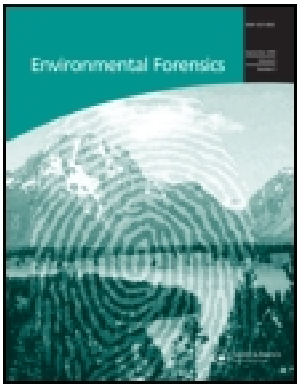

Environmental Forensics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uenf20

A New Approach to Prediction of SO_2 and PM_{10} Concentrations in Istanbul, Turkey: Cellular Neural Network (CNN)

Ülkü Alver Şahin [a] , Osman Nuri Ucan [b] , Cuma Bayat [c] & Orhan Tolluoglu [d]
[a] Istanbul University, Environmental Engineering Department, Istanbul, Turkey
[b] Aydın University, Electrical-Electronics Engineering Department, Istanbul, Turkey
[c] Arel University, Tepekent-Büyükçekmece, Istanbul, Turkey
[d] Air Force Academy, Istanbul, Turkey
Version of record first published: 09 Sep 2011.

To cite this article: Ülkü Alver Şahin , Osman Nuri Ucan , Cuma Bayat & Orhan Tolluoglu (2011): A New Approach to Prediction of SO_2 and PM_{10} Concentrations in Istanbul, Turkey: Cellular Neural Network (CNN), Environmental Forensics, 12:3, 253-269

To link to this article: http://dx.doi.org/10.1080/15275922.2011.595047

Environmental Forensics, 12:253–269, 2011

ISSN: 1527-5922 print / 1527-5930 online
DOI: 10.1080/15275922.2011.595047

A New Approach to Prediction of SO_2 and PM_{10} Concentrations in Istanbul, Turkey: Cellular Neural Network (CNN)

Ülkü Alver Şahin,[1] Osman Nuri Ucan,[2] Cuma Bayat,[3] and Orhan Tolluoglu[4]

[1]*Istanbul University, Environmental Engineering Department, Istanbul, Turkey*
[2]*Aydın University, Electrical-Electronics Engineering Department, Istanbul, Turkey*
[3]*Arel University, Tepekent-Büyükçekmece, Istanbul, Turkey*
[4]*Air Force Academy, Istanbul, Turkey*

This article describes the application of a cellular neural network (CNN) to model air pollutants. In this study, forthcoming daily and hourly values of particulate matter (PM_{10}) and sulphur dioxide (SO_2) were predicted. These air pollutant concentrations were measured at four different locations (Yenibosna, Sarachane, Umraniye and Kadikoy) in Istanbul between 2002 and 2003. Eight different meteorological parameters (temperature, wind speed and direction, humidity, pressure, sunshine, cloudiness, rainfall) recorded at Florya and Goztepe meteorological stations were used to model inputs. First, the results of CNN prediction and statistical persistence method (PER) were compared. Then, CNN and PER outputs were correlated with real time values by using statistical performance indices. The indices of agreement (*d*) for daily mean concentrations were found using CNN and PER prediction models: 0.71–0.80 and 0.71–0.78 for PM_{10}, and 0.81–0.84 and 0.77–0.82 for SO_2 in all air quality measurement stations, respectively. From these values, CNN prediction model are concluded to be more accurate than PER, which is used for comparison. In hourly prediction of mean concentrations with CNN, *d* value is found as 0.78 and 0.92 for PM_{10} and SO_2, respectively. Thus, it was concluded that CNN-based approaches could be promising for air pollutant prediction.

Keywords: cellular neural network (CNN), Air pollution, particulate matter (PM), sulfur dioxide (SO_2), meteorology

Introduction

The main sources of air pollution in Istanbul, Turkey are the combustion of poor quality coal, increased traffic load, and industrial activities. During the winter, sulfur dioxide (SO_2) and particulate matter (PM) are the major air pollutants affecting regional air quality. In the past two decades, many scientists have focused on the air pollution problems in Istanbul (Tayanç, 2000; Saral and Ertürk, 2003; Esen et al., 2005; Im et al., 2008; Gaga et al., 2009; Şahin et al., 2011). SO_2 is formed as a result of burning coal and oil, which consist of sulfur, metal melts, and other industrial outputs. The highest SO_2 concentration is observed in domestic areas and industrial regions especially in winter when poor quality coal is improperly used. PM is formed by the mixture of oil, gasoline, and diesel fuel combustions. When SO_2 and PM concentrations are high, the level of the respiratory and cardiovascular diseases also increases. SO_2 concentration as a harmful pollutant causes acid rain and various corrosion effects on constructions.

Many deterministic and stochastic approaches exist for modeling the concentrations of air pollutants. Various deterministic, dispersion, and statistical models can be studied in this area. Mathematical expressions in deterministic models are insufficient to explain real life-physical and chemical processes. Dispersion models are only effective if many parameters, such as meteorological data and the special emission sources, are considered; thus these models are not easy to implement. Classic statistical models do not reflect the complexity of variables elsewhere. The well-known machine-learning approach is artificial neural networks (ANN). This approach is concerned with the design and development of algorithms that allow computers to empirically learn the behavior of data sets. In many studies, ANNs are applied to predict environmental pollutants (Boznar et al., 1993; Mok and Tam, 1998; Saral and Ertürk, 2003; Chelani et al., 2002; Sahin et al., 2005; Raha, 2007). Gardner and Dorling (1998) published a comprehensive review of studies using an ANN approach for environmental air pollution modeling. Kukkonen et al. (2003) studied five neural network (NN) models, a linear statistical model and a deterministic modeling system for the prediction of urban NO_2 and PM_{10} concentrations. Sahin et al. (2004) used a multi-layer neural network model to predict daily CO concentrations in the European side of Istanbul, Turkey, by using meteorological variables. Kurt et al. (2008) also developed an online air pollution forecasting system for Istanbul using NN. Another NN model developed by Saral and Ertürk (2003) was also used to predict regional SO_2 concentrations. Junninen et al. (2004) applied regression-based imputation, nearest neighbor interpolation, a

Address correspondence to Ülkü Alver Şahin, Istanbul University, Environmental Engineering Department, 34 850, Avcilar, Istanbul, Turkey. E-mail: ulkualver@istanbul.edu.

self-organizing map, a multilayer perceptron model and hybrid methods to simulate missing air quality data. Nagendra and Khare (2006) studied the usefulness of NNs in understanding the relationship between traffic parameters and NO_2 concentrations. Recently, several researchers used NN techniques to predict airborne PM concentrations, including Ordieres et al. (2005), Hooybergs et al. (2005), Perez and Reyes (2006), and Slini et al. (2006). All of these studies reported that ANN could be used to develop efficient air-quality analysis and forward-looking prediction models. In ANNs, however, the training process becomes increasingly complex and requires longer periods of time because the number of weighting coefficients of the ANN rises into the millions due to the complexity of the environmental study.

To reduce the number of weighting coefficients, Chua and Yang (1988) introduced another machine learning approach, the cellular neural network (CNN), in 1988. Because each cell of the CNN is represented by a separate analog processor, and each cell is locally interconnected to its neighbors by matrix A and receives feedback from them by matrix B, this configuration presents a very high-speed tool for parallel dynamic processing of two-dimensional (2D) structures (Cimagalli, 1993; Guzelis and Karamahmut, 1994; Ucan et al., 2001; Grassi and Grieco, 2002; Thai and Cat, 2008). CNN approaches have been applied to air pollution modeling by a number of researchers with excellent results (Ozcan et al., 2007; Thai and Cat, 2008; Sahin et al., 2011).

In this study, a CNN method was applied to predict the daily mean and hourly mean concentrations of PM_{10} and SO_2 pollutants in the Yenibosna, Sarachane, Ümraniye and Kadıköy regions of Istanbul, Turkey. This discussion is organized beginning with the next section in which the study area and database are explained. Then the CNN and PER modeling techniques are defined. To evaluate model prediction, statistical performance indices are explained. Next, the CNN is tested on real data and the results are presented and compared with PER technique. Finally, the results of the study are evaluated.

Material and Methods

Study Area and Data

The study area is the metropolitan city of Istanbul, which is located 41° N and 29° E. The Bosporus Channel separates this city into two areas, the European and the Asian sides. The total area of the both parts of the city is approximately 5700 km^2. More than 12 million people live in Istanbul and more than 40% of Turkey's heavy industry is located in the city. For this reason, air pollution problems are important in Istanbul.

The Greater Istanbul Metropolitan Municipality Directorate of Environmental Protection (IGMM-DEP) has conducted air pollution measurement at 10 observation stations located at various key topographic points around the city since 1992. General Directorate of the Turkish State Meteorological Services (GDTSMS) in Istanbul provided the daily meteorological data.

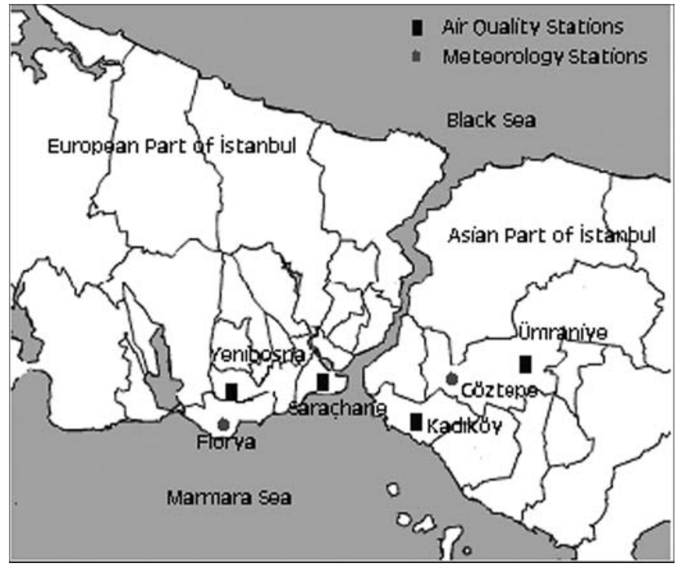

Figure 1. Map of the air quality and meteorology stations in Istanbul, Turkey.

A total of 17 meteorology stations are located in various parts of Istanbul. In this study, SO_2 and PM_{10} concentrations were measured by four stations located in Yenibosna, Sarachane, Göztepe and Kadıköy-Istanbul, and the daily meteorological data was measured by two stations located in Florya and Göztepe-Istanbul as shown in Figure 1.

Figure 1 shows the location of the four air quality measurement (AQM) stations and two meteorology stations. The sampling sites were categorized using criteria proposed by the European Environmental (EU) Agency and shown in Table 1 (Dingenen et al., 2004). Table 1 shows the specific pollution sources near the air quality monitoring stations. Among these criteria are the distance of the stations from large pollution sources such as cities, power plant and major motorways, and the traffic volume.

In this study, SO_2 and PM_{10} data were collected by GIMM-DEP and measured using AF 21 M and MP 101 M sensors, respectively (Environmental Inc.,). We evaluated data measured in Yenibosna, Saraçhane, Göztepe and Kadiköy location of lstanbul during 2002 and 2003. The number of total data units is 5840

Table 1. *Specific pollution sources and category by the European Environmental Agency (EU) of the air pollutant sampling sites.*

	Pollution Sources			Categorized by EU	
AQ Stations	Commercial	Industrial	Traffic	Urban background[1]	Curbside[2]
Yenibosna	x	x	x	x	x
Saraçhane	x	—	x	x	x
Umraniye	x	x	—	x	—
Kadiköy	x	—	—	x	—

[1]Urban background: <2500 vehicles/day within a radius of 50 m.
[2]Kerbside: within street canyons.

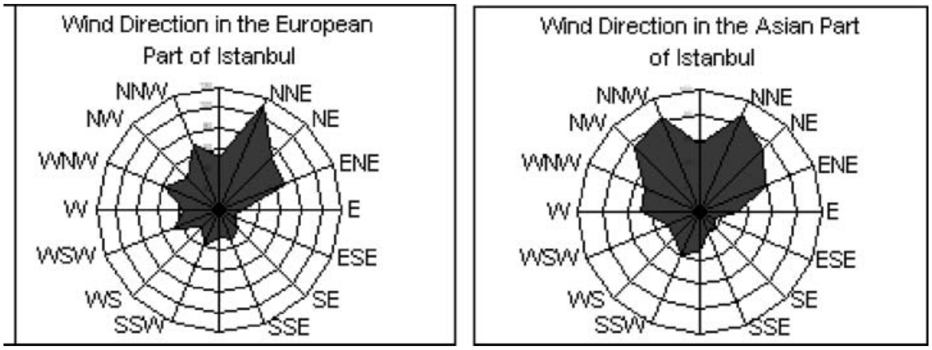

Figure 2. Wind direction in the European and the Asian part of Istanbul, Turkey, during the years 2002 and 2003.

and the number of days is 730 for each measurement point and each pollutant. Also attempted to predict were the hourly mean PM_{10} and SO_2 concentrations measured in Yenibosna AQM Station during February and March of 2003. In total, 1340 hours of pollutant data were used in the study between February 2002 and April 2002. To predict the future air pollutant concentration, data measured in Florya and Göztepe Meteorological Stations were used because of their proximity to AQM stations considered in the study. Meteorological parameters used in the study include dry-bulb temperature that is the ambient temperature measured in whole degrees (°C), cloudiness that is the amount of cloud cover ranging from 0 to 10 where 0 represents clear sky and 10 represents overcast cloudy sky, relative humidity that is the percentage of water vapor in air, atmospheric pressure in mbar, daily duration of sunlight in hour and wind direction in degrees from which the wind is blowing, for example, 90° (East [E]), 180° (South [S]), 270° (West [W]) and 360° (North [N]). Total wind directions are 16 in the model. Figure 2 shows the daily wind directions during the study period (2002–2003). Prevailing wind direction is observed NNE in European part of Istanbul and as between NE and NW in Asian part of Istanbul. These wind directions show that pollutants in the region are derived from the home, industry, and traffic sources. Air movement from Marmara Sea to measurement locations is very poor. In addition, wind speed is the average value of the day measured in m/s. Precipitation amount is a general term used for rainfall, snowfall and hailfall in mm/m^2. The number of total meteorological data used in this study is 11680. Statistical evaluations of all air pollutants and meteorological data pertaining to 2002–2003 are shown in Table 2. The monitoring data is designed to meet the requirements for training and testing CNN. This database, in its form of original time series, is divided into training and test sets taking the odd numbered pattern as training data and even numbered ones as test data.

Structure of Cellular Neural Networks

Most neural networks fall into two main classes: *memoryless neural networks* and *dynamical neural networks*. As in Hopfield networks and CNN, dynamical neural networks are usually designed as dynamic systems in which the inputs are set to constant values and the path approach to a stable equilibrium point depends upon the initial state. A CNN is composed of large-scale nonlinear analog circuits which process signals in real time (Chua and Yang, 1988). The basic unit of a CNN is called a cell, and these units communicate with each other directly only through their nearest neighbors. Adjacent cells can therefore interact directly with each other. Cells not directly connected together affect each other indirectly because of the propagation effects of the continuous real-time dynamics of the CNN. The structure of a 2D 3×3 CNN is shown in Figure 3.

The CNN used in this study consisted of M rows and N columns (MxN). In this structure, ith line and jth column are designated cell (i,j) and denoted by C(i,j). A typical example of a cell is shown in Figure 4. In Figure 4, u_{ij}, y_{ij} and x_{ij} correspond to the input, the output and the state variable of the cell, respectively. The node voltage v_{xij} of $C(i,j)$ is defined as the state of the cell whose initial condition is assumed to have a magnitude less

Table 2. *The minimum, mean and maximum values of meteorological model parameters during the years 2002 and 2003.*

			Minimum		Maximum		Mean	
Meteorologic parameters	Abbreviations	Units	Florya	Goztepe	Florya	Goztepe	Florya	Goztepe
Temperature	T	°C	−2.2	−2.2	31.2	32	14.7	14.7
Wind speed	WS	m/s	0.3	0.2	6.2	7.3	2.2	2.5
Sunshine	S	hour	0	0	13.8	12.9	6.7	6.3
Relative humidity	RH	%	43.3	38.7	95.7	96	72.2	74.8
Pressure	P	mbar	990.9	988.8	1031.4	1032.7	1012.5	012.6
Cloudy	C	M	0	0	10	10	4.4	6.3
Wind direction	WD	North (N), South (S), West (W), East (E)	WSW		NNW		—	
Rainfall	R	mm	0	0	31.8	61.9	—	

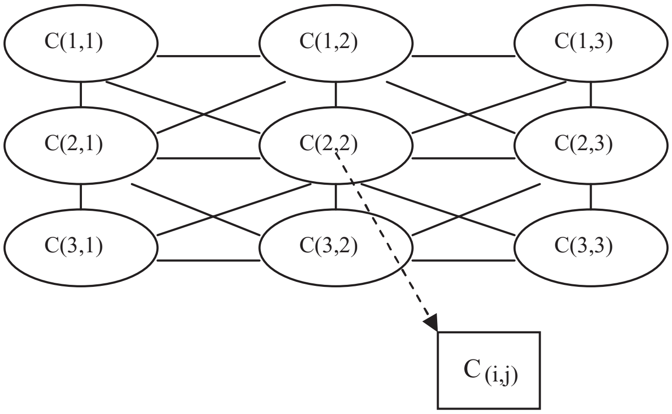

Figure 3. A two-dimensional cellular neural network (CNN) of size 3 × 3. Links between the cells (ellipse) indicate interactions between the linked cells.

than or equal to 1. Each cell contains one independent current source, one linear-capacitor C, two linear resistors Rx and Ry and linear voltage-controlled current sources ($I_{xy}(i,j{:}k,l)$), which are coupled to its neighbor cells via the controlling input voltage and the feedback from the output voltage of each neighboring cell $C(k,l)$. The constant coefficients $A(i,j;k,l)$ and $B(i,j;k,l)$ are known as the cloning templates, and these are the parameters linking cell $C(i,j)$ to its neighbor $C(k,l)$. The equivalent block diagram of a CNN cell is shown in Figure 5. The first-order non-linear equation defining the dynamic of a CNN can be derived as follows (Arena et al., 1997; Hadad and Piroozman, 2007; Thai and Cat, 2008):

The r-neighborhood of a cell $C(i,j)$ in a CNN is defined by:

$$N_r(i,\ j) = \{C(k,l)/\max(/k-i/,\ l-j/ \le r,\ 1 \le i \le M;\quad 1 \le j \le N\} \tag{1}$$

A general form of the cell dynamical equations may be written as:

$$C\frac{dv_{xij}(t)}{dt} = -\frac{1}{R}v_{xij}(t) + \sum_{C(k,l)\in N_r(i,j)} A(i,\ j;k,l)v_{ykl}(t) + \sum_{C(k,l)\in N_r(i,j)} B(i,\ j;k,l)v_{ukl} + I \tag{2}$$

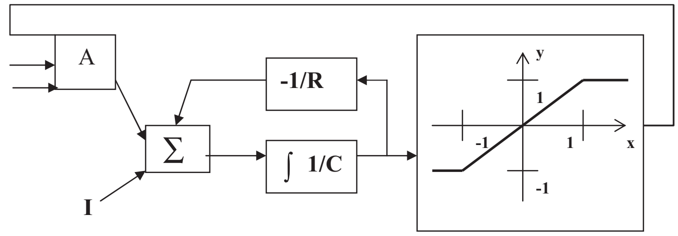

Figure 5. Equivalent block diagram of a cell in a cellular neural network (CNN).

In the CNN system, (A,B,I) are the local connective weighting values of each cell $C(i,j)$ to its neighbors. Each cell of the CNN is represented by a separate analog processor, and each cell is locally interconnected to its neighbors by matrix A and gets feedback from them by matrix B. This configuration results in a very high-speed tool for parallel dynamic processing of 2D structures:

$$A = \begin{bmatrix} a_{-1,-1} & a_{-1,0} & a_{-1,1} \\ a_{0,-1} & a_{0,0} & a_{0,1} \\ a_{1,-1} & a_{1,0} & a_{1,1} \end{bmatrix},\quad B = \begin{bmatrix} b_{-1,-1} & b_{-1,0} & b_{-1,1} \\ b_{0,-1} & b_{0,0} & b_{0,1} \\ b_{1,-1} & b_{1,0} & b_{1,1} \end{bmatrix},\ I \tag{3}$$

The output is related to the state by the nonlinear equation. Characteristic of the output function $v_{yi,j} = f(v_{xi,j})$ are:

$$v_{yij}(t) = \frac{1}{2}\left(\left|v_{xij}(t)+1\right| - \left|v_{xij}(t)-1\right|\right)$$
$$v_{yij} = \begin{cases} -1 & \text{when} \quad v_{xij} < -1 \\ v_{xij} & \text{when} \quad -1 < v_{xij} < 1 \\ 1 & \text{when} \quad v_{xij} > 1 \end{cases} \tag{4}$$

The network behavior of a CNN depends on the initial state of the cells, namely the bias I, and the weighting values of the A and B matrices, which are associated with the connections inside the well-defined neighborhood of each cell. CNNs are arrays of locally and regularly interconnected neurons or cells whose global functionalities are defined by a small number of

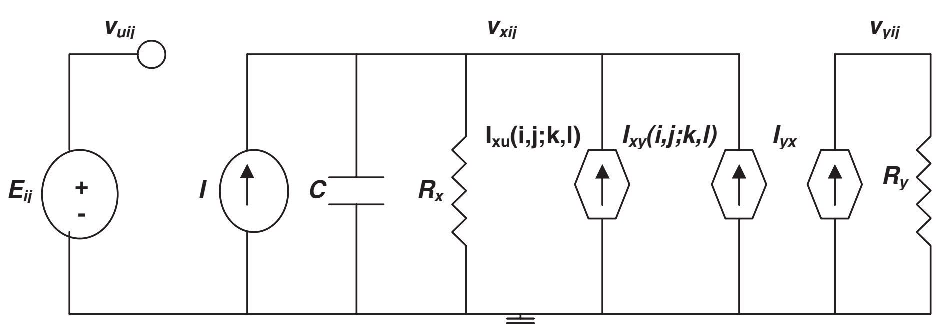

Figure 4. A classic cell scheme for a cellular neural network (CNN).

$$u=\begin{bmatrix} WD_{t} & WD_{t+2} & WD_{t+4} & \cdots & WD_{t+n} \\ RH_{t} & RH_{t+2} & RH_{t+4} & \cdots & RH_{t+n} \\ P_{t} & P_{t+2} & P_{t+4} & \cdots & P_{t+n} \\ T_{t} & T_{t+2} & T_{t+4} & \cdots & T_{t+n} \\ SO_{2t} & SO_{2t+2} & SO_{2t+4} & \cdots & SO_{2t+n} \\ PM_{10t} & PM_{10t+2} & PM_{10t+4} & \cdots & PM_{10t+n} \\ WS_{t} & WS_{t+2} & WS_{t+4} & \cdots & WS_{t+n} \\ C_{t} & C_{t+2} & C_{t+4} & \cdots & C_{t+n} \\ R_{t} & R_{t+2} & R_{t+4} & \cdots & R_{t+n} \\ S_{t} & S_{t+2} & S_{t+4} & \cdots & S_{t+n} \end{bmatrix} \quad y=\begin{bmatrix} WD_{t+1} & WD_{t+3} & WD_{t+5} & \cdots & WD_{t+n+1} \\ RH_{t+1} & RH_{t+3} & RH_{t+5} & \cdots & RH_{t+n+1} \\ P_{t+1} & P_{t+3} & P_{t+5} & \cdots & P_{t+n+1} \\ T_{t+1} & T_{t+3} & T_{t+5} & \cdots & T_{t+n+1} \\ SO_{2t+1} & SO_{2t+3} & SO_{2t+5} & \cdots & SO_{2t+n+1} \\ PM_{10t+1} & PM_{10t+3} & PM_{10t+5} & \cdots & PM_{10t+n+1} \\ WS_{t+1} & WS_{t+3} & WS_{t+5} & \cdots & WS_{t+n+1} \\ C_{t+1} & C_{t+3} & C_{t+5} & \cdots & C_{t+n+1} \\ R_{t+1} & R_{t+3} & R_{t+5} & \cdots & R_{t+n+1} \\ S_{t+1} & S_{t+3} & S_{t+5} & \cdots & S_{t+n+1} \end{bmatrix}$$

a) CNN model training data set

$$u=\begin{bmatrix} WD_{t+1} & WD_{t+3} & WD_{t+5} & \cdots & WD_{t+n+1} \\ RH_{t+1} & RH_{t+3} & RH_{t+5} & \cdots & RH_{t+n+1} \\ P_{t+1} & P_{t+3} & P_{t+5} & \cdots & P_{t+n+1} \\ T_{t+1} & T_{t+3} & T_{t+5} & \cdots & T_{t+n+1} \\ SO_{2t+1} & SO_{2t+3} & SO_{2t+5} & \cdots & SO_{2t+n+1} \\ PM_{10t+1} & PM_{10t+3} & PM_{10t+5} & \cdots & PM_{10t+n+1} \\ WS_{t+1} & WS_{t+3} & WS_{t+5} & \cdots & WS_{t+n+1} \\ C_{t+1} & C_{t+3} & C_{t+5} & \cdots & C_{t+n+1} \\ R_{t+1} & R_{t+3} & R_{t+5} & \cdots & R_{t+n+1} \\ S_{t+1} & S_{t+3} & S_{t+5} & \cdots & S_{t+n+1} \end{bmatrix} \quad y=\begin{bmatrix} WD_{t+2} & WD_{t+4} & WD_{t+6} & \cdots & WD_{t+n} \\ RH_{t+2} & RH_{t+4} & RH_{t+6} & \cdots & RH_{t+n} \\ P_{t+2} & P_{t+4} & P_{t+6} & \cdots & P_{t+n} \\ T_{t+2} & T_{t+4} & T_{t+6} & \cdots & T_{t+n} \\ SO_{2t+2} & SO_{2t+4} & SO_{2t+6} & \cdots & SO_{2t+n} \\ PM_{10t+2} & PM_{10t+4} & PM_{10t+6} & \cdots & PM_{10t+n} \\ WS_{t+2} & WS_{t+4} & WS_{t+6} & \cdots & WS_{t+n} \\ C_{t+2} & C_{t+4} & C_{t+6} & \cdots & C_{t+n} \\ R_{t+2} & R_{t+4} & R_{t+6} & \cdots & R_{t+n} \\ S_{t+2} & S_{t+4} & S_{t+6} & \cdots & S_{t+n} \end{bmatrix}$$

b) CNN model testing data set

Figure 6. Input (*u*) and output (y) matrices of the cellular neural network (CNN) model for training and testing in this study.

parameters (*A, B,* and *I*) that specify the operation of the component cells as well as the connection weights between them. The CNN can also be considered as a nonlinear convolution with the template. Since their introduction in 1988 by Chua and Yang, the CNN has attracted a lot of attention. Not only do these systems have a number of attractive properties from a theoretical point of view, but they also have many well-known applications such as image processing, motion detection, pattern recognition and simulation. Albora et al. (2001) applied this contemporary approach to the separation of regional and residual magnetic anomalies on synthetic and real data. Hadad and Piroozmand (2007) applied the CNN to modeling and solving the nuclear reactor dynamic equations. Here, air pollution parameters were predicted using CNN approach. To evaluate the prediction results of the CNN, statistical performance indices were calculated described as later in text.

Four matrices were built for PM_{10} and SO_2 from the data set during 2002 and 2003 years. These matrices are shown in Figure 6. All have 10 rows and 365 columns to predict daily concentrations, and 10 rows and 670 columns to predict hourly concentrations for each station.

Structure of the Persistence Method

The PER consists of a very simple prediction: Today *(t)* PM_{10} and SO_2 mass concentration will be the same as yesterday *(t-1)*. In this case:

$$y_t = f(y_{t-1}) \tag{5}$$

$$PM_{10(t)} = PM_{10(t-1)}, \quad SO_{2t} = SO_{2,(t-1)} \tag{6}$$

it is not expected to be very accurate. Here, there is no parameter to adjust, and prediction errors are calculated from the performance of the test set using Equations 7–11.

Statistical Performance Indices

To evaluate model prediction objectively in this study, four statistical performance indices are computed: (1) the correlation coefficient (r), and the index of agreement (*d*), (2) the mean bias error (*Bias*), (3) the mean absolute error (*MAE*) and (4) the root mean squared error (*RMSE*). These indices are based on the

Table 3. *Summary statistics of daily PM_{10} and SO_2 concentrations ($\mu g/m^3$) of each stations during the years 2002 and 2003 [winter (summer)]*

Stations	Pollutions	Mean	Standard deviation	Minimum	Maximum	Median
Yenibosna	PM_{10}	65.6 (59.8)	29.9 (21.5)	18.0 (12.7)	212.3 (147.6)	61.3 (58.0)
Saraçhane	PM_{10}	71.9 (65.7)	36.9 (24.3)	13.8 (27.4)	248.0 (177.0)	65.6 (59.5)
Kadiköy	PM_{10}	63.8 (45.5)	38.0 (30.3)	4.2 (4.6)	255.5 (271.1)	59.3 (38.9)
Ümraniye	PM_{10}	59.9 (51.4)	36.9 (21.2)	5.0 (12.5)	226.7 (147.5)	52.5 (47.1)
Yenibosna	SO_2	37.0 (20.7)	24.5 (21.4)	0.0 (0.0)	161.7 (108.0)	32.3 (14.0)
Saraçhane	SO_2	39.5 (15.8)	30.8 (12.6)	0.5 (0.0)	173.2 (73.3)	32.4 (13.8)
Kadiköy	SO_2	24.3 (10.2)	19.1 (10.2)	0.0 (0.0)	133.2 (60.6)	18.7 (7.2)
Ümraniye	SO_2	26.1 (10.0)	25.9 (10.3)	0.0 (0.0)	165.8 (55.7)	19.1 (7.0)

deviations between predicted and original observation values. *RMSE* summarizes the difference between the observed and the imputed concentrations and was used to quantify the average error of the model. Moreover, the *MAE and RMSE* were included in the comparison as more sensitive measures of residual error as. *Bias* is the degree of correspondence between the mean prediction and the mean observation. Lower values of bias are optimal, while bias values <0 indicate underforecasting. Evaluation can also be undertaken by considering measures of agreement, such as the Pearson product moment correlation coefficient (r). The index of agreement is abounded, relative measure that is capable to measuring the degree to which predictions are error-free. The denominator accounts for the deviation of model from the mean of the observations as well as to the deviation of observation from their mean values. In a good model d and r should approach to 1 (Nunnari et al., 2004; Kukkonen et al., 2003). All these indices are formulated as:

$$r = \sqrt{1 - \frac{\sum_{i=1}^{N} (O_i - T_i)^2}{\sum_{i=1}^{N} (O_i - \hat{O})^2}} \tag{7}$$

$$d = 1 - \frac{\sum_{i=1}^{N} (\mathrm{P}_i - \mathrm{O}_i)^2}{\sum_{i=1}^{N} \left(|\mathrm{P}_i - \overline{\mathrm{O}}| + |\mathrm{O}_i - \overline{\mathrm{O}}| \right)^2} \tag{8}$$

$$Bias = \frac{1}{N} \sum_{i=1}^{\mathrm{N}} (\mathrm{O_i} - \mathrm{P_i}) \tag{9}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |O_i - T_i| \tag{10}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (O_i - T_i)^2} \tag{11}$$

where O_i and P_i are the observed and predicted pollution values, respectively, in $i = 1., 2., \ldots, N$ days, while $\hat{O}$ is the mean of the observed time series and N is the total number of observations. In addition, the standard deviations (σ) of the observed time series (O) and predicted time series (P) were calculated.

Results and Discussion

Summary statistics of daily PM_{10} and SO_2 data for seasonal between 2002 and 2003 at the Yenibosna, Saraçhane, Ümraniye, and Kadıköy stations are given in Table 3. The PM_{10} and SO_2 concentrations recorded at Saraçhane stations are the highest value. This station is in the urban area with traffic and residential population and also in the low sea level. In Yenibosna, traffic, industry and residential populations are quite dense. The 5-year average SO_2 concentration measured at the Yenibosna station was 1.5 times higher than the concentration measured at the Umraniye station (Şahin et al., 2011). As given in Table 3, at both monitoring stations the results of SO_2 recorded in winter were 2 times higher than those measured in summer. The 24-hour PM_{10} limit (50 $\mu g/m^3$) was exceeded on many days for all stations. But the 24-hour SO_2 limit (125 $\mu g/m^3$) was exceeded on only a few days for all stations. Before 1995, the average SO_2 level was 250 $\mu g/m^3$ in Istanbul (Tayanç, 2000). After 1995, the use of natural gas instead of coal became more widespread and SO_2 levels have therefore begun to decrease. After 1999, the average SO_2 concentration was 25 $\mu g/m^3$. However, PM_{10} levels have not effectively decreased over this period. No significant difference was reported in PM_{10} pollution levels between winter and summer and also daily. The effect of long distance transport should be considered as well as the anthropogenic pollution sourced from industry, heating and transport (Karaca, 2009; Kındap, 2008). It is known that air pollutants can stay for a long time in atmosphere or can be transported. Therefore, pollution occurring in a certain area might be caused by a local emission source present at the same time or a little while before. In addition, immediate or past meteorological factors are known to be quite effective on the concentrations of pollutant in atmosphere. Consequently, it has come into consideration in the present study that air pollution can be modeled with many temporal and parametric factors (pollutant + meteorology) and a general photograph was formed at time scale. Therefore, in this photograph, a pixel demonstrating the pollutant concentration at t time is estimated by using the effects of itself and other pollutants at t-1 and $t + 1$ time as well as the effects of meteorological parameters.

In this study, PM_{10} and SO_2 concentration values were predicted using CNNs in four different air pollution monitoring stations: Yenibosna, Saraçhane, Kadıköy, and Ümraniye. The

Table 4. Correlation coefficient calculated between all parameters using model study

		Florya Meteorology Stations' Parameters							
AQMS	Pollutants	T	C	RH	P	S	WS	WD	R
Yenibosna	SO_2	−0.306**	0.046	−0.001	0.166**	−0.196**	−0.335**	0.163**	−0.056
	PM_{10}	−0.049	−0.192**	−0.015	0.184**	0.053	−0.332**	−0.029	−0.207**
Sarachane	SO_2	−0.297**	0.115**	0.035	0.323**	−0.274**	−0.219**	0.090*	−0.092*
	PM_{10}	−0.105**	−0.091*	−0.018	0.184**	−0.049	−0.303**	0.002	−0.145**
		Göztepe Meteorology Stations' Parameters							
Ümraniye	SO_2	−0.349**	−0.021	−0.123**	0.318**	−0.196**	−0.419**	0.047	−0.117**
	PM_{10}	−0.094*	−0.219**	−0.134**	0.160**	0.005	−0.425**	0.048	−0.158**
Kadıköy	SO_2	−0.376**	−0.012	−0.139**	0.287**	−0.238**	−0.425**	0.015	−0.134**
	PM_{10}	−0.242**	−0.030	−0.079	0.239**	−0.167**	−0.373**	0.017	−0.080

*: Correlation is significant at the 0.05 level.
**: Correlation is significant at the 0.01 level.

daily future concentrations of these parameters were estimated during 2002–2003 years and the hourly future concentrations were predicted during February and March of 2003. The most important factor in the establishment of the CNN model is neighboring relations. For this reason, we have calculated correlations between meteorological and pollution parameters using the statistical software package SPSS11.5 and results are shown in Table 4. To improve prediction performance, the CNN model was set with side by side high correlation coefficients among the data values.

Wind direction in the European and the Asian part of Istanbul during 2002 and 2003 years is between NW and NE (Figure 2). The pollutions in the study areas are transported mainly from urban center, Bosporus and Black Sea. The high correlation between pollutants and meteorological parameters in the area of the continent of Asia (Ümraniye and Kadıköy) are founded. Especially, the highest negatively correlation between the SO2 concentration and air temperature and wind speed is calculated as 0.349 and 0.425, respectively, in Kadıköy (Table 4). The Asian continent of Istanbul is a densely residential area and the pollutants results from the mainly domestic heating and also traffic. In the European continent, the pollutants may results from the industrial activity in addition to domestic heating and traffic. All of these are differences in the relationship between pollutants and meteorological factors. Wind speed and temperature at the most effective parameter on all the pollutants. Therefore, these parameters when creating CNN matrices were placed close to each other. This system is applied to creating the all CNN matrices.

In the CNN model used in the present study, the elements of input (u) and output (y) matrices are shown in Figure 6. Matrices representing two different views were formed. Data at (t, t + 2, t + 4 …) times were entered to u matrix and thus y matrix representing the following view (t + 1, t + 3, t + 5 …) was estimated. For example, when computing $SO_{2(t+2)}$ in y matrix, we use $SO_{2(t+1)}$, $PM_{10(t+1)}$, $T_{(t+1)}$ and $SO_{2(t+3)}$, $PM_{10(t+3)}$, $T_{(t+3)}$ with the same weight during the computation because it spans the same period of time after and before the $SO_{2(t+2)}$.

We have designed a MATLAB 7.0 code on Pentium IV computers for our CNN model. The CNN training process required approximately 2, 2.55, 3, and 2.15 minutes, respectively, to predict the daily mean concentrations based on data from the Yenibosna, Sarachane, Umraniye, and Kadiköy AQM Stations and 2.45 minutes to predict the hourly mean concentrations based on data from the Yenibosna AQM Station. The processes were stopped when the error reached a value of 2.10^{-4}. Testing of the CNN approach with the optimized A, B and I templates occurred in real time. In training the CNN model using u and y matrices, we obtained A, B and I templates for each study as reported here. To predict the daily mean SO_2 and PM_{10} concentrations in Yenibosna:

$$A = \begin{bmatrix} -0.0015 & 0.0095 & -0.0011 \\ -0.0004 & 1.0257 & -0.0004 \\ -0.0011 & 0.0095 & -0.0015 \end{bmatrix},$$
$$B = \begin{bmatrix} -0.0019 & -0.0014 & -0.0097 \\ -0.0072 & 0.0015 & -0.0072 \\ -0.0097 & -0.0014 & -0.0019 \end{bmatrix},$$
$$I = [0.0015] \tag{12}$$

To predict the daily mean SO_2 and PM_{10} concentrations in Sarachane:

$$A = \begin{bmatrix} -0.0016 & 0.0011 & 0.0052 \\ 0.0015 & 1.0034 & 0.0015 \\ 0.0052 & 0.0011 & 0.0016 \end{bmatrix},$$
$$B = \begin{bmatrix} -0.0020 & -0.0027 & -0.0028 \\ -0.0024 & 0.0012 & -0.0024 \\ -0.0028 & -0.0027 & -0.0020 \end{bmatrix},$$
$$I = [0.0012] \tag{13}$$

To predict the daily mean SO_2 and PM_{10} concentrations in Ümraniye:

$$A = \begin{bmatrix} -0.0051 & 0.0061 & -0.0043 \\ -0.0021 & 1.0133 & -0.0021 \\ -0.0043 & 0.0061 & -0.0051 \end{bmatrix},$$
$$B = \begin{bmatrix} -0.0017 & -0.0019 & -0.0022 \\ -0.0022 & 0.0047 & -0.0022 \\ -0.0022 & -0.0019 & -0.0017 \end{bmatrix},$$
$$I = [0.0047] \tag{14}$$

To predict the daily mean SO_2 and PM_{10} concentrations in Kadıköy:

$$A = \begin{bmatrix} -0.0010 & 0.0064 & -0.0069 \\ -0.0046 & 1.0206 & -0.0046 \\ -0.0069 & 0.0064 & -0.0010 \end{bmatrix},$$
$$B = \begin{bmatrix} -0.0042 & -0.0036 & -0.0036 \\ -0.0034 & 0.0035 & -0.0034 \\ -0.0036 & -0.0036 & -0.0042 \end{bmatrix},$$
$$I = [0.0035] \tag{15}$$

To predict the hourly mean SO_2 and PM_{10} concentrations in Yenibosna:

$$A = \begin{bmatrix} -0.0012 & -0.0017 & -0.0021 \\ -0.0021 & 1.0244 & -0.0046 \\ -0.0021 & -0.0017 & -0.0012 \end{bmatrix},$$
$$B = \begin{bmatrix} -0.0021 & -0.0022 & -0.0021 \\ -0.0020 & -0.0023 & -0.0020 \\ -0.0021 & -0.0022 & -0.0021 \end{bmatrix},$$
$$I = [-0.0023] \tag{16}$$

Here, neighborhood (r) is chosen as 1. To guarantee stability of the CNN, the templates are symmetric. We have replaced the template values obtained in Equation 12–16 with those from Equations (2–3). In the optimization process, all template coefficients were chosen to four decimal precisions. Linear region of the piece wise non-linear function was especially chosen as in Figure 5. Thus, the multilevel CNN outputs were obtained between -1 and $+1$ values. Furthermore, the CNN output values were mapped to real range of 0–250 $\mu g/m^3$ for SO_2 and 0–500 $\mu g/m^3$ for PM_{10} using a metric system. As a result, we have reached precise results that are relatively close to the desired concentrations. Hence, the CNN prediction results were first compared with the PER. Where the PER method consists of a very simple prediction, daily air pollution concentrations of two consecutive days are assumed to be the same.

The correlation coefficients obtained after training the CNN were calculated. The correlation coefficient results were obtained between 0.81 and 0.90 for all CNN models. When all of the model training results are evaluated in general, the correlation coefficient of the CNN training results and real data had much higher than the PER approach. The similar results were reported by Şahin et al. (2011).

The data set was tested using the A, B, I (Equations 12–16) in terms of the CNN model obtained after training. The real daily mean concentrations of SO_2 (Figure 7) and PM_{10} (Figure 8) were compared to the predictions of CNN and PER models in all AQM stations for both 2002 and 2003. Furthermore, statistical evaluations of frequency-residuals versus for CNN and PER were derived as in Figure 9 for the daily mean SO_2 and in Figure 10 for the daily mean PM_{10}. Results were found satisfactory. In all AQM Stations, SO_2 residuals for CNN prediction values alter between -70 and $+110$; however, they were between -90 and $+100$ for PER model. The maximum percentage of "0" residual value for SO_2 is observed in Kadıköy and the minimum is observed in Yenibosna. Different pollution sources may affect this special case. Yenibosna AQM station is affected by all pollution sources, whereas Kadiköy AQM station is affected only by home-originated pollutants (Table 1). It is known that SO_2 is generally derived from domestic heating depending on the change in meteorological conditions. To increase achievement level of CNN model in Yenibosna, other effective parameters such as traffic and industrial activities are added to model structure. The same result is not found for $PM_{10.}$ It was concluded that PM_{10} is resulted from different sources.

The relevant levels of daily mean SO_2 and PM_{10} concentrations, according to EU legislation (see the EC Normative-Council directive 1999/30/EC of 22 April 1999 relating to limit concentration limits for sulfur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air) are 125 $\mu g/m^3$ and 50 $\mu g/m^3$, respectively, and are not to be exceeded more than three and 35 times a year, respectively. The environmental laws in Turkey are being revised according to guidelines European Union. When Draft Air Pollution Control Laws are considered, it will be necessary to assert the EU limit values. Approximately 61% and 30% in Kadıköy, 68% and 71% in Saraçhane, 52% and 43% in Ümraniye, 68% and 65% in Yenibosna of observed PM_{10} concentration values in winter and summer seasons during the 2002 and 2003, respectively, are higher than the limit values. This situation was observed in the CNN model prediction, and the studies yielded predictions with 90% success. However, the exceeding the attention level of SO2 concentration are observed one day in Kadıköy and Yenibosna, five days in Saraçhane, two days in Ümraniye, all in winter. The exceeding the attention level of SO2 were predicted by CNN and PER with 5% and% 66 error in Kadıköy, Yenibosna and Ümraniye, respectively. This situation is not same in Saraçhane. As in Figure 7, however, the high SO_2 concentrations trend close to the CNN prediction trend.

Figure 11 displayed the scatter plots of the observed versus the predicted seasonal (winter and summer) PM_{10} and SO_2 concentration levels at the Air Quality Measurements Stations of Yenibosna, Sarachane, Umraniye, and Kadiköy during 2002-2003 years for CNN model. The correlation coefficients

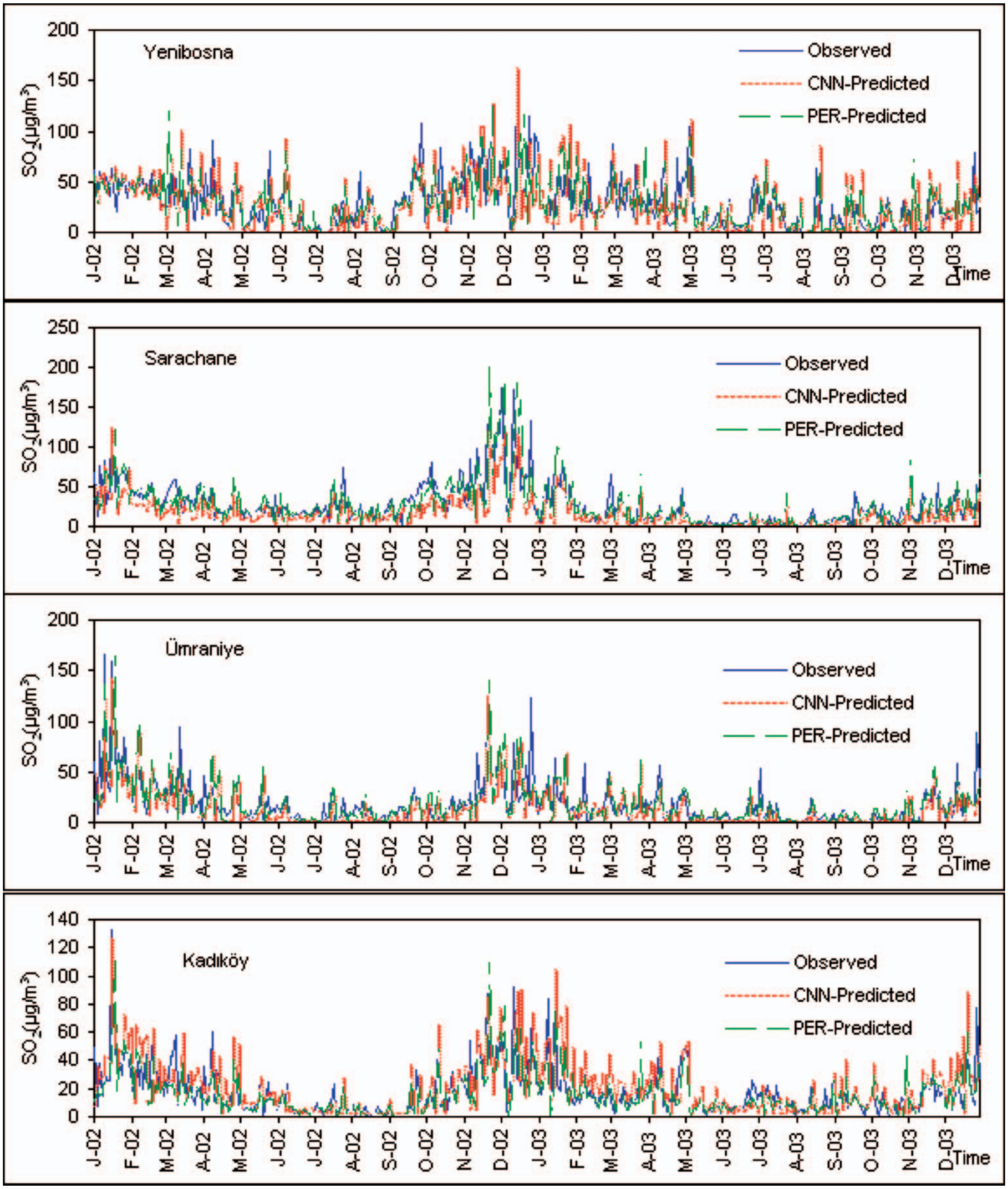

Figure 7. Graphs of the daily mean time series of the observed and predicted concentrations of SO_2 at the Air Quality Measurement Stations of Yenibosna, Sarachane, Ümraniye, and Kadıköy during 2002 and 2003 for two models (cellular neural network [CNN] and statistical persistence method [PER]). (Color figure available online.)

between observed and predicted in winter and summer are varied. Generally, CNN provides the most reliable predictions of the daily SO_2 concentration levels in winter seasons for Saraçhane and Ümraniye stations. The source of SO2 in these areas is only commercial heating. However, SO_2 could emit in the atmosphere from commercial, industry, airport etc. in Yenibosna and from commercial, sea traffic etc. in Kadıköy. Also, CNN provides a little more reliable predictions of the daily PM_{10} concentration levels in winter season for Saraçhane, Ümraniye and Kadıköy. The mean PM_{10} concentrations in summer and winter season are not difference as in Table 3. The main source of PM_{10} in Yenibosna is traffic and industry. These result

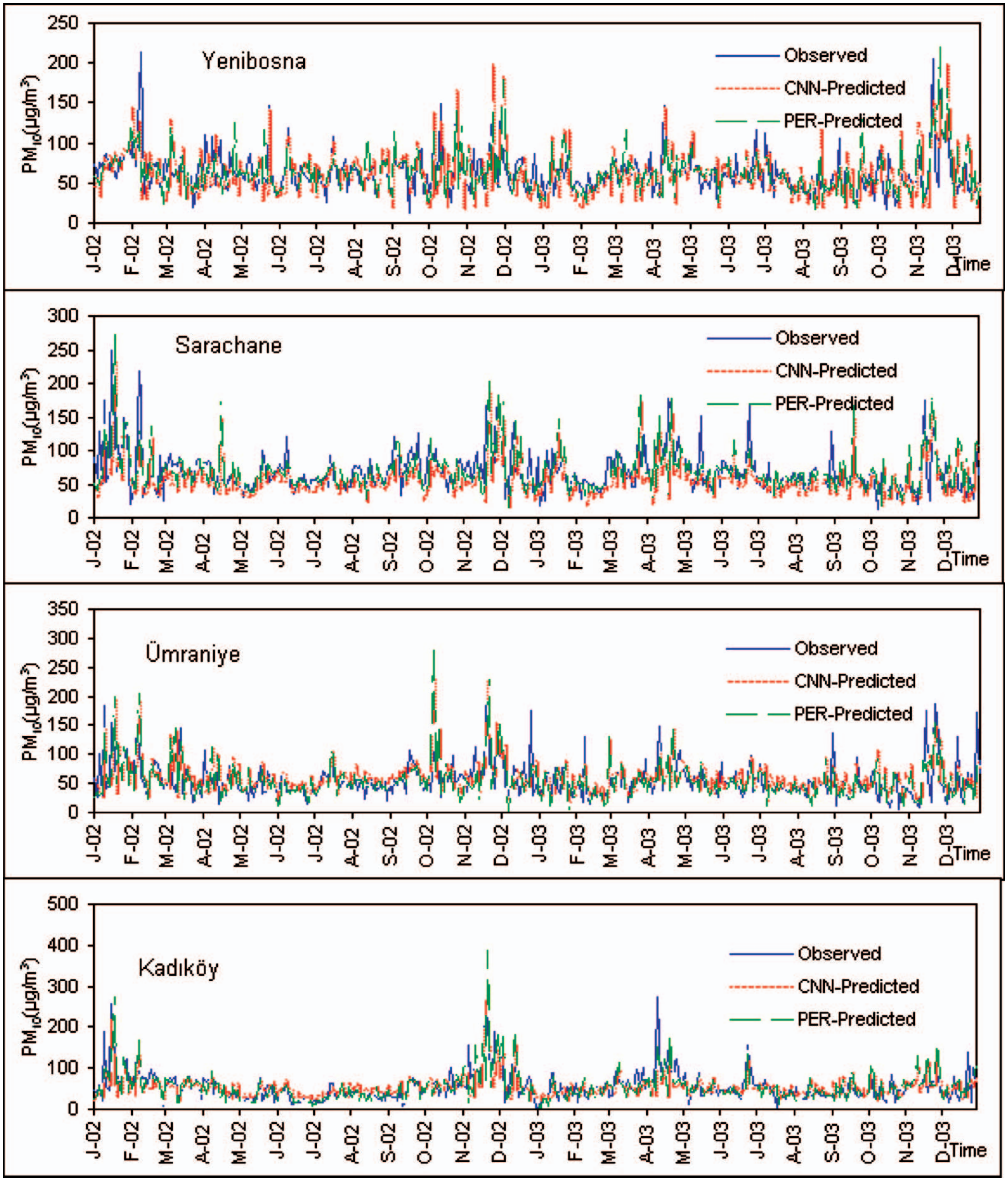

Figure 8. Graphs of the daily mean time series of the observed and predicted concentrations of PM_{10} at the Air Quality Measurement Stations of Yenibosna, Sarachane, Ümraniye and Kadıköy during 2002 and 2003 for two models cellular neural network [CNN] and statistical persistence method [PER]). (Color figure available online.)

shows that the CNN model can predict seasonal SO_2 differences clearly, cannot predict PM_{10} differences. The concentrations of SO_2 in winter are two times higher in summer.

The CNN and PER model results were also checked by calculating five different statistical indices, given in Equation 7–11, which are based on the deviations between predicted values and original observations. The final results of statistical model evaluation for the daily mean SO_2 and PM_{10} concentrations during 2002 and 2003 years are presented in Table 5. For both pollutants, the results are separately presented for each AQM

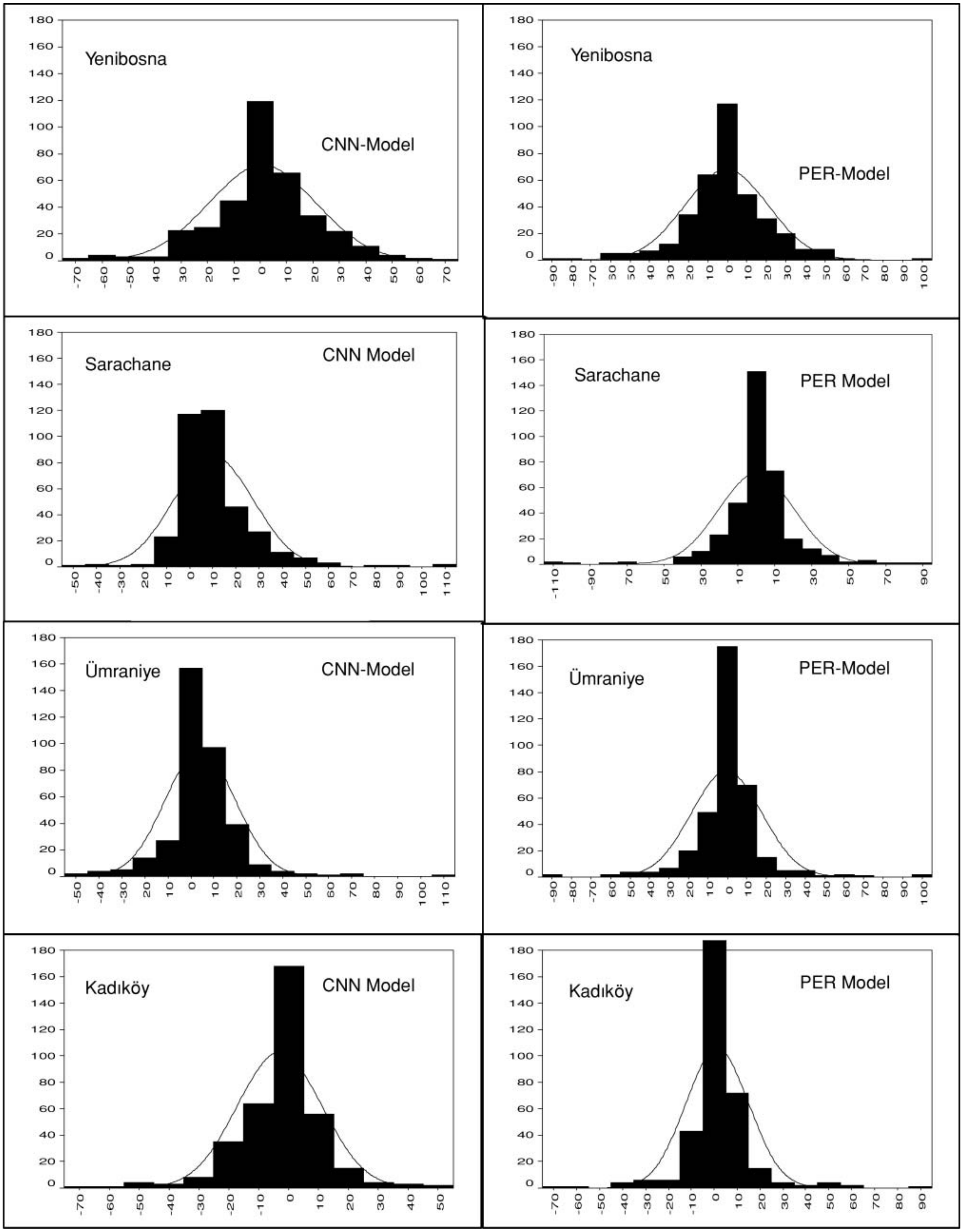

Figure 9. Graphs of the frequency distribution of residuals of the cellular neural network (CNN) and statistical persistence method (PER) model prediction for the daily mean concentrations of SO_2 at the Air Quality Measurement Stations of Yenibosna, Sarachane, Ümraniye, and Kadıköy during the years 2002 and 2003. Residual: observed–predicted.

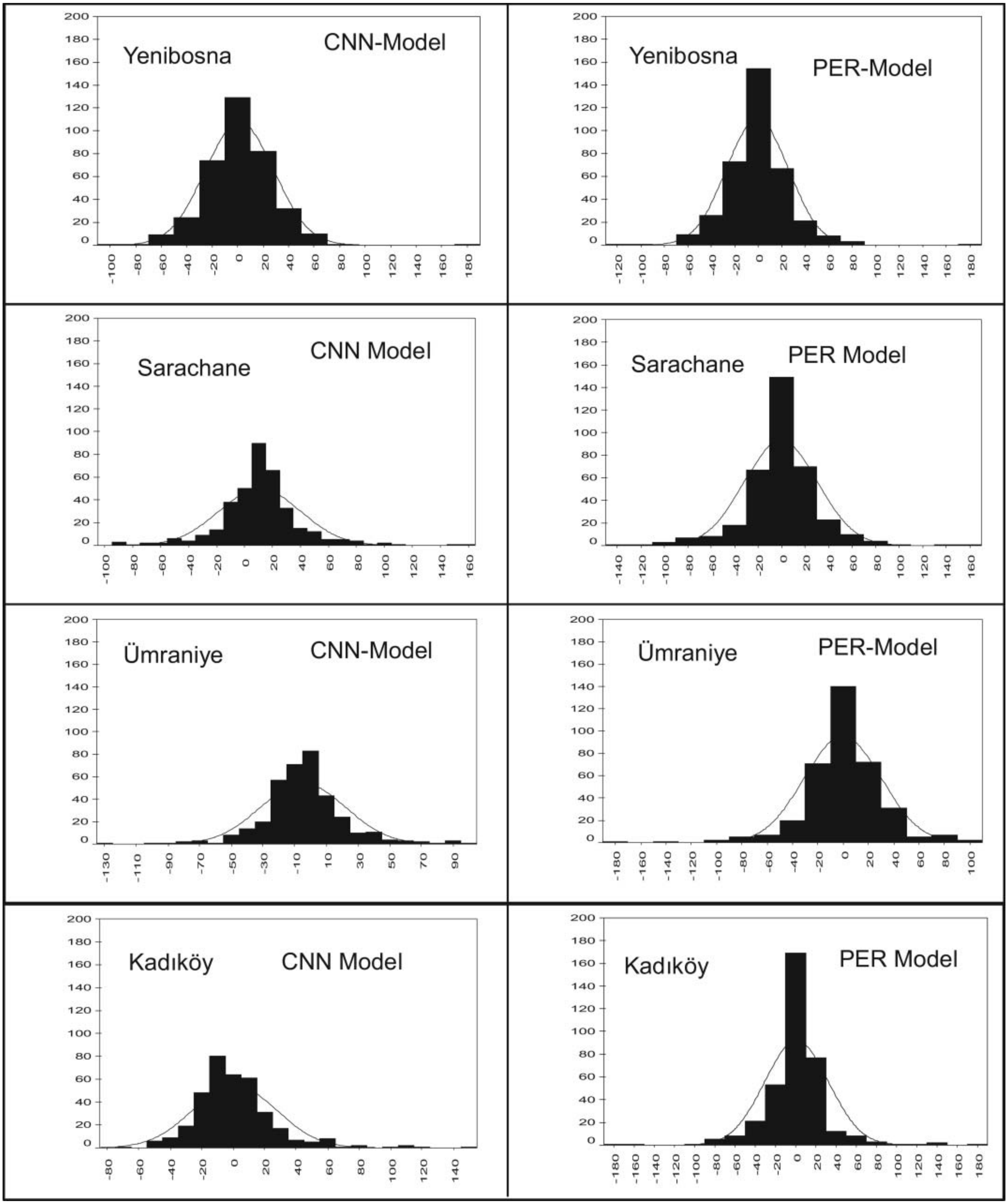

Figure 10. Graphs of the frequency distribution of residuals of the cellular neural network (CNN) and statistical persistence method (PER) model prediction for the daily mean concentrations of PM_{10} at the Air Quality Measurement Stations of Yenibosna, Sarachane, Umraniye, and Kadiköy during the years 2002 and 2003. Residual: observed–predicted.

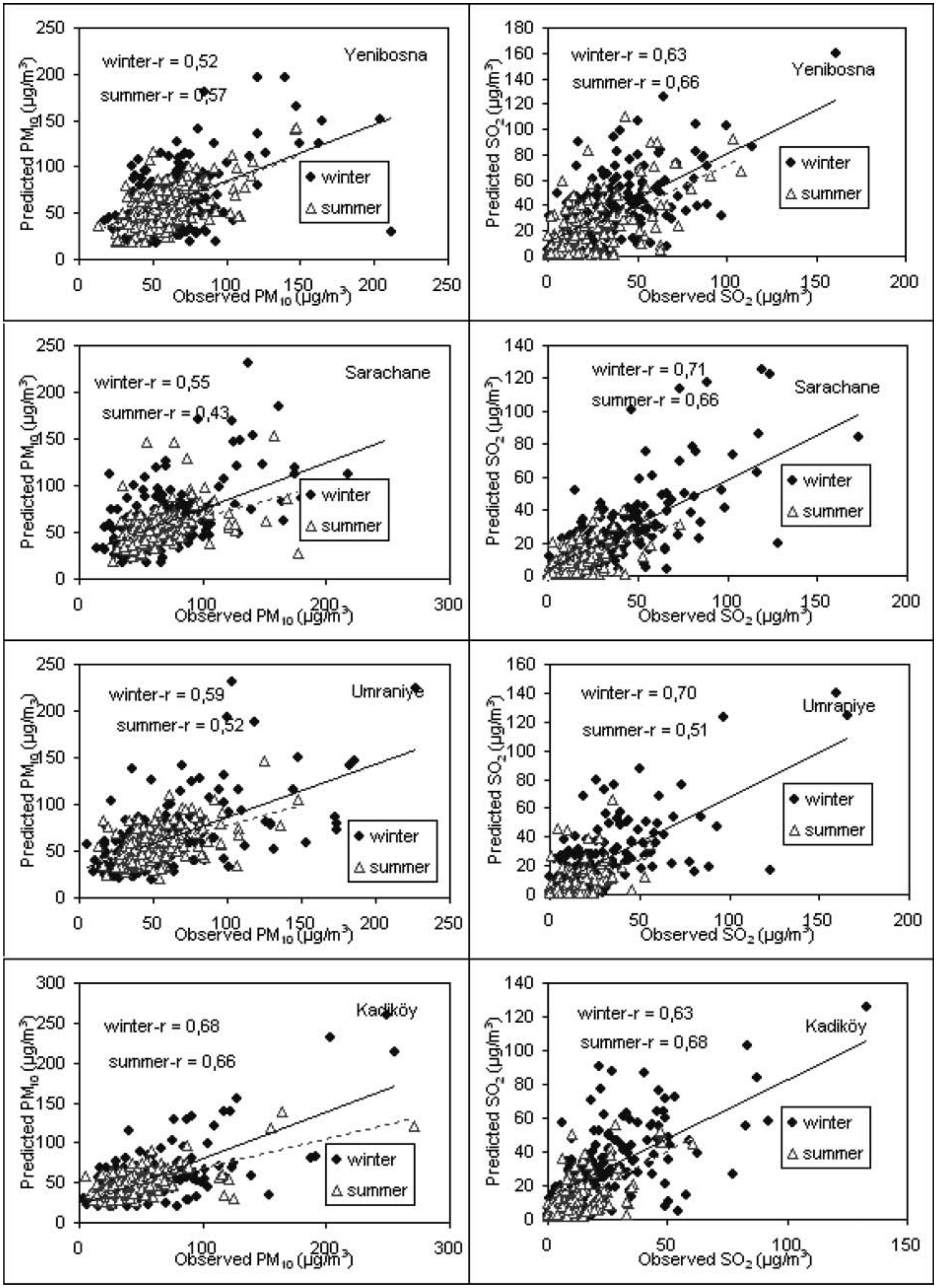

Figure 11. Scatter plots of the observed versus the predicted seasonal (winter and summer) PM10 and SO2 concentration levels at the Air Quality Measurements Stations of Yenibosna, Sarachane, Umraniye, and Kadiköy during the years 2002 and 2003 for the cellular neural network (CNN) model.

Table 5. *Statistical performance indices between daily mean concentrations of the observed and predicted SO_2 and PM_{10} at the four Air Quality Measurement Stations for cellular neural network (CNN) and statistical persistence method (PER) models*

Pollutant	AQMS	Model	Max	Min	Mean	σ	r	d	Bias	MAE	RMSE	Difference[1]—t test (p value)
PM_{10}	Yenibosna	CNN	196.6	17.9	61.2	29.4	0.54	0.73	1.44	19.7	26.8	1.04(0.297)
		PER	231.8	16.3	63.4	26.3	0.51	0.71	−0.76	17.9	26.1	−0.58(0.560)
	Saraçhane	CNN	231.4	16.5	58	27.4	0.52	0.71	10.84	21.8	30.9	7.15(0.000)
		PER	274.3	16.5	69.5	31.2	0.51	0.71	−0.71	20.3	31.0	−0.43(0.667)
	Ümraniye	CNN	231.7	18.8	60.3	28.1	0.58	0.75	−4.73	19.0	27.1	−3.36(0.001)
		PER	287.4	0	56	32.7	0.54	0.72	−0.36	20.4	30.2	−0.24(0.807)
	Kadıköy	CNN	260	19.1	53.6	26.5	0.67	0.80	0.89	18.7	26.4	0.77(0.443)
		PER	385.3	4.2	53.6	37.4	0.63	0.78	0.86	19.1	31.4	0.58(0.559)
SO_2	Yenibosna	CNN	160	1	27.9	26.2	0.67	0.82	0.89	14.4	20.2	0.85(0.394)
		PER	125.5	0.3	29.2	23.2	0.60	0.77	−0.46	14.6	21.3	−0.44(0.657)
	Saraçhane	CNN	125.7	0.4	17.2	19.4	0.76	0.81	10.4	13.1	20.0	11.58(0.000)
		PER	202.7	0	27.4	28.7	0.73	0.82	0.19	12.2	20.3	−0.17(0.861)
	Ümraniye	CNN	140	0.4	14	19.2	0.73	0.84	4.05	9.8	15.5	5.25(0.000)
		PER	163.9	0	18.4	21.3	0.64	0.79	−0.42	10.4	17.9	−0.44(0.658)
	Kadıköy	CNN	125.6	2.1	20.3	19.6	0.71	0.83	−3.02	9.4	14.4	−4.10(0.000)
		PER	113.8	0	15.9	15.8	0.65	0.79	1.30	7.8	13.7	1.79(0.074)

[1]Between observed and predicted data.

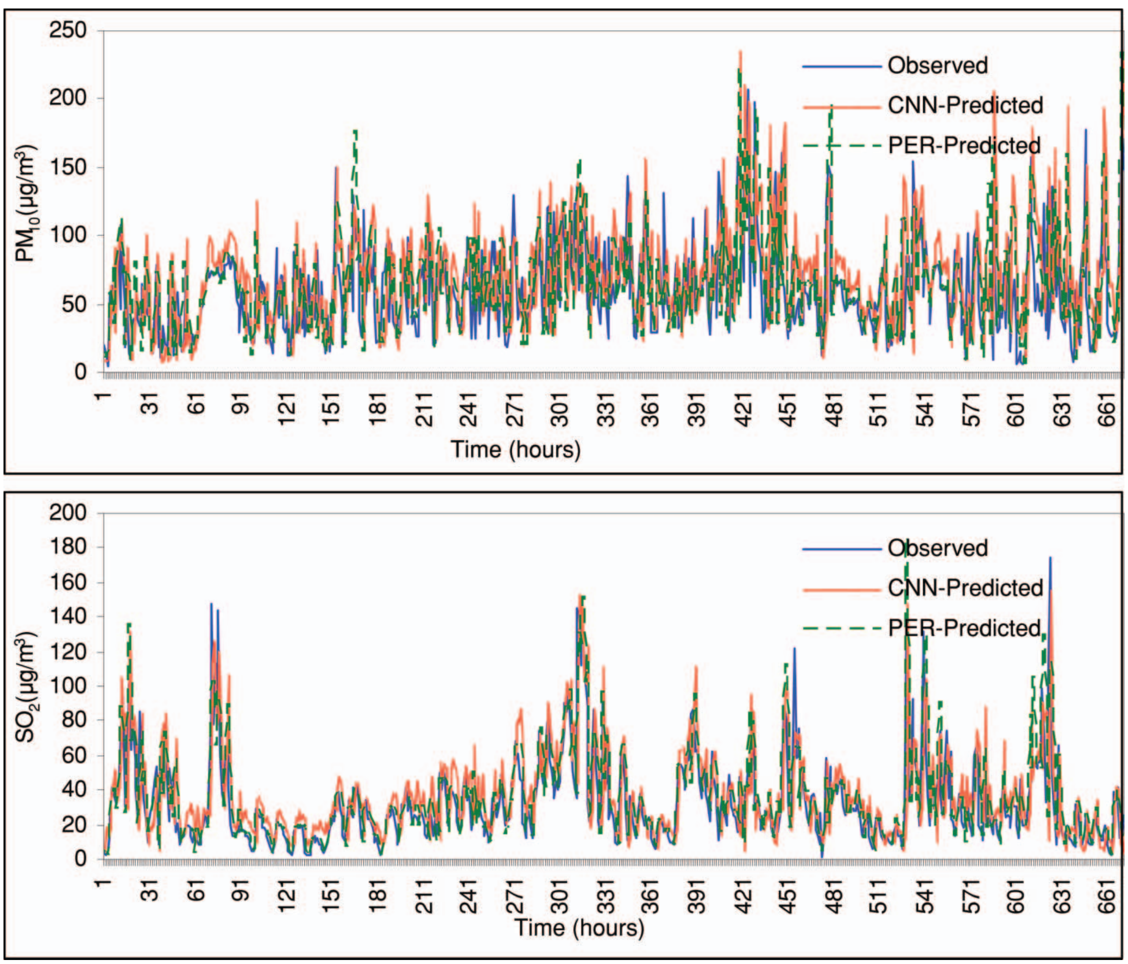

Figure 12. Graphs of the hourly mean time series of the observed and predicted concentrations of PM_{10} and SO_2 at the Yenibosna Air Quality Measurement Stations during February and March 2003 for two models (cellular neural network [CNN] and statistical persistence method [PER]). (Color figure available online.)

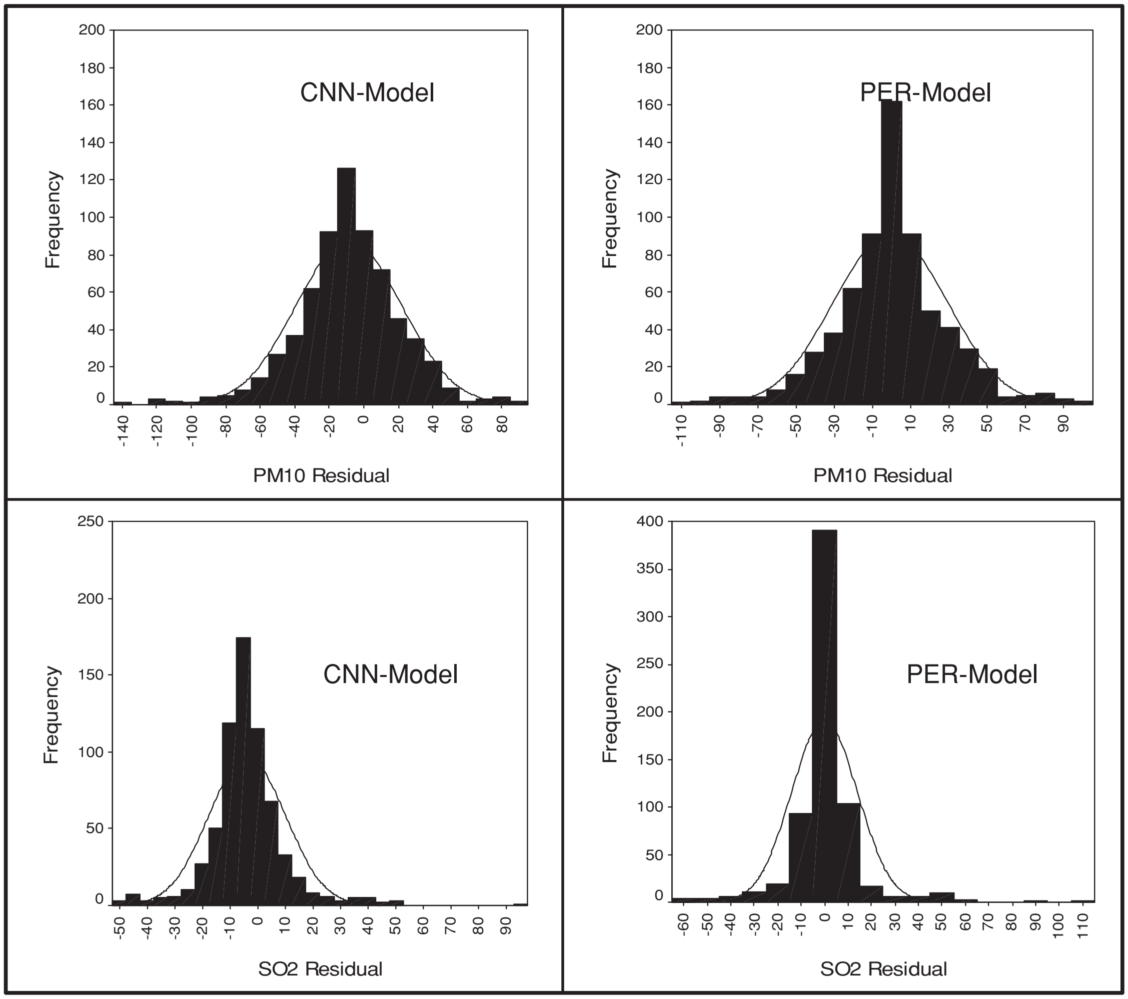

Figure 13. Graphs of the frequency distribution of residuals of the cellular neural network (CNN) and the statistical persistence method (PER) model prediction for the hourly mean concentrations of PM_{10} and SO_2 at the Yenibosna Air Quality Measurement Stations of during February and March 2003. Residual: observed–predicted.

Stations. For PM_{10} and SO_2 in Sarachane AQM Station, CNN model prediction *Bias* value reached approximately to 10. This result demonstrates that the observed concentration is less than the predicted concentration. The indices of agreement in CNN prediction were found between 0.71 and 0.80 for PM_{10}, and 0.81 and 0.84 for SO_2, in all AQM Stations. Nevertheless, the indices of agreement in PER prediction were found less than CNN prediction, between 0.71–0.78 for PM_{10} and 0.77–0.82 for SO_2. For PM_{10}, the maximum index of agreement (*d*) is 0.80 for CNN model and 0.78 for PER model in Kadıköy AQM Station.

The correlation between observed and CNN predicted daily mean PM_{10} and SO_2 data for only commercial site, Kadıköy is 0.67 and 0.71, respectively and for PER predicted data, it is 0.63 and 0.65, respectively. However, the correlation between observed and CNN predicted PM_{10} and SO_2 data for all commercial, industrial and traffic sites, Yenibosna are 0.54 and 0.67, respectively and it is 0.51 and 0.60, respectively for PER predicted data. If all the correlation coefficient (*r*) is evaluated, *r* values of SO_2 prediction become higher than the PM_{10} prediction. For SO_2 pollutant, the correlation coefficient (r) of CNN was found between 0.67 and 0.76 for all AQM stations. In the same durations, 0.60 and 0.73 values are calculated for PER.

The hourly mean future PM_{10} and SO_2 concentrations were estimated during February and March 2003 periods in Yenibosna AQM Stations. CNN and PER, predicted and actual hourly mean concentrations of PM_{10} and SO_2 were compared as in Figure 12. In addition, statistical evaluations of residual frequency distribution in Figures 13 were made for the hourly mean SO_2 and PM_{10}. The final results of statistical model evaluation for the hourly mean SO_2 and PM_{10} concentrations are presented in Table 6. The indices of agreement in CNN prediction were found 0.78 and 0.92 for PM_{10} and SO_2, respectively. Nevertheless, the indices of agreement of PER prediction were found less than CNN prediction, 0.77 for PM_{10} and 0.91 for SO_2.

To compare results of CNN and PER models, the statistical significance was determined by the student's *t* test and shown in Table 5 and Table 6. The daily mean PM_{10} and SO_2 concentrations modeling by the CNN had a statistically significant

Table 6. Statistical performance indices between hourly mean concentrations of the observed and predicted SO_2 and PM_{10} at the Yenibosna Air Quality Measurement Station for cellular neural network (CNN) and statistical persistence method (PER) models.

	PM_{10}		SO_2	
Statistical performance indices	CNN	PER	CNN	PER
Maximum	233.7	234.0	154.0	187.0
Minimum	8.3	6.0	3.9	2.0
Mean	69.1	61.1	36.6	32.9
σ	36.4	32.6	24.7	24.2
R	0.61	0.59	0.85	0.83
d	0.78	0.77	0.92	0.91
Bias	−8.70	−0.76	−3.77	−0.05
MAE	23.5	20.7	9.9	7.8
RMSE	31.4	29.0	14.2	13.9
Difference[1]— t test (p value)	−7.47(0.000)	−0.67(0.503)	−7.11(0.000)	−0.09(0.927)

[1]Between observed and predicted data.

difference from observed mean concentrations at a confidence level 99% in the Saraçhane and the Ümraniye air monitoring stations. Also in Kadıköy, the daily mean SO_2 concentrations modeling by the CNN had a statistically significant difference (Table 5). Such differences for SO_2 are observed particularly in the winter period. Whereas, no significant differences the result of CNN model and observed data for all pollutant were observed in Yenibosna. The PER prediction determined in all stations had no significant differences. The PER approaches do not use meteorological or pollutant parameters for modeling. The concentration differences between the observed and the predicted by CNN would be stemming from the differences in the geographical and the climatic conditions of the regions as well as the changes in the meteorological conditions. In the air monitoring stations close to the meteorological stations (Yenibosna and Kadıköy), the CNN model have predicted more accurately. Air pollutants can be more accurately represented by meteorological parameters. If all parameters measured in the same point, CNN could be more successful.

Conclusion

In this study, the major air pollutants of concern for the city of Istanbul, PM_{10} and SO_2, are estimated using a CNN approach. There are many computational methods available for air pollutant modeling. One of the frequently used methods is the use of an ANN. In ANN, the training process time increases as the problem becomes increasingly complex. To reduce the complexity of the calculations used by the ANN, Chua and Yang introduced CNN in 1988 as a new non-linear, dynamic neural network structure. In a CNN, the correlations between neighboring pixels are modeled by cloning templates with a limited number of elements and using these pixels for solving complex problems.

Here, the daily and hourly mean concentrations of PM_{10} and SO_2 air pollutants in Istanbul are modeled. The forthcoming daily air pollutant values are predicted by CNN during 2002 and 2003 and hourly values during February and March 2003. Comparing the results obtained using CNN model with those obtained using PER technique; we observed that the CNN model provides more reliable predictions. In previous similar ANN modeling studies, the correlation coefficient values ranged between 0.50 and 0.80 (Mok and Tom, 1998; Chelani et al., 2002; Sahin et al., 2005; Hooyberghs et al., 2005; Slini et al., 2005, Kurt et al., 2008). In this study, for CNN model, r value was measured between 0.51-0.63 for the daily mean PM_{10} and 0.60-0.76 for the daily mean SO_2. Additionally, r value was measured as 0.61 for the hourly mean PM_{10} and 0.85 for the hourly mean SO_2 concentrations.

These results show that the CNN modeling technique can be considered a promising approach for air pollutant prediction. We have proposed a new method for modeling the air-pollution problem using a CNN. In addition, we propose to test the ability of CNN models to model other environmental pollution problems. We specifically propose to apply CNN methods to 3D air pollution modeling problems in the future.

Acknowledgement

We are grateful to Istanbul Municipality, Environmental Protection Directorate and Istanbul Regional Directorate of Meteorology for their help in obtaining real data. This work was supported by the Research Fund of the Istanbul University. Project number: T-486/25062004.

References

Albora, A. M., Ucan, O. N., Ozmen, A., and Ozkan, T. 2001. Separation of Bouguer anomaly map using cellular neural network. *Journal of Applied Geophysics* 46:129–142.

Arena, P., Caponetto, R., Fortuna, L., and Manganaro, G. 1997. Cellular neural network to explore complexity. *Soft Computing* 1:120–136.

Boznar M., Lesjak M., and Malker P. 1993. A neural network based method for short-term predictions of ambient SO2 concentrations in highly polluted industrial areas of complex terrain. *Atmospheric Environment* 27B (2): 221–230.

Cimagalli, V. 1993. Cellular neural networks: A review. *Proceedings of the Sixth Italian workshop on parallel architectures and Neural Networks* (May 12–14, Vietri Sul Mare, Italy) Ed. E. Caianiello.

Chelani A. B., Chalapati Rao C. V., Phadke K. M., and Hasan M. Z. 2002. Prediction of sulfur dioxide concentration using artificial neural networks. *Environmental Modeling & Software* 17:161–168.

Chua L. O., and Yang L. 1988. Cellular neural networks: Application. *IEEE Transactions on Circuits and Systems* 35(10): 1273–1290.

Dıngenen, R. V., Raes, F., and Putaud, J.-P. et al. 2004. A European aerosol phenomenology–1; physical characteristics of particulate matter at kerbside, urban, rural and background sites in Europe. *Atmospheric Environment* 38:2561–2577.

Esen, F., Taşdemir, Y., and Cindoruk, S. S. 2005. Evaluation of NOx and O3 concentrations in the atmosphere of Bursa, Turkey. *Environmental Forensics* 6(3): 311–317.

Gaga, E. O., Tuncel, G., and Tuncel, S. G. 2009. Sources and wet deposition fluxes of polycyclic aromatic hydrocarbons (PAHs) in an urban site 1000 meters high in center Anatolia (Turkey). *Environmental Forensics* 10(4): 286–298.

Gardner M. W., and Dorling S. R. 1998. Artificial neural networks (the Multilayer Perceptron)–A review of applications in the atmospheric sciences. *Atmospheric Environment* 32:2627–2636.

Grassi, G., and Grieco L. A. 2002. Object-oriented image analysis via analogic cnn algorithms–Part I: Motion estimation. In *Proceedings of the 7th IEE International Workshop on Cellular Neural Networks and Their Applications*, Frankfurt, Germany, pp. 172–180.

Guzelis, C., and Karamahmut, S. 1994. Recurrent perceptron learning algorithm for completely stable cellular neural networks. In *Proceedings the Third IEEE Int. Workshop on Cellular Neural Network and Applications*, Rome, Italy, 18–21 December, pp. 177–182.

Hadad, K., and Piroozmand, A. 2007. Application of cellular neural network (CNN) method to the nuclear reactor dynamics equations. *Annals of Nuclear Energy* 34:406–416.

Hooyberghs, J., Mensink, C., Dumont, G., et al. 2005. A neural network forecast for daily average PM10 concentrations in Belgium. *Atmospheric Environment* 39(18): 3279–3289.

Im, U., Tayançi M., and Yenigün, O. 2008. Interaction patterns of major photochemical pollutants in Istanbul, Turkey. *Atmospheric Research* 89:382–390.

Junninen, H., Niska, H., Tuppurainen, K., et al. 2004. Methods for imputation of missing values in air quality data sets. *Atmospheric Environment* 38:2895–2907.

Karaca, F., Anil, I., and Alagha, O. 2009. Long-range potential source contributions of episodic aerosol events to PM10 profile of a megacity. *Atmospheric Environment* 43:5713–5722.

Kındap T. 2008. Identifying the trans-boundary transport of air pollutants to the city of istanbul under specific weather conditions. *Water Air Soil Pollution* 189:279–28.

Kukkonen, J., Partanen, L., Karppinen, A., et al. 2003. Extensive evaluation of neural network models for the prediction of NO2 and PM10 concentrations, compared with a deterministic modeling system and measurements in central Helsinki. *Atmospheric Environment* 37:4539–4550.

Kurt, A., Gülbağcı, B., Karaca, F., and Alagha, O. 2008. An online air pollution forecasting system using neural networks. *Environmental International* 34:592–598.

Mok K. M., and Tam S. C. 1998. Short-term prediction of SO2 concentration in Macau with artificial neural networks. *Energy and Buildings* 28:279–286.

Nagendra, S. M. S., and Khare, M. 2006. Artificial neural network approach for modelling nitrogen dioxide dispersion from vehicular exhaust emissions. *Ecological Modelling* 190:99–115.

Nunnari, G., Dorling, S., Schlink, U., et al. 2004 Modelling SO2 concentration at a point with statistical approaches. *Environmental Modelling & Software* 19(10): 887–905.

Ordieres, J. B., Vergara, E. P., Capuz, R. S., and Salazar, R. E. 2005. Neural network prediction model for fine particulate matter (PM2.5) on the US-Mexico border in El Paso (Texas) and Ciudad Juarez (Chihuahua). *Environmental Modeling & Software* 20:547–559.

Ozcan, H., Bilgili, E., Sahin, U., Ucan, O. N, and Bayat, C. 2007. Modelling of tropospheric ozone concentrations using genetically trained multilevel cellular neural networks. *Advances in Atmospheric Science* 24(5): 907–914.

Perez, P., and Reyes, J. 2006. An integrated neural network model for PM10 forecasting. *Atmospheric Environment* 40:2845–2851.

Raha, D. 2007 Comparative study of artificial network modelling in predicting effluent oil and grease from coastal primary and chemically assisted primary sewage treatment plants. *Environmental Forensics* 8:371–381.

Saral, A., and Ertürk, F. 2003. Prediction of ground level SO2 concentrations using artificial neural network. *Water, Air and Soil Pollution* 3:297–306.

Sahin, U., Ucan, O. N., Soyhan, B., and Bayat, C. 2004 Modelling of CO distribution in Istanbul using artificial neural networks. *Fresenius Environmental Bulletin* 13(9): 839–845.

Sahin, U., Ucan, O. N., and Bayat, C. 2005. Modelling of SO2 distribution in Istanbul using artificial neural networks. *Environmental Modeling and Assessment* 10:135–142.

Sahin, Ü, Bayat, C., and Ucan, O. N. 2011. Application of cellular neural network (CNN) to the prediction of missing air pollution data. *Atmospheric Research* 101:314–326.

Slini, T., Kaprara, A., Karatzas, K., and Moussiopoulos, N. 2006. PM10 forecasting for Thessaloniki, Greece. *Environmental Modelling & Software* 21(4): 559–565.

Ucan, O. N., Bilgili, E., and Albora, A. M. 2001. Detection of buried objects on archeological areas using genetic cellular neural network. In *Proceedings of the European Geophysical Society XXVI General Assembly 3*, France, pp. 223–231.

Tayanç, M. 2000. An assessment of spatial and temporal variation of sulphur dioxide levels over Istanbul, Turkey. *Environmental Pollution* 107:61–69.

Thai, V. D., and Cat, P. T. 2008. Modelling air-pollution problem by cellular neural network. In *Proceedings of the 10th International Conference on Control, Automation, Robotics and Vision (December 17–20*, 2008, Hanoi, Vietnam)., pp. 1115–1118.