Satellite-assisted Management of Air Quality

SMAQ

Deliverable D6.1:

Optimal air pollution monitoring network configuration

LIFE ENVIRONMENT PROGRAMME
LIFE05 ENV/GR/000214

TASK 6: Monitoring Network Design

TASK 6.1: Optimisation and design of a multi-pollutant monitoring network

Report Date: 25/09/08
Bibliographical Information

Project: Satellite-assisted Management of Air Quality - SMAQ

Subject: Optimal air pollution monitoring network configuration

LIFE ENVIRONMENT PROGRAMME
Contract No. LIFE05 ENV/GR/000214
Duration of Contract: 01/10/2005 - 31/03/2009
TASK: 6. Monitoring Network Design
TASK: 6.1 Optimisation and design of a multi-pollutant monitoring network
Editing Partner: CERTH
Other Partners: 
Report Date: 25/09/08

Pages: 37 (including figures, tables, attachments)

Key Words: Monitoring, network Air Pollution

Contact Person Editing Partner
Name: Anastasios Karabelas
Phone: +302310498181
Fax: 
e-mail: karabaj@cperi.certh.gr

Authors Editing Partner
Name: D. Sarigiannis
Phone: 39 0332 785373
Fax: +39 0332 785294
e-mail: dimosthenis.sarigiannis@ec.europa.eu
Table of Content

**EXECUTIVE SUMMARY** ............................................................................................................................................ 4

**OBJECTIVES**................................................................................................................................................................. 7

1. **INTRODUCTION** ...................................................................................................................................................... 7

2. **METHODOLOGY DESCRIPTION**............................................................................................................................ 11
   2.1.1. **FIRST OBJECTIVE: REPRESENTATION OF SPATIAL-TEMPORAL PATTERNS** ......................................................... 11
   2.1.2. **SECOND OBJECTIVE: DETECTION OF VIOLATIONS OVER AMBIENT STANDARDS**.............................................. 16
   2.1.3. **THIRD OBJECTIVE: RISKS TO HUMAN HEALTH** ................................................................................................. 18
   2.1.4. **FOURTH OBJECTIVE: DAMAGE TO CULTURAL HERITAGE** ................................................................................ 19
   2.1.5. **FIFTH OBJECTIVE: RECIPROCAL RELATIONSHIP BETWEEN LAND-USE AND POLLUTION** ................................. 20
   2.1.6. **SIXTH OBJECTIVE: MINIMUM COST** ............................................................................................................. 21
   
   Capital cost........................................................................................................................................................................... 21
   Operation/Maintenance Cost.............................................................................................................................................. 22
   Analyzers ............................................................................................................................................................................... 22
   Replacement Cost ................................................................................................................................................................. 23

   2.2. **INFORMATION FUNCTION APPROACH** ........................................................................................................... 24
   Structure of the Information Function ................................................................................................................................ 24

3. **RESULTS AND DISCUSSION** ................................................................................................................................. 29

   Network Reliability and Operability .................................................................................................................................. 33

4. **CONCLUSIONS** ......................................................................................................................................................... 33

**REFERENCES**.................................................................................................................................................................... 34
EXECUTIVE SUMMARY

Air quality is influenced by a variety of factors related to natural and anthropogenic emissions, local meteorological conditions, as well as global climatic variations. It has directly measurable impacts on public and ecosystem health and its effective monitoring and management represents a significant challenge.

Fundamental aim of this study, carried out within the framework of the Satellite-Assisted Management of Air Quality (SMAQ) project carried out in the region of Western Macedonia, Greece, is to provide decision support for air quality management at the conurban up to regional scales based on multi-criteria analysis and interactive optimization. Key objective of the present work is to minimize uncertainty in decision-making regarding operational air pollution control and abatement measures.

This project aims to

- establish performance indicators with regard to the allocation of monitoring stations, and
- implement a new methodology for developing new monitoring networks for effective air quality monitoring network design and implement this concept for streamlining existing networks.

Dispersion models together with ground-based measurements have been used so far in relative methodologies found in the relevant literature. The innovation of this work lies in the use of new techniques of satellite-based Earth observation; this method can provide spatial information on pollution over urban and non-urban areas alike. It therefore reduces the uncertainties and errors currently associated with dispersion models. For a comprehensive pollution investigation, one needs to examine not only the pollution distribution but also the geographical distribution of its causes and of its effects on the environment and public health. Data on these phenomena should be spatially referenced and they are ideally organised in thematic layers in order for them to be easily comparable and updated. This is precisely the role of GIS (Geographical Information System), which is a powerful set of tools for storing, retrieving, transforming and displaying spatial data from the real world.
The interest of the approach is to represent the patterns of regional air quality, at a minimum of an overlap, meeting at the same time multiple objectives and accounting for different types of pollutants. Objectives of monitoring networks designed following the methodology outlined in this report are the following:

- Reliable representation of air pollution spatial-temporal patterns
- Detection of pollution exceedances over ambient air quality standards
- Monitoring air pollution in areas where it represents the major risks to human health
- Monitoring air pollution in areas where it represents the major damage to cultural heritage
- Identifying the reciprocal relationship between land-use and pollution
- Minimization of overall cost of the network.

The operational objective of the application of the methodology developed herein in W. Macedonia is to assess the efficacy and optimize the regional monitoring network of ambient air quality.

The study was implemented in the following stages:

1. The selection of the objectives to be met from the monitoring network. This was carried out through an extensive literature search, in order to take into account all possible objectives used in other similar exercises, including the requirements set in the Air Quality framework directive of the European Commission.
2. The establishment of an Information Function (IF), which would represent the interests of those responsible for regional air quality management and for the attainment and maintenance of the National and EU Ambient Air Quality Standards. The IF combines information on violation of ambient air quality thresholds, land-use, human population density, cultural property density and cost.
3. The structure of the algorithm takes into consideration the dependence between locations in the amount of information they report, by applying linear regression analysis.
4. Application of the method in the region of W. Macedonia and in particular the area delineated by the urban areas of Florina, Ptolemais and Kozani.
5. Establishment of performance indicators:
   - exposure of population to pollution,
• compliance to ambient air pollution standards,
• gain of information by locating a station at a certain place, and
• area coverage.

6. Assessment of the existing air quality monitoring network in W. Macedonia under the same conditions adopted for designing the optimal network configuration.

7. Assessment of the optimal network reliability and operability, in the case of one station failure.

This approach provides a joint solution to the problem of optimum number and configuration of ambient air quality monitors, and it can be used to streamline existing networks. It is therefore a significant addition to the tools available to those responsible for the design and operation of air quality monitoring networks. It provides monitor site assignments designed to maximize monitor coverage of a region as well as maximum gain of information, in terms of exposure of population, land-use and vulnerable cultural receptors. It gives priority to areas with high emission levels without neglecting weaker sources. The method is easy to use and inexpensive in computer resources.

The integrated methodology described herein uses the spatial distribution of primary pollutants calculated by appropriately processing satellite imagery as input regarding the distribution of pollution. It may, however, use any continuous in space distribution of pollutants as input, independent of the technique used to derive it. Other state-of-the-art techniques that may provide similar information include numerical models of atmospheric pollutants dispersion, and spatial interpolation (e.g. kriging) of ground-based point measurements. The MOON method is therefore flexible and can be adapted to the needs and the technical capabilities of the end-user.
OBJECTIVES

Air quality is one of the 13 “quality of life” indicators in the EU. Therefore, it has become an urgent necessity to locate ambient air monitoring stations efficiently, in order to provide information for environmental control programs. This study addresses the problem of establishing adequate performance indicators to assess the allocation of monitoring stations in space and suggests a new methodology for developing new networks of stations for effective air quality monitoring network design and streamlining an existing network.

The main innovative aspect lies in the use of results of a prototype image processing method for high-resolution satellite images, from which information on air pollution distribution over urban and non-urban areas can be derived. The data derived from satellite images, after proper handling, feed an air-pollution GIS (Geographical Information System), which is the most appropriate software platform to produce input data for the optimization algorithm delineated below, as well as for the final representation of the results.

The interest in this case is to represent the patterns of regional air quality, at a minimum of an overlap, meeting at the same time multiple objectives and accounting for different types of pollutants. This approach, which provides a joint solution to the problem of optimum number and configuration of ambient air quality monitors, is shown to be a valuable addition to the methods already available to the air quality monitor network designer.

1. INTRODUCTION

Air pollution sampling site selection is one of the most important and yet most vexing of the problems faced by those responsible for regional and conurban air quality management and for the attainment and maintenance of National Ambient Air Quality Standards. Since one can not hope to monitor air
quality at all locations at all times, selection of sites to give a reliable and realistic picture of air quality becomes a major issue and at the same time a difficult task. The location (configuration) and the number of stations may be based on many factors, some of which may depend on limited resources, federal and state regulations and local conditions. The combination of these factors has made air quality surveys more complex, requiring comprehensive planning to ensure that the prescribed objectives can be attained in the shortest possible time and at the least cost. Furthermore, the choice and siting of the measuring network represents a factor of significant economic relevance for policymakers. In view of the fact that costs of equipment, maintenance and operating personnel are increasing dramatically, the possibility of optimizing monitoring design, is most attractive to the directors of air quality management programs.

The design of a measurement network for air pollution monitoring involves both spatial and temporal sampling considerations (Chen and Seinfeld, 1975). Most newly designed networks, however, use exclusively automatic continuous measuring equipment. These devices produce time-averaged values of pollutant concentrations measured over short intervals (on the order of a few seconds). As a consequence, these values can be considered as instantaneous for all practical purposes. This reduces the original spatial and temporal problem to one of space alone, that is the location of stations.

The strategy for, and design and operation of air quality monitoring networks are determined by the objectives of the monitoring activities. These objectives for air pollution monitoring networks have been reported frequently in the literature (US EPA, 1971; WHO, 1977; Seinfeld, 1972; Ott 1977) and generally fall into one of the following categories.

1. Creation of data bases containing current spatial-temporal air quality in order to permit the analysis of long trends in air pollution or for other research purposes such as the validation of mathematical models for air pollutant diffusion, transport and transformation.
2. Evaluation of the effectiveness of control strategies and activation of episode controls in highest pollution concentration areas.
3. Evaluation of risks to human health with attention to densely populated areas.
4. Determination of the risk of damage to vulnerable receptors (e.g. historic and/or artistic valuable property, vegetation).
5. Development and updating of a data base for land-use planning
6. Compliance with ambient air quality standards (detection of violations of those standards)
7. Control of emissions from singularly important sources, e.g. thermal or nuclear plants.

Generally implied within these objectives is also the minimization of the cost of the network.

In addition to the above objectives, a series of EU directives and decisions sets the requirements on how monitoring and assessment of air pollution should be carried out in the EU member states and associated countries. The present compound-specific Directives for \( \text{SO}_2 \) (89/427/EEC), \( \text{NO}_2 \) (85/203/EEC) and \( \text{O}_3 \) (92/72/EEC) issued during the period 1989-1995, require in principal that all exceedances of the limit values are detected. The Draft Framework Directive on Ambient Air Quality (FWD) indicates that air quality is assessed relative to the limit values, which are in effect at any time. Furthermore, the FWD requires that measurement of AQ is mandatory in agglomerations with population density higher than a certain value per \( \text{km}^2 \), which should be decided by the member states.

Spatial coverage is another prerequisite for AQMN, since the term “assess” involves full description of the air quality. All the aforementioned are the criteria behind the design and establishment of a Regulatory Network, namely of the network aimed to report compliance with the EU air quality directives.

Primary studies on optimization of Air Quality Monitoring Networks (herein referred as AQMN) have been primarily concerned with a few objectives of interest. But the recent literature shows a growing interest for attempting the problem of multiple objectives in the AQMN design. Naito and Ochiai (1981) and Munn (1981) consider that the incorporation of multiple objectives is extremely important to lay the foundation of the future practices towards AQMN optimization. Few methodologies exist in the literature that can accomplish the task of designing a network capable of fulfilling all of the objectives stated above (Modak et al. 1984, Trujillo et al. 1991, Smith et al. 1979). Yet, the interest of air pollution policy makers is shifted towards this direction.
At the same time, it may be essential to discuss, in brief, the philosophy of common monitoring sites. In general, it has been observed that every moderate or large scale multi-pollutant AQMN adheres to a policy of retaining a maximum number of common sites, i.e., sites where a number of pollutants are simultaneously measured. Networks with pollutant specific sites are rare. Economic and assessment considerations (Hickey et al. 1971, Stern 1976) appear to be the principal reasons this. For a common site network, the costs of installation and maintenance of the monitoring sites are shared amongst various pollutants. Hence therefore the total cost of monitoring is greatly reduced. Besides, in many instances, a common-site consideration is preferred due to the following reasons:

1. Exposure assessment is never complete when only one pollutant is measured. It is well known that several pollutants are synergistic and therefore it is mandatory to measure certain pollutants at a common site.

2. If more than one pollutant is measured at a common site, then missing value estimation for a pollutant is possible especially when cross-correlations between the other pollutants are quite high and already established.

Attempts to produce optimal AQMN based on multiple pollutants have, so far, been restricted to a limited number of references. In fact, present multi-pollutant approaches focus only on the combination of SO$_2$ with smoke (Green, 1966; Modak et al. 1984) due to their high degree of affinity. This is up to a certain point justifiable. Every pollutant has its characteristic variability due to its specific emission sources, rates of diffusion as well as transformations; and therefore it is quite logical that the optimal number and configuration are pollutant specific. For instance, pollutant A having higher spatial variability than pollutant B may require more monitors in order to correctly represent its regional patterns.

As yet no references have been cited in the literature combining O$_3$, PM10, PM2.5, PM1 and CO with SO$_2$ and NO$_2$ for the design of an AQMN. This may be attributable to the fact that for the case of O$_3$ for instance, this pollutant behaves as a peripheral one, i.e. has high values in the districts of a city and not within it. Thus high concentrations of SO$_2$ and NO$_2$ are actually associated with low values of O$_3$ and vice versa. The present approach, however, thanks to the aerosol model transformation, enables the
coupling of NO$_2$ and SO$_2$ as criteria pollutants for the AQMN design. Pollution maps of O$_3$ and CO, created through kriging of available O$_3$ and CO concentrations measured from the present monitoring network, serve as an additional objective in the methodology described below. At the same time, pollution maps of SO$_2$ and NO$_2$ based on a 4-year data series (2001-2004) are taken into consideration for the design of the recommended monitoring network. Since none of the available monitoring stations covers the measurement of fine particles (PM2.5, PM1), respective references in the literature are used as guidelines for the integration of fine particles into the optimal monitoring network. Coarse particles (PM10) are already being measured by a few stations in the area; their network will be completed by new ones. This will result in an integrated methodology for optimal AQMN, combining the seven most important air quality pollutants-indices.

### 2. METHODOLOGY DESCRIPTION

Initially, a grid is superimposed over a map of the area of interest. The intersections of the grid lines every 1 km are used as potential sites, having thereafter a surface of 1 km$^2$. The next step was to establish the objectives, on which the optimal configuration of the air pollution monitoring stations would be based and evaluated. The combination of the objectives mentioned in the Introduction Section yielded the following six objectives for our case.

**2.1.1. First Objective: Representation of Spatial-Temporal Patterns**

It is well known that conventional analytical air pollution observations produce time series of “space-deficient” point data, unless taken from an extremely large number of costly stations. Pollutant concentrations at any point in the domain have been, so far, estimated by interpolation from concentration data at locations where stations exist, through spatial correlation analysis and dispersion models. New techniques of Earth observation from satellite, however, can offer spatial information on pollution, therefore reducing the uncertainties and errors associated with dispersion models.
The approach of analysis of satellite images has already been applied with success in the particularly polluted greater Athens Area, with a view to modernise, optimise the siting and extension of the ground-based air pollution network of the city (Programme Life, 1997). Briefly, maps indicating the horizontal distribution of airborne particulate pollution can be produced by using high resolution satellite images, produced from high spatial resolution (HSR) sensors on board satellites. These sensors are sensitive to the optical spectrum, which extends from approximately 0.3 to 14 μm (including UV, visible, near-, mid- and thermal infrared wavelengths) and is termed “optical” because lenses and mirrors can be used to refract and reflect energy. In this spectral domain, the signal recorded by the sensors is emitted or reflected by the Earth and the atmosphere. A change in the composition of the atmosphere (e.g., by the presence of pollution) modifies the signal through the interaction between radiation and the atmospheric components. These mechanisms induce optical atmospheric effects on the images that may affect the signal recorded by the sensor in two ways: geometrically and radiometrically. Contrarily to geometric modifications due to light refraction that are not intense enough to be observable, some of the radiometric modifications linked to light absorption, scattering and backscattering by atmospheric molecules and particles (of diameters 0.1-3 μm) are observable by satellite sensors. Since HSR satellite observations are performed in the so-called atmospheric windows, where the extinction coefficient due to molecular absorption ($K_{a}^{abs}$) is minimum, pollution observations will be linked to particulate absorption ($K_{a}^{abs}$) and to elastic scattering (the non-elastic scattering mechanisms, e.g. Raman, can be observed only in very narrow spectral bands by means of lidars). Elastic scattering depends on the molecular (Rayleigh) scattering coefficient ($K_{a}^{scat}$) and the particulate (Mie) scattering coefficient ($K_{a}^{scat}$). These are both wavelength dependent mechanisms except from Mie scattering when caused by particles too large compared to the observed wavelength (e.g., cloud particles). Comparing these two coefficients, the values of $K_{a}^{scat}$ are high even when measuring in unpolluted atmospheres. Therefore this coefficient is not sensitive enough to detect pollution variations, and thus $K_{a}^{scat}$ and $K_{a}^{abs}$ are most often used as quantitative evaluators of pollution levels. Variations in $K_{a}^{scat}$ and variations in $K_{a}^{abs}$ introduce different optical effects on the images, due to the distinct mechanisms of scattering and absorption. The effect of scattering is a blurring of the image due to
contrast reduction (i.e., dark targets appear brighter and bright targets appear darker) while the effect of absorption is an obscuring of the image due to attenuation of the radiation (i.e., all targets appear darker). The two effects can be mathematically decoupled since “blurring” corresponds to an additive, and “obscuring” to a multiplicative factor. Subsequently, these effects allow the computation of the integrals of $K_{a}^{\text{scat}}$ and of $K_{a}^{\text{abs}}$ from the Earth’s surface to the orbital altitude $z$ for a given wavelength $\lambda$.

This is precisely the particulate optical thickness $\tau_{a}$, and it is:

$$\tau_{a}(z, \lambda) = \int_{0}^{z} (K_{a}^{\text{scat}} + K_{a}^{\text{abs}}) (z', \lambda) \, dz'$$

(2)

The use of the particulate optical thickness, is appropriate in air pollution measurements because:

- The problem of particulate pollution is very sensitive at present especially after new scientific evidence on health effects of small particles (Lipfert, 1994; Pope et al., 1995).
- The presence of particles in the atmosphere always causes a reduction of the optical thickness (Horvath, 1981). This reduction is strongly correlated with the concentration of small particles (Waggoner et al., 1981)
- In photochemical pollution, light extinction is due to particle, while only the yellow-brownish coloration of the smog is due to NO$_2$ (Waggoner et al., 1980).

The magnitude of the optical thickness varies according to the spectral band used. In the UV spectral domain the atmosphere is practically opaque due to Rayleigh scattering and ozone absorption, and in the near-infrared it is too transparent for pollution observations. The visible domain is most appealing for evaluating optical effects of pollution. The mid-infrared can be used to distinguish hot pollution sources (e.g., actively burning fires) inside an observed haze (Lillesand et al. 1994), and the thermal infrared can be taken into account to increase confidence in pollution mapping (Sifakis et al., 1992).

Wavelength dependence of the optical effects of pollution allows the detection of even tenuous pollution, such as smoke wreaths from factories or diluted urban haze. It also allows the differentiation of pollution from natural clouds. The following methods can be used:

- the “ocean method” above clear water, using infrared satellite data (Griggs, 1975);
the “brightness method” above land or water features, using data in the visible spectrum (Fraser et al., 1984);

• the “contrast-reduction method” over land (Tanre, 1988) or over mixture of land with water (Kergomard, 1989);

• a method using long-wavelength visible data exclusively over vegetation features (Kaufman, 1988).

The results of the present work have been based on a method developed within the frame of the EC-funded research projects ICAROS and ICAROS NET. This method determines the nature and concentration of pollutants correlated with certain optical thickness values and is delineated below.

Although numerous chemical species have been identified in secondary aerosols, the most prevalent are sulfate, nitrate, ammonium and water. Therefore, the formation of secondary aerosols from gaseous H$_2$SO$_4$, HNO$_3$, NH$_3$ and H$_2$O has been the subject of much theoretical and experimental investigation. Mathematical models have been developed to study aerosol formation in plumes and in conurban areas. These atmospheric gas-aerosol equilibrium models calculate the partitioning of total sulfate, nitrate, ammonium and water in an air parcel assuming that the system is in thermodynamic equilibrium (as expressed by its minimum Gibbs free energy). Knowledge of temperature, relative humidity and gaseous ammonia concentration is assumed. In most ambient situations, it is expected that assumptions required for equilibrium to hold are valid since the characteristic time for mass transfer to and from an aerosol particle is of the order of a fraction of a second (Seifeld, 1980; Schwartz and Freiberg, 1981).

To begin with, the concentrations of sulfate and nitric acid are estimated through the oxidation of sulfate dioxide and nitrate dioxide. The concentrations of these two pollutants are obtained from ground-based measurements. The transformation of SO$_2$ and NO$_2$ into their respective acids is given by the following equations:

\[
\text{SO}_2 + \text{OH} \xrightarrow{M, k_1} \text{HOSO}_2 \quad k_1=9 \times 10^{-13} \text{ cm}^3/(\text{mole s}) \text{ (at 298 } \text{K})
\]

\[
\text{HOSO}_2 + \text{O}_2 \rightarrow \text{HO}_2 + \text{SO}_3 \text{ (rapid)}
\]
SO$_3$ + H$_2$O$_g$ $\rightarrow$ H$_2$SO$_4$ (rapid)

NO$_2$ + OH $\xrightarrow{M,k_3}$ HNO$_3$ $k_3$ = 1.1$ \times$ 10$^{-11}$ cm$^3$(mole$^{-1}$s$^{-1}$) (at 300 oK)

By introducing the concentrations of these two acids in the ARES thermodynamic equilibrium model (based on an aerosol model MARS developed by Seinfeld et al. 1986) and using also as input variables the ambient relative humidity and temperature, the mass and the chemical composition of atmospheric aerosols containing sulfate, nitrate, ammonium and water are estimated.

The logarithmic transformations of aerosol concentrations of NH$_4^+$, NO$_3^-$, SO$_4^{2-}$ and water are used as input variables in a forward stepwise regression analysis, so as to determine the most important variables correlated with the optical thickness values. The linear regression coefficient between optical thickness values and the aerosol phase of the aforementioned Figure 1. Correlation between observed values of $\delta_{opt}$ and the predicted ones from ln[ANH$_4$] and ln[aerosol] for the Florina-Ptolemais-Kozani axis.

The inorganic species has shown a very good correlation. For the Florina-Ptolemais-Kozani axis, the variance of $\delta_{opt}$ is at 88% (R=0.939) explained by the natural logarithmic values of NH$_4$ and the total aerosol and is given by the following equation:

$$\delta_{opt} = 0.5639 \ln [ANH_4] - 0.1413 \ln[aerosol] \quad (E1)$$
It should be mentioned at this point, that the concentration of aerosol NH$_4$ depends on the concentrations of the other aerosol species, therefore the latter correlation can be considered as indicative of the overall aerosol chemistry. At the same time, the negative coefficient of aerosol in (E2) does not indicate a negative relationship between $\delta_{\text{opt}}$ and the total concentration of aerosol, since the two variables of E1, i.e. ANH$_4$ and aerosol are not independent.

The optical thickness values have been extracted by images obtained by high spatial resolution (HSR) sensors on board in-polar-orbit Earth satellites at the time of their daily overpass (SPOT and LANDSAT). Through an image processing algorithm, the spatial distribution of optical thickness over the whole area is obtained. It is hence necessary to extract the single values of $\delta_{\text{opt}}$, corresponding to the sites of monitoring stations for which ground-based measurements of pollutants concentrations are available. Furthermore, for each image of successful satellite passage, statistical analysis is applied on each site of an existing active monitoring station. To achieve this, satellite data are first geo-referenced and geographically corrected and then imported into a Geographic Information System (GIS), in order for them to be integrated with land cover and land use maps of the areas of interest and transformed in a format suitable for the optimization algorithm.

2.1.2. Second Objective: Detection of Violations over Ambient Standards

Several objective functions have been used in the past to measure the ability of a network to detect violations of standards. The approach taken here considers the potential of a monitoring site for the detection of violations in terms of violation scores. This choice was due to the need to incorporate in the multi-objective optimisation algorithm the semantic distance (the relative significance) of the various alternative solutions, i.e. the alternative station locations in terms of meeting the objectives. A location with a high violation score is then considered to have a high potential for detection of violations. The computation of violation scores is essentially a weighted scoring of optical thickness values, which correspond to concentrations of SO$_2$ and NO$_2$ above prescribed thresholds. These scores are therefore dependent on,
1. The threshold levels.
2. The weighing factors between each threshold range and the weighing function.

A decision on thresholds and weighing factors (i.e. severity) is indeed pollutant-specific and further dependent upon the averaging time and populations concerned. Several weighing functions have been reported such as linear ones, segmented linear, non-linear, segmented non-linear etc. (Ott, 1978). In our case, a segmented nonlinear weighing function proposed by Modak et al.(1985) is used. The violation score for each candidate location is given by the equation,

\[ V_i = \sum_{i=1}^{T} \sum_{k=1}^{N_t} \left( w_{k+1} - w_k \right) \frac{(x_i - x_k)X}{(x_{k+1} - x_k)} \]

where

\( V_i \) = the violation score for the \( i^{th} \) candidate location,

\( w_k \) = the weighing factor corresponding to threshold \( x_k \),

\( x_k \) = the \( k \)-th threshold,

\( X = 0 \) if \( (x_i - x_k) \leq 0 \),

\( X = 1 \) otherwise

\( N_t \) = the total number of thresholds,

\( T \) = the total number of simulated observations,

The weighing function, associated with the violation scores, is important because the severity of threshold violations reported by two monitoring stations may differ.

In order to assist public authorities to manage and reduce health hazards and other risks from air pollutant, the WHO, EU and national authorities publish guidelines and limit values for most of the common pollutants. Based on those guidelines, the values for thresholds for SO\(_2\) and NO\(_2\) were chosen as seen in Table II. The optical thickness (\( \delta_{\text{opt}} \)) values that correspond (via the correlation analysis mentioned above) to those values are also listed. In addition, the air pollution index (weighing factor)
ranging from 0 to 4 according to the severity of threshold exceedance is also given. The dose-response function underlying the assignment of the weighing factors is fitted to a binomial. This mathematical form penalises non-linearly the exceedance of higher threshold values with regard to exceedance of lower ones. In the absence of authoritative epidemiological data regarding air pollution effects on ecological vulnerability and public health, the network design adheres to the precautionary principle in environmental management.

TABLE II
Air Pollution Index assigned to optical thickness values

<table>
<thead>
<tr>
<th>NO₂ (μg/m³)</th>
<th>SO₂ (μg/m³)</th>
<th>Threshold (δopt)</th>
<th>Weighing Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>133</td>
<td>80</td>
<td>0.47</td>
<td>0.5</td>
</tr>
<tr>
<td>200</td>
<td>120</td>
<td>0.91</td>
<td>1.0</td>
</tr>
<tr>
<td>217</td>
<td>130</td>
<td>1.10</td>
<td>2.5</td>
</tr>
<tr>
<td>250</td>
<td>150</td>
<td>1.63</td>
<td>3.0</td>
</tr>
<tr>
<td>300</td>
<td>180</td>
<td>3.03</td>
<td>4.0</td>
</tr>
</tbody>
</table>

2.1.3. Third Objective: Risks to human health

Population exposure could only be assessed via detailed epidemiological studies and personal monitoring and not solely based on the data at fixed outdoor air quality monitoring stations. The EU project EXPOLIS is the first organised attempt to provide a large, population-based exposure study in Europe. Its objectives included:

(a) assessment of exposures of European urban population to major air pollutants;
(b) analysis of the personal and environmental determinants and interrelationships to these exposures;
and,
(c) development of a European database for simulation of air pollution exposures.
Substantial differences have been found between the air pollution levels measured at the fixed outdoor monitors and levels to which people are actually exposed (Silverman et al. 1982). It may not be appropriate then to expect that a fixed outdoor AQMN, optimized for the maximization of population dosage product (Darby et al. 1974) would provide a rational basis for the assessment of health effects due to pollution. It may only be considered as a cursory tool, to place the monitoring stations in the populated areas of probably higher relative concern.

2.1.4. Fourth Objective: Damage to cultural heritage

Damage to cultural property in museums, libraries and historic buildings is caused by environmental pollution, outdoors as well as indoor (EC Cultural Heritage Newsletter on Research, 1997). Many attempts have been made in the past to provide some qualitative and quantitative characterization of the pollution profiles over building surfaces and inside them. As yet, no guidelines for the evaluation of exposure of historic and/or artistic valuable property have been established. Although many issues are now well understood, additional research is needed to determine materials’ susceptibility with respect to composition, relative humidity and pollutants concentrations.

Thereafter, the density of monuments per km² has been considered as an additional objective for optimal location of an air pollution monitoring network. Yet, it should be mentioned that site identification (in terms of exact geographic co-ordinates) of the most important cultural monuments in the area of study was not feasible. The only available information was the number of identified monuments in the area, with no information on the type or the relative importance of the monument. Therefore, a normalized grading system was used, whereby the municipalities of W. Macedonia were sorted with regard to their relative importance as measured by the number of catalogued cultural monuments existing in their territory.
2.1.5. Fifth Objective: Reciprocal relationship between land-use and pollution

The land-use in W. Macedonia was found in the EU land cover database, CORINE. Through these maps, information on the distribution of pollution sources and the land cover can be derived. Land-use categories were divided in three levels. The codes of the last level correspond to the exact type of land cover. In total, there are 44 land-use codes. In the present study, we have considered 13 categories, listed in Table III, as the most important ones and we have assigned a weight to each one of them, according to the importance of placing a monitoring station in the grid-point that they represent.

**TABLE III**

CORINE land cover nomenclature

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Artificial surfaces</td>
<td>1.2. Industrial, commercial and transport units</td>
<td>1.2.1 Industrial or commercial units</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.2.4 Airports</td>
<td>10/1</td>
</tr>
<tr>
<td></td>
<td>1.4. Artificial non-agricultural vegetated areas</td>
<td>1.4.1 Green urban areas</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.4.2 Sport and leisure facilities</td>
<td>0.25</td>
</tr>
<tr>
<td>2. Agricultural areas</td>
<td>2.1. Arable land</td>
<td>2.1.1. Non-irrigated arable land</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.1.2. Permanently irrigated land</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.1.3. Rice fields</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>2.2. Permanent crops</td>
<td>2.2.1. Vineyards</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.2.3. Olive groves</td>
<td>0.3</td>
</tr>
<tr>
<td>3. Forests and semi-natural areas</td>
<td>3.1. Forests</td>
<td>1.1.1. Broad-leaved forest</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.1.2. Coniferous forest</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.1.3. Mixed forest</td>
<td>0.2</td>
</tr>
<tr>
<td>4. Wetlands</td>
<td>4.1. Inland wetlands</td>
<td>4.1.1. Inland marshes</td>
<td>0.2</td>
</tr>
</tbody>
</table>
2.1.6. Sixth Objective: Minimum cost

An important objective in the design of the network is the minimization of its total cost, or rather, the maximization of its cost-effectiveness. The approach undertaken here was to consider that a combination of some stations from the present network and the addition of some new stations would provide the optimal air pollution monitoring network. The overall cost analysis is based on economic data provided by the JRC. The economic assessment of an air monitoring station is basically focused on determining the annual:

- Capital cost
- Maintenance/Operation cost
- Replacement cost.

**Capital cost**

The general cost for station installation and equipment is reckoned to 120,000€ per station. However, assuming that this amount will be covered by bank loan, with an annual rate of interest of 3-5% and 10 years of pay-off time, the annual installment for a station is given by the equation:

\[ A_{eq/\text{year}} = C_{eq} \left( \frac{i(1+i)^n}{(1+i)^n-1} \right) \]

where

- \( C_{eq} \) = capital cost of a new station
- \( C_{eq} \) = Cost of PM, SO\textsubscript{2} and NO\textsubscript{2} analyzers in €, for a fixed station
- \( i \) = annual rate of loan interest
- \( n \) = years for paying off the loan

For a fixed station, the following assumption is made: both SO\textsubscript{2} and NO-NO\textsubscript{2}-NO\textsubscript{x} analyzers are replaced before the implementation of the plan.
The coefficient \[
\left[ \frac{i \cdot (1 + i)^n}{(1 + i)^n} - 1 \right]
\] is called capital recovery coefficient and is used for the estimation of paying off a loan by equal installments every year.

**Operation/Maintenance Cost**

The annual operation and maintenance (O/M) costs are given in detail in Table IV.

### TABLE IV

Annual operational cost data for the current monitoring station network

<table>
<thead>
<tr>
<th>Cost Element</th>
<th>Cost (in €)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrical Power Supply Cost</td>
<td>8000</td>
</tr>
<tr>
<td>Phone fixed cost</td>
<td>30000</td>
</tr>
<tr>
<td>Management fee for authorized personnel (2 persons)</td>
<td>50000</td>
</tr>
<tr>
<td>Routine maintenance by authorized company</td>
<td>50000</td>
</tr>
<tr>
<td>Air coolers maintenance</td>
<td>4000</td>
</tr>
<tr>
<td>Gas calibrator</td>
<td>4000</td>
</tr>
<tr>
<td>Unexpected</td>
<td>2000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>148000</strong></td>
</tr>
<tr>
<td><strong>Analyzers</strong></td>
<td></td>
</tr>
<tr>
<td>NO-NO\textsubscript{2}-NO\textsubscript{x} analyzer</td>
<td>5000</td>
</tr>
<tr>
<td>SO\textsubscript{2} analyzer</td>
<td>5000</td>
</tr>
<tr>
<td>PM analyzer</td>
<td>5000</td>
</tr>
</tbody>
</table>
Assuming an increase in operational cost of 3.5% every year due to inflation, the annual cost for O/M, both for a present station and a new station, is estimated by the equation:

\[
A_{\text{op/year}} = C_{\text{op}} \left[ \frac{(1 + j) \left[ (1 + j)^T - 1 \right]}{T \cdot j} \right]
\]

where

- \(C_{\text{op}} = 690/15 \text{ ML}\)
- \(j = \text{annual increase in operational cost} (=0.035)\)
- \(T = \text{lifetime of project, in years}\)

The lifetime of the project is assumed to be 20 years.

**Replacement Cost**

It is assumed that the useful life of the automatic analyzers for PM, SO\(_2\) and NO\(_2\) are 5 years, which means that every five years, throughout the whole lifetime of the project, the analyzers have to be replaced. The annual cost for the replacement of analyzers, assuming also an annual increase of 3.5% in the cost of automatic analyzers is given by the equation:

\[
A_{\text{rep/year}} = (C_{\text{SO2 analyzer}} + C_{\text{NO2 analyzer}}) \frac{(1 + j)^5 + (1 + j)^{10} + (1 + j)^{15}}{T}
\]

Therefore, the total annual cost for one monitoring station will be given by the sum of

\[
C = A_{\text{eq/year}} + A_{\text{op/year}} + A_{\text{rep/year}}
\]
2.2. Information Function Approach

In order to meet the objectives described above we have developed a new approach, inspired by the work of Modak (1985), in order to determine optimal air quality monitoring network configurations. This approach is based on the establishment of an information function (IF) that encompasses the multiple objectives of the organization responsible to design the network. A general form of the IF could be described as,

$$\text{IF} = f(d_1, d_2, d_3, \ldots, d_n)$$

where $d_1, d_2, d_3, \ldots, d_n$ are the decision variables related to each one of the objectives.

In our case, the IF could be expressed as,

$$\text{IF} = f(V, L, M, P, C)$$

where

V = the violation score  
L = the land-use score  
M = the monuments density  
P = the population density  
C = the cost function

All these variables are calculated separately for each grid-point in the area of interest, which is treated as a candidate location for placing a monitoring station.

Structure of the Information Function

Several forms of IF could be used in order to combine the four aforementioned variables (V, L, M, C). Yet, the form most appropriate for our case should be,
The violation of certain pollutant thresholds is important in populated areas or in areas of historic or artistic interest. This is why we use the product of P and V, which yields the Population Dosage Product (herein referred to as PDP). This condition would imply that if a location has a high violation score, but zero human population density, then the only information gained by placing a station at the respective location would be associated with the type of land-use and the density of the monuments.

Each location reports a certain amount of information. Since the domain shows various degrees of dependence between locations, an advantage of such a dependency must be taken into consideration. Linear Regression Analysis (LRA) is the simplest, yet flexible tool to analyze dependencies or overlaps in such situations. In LRA, similarities among locations are assessed by calculating the linear correlation coefficient between locations. If the correlation coefficient $R$, between two locations is higher than a cut-off value $R_0$, then the locations are considered to be dependent. The optimization problem could therefore be considered as the identification of the optimum number and configuration of monitoring stations, in order to obtain the maximum gain of information (estimated through IF) in the area of interest at a minimum possible overlap between stations.

In the next step, a linear correlation matrix of $21 \times 21$ from the data of optical thickness is constructed for each grid-point in the region. The values in this matrix are the correlation coefficients between the optical thickness values of the central grid-point and the surrounding ones. For those values of $R$ that are higher than the cut-off value $R_0$, the gain of information ($GI$) for each grid-point is given by the equation:

$$GI_i = \sum_{k=1}^{21} R_{ik} \cdot IF_i, \text{ if } R_{ik} \geq R_0$$

$$IF = \frac{(V \cdot P) + L + M}{C}$$
where

\[ i = i^{th} \text{ location of a grid-point in the large grid} \]

\[ k = k^{th} \text{ location of a grid point in the small grid of } 21 \times 21 \]

In this way, the information reported by the dependent grid-points is passed to the central grid-point. By examining the matrix of GI values, the location with the maximum value of GI is selected as the first best monitoring location. The information of those grid-points that are represented by the first monitoring location as well as the information of the first best monitoring location are then set to zero for the calculations to follow. In this way minimum overlap between the first and the following monitoring location is achieved. The matrix of GI values is recalculated and a location is identified, which has now the highest value of GI. This location is then selected as the next best monitoring station. The values of information of the second monitoring station and of the grid-points represented by it are again set to zero and the matrix of GI values is calculated anew. This procedure of sequential selection is repeated till all the grid-points in the region are completely covered or the budgetary constraint is violated.
**Start**

- Superimpose grid of pixels of 1 km² over the area

**Calculate violations of SO₂ and NO₂ in terms of δopt**

**Thresholds for SO₂ and NO₂ (EU regulations)**

**Monuments density**

**Population density**

**Estimate statistical correlation of δopt with NO₂ and SO₂ (via aerosol model)**

**Information function**

\[
IF = \frac{(\text{Violation} \times \text{Population}) + \text{Use of land} + \text{Monuments}}{\text{Cost}}
\]

**Assess similarities among stations by Linear Regression Analysis**

**Estimate gain of information (GI) per pixel**

If information (IF) of surrounding (area 21 x 21 km) pixels with \( R > R^o \) is assigned to it

**Place monitoring station in the location with maximum GI**

Set to zero IF value of previous pixel and of those which represents

**Budgetary constraint violated**

NO

End

**YES**

MOON Information Flowchart: Determination procedure for allocation of monitoring stations
As developed, the algorithm can be used to identify the “best” sequence of grid-points in a systematic manner to ensure maximization of regional information and coverage at a minimum overlap. Therefore, the implementation of this approach can be twofold.

1. To provide the optimal number and configuration of a new monitoring network
2. To streamline the present monitoring network, with the minimum loss of information.

The following flowchart depicts the algorithm for correlating Earth observation-derived aerosol optical thickness with the ground measured concentrations of PM, SO$_2$ and NO$_x$.

Flowchart of the algorithm for computational assessment of pollution profiles via $\delta_{opt}$ (aerosol optical thickness)
3. RESULTS AND DISCUSSION

The present algorithm has been applied to find the best number and the configuration for monitoring PM, sulfur dioxide and nitrogen dioxide concentrations in the Kozani-Ptolemais-Florina axis of W. Macedonia, Greece. These cases are particularly interesting as multiple pollutants (particulate matter, sulfur dioxide and nitrogen dioxide) are involved, and heavy industry coexists with populated areas and agricultural land, resulting in a conflict between the objectives of the network.

Fig.1. Existing monitoring network in W. Macedonia
An extensive evaluation of the effectiveness of the suggested monitoring network was performed, thanks to five performance indicators established in the present work:

1. **Percentage Exposure** has been defined as a ratio of the PDP detected by the AQMN to that of total PDP calculated on summing over all the candidate locations.
2. **Percentage Compliance** denotes a ratio of the violation score detected by the AQMN to that of the total violation score calculated by summing over the violation scores at all candidate locations.
3. Percentage Information is the ratio of the information gained by the AQMN to the total information gain calculated on summing over all candidate locations.
4. **Percentage Coverage** measures the ratio of the number of candidate locations covered by the AQMN, in terms of dependency of those locations to the total number of candidate locations. The dependency among locations is estimated from linear regression analysis of the optical thickness values.
5. **Total Cost** of the AQMN.

These five indicators combine the objectives of both the Regulatory and the EUROAIRNET network regarding pollution in densely populated areas, detection of exceedances (or near exceedances), representative air quality information and coverage of large areas. SO$_2$ and NO$_2$ are selected as first priority indicators of the exposure of all three important receptors of air pollution (population, materials and ecosystems) to be included in EUROAIRNET.

It can be observed from Figure 2, that as the cut-off correlation coefficient $R^0$ is increased from 0.85 to 0.95, the number of monitors required for the total representation of the pattern also increases. For a stipulated budget, the air quality monitoring organization could maintain either a high or a low value $R^0$ based network. A high $R^0$ based network may not necessarily cover the entire region, but the covered region will be well represented. A low $R^0$ based network on the other hand, would offer more coverage of the region, but the covered region may not be satisfactorily represented. The ultimate decision in such a case is of course left to the air quality monitoring organization.

The marginal effectiveness of the monitoring network decreases as the monitors are sequentially selected. In other words, the costs of additional monitors to cover the entire region, say beyond 20% (20 stations) to 22% (30 stations), are rather significant as compared to the accrued benefit. In many
cases, therefore, it is possible that the regional competent authorities would be interested to maintain the AQMN only up to a gain of information of 22.2% for 20 stations.

Fig.2. Relationship between cut-off correlation coefficient, number of monitors and the effectiveness of the optimally located stations in W. Macedonia.
Figure 3 shows the configuration of the optimal locations for $R^0$ equal to 0.90 respectively for a maximum number of 20 stations. The number of stations is of course subject to the budgetary constraint. It should be noted that a network designed for a higher value of $R^0$ is not necessarily a simple extension, in terms of a mere addition of some more monitors, as compared to that of the network obtained at a lower $R^0$ value. It appears therefore, that the expansion of a network to a higher reliability of representation, may not only require an additional number of monitors but perhaps also a relocation of existing ones as well.

Fig.3. Configuration of optimally placed monitors at cut-off correlation coefficient for W. Macedonia $R^0 = 0.9$. In blue are shown the locations of the additional ground monitors that would be required to ensure the coverage, information acquisition, compliance check and exposure levels set as objectives in the optimization utility function up to a total number of twenty stations.

Overall, the optimization results tend to create a monitoring network that extends in both horizontal dimensions. This is a clear departure from the current design of the monitoring network, which is based on providing information along the Florina-Ptolemais-Kozani axis, limiting thus the ability to track
pollution transport in the perpendicular direction with regard to this axis. Yet, results from the satellite-based data show that high pollution levels can be observed in areas not covered adequately (if at all) from the spatial configuration of the currently existing network. The proposed solution by the multi-objective optimization algorithm corrects this and more than doubles the amount of the obtained information gain.

**Network Reliability and Operability**

The reliability of the network in terms of being able to provide adequate air quality monitoring in the case of one station failure was tested. The four performance indicators of monitoring network performance have been reevaluated, assuming that one station fails to work properly and produce data. Results showed that if one out of the 20 stations foreseen fails at a time, the percentage of the performance indexes remains practically unchanged, since the combination of the remaining stations is sufficient to cover the loss of information due to one’s station’s failure. The reliability of the optimal monitoring network for the area under study has been tested for the failure of the 16 most important stations out of the 20 suggested ones. As a general remark, though, it should be mentioned that the precise locations of monitors involve further questions of site security, power availability, and housing of the sampler itself. These questions are beyond the scope of this work.

**4. CONCLUSIONS**

The present approach presented for optimal configuration monitoring networks and streamlining existing networks, is a significant addition to the tools available to those responsible for the design and operation of air quality monitoring networks. It provides monitor site assignments for PM, SO₂ and NO₂, designed to maximize monitor coverage of a region as well as maximum gain of information, in terms of exposure of population, land-use and vulnerable cultural receptors at a minimum cost. At the same time, the optimal configuration of AQMN suggested herein follows the EU Directive on AQ
monitoring and furthermore can provide EUROAIRNET with the representative monitoring stations for W. Macedonia. It gives priority to areas with high emission levels without neglecting weaker sources. The method is easy to use and inexpensive in computer resources. It provides an integrated as well as unique method for designing monitoring networks in Europe, based on the EU Directive’s most important pollutants, in terms of human health and vegetation protection.

REFERENCES


Air Pollution Exposure in European Cities: the EXPOLIS study, Matti Jantunen, Klea Katsouyanni, Helmut Knoeppel, Nino Kuenzli, Erik Lebret, Marco Maroni, Kristina Saarela, Radim Sram, Denis Zmirou, Final report.


European Cultural Heritage Newsletter on Research, Special Issue, June 1997, Vol.10


