EPA-600/4-81-002
February 1981

METHODOLOGY FOR DESIGNING AN OPTIMUM AIR QUALITY MONITORING NETWORK

by

Mei-Kao Liu and Joel Avrin
Systems Applications, Incorporated
950 Northgate Drive
San Rafael, California 94903

James L. McElroy, Project Officer
U.S. Environmental Protection Agency
Environmental Monitoring Systems Laboratory
Las Vegas, Nevada 89114

Contract No. 68-03-2446

ENVIRONMENTAL MONITORING SYSTEMS LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
LAS VEGAS, NEVADA 89114

EPA-600/4-81-002
February 1981

METHODOLOGY FOR DESIGNING AN OPTIMUM AIR QUALITY MONITORING NETWORK

by

Mei-Kao Liu and Joel Avrin
Systems Applications, Incorporated
950 Northgate Drive
San Rafael, California 94903

James L. McElroy, Project Officer
U.S. Environmental Protection Agency
Environmental Monitoring Systems Laboratory
Las Vegas, Nevada 89114

Contract No. 68-03-2446

ENVIRONMENTAL MONITORING SYSTEMS LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
LAS VEGAS, NEVADA 89114

DISCLAIMER

CONTENTS

FIGURES

TABLES

ACKNOWLEDGMENT

We would like to express our sincere thanks to our colleagues, Dr. Martin J. Hillyer, Dr. Philip M. Roth, and Dr. C. Shepherd Burton for many enlightening discussions. From these brainstorming sessions, the approach presented in this report began to take shape. We would also like to thank Dr. George Flatman of the Environmental Monitoring Systems Laboratory of the U.S. Environmental Protection Agency for providing critical and timely advice that helped finalize the statistical aspect of this study.

SECTION 1

SUMMARY

An objective methodology is presented for determining the optimum number and disposition of ambient air quality stations in a monitoring network. The proposed methodology uses climatological information and an air quality simulation model. First, the climatological information is used to generate a limited number of meteorological scenarios representative of the region of interest. For each of the scenarios, the air quality simulation model is employed to produce the corresponding temporal-varying air quality patterns. The air quality patterns serve as the primary data base in a two-step procedure for determining the monitoring network. In the first step, the air quality patterns are collapsed into a single pattern through the use of the figure-of-merit (FOM) concept. For a specific time interval and location, the FOM is determined as the sum over the meteorological scenarios of the products of the pollutant concentrations and the associated probabilities of occurrence. The identification and ranking of the most desirable monitoring locations are achieved using the resultant FOM fields. In the second step, the network configuration is determined on the basis of the concept of a sphere of influence (SOI). The SOI are dictated by a cutoff value in the spatial correlation coefficients between the predicted pollutant concentrations at the monitoring locations identified and the corresponding concentrations at neighboring locations in the region. This cutoff value is related to an estimate of concentration variations that can be accounted for by a given monitoring station. The minimum number of monitoring stations required is then determined by deleting lower-ranked stations whose SOI overlap the SOI of higher-ranked stations and whose SOI provide non-overlapping coverage of less than some fixed percentage of the coverage of the SOI of the higher-ranked stations.

As a demonstration, the siting methodology was applied to the metropolitan Las Vegas area for carbon monoxide monitoring stations. A 10-station network, consisting of 7 stations for average peak concentrations and 3 stations for background concentrations, was selected for a desired minimum detection capability of 50 percent of the concentration fluctuations (95 percent of the time) and hence a cutoff spatial correlation coefficient of 0.8 networks with fewer stations would be selected if smaller minimum detection capabilities of concentration fluctuations are deemed acceptable, and vice versa.

SECTION 2

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 (NAAQS). Additional monitoring may be required 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 sometimes, limited use of analytical tools such as simple Gaussian models (Ludwig and Kealoha 1975). 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 the atmospheric conditions. The design of an optimal monitoring network, therefore, requires an a priori knowledge of these concentration variabilities. An objective methodology for designing such a monitoring network is proposed and demonstrated in this study.

The methodology uses a data base consisting of a comprehensive set of pollutant concentration distributions representative of the region of interest. The practice of using simulated concentration distributions generated using an air quality model 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. The air quality model, of course, produces the distributions by linking the source emissions with the prevailing meteorological conditions.

The actual siting methodology consists of two steps. The goal of the first step is simply to ascertain the most favorable locations for making air quality measurements. To the goal, a concept called the figure-of-merit, introduced in an earlier study (McElroy et al., 1978), is used to facilitate the identification and ranking of potential monitoring sites. Procedures that utilize this concept for such a purpose and that develop the data base for use with the entire methodology are delineated in section 3.

The goal of the second step in the siting methodology is to determine the minimum number of monitoring stations and, hence, the optimum network configuration. The spatial correlation coefficient, which is commonly used in statistics and turbulence research, is employed as the parameter with which to measure the relevance of air quality at one point to that of another point in its neighborhood. By imposing a minimum value for this coefficient, a sphere of influence for each potential monitoring site can be defined. Subsequently, by deleting redundancies among the monitoring stations identified and ranked in the first step, the optimum monitoring network containing a minimum number of monitoring stations thus can be determined. The theoretical framework underlying this approach is described in section 4.

As a demonstration of its utility, the entire siting methodology was applied to the metropolitan Las Vegas area. A relatively inert pollutant species, carbon monoxide, was used as an example. The application of the siting methodology to Las Vegas is discussed in section 5.

SECTION 3

IDENTIFICATION AND RANKING OF POTENTIAL MONITORING SITES

This section discusses the first step of the objective methodology for the design of an optimum air quality monitoring network -- the procedures for the identification and ranking of potential monitoring locations. As discussed in prior studies, the desirability of placing an air quality monitor at a given location is closely related to specific monitoring objectives (Ludwig and Kealoha 1975, Ott 1975, Liu et al. 1977, and Ludwig, Berg, and Hoffman 1976). In general, the primary objective of an air quality monitoring network is to monitor the highest concentrations in the area of interest to ensure compliance with air quality standards. These monitoring sites are labeled by Ott (1975) as the "A"-type stations and by Ludwig, Berg, and Hoffman (1976) as the "street canyon" and "traffic corridor" stations. In addition to this primary objective, other secondary objectives for air quality monitoring also exist. For example, as discussed by Ott (1975) and Ludwig, Berg, and Hoffman (1976), additional stations may be needed either to measure the population exposure in a residential area or to provide the background as baseline concentrations typical of the outlying rural areas. The former, called the "C"-type stations by Ott and the "neighborhood" stations by Ludwig, Berg, and Hoffman, would require additional information, such as demographic data. The latter, called the "E"-type stations by Ott and "regional" stations by Ludwig, Berg, and Hoffman can, however, be incorporated into the present siting algorithm, which is designed primarily for locating the pollutant concentration maxima.

In an earlier study by McElroy et al. (1978), the desirability of placing an air quality monitor at a given location in an urban area was accomplished using a concept called the figure-of-merit (FOM). In its most general form, the FOM can be defined as the sum over an exhaustive or comprehensive set of conditions of the products of an air quality index either observed or predicted and the associated probability of occurrence:

$$\text{FOM} = \sum \text{(Air Quality Index)} \cdot \text{(Probability of Occurrence)} \; . \qquad (1)$$

The summation is to be 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.

The air quality index in Equation (1) can be a composite of several pollutant concentrations, weighted again by the relative importance of the individual species, if it is desirable to design a multiple pollutant-species monitoring network. For example, to locate a site for measuring multiple pollutant species, the air quality index in Equation (1) can be generalized using a composite concentration index proposed by Ott and Thom (1976):

$$I = \sum_{\ell=1}^{N} w_\ell c_\ell \quad , \tag{2}$$

where I denotes the overall air quality index, c_ℓ is the concentration of species ℓ, and w_ℓ is the corresponding weighting factor reflecting the importance of pollutant species ℓ in the assessment of the overall air quality. In general, c_ℓ can be either an observed or predicted concentration. For the sake of simplicity, only one pollutant species is considered in the present study. In this case, the FOM at any location can be defined as the sum of the products of concentration for a specific pollutant and the associated probability of occurrence of the corresponding meteorological conditions which are in turn based on available local climatological data:

$$\text{FOM}(x,y) = \sum_{k=1}^{M} \begin{pmatrix} \text{Concentration at location} \\ \text{(x,y) under meteorological} \\ \text{pattern k} \end{pmatrix} \cdot \begin{pmatrix} \text{Probability of} \\ \text{meteorological} \\ \text{pattern k} \end{pmatrix} . \tag{3}$$

As an alternative, with special emphasis on the detection of maximum concentrations exceeding the NAAQS, the FOM can be defined as a step function of the pollutant concentration in a similar manner:

$$\text{FOM}(x,y) = \sum_{k=1}^{M} \begin{bmatrix} \text{1, if NAAQS or some} \\ \text{fraction thereof is} \\ \text{exceeded at location} \\ \text{(x,y) under meteo-} \\ \text{rological pattern k;} \\ \\ \text{0, if not.} \end{bmatrix} \cdot \begin{pmatrix} \text{Probability} \\ \text{of meteorological} \\ \text{pattern K} \end{pmatrix} . \tag{4}$$

In Equations (3) and (4), the pollutant concentration can be either observed or expected. In the present study, the concentration fields are generated from an air quality simulation model, which plays a central role in the siting methodology by linking the known emissions distribution with air quality patterns for a given meteorological scenario. The simulated temporal-varying air quality patterns, when combined with the corresponding statistics, permit the determination of the FOM as per Equations (3) and (4).

A selection of the most favorable air quality monitoring sites can then be accomplished by ascertaining the noncontiguous peaks in the FOM field and ranking locations according to magnitude of FOM. This process, which can easily be carried out numerically, essentially completes the first step of the present siting methodology.

After the identification of the maxima in the FOM field and their ranking according to the corresponding FOM values, two issues remain to be resolved before the optimum monitoring network can be developed: The first issue is related to the representativeness of air quality data for a selected monitoring station, and the establishment of an area surrounding this station for which the data can be extrapolated. The second issue is concerned with the minimum number of measurement stations needed to obtain sufficient monitoring coverage, as determined by the capability of detection of concentration fluctuations by a given monitoring network. These two issues, apparently interrelated, are addressed in the next section.

SECTION 4

DETERMINATION OF SPHERES OF INFLUENCE AND THE OPTIMUM MONITORING NETWORK

The determination of the minimum number of monitoring stations required appears to be the most crucial element in developing an optimum air quality monitoring network. Intimately related to this element is the determination of the spatial coverage, or the sphere of influence (SOI), for each of the monitoring locations. In this context, the SOI is defined as the surrounding area over which the air quality data for a station can be considered to be representative.

Obviously, the specification of an SOI for any selected site is not unique. Its establishment depends on the method of reconstructing, either through interpolation, or extrapolation, the concentration field from the data obtained for a given site. It is conceivable that different interpolations or weighting methods can yield different SOI if the interpolation error is to be kept to a minimum. For example, a linear interpolation might yield a SOI different from an inverse distance interpolation. In the former case, gradients are assumed to be constant and can be positive or negative. In the latter case, the gradients are not spatially constant and are generally negative; that is, as one progresses outward, the extrapolated values decrease monotonically. In an earlier study, a heuristic approach based on the geometry of the computed FOM field was adopted (Liu and Moore 1980). For comparison purposes, the results of this alternative approach are summarized in Appendix A.

A more rigorous approach is adopted in the present study. This approach is based on the statistical properties of the spatial distributions of the pollutant concentration distributions used in the first step of the siting methodology. Analogous to the study of turbulence in a Eulerian framework, a spatial correlation coefficient is introduced between values of pollutant concentration at a given site and the corresponding values at its neighboring points as a function of radial distance away from the station:

$$r(s_0, s_0 + \Delta s) = \frac{\text{cov}\,[C(s_0),\, C(s_0 + \Delta s)]}{\sqrt{\text{var}\,[C(s_0)]\cdot\text{var}\,[C(s_0 + \Delta s)]}}\,, \qquad (5)$$

where $C(s_0)$ and $C(s_0 + \Delta s)$ can be measured or predicted values at the points s_0 and $s_0 + \Delta s$ 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 provides a measure of the intensity of association between $C(s_o)$ and $C(s_o + \Delta s)$, namely, the concentrations 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 the determination of the SOI. Similar to the characteristics of the correlation coefficient commonly used in the study of turbulent velocity, temperature, and concentration fluctuations, the spatial correlation coefficient is expected to initially decrease from one as the distance increases. Consequently, a cutoff distance of s_c can be found to determine the SOI for a predetermined minimum spatial correlation coefficient as is illustrated in Figure 1.

Thus, in the second step of this siting methodology, the spatial correlation coefficients surrounding each of the potential monitoring sites are evaluated. The computation can be carried out along all radial directions until the spatial correlation coefficient falls below a predetermined minimum or cutoff value. Consequently, the SOI for each of the stations identified in the first step of the current siting methodology can be determined.

The choice of the cutoff value for the spatial correlation coefficient can be determined statistically for a given monitoring site. Assume that $C_1 = (C_{11}, C_{12}, \ldots C_{1n})$ and $C_2 = (C_{21}, C_{22}, \ldots C_{2n})$ denote the pollutant concentrations at the monitoring site and the corresponding pollutant concentrations at a neighboring point, respectively. A computational form of equation (5) for C_1 and C_2 with a sample size, n, is given by

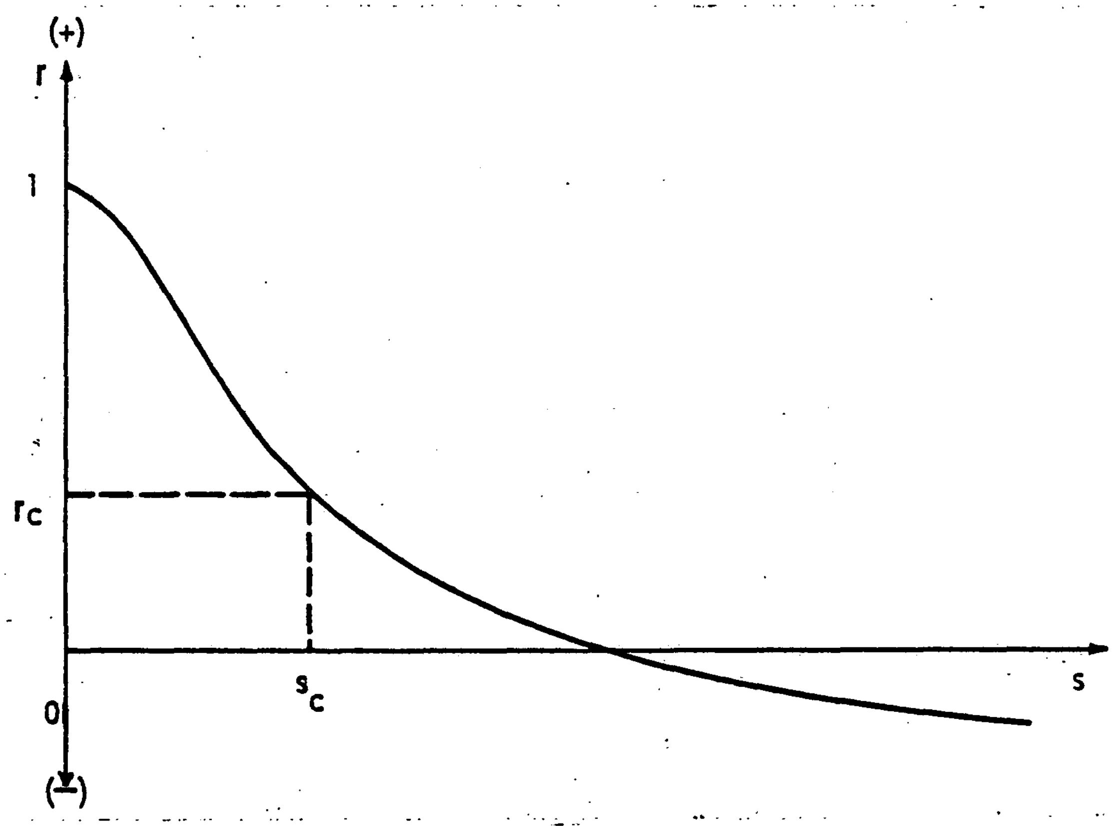


Figure 1. Generalized correlation coefficient as a function of distance.

$$r = \frac{\sum_{i=1}^{n} (C_{1i} - \overline{C_1})(C_{2i} - \overline{C_2})}{\sqrt{\sum_{i=1}^{n} (C_{1i} - \overline{C_1})^2 \sum_{i=1}^{n} (C_{2i} - \overline{C_2})^2}} , \tag{6}$$

where

$$\overline{C_1} = \frac{1}{n} \sum_{i=1}^{n} C_{1i}$$

and

$$\overline{C_2} = \frac{1}{n} \sum_{i=1}^{n} C_{2i}$$

Assuming that these two correlated random variables, C_1 and C_2, are from a bivariate normal distribution, a general expression can be derived for the probability distribution of a correlation coefficient, r, associated with a sample size, n, randomly drawn from an infinite population with a true correlation coefficient ρ. This probability distribution is given by David (1938).

$$p(r|n,\rho) = \frac{(1 - \rho^2)^{\frac{n-1}{2}}}{\pi(n - 3)!} (1 - r^2)^{\frac{n-4}{2}} \frac{d^{n-2}}{d(\rho r)^{n-2}} \left[\frac{\text{arc cos } (-\rho r)}{\sqrt{1 - (\rho r)^2}}\right] \tag{7}$$

The probability integral, E, given by

$$E = \int_{r_1}^{r_2} p(r|n,\rho_0) \, dr , \tag{8}$$

then represents the confidence level of the test hypothesis that $r = \rho_0$ with $r_1 < \rho < r_2$ as the confidence interval.

The probability integral, Equation (8), can be evaluated using a quadrature method (David 1938). In Figure 2, the confidence interval for correlation coefficients at a 95 percent confidence level are reproduced from Tables of the Ordinates and Probability Integral of the Distribution of the Correlation Coefficient in Small Samples (David 1938).

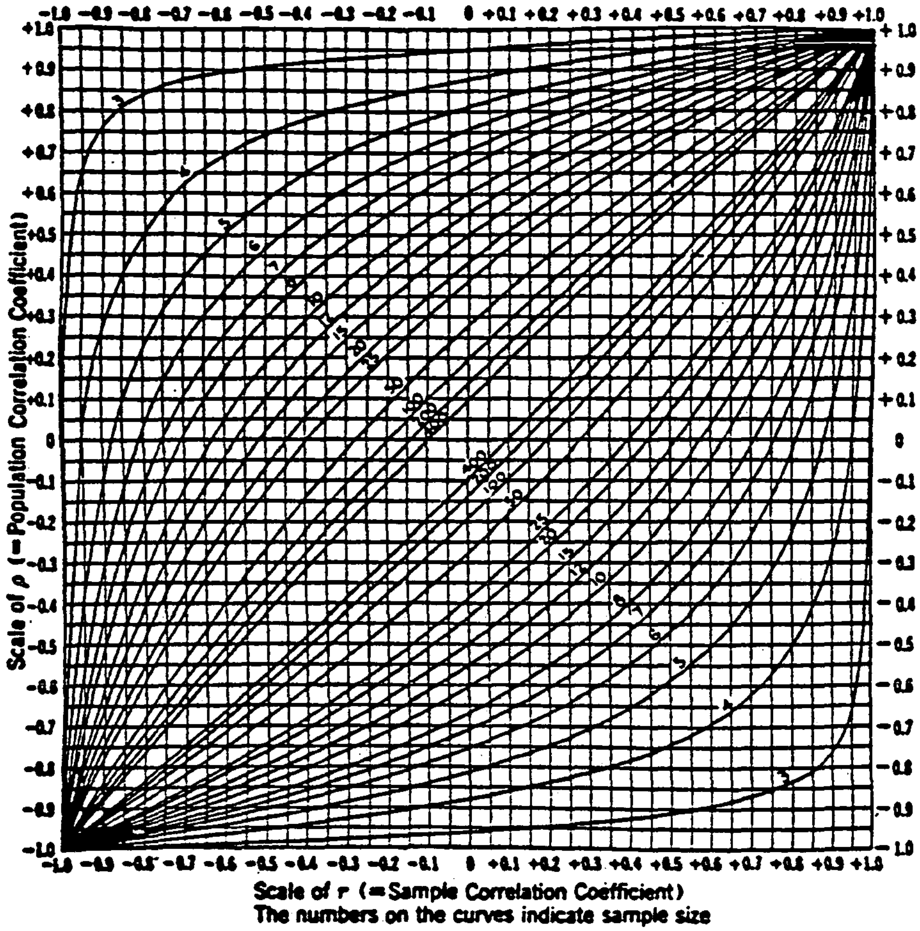


Figure 2. Confidence interval (belts) for the correlation coefficient at 95 percent confidence level. (This figure is reproduced with permission of E. S. Pearson, from David (1938), Cambridge University Press for the Biometrika Trustees).

If a linear relationship is assumed between the variables, then the square of the correlation coefficient represents the fraction of the variance of C_2 which can be accounted for by the variations in C_1 (Ezekiel 1941). Note that the use of r^2 or ρ^2 in this way does not imply a causal relationship between C_1 and C_2 but that a linear association exists between them such that $100\,\rho^2$ is the percentage of concentration variations explained by concentration variations at a potential monitoring site. Thus,

$$\text{variance explained} = \rho_c^2 \qquad (9)$$

Once the minimum desired value of variance explained is selected, the minimum acceptable value of ρ_c can be obtained from equation (9). This value for ρc can be entered into a nomograph such as Figure 2 as a lower bound to obtain a cutoff value for the sample coefficient, r_c, for a specific sample size and hence establish the SOI for each of the ranked stations:

$$A_1 = A(x_1, y_1)$$
$$A_2 = A(x_2, y_2)$$
$$A_3 = A(x_3, y_3)$$
$$\ldots ,$$
$$A_N = A(x_N, y_N) \quad ,$$

where A_i, $1 \leq i \leq N$, is defined by all points (x, y) that lie in the simply connected region containing station (x_i, y_i) enclosed by a contour determined by $r > r_c$. The total areal coverage by the monitoring network for all N stations, as illustrated in Figure 3, is given by

$$A_{Network} = A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_N \,. \tag{10}$$

The determination of the minimum number of monitoring stations required can be then carried out by deleting lower ranking stations whose SOI overlap the SOI of the higher ranking stations and whose SOI provide non-overlapping coverage of less than some fixed percentage of the coverage of the higher ranking stations.

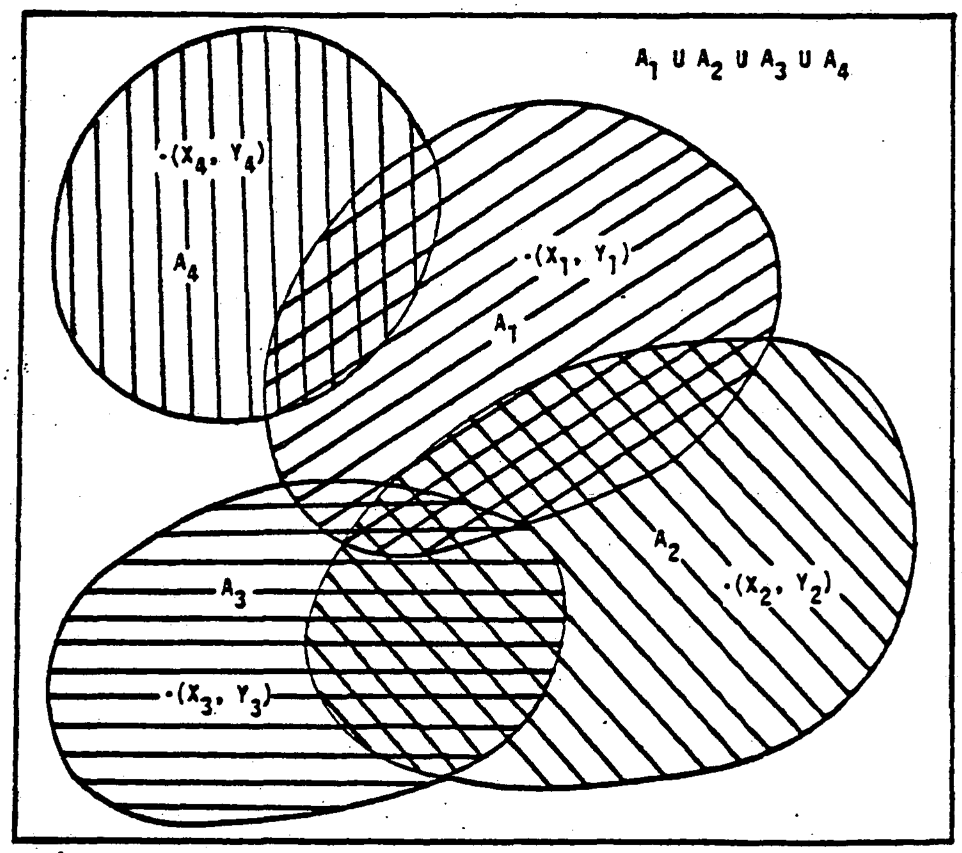


Figure 3. Joint areal coverage for monitoring stations.

SECTION 5

APPLICATION OF THE SITING METHODOLOGY TO THE LAS VEGAS VALLEY

The methodology for designing an optimum network described in the previous two sections was applied to the metropolitan Las Vegas area. Although the procedures and the tools developed for this siting methodology are applicable to chemically inert as well as reactive pollutants, only a relatively inert species -- carbon monoxide -- was considered in the present study. Las Vegas is a relatively isolated, urban community surrounded by desert, with a population of over 300,000. The Las Vegas Valley, located in southern Nevada, is bounded by the Sheep Range and Las Vegas Range to the north, Spring Mountain to the west, Frenchman and Sunrise Mountains to the east, and the McCullough Range to the south (Figure 4). The floor of the valley slopes gently from west to east, from about 980 meters (m) mean sea level (MSL) in the west to about 550 m MSL in the east-southeast. East of the Las Vegas Wash, the terrain gently rises again. The configuration of these surrounding mountains, whose elevations average about a kilometer (km) above the valley floor, imparts a bowl shape to the valley and provides relief passes to the northwest, southwest, and southeast.

A preliminary task in the demonstration consisted of an assessment of the air quality model employed, the SAI Atmospheric Pollution Simulation Model, to reproduce pollutant concentration distributions under a variety of meteorological conditions. For this purpose, the Las Vegas Valley was divided into 1 km x 1 km grids over a 48 km x 70 km modeling region. With aerometric data gathered in a field measurement program, the SAI model was exercised for six days during the winter of 1975-1976. The predicted CO concentrations were compared with field measurements to assess the validity of the model. Good agreement was shown between predicted and measured values for nearly all the cases examined. The model predicted diurnal trends well, but sometimes failed to predict the absolute magnitudes of peak concentrations, especially in the downtown and Las Vegas Wash areas. The highest value concentrations resulted from the afternoon traffic peak (downtown). High concentrations also occurred at the lowest topographic point in the valley (Las Vegas Wash). The occurrence of microscale phenomena (i.e., those of a smaller scale than the model can resolve) or uncertainties in input data usually accounted for discrepancies. Linear correlation coefficients between hourly values of measured and predicted CO concentrations generally ranged between 0.7 and 0.9. These comparisons are as good as any that have been reported over this time interval for models finely tuned to a specific region. A detailed presentation of the field program and the model validation is found in McElroy et al. (1978).

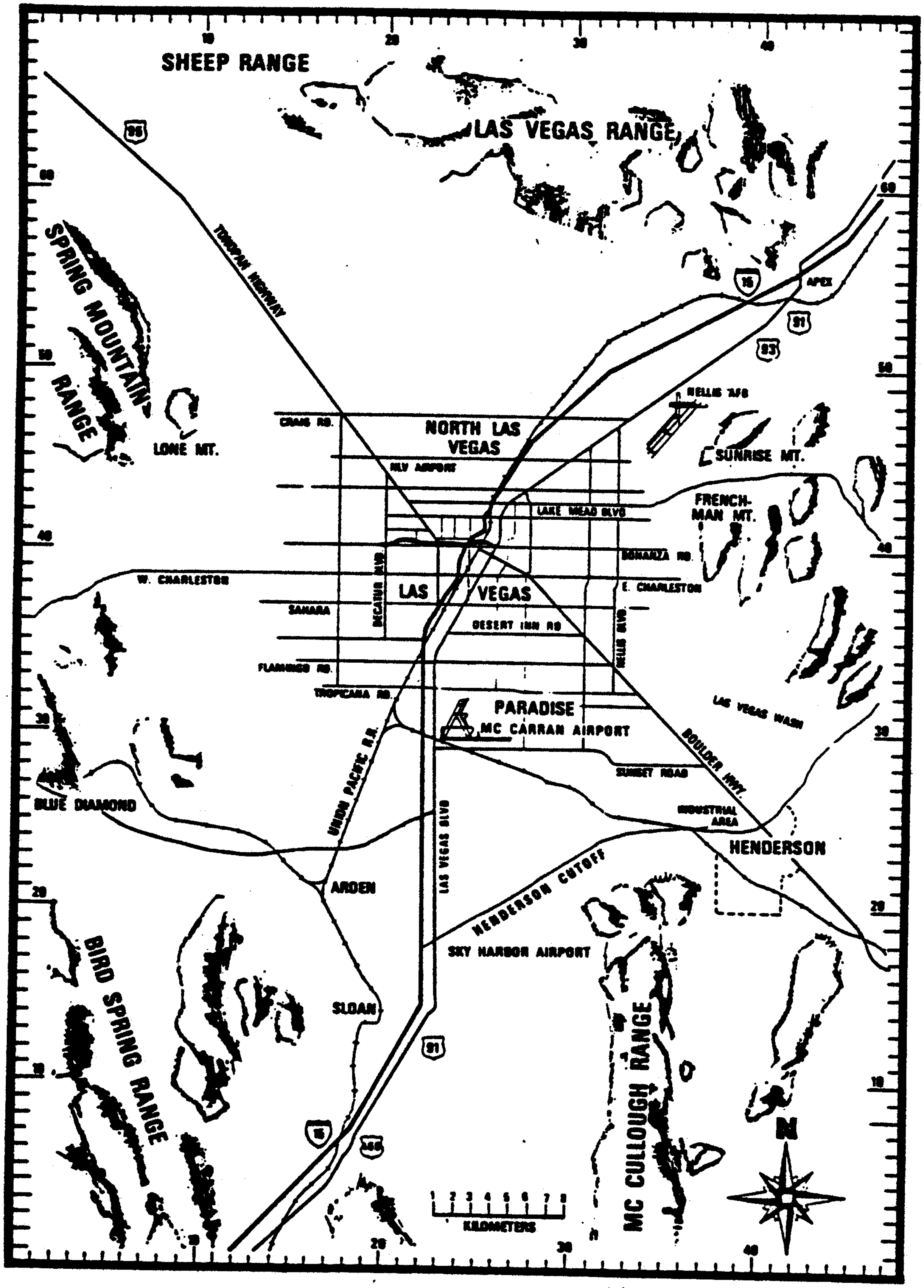


Figure 4. Map of the Las Vegas Valley.

With a validated air quality simulation model in hand, the stage is set for the design of the monitoring network. As indicated earlier, the first step is the identification and ranking of potential monitoring sites and the second step is the determination of the optimum configuration for the monitoring network. These two steps are addressed in the following subsections.

THE FIRST STEP--IDENTIFYING AND RANKING MONITORING SITES

For this purpose, six meteorological scenarios were selected on the basis of historical weather data for the Las Vegas area (McElroy et al. 1978). The SAI Air Pollution Simulation Model was exercised for each of the six meteorological scenarios. In this case, carbon monoxide concentration fields spanning 13 hours, at hourly intervals between 7:00 a.m. and 8:00 p.m., LST were determined. The corresponding FOM fields were subsequently computed using the following expression:

$$\mathrm{FOM}(x,y) = \sum_{k=1}^{6} \begin{pmatrix} \text{Concentration at location} \\ (x,y) \text{ under meteorological} \\ \text{pattern } k \end{pmatrix} \cdot \begin{pmatrix} \text{Probability of} \\ \text{meteorological} \\ \text{pattern } k \end{pmatrix} . \qquad (11)$$

An algorithm developed for identifying potential monitoring sites, as outlined in section 3, was used for searching for the highest values in the FOM field. This algorithm eliminates locations having high FOMs that are adjacent to locations having higher FOMs without an intervening trough. Such locations are considered to be adequately represented by the adjacent location having the higher FOM value. The isolated peaks of the FOM thus selected are chosen as potential candidates for monitoring stations. Because the NAAQS for CO have been specified as one-hour and eight-hour averages, computations for the FOM were carried out for each hour from 7:00 a.m. through 8:00 p.m. and for the two eight-hour periods near the morning and evening traffic peaks. These periods represent the highest CO concentrations either observed or predicted in the Las Vegas area because of the dominant contribution of automotive emissions. The resultant FOMs for the following time periods are shown in Figures 5, 6, and 7.

- One-hour period near the morning traffic peak (7:00 a.m.).
- One-hour period near the evening traffic peak (6:00 p.m.).
- Eight-hour period near the evening traffic peak (12:00 p.m. to 8:00 p.m.).

In these figures, isopleths for the FOM (in parts per million) overlay the selected monitoring locations. The locations are ranked alphabetically according to the magnitude of the FOM. Consequently, a total of more than 40 monitoring stations were identified and ranked. A perusal of these selected monitoring stations shows a pattern of proximity to major Las Vegas roadways, a fact that is not too surprising since traffic is the major emission source for CO. It is, however, interesting to note that all three sets of calculations

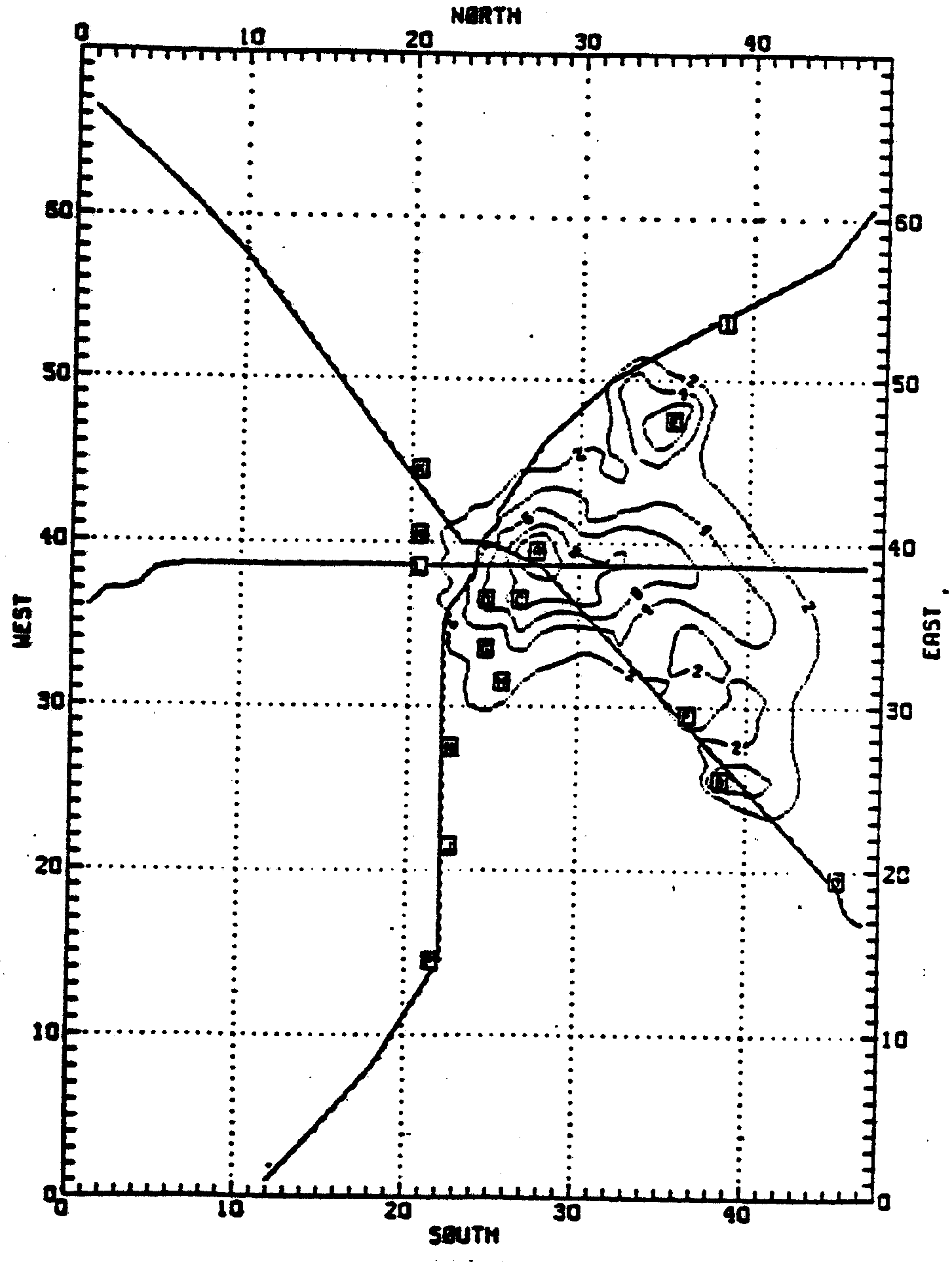


Figure 5. Ranked monitoring locations for CO based on one-hour averaging period near morning traffic peak: Las Vegas Valley.

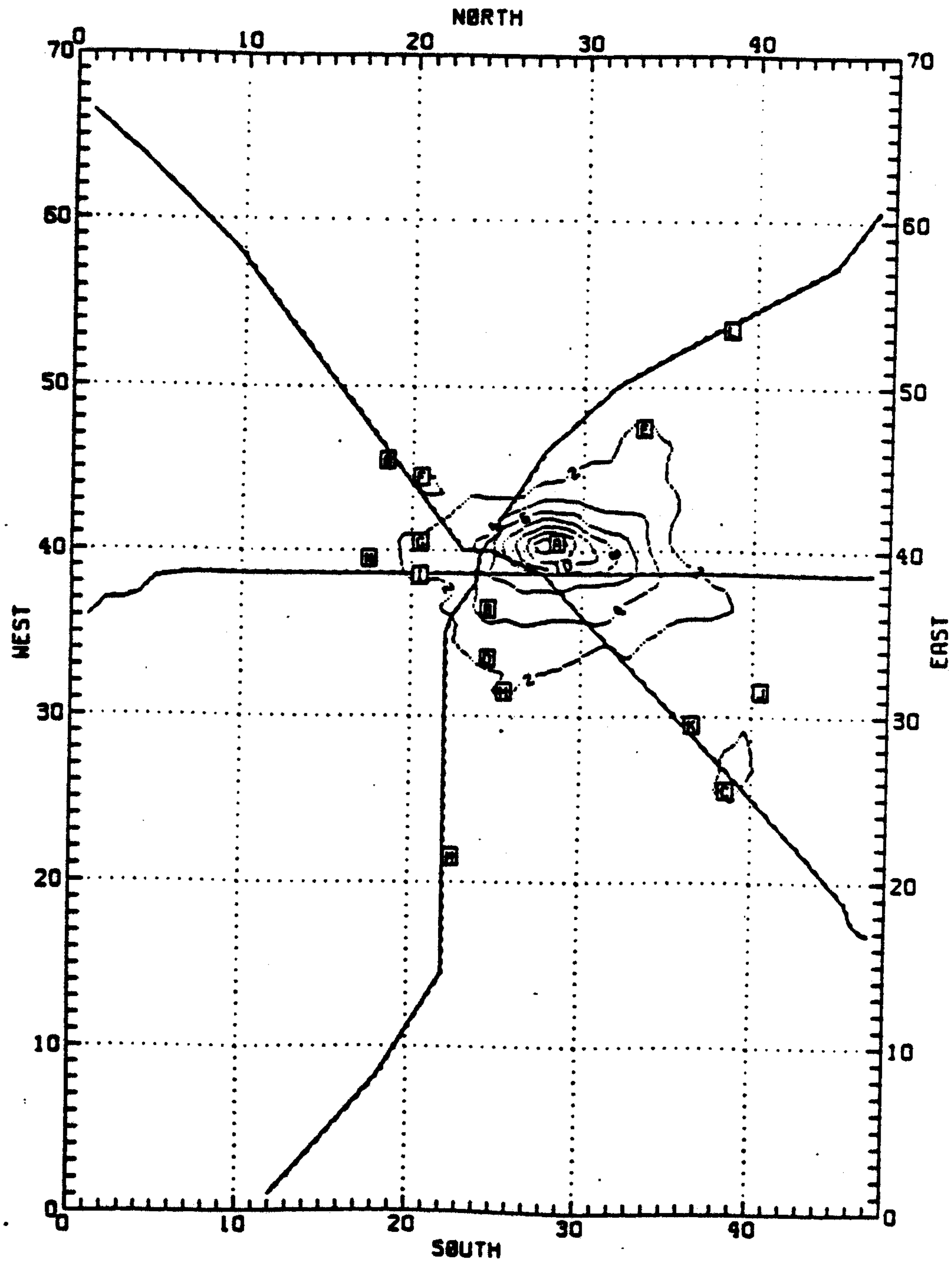


Figure 6. Ranked monitoring locations for CO based on one-hour averaging period near evening traffic peak: Las Vegas Valley.

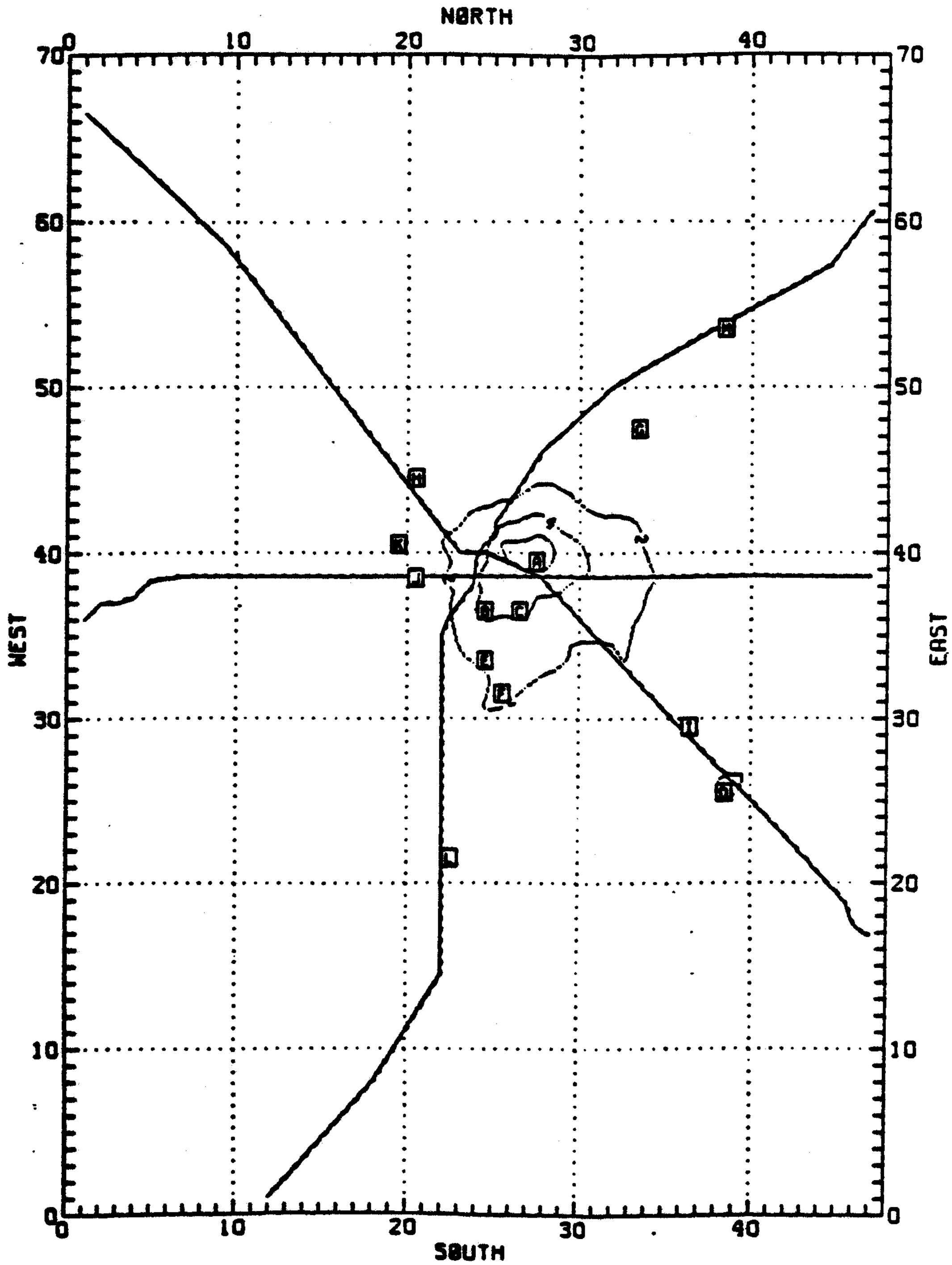


Figure 7. Ranked monitoring locations for CO based on eight-hour averaging period near evening traffic peak: Las Vegas Valley.

identify a location in the vicinity of an existing station on East Charleston (see Figure 4), a location that usually reports the highest CO concentrations in Las Vegas. The one-hour morning maximum (Figure 5) identified locations at Henderson and downtown Las Vegas as the second- and third-ranked locations, whereas the one-hour evening maximums (Figure 6) identified the same two locations, but in reverse order. This finding is probably due to the fact that Las Vegas morning traffic is job-related, whereas the evening traffic is primarily caused by visitors in the downtown area. This result seems to indicate that the FOM methodology developed under the present project can indeed detect subtle diurnal variations in an emissions pattern that may be unique to the Las Vegas area. In the next subsection, an optimum monitoring network for the metropolitan Las Vegas is established from among these stations by statistically determining the minimum number of monitoring stations required.

THE SECOND STEP--DETERMINING THE MINIMUM NUMBER OF STATIONS REQUIRED

As described in section 4, the determination of the optimum network configuration -- the minimum number of monitoring stations required -- is made by using an SOI for each of the ranked monitoring locations. The SOIs are, in turn, determined by the spatial correlation coefficients and the associated cutoff values for a prescribed confidence level.

Concentration fields for the six meteorological scenarios as determined by the model simulation provide the data base for evaluating the spatial correlation coefficients for each of the monitoring locations identified. Prior to the calculation of the spatial correlation coefficients, a smoothing of these concentration fields was accomplished to remove small-scale fluctuations. Similar operations are commonly used in turbulence research and numerical weather prediction (Shuman 1957, and Haltiner 1971). Smoothing in this study was accomplished by the following operation: Assuming that C^g_{xy} is the gth time-smoothed concentration at grid point (x,y), then the (g + 1)-th time-smoothed concentration field is obtained by

$$C^{g+1}_{xy} = \frac{C^g_{xy} + b\overline{C}^g_{xy}}{1 + b} \qquad (12)$$

where

$$\overline{C}^g_{xy} = \frac{1}{4}\left(\overline{C}^g_{x-1,y} + C^g_{x+1,y} + \overline{C}^g_{x,y-1} + C^g_{x,y+1}\right),$$

and b is a weighting factor to be empirically determined. A value of 2 was chosen for b, which is comparable to the well-known two-dimensional Shuman filter (Nelson and Weible 1980).

In the present application, three smoothing operations were sufficient to facilitate further analysis without altering the essential features of the original concentration fields. The data base thus consists 13 hourly smoothed concentration fields for each of the 6 meteorological scenarios for a total of 78 samples. Values of the cutoff sample correlation coefficient to ensure specific minimum values of the population coefficient and variance explained as determined from a numerical version of Figure 2 for a sample size of 78 and 95 percent confidence level are shown in Table 1.

TABLE 1. CUT-OFF SAMPLE CORRELATION COEFFICIENT TO ENSURE MINIMUM VALUES FOR POPULATION CORRELATION COEFFICIENT AND VARIANCE EXPLAINED FOR 95% CONFIDENCE WITH 78 SAMPLES

Sample Correlation Cut-Off Value r_c	Population Correlation Minimum Value ρ_c	Variation Explained Minimum Value ρ_c^2
0.4	0.18	0.03
0.5	0.30	0.1
0.6	0.44	0.2
0.7	0.56	0.3
0.8	0.70	0.5
0.9	0.85	0.7

To demonstrate the utility of the siting methodology, a total of 19 monitoring locations was selected from the highest ranking monitoring locations determined. The 13 highest ranking locations, determined by the eight-hour FOM for the evening traffic peak, were augmented by three locations each from the highest ranking one-hour FOM for the morning and evening traffic peaks, which were not adjacent to the locations already selected. The 19 stations were selected to cover maximum or peak concentrations. In addition, as a further demonstration of the methodology for secondary monitoring objectives discussed earlier, three stations located in the northern, western, and southeastern outskirts of the city were arbitrarily added to measure either the background or baseline air quality in the Las Vegas area. The locations and characterizations of these 22 stations are listed in Table 2.

A specific selection of sample correlation-coefficient isopleths for each of the 22 stations is shown in Appendix B. As described in section 3, the SOI is dictated by the cutoff correlation coefficient. Assuming that the criterion for an optimum network design is its capability for catching at least 50 percent of the concentration variations 95 percent of the time, then Equation (9) yields a minimum value for ρ_c of 0.7.

TABLE 2. IDENTIFICATION OF POTENTIAL MONITORING LOCATIONS IN THE LAS VEGAS VALLEY

Station Identification	x-Coordinate	y-Coordinate	Comments
A	28	40	
B	25	37	
C	27	37	
D	39	26	Locations determined by eight-hour figure-of merit near the evening traffic peak
E	25	34	
F	26	32	
G	34	48	
H	21	45	
I	37	30	
J	21	39	
K	20	41	
L	23	22	
M	39	54	
N	41	32	Locations determined by one-hour figure-of-merit at the evening traffic peak
O	18	40	
P	19	46	
Q	23	48	Locations determined by one-hour figure-of-merit at the morning traffic peak
R	46	20	
S	22	15	
T	4	37	Arbitrarily chosen rural locations
U	28	60	
V	43	15	

With a sample size of 78, Table 1 shows that this value of ρ_c corresponds to cutoff sample correlation coefficient of 0.8. Therefore, this value was used to determine the spheres of influence for each of the 22 stations listed in Table 2. Stations were deleted from the list if the individual areal coverage, after eliminating overlapping regions already covered by higher ranking stations, was less than 10 percent of the coverage of the highest ranked station. As a result, a total of 10 air quality monitoring stations was selected, among which 3 are rural background stations. The locations of these 10 stations and their joint areal coverage (shaded areas) are shown in Figure 8. For each of these stations, the following statistics were compiled to measure the effectiveness of individual stations as well as that of the overall network:

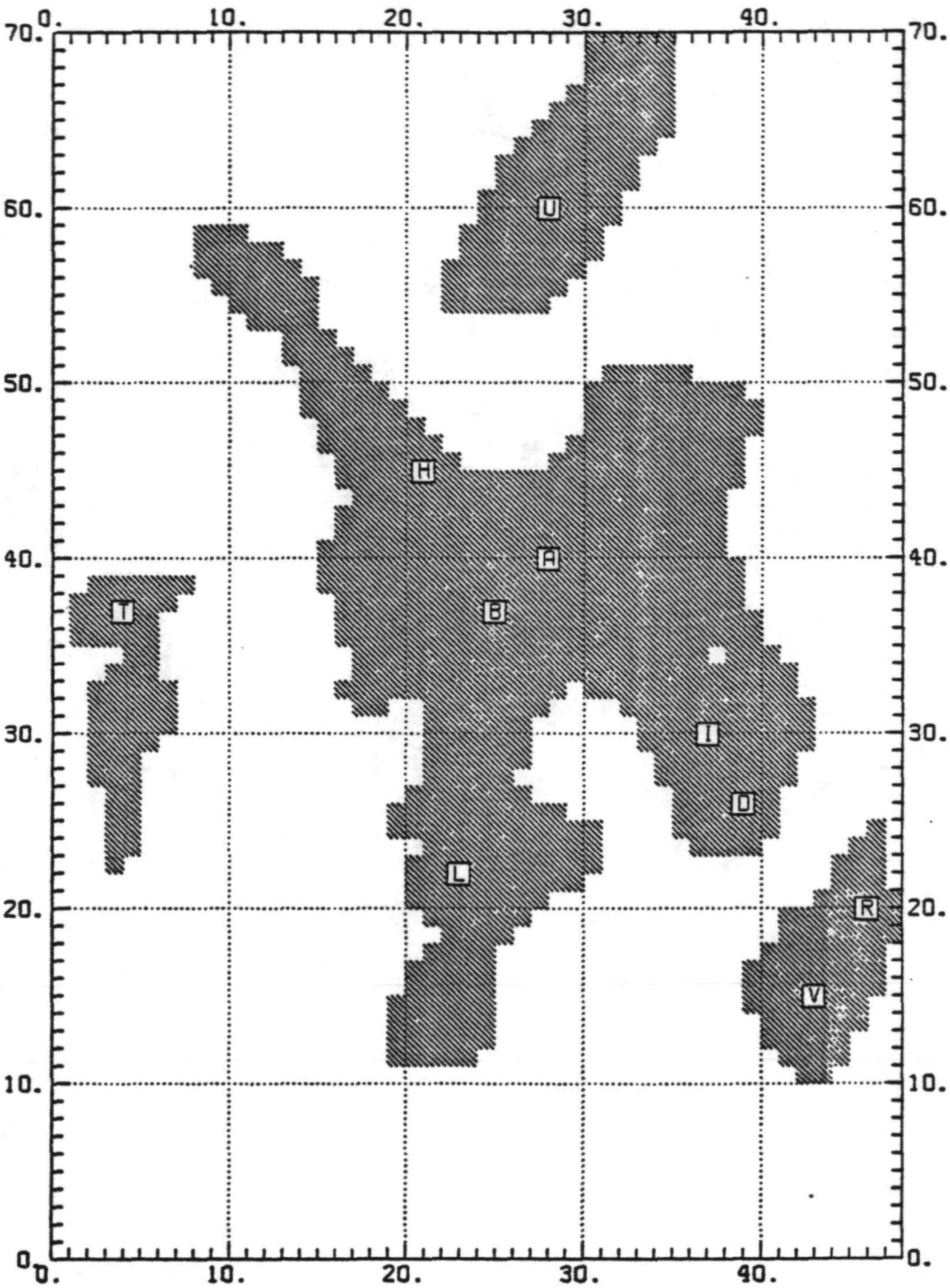


Figure 8. An optimum air quality monitoring network for Las Vegas Valley based on a cutoff sample spatial correlation coefficient of 0.8.

- The fraction of area covered by the individual station based on the total area considered (48 km x 70 km = 3,360 km^2), as determined by SOI.

- The cumulative areal coverage, expressed as the fraction of the total area considered, beginning with the highest-ranked station.

These summary statistics are listed in Table 3.

TABLE 3. SUMMARY STATISTICS FOR SELECTED MONITORING STATIONS IN LAS VEGAS VALLEY BASED ON A CUTOFF SAMPLE SPATIAL CORRELATION COEFFICIENT OF 0.8

Station	x-Coordinate	y-Coordinate	Individual Station Fractional Areal Coverage	Cumulative Fractional 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

As a sensitivity test of the siting methodology, identical computations were made using a cutoff sample correlation coefficient of 0.5. This network configuration and its corresponding joint areal coverage (shaded areas) are presented in Figure 9. A total of 7 air quality monitoring stations were selected by the siting methodology. Among these stations, 2 are rural background stations. Summary statistics for the 7-station network are given in Table 4. It is interesting to note that cumulative areal coverage increases from 26.6 percent of the total region for the 10-station network to 62.3 percent for the 7-station network. However, it should be noted as is shown in Table 1, that the 7-station network can only detect a minimum of 10 percent of the concentration variations, whereas the 10-station network can detect a minimum of 50 percent of the concentration variations for the area within the combined spheres of influence (see Figures 8 and 9).

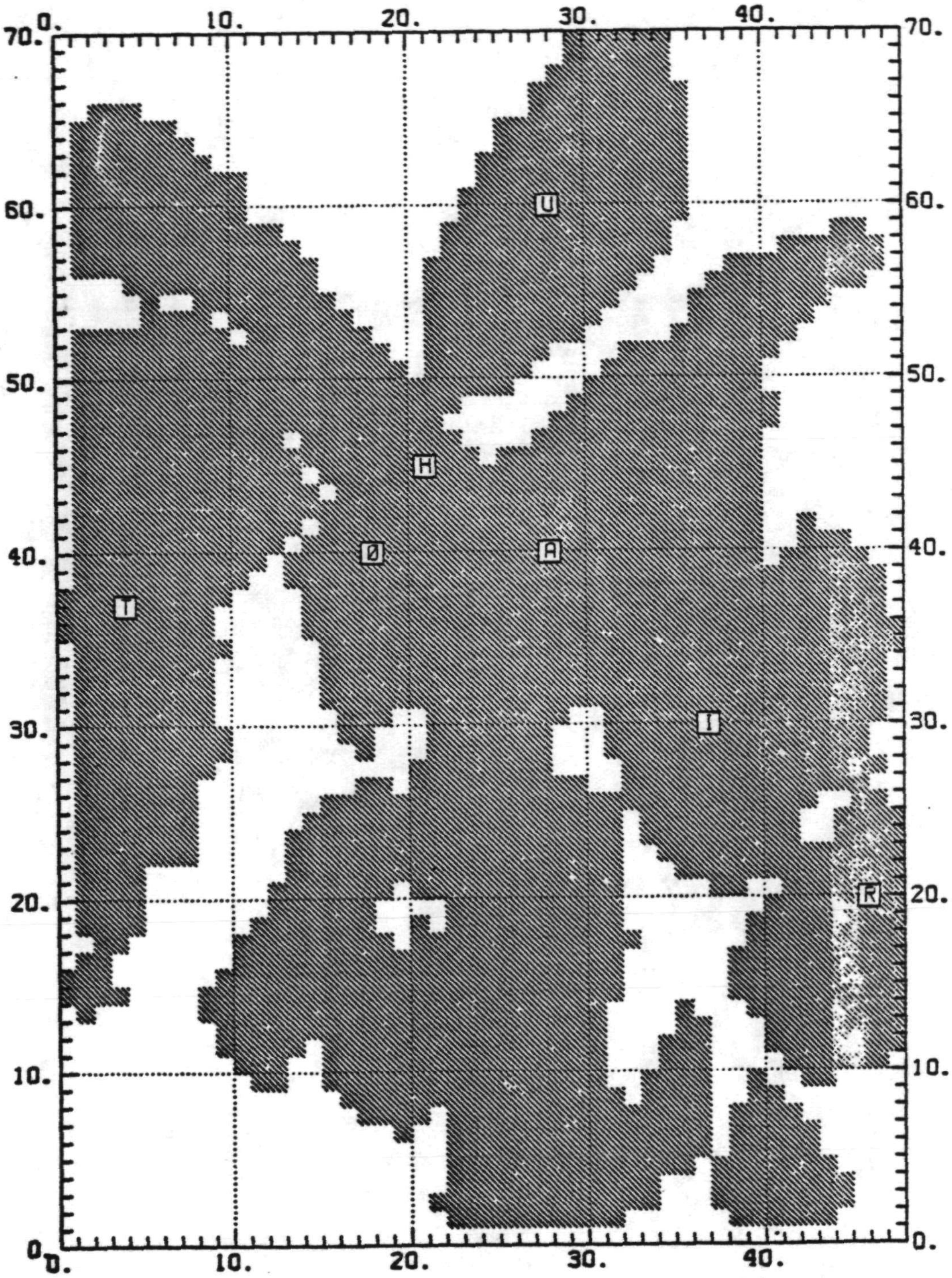


Figure 9. An optimum air quality monitoring network for Las Vegas Valley based on a cutoff sample spatial correlation coefficient of 0.5.

TABLE 4. SUMMARY STATISTICS FOR SELECTED MONITORING STATIONS IN LAS VEGAS VALLEY BASED ON A CUTOFF SAMPLE SPATIAL CORRELATION COEFFICIENT of 0.5

Station	x-Coordinate	y-Coordinate	Individual Station Fractional Areal Coverage	Cumulative Fractional Areal Coverage
A	28	40	0.2604	0.2604
H	21	45	0.0664	0.3268
I	37	30	0.0327	0.3595
O	18	40	0.0429	0.4024
R	46	20	0.0354	0.4378
T	4	37	0.1232	0.5610
U	28	60	0.0622	0.6232

Further observations can be made concerning these monitoring networks as determined by the siting methodology:

- As shown in Figures 8 and 9, all urban stations chosen are located, as expected, along the main transportation corridors that constitute the bulk of carbon monoxide emissions sources (about 80 percent) in the Las Vegas area. Such stations coresond to the "street canyon" and "traffic corridor" stations referred to by Ludwig, Berg, and Hoffman (1976) and the "A"-type stations referred to by Ott (1975).

- The joint areal coverage of these stations (as shown in Figures 8 and 9) tends to shift toward the east, southeast, and south of the major emissions sources, apparently reflecting prevalent local wind directions in Las Vegas for the meteorological scenarios used.

- The western and northern rural background stations designated as stations T and U, were selected in both networks presumably because air quality in the neighborhood of these sites is not significantly affected by the major emissions sources in the metropolitan Las Vegas area.

SECTION 6

CONCLUDING REMARKS

An objective methodology was presented for determining an optimum air quality monitoring network for an urban area. This methodology, in totality, yields both the number and disposition of the monitoring stations. Although the methodology is applicable to all pollutants, it was applied to the Las Vegas Valley using a relatively inert species, carbon monoxide, as an example. The exercise of the methodology for Las Vegas suggests that the design of an air quality monitoring network can be accomplished in a technically rigorous manner.

REFERENCES

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).

Ezekiel, M. (1941), Methods of Correlation Analysis (John Wiley and Sons, Inc., London, England).

Haltiner, G. J. (1971), Numerical Weather Prediction (John Wiley and Sons, Inc., New York, New York).

Liu, M. K. and G. E. Moore (1980), "Development and Application of a Methodology For Air Quality Monitoring Network Design," Proc. of Second Joint Conference on Applications of Air Pollution Meteorology and the Second Conference on Industrial Meteorology, March 1980, New Orleans, Louisiana, pp. 727-733.

Liu, M. K. et al. (1977), "Development of a Methodology for the Design of a Carbon Monoxide Monitoring Network," EPA-600/4-77-019, U.S. Environmental Protection Agency, Las Vegas, Nevada.

Ludwig, F. L. and J. H. S. Kealoha (1975), "Selecting Sites for Carbon Monoxide Monitoring," EPA 450/3-75-077. U.S. Environmental Protection Agency, Research Triangle Park, North Carolina.

Ludwig, F. L., N. J. Berg, and A. J. Hoffman (1976), "The Selection of Sites for Air Pollutant Monitoring," Paper presented at the 69th Annual Meeting of the Air Pollut. Control Assoc., 27 June - 1 July 1976, Portland, Oregon.

McElroy, J. L. et al. (1978), "Carbon Monoxide Monitoring Network Design Methodology," EPA-600/4-78-053, U.S. Environmental Protection Agency, Las Vegas, Nevada.

Nelson, S. P. and M. L. Weibel (1980), "Three-Dimensional Shuman Filter," J. Appl. Meteorol., Vol. 19, pp. 464-469.

Ott, W. R. (1975), "Development of Criteria for Siting Monitoring Stations," Paper presented at 68th Annual Meeting of the Air Pollut. Control Assoc., Boston, Massachusetts.

Ott, W. R. and G. C. Thom (1976), "A Critical Review of Air Pollution Index Systems in the United States and Canada," J. Air Pollut. Control Assoc., Vol. 26, pp. 460-470.

Shuman, F. G. (1957), "Numerical Methods in Weather Prediction: II. Smoothing the Filtering," Mon. Wea. Review, Vol. 85, pp. 357-361.

U.S. Environmental Protection Agency (1971), "Guidelines: Air Quality Surveillance Network," Public Health Service, AP-98.

APPENDIX A

A HEURISTIC APPROACH FOR OPTIMAL MONITORING NETWORK DESIGN

In this appendix, a summary is presented of a heuristic approach used in a preliminary study by Liu and Moore (1980) for determining the optimum monitoring network and its application to metropolitan Las Vegas.

Similar to the statistical approach described in the text, the heuristic approach also consists of two steps. Furthermore, the first step in both approaches is identical. Maxima or peak values of figure-of-merit (FOM) are used to identify and rank potential monitoring stations. In the second step, the two approaches differ principally in the definition of the sphere of influence (SOI).

On the basis of a heuristic approach, the SOI was defined by Liu and Moore (1980) as the monitoring site itself plus those grid locations along a straight line in any direction from the site showing a continually decreasing FOM. An algorithm was developed that employs both a forward mode and a backward mode to screen the most suitable candidates for the monitoring network from the N monitoring locations identified in the first step. The forward mode starts with the highest FOM grid and defines the SOI for each of the N monitoring stations in descending rank of FOM.

The backward mode is used to eliminate lower-ranked stations that do not contribute to the total spatial or FOM coverage of the monitoring network. In the backward mode, the SOIs of the lower-ranked stations are overlayed by the SOIs of the higher-ranked stations. At the end of this operation, the stations that do not appear in the listing are those that are completely contained in the SOIs of the higher-ranked stations. These stations, considered as redundancies, can be thus deleted from the network without affecting the adequacy of coverage. Consequently, an optimum monitoring network is determined.

This heuristic approach was applied to metropolitan Las Vegas using the data base described earlier. Summary statistics generated during the forward and backward modes of the site selecting algorithm are shown in Tables A-1, A-2, and A-3. The station locations are identified in Figures 5, 6, and 7, respectively, in the text of this report. It can be seen from the tables that, though the inclusion of all five stations is necessary on the basis of the one-hour morning maxima, only three stations overlapping the five stations for the morning maxima are required on the basis of the one-hour evening and the eight-hour evening maxima. These five stations, denoted as A to E in Figure 5 of the text, then represent the optimal network for monitoring carbon monoxide in the Las Vegas area, as determined using the heuristic approach.

A comparison of the 5-station network developed using the heuristic approach with the 10-station or 7-station networks developed using the statistical approach shows that both approaches place the majority of the monitoring stations, with a varying degree of emphasis, in downtown Las Vegas and Henderson in the southeast (Figure 4). The principal difference between the two approaches lies in the areal coverage statistics. The areal coverage in the heuristic approach may be grossly inflated because the SOI based on a line of sight can be extended indefinitely. The areal coverage determined by the statistical approach is apparently more realistic. Thus, the statistical approach is favored over the heuristic approach in the determination of the optimum monitoring network.

TABLE A-1. CUMULATIVE COVERAGE STATISTICS FOR THE MONITORING NETWORK FOR CO BASED ON ONE-HOUR AVERAGING PERIOD NEAR MORNING TRAFFIC PEAK--LAS VEGAS VALLEY (7 a.m.)

Station	Peak Value (ppm)	Cumulative Percent of Areal Coverage	Cumulative Percent of Figure-of-Merit Coverage
A	13.763	70.7%	67.7%
B	10.200	87.0	80.5
C	9.492	88.5	83.2
D	8.249	88.6	83.5
E	7.390	98.2	93.7

TABLE A-2. CUMULATIVE COVERAGE STATISTICS FOR THE MONITORING NETWORK FOR CO BASED ON ONE-HOUR AVERAGING PERIOD NEAR EVENING TRAFFIC PEAK--LAS VEGAS VALLEY (6 p.m.)

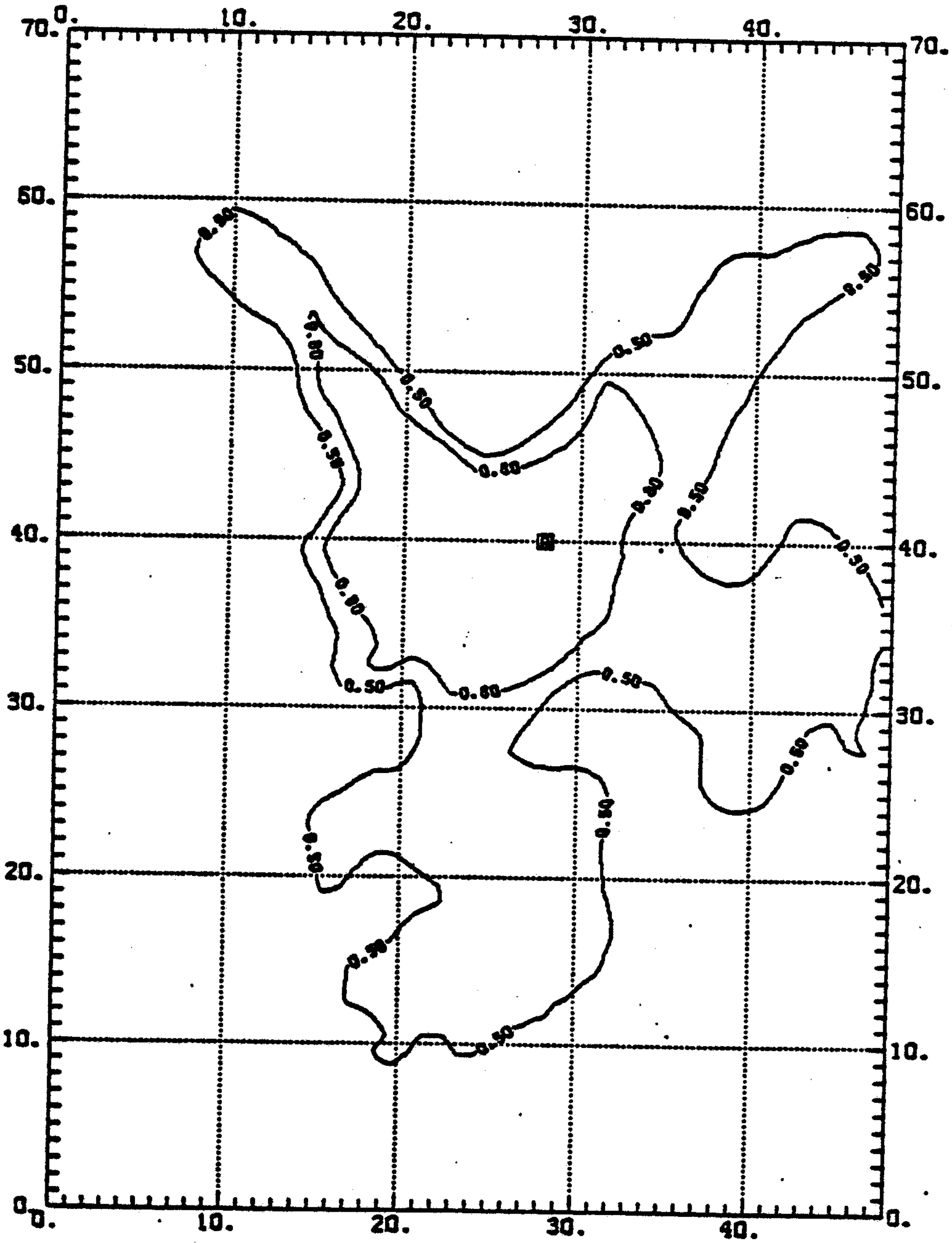
Station	Peak Value (ppm)	Cumulative Percent of Areal Coverage	Cumulative Percent of Figure-of Merit Coverage
A	15.472	89.6%	91.4%
B	6.233	96.8	96.6
C	4.870	99.8	99.9
D	2.796	99.8	99.9
E	2.445	99.9	99.9

TABLE A-3. CUMULATIVE COVERAGE STATISTICS FOR THE MONITORING NETWORK FOR CO BASED ON EIGHT-HOUR AVERAGING PERIOD NEAR EVENING TRAFFIC PEAK--LAS VEGAS VALLEY (12 p.m. - 8 p.m.).

Station	Peak Value (ppm)	Cumulative Percent of Areal Coverage	Cumulative Percent of Figure-of-Merit Coverage
A	7.461	90.0%	91.5%
B	5.209	98.4	98.2
C	4.593	98.9	98.7
D	3.410	100.0	100.0
E	2.672	100.0	100.0

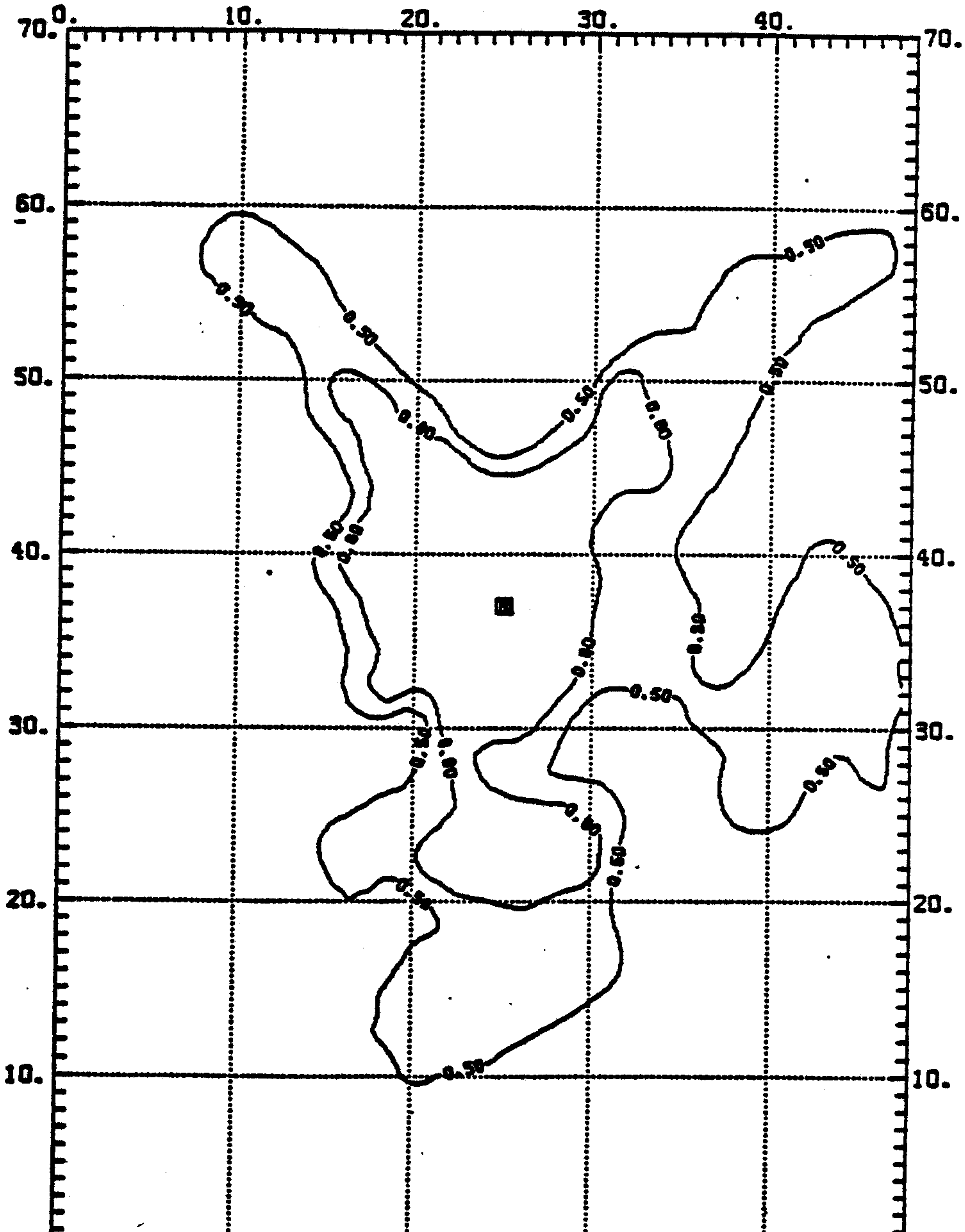
APPENDIX B

SAMPLE SPATIAL CORRELATION-COEFFICIENT ISOPLETHS IN LAS VEGAS VALLEY



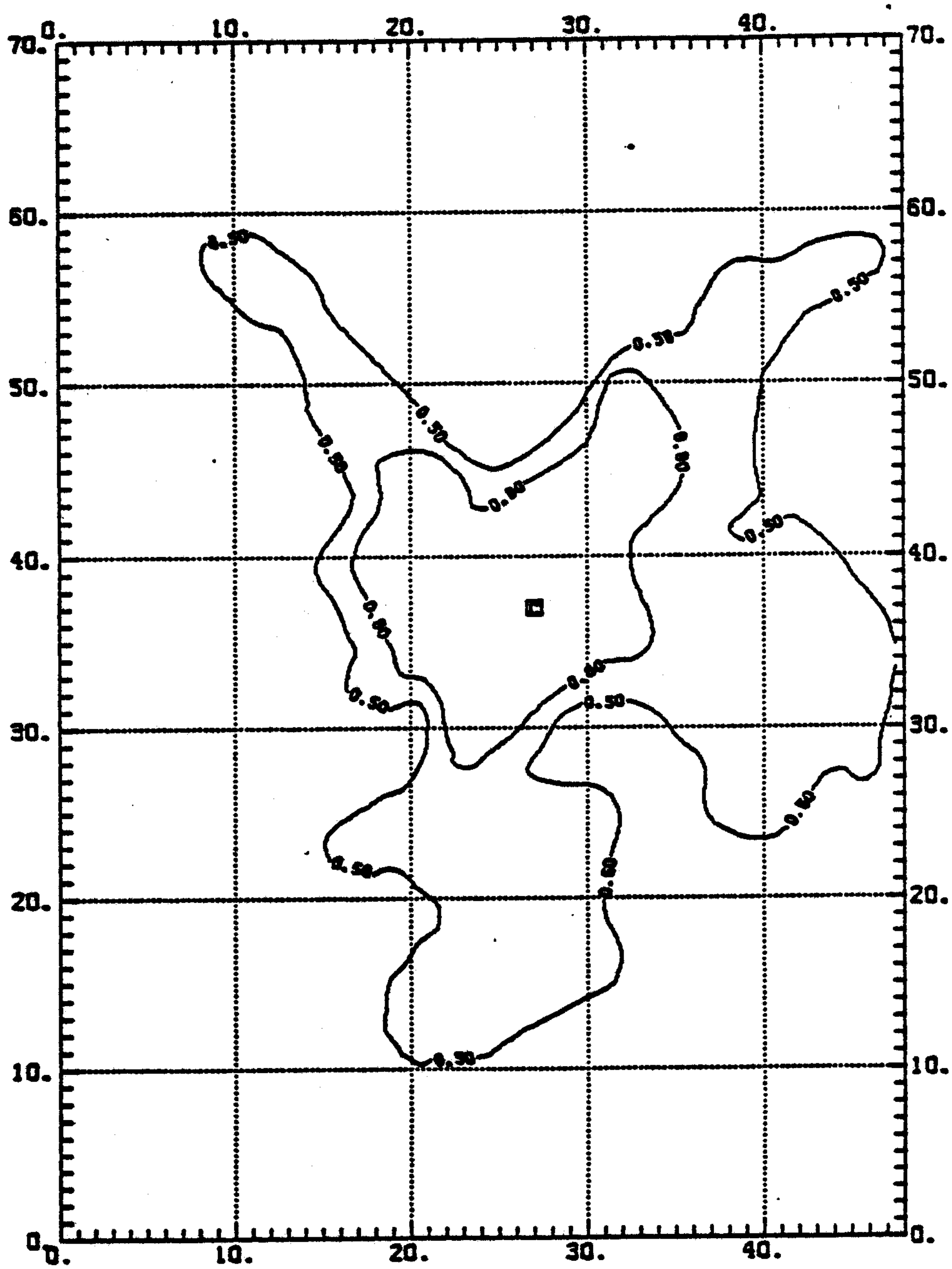
(a) Site A

Figure B-1. Sample spatial correlation-coefficient isopleths in Las Vegas Valley.



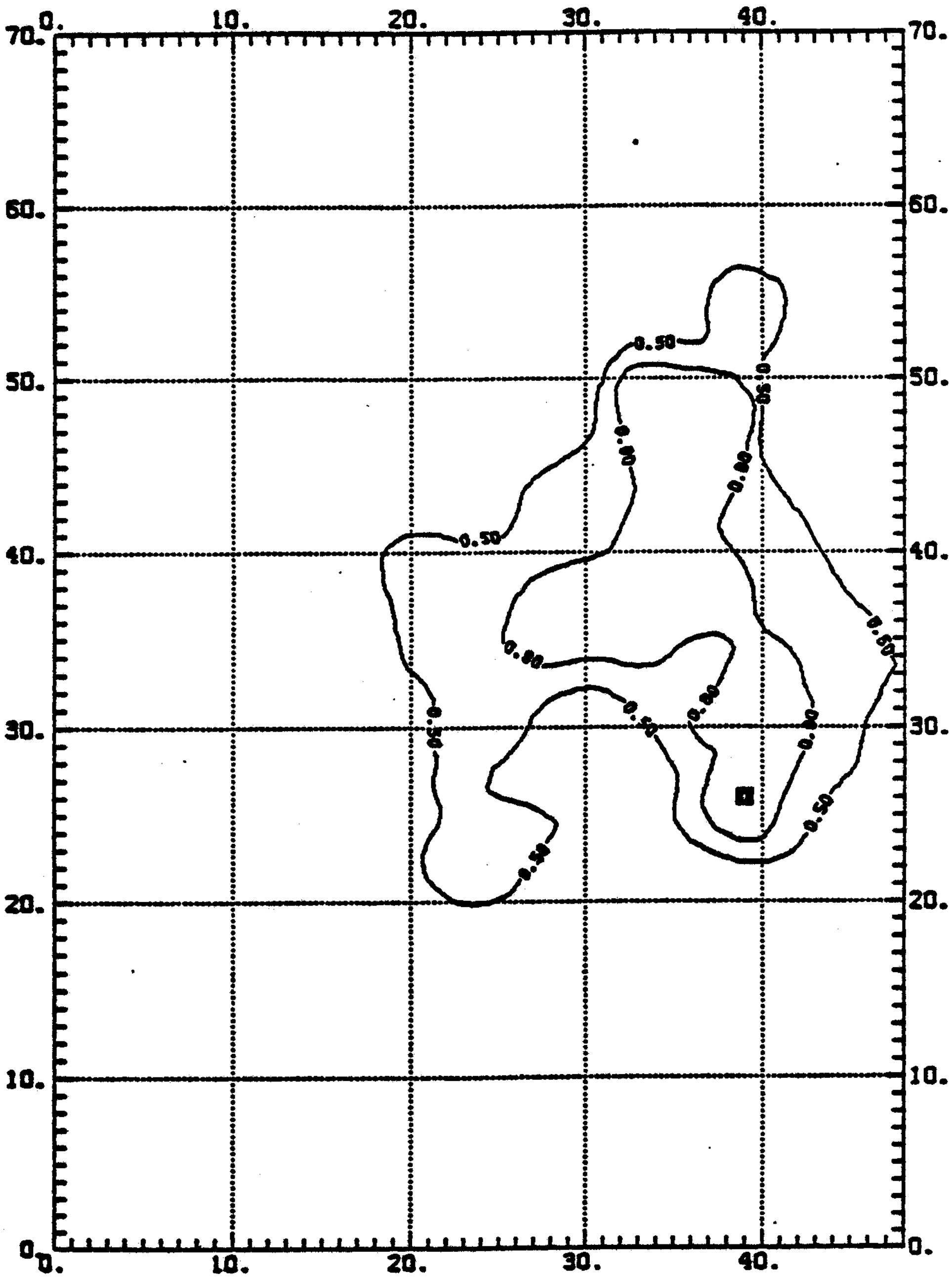
(b) Site 8

Figure B-1. (continued)



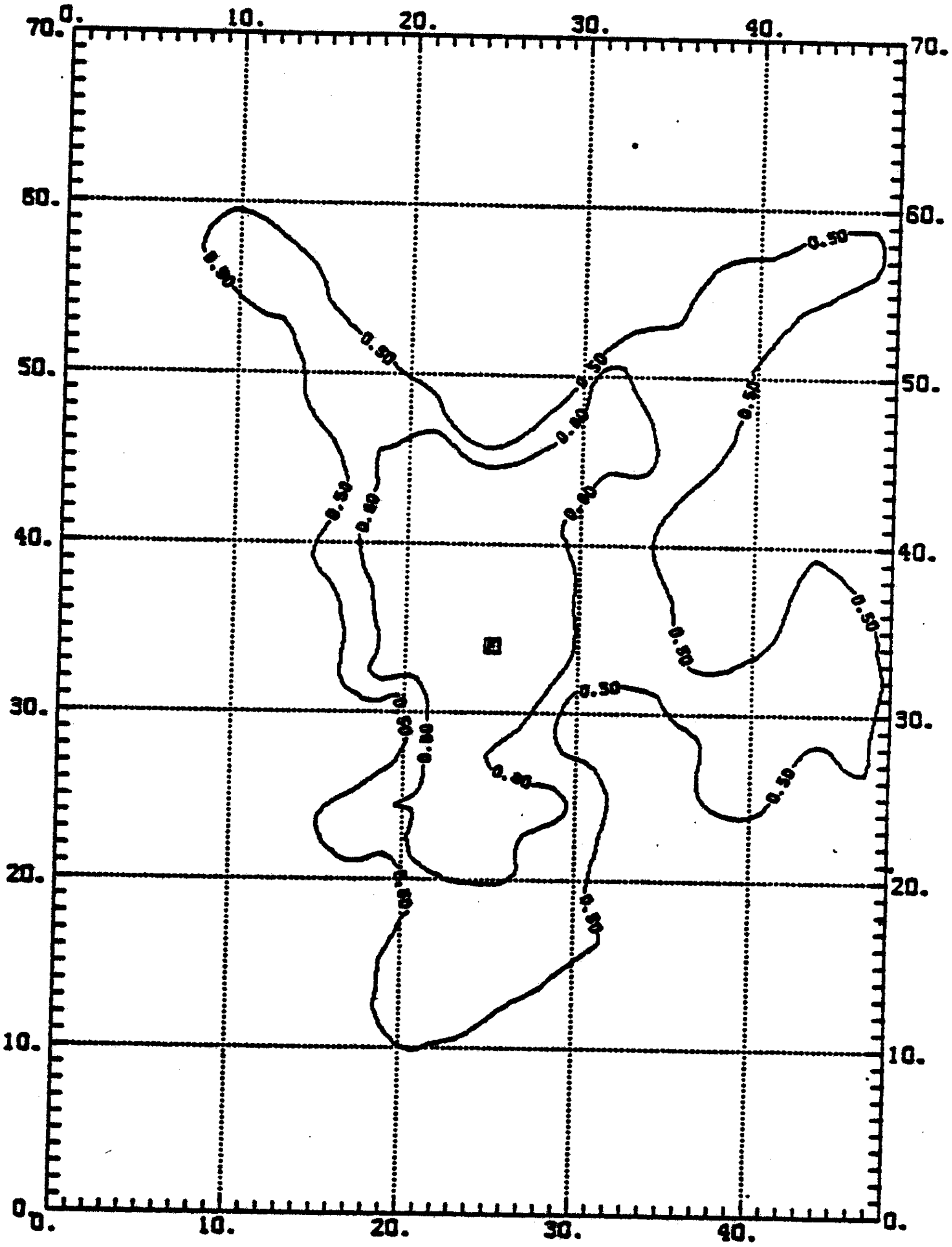
(c) Site C

Figure B-1. (continued)



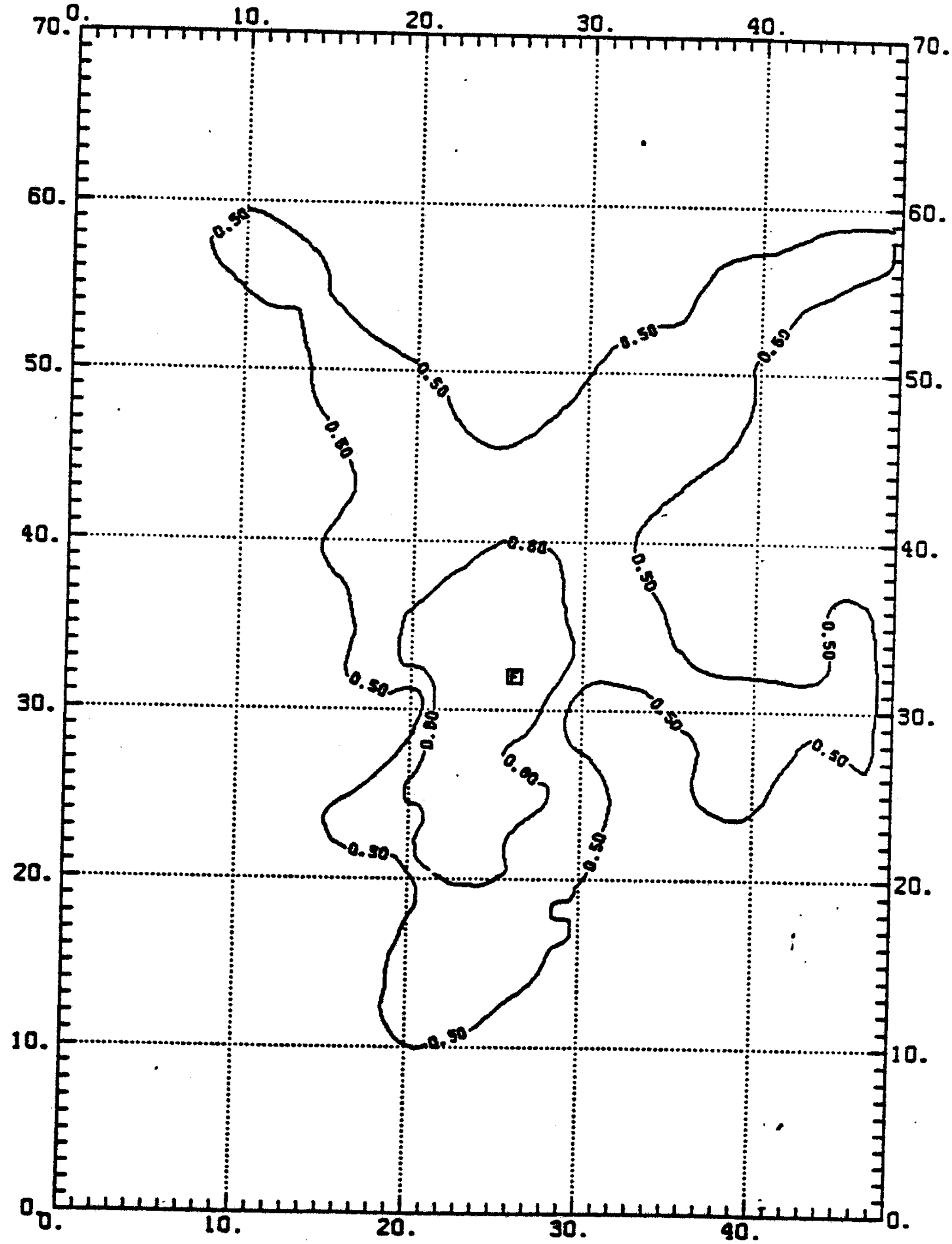
(d) Site D

Figure B-1. (continued)



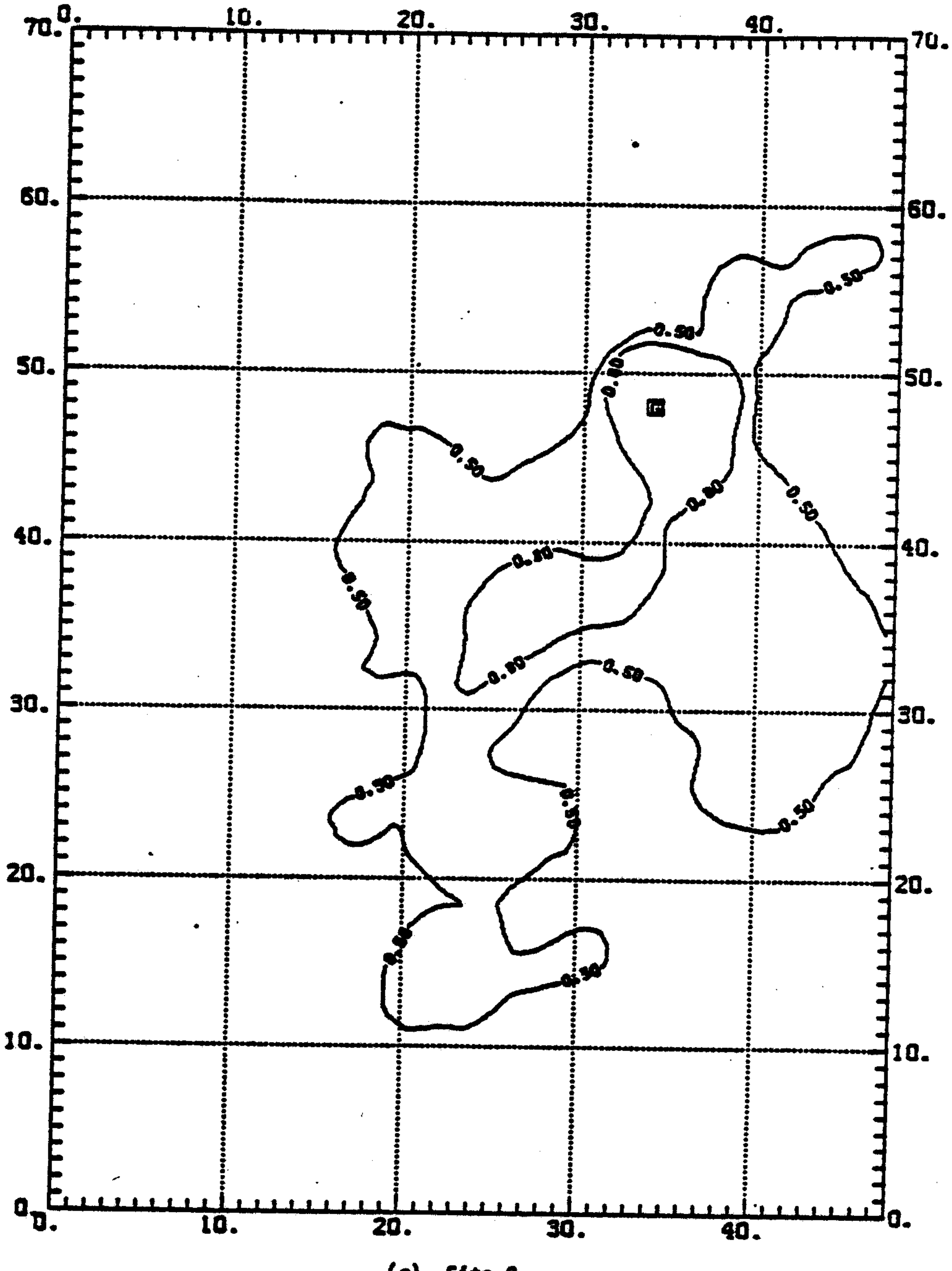
(e) Site E

Figure B-1. (continued)



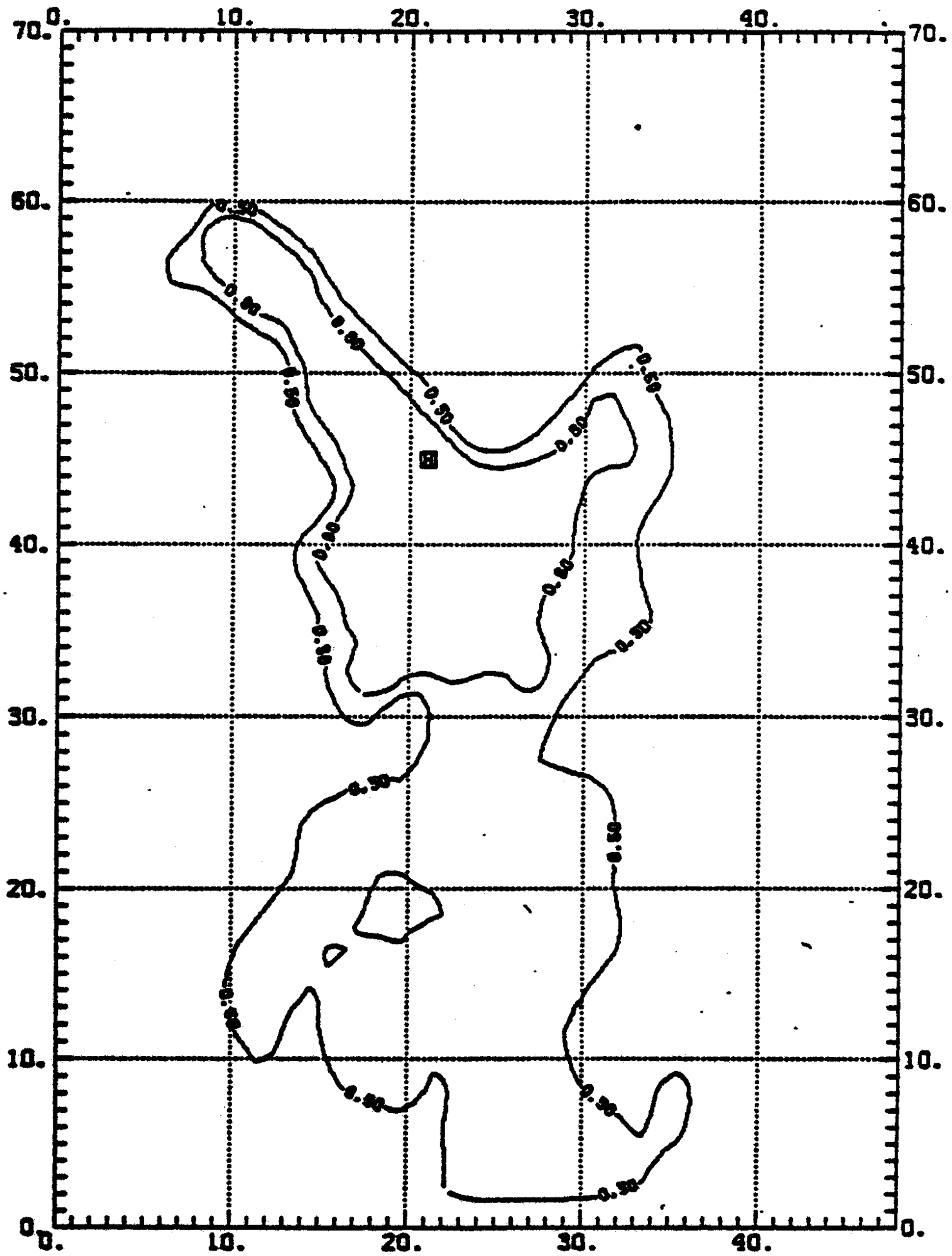
(f) Site F

Figure B-1. (continued)



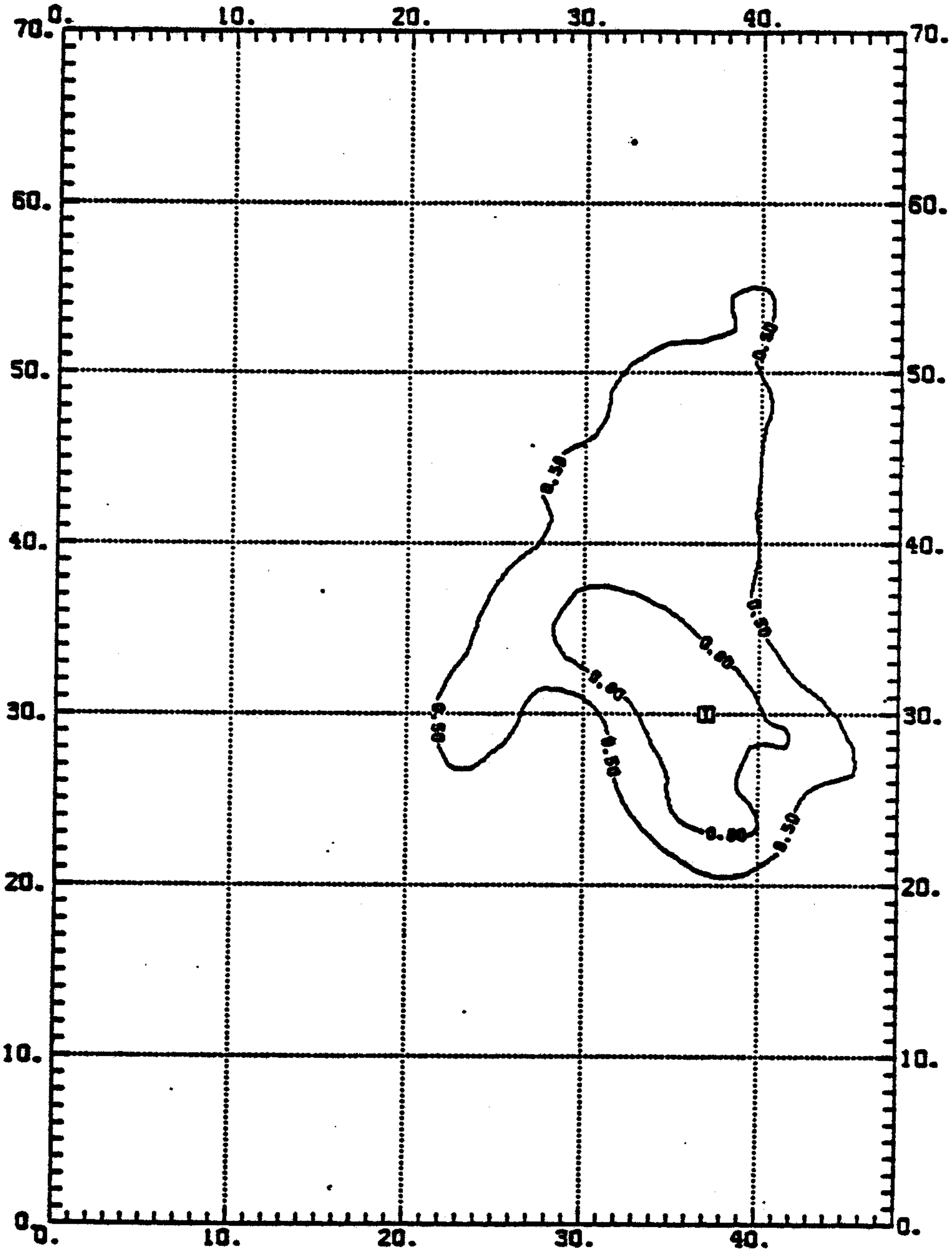
(g) Site G

Figure B-1. (continued)



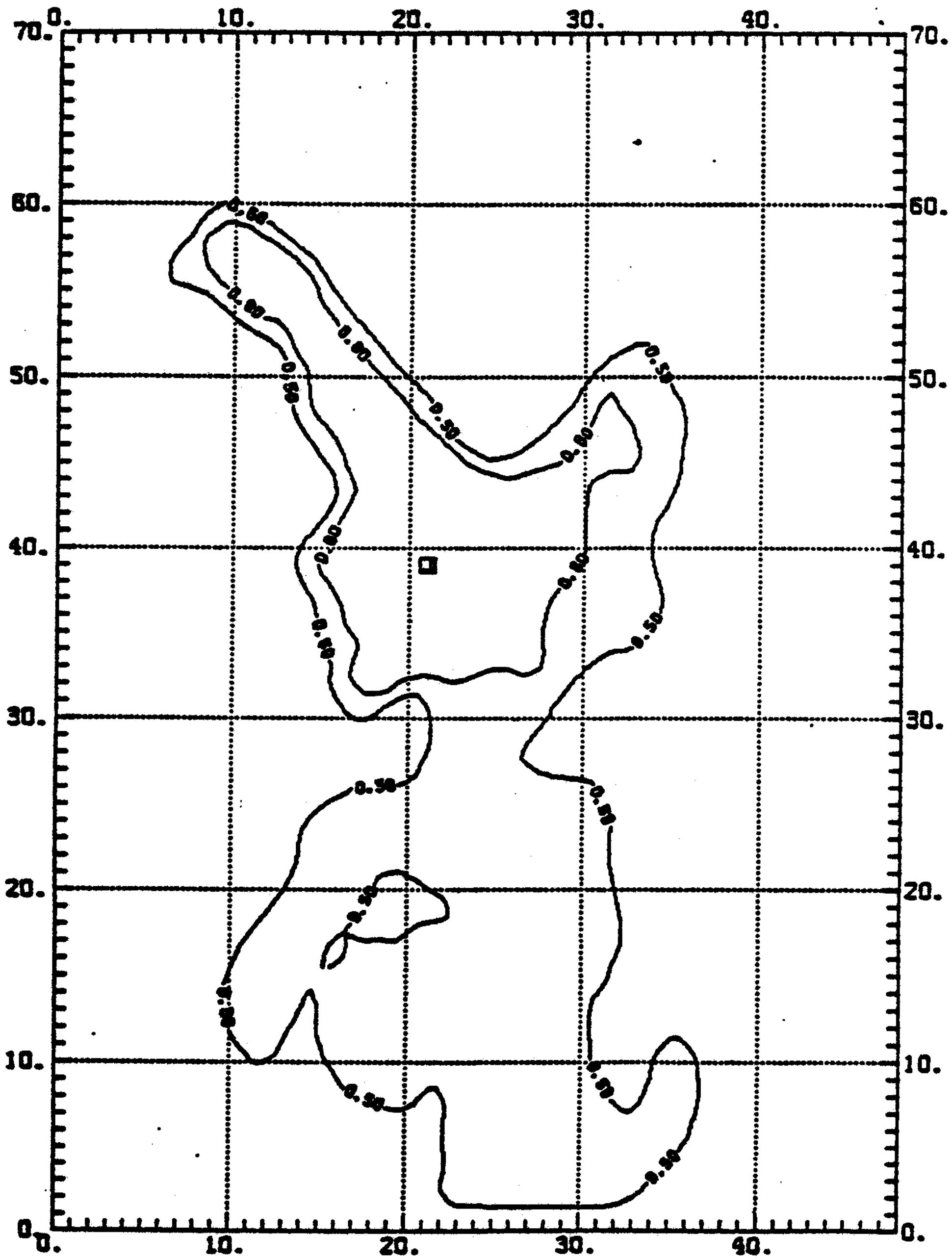
(h) Site H

Figure B-1. (continued)



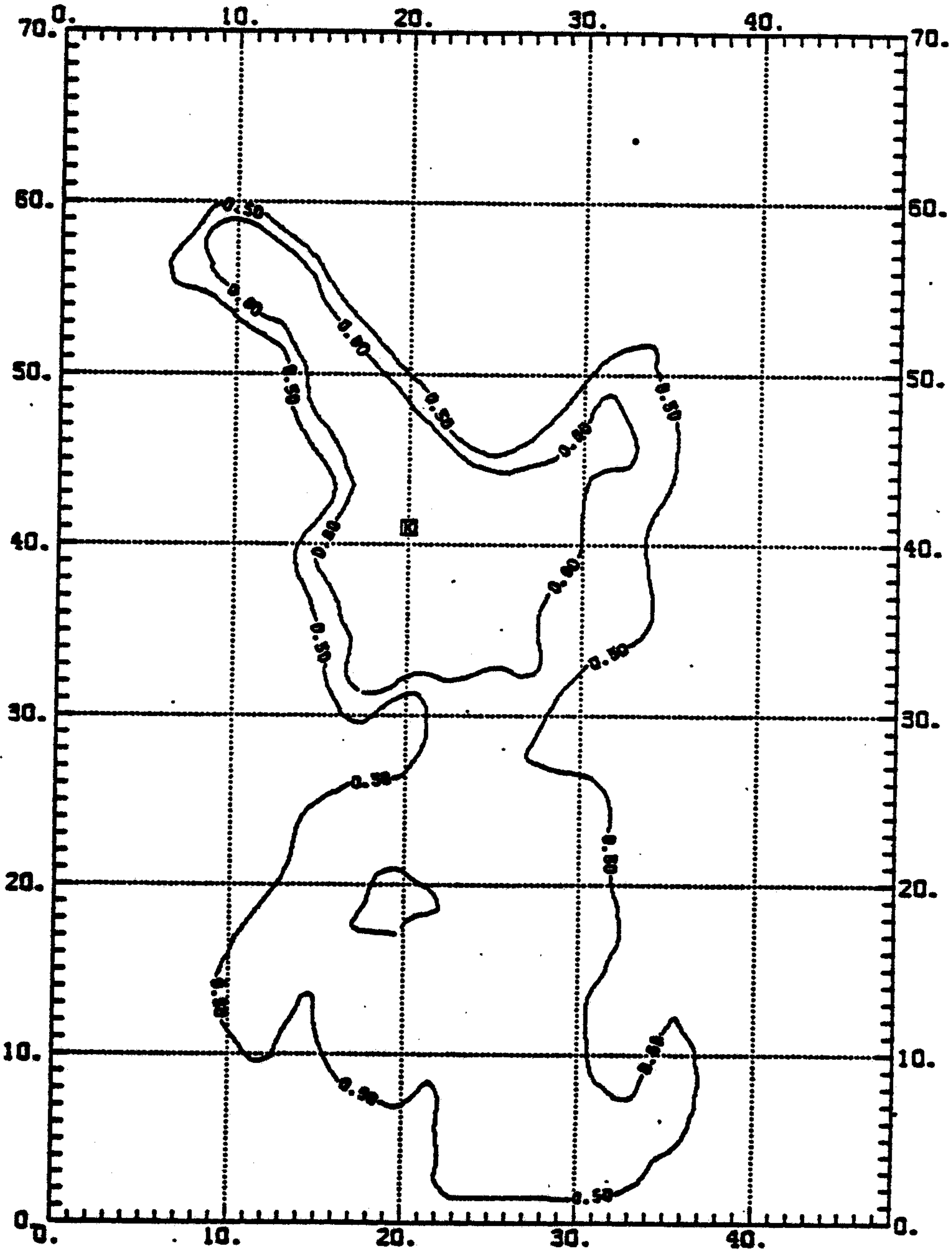
(1) Site I

Figure B-1. (continued)



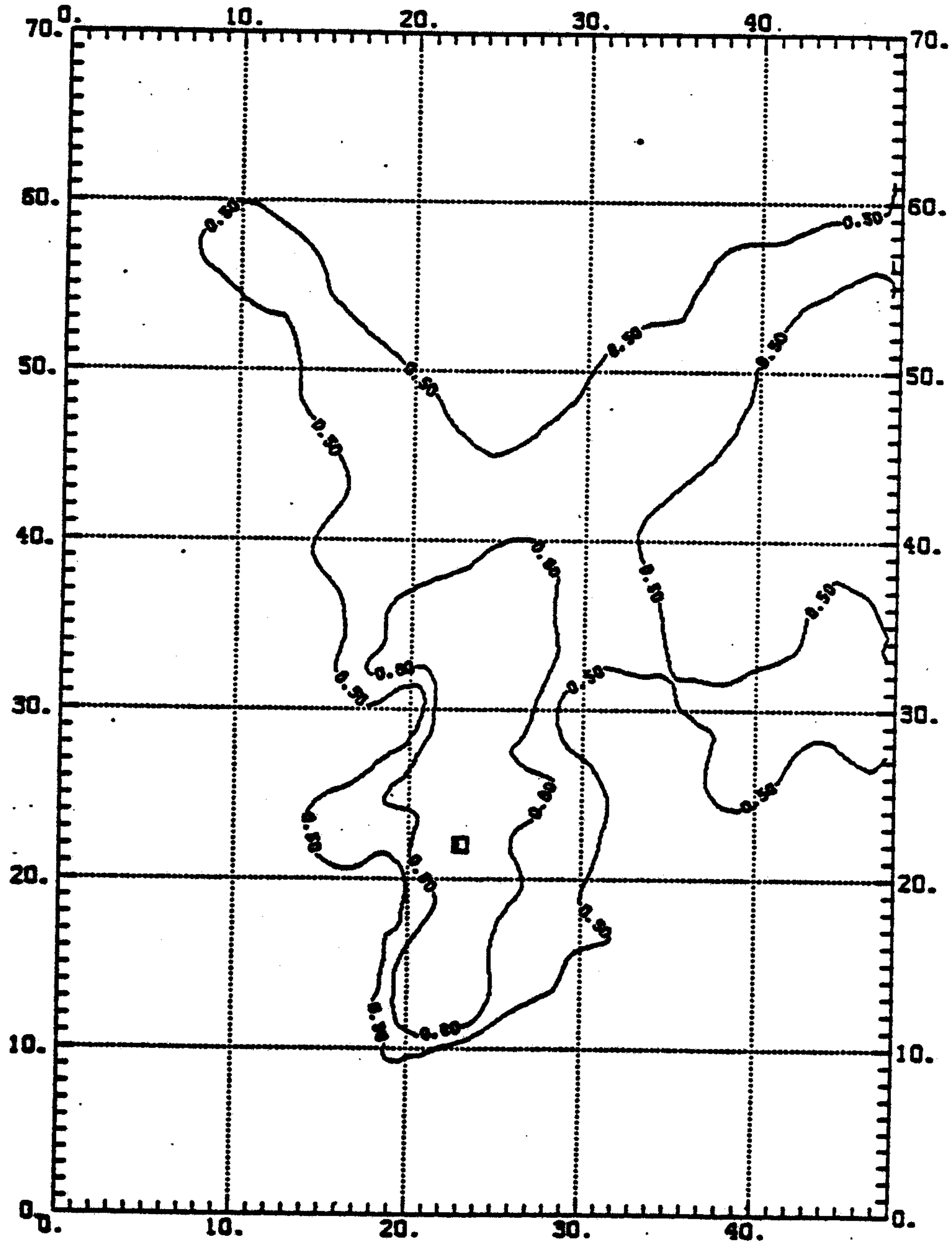
(j) Site J

Figure B-1. (continued)



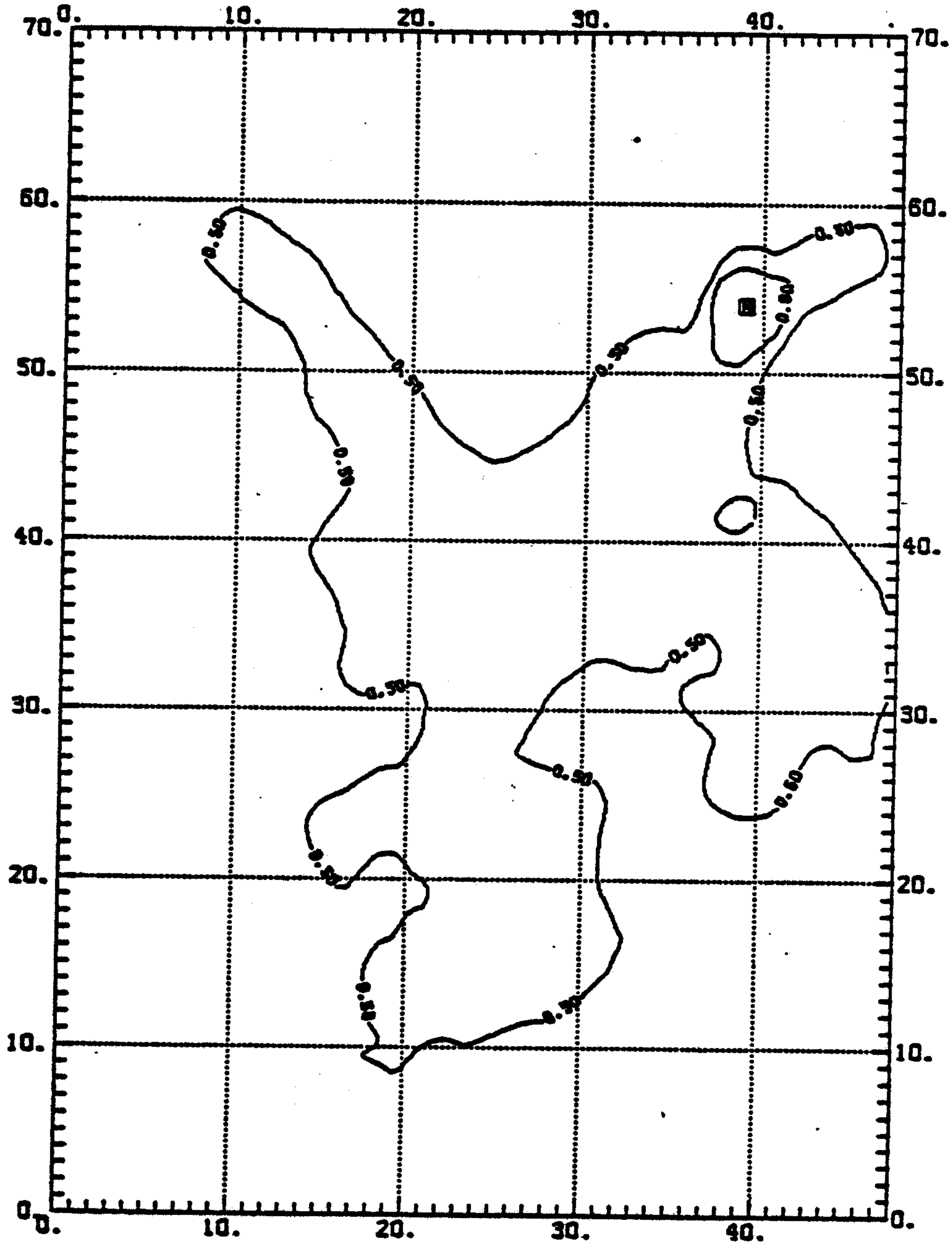
(k) Site K

Figure B-1. (continued)



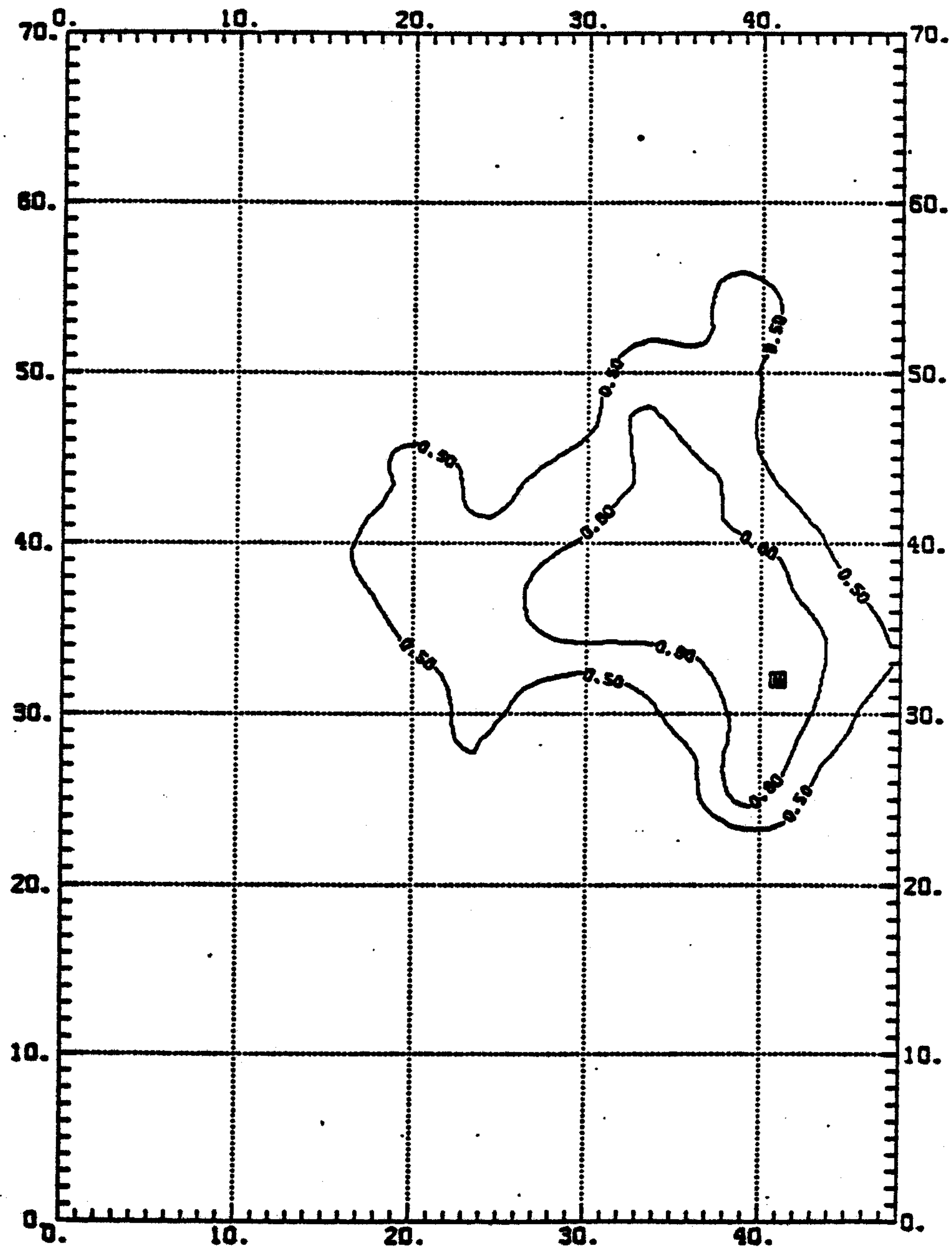
(1) Site L

Figure B-1. (continued)



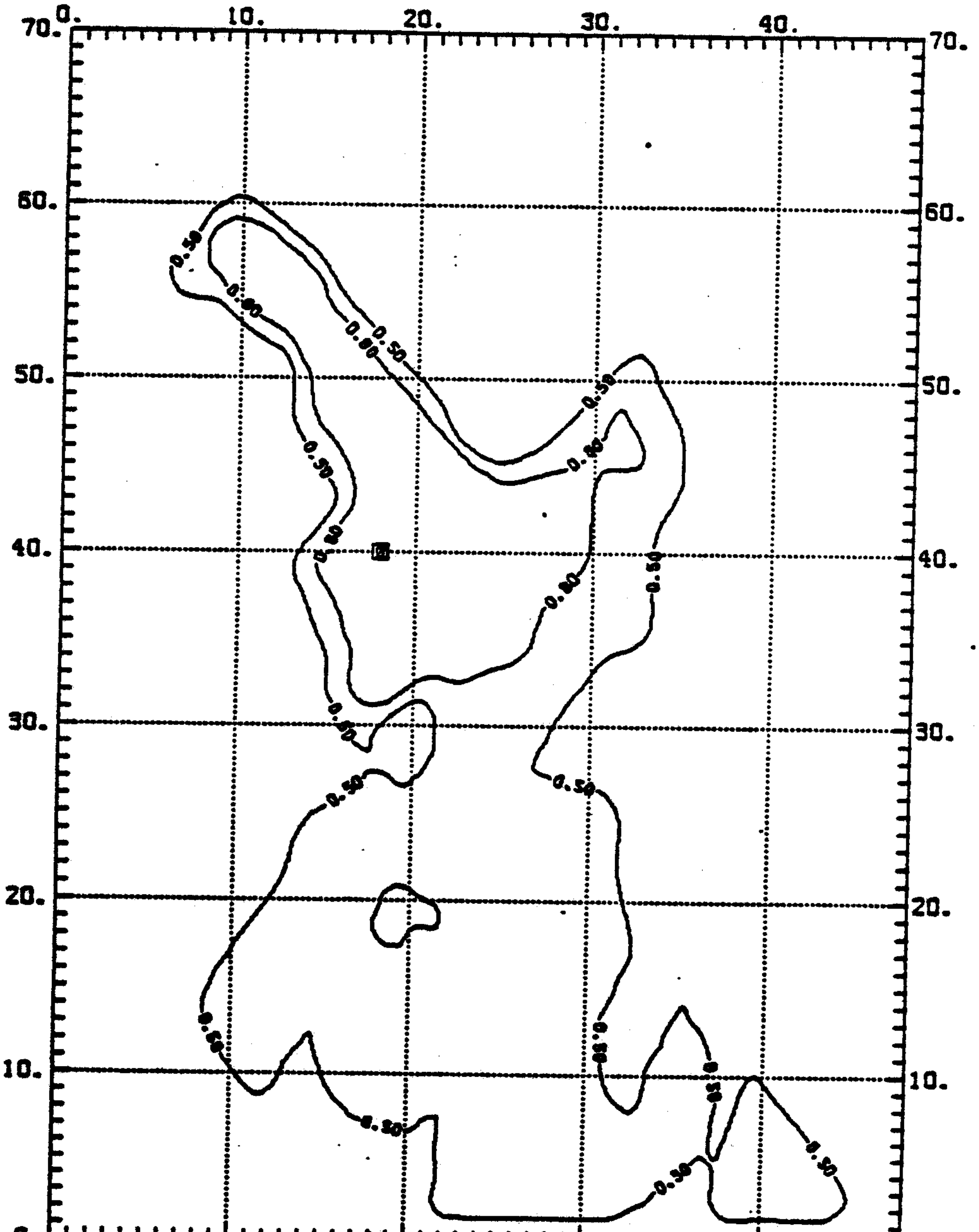
(a) Site M

Figure B-1. (continued)



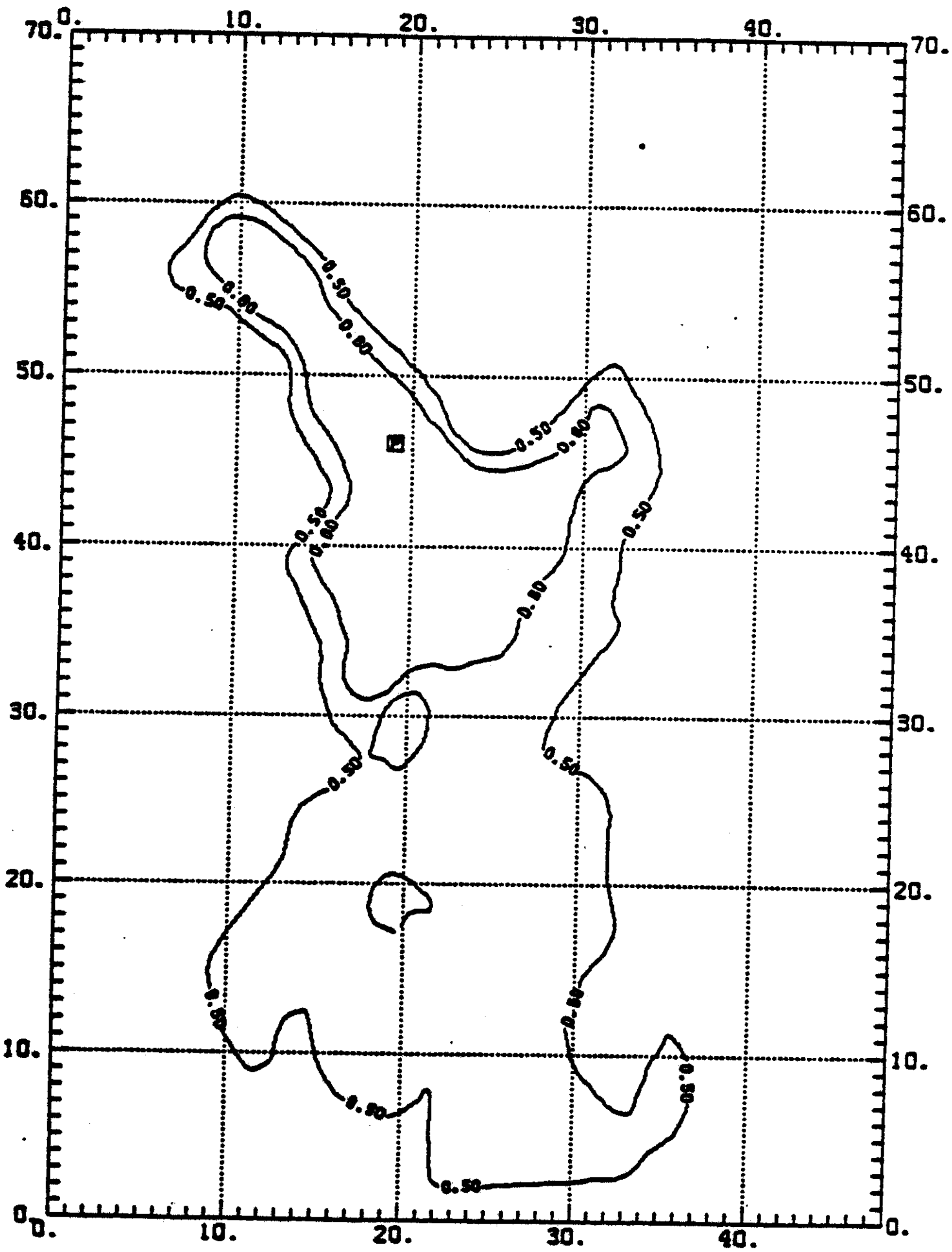
(n) Site N

Figure B-1. (continued)



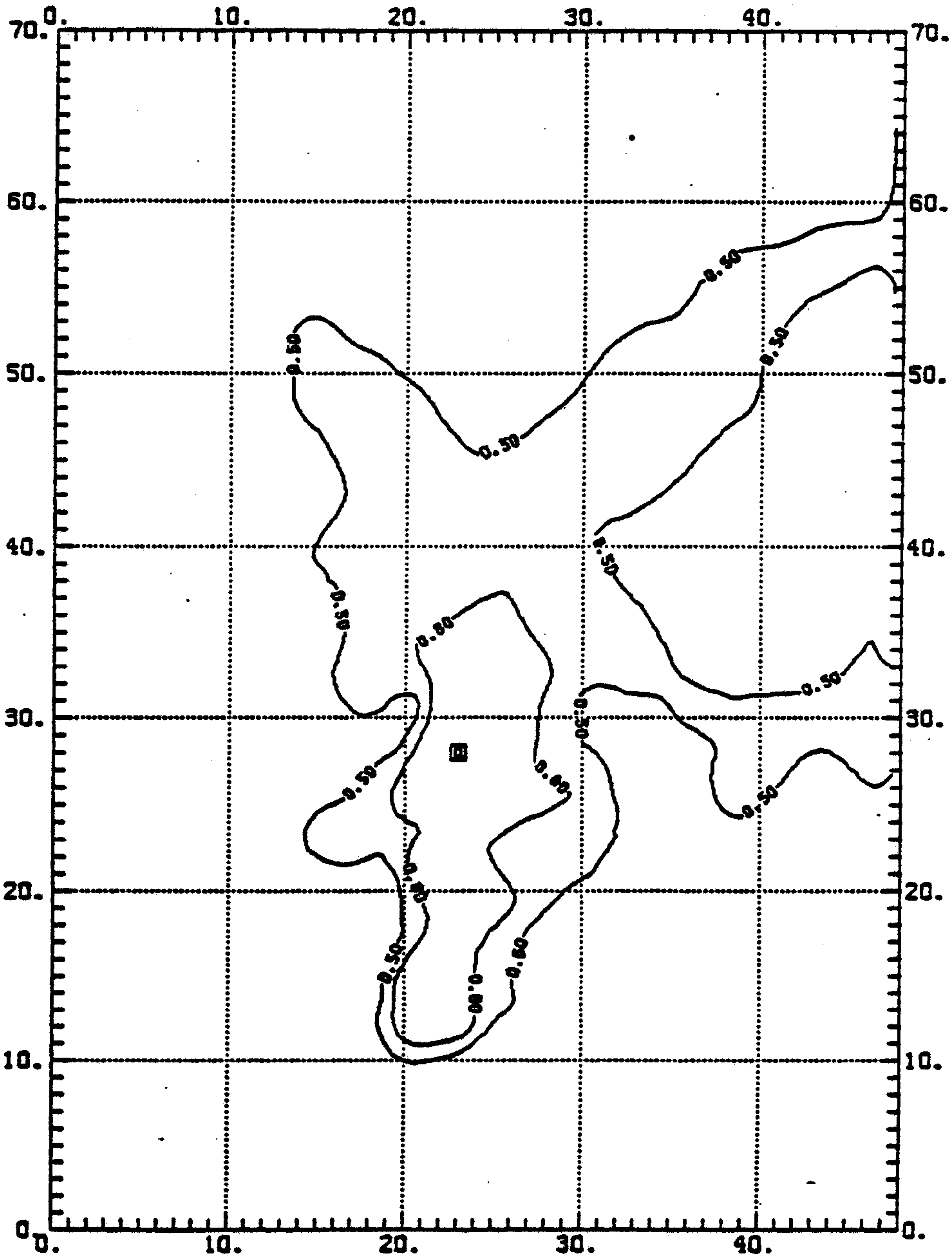
(o) Site 0

Figure B-1. (continued)



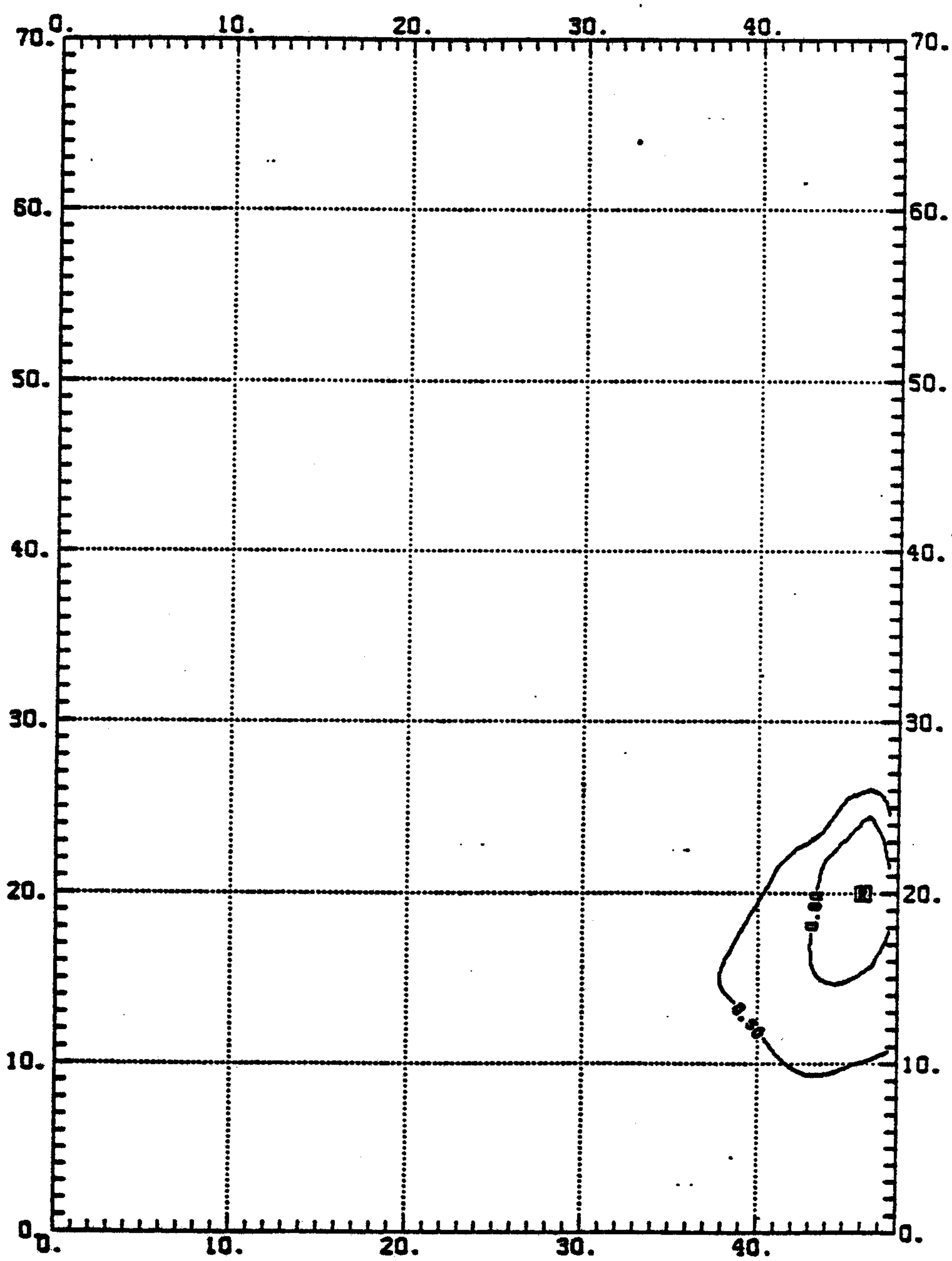
(p) Site P

Figure B-1. (continued)



(q) Site Q

Figure B-1. (continued)



(r) Site R

Figure B-1. (continued)

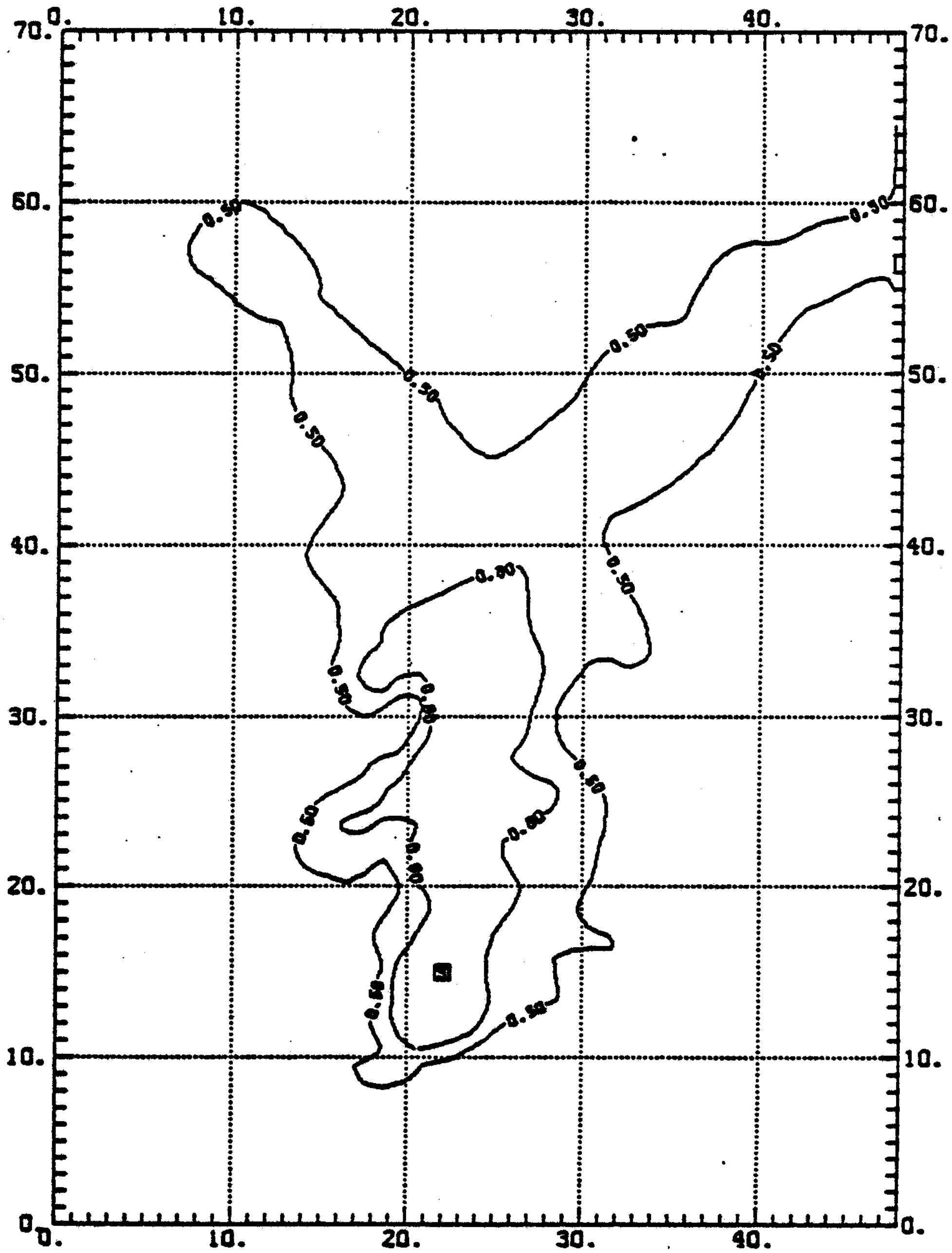
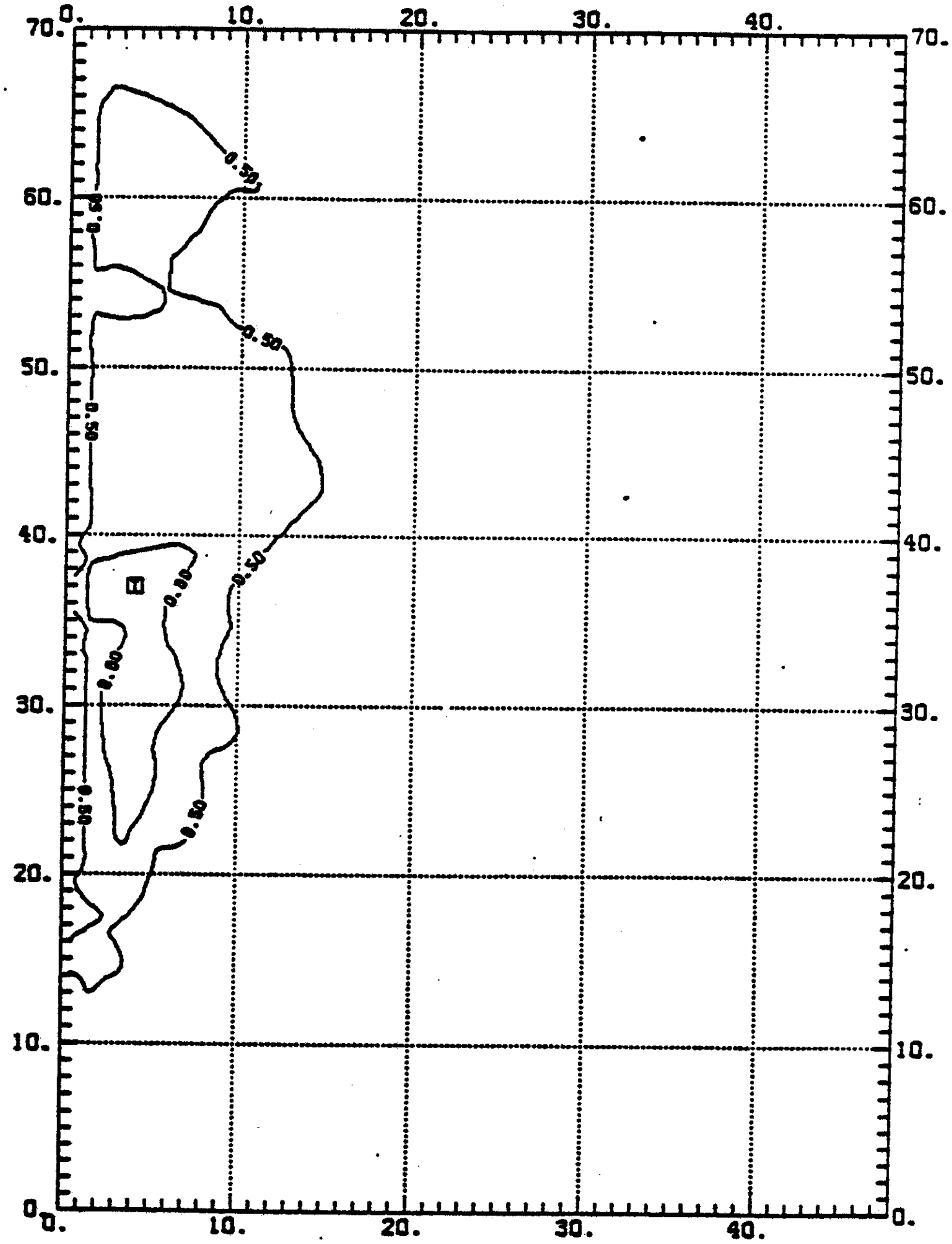
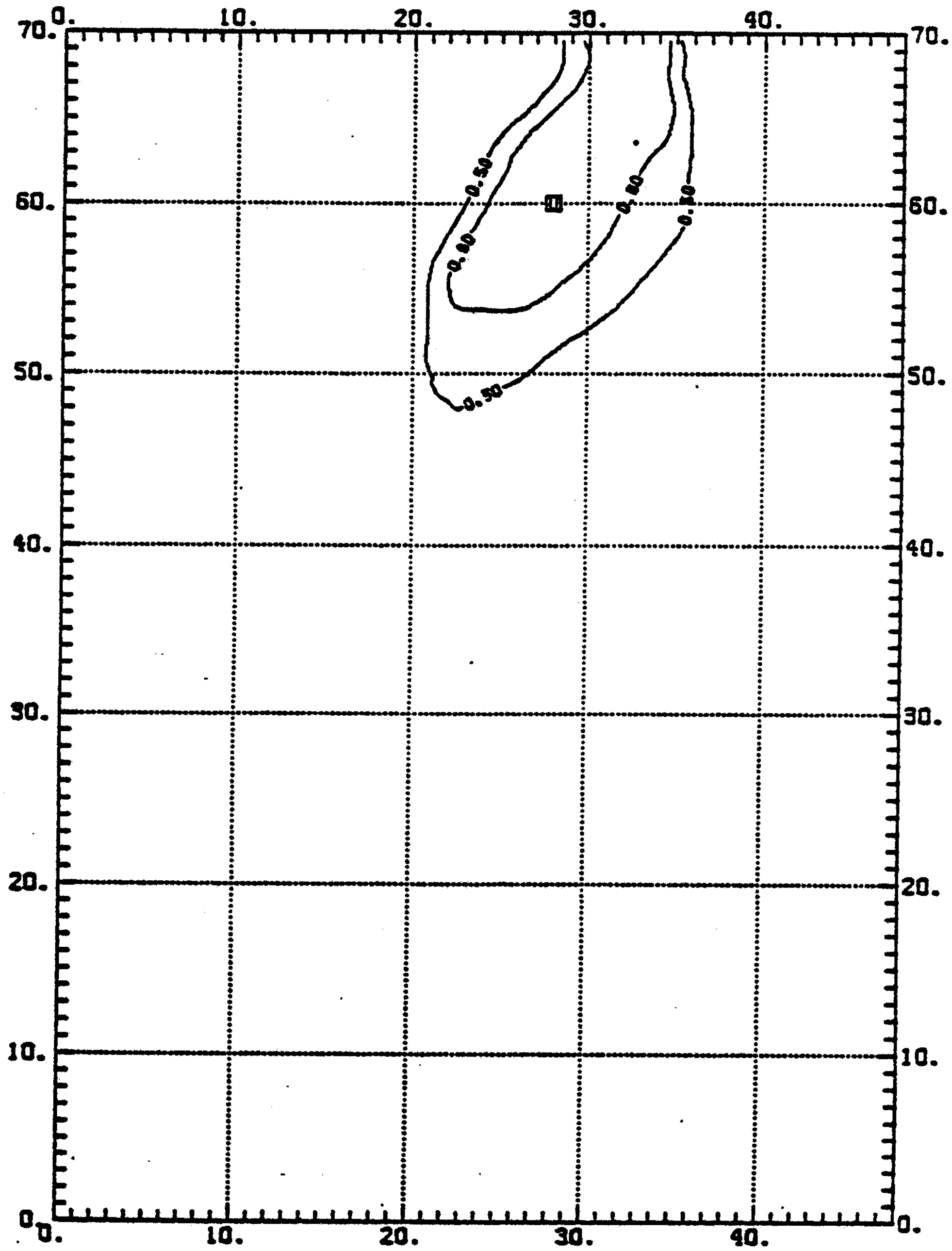


Figure B-1. (continued)



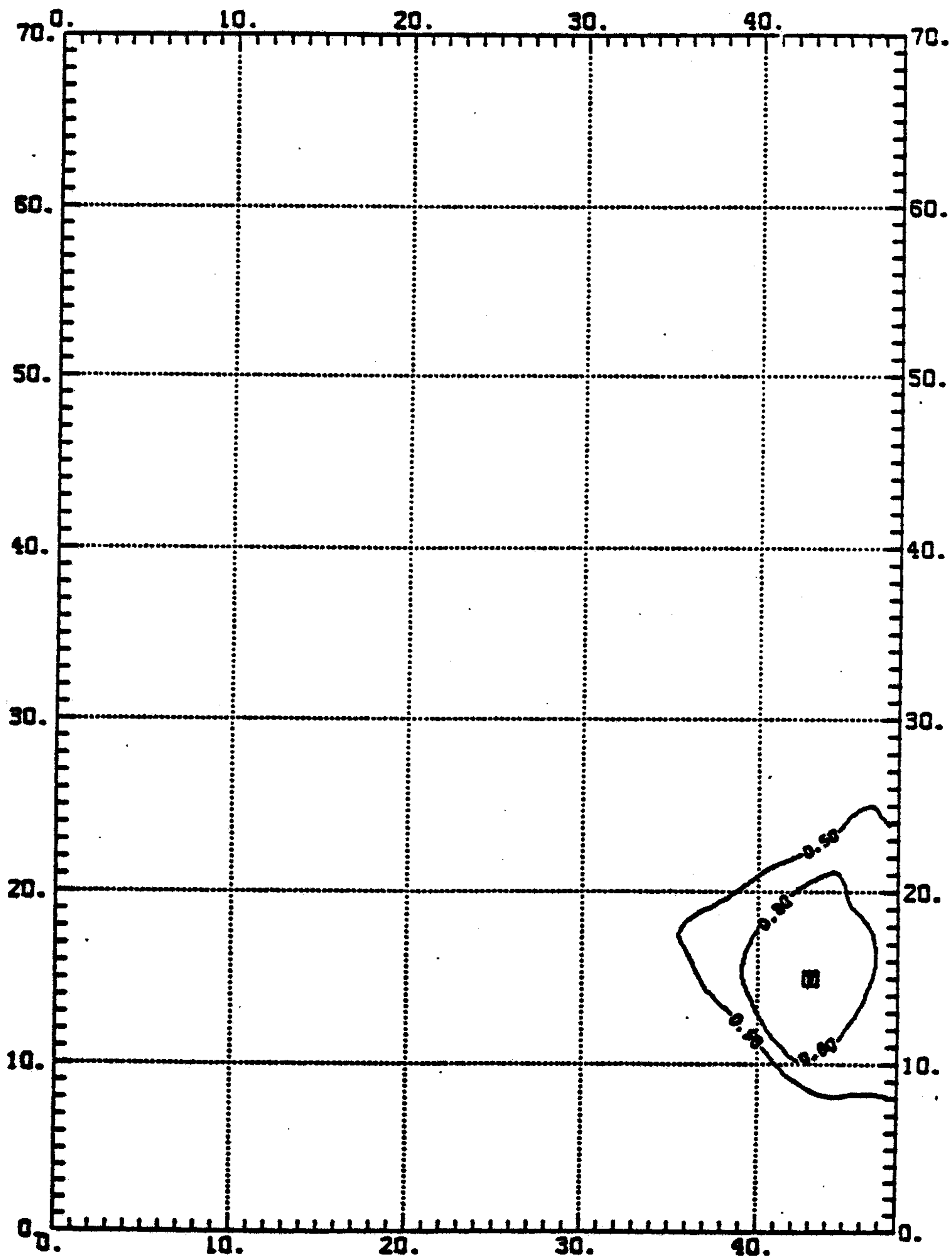
(t) Site T

Figure B-1. (continued)



(u) Site U

Figure B-1. (continued)



(v) Site V

Figure B-1. (continued)

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-600/4-81-002	2.	3. RECIPIENT'S ACCESSION NO.
4. TITLE AND SUBTITLE METHODOLOGY FOR THE DESIGN OF AN OPTIMUM AIR QUALITY MONITORING NETWORK		5. REPORT DATE February 1981 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) M. K. Liu, J. Avrin		8. PERFORMING ORGANIZATION REPORT NO. 158-EF79-146R4
9. PERFORMING ORGANIZATION NAME AND ADDRESS Systems Applications, Incorporated 950 Northgate Drive San Rafael, California 94903		10. PROGRAM ELEMENT NO. 11. CONTRACT/GRANT NO. 68-03-2446
12. SPONSORING AGENCY NAME AND ADDRESS U.S. Environmental Protection Agency--Las Vegas, NV Office of Research and Development Environmental Monitoring Systems Laboratory Las Vegas, Nevada 89114		13. TYPE OF REPORT AND PERIOD COVERED 14. SPONSORING AGENCY CODE EPA/600/07
15. SUPPLEMENTARY NOTES J. L. McElroy, EMSL-LV, Project Officer		

16. ABSTRACT

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.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Meteorology Air quality Air pollution	Design Methodology Monitoring network	55C 68A 91A

18. DISTRIBUTION STATEMENT RELEASE TO PUBLIC	19. SECURITY CLASS *(This Report)* UNCLASSIFIED	21. NO. OF PAGES 60
	20. SECURITY CLASS *(This page)* UNCLASSIFIED	22. PRICE

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE