United States
Environmental Protection
Agency

Environmental Monitoring Systems
Laboratory
Las Vegas NV 89114

Research and Development

EPA-600/S4-81-002 Apr. 1981

Project Summary

Methodology for the Design of an Optimum Air Quality Monitoring Network

Mei-Kao Liu and Joel Avrin

A two-step objective method is presented for determining the optimum number and disposition of ambient air quality stations in a monitoring network. The method uses a data base consisting of a comprehensive set of simulated or measured air quality patterns representative of the region of interest. In the first step, the most desirable monitoring locations are identified and ranked. The minimum number of required locations is determined in the second step through eliminating redundancies among the locations identified in the first step with regard to spatial coverage over the region of interest. As a demonstration, the method is applied to the Las Vegas Valley of Nevada for the pollutant species carbon monoxide.

This Project Summary was developed by EPA's Environmental Monitoring Systems Laboratory, Las Vegas, NV, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Introduction

The Clean Air Act requires state and local agencies to monitor ambient air quality, primarily for documenting an area's compliance with the National Ambient Air Quality Standards. Additional monitoring may be needed to satisfy secondary objectives such as providing background or baseline concentrations. Currently, the determination of the number and location of monitoring stations required in a network is primarily based on subjective considerations, semiquantitative rules supported by experience, or limited use of analytical tools such as simple Gaussian models. Nontechnical considerations, such as convenience and accessibility, are usually the dominant factors in selecting a specific monitoring location within the area of interest. On the other hand, because of the fluctuations in pollutant emission rates and the turbulent nature of the atmosphere, pollutant concentration distributions are highly variable, both in time and space. The concentrations measured at any given site depend on the emission patterns as well as atmospheric conditions. The design of an optimal monitoring network, therefore, requires an a priori knowledge of these concentration variabilities.

Method

A two-step objective methodology for the design of such an air quality monitoring network is proposed in this study. The methodology uses a data base consisting of a comprehensive set of measured or simulated air quality patterns representative of the region of interest. The objective of the first step is simply to ascertain the most favorable locations for making air quality measurements. To this goal, a concept called the figure-of-merit (FOM), introduced in an earlier study, is used to facilitate the identification and ranking of potential monitoring sites.[1] The objective of the second step is to determine the network configuration by deleting redundancies among the monitoring stations identified in the first step. The spatial correlation coefficient between pollutant concentrations at the monitoring site and those at a neighboring point is invoked as the parameter for estimating the relevance that air quality measured at one point has on another point. By using the sphere of influence, as determined by a minimum or cutoff value in the correlation coefficient distributions, the monitoring network, containing a minimum number of monitoring stations, can be obtained.

The figure-of-merit is defined as the sum over a comprehensive set of conditions of the products of an air quality index, either observed or predicted, and the associated probability of occurrence. Herein, the summation is performed over all meteorological scenarios potentially leading to high air pollutant concentrations. The FOM contains weighting by the probabilities of occurrence of scenarios to avoid situations related to extremely rare events or periods. These situations would not necessarily provide the best criteria for determining a permanent or semipermanent site for a monitoring network. In general, the air quality index can be composite of several pollutant concentrations, weighted by the relative importance of the individual species, if it is desirable to design a multiple pollutant species monitoring network. For simplicity, the air quality index used here was the concentration of a single pollutant species. Thus, for any given location (x, y),

$$\mathrm{FOM}\,(x, y) = \sum_{k=1}^{M} \begin{pmatrix}\text{concentration at location} \\ \text{(x, y) under meteorological} \\ \text{pattern k}\end{pmatrix} \cdot \begin{pmatrix}\text{probability of} \\ \text{meteorological} \\ \text{pattern k}\end{pmatrix} . \quad (1)$$

Although the concentration fields for Equation (1) can be either observed or predicted, the practice of using predicted concentration fields was adopted for this study; few, if any, regions have a monitoring network in operation over a sufficient time interval and of sufficient density to yield the requisite concentration distributions.

As is the case herein, the concentration distributions can be developed using an air quality simulation model for meteorological scenarios specified for the region of interest. These distributions, when combined with the corresponding probabilities of occurrence, permit the evaluation of FOM. A selection of the most favorable air quality monitoring sites can then be accomplished by ascertaining the noncontiguous peaks in the resultant FOM field and ranking the locations according to magnitude of FOM.

The approach for establishing the minimum number of monitoring stations from those identified and ranked in the first step involves the concept of the sphere of influence (SOI) which for a given monitoring station is determined by the statistical properties of the pollutant concentration distributions used in the first step. Analogous to the study of turbulence in an Eulerian framework, a spatial correlation coefficient is introduced between the values of pollutant concentrations at a potential monitoring site and the corresponding values at its neighboring points as a function of radial distance away from the station:

$$r(s_o, s_o + \Delta s) = \frac{\mathrm{cov}\,[c(s_o), c(s_o + \Delta s)]}{\sqrt{\mathrm{var}\,[c(s_o)] \cdot \mathrm{var}\,[c(s_o + \Delta s)]}} , \quad (2)$$

where the $c(s_o)$ and $c(s_o + \Delta s)$ can be measured or predicted concentrations at the points s_o and $s_o + \Delta s$ and $\Delta s = [(\Delta x)^2 + (\Delta y)^2]^{1/2}$. The symbols cov and var denote covariance and variance, respectively.

Statistically, the correlation coefficient measures the linear association between $c(s_o)$ and $c(s_o + \Delta s)$, namely, the concentrations measured at the monitoring sites and those at its neighboring points. This coefficient, lying between -1 and +1, thus by itself furnishes an ideal dimensionless tool for determining the SOI. The spatial correlation coefficient is expected to initially decrease from one as the distance increases. Consequently, the SOI can be delineated on the basis of a predetermined minimum or "cutoff" value for the correlation coefficient.

The square of the correlation coefficient represents the fraction of the variance of one variable explained by the other variable.[2] One hundred times this value then can be considered as the percentage of concentration variations explained by concentration variations at a potential monitoring site. In practice, values of pollutant concentrations are generally taken from a finite sample of a larger population. Thus, after a value for the minimum desired amount of variance explained is selected, the minimum acceptable value of the population correlation coefficient can be computed using the above-mentioned relationship. Procedures and tables developed by F. N. David can in turn be used with the minimum acceptable value for this coefficient to yield a cutoff value for the sample correlation coefficient for a specific sample size and confidence level.[3]

Once the cutoff value for the sample correlation coefficient is determined, a SOI can be subsequently developed for each of the ranked potential monitoring sites. The determination of the minimum number of monitoring stations required can then be carried out by deleting lower-ranked stations whose SOI overlap the SOI of higher-ranked stations and whose SOI provide nonoverlapping coverage of less than some fixed percentage of the coverage of the higher-ranked stations.

Application

The siting methodology was applied as a demonstration to the Las Vegas Valley of Nevada (see Figure 1) for the pollutant species carbon monoxide. The simulated concentration fields were obtained by exercising the SAI Atmospheric Pollution Simulation Model for six different meteorological scenarios for Las Vegas.[1] Nineteen stations covering concentration maxima and three stations covering background concentrations in rural areas were considered. For example, a 10-station network, among which 3 stations are for background concentrations, was selected for a desired minimum detection capability of 50 percent of concentration fluctuations (95 percent of the time) and hence a cutoff sample correlation coefficient of 0.8. The network configuration and joint areal coverage (shaded areas) are shown in Figure 2. The fraction of total area covered by each station and the cumulative areal coverage, beginning with the highest-ranked station, are also tabulated in this figure. Stations identified by T, U, or V are the background stations.

Networks with fewer stations than in the above network would be selected if smaller minimum detection capabilities of concentration fluctuations are deemed acceptable, and vice versa. Background stations could, of course, be deleted for networks with the sole objective of discerning peak concentrations.

References

1. McElroy, J. L. et al. (1978), "Carbon Monoxide Monitoring Network Design Methodology," EPA-600/4-78-053, Environmental Monitoring Systems Laboratory, U.S. Environmental Protection Agency, Las Vegas, Nevada.
2. Ezekiel, M. (1941), Methods of Correlation Analysis, John Wiley and Sons, Inc., London, England.
3. David, F. N. (1938), *Tables of the Ordinates and Probability Integral of the Distribution of the Correlation Coefficient in Small Samples*, The Biometrika Office, Cambridge University Press, Cambridge, England.

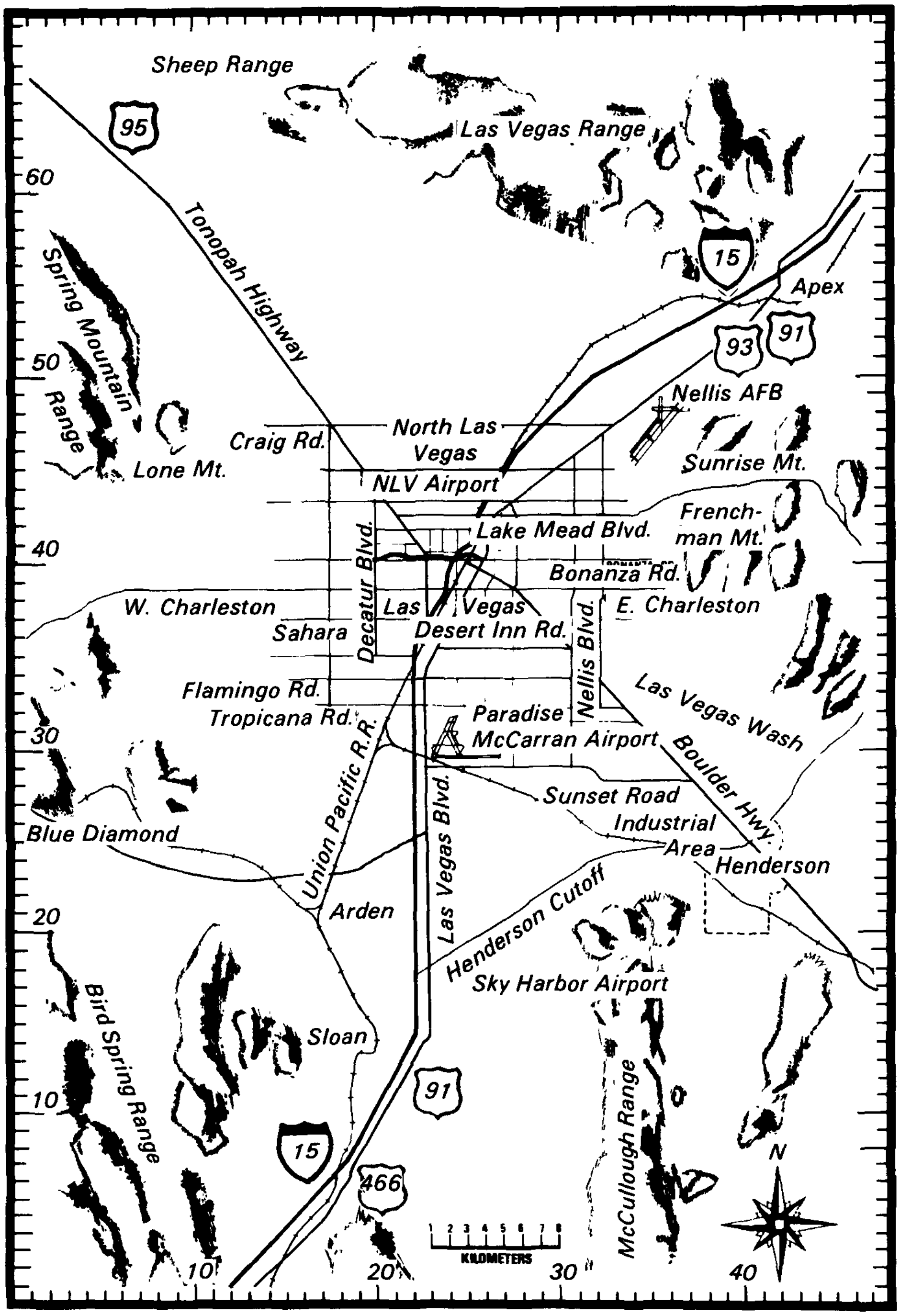
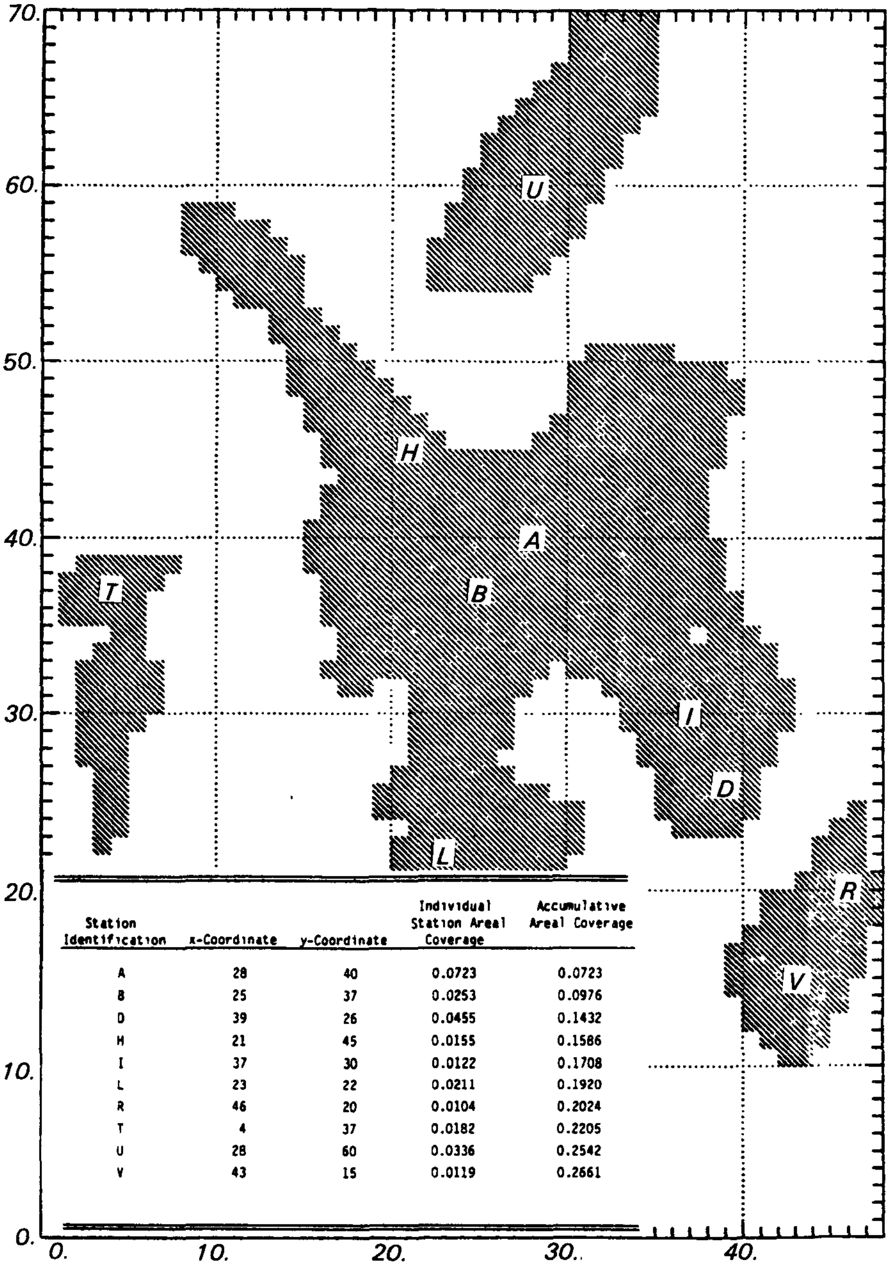


Figure 1. *Map of the Las Vegas Valley.*



Station Identification	x-Coordinate	y-Coordinate	Individual Station Areal Coverage	Accumulative Areal Coverage
A	28	40	0.0723	0.0723
B	25	37	0.0253	0.0976
D	39	26	0.0455	0.1432
H	21	45	0.0155	0.1586
I	37	30	0.0122	0.1708
L	23	22	0.0211	0.1920
R	46	20	0.0104	0.2024
T	4	37	0.0182	0.2205
U	28	60	0.0336	0.2542
V	43	15	0.0119	0.2661

Figure 2. *An optimum air quality monitoring network based on a cut-off sample correlation coefficient of 0.8 — Las Vegas Valley.*

Mei-Kao Liu and Joel Avrin are with Systems Applications, Inc., San Rafael, CA 94903.
***James L. McElroy** is the EPA Project Officer (see below).*

The complete report, entitled "Methodology for the Design of an Optimum Air Quality Monitoring Network," (Order No. PB 81-171 191; Cost: $8.00, subject to change) will be available only from:
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: 703-487-4650

The EPA Project Officer can be contacted at:
Environmental Monitoring Systems Laboratory
U.S. Environmental Protection Agency
Las Vegas, NV 89114

☆ U.S GOVERNMENT PRINTING OFFICE. 1981 -757-012/7051

United States
Environmental Protection
Agency

Center for Environmental Research
Information
Cincinnati OH 45268

PS 0000329
U S ENVIR PROTECTION AGENCY
REGION 5 LIBRARY
230 S DEARBORN STREET
CHICAGO IL 60604