THE EXPERIMENTAL DESIGN OF AN OBSERVATION NETWORK: SOFTWARE AND EXAMPLES

V. Fedorov, S. Leonov, M. Antonovski, S. Pitovranov

January 1987
WP-87-05

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

Foreword

One important theme within the IIASA Environmental Program is the optimization of monitoring systems. This Working Paper is a contribution to that activity.

Dr. Leonov was an associate research scholar at IIASA from November 1986 through January 1987. His arrival triggered some particularly fruitful discussions on the design of monitoring systems, and led to the development of the system reported in this paper. Although the example used to test the approach is rather elementary from the standpoint of a pollution control agency, the ideas presented provide a springboard for further studies at IIASA and in Dr. Leonov's home institution in Moscow.

Acknowledgement

We are most grateful to Professor R.E. Munn for his constructive criticism and recommendations which were extremely helpful in the preparation of this paper.

THE EXPERIMENTAL DESIGN OF AN OBSERVATIONAL NETWORK: SOFTWARE AND EXAMPLES

V. Fedorov, S. Leonov, M. Antonovski, S. Pitovranov

I. INTRODUCTION

Because of increasing anthropogenic impacts on the biosphere, there has been a rapid development of environmental pollution monitoring systems. During the last 15–20 years many monitoring programs have been initiated on the national (including countries such as Canada, Great Britain, USA, USSR and others) and international levels (including the European Monitoring and Evaluation Program (EMEP) under the auspices of the European Commission for Europe, and the Background Air Pollution Monitoring Network under the auspices of the WMO). Environmental monitoring systems have been defined in several ways. An Intergovernmental Working Group defined monitoring as "a system of continued observation, measurement and evaluation for defined purposes" (International Working Group 1971). Yu. Izrael (1974) has provided the following definition:

- observations of the state of the environment and factors affecting the environment;
- assessment of the state of the environment and impact factors;
- prediction and assessment of the future state of the environment.

According to Izrael, observations are only the first stage of a monitoring system.

Anthropogenic impacts on the biosphere have different spatial and temporal scales. That is why monitoring systems also have to be classified by different spatial levels, here we define the levels as local, regional and global.

Local monitoring systems are usually required in urban areas (city, industrial complex) to detect anthropogenic impacts on human health and the environment. The geographic scale is of the order of several kilometers to one hundred kilometers. The monitoring systems assess environmental quality, which is usually based on such criteria as maximum permissible pollutant concentrations in air, drinking water, etc. Also they provide data for the assessment of economic losses resulting from environmental pollution. Usually, it is possible to define relationships between emissions and concentrations of pollutants on local levels.

Regional monitoring systems serve a similar purpose, but operate over a considerably larger scale (approximately 1000 km). A major problem is that source–receptor pollution relationships are much more complex on the regional level than they are on the local level.

The goal of global monitoring systems is to observe and assess changes in the biosphere on a planetary or hemispheric scale. In this case anthropogenic impacts are observed as averages over the entire world community. Accordingly, source–receptor linkages are even more complicated than at the regional level. The impacts of anthropogenic sources are mediated by processes having the following specific features (Rovinski and Wirsma, 1986). First, such processes are

associated with pollutants having small concentrations which cause large-scale cumulative effects. Second, these processes are always associated with long-range transport of pollutants, mainly in the atmosphere, but sometimes in the hydrosphere. Third, global impacts from anthropogenic activities have long lead times before they are noticed. Therefore, it seems that even in future studies, global monitoring programs will focus on statistical analysis of observation data (see, for example, Izrael, Rovinski, Antonovski et al. 1985), and not on defining source–receptor relationships.

According to the above-mentioned definition of Izrael, the objectives of monitoring include prognoses of environmental states. The tool for such forecasts is mathematical models of different kinds (from partial differential equations to simple regression equations). There are many obstacles to the application of mathematical models in environmental sciences. We mention here only a few of them: the lack of knowledge of processes and physical parameters of ecosystems; the complexity of physical processes in environmental systems; the stochastic character of various processes (for example, atmospheric turbulence); measurements of different environmental parameters with large errors; etc.

The goal of this study is to suggest a procedure which can be useful for optimizing the design of monitoring networks. The approach and the numerical algorithms presented in this paper are to some extent a continuation of that by Fedorov, 1986. Here, however, emphasis will be placed on the technical aspects of the numerical algorithms relating to the construction of optimal observational networks.

The proposed versions of algorithms are oriented to solving problems of the optimal location of observation stations in a given region. Their generalization to other optimal experimental design problems should be straightforward. Any principal changes to be made would be in the block describing "controllable area." As anticipated readers will be from the environmental pollution or meteorological paradigms, all illuminative examples are related to air pollution monitoring.

We wish to emphasize two main assumptions which are crucial to the approach.

1. The optimal design of an observational network is ***model oriented.*** It is assumed that the trends of the observed values can be (at least approximately) described by a mathematical model containing unknown parameters. For instance, it could be a Gaussian plume model where diffusion coefficients, an affective stack height and emission intensity have to be estimated (S. Hanna et al., 1982).
2. All uncertainties (observational errors, fluctuations of processes under investigation, small irregularities, deviations of the model from the "true" behavior, etc.) are absorbed by additive errors, which are assumed to be random.

   Assumptions about the randomness of errors are crucial to the whole approach because all objective functions (both in analysis and design) are formulated as expectations of some deviations of estimators from the "true" values. Most frequently it is the variance of an estimator or the variance-covariance matrix and some functions of it in multidimension cases.

   A summary of 1 and 2 leads us to the following presentation:

$$y_i = \eta(x_i, \vartheta) + \varepsilon_i \ , \tag{1}$$

where y_i is an observed value at an i-th station; x_i are the coordinates of this station; $\eta(x, \vartheta)$ is an *a priori* given function (for instance, it could be a concentration of some pollutant); ε_i is a random value with zero mean ($E(\varepsilon_i) = 0$). Usually, $\eta(x, \vartheta)$ is called the "response function."

The algorithms presented in this paper are oriented to the case where errors of observations are uncorrelated: $E[\varepsilon_i \varepsilon_j] = \delta^2 \lambda^{-1}(x_i)\delta_{ij}$, where $\lambda(x)$ is the so-called "effectiveness function" reflecting the accuracy of observations at the given point x.

It is assumed throughout this paper that y_i is a scalar. The generalization for more complicated situations, for instance y_i either a vector or a function of time, is straightforward (compare with Fedorov, 1972, Ch.5; Mehra and Lainiotis, 1976).

One can apply the method to a vector case when the concentration of several pollutants have to be observed. If the dynamics of some environmental characteristics are of interest then it becomes necessary to consider responses belonging to some functional space.

II. OPTIMALITY CRITERIA

The algorithms presented in this paper comprise two main types of optimality criteria: the first is related to the variance-covariance matrix of estimated parameters, while the second is based on variance characteristics of the response function estimators. Details can be found in Fedorov, 1972; Silvey, 1980; Atkinson & Fedorov (to be published).

Table 1 contains optimality criteria which can be handled with the help of the software described later. The criteria used in the statistical literature are in the first column of the table; formal definitions of optimality criteria are in the second; and the corresponding dual optimization criteria are formulated in the third column (see, for instance, Atkinson & Fedorov (to be published)).

TABLE 1.

Optimality criteria	$\Phi(D)$	$\lambda(x)f^T(x)\dot{\Phi}f(x) - tr\,\dot{\Phi}D = -\varphi(x,\xi)$
D-criterion	$\ln\|D\|$	$\lambda(x)d(x,\xi) - m$,
generalized D-criterion	$\ln\|A^TDA\|$	$\lambda(x)f^T(x)DA[A^TDA]^{-1}A^TDf(x) - s$, $s = \mathrm{rank}A$
A-criterion	$tr\,D$	$\lambda(x)f^T(x)D^2f(x) - tr\,D$,
linear criterion	$tr\,AD$, $A \geq 0$	$\lambda(x)f^T(x)DADf(x) - tr\,AD$,
α-criterion	$\int_Z d(x,\xi)\omega(x)dx$	can be transformed to the previous case with $A = \int_Z f(x)f^T(x)\omega(x)dx$,
extrapolation	$d(x_0,\xi)$	$\lambda(x)\{f^T(x)Df(x_0)\}^2 - d(x_0,\xi)$, $A = f(x_0)f^T(x_0)$.

Theoretically all of the algorithms discussed are valid for the case of linear parametrization:

$$\eta(x,\vartheta) = \vartheta^T f(x)\,, \tag{2}$$

where $f(x)$ is a vector of given functions. How to handle nonlinear models will be considered in Example 3.

The following notations are used in Table 1:

- $D = D(\xi) = ND(\hat{\vartheta})$, where $D = D(\xi)$ is a normalized variance-covariance matrix, $D(\hat{\vartheta})$ is a variance-covariance matrix of the least square estimator $\hat{\vartheta}$, $D^{-1}(\xi) = M(\xi) = \int_X \lambda^{1/2}(x) f(x) f^T(x) \xi(dx)$ or $= \sum_{i=1}^{n} p_i \lambda^{1/2}(x_i) f(x_i) f^T(x_i)$.

- ξ is a design, i.e., $\xi = \{p_i, x_i\}_{i=1}^{n}$, where p_i is a fraction of observations which has to be located at a point x_i; p_i could be the duration, frequency or the precision of observation;

- m is a number of unknown parameters (dimension of ϑ) ;

- $\dot{\Phi} = \partial\Phi / \partial D = \{\partial\Phi / \partial D_{\alpha\beta}\}_{\alpha,\beta=1}^{m}$;

- $d(x,\xi) = f^T(x) D f(x)$ is a normalized variance of the estimator $\eta(x,\hat{\vartheta})$ at a given point x ;

- X is a controllable region, $x_i \in X$;

- A is a utility matrix, usually reflecting the significance of some parameters or their linear combinations;

- $\omega(x)$ is a utility function, usually reflecting the interest of a practitioner in the value of the response function at a point x.

The existence of a nonsingular optimal design is assumed for all optimality criteria in Table 1. Singular optimal designs (i.e. an information matrix $M(\xi^*)$ is singular, $|M(\xi^*)| = 0$, in the regular case $D(\xi) = M^{-1}(\xi)$) can occur when rank $A < m$. In practice one can easily avoid singular designs applying to the regularized version of the initial problem (see Fedorov, 1986, section 2):

$$\Phi_\rho[D(\xi)] = \Phi[\{(1-\rho)M(\xi) + \rho M(\xi_0)\}^{-1}]\,, \tag{3}$$

where $|M(\xi_0)| \neq 0$.

Objective function (3) can also be used in cases where it is necessary to complement existing networks defined by ξ_0 by some new observational stations. D- and A-criteria are usually used when all unknown parameters are equally of interest. The first one is preferable, being invariant to linear transformation of unknown parameters (for instance when one needs to rescale some of them). This assertion can be easily verified with the help of the corresponding dual optimization problems (Table 1, column 3). When some parameters are more important than others, matrix A is usually chosen diagonal with elements $A_{\alpha\alpha}(\alpha = \overline{1,m})$ reflecting the

significance of the corresponding parameters $\vartheta_\alpha (\alpha = \overline{1,m})$.

The last two criteria can be used when an experimenter is interested in the explicit estimation of a response function $\eta(x,\vartheta)$. For instance, if there are points $x_1, x_2, \ldots, x_j$ of special interest, then $\omega(x) = \sum_{k=1}^{j} \delta(x - x_k)$, where $\delta(x - x')$ is δ-function, and $\Phi(D) = \sum_{k=1}^{j} d(x_k, \xi)$.

III. FIRST-ORDER ITERATIVE ALGORITHM

III-1. *The algorithm*

We start with the iterative algorithm of the following form (for details see Fedorov 1986):

$$\xi_{s+1} = (1-\alpha_s)\xi_s + \alpha_s \xi(x_s) , \tag{4}$$

where

ξ_s is a current design on a step s, $\xi_s = \{x_{is}, p_{is}, i = \overline{1,n_s}\}, \sum_{i=1}^{n_s} p_{is} = 1$,

$X_s = \{x_{is}; i = \overline{1,n_s}\}$ is a supporting set of the design;

$\xi(x_s)$ is a design with the measure totally located at a point x_s.

For designs with a finite number of supporting points, formula (4) means that

$p_{i,s+1} = (1-\alpha_s)p_{is}$, $i \neq i^*$, $\quad p_{i^*,s+1} = (1-\alpha_s)p_{i^*s} + \alpha_s$, if $x_s = x_{i,s}$

or

$p_{i,s+1} = (1-\alpha_s)p_{i,s}$, $i = \overline{1,n_s}$; $\quad p_{n_s+1,\,s+1} = \alpha_s$, if $x_s \neq x_{i,s}$.

The algorithm provides so-called forward and backward procedures. In the backward procedure, the "least informative" points are deleted from the current design, while conversely the forward procedure includes the new, "most informative" ones.

III-2. *Selection of* $\{x_s\}$ *and* $\{\alpha_s\}$.

For the forward procedure:

$$x_s = x_s^+ = \operatorname*{Argmin}_{x \in X} \varphi(x, \xi_s) ,$$

$$\alpha_s = \gamma_s .$$

For the backward procedure:

$$x_s = x_s^- = \operatorname{Arg}\max_{x \in X_s} \varphi(x, \xi_s) ,$$

$$\alpha_s = \begin{cases} -\gamma_s, & p_j^* \geq \gamma_s \\ -p_s^* / (1-p_s^*), & p_s^* < \gamma_s, \end{cases}$$

$p_s^* = p(x_s^-)$ is a weight for a point x_s^-.

The algorithm provides three choices of gain sequence $\{\gamma_s\}$:

(a) $\gamma_s = \dfrac{1}{n_0+s}$, $s=1,2,\ldots$; n_0 is a number of supporting points in an initial design. With this choice of γ_s, one can simulate the subsequent inclusion (deletion) of the most (least) informative stations.

(b) γ_s is defined by the steepest descent method, which provides the largest decrease of the objective functions in the chosen direction $\xi(x)$.

(c) $\gamma_s \equiv C_0$, where C_0 is a small constant (0.01 $\div$ 0.1) which is defined by a user. This sequence does not satisfy traditional conditions

$$\lim_{s \to \infty} \gamma_s = 0, \quad \sum_s \gamma_s = \infty, \quad \sum_s \gamma_s^2 < \infty ,$$

which are usually implied to prove the convergence of the iterative algorithms, but may be useful for the construction of the discrete designs.

Numbers of steps (length of excursion) for the forward and backward procedures (*nfor* and *nbac* respectively) are defined by the user.

III-3. *D-criterion.*

The algorithm "DOPT" is oriented for the construction of D-optimal designs providing the minimum of the determinant $|D(\vartheta)|$, $D(\vartheta)$ is a covariance matrix of the parameters' estimators. Geometrically the minimization of the determinant is equivalent to the minimization of the volume of the ellipsoid of concentration for the parameters' estimators. Simultaneously the algorithm minimizes $\sup_{x \in X} \lambda(x) d(x, \xi)$ (see Table 1) securing an effective estimation of the response function over set X. Moreover, in the case of normally distributed errors ε_i D-optimal design ensures the best value of the noncentrality parameter when the hypothesis

$$\sup_{x \in X} \eta^2(x, \vartheta_t) \geq \delta , \quad \delta > 0 ,$$

is tested (see Fedorov, 1986).

A function $\varphi(x, \xi_s)$ has the following presentation for the D-criterion (see Table 1):

$$-\varphi(x, \xi_s) = \lambda(x) d(x, \xi) - m ,$$

where $d(x,\xi_s) = f^T(x)D^{-1}(\xi_s)f(x)$, see also page 4.

The formulae for sequential recomputation of the covariance matrix and the determinant are

$$D(\xi_{s+1}) = (1-\alpha_s)^{-1}\left[D(\xi_s) - \frac{\alpha_s D(\xi_s)f(x)f^T(x)D(\xi_s)}{1-\alpha_s+\alpha_s d(x,\xi_s)}\right], \tag{5}$$

$$|D(\xi_{s+1})| = (1-\alpha_s)^{1-m}[1-\alpha_s+\alpha_s d(x,\xi_s)]^{-1}\cdot|D(\xi_s)| . \tag{6}$$

III-4. *Some notes on the algorithm.*

Stopping rule

The calculations are terminated if

(a) the convergence criterion is attained for the forward procedure: $\frac{|\varphi(x_s^+)|}{m} < \delta$, where δ is defined by a user (this means that the value of the directional derivative is small enough and, subsequently, ξ_s is close enough to the optimal design).

(b) a given number of iterations is attained.

Merging of supporting points in the forward procedure

Let h_k be a size of the k-th grid element defined during the mapping of X, $k = \overline{1,L}$; L is a dimension of controllable region X. If

$$|x_{i,k} - x_{s,k}^+| < C_{\text{mer}}\, h_k;\ x_i \in X_s\ ,\quad k=\overline{1,L}\ ,$$

then a point x_i is merged with a point x_s^+, constant C_{mer} being defined by a user.

Deleting of points with small weights in the forward procedure.

If for some i, $p_{i,s} < \delta$, then a point $x_{i,s}$ is deleted from the design. The covariance matrix and criterion value are recomputed, formulae (5), (6) being used with $\alpha = -p_{i,s}$ and

$$p_{j,s+1} = p_{j,s}/(1-p_{i,s})\ ,\quad j \neq i\ .$$

Both latter procedures help to avoid designs with a large number of supporting points.

IV. OPTIMIZATION ALGORITHM OF THE EXCHANGE TYPE

The algorithm has the form

$$\xi_{s+1} = \xi_s + \alpha_s \xi(x_s) \tag{7}$$

where α_s can be either positive or negative.

From a computational point of view, the main difference in algorithm (7) from the one described in Section 3 is that the whole design is not recomputed at each step; all modifications concern only newly included ($\alpha_s > 0$) or deleted ($\alpha_s < 0$) points, which explains the origin of the term "exchange" in the title of the algorithm (see also Fedorov, 1986). The various modifications of the "exchange type" algorithm are particularly useful when some subset of an initial design has to be included in the final design (some prescribed observational stations have to be included in the final observational network). The algorithm can be easily adapted to solve the regularized versions of the originally singular design problems conserving some "regular" fractions of an initial design.

The presented software contains three modifications of the exchange procedure.

(1) *Deleting the least informative points from the initial design.*

The backward procedure is executed (some points are deleted) with $\alpha_s = -1/n_0$, n_0 is the number of points in the initial design

$$x_s = x_s^- = \operatorname{Arg}\max_{x \in X_s} \varphi(x, \xi_s) .$$

A number of steps for deleting (*miter = nbac*) is chosen by a user. All points in the final design (normalized to: $\Sigma p_i = 1$) have equal weights:

$$p_i = 1/(n_0 - nbac), \quad nbac < n_0 .$$

This procedure can be used, for instance, when it is necessary to find and remove a given number of the least informative stations (see example 1(b).

(2) *Inclusion of the most informative points.*

The forward procedure is executed with

$$\alpha_s = 1/n_0 ,$$

$$x_s = x_s^+ = \operatorname{Argmin}_{x \in X} \varphi(x, \xi_s);$$

a number of steps for including (*miter = nfor*) is chosen by a user; final weights are equal to

$$p_i = 1/(n_0 + nfor).$$

For both of the above procedures, the normalization of the covariance matrix is carried out during the last step, as otherwise $\sum_i p_{i,s} < 1$ for the backward procedure, and $\sum_i p_{i,s} > 1$ for the forward procedure.

Normalization is not executed during the intermediate steps in order to make tangible either the decrease of the determinant $|D(\xi_s)|$ due to the deletion of the observational stations or its increase due to the inclusion of stations.

(3) *Standard exchange procedure.*

Forward and backward procedures are executed subsequentially, the initial procedure being chosen by a user. Here, the number of steps for the forward and backward procedures are equal:

$nfor = nbac = nn$,

and the maximal number of iterations has the form

$miter = 2nn \cdot k_0$; k_0 is an integer.

A measure ξ is a probability measure at the end of the "large iteration", the length of which is equal to $2nn$; this fact explains the choice of a value $miter$.

The choice of $\{x_s\}$ is as described above,

$$\alpha_s = \begin{cases} \gamma_s, & \text{forward procedure} \\ -min(\gamma_s, p_s^*), & \text{backward procedure} \end{cases} \qquad (8)$$

There are two variants for the choice of gain sequence $\{\gamma_s\}$:

(a) $\gamma_s = \dfrac{1}{n_0+1+l}$, $s=1,2,\ldots$; l is an integer part of $(s-1)/2nn$;

γ_s changes after executing both forward and backward procedures, i.e., it is a "large iteration";

(b) $\gamma_s = C_0$, C_0 is a constant defined by the user.

The popular Mitchell algorithm (Mitchell, 1974) can be considered as a particular case of this version. It is well known that the Mitchell algorithm does not generally converge to an optimal solution.

Some notes on the algorithm.

Stopping rule

(a) The convergence criterion is attained at the last step of the forward procedure

$$\frac{|\varphi(x_s^+)|}{m} < \delta, \quad \text{or}$$

(b) the maximal number of iterations is attained.

Merging of points for forward procedure

This is the same as described in Section III.

Fixing of initial points.

Some points in an initial design may be fixed by a user (locations and weights).

Ensuring that the weights are positive.

If for the backward procedure $p_s^* < \gamma_s$ for some s (see (8)), then for the subsequent forward procedure

$$\alpha_{s'} = p_{s'}^*, \text{ for some s' .}$$

This modification is necessary to keep ξ_s in the set of probability measures at the end of the "large iteration".

V. LINEAR OPTIMALITY CRITERIA

Algorithms LINOPT and LINEX are intended for the construction of linear optimal designs providing minima of the value

$$tr\ AD(\xi_s)\ ,$$

where A is a utility matrix (it is a symmetric nonnegative definite $(m \times m)$ matrix) chosen by the user according to his needs.

The major difference in the algorithms LINOPT (first-order iterative algorithm) and LINEX (optimization algorithm of the exchange type) from DOPT and DOPTEX respectively, is that the function $\varphi(x,\xi_s)$ has the following presentation:

$$-\varphi(x,\xi_s) = \lambda(x)\, f^T(x) D(\xi_s) AD(\xi_s) f(x) - tr\ AD(\xi_s)\ .$$

VI. USER'S GUIDE

1. *Mapping of a controllable region X.* Program MAP is intended for mapping of a controllable region X. The current version of the program handles one- and two-dimensional regions but generalization to larger dimensions is not difficult. The region X is defined on a uniform grid with given densities for each variable. Such a presentation of X is explained by the fact that usually a user deals with irregular regions which cannot be described analytically (non-convex, with subregions where the location of observational stations is impossible, for example, lakes, densely populated areas, etc.). Two output files are created by the program:

"REG.DAT" contains the data in its original scale,

"SCALE.DAT" contains the normalized data $(-1 \le x_i^{(nor)} \le +1,\ i=1,L)$.

2. Programs DOPT, DOPTEX, LINOPT, LINEX, utilize three files:

"OUT.DAT" is for output information (see examples),

"REG.DAT" contains the design's grid,

"DES.DAT" contains an initial design.

3. The structure of a vector of basic functions $f(x)$ must be set in the subroutine *resp*:

$$resp\ (m, f, L, x),$$

where m is the number of unknown parameters,

$$f = (f_1(x), \ldots, f_m(x))^T,\quad x = (x_1, \ldots, x_L)^T\ .$$

If the effectiveness function $\lambda(x)$ is not constant, then instead of $f_\alpha(x)$ the function $\lambda^{1/2}(x) f_\alpha(x)$ has to be programmed, $\alpha = \overline{1,m}$.

4. All auxiliary subroutines (matrix inversion, calculation of the initial determinant, minimization of a function $\varphi(x,\xi_s)$ etc.) for programs DOPT,DOPTEX and LINOPT,LINEX are saved in the files "SUBD.F" and "SUBL.F", respectively.

VII. Examples

Example 1. Linear parametrization, D-criterion.

To illustrate the possibilities of the proposed software, let us consider a comparatively simple example based on air pollution data from Modak and Lohani, 1985. The particular example we shall use is shown in Figure 1, which gives isopleths of monthly mean values of SO_2 concentration for 9am in Taipei City, Taiwan. The original network contains eleven observing stations (see Figure 2). To begin, the underlying model was chosen as a polynomial of the second degree with uncorrelated random additive errors:

$$y_i = \vartheta_1+\vartheta_2 x_{1i} + \vartheta_3 x_{1i}^2+\vartheta_4 x_{2i}+\vartheta_5 x_{2i}^2+\vartheta_6 x_{1i} x_{2i}+\varepsilon_i \ , \qquad (9)$$

where (x_{1i}, x_{2i}) are coordinates of the i-th station. Of course, this model is too simple for a good approximation of the pattern presented in Figure 1, but because of its simplicity one can easily understand the main features of the software.

The optimality criterion was taken equal to the normalized determinant of variance-covariance matrix (*D*-criterion).

(a) *Completely new network.* The purpose of this algorithm is to find the "best observation" network under the assumption that there are no constraints on the number of stations and their locations except that the stations have to be within the city's area.

The ratio of determinants for the original and optimal locations is greater than 10^4 (see Printout 1). One can observe (Figure 3) a typical (for the conventional optimal design) location of observation stations: most of them have to be on the boundary of the area and only a few (in our case only one) inside it. This should be compared with the result by Modak and Lohani, 1985, p.14, based on the so-called "minimum spanning tree" algorithm, where observing stations are mainly located inside the area.

However, a comparison of results is conditional since the authors did not report the model used for the monthly averaged concentration of SO_2.

For illustrative reasons both DOPT and DOPTEX programs were used to construct the optimal allocation of observation stations and naturally they led to the same (up to computational accuracy) results.

The optimal network consists of seven stations (the model contains six unknown parameters). Usually the number of observing stations is equal to the number of unknown parameters. The seventh point appears here due to some peculiarities in the controllable region.

One can see that the variances of all parameters (except the first one whose variance does not depend upon the allocation of stations) are reduced 10-20 times.

Theoretically the optimal design assumes that the accuracy of observations at the various points is different. Sometimes this demand is not realistic in practice but it is easy to verify theoretically that the design characteristics are quite stable under variation of weights (see Fedorov and Uspensky, 1975, p.56). The calculations confirm this fact for our example. For instance, from the optimal design (see Printout 1), point 1 with small weight (~0.054) was removed from the design and for all others the weights were chosen equal 1/6 (so called saturated design: number of observation = number of unknown parameters). The ratio of the determinants of the variance-covariance matrix for the newly constructed design and *D*-optimal designs was found to be equal to ~1,2. In terms of variances, the discrepancy ($\sim\sqrt[6]{1,2}$) is negligible.

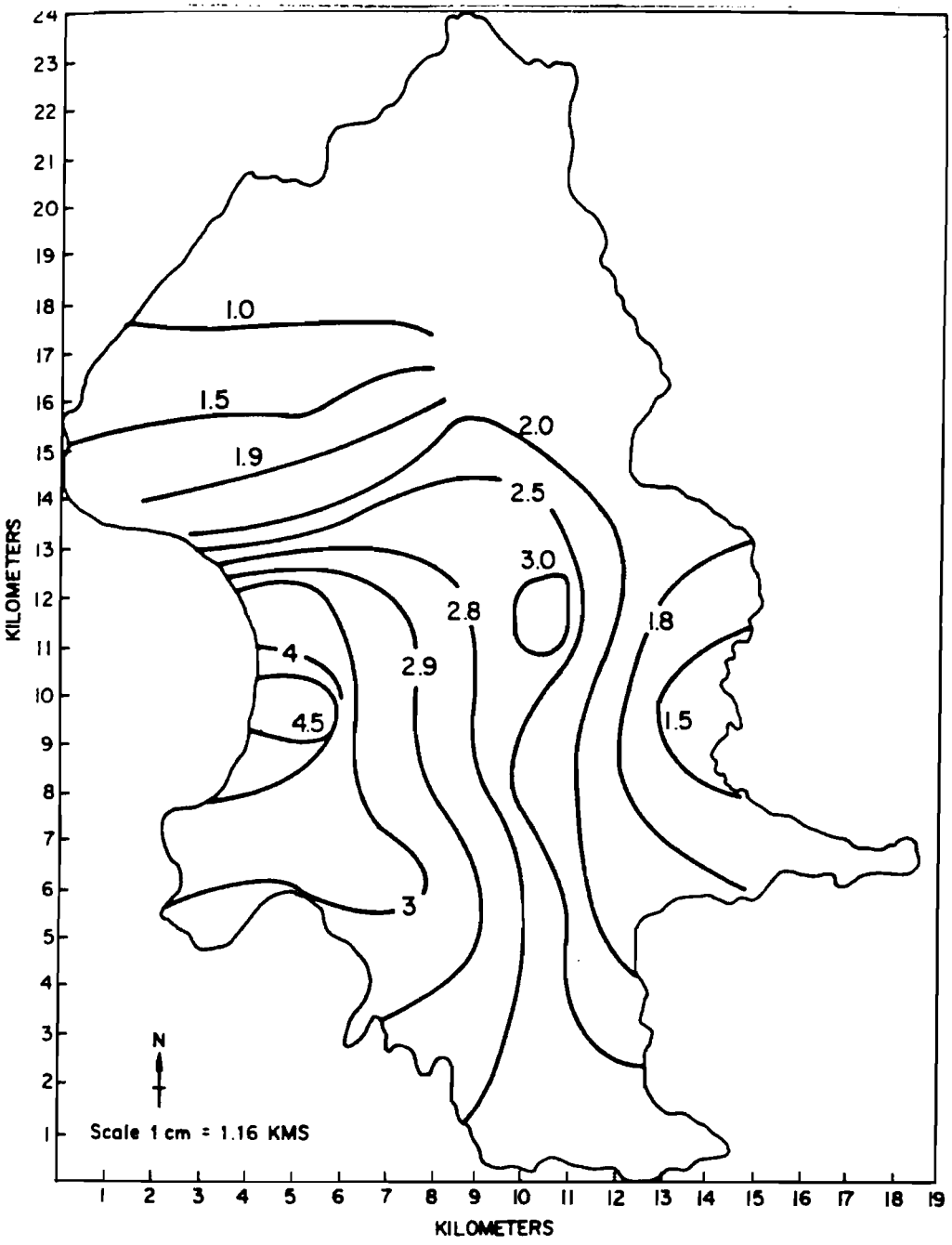

Figure 1. Monthly average (January, 1981) of SO_2 (in 0.1 ppm) at 9am for Taipei City.

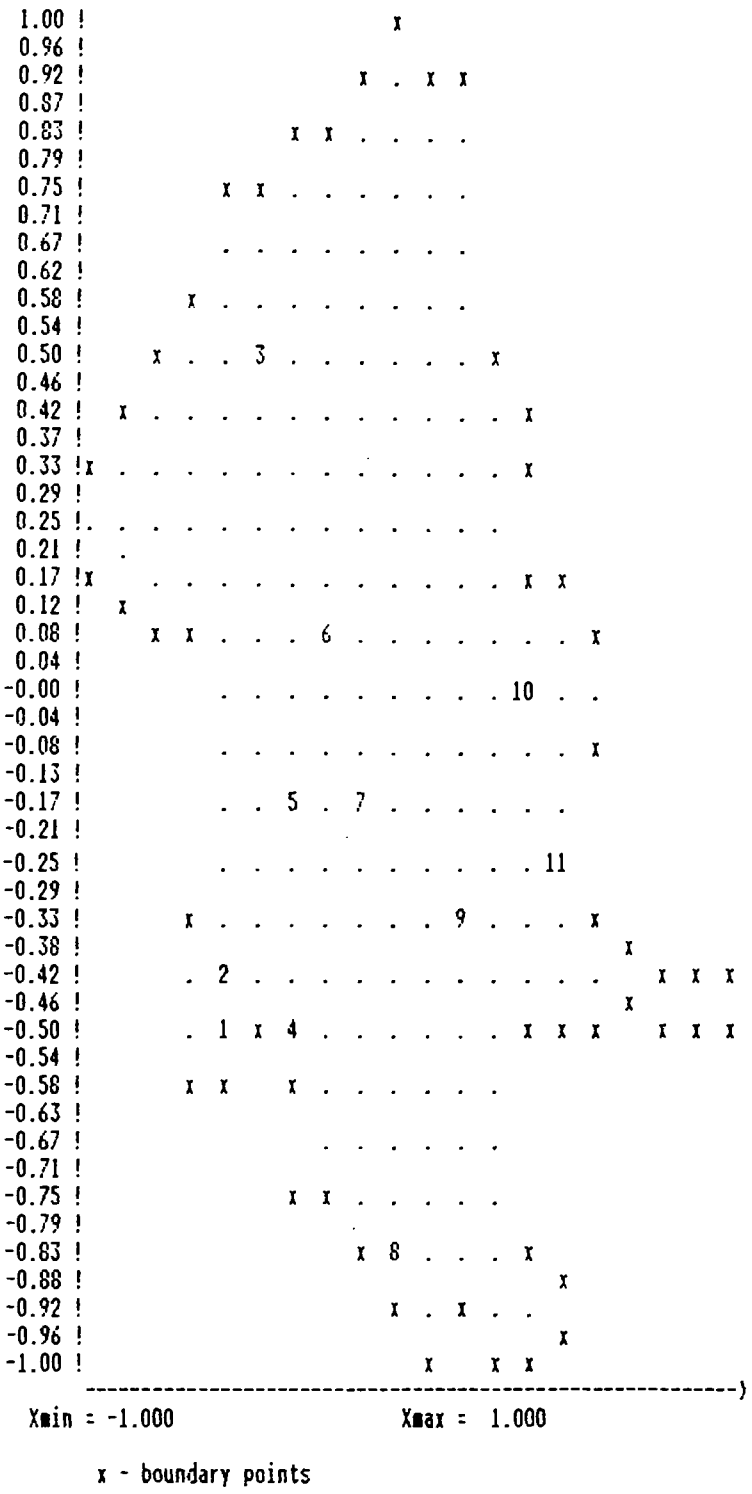

Figure 2. Configuration of the existing observation network.

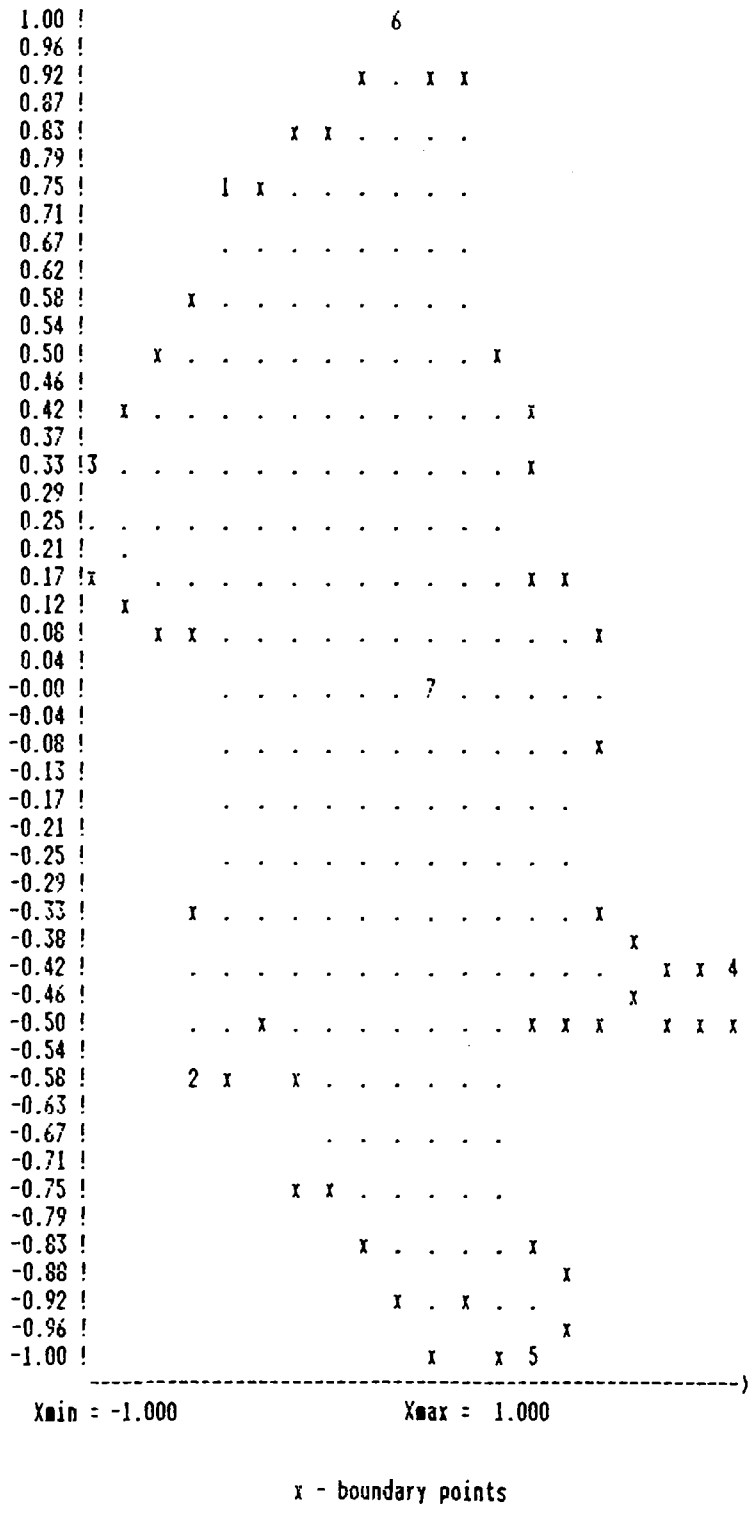

Figure 3. *D*-optimal observation network.

(b) Removal of the "least informative" stations. In practice one may need to reduce the number of observation stations avoiding any relocations. In this case the deletion procedure has to be used; it can be executed by either program. Here, DOPTEX is preferable from the computational point of view.

To be specific, the necessity of removing four stations was assumed. Stations 4, 5, 6, 9 (compare Figures 2 and 4) were subsequently deleted. The non-normalized determinant of the variance-covariance matrix decreased 5.5 times but the normalized matrix increased 2.7 times. In other words, the decrease in the non-normalized determinant confirms that any deletion of data (assuming their correctness) decreases the final information but the increase in normalized determinant indicates that the deletion was done in the sense of the chosen criterion optimally: the "effectiveness" of the rest of the observation stations significantly increased (see printout 2).

(c) Addition of new stations. The "inclusion" algorithm was used to find the location of three new stations - they appeared on the boundary of the region. The non-normalized determinant increased $2.3 \cdot 10^3$ times and the normalized one increased 5.10^2 times, i.e., in this case both the total information and its effectiveness increased (see printout 3 and Figure 5).

(d) *Optimal observation network containing some stations with fixed positions.* When creating a new observation network, one can face the necessity of including in it some N_0 (for instance, well equipped) existing stations. If the total number N of stations is given, then one has to consider the following design problem

$$\xi_0^* = \operatorname{Arg} \min_{\xi} \Phi[(1-N_0/N)\xi + (N_0/N)\xi_0] , \qquad (10)$$

where ξ_0 describes the location and accuracy of an existing station required to be in the planned network. The solution of (10) can be computed with the help of programs DOPTEX and LINEX.

The results of the calculations for D-criterion and model (9) are presented in Printout 4 and Figure 6.

Example 2. Linear parametrization, A-criterion. Theoretically the optimal location of observational stations depends upon the chosen criterion of optimality. In practice the dependence is usually negligible. To confirm this fact, let us consider the A-criterion when the quality of a location is characterized by the average variance of the parameter estimators: $\Phi = m^{-1} \sum_{\alpha=1}^{n} D_{\alpha\alpha} = m^{-1} tr\, D$. The results of the calculation (program LINOPT) for model (9) are presented in Printout 5. The allocation of all observation stations coincides (up to computation accuracy) with that for the D-optimal allocation, see Figure 3. The major traceable difference is in the "weights": the points which are closer to the origin have the greater weights (i.e. the accuracy (or number of repetitions) of observations has to be greater for the "central points").

Example 3. Nonlinear parametrization, D-criterion.

All previous considerations were based on equation (9) which cannot describe all details of the patterns presented in Figure 1. Unfortunately Modak and Lohani, 1985, did not report the model used to construct isopleths drawn on their figure. Therefore, we applied the following nonlinear response function trying to approximate those isopleths:

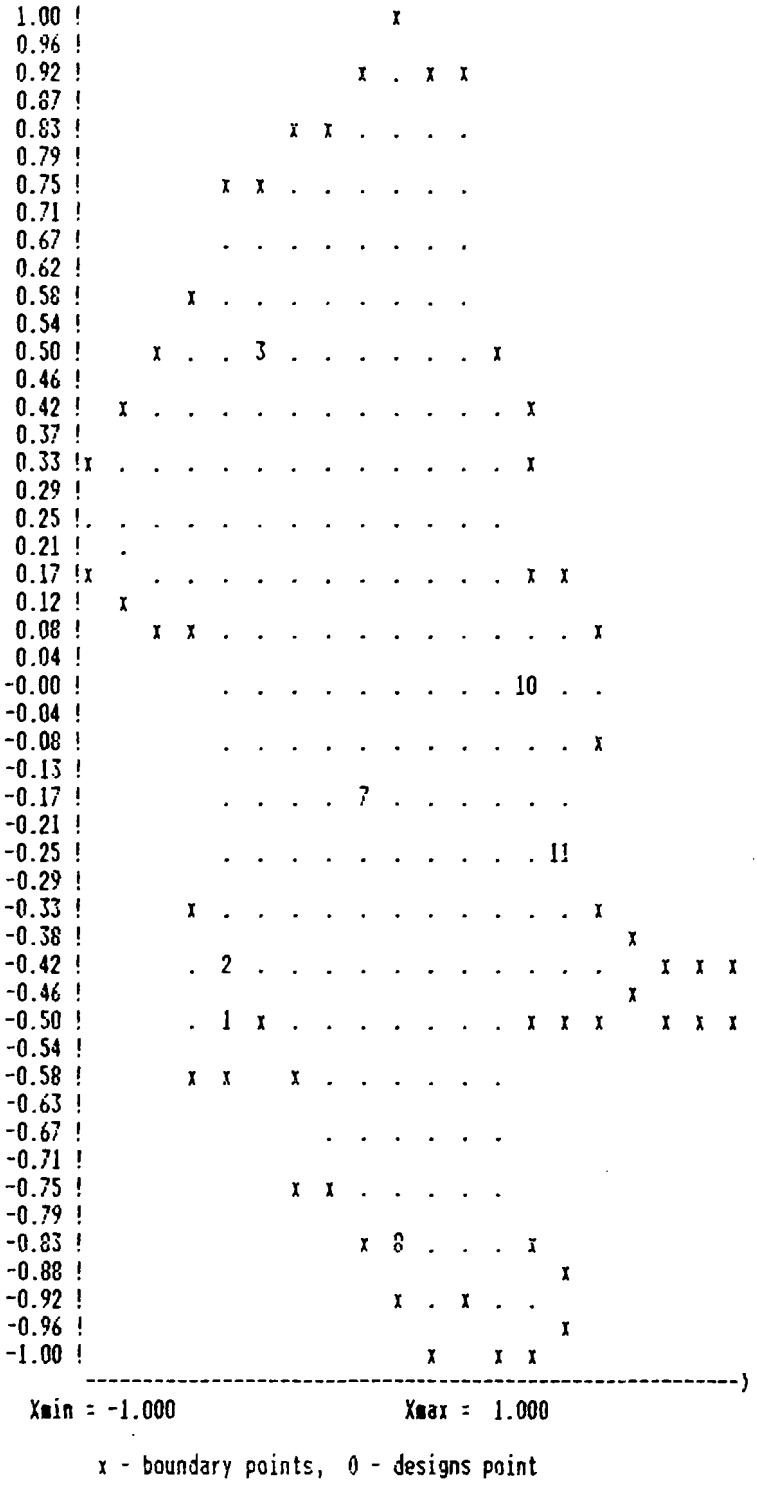

Figure 4: The observation network after deletion of four stations.

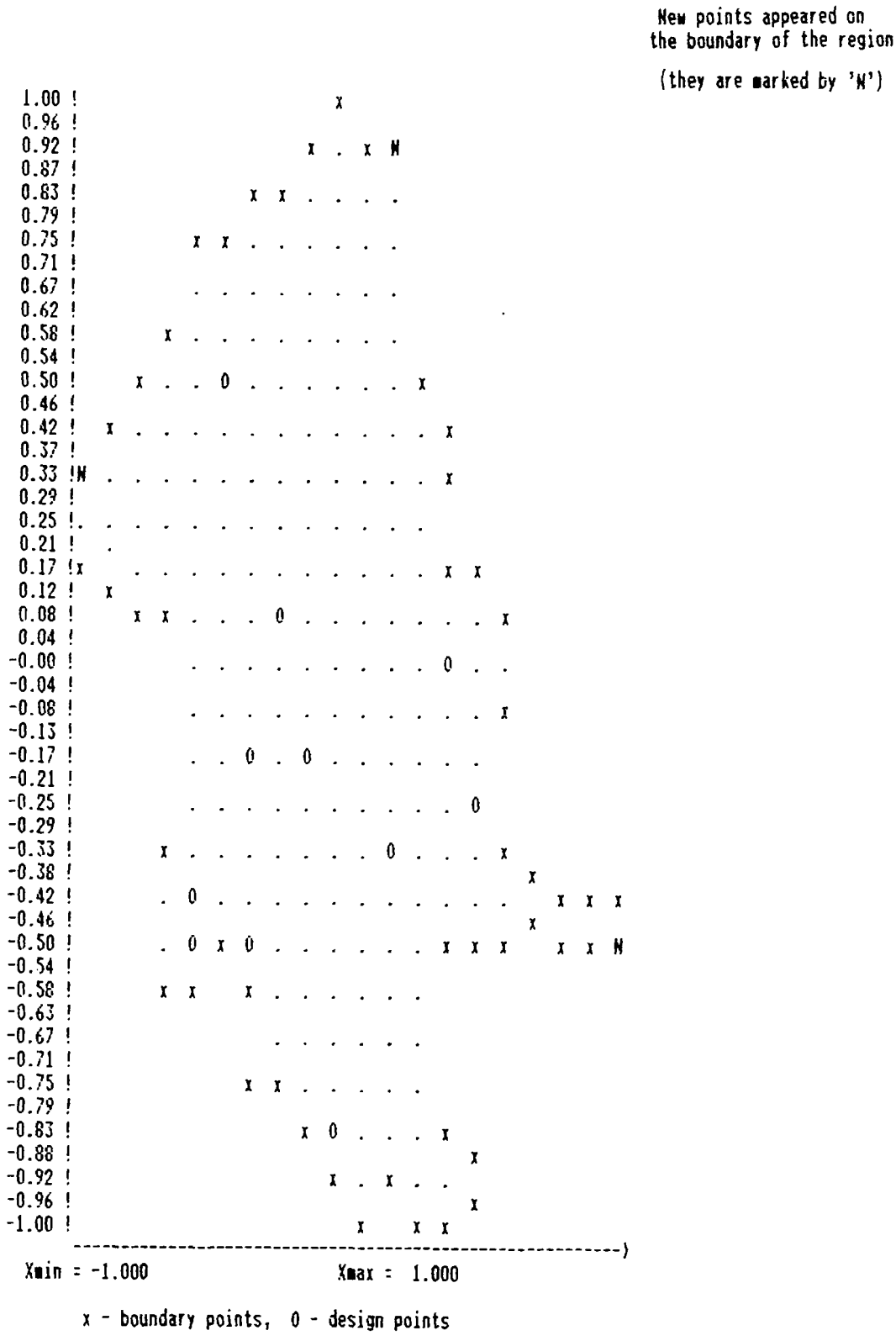

Figure 5: The observation network with three new stations.

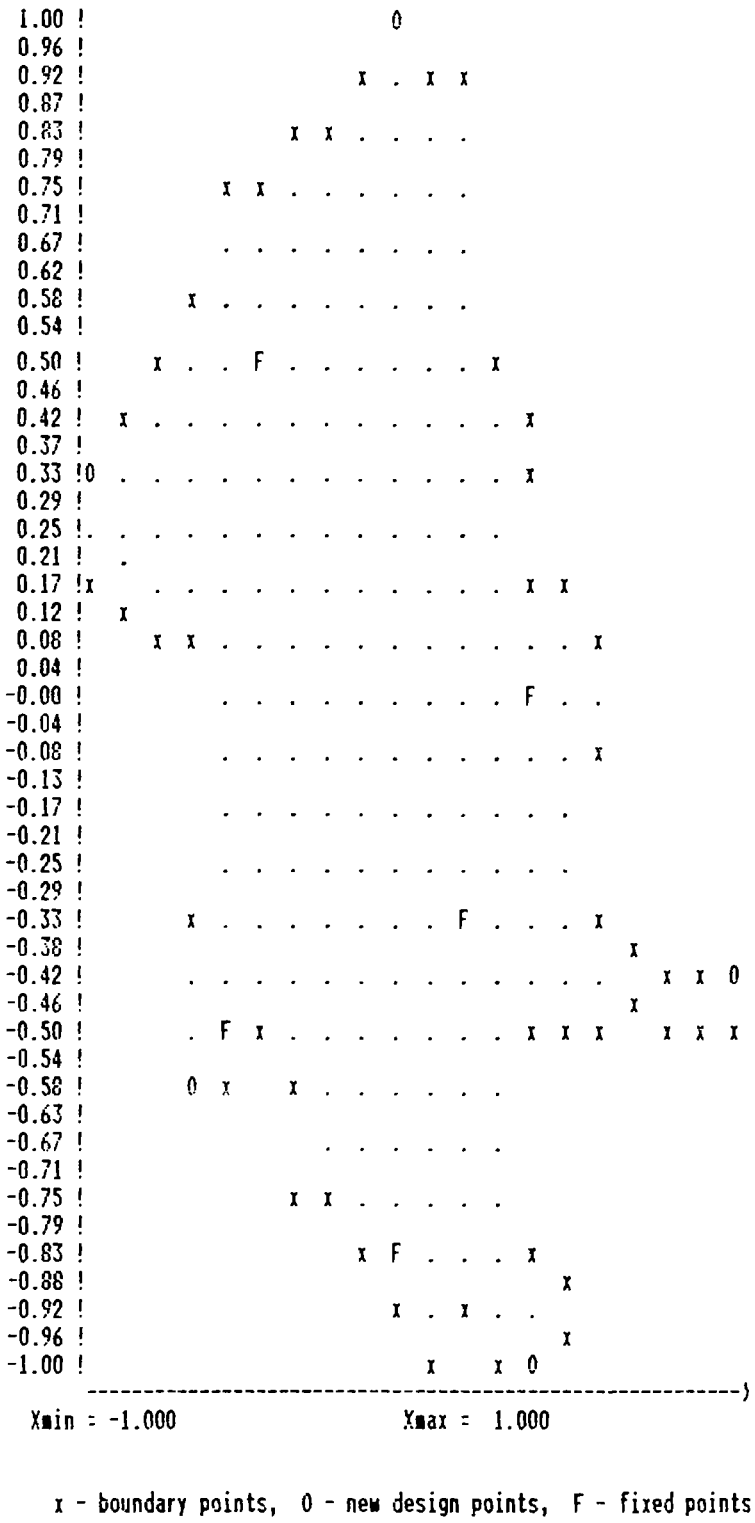

Figure 6: *D*-optimal observation network with five fixed stations.

$$\eta(x,\vartheta) = \sum_{j=1}^{2} \vartheta_{1j} \exp\{-1/2[\vartheta_{2j}(x_1-a_{1j})^2 + 2\vartheta_{3j}(x_1-a_{1j})(x_2-a_{2j}) + \vartheta_{4j}(x_2-a_{2j})^2]\} \qquad (11)$$

where a_j describes the location of maximums.

The least square method was used to fit this response. The values of concentrations corresponding to the uniform grid were taken as "observations" to restore $\vartheta^T = (\vartheta_{11}, \ldots, \vartheta_{41}, \ldots, \vartheta_{12}, \ldots, \vartheta_{42})$. The corresponding "estimates" (for normalized x_i: $-1 \le x_i \le 1$; $i=1,2$) were found to be $\hat{\vartheta}^T$ = (2.83, 4.42, -1.78, 15.3, 1.37, 0.54, -0.50, -2.00). Only the first exponent has a bell shape, while the second contains the negative coefficient ϑ_{42} corresponding to the quadratic term; this fact tends to confirm that there are more than two pollution emission centres.

One can see that in this example, unlike the linear case, we are concerned with the values of the parameters' estimates. The reason is that in the linear case the variance-covariance matrix does not depend upon estimated parameters while in the nonlinear case (see Fedorov and Uspensky, 1975) this matrix (or more accurately its asymptotic value) depends upon the true values of the unknown parameters ϑ_t:

$$\lim_{N\to\infty} ND(\hat{\vartheta}_N) = M^{-1}(\vartheta_t,\xi),$$

where $M(\vartheta,\xi) = \int f(\vartheta,x) f^T(\vartheta,x)\xi(dx)$, $f(\vartheta,x) = \dfrac{\partial\eta(x,\vartheta)}{\partial\vartheta}$,

where N is the number of observations and ξ is a limit point. Optimal designs formally defined by (2) will also depend upon ϑ_t, which are naturally unknown *a priori*. In this situation the following procedure is recommended:

- a user has to choose some probable (reasonable, admissible, etc.) values of ϑ and define intervals which will almost certainly contain true values of unknown parameters;
- for boundary points of these intervals, optimal designs have to be computed with the help of one of the above described programs;
- if the corresponding designs differ greatly from each other, an "average" design has to be constructed.

Fortunately optimal designs are rather stable to the variation of parameters and therefore the latter procedure can be avoided.

In our example, the vector $\hat{\vartheta}$ was taken as a central point and intervals were taken equal to $\pm 0.1\cdot\hat{\vartheta}_a$. Printout 6 and Figure 7 contain information on the D-optimal design ξ_c^* for the central point. All designs (optimal allocations) practically coincide with ξ_c^* and only in some of them do one or two additional points appear with small weights. These additions were removed and subsequently the corresponding determinants were computed. The ratio of determinants for the modified designs and optimal design fluctuated between 1.02 and 1.09. That is negligible for practical needs. Therefore, ξ_c^* can be used as a design, defining an optimal observation network for the nonlinear model (11).

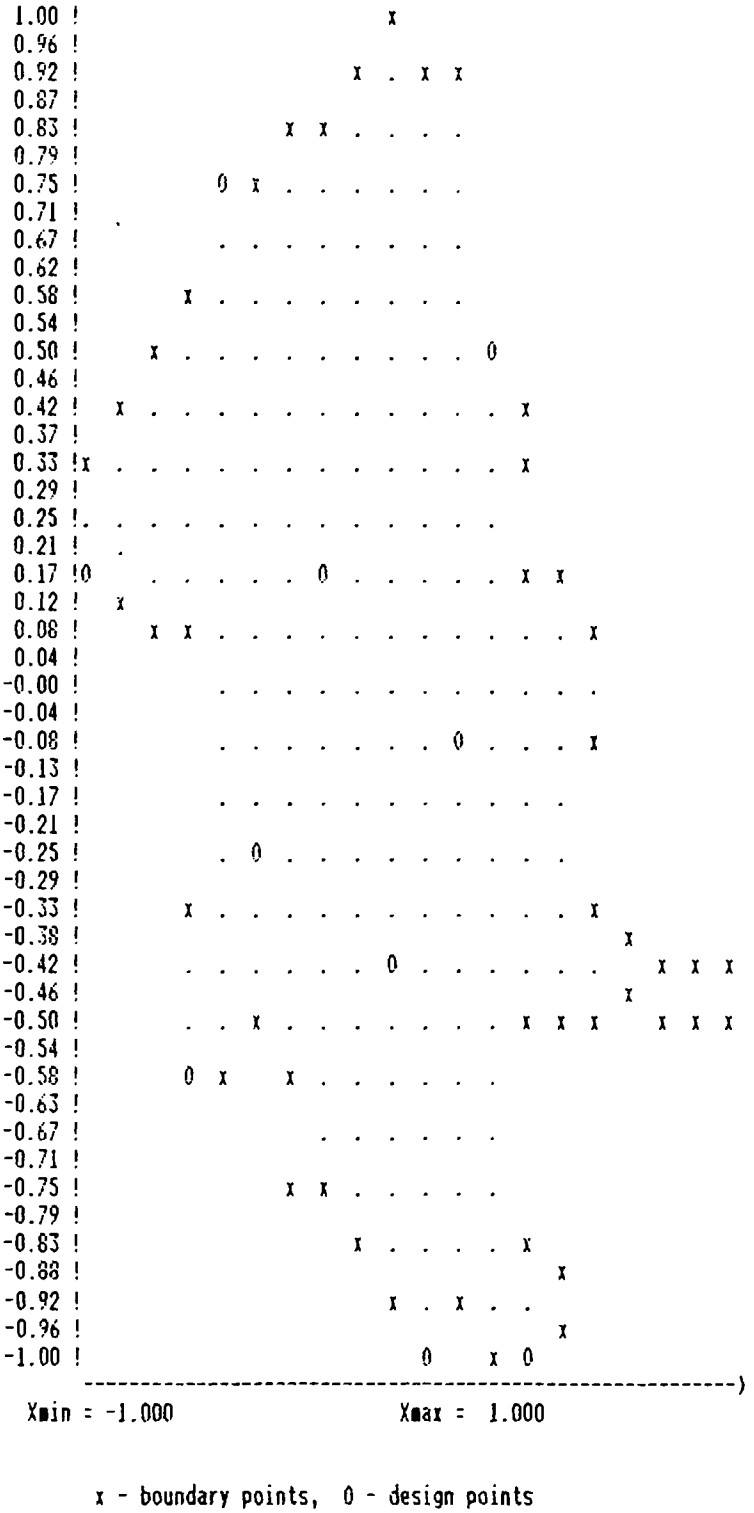

Figure 7: *D*-optimal observation network for nonlinear model (11).

REFERENCES

Atkinson, A.C. & V.V. Fedorov (1987) Optimum Design and Experiments. Encyclopedia of Statistical Sciences. New York: Wiley & Sons (to be published).

Fedorov, V.V. (1972) Theory of Optimum Experiments. New York: Academic Press, pp.292.

Fedorov, V.V. & A.B. Uspensky (1975) Numerical Aspects of Design and Analysis of Experiments. Moscow, Moscow State University, p.167.

Fedorov, V.V. (1986) Optimal Design of Experiments: Numerical Methods. WP-86-55. Laxenburg, Austria, International Institute for Applied System Analysis

Intergovernmental Working Group (1971) Proceedings of Meeting of Intergovernmental Working Group on Environmental Monitoring (IGNG, Geneva).

Izrael, Yu.A. (1974) Global observation system. Prediction and assessment of changes in environmental state. Based on monitoring. Meteorology: Hydrology, 7, 3-8 (in Russian).

Izrael, Yu.A, F.Yu. Rovinsky, M.Yu. Antonovski et al (1984) On the statistical substantiation of regional atmospheric pollution components in a background area. DAN SSSR 276(2):334-337 (in Russian).

Hanna, S.R., G.A. Briggs & R.P. Hosker (1982) Handbook on Atmospheric Diffusion. Technical Information Center, U.S. Department of Energy, p.102.

Mitchell, T.J. (1974) An Algorithm for Construction of D-Optimal Experimental Designs. *Technometrics*, 16:203-210.

Mehra, R.K. & D.G. Lainiotis (1976) Systems Identification: Advances and Case Studies. New York: Academic Press, p.593.

Modak, P.M. & B.N. Lohani (1985) Optimization of Ambient Air Quality Monitoring Networks. *Environmental Monitoring and Assessment*, Part I-II, 5:1-53.

Rovinsky, F.Ya, and G.B. Wiersma (1986) Procedures and Methods of Integrated Global Background Monitoring of Environmental Pollution, WMO. (Forthcoming.)

Silvey, S.D. (1980) Optimal Design. London: Chapman and Hall, p.86.

USER'S GUIDE

INSTRUCTIONS FOR PROGRAM MAP - MAPPING OF A CONTROLLABLE REGION R(X)

SCREEN	COMMENTS
1. SPACE DIMENSION - ? (L)	L is a number of controllable variables
2. X1(min), X1(max) - ? (X1min, X1max)	X1min, X1max are the minimal and maximal values of the first coordinate
3. GRID FOR X1 ? (NX1)	Interval (X1(min), X1(max)) is divided into NX1 parts, rx defines an initial grid for X1: rx = (X1(max) - X1(min))/NX1
Messages 4 - 7 appear if L = 2 .	
4. X2(min), X2(max) - ? (X2min, X2max)	X2min, X2max - minimal and maximal values of the second coordinate
5. GRID FOR X2 ? (NX2)	Interval (X2(min), X2(max)) is divided into NX2 parts, ry defines an initial grid for X2: ry = (X2(max) - X2(min))/NX2
Message 6 appears for all X1 = x , belonging to the grid.	
6. X1 = x, BOUNDS FOR X2 ? (Y1, Y2)	Y1 and Y2 are bounds of the 2-nd coordinate for current value x of the 1-st coordinate
7. NEW BOUNDS FOR X2 : yes - 1, no - 0 (INEW)	INEW = 1 - go to (6) with the same value x [if for a given x the set R(x) is not convex, R(x) = { y : a pair (x,y) belongs to the controllable region }] INEW = 0 - go to (6) with new value x , x = x(new) = x(old) + rx

INSTRUCTIONS FOR PROGRAM DOPT -

FIRST - ORDER OPTIMIZATION ALGORITHM FOR D - CRITERION

SCREEN	COMMENTS
1. SPACE DIMENSION - ? (L)	L is a number of controllable variables
2. CONSTANT FOR CONVERGENCE CRITERION - ? (EPS)	EPS - a constant for testing convergence of the algorithm
3. NUMBER OF ESTIMATED PARAMETERS - ? (M)	M - number of parameters (M must correspond to subroutine RESP , where a response function is calculated)
4. NUMBER OF POINTS IN INITIAL DESIGN - ? (N0)	N0 - number of supporting points in an initial design
5. DESIGN INPUT: FILE - 1, MONITOR - 2 (IDES)	IDES = 1 - initial design is saved in the file 'DES.DAT'; IDES = 2 - initial design is defined on the screen
Messages 6,7 appear if IDES = 2 (i = 1,....,N0)	
6. Point i , coordinates - ? (X(i,k), k = 1,L)	X(i,k) - coordinates of an i-th point in an initial design
7. Weight for point i ? (P(i))	P(i) - weight of an i-th point
8. GRAPHICAL PRESENTATION OF INITIAL DESIGN: yes - 1, no - 0 (IDO)	IDO = 1 - subroutine GRAPH is executed for initial design
Message 9 appears if initial information matrix is singular.	
9. SINGULAR INFORMATION MATRIX: NEW ATTEMPT: yes - 1, no - 0 (IDD)	IDD = 1 - go to (4) (new initial design is formed) IDD = 0 - STOP
10. SELECTION OF GAIN SEQUENCE: 1 - alfa(s) = const , 2 - alfa(s) = 1/s , 3 - alfa(s) is steepest descent sequence (IALF)	IALF = 1 - gain sequence is constant IALF = 2 - gain sequence is 1/s ; IALF = 3 - steepest descent method is executed

Message 11 appears if IALF = 1 .

11. CONSTANT FOR GAIN SEQUENCE (0.01 - 0.05) (ALFA)	ALFA is the chosen constant for gain sequence (0.01 - 0.05 - recommended values)
12. NUMBER OF ITERATIONS ? (MITER)	MITER - maximal number of iterations
13. CONSTANT FOR MERGING OF SUPPORTING POINTS - ? (CMER)	CMER is an internal constant (see section 3)
14. FORWARD LENGTH OF EXCURSION (NFOR)	NFOR - number of steps for forward procedure
15. BACKWARD LENGTH OF EXCURSION (NBAC)	NBAC - number of steps for backward procedure
16. INITIAL PROCEDURE: forward - 1, backward - 2 (IPRO)	The algorithm starts with: - forward procedure if IPRO = 1, - backward procedure if IPRO = 2.
17. STEPWIZE INFORMATION : yes - 1, no - 0 (IINF)	IINF = 1 - intermediate information is saved in the file 'OUT.DAT' and shown on the monitor (current design, value of the determinant etc)

Message 18 appears if L = 2 .

18. GRAPHICAL PRESENTATION OF DESIGN: yes - 1, no - 0 (IGR)	IGR = 1 - subroutine GRAPH is executed for final design
19. SCALING OF DESIGN: yes - 1, no - 0 (ISC)	ISC = 1 - scaling of final design is carried out

Messages 20 - 22 appear if ISC = 1.

20. X1(min), X1(max) - ? (X1min, X1max)	X1min, X1max - minimal and maximal values of the 1-st coordinate

Message 21 appears if L = 2 .

21. X2(min), X2(max) - ? (X2min, X2max)	X2min, X2max - minimal and maximal values of the 2-nd coordinate
22. GRAPH IN REAL SCALE: yes - 1, no - 0 (IGRS)	IGRS = 1 - subroutine GRAPH is executed for final design in real scale.

INSTRUCTIONS FOR PROGRAM DOPTEX -
OPTIMIZATION ALGORITHM OF THE EXCHANGE TYPE FOR D - CRITERION

SCREEN	COMMENTS

Messages 1 - 9 coincide with those for program DOPT .

10. NUMBER OF FIXED POINTS IN INITIAL DESIGN (MFIX)	The first MFIX points in initial design are fixed
11. CHOICE OF THE ALGORITHM: 1 - DELETING PROCEDURE, 2 - INCLUDING PROCEDURE, 3 - STANDARD EXCHANGE PROCEDURE (IALG)	IALG = 1 - deleting (ONLY !) is executed; IALG = 2 - including (ONLY !) is executed; IALG = 3 - exchange procedure with including and deleting is executed

Message 12 appears if IALG = 1.

12. NUMBER OF POINTS FOR DELETING (MBAC)	MBAC - number of steps for deleting procedure (MBAC must be less than M0 - M !!!)

Message 13 appears if IALG = 2.

13. NUMBER OF POINTS FOR INCLUDING (MFOR)	MFOR - number of steps for including procedure

Messages 14 - 19 appear if IALG = 3.

14. SELECTION OF STEPSIZE SEQUENCE: 1 - alfa = const, 2 - alfa = 1/s (IALF)	IALF = 1 - stepsize is constant; IALF = 2 - stepsize is of the form 1/s , s = M0+1, M0+2,...

Message 15 appear if IALF = 1.

15. CONSTANT FOR STEPSIZE - ? (ALFA)	ALFA is a constant stepsize for the algorithm

16. NUMBER OF ITERATIONS - ? (MITER) — MITER - maximal number of iterations

17. CONSTANT FOR MERGING OF SUPPORTING POINTS (CMER) — CMER is an internal constant of the algorithm (see section 3)

18. LENGTH OF EXCURSION - ? (forward and backward) (NFOR) — NFOR - number of steps for forward and backward procedures (Attention : MITER = 2*NFOR*K, K - integer !!!)

Messages 19 - 25 coincide with messages 16 - 22 for program DOPT .

Instructions for programs LINOPT and LINEX are almost the same as for programs DOPT and DOPTEX respectively. There appears one additional message (after message 3).

3*. UTILITY MATRIX (upper triangular part)
(ut(1,1), ut(1,2), ... , ut(1,n)
ut(2,2), ... , ut(2,n)
.....
ut(n,n))

{ ut(i,j) } is a symmetric utility matrix,
j = i,...,n ; i = 1,...,n .

INSTRUCTIONS FOR SUBROUTINE GRAPH -

GRAPHICAL PRESENTATION OF THE DESIGN

SCREEN	COMMENTS
1. Number of divisions for X1 ? (MX)	The graph has MX positions for the first coordinate and MY positions for the second coordinate
2. Number of divisions for X2 ? (MY)	

Printout 1: Completely new network, D-criterion.

```
      Example for program DOPT: output file 'OUT.DAT'
      -----------------------------------------------

SPACE DIMENSION - ?                          Number of controllable variables :
  2                                                    L = 2

CONSTANT FOR CONVERGENCE CRITERION - ?       Constant for convergence
  0.01000                                       test:  eps = 0.01

NUMBER OF ESTIMATED PARAMETERS - ?           Number of parameters :  m = 6
 6

NUMBER OF POINTS IN INITIAL DESIGN - ?       Number of points in initial
 11                                             design : N0  = 11

   *********

INITIAL DESIGN

point   weight   coordinates                  Supporting points of the
   1.   0.091     -0.5789  -0.5000          initial design and their weights
   2.   0.091     -0.5789  -0.4167
   3.   0.091     -0.4737   0.5000
   4.   0.091     -0.3684  -0.5000
   5.   0.091     -0.3684  -0.1667
   6.   0.091     -0.2632   0.0833
   7.   0.091     -0.1579  -0.1667
   8.   0.091     -0.0526  -0.8333
   9.   0.091      0.1579  -0.3333
  10.   0.091      0.3684   0.
  11.   0.091      0.4737  -0.2500
   *********

INITIAL INFORMATION MATRIX                     Initial information matrix:
  1.000                                      since the matrix is symmetric,
 -0.167   0.150                              only its low triangular part
  0.150  -0.041   0.035                            is shown
 -0.235   0.038  -0.032   0.169
  0.169  -0.042   0.023  -0.076   0.065
  0.038  -0.032   0.012  -0.042   0.011   0.023


DETERMINANT OF INITIAL INFORMATION MATRIX      The value of the optimality
                                               criterion: in this example
  8.08280e-09                                  it is the determinant of the
                                                  information matrix


INITIAL COVARIANCE MATRIX                      Initial covariance matrix:
  6.145                                      for example,  d(2) = 30.835
  5.487  30.835                              is the variance of the initial
-21.406  -7.068 123.469                      estimator for the 2-nd parameter
 13.228  39.169 -39.831  83.318
  5.754  38.872 -23.188  74.240  96.087
 29.515  89.311 -98.286 167.743 144.397 401.929
```

```
SELECTION OF GAIN SEQUENCE
1 - alfa(s) = const
2 - alfa(s) = 1/s
3 - alfa(s) is steepest descent sequence
   3
```

Steepest descent sequence is chosen

```
NUMBER OF ITERATIONS - ?
 100
```

Maximal number of iterations is 100

```
CONSTANT FOR MERGING OF
 SUPPORTING POINTS
  2.00000
```

```
FORWARD LENGTH OF EXCURSION - ?
 3
BACKWARD LENGTH OF EXCURSION - ?
 3
```

3 steps are executed for forward and backward procedures. Here NFOR and NBAC may differ from each other

```
INITIAL PROCEDURE:
forward - 1, backward - 2
 1
```

The algorithm starts with forward procedure

```
   *********
ITERATION  no.
 100
```

! Final information !

```
CONVERGENCE CRITERION VALUE
  0.04577

         ********** FINAL DESIGN ************

point   weight   coordinates
   1.   0.054     -0.5789   0.7500
   2.   0.172     -0.6842  -0.5833
   3.   0.130     -1.0000   0.3333
   4.   0.162      1.0000  -0.4167
   5.   0.157      0.3684  -1.0000
   6.   0.165     -0.0526   1.0000
   7.   0.159      0.0526   0.

FINAL COVARIANCE MATRIX
  6.006
  0.176   3.492
 -6.221   1.419  11.718
 -0.155   0.922   0.982   2.664
 -6.452  -0.314   5.953   0.193  10.418
 -2.738   3.628   8.524   0.632   4.113  19.606

VALUE OF THE DETERMINANT
  1.13645e-04
```

Final design

Final covariance matrix

Final value of the determinant (compare with the initial one: 8.08e-09)

Printout 2: Removal of the "least informative" stations.

```
TYPE OF ALGORITHM:
 1 - DELETING PROCEDURE
 2 - INCLUDING PROCEDURE
 3 - STANDARD EXCHANGE PROCEDURE

   1

NUMBER OF POINTS FOR DELETING - ?
 4

                                                    Number of steps for
                                                      deleting is 4

                                                    ------------------
   *********                                        ! Final information !
                                                    ------------------

ITERATION  no.
 4                                                  4 steps of deleting were executed

          **********  FINAL DESIGN  ************

point   weight   coordinates
   1.   0.143     -0.5789  -0.5000                  Here all the points have
   2.   0.143     -0.5789  -0.4167                    equal weights ( 1/7 )
   3.   0.143     -0.4737   0.5000
   4.   0.143     -0.1579  -0.1667
   5.   0.143     -0.0526  -0.8333
   6.   0.143      0.3684   0.
   7.   0.143      0.4737  -0.2500

UNNORMALIZED COVARIANCE MATRIX
 10.666
 -0.971  45.565
-36.067  12.813 171.325
  7.688  57.472 -23.651 114.453
 -7.060  69.305  15.910 115.633 161.417
 19.268 124.806 -69.439 228.493 226.590 526.796

UNNORMALIZED DETERMINANT
  1.45874e-09


FINAL COVARIANCE MATRIX                             Final covariance matrix
  6.788
 -0.618  28.996
-22.952   8.154 109.025
  4.893  36.573 -15.050  72.834
 -4.493  44.103  10.125  73.585 102.720
 12.262  79.422 -44.189 145.405 144.194 335.234

VALUE OF THE DETERMINANT                            The determinant and covariance
  2.19582e-08                                       matrix did not significantly improve:
                                                      new points were not included
```

Printout 3: Addition of new stations.

```
TYPE OF ALGORITHM:
 1 - DELETING PROCEDURE
 2 - INCLUDING PROCEDURE
 3 - STANDARD EXCHANGE PROCEDURE

   2

NUMBER OF POINTS FOR INCLUDING - ?
 3


   *********
ITERATION  no.
 3

          **********  FINAL DESIGN  ************

point   weight   coordinates
   1.   0.071     -0.5789  -0.5000
   2.   0.071     -0.5789  -0.4167
   3.   0.071     -0.4737   0.5000
   4.   0.071     -0.3684  -0.5000
   5.   0.071     -0.3684  -0.1667
   6.   0.071     -0.2632   0.0833
   7.   0.071     -0.1579  -0.1667
   8.   0.071     -0.0526  -0.8333
   9.   0.071      0.1579  -0.3333
  10.   0.071      0.3684   0.
  11.   0.071      0.4737  -0.2500
  12.   0.071      0.1579   0.9167
  13.   0.071      1.0000  -0.5000
  14.   0.071     -1.0000   0.3333


UNNORMALIZED COVARIANCE MATRIX
  2.487
  0.531   4.928
 -2.862   3.767  17.030
  0.831   1.808   2.613   4.959
 -3.647  -2.356  -1.809  -2.407  15.427
 -0.965   9.019  21.408   6.406  -8.025  50.258

UNNORMALIZED DETERMINANT
  1.82389e-05


FINAL COVARIANCE MATRIX
  3.165
  0.675   6.272
 -3.643   4.795  21.674
  1.058   2.301   3.326   6.311
 -4.642  -2.998  -2.302  -3.064  19.635
 -1.228  11.479  27.246   8.153 -10.213  63.965


VALUE OF THE DETERMINANT
  4.29129e-06
```

```
Now including procedure
     is chosen

Number of steps is 3

------------------
! Final information !
------------------

3 steps were executed

As for deleting procedure,
all the points have equal
   weights ( 1/14 )

Final covariance matrix

Final value of the
   determinant
```

Printout 4: Optimal observation network with fixed stations.

```
NUMBER OF FIXED POINTS IN INITIAL DESIGN
 5
CHOICE OF ALGORITHM:
 1 - DELETING PROCEDURE
 2 - INCLUDING PROCEDURE
 3 - STANDARD EXCHANGE PROCEDURE
 3
SELECTION OF STEPSIZE SEQUENCE
1 - alfa(s) = const
2 - alfa(s) = 1/s
   2

NUMBER OF ITERATIONS - ?
  80

CONSTANT FOR MERGING OF SUPPORTING POINTS
  2.00000

LENGTH OF EXCURSION - ? (forward and backward)
 4
INITIAL PROCEDURE: forward - 1, backward - 2
 1



   *********
ITERATION  no.
 80

CONVERGENCE CRITERION VALUE
  0.06620

        ********** FINAL DESIGN *************

point   weight   coordinates
   1.   0.091     -0.5789  -0.5000
   2.   0.091     -0.4737   0.5000
   3.   0.091     -0.0526  -0.8333
   4.   0.091      0.1579  -0.3333
   5.   0.091      0.3684   0.
   6.   0.143      1.0000  -0.4167
   7.   0.106     -0.0526   1.0000
   8.   0.110     -0.6842  -0.5833
   9.   0.094     -1.0000   0.3333
  10.   0.093      0.3684  -1.0000


FINAL COVARIANCE MATRIX
  5.579
  0.298   3.941
 -6.583   1.466  14.476
  0.335   0.960   0.565   3.148
 -6.736  -0.286   7.137   0.413  12.716
 -3.877   5.138  13.120   1.043   5.575  29.529

VALUE OF THE DETERMINANT
  5.26021e-05
```

Printout 5: Completely new network, A-criterion.

```
ITERATION  no.
 100

CONVERGENCE CRITERION
  0.80579

          **********  FINAL DESIGN  ************

point   weight   coordinates
   1.   0.240     -0.6842  -0.5833
   2.   0.104     -0.5789   0.7500
   3.   0.128      0.1579   0.9167
   4.   0.123      0.3684  -1.0000
   5.   0.051     -1.0000   0.3333
   6.   0.117      1.0000  -0.4167
   7.   0.237     -0.0526   0.

FINAL COVARIANCE MATRIX
  4.112
  0.777   4.710
 -4.269   1.014  12.164
  0.211   1.132   1.493   3.169
 -4.730  -0.857   2.951  -0.519  10.331
 -1.661   4.375   6.991   1.302   3.228  18.299

VALUE OF THE DETERMINANT      0.00006880

VALUE OF THE CRITERION -  trace ( UTIL * D )
  52.7850
```

Printout 6: *D*-optimal observation network for nonlinear model (11).

```
SPACE DIMENSION - ?
   2

CONSTANT FOR CONVERGENCE CRITERION - ?
  0.02500

NUMBER OF ESTIMATED PARAMETERS - ?
 8

NUMBER OF POINTS IN INITIAL DESIGN - ?
 11


   *********
INITIAL DESIGN
point   weight   coordinates
   1.   0.091     -0.5789  -0.5000
   2.   0.091     -0.5789  -0.4167
   3.   0.091     -0.4737   0.5000
   4.   0.091     -0.3684  -0.5000
   5.   0.091     -0.3684  -0.1667
   6.   0.091     -0.2632   0.0833
   7.   0.091     -0.1579  -0.1667
   8.   0.091     -0.0526  -0.8333
   9.   0.091      0.1579  -0.3333
  10.   0.091      0.3684   0.
  11.   0.091      0.4737  -0.2500
   *********

INITIAL INFORMATION MATRIX
    0.281
   -0.024     0.010
   -0.011     0.001     0.003
   -0.013     0.001     0.001     0.001
    0.484    -0.070    -0.015    -0.032     1.528
   -0.080     0.004     0.004     0.006    -0.188     0.044
   -0.091    -0.000     0.003     0.008    -0.218     0.043     0.105
   -0.040     0.003    -0.001     0.005    -0.300     0.026     0.057     0.108


DETERMINANT OF INITIAL INFORMATION MATRIX
  4.60934e-16

INITIAL COVARIANCE MATRIX
  204.087
 -854.353  4071.180
  220.849  -987.489   726.125
-1613.658  7300.608 -2165.822 16616.678
 -249.994  1126.751  -276.691  2096.874   322.813
 -250.758  1200.176  -288.780  1915.018   340.029   449.828
  169.281  -725.331   190.721 -1450.266  -212.839  -227.304   163.815
 -551.805  2466.506  -588.549  4587.538   709.356   743.261  -475.257  1573.429

   *********
```

```
   *********
ITERATION  no.
 100

CONVERGENCE CRITERION VALUE
 -0.01211

          ********** FINAL DESIGN ************

point   weight   coordinates
   1.   0.069     -0.5789   0.7500
   2.   0.087      0.0526  -1.0000
   3.   0.121      0.1579  -0.0833
   4.   0.116     -0.4737  -0.2500
   5.   0.134     -0.6842  -0.5833
   6.   0.066      0.3684  -1.0000
   7.   0.105      0.2632   0.5000
   8.   0.114     -0.0526  -0.4167
   9.   0.075     -1.0000   0.1667
  10.   0.113     -0.2632   0.1667

FINAL COVARIANCE MATRIX
   17.717
  -12.204   273.555
    6.131   -64.824   314.118
  -58.605   207.324  -317.184  2815.108
   -9.165    33.713    -7.270    94.892    11.031
    2.249    21.348    11.376   -88.175     3.659    26.831
    2.307     2.827    -3.053   -41.319    -0.993     0.325     6.663
  -13.881    50.596   -10.949   138.098    16.540     4.171    -0.756    26.918

VALUE OF THE DETERMINANT
  1.01261e-12

          SCALING OF FINAL DESIGN
X1(min), X1(max) - ?
 0.   19.0000
X2(min), X2(max) - ?
 0.   24.0000
   *********

 FINAL DESIGN (in real scale)
point   weight   coordinates
   1.   0.069      4.0005  21.0000
   2.   0.087      9.9997   0.
   3.   0.121     11.0001  11.0004
   4.   0.116      4.9999   9.0000
   5.   0.134      3.0001   5.0004
   6.   0.066     12.9998   0.
   7.   0.105     12.0004  18.0000
   8.   0.114      9.0003   6.9996
   9.   0.075      0.       14.0004
  10.   0.113      6.9996  14.0004
```