We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Spatial Interpolation Methodologies in Urban Air Pollution Modeling: Application for the Greater Area of Metropolitan Athens, Greece

Despina Deligiorgi and Kostas Philippopoulos
National and Kapodistrian University of Athens
Greece

1. Introduction

Air pollution in urban environments has serious health and quality of life implications. A wide variety of anthropogenic air pollution sources increase the levels of background air pollutant concentrations, leading to the deterioration of the ambient air quality. Principal sources of urban air pollution are vehicular traffic, industrial activity and in general fossil fuel combustion, introducing a mixture of chemical components, particulate matter and biological material into the atmosphere. The deterioration of urban air quality is considered worldwide one of the primary environmental issues and current scientific evidence associate the exposure to ambient air pollution with a wide spectrum of health effects like cardiopulmonary diseases, respiratory related hospital admissions and premature mortality (Analitis et al. 2006; Ito et al., 2005; Samet et al., 2000). Direct measurements of sensitive population groups' exposure to air pollution are scarce and therefore methods of accurate point and areal air quality estimations are prerequisite. This fact highlights the importance of generating accurate fields of air pollution for quantifying present and future health related risks.

In the field of air pollution modeling, two different approaches have been adopted by the scientific community, differentiated by their applied fundamental principles. The first approach involves the numerical simulation of atmospheric dispersion based on the current understanding of physics and chemistry that govern the transport, dispersion and transformation of pollutants in the atmosphere. The modeling process typically requires a set of parameters such as meteorological fields, terrain information along with a comprehensive description of pollution sources. An alternative approach is based on statistical analysis of pollutant concentrations collected from air quality monitoring networks commonly deployed in urban areas. The reasoning of the statistical approach is that physical processes are likely to induce correlations in air quality data collected over space and time. Statistical models generate predictions by exploiting these spatio-temporal patterns, enabling the estimation of pollutant concentrations in unmonitored locations.

The chapter's main objective is to present and review the statistical spatial interpolation methodologies which are commonly employed in the field of air pollution modeling. An additional scope of the chapter is to compare and evaluate the accuracy of the interpolation methods for point estimations, using data from a real urban air quality monitoring network located at the greater area of metropolitan Athens in Greece.

2. Spatial interpolation methodologies

Spatial interpolation is the procedure of estimating the values of the variable under study at unsampled locations, using point observations within the same region. Statistical interpolation methodologies are applied in air pollution modeling for estimating the spatial distribution of pollutants, based on data provided from an existing air quality monitoring network. Such networks consist of a number of irregularly distributed stations and therefore the issue of air pollution spatial interpolation is essentially a problem in the field of scattered data approximation. The monitoring sites are typically located to detect high concentrations and although such a configuration is appropriate for identifying the maximum potential exposure and risk, in many cases it fails to describe the spatial variability of air pollution. Recent developments in the design and the modification of existing air quality monitoring networks address this issue (Lozano et al., 2009; Mofarrah & Husain, 2010; Nejadkoorki, 2011) and propose methodologies for an optimum configuration of monitoring sites used for more reliable and cost-effective spatial predictions.

The spatial interpolation methodologies addressed in this review consist of some of the most widely used interpolation schemes and can be classified according to their inherent general features. The following classifications and characteristics have been adopted in a number of spatial interpolation reviews (Eberly et al., 2004; Li & Heap 2008) as they serve the purpose of illustrating the suitability of each methodology for the spatial process of interest.

- *Point – Areal*. Point based interpolation methodologies produce estimates in discrete unsampled locations based on point observations of the variable under study, whereas the areal based schemes estimations refer to a whole predefined area or zone.
- *Global – Local*. Global methodologies utilize all available data of the variable under study and develop a global estimation function for the study area, while local schemes segment the area and operate independently within a window, using only a fraction of the available data.
- *Exact – Approximate*. The methodologies which produce an estimation surface that essentially passes from the observed data are called exact interpolators, while approximate interpolators make use of the available data to produce the estimation surface, which is not required to predict exactly the observed values.
- *Stochastic – Deterministic*. Stochastic methodologies incorporate the concept of randomness providing therefore additionally the uncertainty associated with the prediction, while deterministic methods make no use of the probability theory.
- *Gradual – Abrupt*. Depending on the degree of smoothness of the resulting estimation surface the methodologies are classified as gradual or abrupt interpolators.

The reviewed methodologies include the nearest neighbor, the triangulated irregular network interpolators, the natural neighbor, which utilize the tessellation of the spatial field of interest, the inverse distance weighted methodologies, the radial basis function and thin plate spline schemes and the geostatistical Kriging set of interpolation methodologies. Additionally the artificial neural network approach to spatial interpolation is presented. The majority of the interpolation schemes are based on the weighted average approach, differentiating in the procedure used to calculate the interpolation weights (w_i).

2.1 Nearest neighbor

The nearest neighbor scheme is the simplest methodology of multivariate interpolation and it is based on the Voronoi diagram (Dirichlet or Thiessen tessellation). Given a set of distinct

points in the Euclidean plane, for constructing the Voronoi diagram, all locations in that space are associated with the closest member of the point set. The result is a tessellation of the plane into a set of the regions associated with members of the point set (Okabe et al., 2000). A sample Voronoi diagram is illustrated in Fig. 1, which corresponds to the partition of the area of study for the case of the air quality monitoring network in Athens, which is used in this work for the evaluation of the interpolation methodologies in air pollution modeling. The nearest neighbor scheme utilizes the Voronoi diagram and for point predictions all estimations within a region are set as equal to the nearest point or to the average its surrounding points.
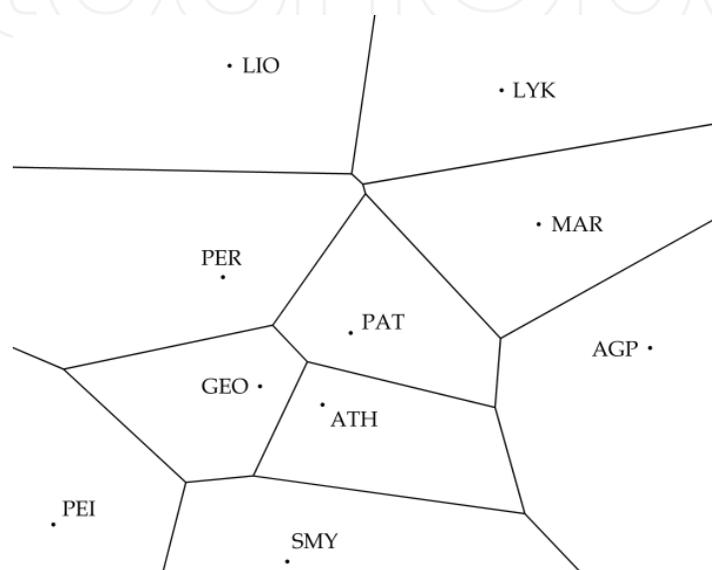

Fig. 1. Voronoi diagram for the air monitoring network of Athens, Greece

The nearest neighbor methodology is selected for its computational simplicity and its main weakness is that each prediction is based on just one measurement and therefore information from other neighboring points is ignored (Webster & Oliver, 2007). Furthermore, it may be more suitable for data that are more discrete than continuous in nature (Eberly et al., 2004). These drawbacks limit the use of the nearest neighbor scheme in air pollution modeling which has been used in spatial interpolation comparative studies (Deligiorgi et al., 2010; Marshal et al., 2008; Shao et al., 2007) and for air quality data imputation (Junninen et al., 2002).

2.2 Triangulated irregular network

The triangulated irregular network (polynomial interpolation) is a group of methodologies that utilize the triangulation of the input space for estimating the values of the variable under study in unsampled locations. Delaunay triangulation is typically employed, which assigns triangles in the dataset by employing the criterion that no data point is positioned inside the circumcircle of any other triangle (Watson & Philip, 1984). Delaunay triangulation maximizes the minimum angle of all the triangle angles and thus it avoids extremely sharp and obtuse angles. The Voronoi tessellation is the geometrical dual of the Delaunay triangulation, obtained by perpendicularly bisecting the edges of the triangulation. In the case of the air quality monitoring network in Athens, both diagrams are presented in Fig. 2. Triangulated irregular network estimates for a target point within one of the formed

triangles is performed by fitting a polynomial function to the values at the vertices of the formed triangles. The predictive surface is typically generated by using a least-squares regression fit that minimizes squared differences between the surface and measured points (Eberly at al., 2004).
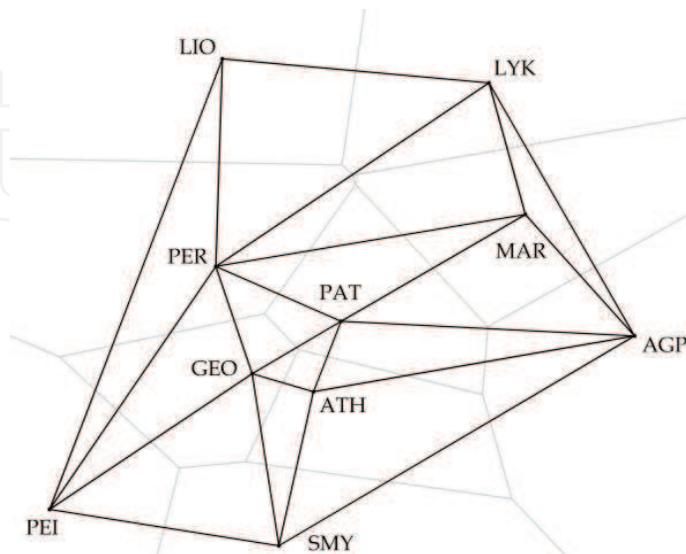

Fig. 2. Delaunay triangulation (black lines) and its dual Voronoi tessellation (gray lines) for the air quality monitoring station in Athens

2.3 Natural neighbor
The natural neighbor interpolation scheme is also based on the Voronoi tessellation and is a local interpolant, where the estimated function is a linear appropriately weighted average of the nearby data points. The selection of the nearby sites (natural neighbors) and the interpolation weights is based on a methodology proposed by R. Sibson (Sibson, 1981). Initially for a given set of irregular points $S = \{S_1 \ldots S_n\}$, the original Voronoi diagram is constructed (Fig. 1). Subsequently, the target interpolation point q is inserted in the diagram, altering the Voronoi tessellation, which for the case of the air monitoring network in Athens is illustrated in Fig. 3. The point q in the new Voronoi diagram is associated with the region $V(q)$ and the interpolation weights w_i are calculated from (Fan, 2005):

$$w_i = \frac{Area(V(q) \cap V(s_i))}{Area(V(q))} \tag{1}$$

where s_i the natural neighbors of q and $V(s_i)$ their associated regions in the original Voronoi diagram. The natural and the nearest neighbor schemes belong to the polygonal interpolation group of methodologies and are local, deterministic and abrupt schemes.

2.4 Inverse distance weighting
The Inverse distance weighting class of interpolation schemes are developed by D. Shepard and are based on the assumption that the degree of influence of the nearby sample points should be greater than the effect of more distant points. The interpolant is also a weighted average of the sample point values and the weights w_i are expressed as an inverse function of distance according to:
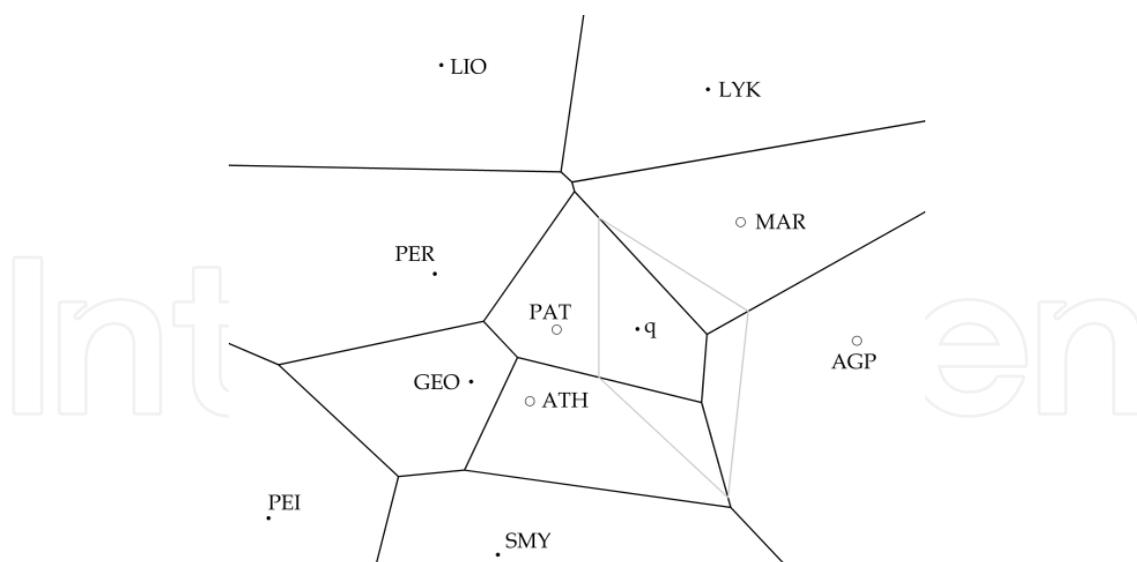
Fig. 3. Voronoi diagram of the air quality monitoring network in Athens, Greece (black lines), natural neighbors (open circle points) and the Voronoi region of *q* (defined from the gray lines)

$$w_i = \frac{1/d_i^p}{\sum\limits_{i=1}^{N} 1/d_i^p}$$ (2)

where *N*, the number of the sample points, d_i the distance of the target point from each of the sample points and *p* the power parameter. The selection of *p* is arbitrary (typically equal to 2) and expresses the relative importance of the nearby points in relation to the distant points. As *p* becomes larger the weight assigned to points in large distances becomes smaller. Especially for *p*→0, the IDW estimate is equal to the spatial average of all sampled points, and when *p*→∞, the methodology is identical with the nearest neighbor scheme. The IDW schemes therefore can be global or local and are deterministic, exact and gradual interpolation methods. Although IDW methodology is one of the most widely used interpolation schemes, it has several limitations, mainly due to the fact that the spatial relationship between two locations is not simply a function of distance and in many cases this distance-decay relationship is not constant over space (Lu & Wong, 2008). In air pollution modeling the IDW method is a popular alternative to Kriging methodology in various scales. It is applied operationally by the Environmental Protection Agency (EPA) for generating real-time O_3, PM_{10} and Air Quality Index spatial predictions in nationwide scales (Eberly et al., 2004).

2.5 Radial Basis Functions
The use of Radial Basis Functions (RBF) as estimators for interpolating irregularly distributed data is introduced by R.L. Hardy (Hardy, 1971) and since then they have been widely used in various disciplines in multivariate scattered data interpolation. The method is based on the assumption that the estimated surface can be approximated to any degree of exactness by the summation of regular mathematically defined surfaces (Hardy, 1978). The RBF interpolation schemes are global, exact, deterministic interpolators. The RBF approximation function can be expressed by:

$$y(x) = \sum_{i=1}^{N} w_i \varphi(\|x - c_i\|) \tag{3}$$

where the estimation function y is a linear aggregate of N radially symmetric functions φ for a given norm $\|\cdot\|$, with Euclidean norm being the most common choice (Baxter, 1992). The data points are called the centers (c_i) of the RBF interpolant and each φ is associated with one of these centers. The coefficients w_i are chosen so that the interpolation conditions are satisfied and estimated by solving the system of linear equations (Rousos & Baxter, 2005):

$$Aw = f \tag{4}$$

where

$$A_{ij} = \varphi(\|x_i - x_j\|) \tag{5}$$

is the interpolation matrix, w the vector of coefficients and f the vector of function values. Some common φ functions are the linear and cubic:

$$\varphi(r) = r \tag{6}$$

$$\varphi(r) = r^3 \tag{7}$$

the Gaussian, the multiquadric and the inverse multiquadric::

$$\varphi(r) = e^{-\frac{r^2}{2c^2}} \tag{8}$$

$$\varphi(r) = (r^2 + c^2)^{0.5} \tag{9}$$

$$\varphi(r) = (r^2 + c^2)^{-0.5} \tag{10}$$

The parameter c, in the Gaussian, the multiquadric and the inverse multiquadric RBFs is called the shape parameter and can be chosen arbitrarily, with a reasonable choice being in the region of the average euclidean distance between the centers. In many cases the quality of approximation is very sensitive to the shape parameter. The main advantage of the RBF interpolation schemes is that they require generally few point observations but relatively well distributed centers in order to create a good approximation of the studied process. The method in general doesn't perform well for a process with abrupt changes in small distances and as an exact method, for processes associated with large observation errors.

2.6 Thin plate splines

Thin plate splines were initially introduced for interpolating meteorological fields (Wahba & Wendelberger, 1980) and can be considered as a special form of RBF based interpolation. The physical analog of the term 'thin plate spline', refers to the bending of a thin sheet of metal so it would pass through a predefined set of points. The Spline interpolation schemes in the case of scattered data interpolation is used to find the minimum curvature surface that passes through a set of irregularly spaced data points (Sandwell, 1987). The two fundamental conditions a thin plate spline satisfies are the exactness and smoothness (Myers, 1994):

Spatial Interpolation Methodologies in Urban Air Pollution Modeling:
Application for the Greater Area of Metropolitan Athens, Greece

347

$$\sum_{i=1}^{N}[f(x_i)-f^*(x_i)]^2 = 0 \tag{11}$$

$$\int_{R}[Df^*(x)]^2 dx \quad is \quad minimized \tag{12}$$

where D is a second order differential operator and R the region of interest. An alternative restriction for a thin plate spline is the minimization of the quantity (Myers, 1994):

$$\sum_{i=1}^{N}[f(x_i)-f^*(x_i)]^2 + \lambda \int_{R}[Df^*(x)]^2 dx \quad is \quad minimized \tag{13}$$

This expression represents a tradeoff between the exactness and smoothness of the solution. The interpolant obtained from above restriction no longer interpolates the noisy data, but smooths it, with the smoothing parameter λ controlling the tradeoff between exactness and smoothness (Wahba, 1990).

2.7 Kriging
Kriging interpolation schemes are the fundamental tools in the field of spatial statistics, developed by the founder of geostatistics G. Matheron and named after D.G. Krige who proposed the technique for mining applications (Krige, 1951). The main difference of the Kriging schemes from the above non-geostatistical methodologies is the assumption that spatial correlation structure of the process under study is known and it is estimated from the observed data. The Kriging interpolation schemes are stochastic, local, gradual and exact interpolators. The Kriging methodology includes two stages: the analysis of the spatial variation and the estimation of the target variable, which is also based on the weighted average approach.

The analysis of the spatial variation is performed through the variogram assessment. The observed data are initially used for the estimation of the semivariance from:

$$\hat{\gamma}(h) = \frac{1}{2n(h)}\sum_{i=1}^{n}(z(x_i)-z(x_i+h))^2 \tag{14}$$

where z, the variable under study, h the lag distance between the pre-ordered data and $n(h)$ the number of the paired data at distance h. The plot of the semivariance versus the lag distance is the empirical variogram, which is subsequently modeled by a parametric model. The most common models are the linear, the circular, the spherical, the exponential, the Gaussian or a nested model of one or more of the above. The characteristic parameters of the variogram are the sill, which is the semivariance value at the plateau of the variogram, representing the distance where there is no correlation between the observations, the range, which is the lag where the semivariance is equal to the sill and the nugget which represents theoretically the minimum variance and is the semivariance value at zero lag distance (Fig.4).

The Kriging scheme is also known as the best linear unbiased estimator, and its estimates are based on the variogram model and the values and location of the measured points (Eberly et al., 2004). The Kriging interpolation weights are chosen using the modeled

variogram so that the estimate is unbiased and the estimation variance is less than any other linear combination of the observed values. The most common Kriging methods are the Simple Kriging, which assumes a known constant mean, Ordinary Kriging, which assumes that there is an unknown constant mean, estimated from the data and Universal Kriging which assumes that there is a trend in the surface that partly explains the data's variations (Eberly et al., 2004).
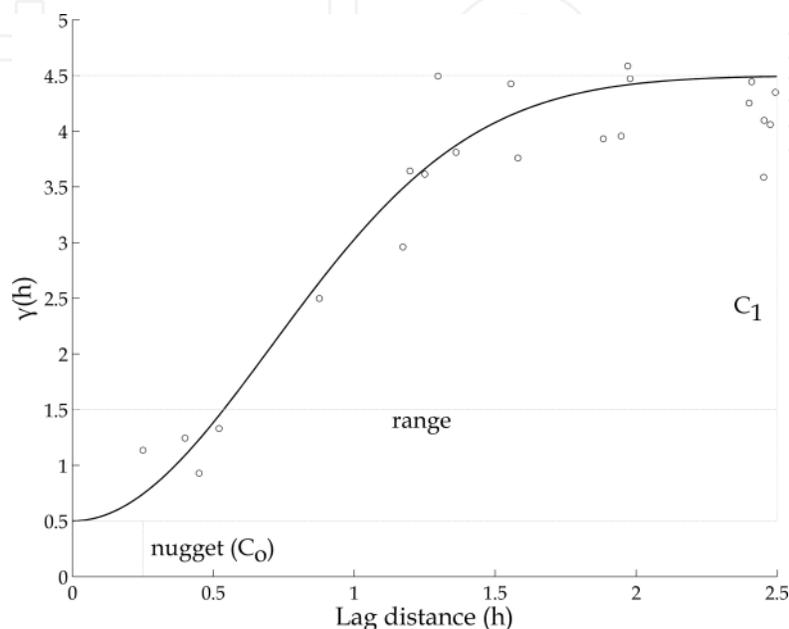

Fig. 4. Example of an empirical variogram with the fitted theoretical model (Gaussian)

2.8 Artificial Neural Networks

Recent advances in artificial intelligence propose an additional approach to the problem of point spatial interpolation of irregularly and gridded data An Artificial Neural Network (ANN) is a mathematical model inspired from the brain physiology to take advantage of their advanced problem solving abilities. Biological brains contain specialized cells called neurons, which interconnect and form highly complex networks. A biological neuron is consisted of a cell body (soma), the dendrites, which extend out of the cell body and an axon which is the transmitter of the neuron. In terms of information processing, the basic operation of an artificial neuron, is the summation of all the incoming signals through its dendrites. A pulse, in accordance with the biological functioning, is sent along the axon if sufficient input signal is received to stimulate the neuron to its threshold level. This simplified neuron operation (McCulloch-Pitts model) can be modeled by the data flow diagram presented in Fig. 5.

The output of the neuron in Fig. 5 is calculated using the following equation:

$$y = f(\sum_{i=1}^{n} x_i w_i - \theta) \tag{15}$$

where f is the binary threshold function. The abilities of the McCulloch-Pitts neuron model can be extended by replacing the activation functions and depending on the application, linear, sigmoid, tangent hyperbolic and other functions can be used.
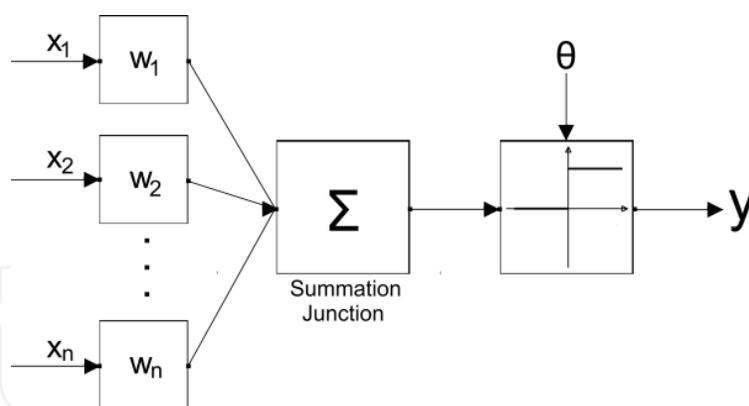
Fig. 5. Simplified McCulloch-Pitts model

An ANN is set of interconnected artificial neurons which can be classified according to their topology. The most commonly employed ANNs are the Feed Forward Networks, which have no feedback connections and their most important characteristic is the lack of memory. The output of such a network depends solely on its current inputs and weights and therefore the order in which the elements of a time series are presented to the network has no effect in its output. Feed forward ANN's are universal approximators, theoretically capable of estimating any measurable input-output function to any desired degree of accuracy (Hornik et al., 1989). In Fig. 6 their most common topology is presented, which contains three layers of neurons.
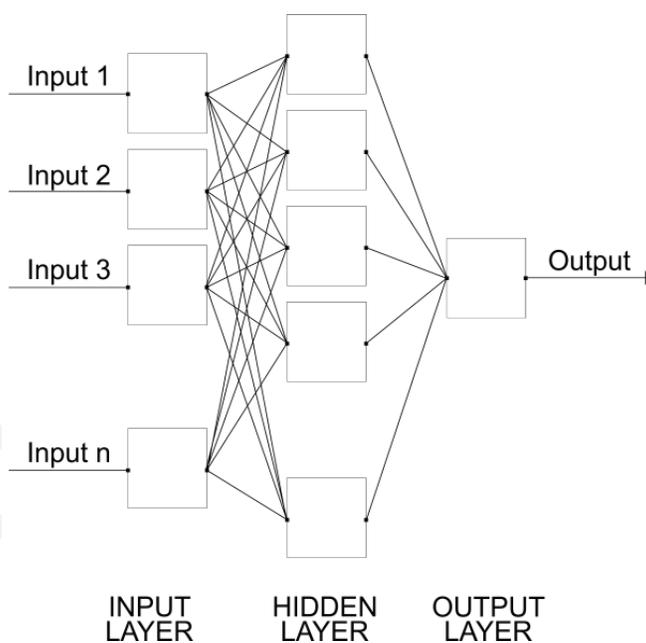
Fig. 6. Feed Forward Artificial Network

The input layer contains a set of artificial neurons with linear activation functions, while in the hidden layer, non-linear activation functions are used. The third layer is the output layer and depending on the application, linear or non-linear activation functions can be used. Besides feed forward ANNs, several topologies exist with feedback connections and their most important feature is the existence of memory. Such networks are called recurrent and their output depends on the current and previous inputs, since these define the state of its memory.

The behavior of an ANN, besides its topology, is also affected by the input neuron weights. The process of determining the weights is called training, which corresponds to the natural learning process in biological brains. Training requires the existence of a training data set, which is composed of input vectors along with the desired output from the network. In general, training can be considered as an optimization problem (selecting the set of weights that minimizes some error measure) and therefore any optimization algorithm can be applied, with backpropagation commonly employed. A successfully trained network is able to generalize, in terms of producing an output for an input that was not presented during training. Such an output can be considered as an interpolation or prediction, depending on the network input. Several issues need to be addressed during training in order to obtain an ANN with satisfactory generalization ability. A common problem, related to the entire family of gradient descent based training algorithms like backpropagation, is that a suboptimal set of weights can be initially determined. A technique to overcome this issue is the multiple training of the ANN, by altering the initial weights values. Another common problem related with training is directly related to the extreme flexibility of ANNs. Specifically, as the number of neurons increases so does the ability of the network to exploit patterns in the training set, irrelevant to the statistical dependence between input and output, such as noise (Blackwell et al., 2009). This effect is known as overtraining and the result is highly degraded generalization ability for the trained ANN. Overtraining can be resolved using early stopping, which requires the partition of the original training set to the normal training set and the validation set, which is continually used during training to evaluate the generalization ability of the ANN. A decrease in this ability will trigger a premature ending of the training process in order to avoid overtraining.

Besides these inherent drawbacks, the wider acceptance of ANNs for statistical interpolation and prediction is limited due to the requirement of a representative training set. This fact differentiates the ANNs from the previous interpolation methodologies.

3. Application of spatial interpolation methods

The accuracy assessment for the aforementioned spatial interpolation methodologies is performed for the greater area of metropolitan Athens in Greece. The area under study is the Athens basin, which is bounded by Mount Aegaleo to the west, Mount Parnitha to the north, Mount Pentelikon to the northeast, Mount Hymmetus to the east and the Saronic Gulf coastline to the southwest. These geophysical features along with the prevailing meteorological conditions, land–use types and pollution sources, create a complex pattern of air pollution spatial distribution. The main sources of pollution in the area are vehicular traffic, industrial activities and central heating emissions. The dispersion conditions at the Athens basin have been extensively studied focusing on air pollution episodes in relation to atmospheric conditions (Asimakopoulos et al., 1992; Lalas et al., 1983; Kassomenos et al. 1995; Güsten et al. 1988).

The accuracy assessment of the interpolation methodologies is carried out for Nitrogen Dioxide (NO_2) which is a primary pollutant and tropospheric Ozone (O_3) which is secondary photochemical pollutant. The analysis is carried out for their mean daily concentrations from the 1st of January 2009 to the 31st of December 2009. The air quality monitoring stations used for the purpose of this study are operated by the Hellenic Ministry of Environment, Energy and Climate Change (MEECC) and their locations are presented in Fig. 7.

The air quality data provided from the MEECC database are mean hourly concentrations of NO_2 and O_3, which are further averaged to provide the mean daily concentrations. Although the mean daily concentrations do not provide information regarding the diurnal pattern of each pollutant, a daily time scale is appropriate in epidemiological studies and for the assessment of the spatial interpolation methodologies. The hourly data availability and the characteristics of each station are presented in Table 1. Calendar days with missing mean daily observations from one or more stations are excluded from the dataset, resulting to a series of 227 and 269 days with concurrent observations for NO_2 and O_3 respectively. As it is observed the data availability for both pollutants is high for all stations except for Marousi and Peristeri and for NO_2 at Piraeus.
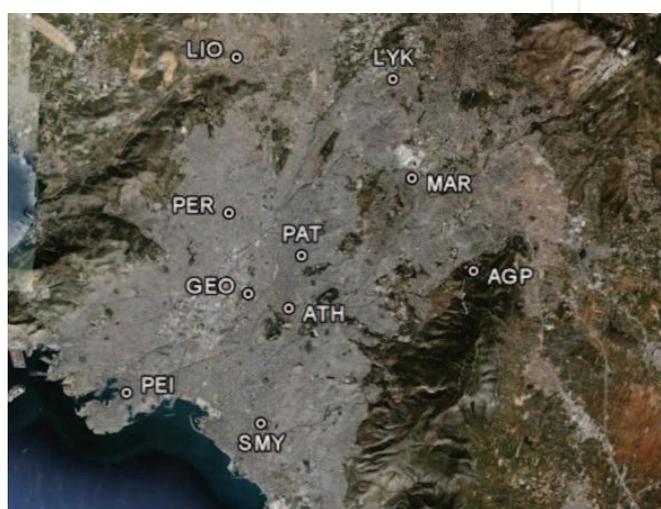
Fig. 7. Air quality monitoring network and experimental area

Station	ID	Station Type	Area Type	Data Completeness (%)	
				NO_2	O_3
Aghia Paraskevi	AGP	Background	Suburban	96.54	99.02
Athinas	ATH	Traffic	Urban	99.89	99.93
Geoponiki	GEO	Industrial	Suburban	97.73	91.67
Liosia	LIO	Background	Suburban	93.18	96.20
Lykovrisi	LYK	Background	Suburban	99.87	99.85
Marousi	MAR	Traffic	Urban	87.47	90.11
Patision	PAT	Traffic	Urban	98.74	98.47
Peristeri	PER	Background	Urban	89.26	89.19
Piraeus	PEI	Traffic	Urban	86.64	95.43
Nea Smyrni	SMY	Background	Urban	95.08	96.44

Table 1. Characteristics of the air quality monitoring stations and data availability

3.1 Descriptive statistical analysis

The analysis of the mean concentration levels for the examined pollutants reveals the increased spatial variability at the area of study. For NO_2 the maximum concentrations are observed for the urban station at Patision (91.23μgr/m³) and the minimum for the suburban station at Aghia Paraskevi (18.31μgr/m³). The spatial variability for O_3 is also high at the

Athens basin, but with the exact opposite distribution. Maximum concentrations are recorded for the suburban station of Aghia Paraskevi (84.91μgr/m³) and minimum at Patision station (24.46μgr/m³). The seasonal evolution of the studied pollutants is presented in Fig. 8 for the three stations situated at the urban core of metropolitan Athens (ATH, GEO, PAT) and for the suburban station at Marousi. It is clear that the values fall in different ranges and have different characteristics. Ozone maxima are observed at the suburban regions at the Athens basin, in locations where lower NO_2 concentrations are reported. Furthermore, O_3 exhibits a clear seasonal cycle, with maximum concentrations observed during the warm period of the year, while NO_2 concentrations remain relatively stable throughout the year, with a minimum at August, due to the decreased vehicular traffic.
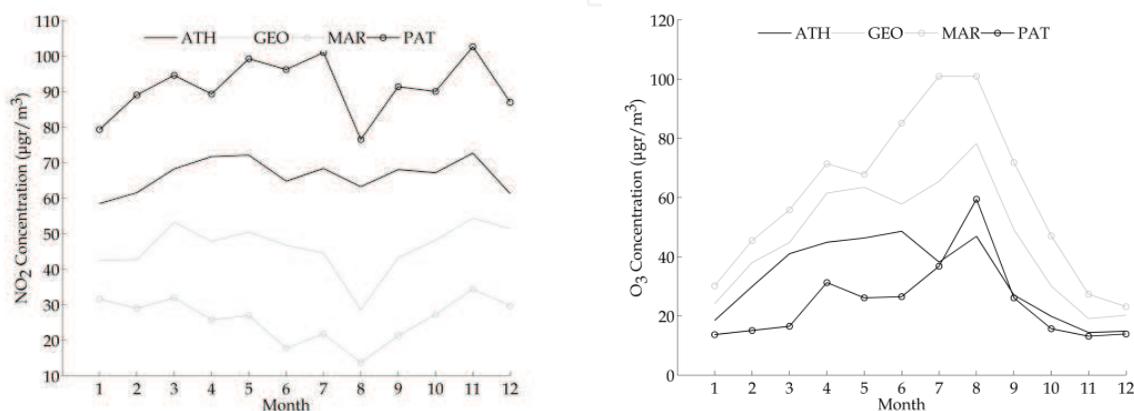

Fig. 8. Seasonal evolution of NO_2 and O_3 for the urban stations at Athinas, Geoponiki and Patision and for the suburban station at Marousi

The quality of the spatial prediction obtained by the various methods can be partly explained from the spatial correlation between the recorded time-series, which is presented in Table 2. The maximum correlation coefficient is observed between MAR and LYK for both pollutants.

	Nitrogen Dioxide (NO₂)					Ozone (O₃)				
	ATH	GEO	MAR	PAT	PER	ATH	GEO	MAR	PAT	PER
AGP	0.47	0.71	0.75	0.61	0.66	0.70	0.83	0.86	0.70	0.87
ATH	1.00	0.79	0.65	0.59	0.70	1.00	0.84	0.74	0.54	0.70
GEO	0.79	1.00	0.88	**0.84**	0.87	0.84	1.00	0.90	**0.81**	0.88
LIO	0.71	**0.90**	0.85	0.80	**0.89**	0.47	0.77	0.80	0.78	0.86
LYK	0.58	0.87	**0.92**	0.80	0.85	0.70	0.89	**0.94**	0.80	**0.94**
MAR	0.65	0.88	1.00	0.80	0.85	0.74	0.90	1.00	0.80	0.91
PAT	0.59	0.84	0.80	1.00	0.74	0.54	0.81	0.80	1.00	0.76
PER	0.70	0.87	0.85	0.74	1.00	0.70	0.88	0.91	0.76	1.00
PEI	**0.82**	0.58	0.38	0.37	0.50	0.82	0.68	0.54	0.42	0.55
SMY	0.77	0.85	0.78	0.68	0.84	**0.84**	**0.93**	0.90	0.76	0.91

Table 2. Correlation coefficient between the target and the remaining stations

3.2 Methodology

The assessment of the spatial interpolation methodologies for the greater area of metropolitan Athens includes thirteen schemes in total, which are presented in detail at

Spatial Interpolation Methodologies in Urban Air Pollution Modeling:
Application for the Greater Area of Metropolitan Athens, Greece

353

Table 3. The accuracy assessment of the evaluated interpolation methodologies is performed using the leave-one-out cross-validation methodology. According to this technique, one station is selected as the target station and an interpolation methodology is employed in order to estimate air pollution concentrations at its coordinates. The predictive accuracy of the interpolation scheme is evaluated against the observed concentrations. The leave-one-out cross-validation technique is applied multiple times for NO_2 and O_3 for the target stations situated at Athinas (ATH), Geoponiki (GEO), Marousi (MAR), Patision (PAT) and Peristeri (PER).

Interpolation Methodology	Abbreviation
Nearest Neighbor	NN
Triangulated Irregular Network	
• Linear Interpolation	Lin
• Cubic Interpolation	Cub
Natural Neighbor	NaN
Inverse Distance Weighted	
• Inverse distance weighted (p=1)	IDW
• Inverse distance squared weighted (p=2)	ID^{2}W
Kriging	
• Ordinary Kriging	OK
Radial Basis Functions Interpolation	
• Linear RBFs	RBFlin
• Gaussian RBFs	RBFg
• Multiquadric RBFs	RBFmq
Splines	
• Thin-plate splines	SPtp
• Biharmonic Splines	SPbh
Artificial Neural Networks	
• Feed Forward Neural Networks	ANN

Table 3. Spatial interpolation methodologies

The application of the NN, Lin, Cub, NaN, IDW, ID^{2}W, RBFlin, SPtp and SPbh schemes is straightforward, following the theoretical background presented at the second section of this chapter. Regarding the Ordinary Kriging methodology, an automated approach is adopted, where isotropy is assumed. During the phase of variogram modeling the circular, spherical, pentaspherical, exponential and Gaussian models are used. The best model for each day and station is selected using the coefficient of determination (R^2), and subsequently it is employed in the OK spatial interpolation phase. For the selected time frame, during the analysis of the spatial variation, in 56.04% of the days for NO_2 and in 56.13% of the days for O_3 the circular model is selected.

Regarding the ANN, the additional dataset required for training consists of daily NO_2 and O_3 concentrations from 2000 to 2008, acquired from the MEECC database. This dataset was randomly divided into training and validation sets, with the 80% of its values used for training and the remaining 20% for validation. The test set in this case, is also the mean daily NO_2 and O_3 concentrations from 2009. The validation set is used for avoiding overtraining

and for selecting the optimum ANN architecture. Multiple ANNs are trained with varying numbers of hidden layer neurons (from 10-50) and the optimum architecture is the one that minimizes the Mean Absolute Error (MAE) on the validation set. In all cases for avoiding the inherent drawback of gradient decent algorithms, training is performed multiple times (50 repetitions). The selected architecture for each station and pollutant is presented in Table 4.

	ATH	GEO	MAR	PAT	PER
NO_2	9-32-1	9-34-1	9-32-1	9-31-1	9-38-1
O_3	9-36-1	9-26-1	9-37-1	9-37-1	9-33-1

Table 4. ANN architecture, number of input, hidden and output layer neurons

In the case of RBFg and RBFmq, the shape parameter was not chosen arbitrarily as the average Euclidean distance of all station pairs. A sensitivity assessment of the effect of the parameter c on the predictive accuracy of RBFg and RBFmq for each station and pollutant is performed using the additional dataset from 2000-2008. The optimum value of the parameter c is the one that minimizes the *MAE* and the results are presented in Table 5.

		ATH	GEO	MAR	PAT	PER
NO_2	RBFq	3	0.75	3.1	4.5	2
	RBFmq	6	4.5	1	6	2.8
O_3	RBFq	1.9	3.5	2	6	1.4
	RBFmq	4	4.5	2	10	2

Table 5. Values of shape parameter c in km

The performance evaluation of each interpolation scheme is based on a set of correlation and difference statistical measures (Willmott, 1982), along with the application of Wilcoxon's sign rank test. The proposed measures for assessing the accuracy of the interpolation schemes include the correlation coefficient (R) and coefficient of determination (R^2), the Mean Absolute Error (*MAE*), which is proposed as the optimal model performance error for dimensioned evaluations (Willmott & Matsuura, 2005), the Mean Absolute Percentage Error (MAPE), the Root Mean Square Error (*RMSE*), the Mean Bias Error (*MBE*) and the Index of agreement d (Willmott, 1982), which is a descriptive measure that scales with the magnitude of the variables, retains mean information and does not amplify outliers (Vicente-Serrano et al., 2003).

According to the above statistical criteria and in combination with data display graphics like scatterplots the best performing scheme is selected. In each case the best performing scheme is statistically tested for its equal predictive accuracy versus the rest interpolation schemes using the Wilcoxon's signed-rank test (Diebold & Mariano, 1995):

$$S_{3a} = \frac{\sum_{t=1}^{T} I_+(d_t) rank(|d_t|) - \frac{T(T+1)}{4}}{\sqrt{\frac{T(T+1)(2T+1)}{24}}} \tag{16}$$

where

Spatial Interpolation Methodologies in Urban Air Pollution Modeling:
Application for the Greater Area of Metropolitan Athens, Greece

355

$$d_t = \left| S_{it} - S'_{it} \right| - \left| S_{it} - S'_{itj} \right| \tag{17}$$

and

$$I_+(d_t) = \begin{cases} 1, d_t > 0 \\ 0, otherwise \end{cases} \tag{18}$$

where, T the number of the concurrent observations, S_{it} the concentration at the i station at time t, S'_{it} the prediction of the best performing interpolation scheme according to statistical measures, and S'_{itj} the prediction of the j scheme at station i and at time t. The null hypothesis of equal predictive accuracy between the methodologies is evaluated by comparing the test statistic (S_{3a}) against the t-distribution. Superior forecasts are indicated by the sign on the test statistic, which is determined by the order which the error terms are subtracted from each other in d_t (Snell et al., 2000).

3.3 Results

A general remark is that the performance of the various interpolation schemes varies significantly, depending on the target station and therefore a universal method cannot be selected for the entire field of interest as the optimum interpolation scheme. Furthermore, the air pollutant of concern affects the performance of spatial interpolation methods, attributed to the differentiations in transport mechanisms and chemical transformations.

Initially, the results for the primary pollutant NO_2 are discussed. The main performance statistics are the MAE and MAPE, which are presented in Table 6. A direct finding is that the ANN scheme for NO_2 consistently outperforms all interpolation methods for all target stations. Among these stations, the most difficult to be predicted is PAT station since all methods produce predictions with significant errors. However, even for this station, ANN predictions are distinctively better (MAPE = 15.67%) compared to other methods with MAPE values ranging from 28.22% for RBFg to 55.10% for IDW scheme. The best ANN predictions for NO_2 are obtained for GEO and MAR having MAPE values 5.9% and 6.1% respectively. The results of the other methods for these two stations are also superior compared to the rest of the target stations, although they cannot compare favorably against ANN predictions. The above results are in accordance with the rest statistical measures. The explained variation by ANN, as it expressed by R^2, has a mean 77.29% for all target stations, varying from 59.04% for PAT to 90.46% for GEO station. Furthermore, high d values (greater than 0.92) are reported for all stations except PAT. Regarding the average model bias, the low MBE values are obtained for ANNs (ranging from -1.34µgr/m^3 to 0.33µgr/m^3) provide strong indication of its superior predictive abilities compared to the other methods, which systematically underestimate or overestimate the observed concentrations.

Regarding O_3 predictions, ANN outperforms the rest interpolation schemes for three target stations (namely ATH, PAT, PER) although its predictive ability is generally lower compared to NO_2 predictions (slightly higher MAPE values). For the stations at GEO and MAR the best results are obtained from RBFg and NN respectively (Table 7). The unexpected superior performance of the simplistic NN scheme can be attributed to the extremely high spatial correlation ($R = 0.94$) between MAR and its nearest neighbor station, located at Lykovrisi. It should be noted that even for these two stations, ANN is among the best performing schemes. In detail, for ATH, PAT and PER the ANN's d values are greater

than 0.88 and the *RMSE* is lower than 13.89μgr/m^3. The higher prediction errors for all schemes are also observed for the urban PAT station, indicating the increased spatial variability of air pollution in the urban core of metropolitan Athens.

	ATH		GEO		MAR		PAT		PER	
	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE
NN	22.24	25.03	22.24	25.03	9.02	10.83	25.43	31.04	10.34	13.33
Lin	12.93	14.89	16.19	18.45	10.85	12.18	39.94	43.08	10.33	13.09
Cub	16.00	19.09	22.16	24.35	19.73	20.73	34.96	38.75	10.23	12.99
NaN	12.06	15.47	15.92	17.97	10.83	12.16	41.72	44.60	12.95	15.20
IDW	30.07	31.86	8.77	10.66	25.95	27.33	52.87	55.10	11.52	14.65
ID2W	32.72	34.47	12.56	14.50	27.89	29.70	52.52	54.88	12.29	16.05
OK	16.22	19.32	15.40	17.82	20.61	21.67	37.30	40.69	12.19	14.63
RBFlin	12.97	16.43	20.64	22.20	15.89	17.03	42.37	45.04	20.50	22.99
RBGg	13.20	16.38	15.47	18.58	8.37	9.65	22.53	28.22	13.69	16.03
RBFmq	12.40	15.04	15.77	18.92	22.45	23.73	23.02	28.87	17.15	21.31
SPtp	12.45	15.37	18.96	21.41	27.09	28.38	34.80	38.68	18.19	22.31
SPbh	12.48	15.41	19.33	21.74	33.39	34.80	34.51	38.41	20.00	24.34
ANN	**6.77**	**8.77**	**4.57**	**5.93**	**4.63**	**6.10**	**11.80**	**15.67**	**8.79**	**11.73**

Table 6. MAE and MAPE (%) performance statistics for NO$_2$

	ATH		GEO		MAR		PAT		PER	
	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE
NN	15.12	19.88	15.12	19.88	**7.42**	**11.04**	15.38	19.24	19.21	23.61
Lin	12.22	16.28	7.65	10.95	9.00	12.63	21.87	26.33	19.79	23.28
Cub	12.06	15.21	11.81	15.63	8.60	11.73	17.87	23.00	21.56	25.10
NaN	14.06	18.57	7.11	10.13	9.01	12.64	23.74	27.88	22.80	26.39
IDW	26.31	30.92	12.77	15.08	13.71	17.82	32.89	35.63	11.20	14.37
ID2W	26.83	31.92	14.89	17.49	15.27	19.80	32.46	35.14	9.83	12.28
OK	13.81	18.41	6.98	9.97	11.99	15.02	26.12	30.14	19.56	24.01
RBFlin	12.93	17.43	9.11	12.27	8.42	11.90	24.22	28.23	30.26	33.68
RBGg	10.42	14.11	**6.15**	**7.78**	8.85	12.25	17.00	21.59	22.00	25.28
RBFmq	10.70	13.91	10.04	14.05	8.56	11.92	16.54	21.33	23.02	28.32
SPtp	10.90	14.87	9.52	13.32	9.91	12.89	18.66	23.78	26.50	31.00
SPbh	10.91	14.96	9.63	13.37	10.12	13.19	18.57	23.74	28.28	32.77
ANN	**8.07**	**10.42**	9.00	12.35	8.85	12.89	**10.47**	**13.89**	**9.13**	**12.27**

Table 7. MAE and MAPE (%) performance statistics for O$_3$

The scatter diagrams produced for each case are in accordance with the results from the statistical measures and some representative cases are presented in Fig. 9. Limited dispersion along the x=y axis is observed for the best performing schemes (for NO$_2$ at GEO for ANN, for O$_3$ at GEO for RBFg and for O$_3$ at MAR for NN) while systematic under-prediction is observed for the IDW scheme for NO$_2$ at GEO (MBE = -8.23μgr/m^3).

The methodologies based on the tessellation of the spatial field of interest (NN, Lin, Cub, NaN) along with the schemes that use simple functions of distance (IDW, ID^{2}W) for the

estimation of the spatial distribution of air pollution, in most cases fail to describe adequately the spatial variation of NO_2 and O_3. The spatial setting of the monitoring stations especially for the primary pollutant NO_2 is not dense enough to provide adequate information for the efficient modeling of the air pollution levels. Regarding O_3, especially for the MAR station an increased performance of the simplified schemes is observed attributed to the existence of a highly representative station in its proximate region (LYK).
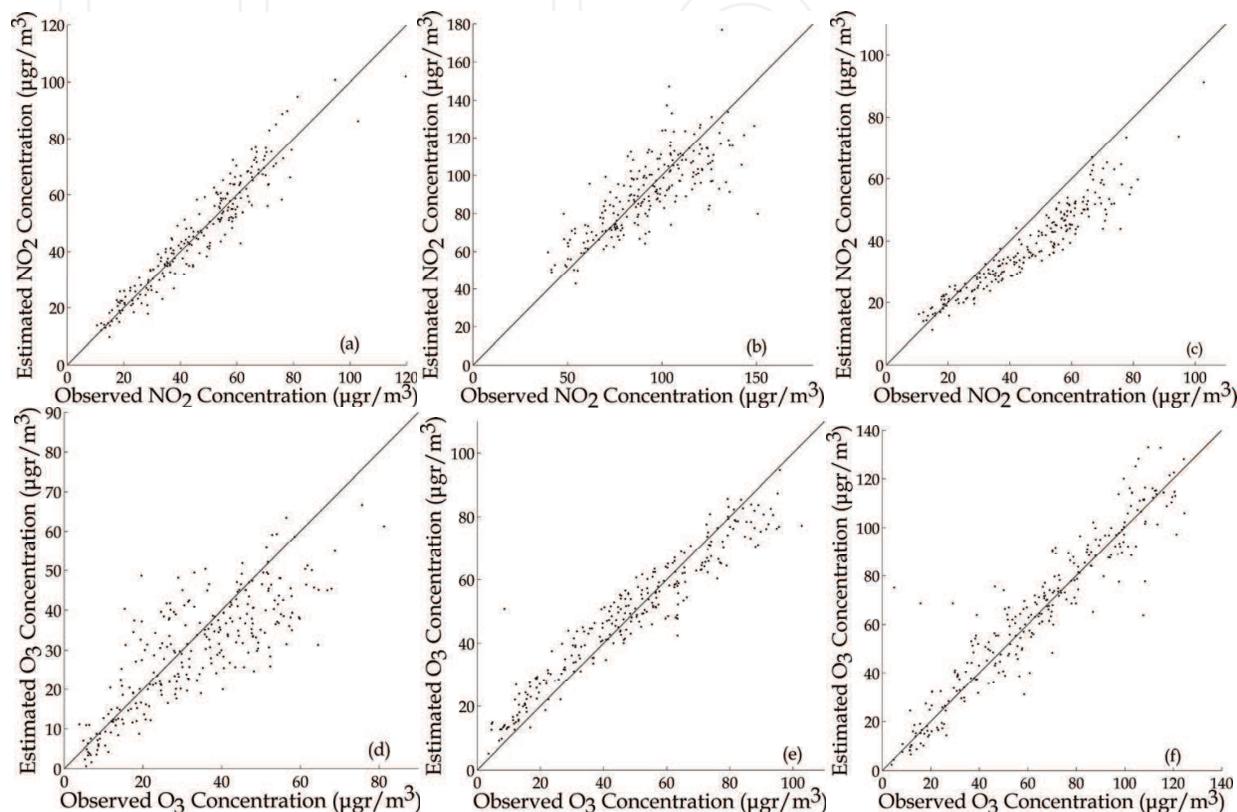
Fig. 9. Scatter diagrams of observed versus estimated concentrations for (a) NO_2 at GEO for ANN, (b) NO_2 at PAT for ANN, (c) NO_2 at GEO for IDW, (d) O_3 at ATH for ANN, (e) O_3 at GEO for RBFg and for (f) O_3 at MAR for NN.

The results of the OK methodology are greatly affected by the sampling configuration of the monitoring network. In most cases, except for O_3 at Geoponiki the OK prediction is comparable to more simplistic models, giving evidence that the scheme cannot give highly accurate results at every point of interest in the area under study. The performance of the methodology could however be increased taking into account, during the variogram analysis, the effect of meteorological factors and topography that induce the anisotropy in the spatial distribution of both primary and secondary pollutants. In our case anisotropism is ignored to simplify and automate the model fitting. The Gaussian scheme from the family of the RBF interpolants is found to be in average the best alternative scheme to ANNs with d values ranging from 0.62 at PAT to 0.89 at MAR and from 0.66 at PAT to 0.96 at GEO for NO_2 and O_3 respectively.

The results of the Wilcoxon's test statistic are presented in Table 8 and are used for assessing the statistical significance of the aforementioned results. Regarding NO_2 and for all stations the predictive ability of the ANN is assessed versus the alternative schemes, while for O_3 the

ANN is assessed for ATH, PAT and PER, the RBFg for GEO and the NN for MAR. The hypothesis of ANN's equal predictive accuracy for NO_2 is rejected in all cases at $p<0.05$ and in all cases except for NN, Lin and Cub for PER at $p<0.01$. The ANN predictive accuracy for O_3 is statistical significantly superior from the rest interpolation schemes at ATH, PAT and PER at $p<0.05$ and in all cases except for RBFg at ATH and PER at $p<0.01$. Regarding the RBFg at GEO and the NN at MAR, the null hypothesis of equal predictive accuracy is accepted at $p<0.05$ for the Lin and the OK models and for the RBFlin and ANN models respectively.

	Nitrogen Dioxide (NO_2)					Ozone (O_3)				
	ATH	GEO	MAR	PAT	PER	ATH	GEO	MAR	PAT	PER
NN	-11.72	-12.61	-9.01	-9.45	**-2.52**	-7.40	-9.48	--	-5.97	-7.40
Lin	-10.09	-11.66	-10.48	-12.51	**-2.49**	-4.80	**-2.33**	-3.33	-9.92	-4.80
Cub	-10.52	-12.61	-12.93	-12.01	**-2.43**	-5.90	-7.87	-3.36	-7.06	-5.90
NaN	-6.33	-11.70	-10.46	-12.61	-6.25	-6.25	-1.71	-3.36	-11.00	-6.25
IDW	-12.89	-8.62	-13.00	-13.03	-3.78	-12.44	-9.91	-7.89	-13.76	-12.44
ID^2W	-12.96	-11.98	-12.97	-13.03	-4.34	-12.30	-11.06	-8.61	-13.72	-12.30
OK	-9.42	-11.41	-12.82	-12.30	-5.16	-5.83	**-1.93**	-7.97	-11.83	-5.83
RBFlin	-7.05	-12.67	-12.51	-12.64	-11.05	-2.89	-5.76	**-2.54**	-11.31	-2.89
RBFg	-7.50	-11.37	-8.63	-8.39	-6.82	**-2.46**	--	-3.23	-7.07	**-2.46**
RBFmq	-8.08	-11.42	-12.99	-8.62	-9.01	-3.53	-5.38	-2.96	-6.51	-3.53
SPtp	-7.58	-12.18	-13.06	-11.98	-9.59	-2.89	-4.85	-5.41	-7.58	-2.89
SPbh	-7.61	-12.25	-13.06	-11.94	-10.29	-2.81	-5.18	-5.47	-7.51	-2.81
ANN	--	--	--	--	--	--	-5.32	**-1.89**	--	--

Table 8. Wilcoxon's sign rank test statistics

4. Conclusion

The estimation of pollution fields, especially in densely populated areas, is an important application in the field of environmental science due to the significant effects of air pollution to public health. In this chapter the spatial interpolation methodologies, commonly employed for air pollution point estimations, are reviewed and their accuracy is assessed for five target sites located at the greater area of metropolitan Athens. In general, an optimal interpolation scheme that outperforms the rest methodologies at the area of study is not established. The errors associated with all interpolation schemes are related to the pollutant of concern and mainly to the spatial configuration of the monitoring network.

The Artificial Neural Network approach to point spatial interpolation is found to be specifically promising, compared to the traditional interpolation techniques. For the primary (NO_2) and secondary (O_3) pollutants addressed in this study, ANNs are found to be statistical significantly superior in all cases except for one case where their predications have equal predictive accuracy with the RBFg scheme and for O_3 at the Marousi station, where

Spatial Interpolation Methodologies in Urban Air Pollution Modeling:
Application for the Greater Area of Metropolitan Athens, Greece

359

the simplistic Nearest Neighbor model outperforms the trained ANN scheme. In this context the ANNs are proposed as an alternative method to spatial interpolation, with increasingly important applications where point air quality estimations are prerequisite. The performance of the ANNs can be further improved, since there is no immediate restriction to the number of their inputs, by including parameters related to the atmospheric dispersion and the chemistry of the pollutant under consideration. The critical requirement of a representative dataset during training is the basis for their limited operational applications in air quality spatio-temporal forecasting.

5. Acknowledgements

6. References

Analitis, A.; Katsouyanni, K.; Dimakopoulou, K.; Samoli, E.; Nikoloulopoulos, A,K.; Petasakis, Y.; Touloumi, G.; Schwartz, J.; Anderson, H.R.; Cambra, K.; Forastiere, F.; Zmirou, D.; Vonk, J.M.; Clancy, L.; Kriz, B.; Bobvos, J. & Pekkanen, J. (2006). Short-term effects of ambient particles on cardiovascular and respiratory mortality, *Epidemiology*, Vol.17, No.2, (March2006), pp. 230-233, ISSN 1531-5487

Asimakopoulos, D.; Deligiorgi, D.; Drakopoulos, C.; Helmis, C.; Kokkori, K.; Lalas, D.; Sikiotis, D. & Varotsos, C. (1992). An experimental study of nighttime air-pollutant transport over complex terrain in Athens. *Atmospheric Environment*, Vol.26, No.1, (March 1992), pp. 59-71, ISSN 1352-2310

Baxter, B. (1992). *The interpolation theory of radial basis functions*, *Ph.D. Thesis*, Trinity College University of Cambridge, Cambridge, United Kingdom

Blackwell, W.J. & Chen F.W. (2009). *Neural Networks in Atmospheric Remote Sensing*, Massachusetts Institute of Technology, ISBN 978-1-59693-372-9, Lexington, USA

Deligiorgi, D.; Philippopoulos, K.; Thanou, L. & Karvounis, G. (2010). A comparative Study of Three Spatial Interpolation Methodologies for the Analysis of Air Pollution Concentrations in Athens, Greece, *7th International Conference of the Balkan Physical Union*, American Institute of Physics (AIP) Conference Proceedings Vol. 1203, pp. 445-450, ISBN 978-0-7354-0740-4

Diebold, F.X. & Mariano R.S. (1995).Comparing Predictive Accuracy, *Journal of Business & Economic Statistics*, Vol.13, No.3, (July 1995), pp. 253-263, ISSN 1537-2707

Eberly, S.; Swall, J.; Holland, D.; Cox, B.; Baldridge, E. (2004). *Developing Spatially Interpolatied Surfaces and Estimating Uncertainty*, United States Environmental Protection Agency

Fan, Q.; Efrat, A.; Koltun, V.; Krishnan, S. & Venkatasubramanian, S. (2005). Hardware-Assisted Natural Neighbor Interpolation, *Proceedings of ALENEX/ANALCO 2005 7th Workshop on Algorithm Engineering and Experiments*, pp. 111-120, ISBN 0-89871-596-2, Vancouver, Canada, 22 January, 2005

Gusten, H.; Heinrich, G.; Cvias, T.; Klasinc, L.; Ruscic, B.; Lalas, D. & Petrakis, D. (1988). Photochemical formation and transport of ozone in Athens, Greece. *Atmospheric Environment*, Vol.22, No.9, pp. 1855-1861, ISSN 1352-2310

Hardy, R.L. (1971). Multiquadric Equations of Topography and Other Irregular Surfaces, *Journal of Geophysical Research*, Vol.76, No.8, pp.1905-1915, ISSN 0148–0227

Hardy, R.L. (1978). *The Application of Multiquadric Equations and Point Mass Anomaly Models to Crustal Movement Studies*, National Oceanic and Atmospheric Administration

Hornik, K.; Stinchcombe, M. & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, Vol.2, No.5, pp. 359-366, ISSN 0893-6080

Ito, K.; De Leon, S. F. & Lippmann, M. (2005). Associations Between Ozone and Daily Mortality: Analysis and Meta-Analysis, Epidemiology, Vol.16, No.4, (July 2005), pp. 446-457, ISSN 1531-5487

Junninen, H.; Niska, H.; Tuppurainen, K.; Ruuskanen, J. & Kolehmainen, M. (2004). Methods for imputation of missing values in air quality data sets. *Atmospheric Environment*, Vol.38, No.18, (June 2004), pp. 2895-2907, ISSN 1352-2310

Kassomenos, P.; Kotroni, V. & Kallos, G. (1995). Analysis of Climatological and Air Quality Observations from Greater Athens Area. *Atmospheric Environment*, Vol.29, No.24, (December 1995), pp.3671- 3688, ISSN 1352-2310

Krige, D.G. (1951). A statistical approach to some basic mine valuation problems on the Witwasersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, Vol.52, No.6, pp. 119-139

Lalas, D.; Asimakopoulos, D. & Deligiorgi D. (1983). Sea-breeze circulation and photochemical pollution in Athens, Greece. *Atmospheric Environment*, Vol.17, No.9, pp. 6121-1632, ISSN 1352-2310

Li, J. & Heap A. D. (2008). *A Review of Spatial Interpolation Methods for Environmental Scientists*, Geoscience Australia, Retrieved from http://www.epa.gov/airtrends/studies.html

Lozano, A.; Usero, J.; Vanderlinden, E.; Raez, J.; Contreras, J.; Navarrete, B. & El Bakouri, H. (2009). Design of air quality monitoring networks and its application to NO_2 and O_3 in Cordova, Spain. *Microchemical Journal*, Vol. 93, No. 2, (November 2009), pp. 211-219, ISSN 0026-265X

Lu, G.Y. & Wong D.W. (2008). An adaptive inverse-distance weighting spatial interpolation technique, *Computers & Geosciences*, Vol.34, No.9, (September 2008), pp. 1044-1055, ISSN 0098-3004

Marshal, J.D.; Nethery, E. & Brauer, M. (2008). Within-urban variability in ambient air pollution: Comparison of estimation methods, *Atmospheric Environment*, Vol.42, No.6, (February 2008), pp. 1359-1369, ISSN 1352-2310

Mofarrah, A. & Husain, T. (2010). A holistic approach for optimal design of air quality monitoring network expansion in an urban area. *Atmospheric Environment*, Vol.44, No.3, (January 2010), pp. 432-440, ISSN 1352-2310

Myers D.E. (1994). Spatial interpolation: an overview. *Geoderma*, Vol.62, No.1-3, (March 1994), pp. 17-28, ISSN 0016-7061

Spatial Interpolation Methodologies in Urban Air Pollution Modeling:
Application for the Greater Area of Metropolitan Athens, Greece

361

Nejadkoorki, F.; Nicholson, K. & Hadad, K. (2011). The design of long-term air quality monitoring networks in urban areas using a spatiotemporal approach. *Environmental Monitoring and Assessment*, Vol.172, No.1-4 , (January 2011), pp. 215-223, ISSN 1573-2959

Okabe, A.; Boots, B.; Sugihara, K. & Chiu, S.N. (2000). *Spatial Tessellations Concepts and Applications of Voronoi Diagrams*, John Wiley & Sons Ltd, ISBN 0-471-986354, Chichester, England

Roussos, G. & Baxter, B. (2005). Rapid evaluation of radial basis functions. *Journal of Computational and Applied Mathematics*, Vol.34, No.9, (September 2008), pp. 1044-1055, ISSN 0098-3004

Samet, J.M.; Dominici, F.; Curriero, F.C.; Coursac, I. & Zeger, S.L. (2000). Fine particulate air pollution and mortality in 20 U.S. cities 1987-1994, *The New England Journal of Medicine*, Vol.343, No.24, (December 2000), pp. 1742-1749, ISSN 1533-4406

Sandwell, D. (1987). Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data. *Geophysical Research Letters*, Vol.14, No.2, (February 1987), pp. 139-142, ISSN 0094-8276

Shao, X.; Stein, M. & Ching, J. (2007). Statistical comparisons of methods for interpolating the output of a numerical air quality model. *Journal of Statistical Planning and Inference*, Vol.137, No.7, (July 2007), pp. 2277-2293, ISSN 0378-3758

Sibson, R. (1981). A brief description of natural neighbor interpolation, In: *Interpolating multivariate data*, Barnett V., pp. 21-36, John Wiley & Sons Ltd, ISBN 978-0471280392, Chichester, England

Snell, S.E.; Gopal, S. & Kaufmann R.K. (2000). Spatial Interpolation of Surface Air Temperatures Using Artificial Neural Networks: Evaluating Their Use for Downscaling GCMs, *Journal of Climate*, Vol.13, No.5, (March 2000), pp. 886-895, ISSN 1520-0442

Vicente-Serrano, S.M.; Saz-Sanchez, M.A. & Cuadrat J.M. (2003). Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature, *Climate Research*, Vol.24, No.2, (July 2003), pp. 161-180, ISSN 1616-1572

Wahba G. (1990). *Spline Models for Observational Data*, Society for Industrial and Applied Mathematics, ISBN 978-0-898712-44-5, Philadelphia, USA

Wahba, G. & Wendelberger, J. (1980). Some new mathematical methods for variational objective analysis using spiline and cross validation. *Monthly Weather Review*, Vol.108, No. 8, (August 1980), pp. 1122-1143, ISSN 1520-0493

Watson, D.F. & Philip, G.M. (1984). Systematic Triangulations, *Computer Vision, Graphics, and Image Processing*, Vol.26, No.2, (May 1984), pp. 217-223, ISSN 0734-189X

Webster, R. & Oliver, M.A. (2007). *Geostatistics for Environmental Scientists*, John Wiley & Sons Ltd, ISBN 978-0-470-02858-2, Chichester, England

Willmott C.J. & Matsura K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, Vol.30, No.1, (December 2005), pp. 79-82, ISSN 1616-1572

Willmott C.J. (1982). Some Comments on the Evaluation of Model Performance. *Bulletin of the Maerican Meteorological Society*, Vol.63, No.11, (November 1982), pp. 1309-1313, ISSN 1520-0477
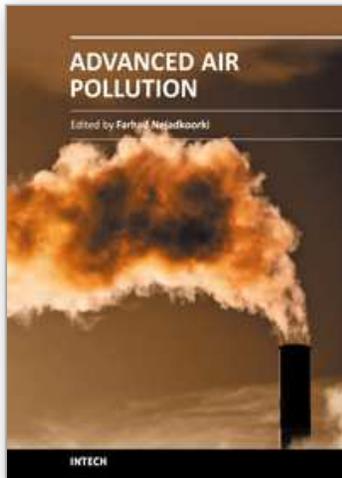
Advanced Air Pollution

Edited by Dr. Farhad Nejadkoorki

Leading air quality professionals describe different aspects of air pollution. The book presents information on four broad areas of interest in the air pollution field; the air pollution monitoring; air quality modeling; the GIS techniques to manage air quality; the new approaches to manage air quality. This book fulfills the need on the latest concepts of air pollution science and provides comprehensive information on all relevant components relating to air pollution issues in urban areas and industries. The book is suitable for a variety of scientists who wish to follow application of the theory in practice in air pollution. Known for its broad case studies, the book emphasizes an insightful of the connection between sources and control of air pollution, rather than being a simple manual on the subject.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds