

RESEARCH ARTICLE

Short-term forecasting of air pollution index in Belgrade, Serbia

Mirela Đurić | Dragana Vujović 

Faculty of Physics, Department of
Meteorology, University of Belgrade,
Belgrade, Serbia

Correspondence

Dragana Vujović, Faculty of Physics,
Department of Meteorology, University of
Belgrade, Dobračina 16, 11000 Belgrade,
Serbia.

Email: dvujovic@ff.bg.ac.rs

Abstract

Forecasting air pollution in big cities is of great importance, and there are various types of air pollution indices (APIs). In the present study, the level of air pollution in winter 2011 at four point-locations in Belgrade, Serbia, was measured using wind speed data and non-standard and standard air temperature data, measured at 10 and 2 m, respectively. Using multiple linear regression (MLR) analysis, equations to forecast the API were obtained. This forecast was verified using data from the winter of 2012/2013. The results obtained are well aligned with the monitored API and verified by the root mean square error (RMSE). It is shown that standard meteorological measurements representative of the city can accurately predict the API at individual point-locations as well as using temperature and wind speed measured at each respective location. Three locations, which measured SO₂, NO₂, PM₁₀, O₃ and CO, showed poor air quality for > 78% of days observed. For the fourth location, the final estimate of prevailing air quality could not be calculated due to the absence of PM₁₀ measurements. Forecasting the API on a short-term scale can be of great help for long-term air-quality improvements.

KEYWORDS

air pollution forecast, air pollution index, air quality, air temperature, meteorology, wind speed

1 | INTRODUCTION

Nowadays more than half the human population lives in urban areas. According to the 2018 Revision of the *World Urbanization Prospects* (UN, 2018), 4.2 billion people live in urban areas, with an expectation of 68% of all mankind living in cities by 2050. Consequently, the number of vehicles in cities will increase, as well as the concentration of air pollutants. This means that the unfavourable effects of air pollution will impact an increasingly large portion of the human population, causing multiple health problems, such as cardiovascular disease, lung

cancer and acute respiratory infections (Kampa and Castanas, 2008). Anderson (2009) correlated an episode of heavy air pollution in London in 1952 with excess deaths. Within three weeks of this episode, there were more than 4,000 deaths. Because influenza was not prevalent during that period, it was concluded that the cause of these deaths was due to air pollution. Gurjar *et al.* (2010) evaluated the health risks in megacities due to air pollution and concluded that South Asian megacities have high health risks due to high concentrations of total suspended particulates, which are more damaging to health than gaseous pollutants. Kampa and

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. Meteorological Applications published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.

Castanas (2008, 362) stated that “short- and long-term exposures have also been linked with premature mortality and reduced life expectancy”. The World Health Organization (WHO) concluded that air pollution causes disease and more than 2 million premature deaths (WHO, 2006) each year. Hence, many cities monitor air quality at several locations, particularly near crowded traffic lines.

Numerous studies have examined the relationship between urban air pollution and meteorological variables (Yadav *et al.*, 2016; Squizzato *et al.*, 2017; Zhang *et al.*, 2017a) and several air-quality indices are in use (Stieb *et al.*, 2012). Their common objective is to facilitate the interpretation of air quality and to help the general public adjust to changes in air quality. Plaia and Ruggieri (2011) reviewed several air pollution indices (APIs). Cogliani (2001) developed an API for the purpose of forecasting air pollution for the next day in three cities in Italy. Deniz Genc *et al.* (2010) used and calculated the APIs for 10 stations in Ankara, Turkey. Kumar and Goyal (2011) used a principal component regression technique for short-term daily forecasts of API in Delhi, India.

However, there is not a lot of this kind of research in Serbia. Lazić *et al.* (2016) conducted one study for the most populated city in Serbia: Belgrade. They used the Operational Street Pollution Model (Berkowicz, 2000) to measure the air pollution within the urban city canyon. However, their study was limited to a 10 week period in summer when pollution is generally less intense as compared with the cold part of the year. Vujović and Todorović (2017) assessed emissions of different pollutants released by air traffic along with weather conditions over a period of eight years. They concluded that “the number of days with poor air quality increased during the cold part of year, especially during anticyclones when temperature inversions form near the surface” and wind speeds are low (p. 93); consequently, emitted pollutants are not adequately dispersed. During the cold part of the year, high pressure, light winds, temperature inversions up to 700 or 1,000 m in height, fog and drizzle are common conditions in Belgrade. These are the most unfavourable weather conditions from the standpoint of local air quality. This weather situation can last for days, bringing warnings of excessive air pollution and providing the motivation for this research.

Both emissions of pollutants and meteorological factors determine air quality at point-locations in a city. The goal of this research was to find an appropriate way to forecast the level of air pollution in Belgrade on a short-term scale. In that way, city authorities would have enough time to act regarding the improvement of air

quality at locations of concern in the city, and consequently protect the health of the citizens.

2 | MATERIALS AND METHODS

2.1 | Data source

Belgrade is the capital of Serbia. According to the Census Atlas (2014), Belgrade has 1,659,440 inhabitants (513 inhabitants·km⁻²). In the last 30 years, it has experienced rapid urbanization. As a result of the increase in the number of inhabitants, the number of motor vehicles more than doubled from 292,449 in 1985 to 595,053 in 2014 (source: Institute for Informatics and Statistics of the City of Belgrade). The major sources of air pollution in Belgrade are electricity production, heating and industry; for oxides of nitrogen, the major source is traffic (EIS, 2012, 2013).

The Environmental Protection Agency of Serbia, a division of the Ministry of Environmental Protection of the Republic of Serbia, provided the data for air pollution used in the present study. Data were measured at four automatic stations situated in residential areas in Belgrade: Novi Beograd, Stari Grad, Mostar and Vračar (Figure 1). These four stations fall under the local network of stations for air-quality monitoring in Belgrade. The locations of the stations are shown in Figure 1. Stations Mostar and Vračar are located near crowded motorways. Stari Grad is in the old city centre with a high number of individual heating system units (as well as Vračar), while Novi Beograd is in the new part of the city, with high buildings, a flat terrain and heavy traffic. Table 1 displays the coordinates, latitudes and longitudes as well as the elevations of the stations.

All stations measured hourly concentrations of sulphur dioxide (SO₂), nitrogen dioxide (NO₂), coarse particulate matter (PM₁₀), ozone (O₃) and carbon monoxide (CO), with the exception of Vračar station, which did not measure PM₁₀ concentrations. An expert team from the Environmental Protection Agency conducted a logical control of the data. All the stations provided meteorological measurements as well as hourly values of air temperature (non-standard measurements) and wind speed, both at 10 m height.

However, we had and the standard meteorological measurements of hourly values of air temperature at 2 m height as well as wind speed at 10 m height. The source of these meteorological data was Belgrade Meteorological Observatory (BMO) at Karađorđev Park, which falls under the network of synoptic stations of the Republic Hydrometeorological Service of Serbia, which is by its

FIGURE 1 Locations of the automatic stations for air quality monitoring in Belgrade, Serbia



TABLE 1 Description of four monitoring stations in Belgrade

Station	Type	Measured pollutants	Observation	Longitude	Latitude	Elevation (masl)
Novi Beograd	Residential	SO ₂ , NO ₂ , PM ₁₀ , O ₃ , CO	January–March, October–December 2011; winter 2012/2013	44 ° 48 ' 11 " E	20 ° 24 ' 0 " N	74
Stari Grad	Residential	SO ₂ , NO ₂ , PM ₁₀ , O ₃ , CO	January–March, October–December 2011; winter 2012/2013	44 ° 49 ' 16 " E	20 ° 27 ' 32 " N	103
Mostar	Urban traffic	SO ₂ , NO ₂ , PM ₁₀ , O ₃ , CO	January–March, October–December 2011; winter 2012/2013	44 ° 47 ' 55 " E	20 ° 27 ' 0 " N	83
Vračar	Urban traffic	SO ₂ , NO ₂ , O ₃ , CO	March, October–December 2011; winter 2012/2013	44 ° 47' 49 " E	20 ° 28 ' 31 " N	141

status a location representative of Belgrade city. The location of BMO is also shown in Figure 1.

The analyser GRIMM EDM 180 aerosol spectrometer was used to measure PM₁₀, while the Teledyne analysers measured the other pollutants. The Teledyne API model 100E gas analyser measured SO₂ concentrations, which were determined by measuring the fluorescence of SO₂ at a wavelength of 330 nm. The Teledyne API Model 200A gas analyser measured the concentration of nitrogen oxides using the chemiluminescence detection of NO₂. The model T300M CO analysers measured the concentration of CO using non-dispersive infrared spectroscopy. The API model 400 A ozone analyser was used to compare the light absorption of the sample air with the light absorption of the sample scrubbed to remove the O₃. The concentration of O₃ was then calculated using the Beer–Lambert Law.

The period selected for all stations was winter because of its high air pollution. The two different halves of

winter were studied: January–March and October–December 2011. Due to inoperative equipment at Mostar station, the first winter was limited to March. For verification of the obtained results, data from the winter 2012/2013 were used.

2.2 | Calculating the air pollution index (API)

On a daily basis, the API of a single station is derived from hourly measured concentrations of air pollutants. Air Protection Law Regulation (2010) prescribes daily threshold and tolerance levels for SO₂, NO₂, PM₁₀ and CO concentrations. Law Order (2006) prescribes the daily threshold level of O₃. Its daily tolerance level is considered to be equal to the threshold. The tolerance level is a threshold raised by the tolerance threshold, that is, the percentage of excess allowed above the threshold under

TABLE 2 Daily threshold and tolerance levels of the measured air pollutants

Air quality	SO ₂ (µg·kg ⁻¹)	NO ₂ (µg·kg ⁻¹)	PM ₁₀ (µg·kg ⁻¹)	CO (mg·kg ⁻¹)	O ₃ (µg·kg ⁻¹)
Good	< 125	< 85	< 50	< 85	< 5
Acceptable		85–125	50–75		5–10
Poor	> 125	> 125	> 75	> 85	> 10

Note: Air Protection Law Regulation (2010).

prescribed conditions. The threshold is the highest concentration allowed for the air pollutant in order to avoid harmful effects on human health and the environment. According to pollutant thresholds and tolerances, air quality is defined as good (clean or slightly polluted air), acceptable (moderately polluted) or poor (over-polluted air) (Table 2). Air quality is considered poor if only one pollutant's concentration exceeds its tolerance at any given station.

The concentrations of five air pollutants were measured at Novi Beograd, Stari Grad and Mostar stations. Following Cogliani (2001), if the concentration of a single pollutant was low, the score was 1; if the concentration was acceptable, the score was 6; and if the concentration was high, the score was 31. Numerical scores of one, six and 31 were selected to avoid superposing of the scores. If the concentration was missing, the score was 0. According to the criteria presented in Table 2, a capital letter indicates the air quality at a given station with “G” for good air quality, “A” for acceptable and “P” for poor. For all five pollutants: 5 was good air quality, 10–30 were acceptable air quality and 35–155 were poor air quality.

Subsequently, new numerical values were attributed to each station to represent the station's daily API. When the air quality score was good, the station index was attributed the value of 0. For an acceptable air quality score, the station index was attributed the value of 1. When the air quality score was poor, the station index was attributed the value of 2.

The method was the same for Vračar station at which four air pollutant concentrations were measured. PM₁₀ data were not collected at this site, so there is a slight difference in the total score compared with the other stations. Because of this, a score of 4 represents good air quality, 9–24 is acceptable and 34–124 when air quality was poor.

Each API calculated with this method was dependent on the number of air pollutants measured. If the number of air pollutants differed from station to station, the stations' APIs were not comparable. This means that Vračar's API could not be compared with the indices of the other three stations due to the absence of PM₁₀ measurements at Vračar.

2.3 | Selection of the correlation variables

The selection of correlation variables was made according to the method, analysis and results provided by Cogliani (2001) and Genc *et al.* (2010), as well as data from the local network for automatic air quality monitoring and BMO. Three independent variables were selected: diurnal temperature range (ΔT), daily average wind speed (V_{av}) and the previous day's API (I_{d-1}). ΔT is the difference between daily maximum and minimum temperatures. Although the relation between ΔT and human health and mortality is well documented (Ge *et al.*, 2013; Zhang *et al.*, 2017b), the present understanding of the relation between ΔT and air pollution can be improved. The API derived from the previous day's concentrations is justifiably used as the basis for estimations of the current day's air quality (Cogliani, 2001; Deniz Genc *et al.*, 2010; Kumar and Goyal, 2011), and it was shown to be a reasonably good replacement for motor vehicle traffic intensity, which produced the air pollution (Cogliani, 2001). A station's daily API is a dependent variable calculated from measured or forecasted independent variables using multiple linear regression (MLR) analysis. Relative humidity and total rainfall did not show statistically significant correlations with daily API, with exception of relative humidity for Vračar (significant at the 0.01 level).

2.3.1 | Multiple linear regression (MLR)

The MLR is a common variation of linear regression (Wilks, 2006) in which there is a single predictand and more than one predictor (in the present case, ΔT , V_{av} and I_{d-1}). Each predictor has its own coefficient. The general form of the prediction equation with regression coefficients illustrates the relation between the calculated station's daily API (I_c) and other variables:

$$I_c = a \Delta T + b V_{av} + c I_{d-1} + d \quad (1)$$

where I_{d-1} is the daily API obtained from the previous day's concentrations of pollutants; ΔT is a diurnal

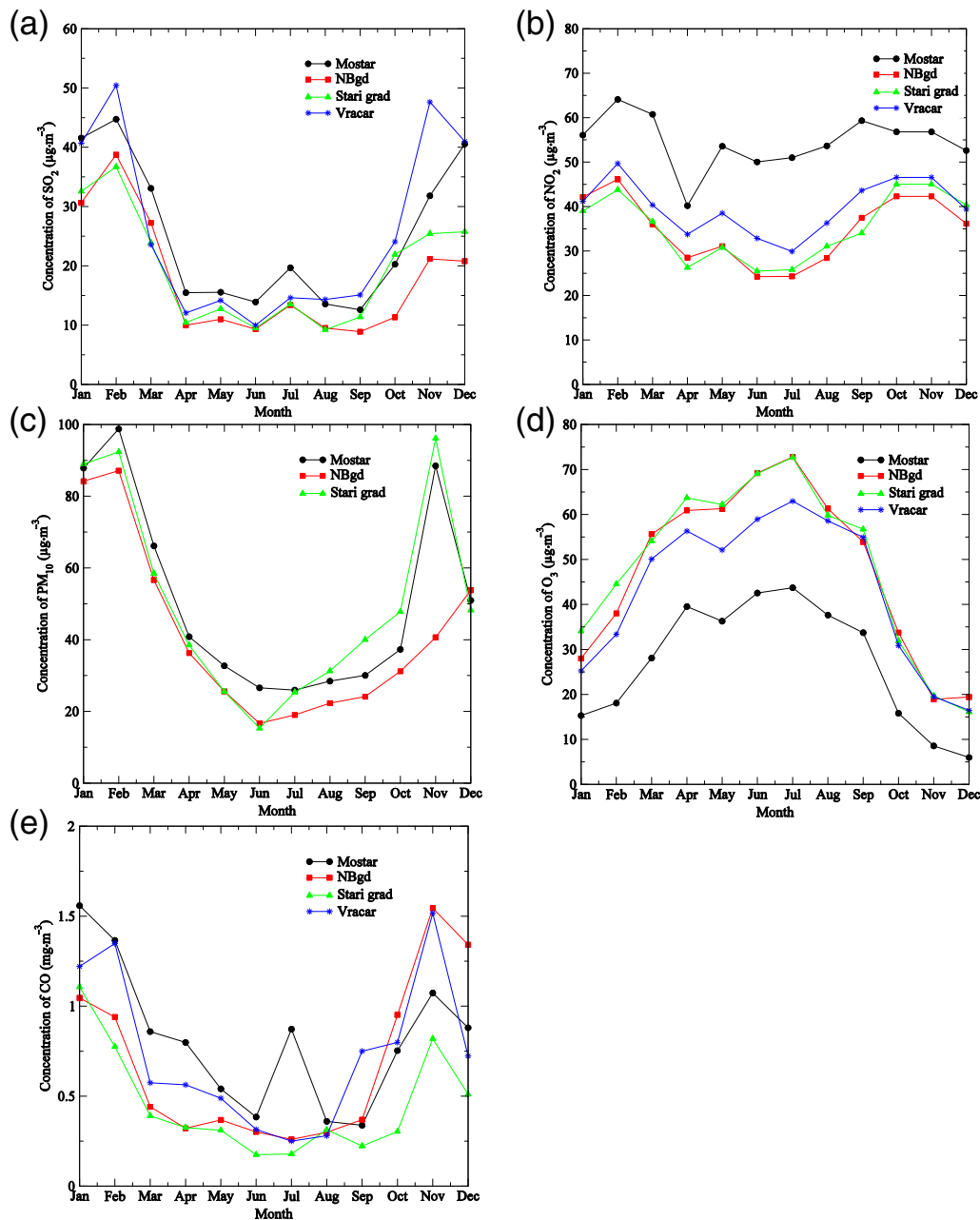


FIGURE 2 Average monthly concentrations of measured air pollutants for 2011: (a) SO_2 , (b) NO_2 , (c) PM_{10} , (d) O_3 and (e) CO at the Mostar, Novi Beograd, Stari Grad and Vračar stations

temperature range ($^{\circ}\text{C}$); and V_{av} is daily average wind speed (m s^{-1}). The regression constant (or intercept) is denoted by d . The regression coefficients a , b , c , and the regression constant are often called the regression parameters. They are calculated by minimizing the sum of squared residuals. For this, the statistical software SPSS (Statistical Package for the Social Sciences), Version IBM SPSS was used. To implement the MLR, the residuals should be normally distributed, the independent variables should not be highly correlated with each other (no multicollinearity) and variance of error terms

should be equally distributed. In order to check whether the residuals of the regression follow a normal distribution, a normal predicted probability plots was examined. A condition index > 15 can indicate multicollinearity problem, while > 30 is a very strong sign for problems with multicollinearity.

2.3.2 | Root mean square error (RMSE)

An API forecast was assessed by determining the mean square error. Forecasted values deviate from observed

$SO_{2av} (\mu g \cdot m^{-3})$	Novi Beograd	Stari Grad	Mostar	Vračar
Mean	17.6654	19.4222	25.2211	25.6318
Median	12.3615	17.7311	19.9362	19.331
Standard deviation	10.0085	9.5399	12.2725	15.0475
Sample variance	100.17	91.01	150.6136	226.4284
Minimum	8.8899	9.2163	12.5915	9.9658
Maximum	38.7426	36.6874	44.7371	50.4474
W/S ratio	4.36	3.98	3.55	5.06
$NO_{2av} (\mu g \cdot m^{-3})$				
Mean	34.9165	35.2674	53.8133	39.9332
Median	36.0423	35.3263	54.8318	40.3472
Standard deviation	7.5097	7.4474	8.2250	6.3424
Sample variance	56.3952	55.4636	67.6502	40.2264
Minimum	24.2344	25.4637	31.1551	29.9063
Maximum	46.1327	45.0174	64.0899	49.6776
W/S ratio	1.90	1.77	2.06	1.66
$PM_{10av} (\mu g \cdot m^{-3})$				
Mean	41.4680	50.6589	51.1623	
Median	33.7361	43.9150	39.0666	
Standard deviation	24.2359	27.8120	27.0814	
Sample variance	587.3801	773.5078	733.4044	
Minimum	16.7055	15.3077	25.9480	
Maximum	87.1219	96.1473	98.7640	
W/S ratio	5.21	6.28	3.81	
$O_{3av} (\mu g \cdot m^{-3})$				
Mean	47.7585	48.7141	27.0742	43.2783
Median	54.7722	55.4055	30.8182	51.0956
Standard deviation	19.1704	19.1433	13.6314	16.9479
Sample variance	367.5042	366.4661	185.8163	287.2301
Minimum	18.9107	16.1043	5.9841	16.4653
Maximum	72.7973	72.6932	43.7243	62.9677
W/S ratio	3.85	4.51	7.31	3.82
$CO_{av} (mg \cdot m^{-3})$				
Mean	0.6819	0.4527	0.8067	0.7357
Median	0.4051	0.3192	0.8295	0.6485
Standard deviation	0.4578	0.2950	0.3905	0.4224
Sample variance	0.2095	0.0870	0.1525	0.1784
Minimum	0.2601	0.1746	0.3173	0.2503
Maximum	1.5448	1.1074	1.5578	1.5166
W/S ratio	5.94	6.34	4.91	6.06

Note: W/S ratio = winter-to-summer ratio.

TABLE 3 Descriptive statistics of monthly averaged concentration of pollutants

APIs and the difference between them is residual. The lower the RMSE, the better the API forecast. Thus, the RMSE is a good measure of how accurately the model predicts the variable.

2.3.3 | Pearson correlation

To find the degree of association between two variables, the Pearson product-moment coefficient of linear

TABLE 4 Correlation coefficients and statistical significance for all point-locations for: (a) January–March 2011 (for Mostar station only March 2011); and (b) October–December 2011

	I_c		I_c		I_c		I_c	
	Novi Beograd	<i>p</i> -value	Stari Grad	<i>p</i> -value	Mostar	<i>p</i> -value	Vračar	<i>p</i> -value
(a)								
I_c	1.000		1.000		1.000		1.000	
I_{d-1}	0.419	0.000	0.439	0.000	0.291	0.060	0.513	0.000
V_{av}	−0.320	0.001	−0.301	0.002	−0.600	0.000	−0.151	0.078
ΔT	0.343	0.001	0.272	0.005	0.517	0.002	0.526	0.000
(b)								
I_c	1.000		1.000		1.000		1.000	
I_{d-1}	0.518	0.000	0.519	0.000	0.490	0.000	0.510	0.000
V_{av}	−0.514	0.000	0.288	0.000	−0.358	0.000	−0.314	0.001
ΔT	0.377	0.000	−0.448	0.003	0.190	0.035	0.546	0.000

Note: I_c : calculated air pollution index (API); I_{d-1} : previous day's API; V_{av} : daily average wind speed; and ΔT : diurnal temperature range.

TABLE 5 Point-location air quality percentages

Novi Beograd	January–March 2011	October–December 2011
Good	6.7	39.1
Acceptable	14.4	10.9
Poor	78.9	50
<i>Stari Grad</i>		
Good	5.6	16.3
Acceptable	10	18.5
Poor	84.4	65.2
<i>Mostar</i>		
Good	6.5	15.2
Acceptable	12.9	19.6
Poor	80.6	65.2
<i>Vračar</i>		
Good	51.1	58.7
Acceptable	16.7	16.3
Poor	32.2	25

correlation, or Pearson correlation, was used. It is between -1 (a perfect negative linear relationship between variables) and 1 (a perfect positive linear relationship). What is often lost sight of is that the correlation coefficient does not provide any explanation about the relationship between variables, at least in some physical or causative sense (Wilks, 2006). If the correlation coefficient is > 0.5 , the correlation is good.

3 | RESULTS AND DISCUSSION

3.1 | Seasonal variations in pollutant concentrations

The analysis began by observing the behaviour of air pollutants throughout the year. Average monthly concentrations obtained from hourly concentrations measured at four automatic air pollution monitoring stations in Belgrade are displayed in Figure 2, with descriptive statistics given in Table 3. Observed values for pollutants SO_2 , PM_{10} and CO have variations, ranging from summer lows to winter highs. The opposite is the case for variations in O_3 , which range from winter lows to summer highs. As for NO_2 , there are no such specific seasonal differentiations. These seasonal variations are in a tight relation with meteorological variables as well as with the number of motor vehicles and combustion heating systems in operation. The winter heating season in Belgrade occurs between mid-October and mid-April. Heating systems are considered operationally inactive between mid-April and mid-October. Therefore, winters are characterized by greater emissions of air pollutants and specific meteorological conditions such as frequent temperature inversions, lower mixing height and anticyclonic atmospheric circulation. Air pollution concentration easily reaches peak values under conditions of increased emissions and inadequate surface level dispersion.

SO_2 is a combustion product of fossil fuel. Its winter-to-summer average monthly concentration ratios (W/S

TABLE 6 Regression parameters for the equation $I_c = a \Delta T + b V_{av} + c I_{-1} + d$ for all stations

Temperature and wind measurement		<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>R</i> ²
<i>Novi Beograd</i>						
January–March 2011	AS	0.024	−0.163	0.329	1.282	0.250
October–December 2011		0.051	−0.589	0.364	1.260	0.457
January–March 2011	BMO	0.036	−0.085	0.339	1.129	0.286
October–December 2011		0.061	−0.155	0.436	0.573	0.405
<i>Stari Grad</i>						
January–March 2011	AS	0.022	−0.288	0.364	1.317	0.261
October–December 2011		0.033	−0.521	0.401	1.140	0.369
January–March 2011	BMO	0.023	−0.097	0.376	1.219	0.296
October–December 2011		0.041	−0.204	0.337	1.151	0.442
<i>Mostar</i>						
March 2011	AS	0.040	−0.491	0.197	2.100	0.523
October–December 2011		0.014	−0.114	0.406	0.995	0.252
March 2011	BMO	0.015	−0.205	0.066	2.045	0.185
October–December 2011		0.033	−0.210	0.352	1.214	0.424
<i>Vračar</i>						
January–March 2011	AS	0.097	−0.106	0.348	0.249	0.386
October–December 2011		0.108	−0.160	0.336	0.020	0.461
January–March 2011	BMO	0.081	−0.088	0.359	0.301	0.402
October–December 2011		0.110	−0.112	0.291	0.065	0.507

Note: AS: automatic station; and BMO: Belgrade Meteorological Observatory.

	<i>I</i> _{c_Vračar} January–March (only March for Mostar)	<i>I</i> _{c_automat}	<i>I</i> _{c_Vračar} October–December	<i>I</i> _{c_automat}
Novi Beograd	0.469971	0.487065	0.681850	0.654319
Stari Grad	0.444235	0.455050	0.567146	0.602803
Mostar	0.462669	0.509612	0.564980	0.600544

TABLE 7 Root mean square error (RMSE) for the forecasted air pollution index (API) for January–March and October–December 2011

Note: *I*_{c_Vračar}: API calculated using concentrations measured at automatic stations and meteorological data from Belgrade Meteorological Observatory (BMO) at Vračar; *I*_{c_automat}: API calculated using both concentrations and meteorological data measured at automatic stations.

ratio, the maximum concentration divided by the minimum concentration), ranged from 3.5 to 5.1. All *W/S* ratios are given in Table 3. Higher winter concentrations are due to an increase in the number of active heating units. Lower summer concentrations are due to the inactivity of heating units and more favourable meteorological conditions with less frequent temperature inversions. A *W/S* ratio of 5.1 was obtained at Vračar station located in a part of the city where the number of individual households with heating system units is large compared with parts of the city heated by city heating plants. The smallest *W/S* ratio of 3.5 was obtained at Mostar station (Figure 2a).

NO₂ is mostly attributed to motor vehicle exhausts. The *W/S* ratios of average monthly concentrations ranged from 1.6 to 2.1. As can be seen in Figure 2b, there are no significantly differentiated winter highs or summer lows. Still, higher concentrations measured during winter are related to specific meteorological conditions. The highest NO₂ concentration was at Mostar station due to its location near the highroad.

PM₁₀ is a primary air pollutant considered to have higher concentrations during winter due to home heating and industrial enterprise. Average monthly *W/S* ratios ranged from 3.8 (Mostar) to 6.3 (Stari Grad). Vračar station lacks PM₁₀ measurements. Three other station's

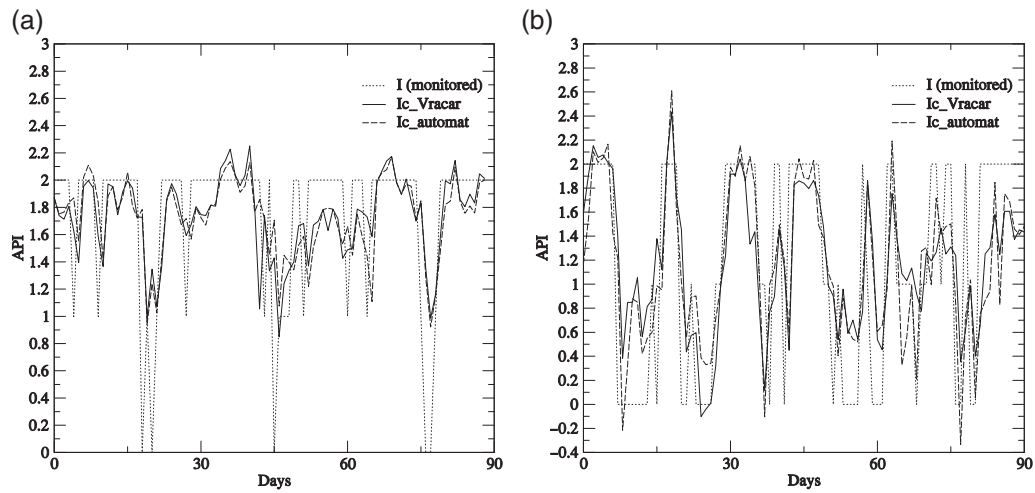


FIGURE 3 Novi Beograd: (a) January–March and (b) October–December 2011

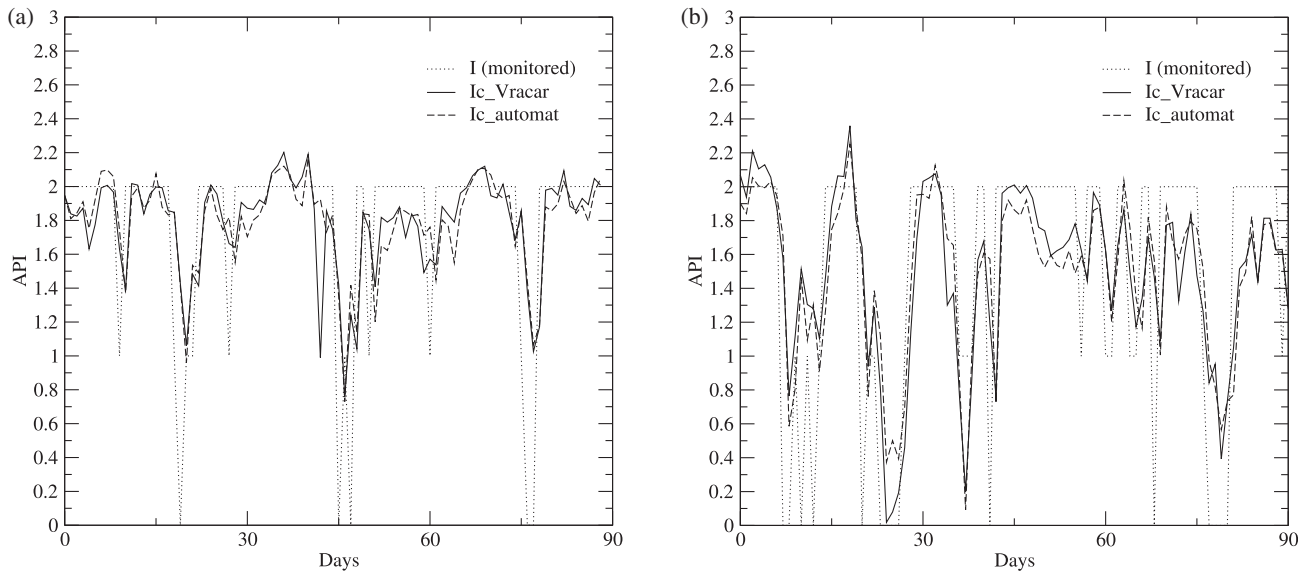


FIGURE 4 Stari Grad: (a) January–March and (b) October–December 2011

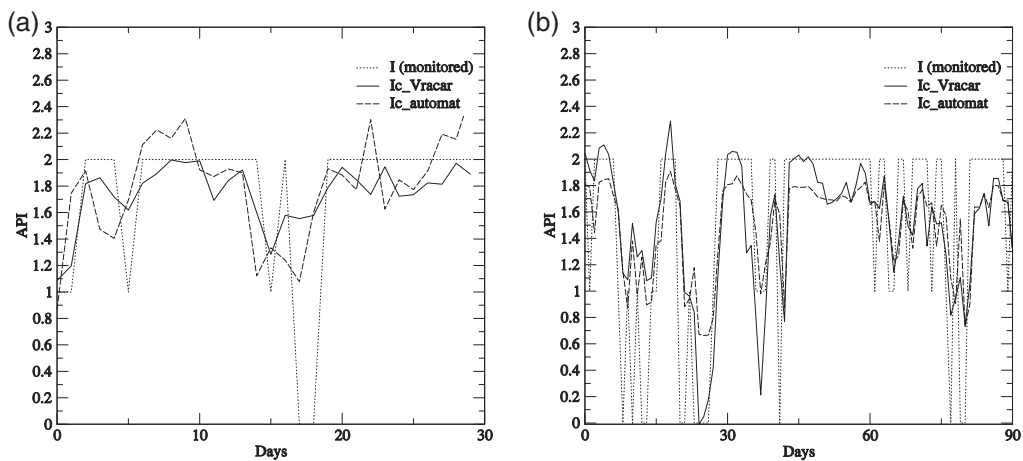


FIGURE 5 Mostar: (a) January–March and (b) October–December 2011

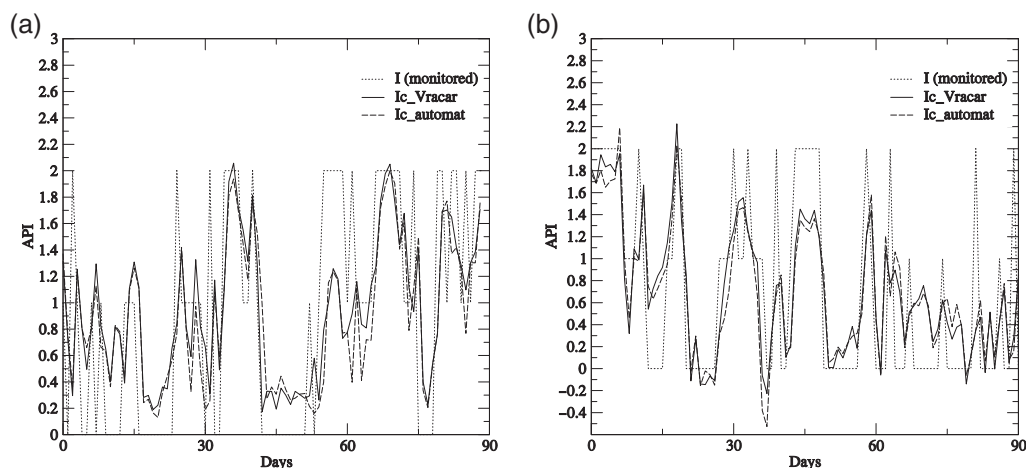


FIGURE 6 Vračar: (a) January–March and (b) October–December 2011

measurements show well-differentiated winter highs and summer lows (Figure 2c).

Surface O_3 is not a primary air pollutant. It is considered a secondary photochemical air pollutant. Therefore, as expected, winter lows and summer highs are well differentiated due to changes in atmospheric circulation and solar radiation (Figure 2d). Summer-to-winter average monthly concentration ratios ranged between 3.8 (Novi Beograd and Vračar) and 7.3 (Mostar). This indicates higher values of surface-level O_3 during summer despite more favourable (from the standpoint of air quality) meteorological conditions such as thermal convection.

CO is a fossil fuel combustion and motor vehicle exhaust product. Therefore, average monthly concentrations show well-differentiated winter highs and summer lows (Figure 2e). Its W/S ratios ranged between 4.9 (Mostar) and 6.3 (Stari Grad).

3.2 | Correlation analysis

The Pearson correlation coefficients are given in Table 4. Values were obtained from pollutant concentrations and meteorological data measured at automatic stations and analysed using the MLR. Values of statistical significance led to the conclusion that the correlation between dependent and independent variables could not be ascribed to a series of random events.

The API was positively correlated to the ΔT for all stations (with the exception of Stari Grad for October–December 2011). A higher ΔT in winter reflects cloudless weather conditions, usually with weak winds and the occurrence of fog. That is a typical condition for air pollution increase. The API was positively correlated

with the previous day's API, too, for all stations. It proved to be the predictor most strictly correlated with the API, except for Vračar, and for Mostar in January–March. Only for Vračar, in both winter half-seasons, is the API the best correlated with ΔT . Negative correlations were established between the stations' daily API and daily average wind speed, which was expected considering the wind's dispersion of air pollutants. The opposite was true for Stari Grad for October–December 2011, at which time the correlation was positive. The API is most strictly correlated with the wind speed only for Mostar (Table 4a). This is a result of Mostar's location: its topography is channelled in a south-east direction, which is the direction of a strong local wind named *koshava* (*košava* in Serbian), usually observed in coldest part of year over a large part of Serbia (Romanić *et al.*, 2016; Romanić, 2019). Correlations showed high statistical significance at 95% confidence intervals using the MLR analysis. The only one that is not statistically significant within a 95% confidence interval is that between the API and V_{av} for Vračar, January–March 2011. Meteorological variables that did not show a high correlation with the API included relative humidity and rain (Cogliani, 2001).

The Pearson correlation coefficients calculated for pollutant concentrations measured at automatic stations and meteorological data measured at BMO are similar or slightly better than previous ones (data not shown).

3.3 | Observed air quality

The air-quality appraisal was determined at all four point-locations. Table 5 displays the air quality (%) found

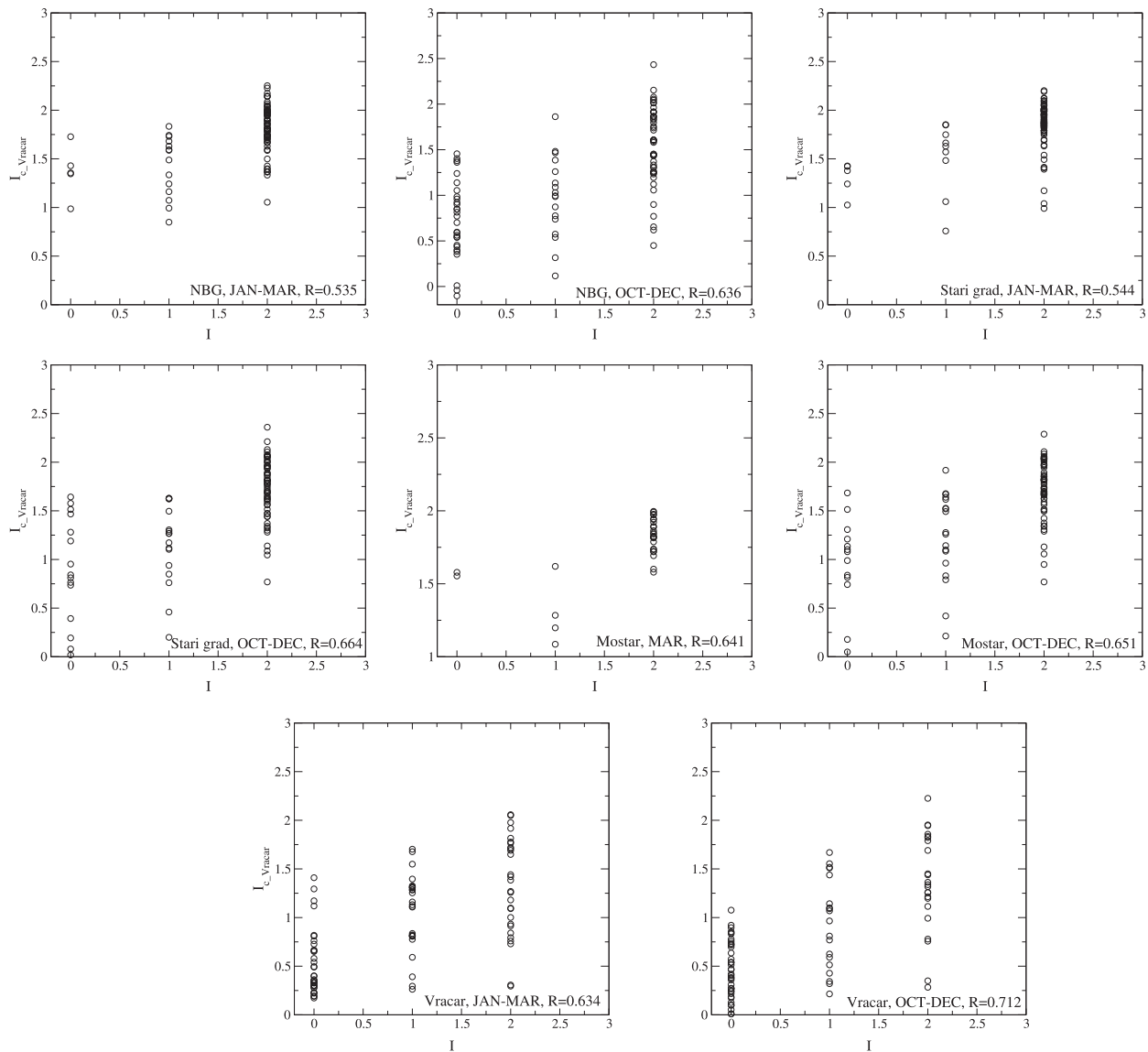


FIGURE 7 Correlation plots between I_{c_Vracar} and I for all locations with correlation co-efficients

at these locations for specific periods. The air-quality appraisals were performed in correlation with the API dependant only on observed pollutant concentrations.

Table 5 reveals the prevalence of poor air quality in two winter seasons at stations Novi Beograd, Stari Grad and Mostar. From the aspect of air pollution, January–March was less favourable than October–December 2011 (84.4% of days observed), so the air quality at Stari Grad was considered poor, followed by Mostar (80.6%) and Novi Beograd (78.9%) in January–March. A final estimate of prevailing air quality at point-location Vračar cannot be given due to absence of PM_{10} measurements.

The air quality assessment combined with the observed API results show that the area represented by Stari Grad station is the most polluted among the

comparable three. The reason for that is the much larger number of individual heating units and its low elevation. Consequently, on cold winter days, temperature inversions are frequent, as well as heavy fog (Veljović *et al.*, 2015).

3.4 | Regression parameters

Linear regression equations were obtained for winter. They were calculated for two cases: one with meteorological data measured at automatic stations; the other with meteorological data measured in the standard way, at BMO. The idea was to determine whether air quality data at point-locations could be analysed using meteorological data from a single, representative

TABLE 8 Correlation between I and $I_{c_automat}$ and between I and I_{c_ver}

	$I \sim I_{c_automat}$ January–March (only March for Mostar)	$I \sim I_{c_ver}$	$I \sim I_{c_automat}$ October–December	$I \sim I_{c_ver}$
Novi Beograd	0.484 ^a (0.000)	0.590 ^a (0.000)	0.673 ^a (0.000)	0.168 (0.110)
Stari Grad	0.511 ^a (0.000)	0.578 ^a (0.000)	0.608 ^a (0.000)	0.535 ^a (0.000)
Mostar	0.502 ^a (0.005)	0.436 ^a (0.014)	0.609 ^a (0.000)	0.511 ^a (0.000)
Vračar	0.580 ^a (0.000)	0.588 ^a (0.000)	0.662 ^a (0.000)	0.498 ^a (0.000)

Notes: Numbers in parentheses are the significance.

^aCorrelation is significant at the 0.01 level (two-tailed).

^a $I_{c_automat}$: air pollution index (API) calculated using both concentrations and meteorological data measured at automatic stations during the winter 2011; I_{c_ver} : API forecasted from 2011 equations, but with 2012/2013 measured data from the automatic stations; and I_{c_ver} : API forecasted from 2011 equations, but with 2012/2013 measured data from the automatic stations.

TABLE 9 Correlation between I and $I_{c_Vračar}$ and correlation between I and I_{c_Vrver}

	$I \sim I_{c_Vračar}$ January–March (only March for Mostar)	$I \sim I_{c_Vrver}$	$I \sim I_{c_Vračar}$ October–December	$I \sim I_{c_Vrver}$
Novi Beograd	0.535 ^a (0.000)	0.399 ^a (0.000)	0.636 ^a (0.000)	0.583 ^a (0.000)
Stari Grad	0.544 ^a (0.000)	0.481 ^a (0.000)	0.664 ^a (0.000)	0.657 ^a (0.000)
Mostar	0.641 ^a (0.000)	0.585 ^a (0.001)	0.651 ^a (0.000)	0.814 ^a (0.000)
Vračar	0.634 ^a (0.000)	0.497 ^a (0.001)	0.712 ^a (0.000)	0.695 ^a (0.000)

Notes: Numbers in parentheses are the significance.

^aCorrelation is significant at the 0.01 level (two-tailed).

^a $I_{c_Vračar}$: air pollution index (API) calculated using concentrations measured at automatic stations and meteorological data of Belgrade Meteorological Observatory (BMO) at Vračar during winter 2011; and I_{c_Vrver} : API calculated using concentrations measured at automatic stations and meteorological data measured at BMO from a 2012/2013 sample.

station. Specific regression equations for all point-locations are given in Table 6. The residuals are normally distributed, according to the normal predicted probability plots. Collinearity check was performed for all linear regression equations. The condition index was far below 15 in all cases, so the absence of multicollinearity was concluded.

3.5 | Air pollution forecast assessment

Using the data acquired by non-standard meteorological measurements of the automatic stations for forecasting API and the data acquired by standard meteorological measurements at BMO, the RMSE of the two can be calculated and a conclusion rendered, that is, data acquired by standard meteorological measurements can be used for short-term forecasting of API at point-locations. Table 7 shows the RMSE of the API forecasted using standard meteorological measurements from BMO located at Vračar, $I_{c_Vračar}$, and the data from non-standard meteorological

measurements at automatic stations. The $I_{c_automat}$ RMSE for the API acquired from standard meteorological measurements at BMO is less than the API RMSE acquired from the automatic station data at all locations and time periods, except Novi Beograd in October–December 2011. These results show that use of meteorological data acquired by standard meteorological measurements (i.e. BMO data) from Vračar to forecast API at all point-locations is justifiable.

Figures 3–6 display the three different APIs for the observed time periods. I is the station's daily API obtained from measured pollutant concentrations at the automatic station; $I_{c_Vračar}$ is the station's daily API calculated from concentrations measured at the automatic station plus meteorological data measured at BMO; and $I_{c_automat}$ is the station's daily API calculated from concentrations and meteorological data measured at the automatic station. Both calculated APIs follow the observed values very well. Figure 7 presents the correlation between $I_{c_Vračar}$ and I for all stations. The best fit was for Vračar, October–December 2011, with a correlation coefficient of

0.712; the worst was for Novi Beograd, March 2011, with a coefficient of 0.535.

3.6 | Verification

The calculated correlation between the station's daily API obtained from measured concentrations of pollutant, I , and the calculated daily API, I_c , was verified at all point-locations and during all designated time periods using data from winter 2012/2013. Namely, the requirement for using the obtained equations for the forecasting of air pollution is that the I_c calculated for one period should keep a similar correlation with I when ΔT , I_{d-1} and V_{av} were measured outside that period. Correlations between I and the daily API forecasted from the 2011 $I_{c_automat}$ (2011 coefficients and data), and correlations between I and the API forecasted from 2011 coefficients and 2012/2013 data from automatic stations, I_{c_ver} , are shown in Table 8. Comparing these correlations, no relevant changes in their values are indicated, except for station Novi Beograd for October–December 2011. This confirms the validity of the correlation equations obtained for the winter in 2011.

It can be concluded that the application of these equations for forecasting API at these point-locations, even when air pollution concentrations and meteorological data were not measured for several periods in 2011, was affirmed, except for station Novi Beograd for October–December when the correlation coefficient was substantially different.

However, the air pollution forecast assessment and the RMSE affirmed the use of BMO data at all point-locations to forecast the API, and made it necessary to verify the validity of the correlation equations obtained this way. The correlations were calculated between I and $I_{c_Vračar}$ (the station's daily API forecasted from 2011 automatic air pollution and BMO meteorological data) and I and I_{c_Vrver} (the station's daily API forecasted from 2012/2013 automatic air pollution and BMO data, using correlation co-efficients obtained from the 2011 sample). Table 9 shows these results. The correlation coefficients are similar and slightly higher for winter 2011 as expected (with the exception of Mostar in March), for which the coefficients in the MLR are originally calculated. However, by comparing these API correlations, the validity of the correlation equations obtained from the 2011 sample can be affirmed.

4 | CONCLUSIONS

Combining air pollution concentrations and meteorological variables in relation to each other and using

statistical multiple linear regression (MLR) methods, empirical equations for forecasting the air pollution index (API) in areas represented by specific automatic stations were obtained. Since the air temperature measured at the automatic stations was not obtained using standard meteorological measurements (i.e. it was measured at a height of 10 masl), Belgrade Meteorological Observatory (BMO) data obtained by standard meteorological measurements were used for the purpose of API forecasting because these data are considered representative of the city of Belgrade. These data were proven to produce slightly more precise APIs, with lower root mean square errors (RMSEs). That means measurements of temperature and wind speed representative of the whole city of Belgrade, as well as the use of temperature and wind speed data measured at each respective station, can accurately predict the API at individual monitoring stations. This may be counterintuitive, but it is perhaps the most interesting finding in the study.

Differences in correlation between the observed and calculated API, I and I_c at different point-locations can be ascribed to the location of the station, which should be more or less representative of the surrounding area. Technical issues such as equipment failure may also affect the reliability of the measurements. In order to improve the API forecasting, it is necessary to select station locations wisely in order to monitor air quality effectively and easily maintain equipment.

Using forecasted meteorological variables obtained from numeric weather forecasting models enables the forecasting of API. The API forecast can be used in efforts to lower the concentration of air pollutants in certain areas of the city and to protect citizens from unhealthy air. Since motor vehicle traffic is a great source of air pollution, practical use can be gained by regulating and reducing the number of vehicles that circulate in certain parts of the city when conditions unfavourable to human health arise. During winters, when air pollution is highest, API forecasting may encourage more ecological solutions, leading to short- and long-term air-quality improvements.

ACKNOWLEDGEMENTS

The Ministry of Science of Serbia (grant number 176013) supported this research. The authors are grateful to the Department of Air Quality Control of the Environmental Protection Agency of the Republic of Serbia, and the Hydrometeorological Service of the Republic of Serbia for providing the data in this research.

ORCID

Dragana Vujović  <https://orcid.org/0000-0002-7738-7018>

REFERENCES

- Air Protection Law Regulation (2010). Uredba o uslovima za monitoring i zahtevima kvaliteta vazduha, Službeni glasnik Republike Srbije, No.11/2010. (in Serbian).
- Anderson, H.R. (2009) Air pollution and mortality: a history. *Atmospheric Environment*, 43, 142–152.
- Berkowicz, R. (2000) OSPM—a parameterised street pollution model. *Environmental Monitoring and Assessment*, 65, 323–331.
- Census Atlas. (2014) *Census of Population, Households and Dwellings in the Republic of Serbia in 2011*. Belgrade, Serbia: The Statistical Office of the Republic of Serbia, p. 128. Available at: <http://publikacije.stat.gov.rs/G2014/Pdf/G20144012.pdf> [Accessed 24th August 2020].
- Cogliani, E. (2001) Air pollution forecast in cities by an air pollution index highly correlated with meteorological variables. *Atmospheric Environment*, 35, 2871–2877.
- EIS. (2012) *Environment Report in the Republic of Serbia for 2011 Year*. Belgrade, Serbia: Serbian Environmental Protection Agency, p. 193. (in Serbian). Available at: <http://www.sepa.gov.rs/download/Izvestaj2011.pdf> [Accessed 24th August 2020].
- EIS. (2013) *Environment Report in the Republic of Serbia for 2012 Year*. Belgrade, Serbia: Serbian Environmental Protection Agency, p. 168. (in Serbian). Available at: http://www.sepa.gov.rs/download/Izvestaj_2012.pdf [Accessed 24th August 2020].
- Ge, W.Z., Xu, F., Zhao, Z.H., Zhao, J.Z. and Kan, H.D. (2013) Association between diurnal temperature range and respiratory tract infections. *Biomedical and Environmental Science*, 26(3), 222–225. <https://doi.org/10.3390/ijerph14080891>.
- Deniz Genc, D., Yesilyurt, C. and Tuncel, G. (2010) Air pollution forecasting in Ankara, Turkey using air pollution index and its relation to assimilative capacity of the atmosphere. *Environmental Monitoring and Assessment*, 166, 11–27.
- Gurjar, B.C., Jain, A., Sharma, A., Agarwal, A., Gupta, P., Nagpure, A. S. and Lelieveld, J. (2010) Human health risks in megacities due to air pollution. *Atmospheric Environment*, 44, 4606–4613.
- Kampa, M. and Castanas, E. (2008) Human health effects of air pollution. *Environmental Pollution*, 151(2), 362–367. <https://doi.org/10.1016/j.envpol.2007.06.012>.
- Kumar, A. and Goyal, P. (2011) Forecasting of daily air quality index in Delhi. *Science of Total Environment*, 409, 5517–5523.
- Law Order (2006). Pravilnik o graničnim vrednostima, metodama merenja imisija, kriterijumima za uspostavljanje mernih mesta i evidenciji podataka, Službeni glasnik Republike Srbije, 19/2006. (in Serbian).
- Lazić, L., Aničić Urošević, M., Mijić, Z., Vuković, G. and Ilić, L. (2016) Traffic contribution to air pollution in urban street canyons: Integrated application of the OSPM, moss biomonitoring and spectral analysis. *Atmospheric Environment*, 141, 347–360.
- Plaia, A. and Ruggieri, M. (2011) Air quality indices: A review. *Reviews in Environmental Science and Biotechnology*, 10, 165–179. <https://doi.org/10.1007/s11157-010-9227-2>.
- Romanić, Đ., Čurić, M., Lompar, M. and Jovičić, I. (2016) Contributing factors to Koshava wind characteristics. *International Journal of Climatology*, 36, 956–973.
- Romanić, Đ. (2019) Local winds of Balkan Peninsula. *International Journal of Climatology*, 39(1), 1–17. <https://doi.org/10.1002/joc.5743>.
- Squizzato, S., Cazzaro, M., Innocente, E., Visin, F., Hopke, P.K. and Rampazzo, G. (2017) Urban air quality in a mid-size city-PM_{2.5} composition, sources and identification of impact areas: from local to long range contributions. *Atmospheric Research*, 186, 51–62.
- Stieb, D.M., Burnett, R.T., Smith-Doiron, M., Brion, O., Shin, H.H. and Economou, V. (2012) A new multipollutant, no-threshold air quality health index based on short-term associations observed in daily time-series analyses. *Journal of Air Waste Management*, 58(3), 435–450. <https://doi.org/10.3155/1047-3289.58.3.435>.
- UN, 2019 (2018). World Urbanization Prospects: The 2018 Revision. United Nations, Department of Economic and Social Affairs, Population Division.
- Veljović, K., Vujović, D., Lazić, L. and Vučković, V. (2015) An analysis of fog events at Belgrade international airport. *Theoretical and Applied Climatology*, 119, 13–24. <https://doi.org/10.1007/s00704-014-1090-6>.
- Vujović, D. and Todorović, N. (2017) An assessment of pollutant emissions due to air traffic at Nikola tesla international airport, Belgrade, and the link between local air quality and weather types. *Transportation Research Part D—Transport and Environment*, 56, 85–94.
- WHO. (2006) Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulphur dioxide: global update 2005. In: *Summary of Risk Assessment*. Geneva, Switzerland: World Health Organization. Available at: http://whqlibdoc.who.int/hq/2006/WHO_SDE_PHE_OEH_06.02_eng.pdf [Accessed 21th May 2018].
- Wilks, D.S. (2006) *Statistical Methods in the Atmospheric Sciences*. London, England: Academic Press, p. 627.
- Yadav, R., Sahu, L.K., Beig, G. and Jaaffrey, S.N.A. (2016) Role of long-range transport and local meteorology in seasonal variation of surface ozone and its precursors at an urban site in India. *Atmospheric Research*, 176–177, 96–107.
- Zhang, H., Wang, Y., Park, T.-W. and Deng, Y. (2017a) Quantifying the relationship between extreme air pollution events and extreme weather events. *Atmospheric Research*, 188, 64–79.
- Zhang, Y., Yu, C., Yang, J., Zhang, L. and Cui, F. (2017b) Diurnal temperature range in relation to daily mortality and years of life lost in Wuhan, China. *International Journal of Environmental Research Public Health*, 14(8), 891. <https://doi.org/10.3390/ijerph14080891>.

How to cite this article: Đurić M, Vujović D. Short-term forecasting of air pollution index in Belgrade, Serbia. *Meteorol Appl.* 2020;27:e1946. <https://doi.org/10.1002/met.1946>

APPENDIX

API = air pollution index

ΔT = diurnal temperature range

RMSE = root mean square error

MLR = multiple linear regression

BMO = Belgrade Meteorological Observatory

I = monitored air pollution index

I_c = calculated air pollution index

I_{d-1} = previous day's air pollution index

$I_{c_Vračar}$ = API calculated using concentrations measured at automatic stations and meteorological data of BMO at Vračar during winter 2011

$I_{c_automat}$ = API calculated using both concentrations and meteorological data measured at automatic stations during winter 2011

I_{c_Vrver} = API calculated using concentrations measured at automatic stations and meteorological data measured at BMO from a 2012/2013 sample

I_{c_ver} = API forecasted from 2011 equations, but with 2012/2013 measured data from the automatic stations