sons: Hoboken, NJ, USA, 1996.

34. Renzi, M.; Stafoggia, M.; Faustini, A.; Cesaroni, G.; Agabiti, N.; Forastiere, F. Health effects of air pollution in rome in december 2015. *Epidemiol. Prev.* **2016**, *40*, 29–32. [PubMed]
35. Hu, D.; Wu, J.; Tian, K.; Liao, L.; Xu, M.; Du, Y. Urban air quality, meteorology and traffic linkages: Evidence from a sixteen-day particulate matter pollution event in december 2015, beijing. *J. Environ. Sci.* **2017**. [CrossRef]
36. Rashki, A.; deW Rautenbach, C.; Eriksson, P.G.; Kaskaoutis, D.G.; Gupta, P. Temporal changes of particulate concentration in the ambient air over the city of zahedan, iran. *Air Qual. Atmos. Health* **2013**, *6*, 123–135. [CrossRef]
37. Hassani, A.; Hosseini, V. An assessment of gasoline motorcycle emissions performance and understanding their contribution to tehran air pollution. *Transp. Res. Part D* **2016**, *47*, 1–12. [CrossRef]
38. Yazdi, M.N.; Delavarrafiee, M.; Arhami, M. Evaluating near highway air pollutant levels and estimating emission factors: Case study of tehran, iran. *Sci. Total Environ.* **2015**, *538*, 375–384. [CrossRef] [PubMed]
39. Asghari, M.; Nematzadeh, H. Predicting air pollution in tehran: Genetic algorithm and back propagation neural network. *J. AI Data Min.* **2016**, *4*, 49–54.
40. Akbari, M.; Samadzadegan, F.; Weibel, R. A generic regional spatio-temporal co-occurrence pattern mining model: A case study for air pollution. *J. Geogr. Syst.* **2015**, *17*, 249–274. [CrossRef]