Journal of Elementology ISSN 1644-2296

Asadi A., Goharnejad H., Niri M.Z 2019.
Regression modelling of air quality based on meteorological parameters and satellite data.
J. Elem., 24(1): 81-99. DOI: 10.5601/jelem.2018.23.1.1599

RECEIVED: 24 January 2018
ACCEPTED: 7 August 2018

ORIGINAL PAPER

REGRESSION MODELLING OF AIR QUALITY BASED ON METEOROLOGICAL PARAMETERS AND SATELLITE DATA

Alireza Asadi[1], Hamid Goharnejad[1], Mahmoud Zakeri Niri[2]

[1]Department of Civil Engineering, Environmental Sciences Research Center Islamshahr Branch, Islamic Azad University, Islamshahr, Iran
[2]Young Researchers and Elite Club, Islamshahr Branch Islamic Azad University, Islamshahr, Iran

Abstract

Although field monitoring can provide an accurate measurement of pollution, these measurements are of a limited spatial coverage. On the contrary, satellite-based observations can provide Aerosol Optical Depth (AOD) products with higher spatial resolution and continuous spatial coverage; however these products cannot directly measure the pollution concentration. In this study, the potential of a Moderate-Resolution Imaging Spectroradiometer (MODIS) sensors was investigated to evaluate the air quality parameters, after which water consumption in the studied area was considered. For this purpose, linear regression analysis was used in order to develop a relationship among MODIS-AOD, metrological data (relative humidity, temperature, precipitation, and wind speed) and air pollution data (CO, O_3, NO_2, SO_2, $PM_{2.5}$) gathered 22 monitoring stations from 2012 to 2016. Among the 5 years of pollution data collection, the period of 2012 to 2014 was used for the model calibration and the period of 2015 to 2016 was used for the validation of the model. The results indicated that the regression models were of the best performance during spring (R^2 = 0.901 for CO), moderate performance during winter (R^2 = 0.674 for CO) and autumn (R^2 = 0.694 for CO), and weak performance during summer (R^2 = 0.181 for SO_2). The results of the validation process also showed that the maximum determination factor (R^2 = 0.83) was obtained during spring season and for $PM_{2.5}$ and the least (R^2 = 0.18) was obtained during summer and for SO_2. Meanwhile, the assessment of water consumption demonstrated that there is significant relationship between water consumption and the concentration of pollution parameters.

Keywords: MODIS-AOD, meteorological parameters, air pollution, linear regression model, water consumption.

Hamid Goharnejad, Assistant Professor, Department of Civil Engineering, Environmental Sciences Research Center, Islamshahr Branch, Islamic Azad University, Islamshahr, Iran, e-mail: H.Goharnejad@iiau.ac.ir; Tel.: +98-912-714-4265

INTRODUCTION

Investigation of the pattern of air pollution emissions due to spatial and temporal variability is one of the most challenging issues. In addition, the spatial and temporal variations of pollution parameters are much higher than the other components of the atmosphere. Therefore, the spatial distribution of air pollution emission is required to be estimated with a high precision in order to become aware of their influence on the weather and climate, biosphere and human health. For this purpose, satellite images could be used for preparing the quantitative data of pollution and their changes in the time and space (Wang, Christopher 2003, Lee et al. 2017, Tang et al. 2017).

Many studies show that an increase in amounts of polluting particles in the atmosphere can lead to cardiovascular and respiratory diseases (Beelen et al. 2014, Chu et al. 2003, Wong et al. 2015, Jin et al. 2017). The air pollution parameters are today monitored by using ground stations, which provide accurate point measurements for their adjacent area. These values are then attributed to the entire region by extrapolation. Nevertheless, the measuring stations are usually dispersed and mostly located in urban areas. Therefore, a regional study of air pollution parameters in the areas without measuring station is difficult and even impossible.

Over the recent years, it has been attempted to predict pollution parameters using Remote Sensing (RS) data. It has been found through many studies that the data obtained from aerosols using the RS, especially the AOD, have a close relationship with the pollution concentration. Therefore, this parameter can be used for estimating the pollution concentration (Engel-Cox et al. 2004, Gupta et al. 2006, Brauer et al. 2012, Cohan, Chen 2014, Raoufi et al. 2018). Such factors as the vertical distribution of particles, the combination of particles, and their size have affected the relation between AOD and pollution concentration (Gupta et al. 2009, Jathar et al. 2014, Kuwayama et al. 2015, Franklin et al. 2017, Liu et al. 2017). The effects have been investigated empirically in order to find a valid relationship between AOD and the concentration of different pollutants (Wang et al. 2014, Dhyani et al. 2017). To overcome the lack of necessary data, statistical models have been also used to eliminate these effects in order to find a more precise relationship between the two parameters of AOD and the concentration of pollution (Liu et al. 2004). Most of the studies have focused on the development of a simple relationship in the form of a linear regression equation between AOD and various pollutants (Liu et al. 2005, Kumar et al. 2007, Liu et al. 2015, Franklin et al. 2017). To improve the relationship between AOD and pollution parameters, meteorological parameters have also been used in some studies (Gupta et al. 2009). Other scientists exploring atmospheric physics have concluded that weather conditions such as wind velocity, relative humidity, and atmospheric temperature and pressure can disrupt the composition of AOD and $PM_{2.5}$, and the surface wind velocity, air temperature, and the boundary layer elevation are important predictors for AOD – $PM_{2.5}$ models (Kumar et al. 2007, Kelly et al. 2017).

Kumar et al. (2007) used AOD data from a MISR sensor of Terra satellite to develop a spatial-temporal model for estimating $PM_{2.5}$ and PM_{10} concentrations in southern California. For this purpose, they used the data from a measuring station to validate the model parameters. They concluded that the model combined of AOD, relative humidity, and wind speed as independent variables, is a good predictor for $PM_{2.5}$ (R^2 = 0.67), whereas, PM_{10} can be better predicted with the help of AOD and dew point (R^2 = 0.76).

There are many effective factors in the creation and emission of air pollutants over the industrial areas and metropolies such as Tehran. Industrial development, population growth, urbanization, and the subsequent development of heavy traffic have made air pollution in Tehran city as one of the most important challenges. Reports provided by Tehran Air Control Company in 2015 about the main air pollutants (CO, O_3, NO_2, SO_2, $PM_{2.5}$), as the fifth published document, have determined the air quality of Tehran city in terms of the time and space aspects of the concentration of pollutants (TAQCC. 2016). The study was based on data from 22 active stations existing in Tehran city. Accordingly, in 2015, there were 21 clean air days (6%), 233 healthy-air days (64%), 105 unhealthy days for sensitive social groups (29%), 5 unhealthy days for the general population (more than 1%) and 1 very unhealthy day (less than 1%) in Tehran city. According to the report, despite the increase in the number of clean days in comparison with previous years, 30% of the days in 2015 were still beyond the healthy standard limit. In 2015, the lowest and the highest number of polluted days occurred in the summer (especially August) and winter (November and December) seasons, respectively. In the recent years, the particles with a diameter of less than 2.5 micron ($PM_{2.5}$), have been the main pollutant in Tehran city, and during 2015, there were 111 days with inappropriate conditions regarding $PM_{2.5}$ pollutants. Based on this report, the concentration of pollutants in Tehran city over 2007-2015 showed an almost downward trend. However, the pollutants O_3 and NO_2, have been relatively increasing since 2014. According to the measurements carried out at the Tehran air quality monitoring stations, aerosols with a diameter of less than 2.5 micron ($PM_{2.5}$) are responsible for unfavorable conditions, and all polluted days in year 2015 occurred due to the increased concentration of this pollutant. The purpose of this study is to estimate the parameters of air quality in Tehran (CO, O_3, NO_2, SO_2, $PM_{2.5}$) by applying established regression relations among these data, MODIS-AOD data, and meteorological parameters (precipitation, relative humidity, temperature, and wind speed) measured at the proximate meteorological stations.

MATERIAL AND METHODS

With the aim of finding regression relations among the AOD, meteorological parameters and pollution parameters, linear regression relations were seasonally developed from January 2012 to September 2016. To achieve this purpose, daily air quality data (measured at 22 active stations) and MODIS-AOD daily data and meteorological parameters between the mentioned time period were used. The data in this time period were separated between seasonal intervals, and regression relations were developed for each season. After selecting the best regression relation for each pollution parameter, the spatial and temporal distribution map of these parameters was extracted for the most polluted days of each season. Finally, in order to investigate the effects of the amount of pollution on urban water consumption, the relationship between weekly water consumption in Tehran and air pollution in this city were studied through the development of regression relations.

The used algorithm and method are presented in Figure 1. It shows a flowchart of the steps involved in performing this research.

Description of the studied area

In this study, Tehran, the capital of Iran with more than 15 million people and covering about 730 km^2, was studied. The geographical location of Teheran is identified by longitude 51°02′ E to 51°36′ E and latitude 35°24′ N to 35°50′ N, which equals to an approximate length of 50 km and an approximate width of 30 km. The height of the city varies from 1050 m.a.s.l in the south to 2000 m.a.s.l in the north. Tehran is bordered with mountainous areas from the north and desert areas from the south, hence, there are different weather conditions in the south and north. The northern regions have cold and dry weather, and the southern areas are hot and dry. Figure 2 depicts the location of Tehran and the city's air monitoring and measurement stations that are used in this study.

Data used

Measured data of air pollutants

This study was based on Tehran's air pollution data collected at 22 active measuring and monitoring stations between 2012 to 2016. Such parameters as the concentration of major air pollutants including carbon monoxide (CO), ozone (O_2), nitrogen oxides (NO_2, NO_x, NO), sulfur dioxide (SO_2), aerosols with a diameter of less than 10 microns (PM_{10}), and particles with a diameter of less than 2.5 microns ($PM_{2.5}$) are continuously measured in these stations. The locations of the active stations are shown in Figure 2.

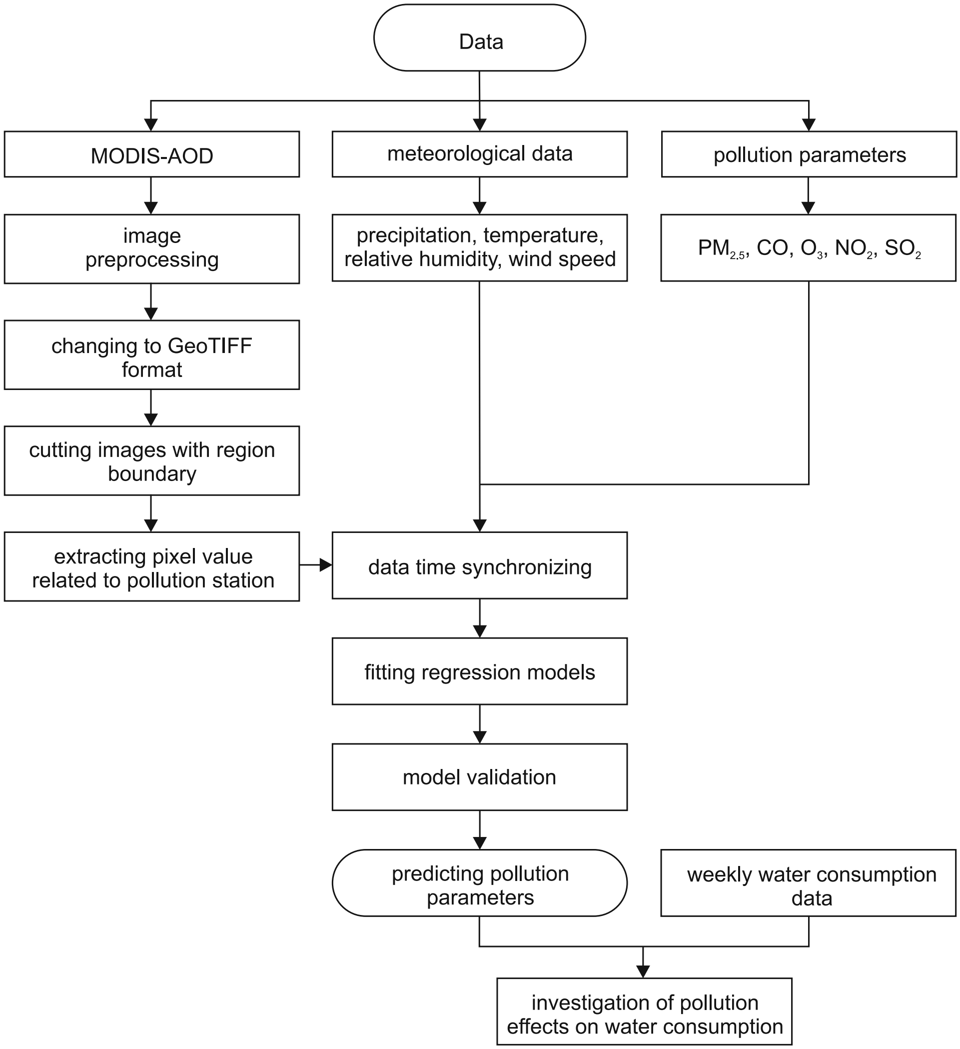

Fig. 1. Steps involved in performing this research

MODIS sensor data

'MODIS (or Moderate Resolution Imaging Spectroradiometer) is a key instrument aboard the Terra (originally known as EOS AM-1) and Aqua (originally known as EOS PM-1) satellites (HE et al. 2017). Terra's orbit around the Earth is timed so that it passes from north to south across the equator in the morning, while Aqua passes south to north over the equator in the afternoon. Terra MODIS and Aqua MODIS are viewing the entire Earth's surface every 1 to 2 days, acquiring data in 36 spectral bands, or groups of wavelengths (see MODIS Technical Specifications). These data will

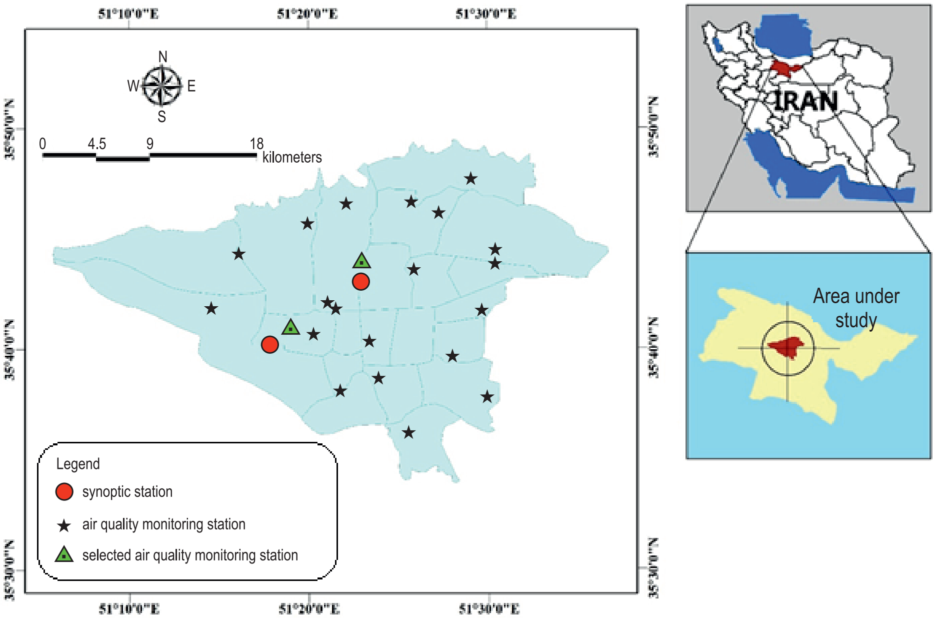

Fig. 2. Location of Tehran city, air quality measurement stations and synoptic stations

improve our understanding of global dynamics and processes occurring on the land, in the oceans, and in the lower atmosphere. MODIS is playing a vital role in the development of validated, global, interactive Earth system models able to predict global change accurately enough to assist policy makers in making sound decisions concerning the protection of our environment' (https://modis.gsfc.nasa.gov/about/).

Meteorological data and water consumption

Meteorological data including precipitation, relative humidity, air temperature, and wind speed were used to develop regression models. For this purpose, among the available 5 meteorological stations in Tehran, the data of 2 stations (Mehrabad and Geophysics stations) were selected as these were closer to the air monitoring stations (Shad Abad and Tarbiat Modarres stations), and the daily data from these 2 stations between 2012 and 2016 were used to develop the regression models. The locations of both stations is shown in Figure 2. The weekly water consumption data were also obtained from the Tehran Water Organization during the time period from 2012 to 2016.

Satellite images pre-processing

The data of a MODIS sensor can be downloaded from the NASA's website as Hierarchical Data Format (HDF). Each file contains atmospheric parameters that are measured by the MODIS sensor. It is necessary to first convert the HDF format to the GeoTIFF format and then use these data in ArcGIS. In this study, due to the number of 1706 AOD images and the time-consuming processes, a model builder tool in ArcGIS software was used to facilitate the processing and generate the necessary inputs. In this way, using the Iterate Raster tool and batch processing, all AOD images were converted to the GeoTIFF format, and the values of each image corresponding to the air pollution monitoring station were extracted from them using the 'extract multi-values to points' tool.

Developing regression models

To develop regression models among pollution parameters, meteorological parameters and AOD, it is rnecessary to develop regression equations for each pollution parameter separately with AOD and air pollution parameters. In this study, the correlation between MODIS-AOD and meteorological parameters (precipitation, relative humidity, temperature, and wind speed) were investigated as independent variables along with air quality parameters (CO, O_3, NO_2, SO_2, $PM_{2.5}$) as dependent variables. In the case of a high correlation, RS data along with meteorological data will be used to prepare the map of air pollution parameters in Tehran city. To achieve this, the time period from 2012 to 2016 was divided into two parts. The period of 2012 to 2014 served for the development of regression models and the period of 2015 to 2016 was used to validate the results. Since seasonal variation is considered, the analyses were conducted for each season. After the seasonal division of data, all data were entered into SPSS and statistical analyses were conducted on the data. In order to include the impact of the meteorological parameters on the regression equations for each pollution parameter and for each season of the year, the relationship between each pollution parameter with AOD was calculated and then, using the stepwise regression approach, the meteorological parameters were entered into the regression model to evaluate the ability of these parameters to explain the dependent variable. Each parameter that is entered into the model and increases R^2 of the developed model can explain the dependent variable, otherwise it will not affect the dependent variable and will be removed from the calculation. With this assumption, the regression equations for each season were developed between the pollution parameters, the AOD, and the meteorological parameters. Finally, a seasonal variation map of the parameters in the period of 2012 to 2016 was drawn for spatial and temporal analyses of the pollution.

Model performance evaluation indices

Four different criteria including the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean bias Error (MBE) and determination coefficient (R^2) were used to examine the accuracy of the regression model prediction. Equations (1) to (4) show how to calculate these four statistical parameters. The less the RMSE value, the higher the precision of the model will be obtained. In addition, when the MAE is closer to zero, the model used is more accurate in the prediction. The amount of MBE, which represents the bias amount of the prediction, should be close to zero in a fairly accurate model. Also R^2 closer to 1 shows good performance of the model.

$$\mathrm{MAE} = \frac{1}{n}\sum_{n=1}^{n} | O_i - P_i | \tag{1}$$

$$\mathrm{MBE} = \frac{1}{n}\sum_{n=1}^{n} (O_i - P_i) \tag{2}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{n=1}^{n} (O_i - P_i)^2} \tag{3}$$

$$R^2 = \left[\frac{\sum_{i=1}^{n} (O_i - O_{ave})(P_i - P_{ave})}{\sqrt{\sum_{i=1}^{n} (O_i - O_{ave})^2 (P_i - P_{ave})^2}} \right]^2 \tag{4}$$

Where, P_i denotes the values predicted by the model, O_i is the measured values, P_{ave} and O_{ave} are respectively the average of the predicted and measured values, and n is the number of observations.

Some other statistical parameters were used in this study. R-square change is the improvement in R-square when the second predictor is added. The R-square change is tested with an F-test, which is referred to as the F-change. A significant F-change means that the variables added in that step significantly improved the prediction. Degrees of freedom (df1 and df2) of an estimate are the number of independent pieces of information that went into calculating the estimate and finally, a significant F-change means that the variables added in that step significantly improved the prediction.

RESULTS AND DISCUSSION

Developing regression models

The development of regression models in winter, spring, summer, and autumn was done by step-by-step entering of meteorological parameters into the model developed between the AOD and each of the air pollution parameters. The summary of each regression model in winter, spring, summer, and autumn is depicted in Table 1 to 4 respectively.

Tables 1 to 4 present the statistical summary of each regression model in winter, spring, summer, and autumn respectively.

As can be seen in Table 1, the regression models are of a R^2 varied from 0.413 for SO_2 gas to 0.674 for CO gas for CO gas. The values of variations in F statistics, which are less than 0.05 in different stages of the step-by-step

Table 1

Statistical summary of regression models developed between AOD and meteorological parameters as independent variables and pollution parameters as a dependent variable in winter

Dependent variable	*R*	*R* square	Adjusted *R* square	Std. error of the estimate	Change statistics				
					R square change	*F* change	df1	df2	sig. *F* change
$PM_{2.5}$	0.733	0.538	0.522	10.676	0.000	0.087	1	144	0.000
CO	0.821	0.674	0.663	4.919	0.001	0.349	1	144	0.000
O_3	0.676	0.457	0.438	7.145	0.006	1.622	1	144	0.005
NO_2	0.711	0.506	0.488	5.627	0.001	0.219	1	144	0.040
SO_2	0.643	0.413	0.393	4.068	0.004	0.965	1	144	0.028

$$\mathrm{PM}_{2.5} = 166.694 \times \mathrm{AOD} + 0.230 \times RH - 0.306 \times P + 13.898 \tag{5}$$

$$\mathrm{CO} = 119.389 \times \mathrm{AOD} - 0.203 \times T - 0.302 \times WS + 8.400 \tag{6}$$

$$\mathrm{O}_3 = 121.183 \times \mathrm{AOD} - 0.305 \times RH + 0.314 \times T - 0.1998 \times P - 1.001 \times \times WS + 18.782 \tag{7}$$

$$\mathrm{NO}_2 = 94.337 \times \mathrm{AOD} - 0.106 \times T - 0.279 \times WS + 9.126 \tag{8}$$

$$\mathrm{SO}_2 = 44.157 \times \mathrm{AOD} - 0.039 \times RH + 0.157 \times T + 0.154 \times P - 0.440 \times \times WS + 6.227 \tag{9}$$

Table 2

Statistical summary of regression models developed between AOD and meteorological parameters as independent variables and pollution parameters as a dependent variable in spring

Dependent variable	*R*	*R* square	Adjusted *R* square	Std. error of the estimate	Change Statistics				
					R square change	*F* change	df1	df2	Sig. *F* change
$PM_{2.5}$	0.913	0.833	0.829	10.162	0.833	222.833	4	179	0.000
CO	0.949	0.901	0.899	4.957	0.901	819.360	2	181	0.000
O_3	0.867	0.752	0.745	7.578	0.002	1.455	1	178	0.229
NO_2	0.906	0.822	0.819	5.490	0.822	276.266	3	180	0.000
SO_2	0.832	0.692	0.687	3.888	0.692	134.839	3	180	0.000

$$\mathrm{PM}_{2.5} = 185.892 \times \mathrm{AOD} - 0.273 \times T + 0.385 \times P + 14.804 \tag{10}$$

$$\mathrm{CO} = 124.038 \times \mathrm{AOD} - 0.331 \times T - 0.156 \times P + 7.636 \tag{11}$$

$$\mathrm{O}_3 = 123.305 \times \mathrm{AOD} - 0.089 \times RH - 0.167 \times T - 0.257 \times P - 1.334 \times \times WS + 15.488 \tag{12}$$

$$\mathrm{NO}_2 = 88.822 \times \mathrm{AOD} + 0.178 \times RH - 0.088 \times T - 0.314 \times P - 0.312 \times \times WS + 7.637 \tag{13}$$

$$\mathrm{SO}_2 = 39.500 \times \mathrm{AOD} + 0.136 \times P - 0.426 \times WS + 7.792 \tag{14}$$

Table 3

Statistical summary of regression models developed between AOD and meteorological parameters as independent variables and pollution parameters as a dependent variable in summer

Dependent variable	R	R square	Adjusted R square	Std. error of the estimate	Change Statistics				
					R square change	F change	df1	df2	Sig. F change
$PM_{2.5}$	0.648	0.420	0.411	10.306	0.420	43.485	3	180	0.000
CO	0.734	0.539	0.531	4.893	0.539	70.164	3	180	0.000
O_3	0.560	0.314	0.295	7.180	0.013	3.245	1	178	0.073
NO_2	0.648	0.420	0.403	5.412	0.001	0.313	1	178	0.077
SO_2	0.425	0.181	0.167	4.083	0.181	13.236	3	180	0.000

$$PM_{2.5} = 185.892 \times \mathrm{AOD} - 0.273 \times T + 0.385 \times P + 14.804 \quad (15)$$

$$CO = 124.038 \times \mathrm{AOD} - 0.331 \times T - 0.156 \times P + 7.636 \quad (16)$$

$$O_3 = 123.305 \times \mathrm{AOD} - 0.089 \times RH - 0.167 \times T - 0.257 \times P - 1.334 \times WS + 15.488 \quad (17)$$

$$NO_2 = 88.822 \times \mathrm{AOD} + 0.178 \times RH - 0.088 \times T - 0.314 \times P - 0.312 \times WS + 7.637 \quad (18)$$

$$SO_2 = 39.500 \times \mathrm{AOD} + 0.136 \times P - 0.426 \times WS + 7.792 \quad (19)$$

Table 4

Statistical summary of regression models developed between AOD and meteorological parameters as independent variables and pollution parameters as a dependent variable in autumn

Dependent variable	R	R square	Adjusted R square	Std. error of the estimate	Change Statistics				
					R square change	F change	df1	df2	Sig. F change
$PM_{2.5}$	0.692	0.479	0.462	11.136	0.479	29.678	3	97	0.000
CO	0.833	0.694	0.681	5.106	0.002	0.539	1	96	0.465
O_3	0.637	0.406	0.387	7.730	0.406	22.054	3	97	0.000
NO_2	0.737	0.543	0.519	5.185	0.006	1.321	1	95	0.253
SO_2	0.623	0.389	0.356	4.140	0.017	2.597	1	95	0.110

$$PM_{2.5} = 147.045 \times \mathrm{AOD} + 0.145 \times RH + 2.064 \times WS + 9.777 \quad (20)$$

$$CO = 119.653 \times \mathrm{AOD} + 0.094 \times RH + 0.109 \times T - 0.251 \times P + 6.175 \quad (21)$$

$$O_3 = 94.085 \times \mathrm{AOD} + 0.091 \times RH + 0.424 \times WS + 9.995 \quad (22)$$

$$NO_2 = 90.696 \times \mathrm{AOD} - 0.206 \times RH + 0.189 \times T + 0.133 \times P + 0.73 \times WS + 9.307 \quad (23)$$

$$SO_2 = 63.110 \times \mathrm{AOD} - 0.048 \times RH - 0.248 \times T - 0.237 \times P + 0.817 \times WS + 5.453 \quad (24)$$

approach, show that the parameters entered into the regression model can well explain the changes of the dependent variable. Equations (5) to (9) show the relationships developed for each parameter of air pollution.

Where, *RH* is relative humidity (%), *P* is the amount of precipitation (mm), *T* is surface temperature (degree Celsius), and *WS* denotes the wind speed (m s^{-1}).

As can be seen in Table 2, the regression models are of a R^2 varied from 0.692 for SO_2 gas to 0.901 for CO gas. The values of variations in *F* statistics, which are also less than 0.05, show that the independent parameters entered into the regression model can well explain the changes of the dependent variable. Equations (10) to (14) show the relationships developed for each parameter of air pollution in the spring.

In summer, regression models have shown relatively weak results with R^2 = 0.181 and 0.539 for parameters SO_2 and CO respectively. Accordingly, the spatial distribution of pollution using these models will be relative in summer. The variations of *F* statistics, having values less than 0.05, show that the parameters entered into the regression model can well explain the changes of the dependent variable. Equations (15) to (19) show the relationships developed for each parameter of air pollution in the summer.

Table 4 shows the statistical summary of the models developed for autumn, which indicates that the models have correlations of R^2 = 0.389 to R^2 = 0.694 for SO_2 and CO, respectively. The developed models do not provide good precision for the prediction of SO_2, while they are of an acceptable accuracy for NO_2 and CO. The variations of *F* statistics, which are also less than 0.05 in different stages of the step-by-step approach, show that the parameters entered into the regression model can explain the changes of the dependent variable. Equations (20) to (24) present the relationships developed for each parameter of air pollution in the autumn.

Verification of the regression models

At this stage of the study, the regression models obtained for each parameter in each season were verified with the parameters of the pollution parameters measured in the time period of years 2015 to 2016 in order to allow the regression models to be evaluated for the time interval other than the calibration interval. Table 5 shows the summary of the statistics calculated in the verification stage. According to this table, it can be seen that the best R^2 is calculated in spring (0.83) for $PM_{2.5}$ and its lowest value (0.18) is calculated in the summer and for SO_2. The values of the RMSE statistic are also of the lowest value in spring for SO_2 (3.84), and of the highest value in autumn and for $PM_{2.5}$ (10.9). The negative values of the MBE statistic for most of the parameters in each season show that the developed regression models tend to overestimate each parameter. Only in the case of SO_2 gas and in winter, the value of MBE=1.42 represents the tendency of the developed

Table 5

Statistical summary of the verification of regression models for each season between time period of 2015 to 2016

Season	Department variable	R^2	RMSE	MBE	MAE	df1	df2	Sig. F change
Winter	$PM_{2.5}$	0.53	10.39	0.02	9.06	1	144	0.000
	CO	0.67	4.83	-0.01	4.18	1	144	0.000
	O_3	0.45	7.00	0.04	6.10	1	144	0.005
	NO_2	0.50	5.50	0.03	4.82	1	144	0.040
	SO_2	0.40	4.22	1.42	3.55	1	144	0.028
Spring	$PM_{2.5}$	0.83	9.97	-0.01	8.58	4	179	0.000
	CO	0.80	4.99	-0.57	4.33	2	181	0.000
	O_3	0.75	7.44	0.03	6.50	1	178	0.229
	NO_2	0.82	5.42	-0.01	4.67	3	180	0.000
	SO_2	0.69	3.84	0.02	3.27	3	180	0.000
Summer	$PM_{2.5}$	0.42	10.17	-0.09	8.74	3	180	0.000
	CO	0.53	4.82	-0.02	4.08	3	180	0.000
	O_3	0.31	7.07	0.00	6.06	1	178	0.073
	NO_2	0.41	5.53	0.42	4.69	1	178	0.077
	SO_2	0.18	4.02	0.05	3.49	3	180	0.000
Autumn	$PM_{2.5}$	0.47	10.90	-0.03	9.57	3	97	0.000
	CO	0.69	4.49	-0.20	4.27	1	96	0.465
	O_3	0.40	7.64	-0.99	6.78	3	97	0.000
	NO_2	0.54	5.02	0.01	4.34	1	95	0.253
	SO_2	0.38	4.00	-0.03	3.41	1	95	0.110

regression model for underestimation for this gas. In summer, this statistic with the value equal to zero for O_3 gas indicates that the regression equation developed for this gas has higher accuracy among other equations, in terms of MBE values. The MAE statistics for the SO_2 parameter in spring and for $PM_{2.5}$ in the winter have values of 3.27 and 9.06, respectively, which are the lowest and highest values of this statistic among other parameters and seasons of the year. In general, and according to Table 5, due to having two statistical parameters superior than the four mentioned statistical parameters, the regression model of estimating the SO_2 parameter in spring can be the best developed regression model among the other models.

Spatial Distribution of Pollution Parameters

The spatial and temporal distributions of pollution parameters in Tehran from 2012 to 2016 were plotted for each parameter, separately. Figures 3 to 7 show the spatial distribution of each of the five pollution parameters, ob-

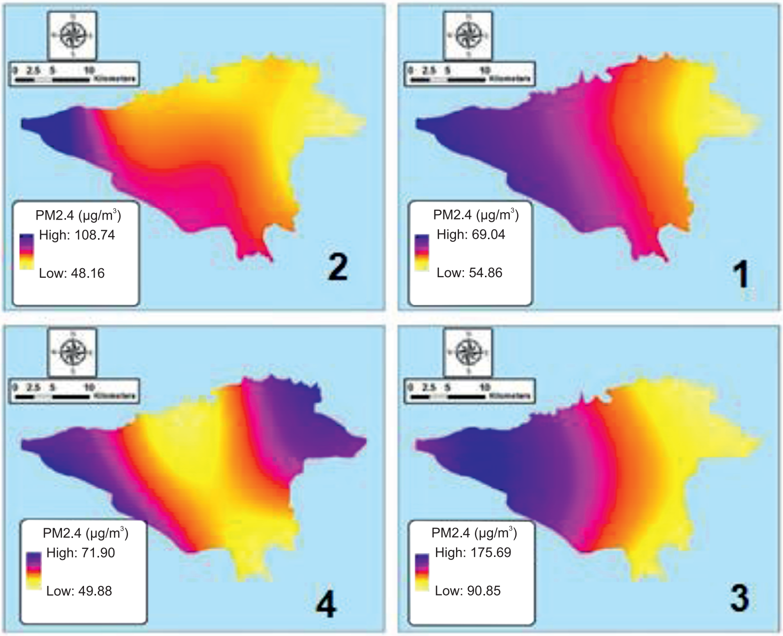

Fig. 3. Spatial distribution of parameter $PM_{2.5}$ in seasons: 1 – winter, 2 – spring, 3 – summer, 4 – autumn

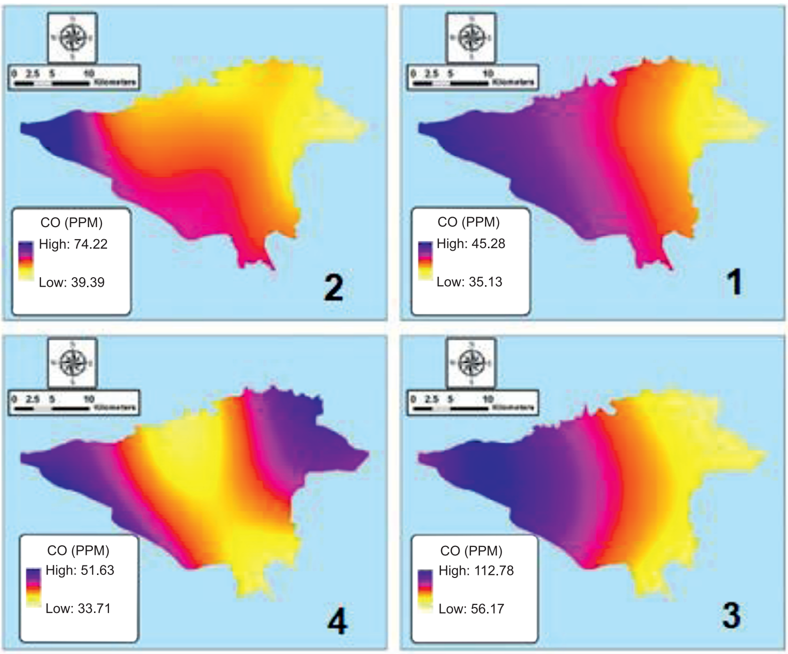

Fig. 4. Spatial distribution of parameter CO in seasons: 1 – winter, 2 – spring, 3 – summer, 4 – autumn

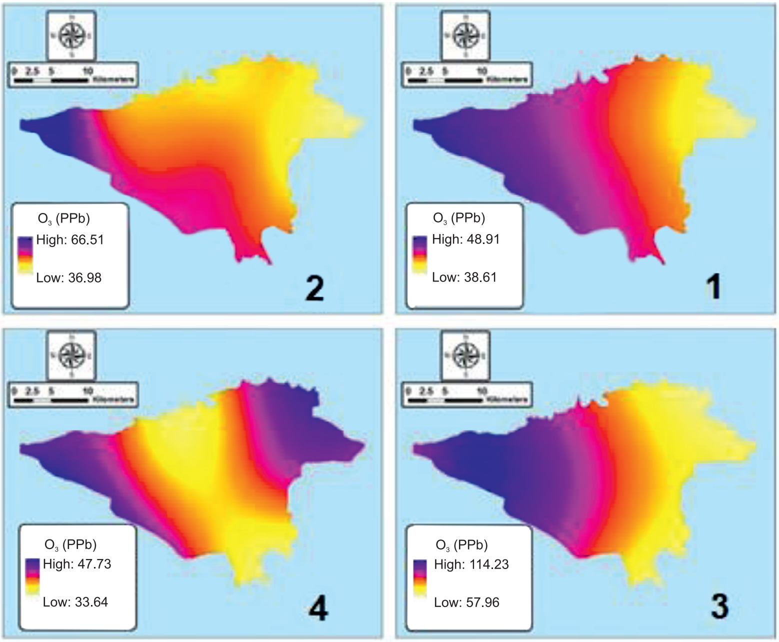

Fig. 5. Spatial distribution of parameter O_3 in seasons: 1 – winter, 2 – spring, 3 – summer, 4 – autumn

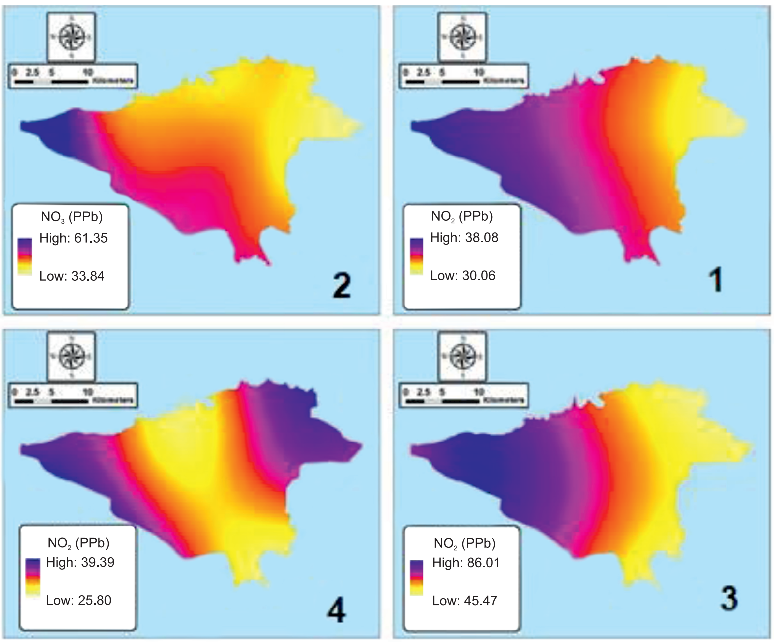

Fig. 6. Spatial distribution of parameter NO_2 in seasons: 1 – winter, 2 – spring, 3 – summer, 4 – autumn

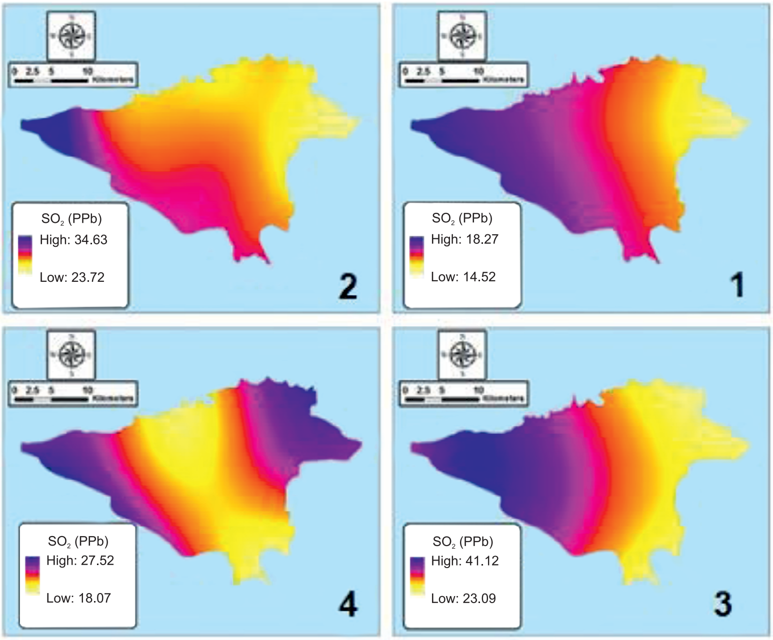

Fig. 7. Spatial distribution of parameter SO_2 in seasons: 1 – winter, 2 – spring, 3 – summer, 4 – autumn

tained by using the regression equations. The spatial distribution of the concentration of $PM_{2.5}$ in summer are illustrated in Figure 3 and show that western regions of Tehran exceed the standard limits (150.4 μg m^{-3}). In the other seasons, the concentration of aerosols does not exceed the standard limit, and winter is of the lowest concentration of these particles. The results indicate that during this time period, the minimum and maximum immission values of carbon monoxide are equal to 33.71 and 112.78 ppm in autumn and summer respectively, hence, Tehran generally presents unhealthy conditions (standard limit = 15.4 ppm). Generally, the highest and lowest concentrations of CO gas occurred in summer and autumn respectively (Figure 4). Based on the available standards that recommended the normal limit of O_3 gas equal to 75 ppb, this pollutant gas occurred excessively in summer and in the western regions of Tehran (Figure 5). In other seasons, the concentration of this gas is in the normal conditions. In the case of NO_2 pollutant, according to the available standards (annual standard value of 21 ppb), the concentration of this pollutant with a minimum and maximum values of 25.80 ppb and 86.01 ppb is also on an unhealthy level in all parts of Tehran city (Figure 6). The SO_2 pollutant is also on an unhealthy level in all seasons of the year, considering the standards available for this gas (7 ppb). The lowest and highest concentrations of this gas are respectively observed in winter (14.52 ppb) and summer (41.12 ppb). In terms of spatial distribution, the concentration of all gases in the western regions are higher than that in the eastern regions. The spatial distribution is continued from the west to the central part of Tehran in summer, and the concentration of pollution is gradually reduced eastwards. In spring, a small part of the western part of the city is highly polluted, and the central and eastern parts are less polluted. Owing to the inclusion of the effects of meteorological factors, the results of this study are more reliable than models that have been developed solely from the linear relationship obtained by the AOD.

The effects of the pollutants on water consumption

For investigating the effects of the mentioned pollutants on water consumption, weekly water consumption data in Tehran city in years 2015 and 2016 were used. To establish a relationship between the volume of water consumed and the amounts of each pollutant, the average weekly pollution parameters were calculated at the measurement stations of Tehran. Table 6

Table 6

Pearson correlation coefficient between weekly water consumption and pollutant parameters

Pollutant parameter	Water use
CO-	0.47
O_3	0.50
NO_2	0.54
SO_2	0.56
$PM_{2.5}$	0.57

illustrates the Pearson correlation coefficient between weekly water consumption and pollutant parameters (CO, O_3, NO_2, SO_2, $PM_{2.5}$). As shown in this table, the lowest correlation, at 0.47, was found for the relationship with CO gas and the highest correlation was observed for $PM_{2.5}$ at 0.57. However, all the values are similar and approximate 0.5. This low correlation can be due to the fact that water consumption is a function of some other factors. For example, on holidays, when traffic in the city is lower and air pollution is relatively decreased, the urban water consumption increases as a result of increased health activities such as washing and bathing. This lack of correlation is more noticeable in long-term holidays such as the holidays of Nowruz celebration. Meanwhile, in this research, due to the lack of daily water consumption data, weekly data were applied, whereas a daily data set may improve the correlation coefficient.

CONCLUSIONS

In this study, several regression relations were developed to estimate the air pollution parameters (CO, O_3, NO_2, SO_2, $PM_{2.5}$) using MODIS-AOD data and meteorological parameters (precipitation, relative humidity, temperature and wind speed). Correlation analysis between different parameters revealed that the highest and lowest correlations between AOD and relative humidity, and wind speed and SO_2 are observed in winter with correlation values of $R^2 = 0.783$ and $R^2 = 0.269$, respectively. In spring, the highest and lowest correlation between AOD and relative humidity, and precipitation and O_3 are seen with correlation values of $R^2 = 0.903$ and $R^2 = 0.632$, respectively. In summer, the highest and lowest correlations between AOD and CO, and wind speed and O_3 are obtained with correlation values equal to $R^2 = 0.723$ and $R^2 = 0.087$, respectively. Also in autumn, the highest and lowest correlations between AOD and CO, and wind and precipitation can be seen with correlation values respectively equal to $R^2 = 0.831$ and $R^2 = 0.263$. The development of regression models in different seasons indicated that these models can well estimate the pollution parameters. In winter, the best result was obtained for the CO parameter with $R^2 = 0.674$ and the lowest correlation was obtained for SO_2 with $R^2 = 0.413$. In spring, the regression models are of high values of R^2 varied from 0.692 for SO_2 gas to 0.901 for CO gas. In summer, regression models showed relatively poor results with $R^2 = 0.181$ for the parameter SO_2 to $R^2 = 0.539$ for the parameter CO. In autumn, the models had correlations of $R^2 = 0.389$ to $R^2 = 0.694$ for SO_2 and CO, respectively. The developed models did not provide good precision for the prediction of SO_2, while having an acceptable accuracy for NO_2 and CO. The verification of the results also showed that the best R^2 is calculated in spring (0.83) for $PM_{2.5}$ and its lowest (0.18) is obtained in summer and for SO_2. The RMSE statistics also yielded the lowest values in spring and for SO_2 (3.84) and the highest value was obtained in autumn and for $PM_{2.5}$ (10.9). The negative values of the MBE statistic for most of the parameters in each season showed that the developed regression models tend to overestimate each parameter. Also, the MAE statistic for the SO_2 parameter in spring and for $PM_{2.5}$ in winter have values of 3.27 and 9.06, respectively, which were the smallest and highest values of this statistic among other parameters and seasons. Altogether, the regression model for the estimation of SO_2 parameter in spring, due to having two statistical parameters superior than the four mentioned statistical parameters, was the best developed regression model. The results of this research indicate that the meteorological variables such as relative humidity and wind speed can increase regression relations between AOD and the parameters of air pollution in Tehran. By having an appropriate global coverage, especially in places without air quality measurement networks, RS data are very useful. A regional approach could be the most appropriate method because of complex chemical and physical parameters that affect the pollution concentration. The results are in agreement with the results of the previous research work (Emili et al. 2010, Emili et al. 2011, Li et al. 2017). In the future, the improvement of RS data adopted to testing the atmosphere, as well as the increased spatial and temporal resolution of data, could lead to more accurate prediction of air pollution parameters based on RS data.

REFERENCES

Beelen R., Raaschou-Nielsen O., Stafoggia M., Andersen ZJ., Weinmayr G., Hoffmann B., Wolf K., Samoli E., Fischer P., Nieuwenhuijsen M., Vineis P., Xun WW., Katsouyanni K., Dimakopoulou K., Oudin A., Forsberg B., Modig L., Havulinna AS., Lanki T., Turunen A., Oftedal B., Nystad W., Nafstad P., De Faire U., Pedersen NL., Östenson CG., Fratiglioni L., Penell J., Korek M., Pershagen G., Eriksen KT., Overvad K., Ellermann T., Eeftens M., Peeters PH., Meliefste K., Wang M., Bueno-De-Mesquita B., Sugiri D., Krämer U., Heinrich J., De Hoogh K., Key T., Peters A., Hampel R., Concin H., Nagel G., Ineichen A., Schaffner E., Probst-Hensch N., Künzli N., Schindler C., Schikowski T., Adam M., Phuleria H., Vilier A., Clavel-Chapelon F., Declercq C., Grioni S., Krogh V., Tsai MY., Ricceri F., Sacerdote C., Galassi C., Migliore E., Ranzi A., Cesaroni G., Badaloni C., Forastiere F., Tamayo I., Amiano P., Dorronsoro M., Katsoulis M., Trichopoulou A., Brunekreef B., Hoek G. 2014. *Effects of long-term exposure to air pollution on natural-cause mortality: An analysis of 22 European cohorts within the multicentre ESCAPE project*. Lancet, 383(9919): 785-795. DOI: 10.1016/S0140-6736(13)62158-3

Brauer M., Amann M., Burnett R. T., Cohen A., Dentener F., Ezzati M., Henderson S., Krzyzanowski M., Martin R. V., Dingenen R. V., Donkelaar A. v., Thurston G. D. 2012. *Exposure assessment for estimation of the global burden of disease attributable to outdoor air pollution*. Environ. Sci. Technol., 46(2): 652-660. *http://doi.org/10.1021/es2025752*

Chu D.A., Kaufman Y.J., Zibordi G., Chern J.D., Mao J., Li C.C., Holben B.N. 2003. *Global monitoring of air pollution over land from the Earth Observing System Terra Moderate Resolution Imaging Spectroradiometer (MODIS)*. J. Geophys. Res.: Atmospheres, 108(D21).

Cohan D.S., Chen R. 2014. *Modeled and observed fine particulate matter reductions from state attainment demonstrations*. J. Air Waste Manage. Assoc., 64(9): 995-1002.

Dhyani R., Sharma N., Maity A.K. 2017. *Prediction of PM 2.5 along urban highway corridor under mixed traffic conditions using CALINE4 model*. J. Environ. Manage., 198: 24-32.

Emili E., Popp C., Petitta M., Riffler M., Wunderle S., Zebisch M. 2010. *PM_{10} remote 12 sensing from geostationary SEVIRI and polar-orbiting MODIS sensors over the complex 13 terrain of the European Alpine region*. Remote Sens. Environ., 114(11): 2485-2499.

Emili E., Lyapustin A., Wang Y., Popp C., Korkin S., Zebisch M., Wunderle S., Petitta M. 2011. *15 M.: High spatial resolution aerosol retrieval with MAIAC: Application to mountain regions*. J. Geophys. Res.: Atmos., 116, D23211. DOI: 10.1029/2011JD016297, 17.

Engel-Cox J.A., Holloman C.H., Coutant B.W., Hoff R.M. 2004. *Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality*. Atmos. Environ., 38(16): 2495-2509.

Franklin M., Kalashnikova O.V., Garay M.J. 2017. *Size-resolved particulate matter concentrations derived from 4.4 km-resolution size-fractionated Multi-angle Imaging SpectroRadiometer (MISR) aerosol optical depth over Southern California*. Remote Sens. Environ., 196: 312-323.

Gupta P., Christopher S. A., Wang J., Gehrig R., Lee Y., Kumar N. 2006. *Satellite remote sensing of particulate matter and air quality assessment over global cities*. Atmos. Environ., 40(30): 5880-5892. DOI: 10.1016/j.atmosenv.2006.03.016

Gupta P., Christopher S.A. 2009. *Particulate matter air quality assessment using integrated* *surface, satellite, and meteorological products: Multiple regression approach*. J. Geophys. Res.: Atmos., 114(D14).

He Q., Zhang M., Huang B., Tong X. 2017. *MODIS 3 km and 10 km aerosol optical depth for China: Evaluation and comparison*. Atmos. Environ., 153: 150-162.

Jathar S.H., Gordon T.D., Hennigan C.J., Pye H.O.T., Pouliot G.A., Adams P.J., Donahue N.M., Robinson, A.L. 2014. *Unspeciated organic emissions from combustion sources and their influence on the secondary organic aerosol budget in the United States*. Proc. Nat. Acad. Sci., 111(29): 10473-10478.

Jin Q., Fang X., Wen B., Shan A. 2017. *Spatio-temporal variations of PM2. 5 emission in China from 2005 to 2014*. Chemosphere, 183: 429-436.

Kelly J.T., Reff A., Gantt B. 2017. *A method to predict PM 2.5 resulting from compliance with national ambient air quality standards*. Atmos. Environ., 162: 1-10.

Kuwayama T., Collier S., Forestieri S., Brady J.M., Bertram T.H., Cappa C.D., Zhang Q., Kleeman M.J. 2015. *Volatility of primary organic aerosol emitted from light duty gasoline vehicles*. Environ. Sci. Technol., 49(3): 1569-1577. DOI: 10.1021/es504009w

Lee E., Županski M., Županski D., Park S.K. 2017. *Impact of the OMI aerosol optical depth on analysis increments through coupled meteorology – aerosol data assimilation for an Asian dust storm*. Remote Sens. Environ., 193: 38-53.

Li Z., Che W., Frey H.C., Lau A.K., Lin C. 2017. *Characterization of PM 2.5 exposure concentration in transport microenvironments using portable monitors*. Environ. Pollut., 228: 433-442.

Liu Y., Park R.J., Jacob D.J., Li Q., Kilaru V., Sarnat J.A. 2004. *Mapping annual mean ground level $PM_{2.5}$* concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. J. Geophys. Res., 109, D22206. DOI:10.1029/2004JD005025.

Liu Y., Sarnat J.A., Kilaru V., Jacob D.J., Koutrakis P. 2005. *Estimating ground-level PM2. 5 in the eastern United States using satellite remote sensing*. Environ. Sci. Technol., 39(9): 3269-3278.

Liu W., Li X., Chen Z., Zeng G., León T., Liang J., Huang G., Gao Z., Jiao S., He X., Lai M. 2015. *Land use regression models coupled with meteorology to model spatial and temporal variability of NO 2 and PM 10 in Changsha*, China. Atmos. Enviro., 116: 272-280.

Liu Y., Hong Y., Fan Q., Wang X., Chan P., Chen X., Lai A., Wang M., Chen X. 2017. *Source--receptor relationships for PM 2.5 during typical pollution episodes in the Pearl River Delta city cluster, China*. Sci. Total Environ., 596: 194-206.

Raoufi S.S., Goharnejad H., Zakeri Niri M. 2018. *Air pollution effects on climate and air temperature of Tehran city using remote sensing data,* Asian J. Water Environ. Pollut., 15(2): (2018): 1-9. DOI: 10.3233/AJW-180012

Tehran Air Quality Control Company. 2016. Tehran Annual Air Quality Report, 265 p., http://air.tehran.ir/

Tang W., Yang K., Qin J., Niu X., Lin Ch., Jing X. 2017. *A revisit to decadal change of aerosol optical depth and its impact on global radiation over China*. Atmos. Environ., 150: 106-115.

Wang A., Song Q., Tu G., Wang H., Yue Y., Yao Q. 2014. *Influence of flue gas cleaning system on characteristics of $PM_{2.5}$ emission from coal-fired power plants*. Int J Coal Sci Technol., 1: 4. https://doi.org/10.1007/s40789-014-0001-x

Wang J., Christopher S.A. 2003. *Intercomparison between satellite derived aerosol optical thickness and PM2. 5 mass: implications for air quality studies*. Geophys. Res. Letters, 30(21).

Wong C.M., Lai H.K., Tsang H., Thach T.Q., Thomas G.N., Lam K.B., Chan K.P., Yang L., Lau A.K., Ayres J.G., Lee S.Y., Chan W.M., Hedley A.J., Lam T.H. 2015. *Satellite-based estimates of long-term exposure to fine particles and association with mortality in elderly Hong Kong residents*. Environ Health Perspect, 123: 1167-1172. http://dx.doi.org/10.1289/ehp.1408264