WORKING PAPER

DESIGN OF AN OBSERVING NETWORK: COMPARISON OF TWO APPROACHES.

V. Fedorov
W. Mueller

June 1987
WP-87-55

IIASA
International Institute
for Applied Systems Analysis

DESIGN OF AN OBSERVING NETWORK: COMPARISON OF TWO APPROACHES.

V. Fedorov
W. Mueller

June 1987
WP-87-55

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

Preface

This report gives the results of some collaborative research on the design of observing networks, undertaken by Valeri Fedorov of the IIASA Environment Program and Mr. Werner Mueller, a graduate student in the Department of Statistics at the University of Vienna. During the winter of 1986-87, Mr. Mueller was a guest Research Aassistant at IIASA, working with Professor Fedorov on various aspects of the optimal design of environmental monitoring networks.

I am particularly pleased with this example of collaboration between the IIASA Environment Program and the academic community in Austria. It is the kind of activity that should be strongly encouraged.

R.E. Munn
Head, Environment Program

DESIGN OF AN OBSERVING NETWORK: COMPARISON OF TWO APPROACHES.

V. Fedorov and W. Mueller

1. INTRODUCTION

In the last decade, the necessity to use mathematical methods for optimizing observing networks was discussed intensively in many publications related to environmental monitoring (see, for instance, Der Megreditchan, 1985; Modak and Lohani, 1985; Munn, 1981). The methods described in these publications have a mainly empirical character. They provide a practitioner with recommendations or formulae for the sequential deleting or inclusion of stations. The corresponding procedures are optimal at every step. The less (most) informative observing station is deleted from (included in) the monitoring network. Their global optimality, e.g., that the network is optimal was not investigated. All methods which the authors found in environmental or methodological journals and publications are based on the analysis of the covariance structure of an observed field, which is assumed to be random. The estimation of the covariance structure is quite a difficult problem, both from the theoretical and experimental points of view.

It appears that network optimization theory developed independently from the optimal experimental design theory, which is a well elaborated sector of mathematical statistics. This theory can also be used for the optimization of observing networks (see for instance Fedorov et al, 1987).

The main objective of this paper is a comparison of the two above mentioned approaches. As a testground for this comparison, the random fields generated by the second-kind regression model were chosen. (These fields can be investigated either within random fields theory or in the framework of regression analysis theory).

Section II contains a short survey of the statistical theory related to the analysis and design of experiments described by the second kind regression models.

In Section III, the properties of the estimators and the network design procedure when applied to the fields generated by the second-kind regression models are analyzed in detail.

The concluding section is devoted to a comparative analysis of the two approaches.

II. SECOND-KIND REGRESSION MODEL (ANALYSIS AND EXPERIMENTAL DESIGN)

In many experimental situations the following assumption seems to be reasonable: at any given j-th time interval (say, season or decade) the observed values are described by response functions that have the same structure differing only by intrinsic parameters:

$$y_{ij} = \eta(\vartheta_j, x_i) + \varepsilon_{ij} \ , \ i = \overline{1,n} \ , \ j = \overline{1,k} \tag{1}$$

y_{ij} are observed values, the vector $x_i \in R^L$ stands for the location of the i-th observing station, ϑ_j is a vector of parameters at the j-th set of observations, ε_{ij} is a random value (usually it is the error of an observation y_{ij}).

In this paper it will be assumed additionally that

- y_{ij} are scalar values;
- $\eta(\vartheta, x)$ is a linear function of parameters ϑ_j, i.e., $\eta(\vartheta_j, x) = \vartheta_j^T f(x)$, where $\vartheta_j \in R^m$ and $f(x)$ is a vector $(m \times 1)$ of priori given functions;
- ϑ_j is a random vector, $E(\vartheta_j) = \vartheta_0$ and $E[(\vartheta_j - \vartheta_0)(\vartheta_j - \vartheta_0)] = D_0$;
- $E[\varepsilon_{ij} / j] = 0$, $\sim E[\varepsilon_{ij}\varepsilon_{i'j'} / j, j',] = \sigma^2 \delta_{jj'}$, $\delta_{jj} = 1$, $\delta_{jj'} = 0$, $j \neq j'$, $E[\varepsilon_{ij}\varepsilon_{i'j'} / i, i'] = 0$, $i, i' = \overline{1,n}$, $j, j' = \overline{1,k}$.

Here $E[\cdots]$ and $E[.../...]$ are "mean" and "conditional mean" values correspondingly.

In what follows two other presentations of (1) will be useful:

$$y_j = F^T \vartheta_j + \varepsilon_j \ , \tag{2}$$

$$y_j = F^T \vartheta_0 + v_j = F^T \vartheta_0 + \{F^T(\vartheta_j - \vartheta_0) + \varepsilon_j\}, \tag{3}$$

where $v_j = F^T(\vartheta_j - \vartheta_0) + \varepsilon_j$, $y_j^T = (y_{1j}, \ldots, y_{nj})$, $F = \{f(x_1), \ldots, f(x_n)\}$.

The most crucial assumption from the above set is that the fluctuation of responses is modelled by the randomness of parameters ϑ_j. In the majority of publications related to the optimization of an observing network the response $y_j(x)$ is assumed to be a random field. From (3) it is evident that second kind regression models lead to finite dimension parameterization of this field with an average trend $F^T \vartheta_0$ and a correlation structure defined by

$$E[(y_j - F^T\vartheta_0)(y_j - F^T\vartheta_0)^T] = \sigma^2 I + F^T D_0 F \ , \tag{4}$$

where I is the identity matrix.

The last remark hints that the estimation problem for (1) can be handled either with the help of the traditional regression (or model fitting) approach or the filtration theory approach.

The estimation theory for problem 1 is well developed (see, for instance, Rao, 1975 and Fedorov and Uspensky, 1977); therefore the essential results used in the following section will be only briefly surveyed.

(a) *The vector of mean values ϑ_0 and variance-covariance matrix D_0 are known, parameters ϑ_j have to be estimated.*

In this paper we restrict ourselves to the class of linear estimators

$$\tilde{\vartheta}_j = L_1 y_j + L_2 \vartheta_0 \ . \tag{5}$$

The best linear unbiased estimator $\hat{\vartheta}_j$ is defined as a solution of the following minimization problem:

$$\hat{\vartheta}_j = Arg \min_{L_1, L_2} E[(\tilde{\vartheta}_j - \vartheta_0)(\tilde{\vartheta}_j - \vartheta_0)^T] \ , \tag{6}$$

subject to the unbiased-ness constraints $E(\tilde{\vartheta}_j - \vartheta_0) = 0$.

Problem (6) is multiobjective and the assertion that $\hat{\vartheta}_j$ is a solution of it means that the variance-covariance matrix $D(\hat{\vartheta}_j)$ of $\hat{\vartheta}_j$ is the least one in the class of linear unbiased estimators:

$$D(\hat{\vartheta}_j) = E[(\hat{\vartheta}_j - \vartheta_0)(\hat{\vartheta}_j - \vartheta_0)^T] \leq E[(\tilde{\vartheta}_j - \vartheta_0)(\tilde{\vartheta}_j - \vartheta_0)^T] = D(\tilde{\vartheta}_j) \tag{7}$$

or $D(\hat{\vartheta}_j) = D(\tilde{\vartheta}_j) - \delta$, where δ is a nonnegatively defined matrix.

The straightforward calculations give

$$\hat{\vartheta}_j = (D_0^{-1} + M)(D_0^{-1}\vartheta_0 + \sigma^{-2}F y_j), \tag{8}$$

where

$$M = \sigma^{-2}FF^T = \sigma^{-2}\sum_{i=1}^{n} f(x_i)f^T(x_i), \tag{9}$$

and

$$D(\hat{\vartheta}_j) = (D_0^{-1} + M)^{-1} \tag{10}$$

It is evident that estimator (8) is better (in the sense of (7)) than the usual least square estimator

$$\bar{\vartheta}_j = \sigma^{-2}M^{-1}F y_i \tag{11}$$

which has a greater variance-covariance matrix

$$D(\bar{\vartheta}_j) = M^{-1} \geq (D_0^{-1} + M)^{-1}, \tag{12}$$

but it has to be emphasized that in $\bar{\vartheta}_j$ the information about ϑ_0 and D_0 is not used.

For better understanding of the interrelation between $\hat{\vartheta}_j$ and $\bar{\vartheta}_j$ it is useful to note that

$$\hat{\vartheta}_j = Arg \min_{\vartheta_j} [\sum_{i=1}^{n} \sigma^{-2}(y_{ij} - \vartheta_j^T f(x_i))^2 + (\vartheta_j - \vartheta_0)^T D_0^{-1}(\vartheta_j - \vartheta_0)],$$

while

$$\bar{\vartheta}_j = Arg \min_{\vartheta_j} \sum_{i=1}^{n} \sigma^{-2}(y_{ij} - \vartheta_j^T f(x_i))^2 .$$

(b) *Variance-covariance matrix D_0 is known, ϑ_0 and ϑ_j have to be estimated.*

If $\tilde{\vartheta}_0 = L_0 y$, where $y^T = (y_1^T, \ldots, y_k^T)$ and

$$\hat{\vartheta}_0 = Arg \min_{L_0} E[(\tilde{\vartheta}_0 - \vartheta_0)(\tilde{\vartheta}_0 - \vartheta_0)^T], \; E(\tilde{\vartheta}_0 - \vartheta_0) = 0, \tag{13}$$

then rather tedious but relatively simple calculations lead to

$$\hat{\vartheta}_0 = k^{-1}\sum_{j=1}^{k} \bar{\vartheta}_j , \tag{14}$$

where $\bar{\vartheta}_j$ is defined by (11) and

$$D(\hat{\vartheta}_0) = k^{-1}(D_0+M^{-1})^{-1}\,. \tag{15}$$

The best linear unbiased estimator for ϑ_j coincides with $\bar{\vartheta}_j$

(c) Neither ϑ_0 or D_0 are known.

In this case it is impossible to construct estimators similar to (6) or (13) and the general recommendation is (see, for instance, Fisk 1967; Rao, 1975) that one can use the formulae from section (a) and (b) if ϑ_0 and D_0 are substituted by:

$$\hat{\vartheta}_0 = k^{-1}\sum_{j=1}^{k}\bar{\vartheta}_j \text{ and } \hat{D}_0 = (k-1)^{-1}\sum_{j=1}^{k}(\bar{\vartheta}_j-\hat{\vartheta}_0)(\bar{\vartheta}_j-\hat{\vartheta}_0)^T \tag{16}$$

or any other consistent estimators. The consistency will guarantee the asymptotical fulfillment of (10) and (15), for instance.

In all three cases the prognosis for response function $\vartheta^T f(x)$ at any given point x based on formulae $\tilde{y}(x) = \tilde{\vartheta}^T f(x)$, where $\tilde{\vartheta}$ can coincide with any discussed estimator. The variance of $\tilde{y}(x)$ equals

$$d(x) = f^T(x)D(\tilde{\vartheta})f(x), \tag{17}$$

where $D(\tilde{\vartheta})$ is the corresponding variance-covariance matrix.

As soon as one knows the dependence of the accuracy of the estimator of parameters ϑ_j and ϑ_0 (see (9), (10), (12), or the accuracy of $\tilde{y}(x)$ upon the locations of observing stations the problem of their optimal choice can be immediately formulated:

$$\xi_n = (x_1, \ldots, x_n),$$
$$\xi_n^* = Arg \min_{\xi_n} \Phi[D(\tilde{\vartheta},\xi_n)], \tag{18}$$

where $D(\tilde{\vartheta}_1\xi_n)$ is defined by (10), (12) or (15), and Φ is a criterion of optimality (see, for instance, Fedorov 1972; Fedorov et al, 1987). In this paper for the sake of simplicity only the D-criterion (or related criteria) will be considered:

$$\Phi(D) = ln\,|D|\,.$$

In what follows we shall need three versions of the D-criterion:

$$\Phi_1 = -ln\,|M|\,,\ \Phi_2 = -ln\,|D_0^{-1}+M|\,,\ \Phi_3 = -ln\,|D_0+M^{-1}|$$

In spite of their similarity the corresponding optimal locations are oriented to different purposes. In the first case we are trying to increase the effectiveness of estimation of ϑ_j, i.e., the current situation (say, on j-th day) is of prime interest. In the third case the main objective is the estimation of an average situation (ϑ_0 or $\vartheta_0^T f(x)$ could be annual averages). The second case is similar to the first one but the prior information ϑ_0 and D_0 is taken into account.

From the optimal experimental design theory (see, for instance, Fedorov 1972) it follows that the locations of observing stations have to coincide with maxima of function $\varphi(x,\xi_k)$ defined for the criteria considered in Table 1.

Table 1.

Criterion Φ	$\varphi(x_1\xi_k)$
$\ln\|M(\xi_k)\|^{-1}$	$f^T(x)M^{-1}(\xi_k)f(x)$
$\ln\|D_0^{-1}+M(\xi_k)\|^{-1}$	$f^T(x)[D_0^{-1}+M(\xi_k)]^{-1}f(x)$
$\ln\|D_0+M^{-1}(\xi_k)\|$	$f^T(x)M^{-1}(\xi_k)[D_0+M^{-1}(\xi_k)]^{-1}M^{-1}(\xi_k)f(x)$

One of the simplest numerical procedures for finding an optimal location of observing stations (optimal design in statistical publications) could be the following one:

[1] There is $\xi_{ks} = (x_{1s}, \ldots, x_{ks})$.
Matrix $M(\xi_{ks}) = \sum_{i=1}^{k} f(x_{is})f^T(x_{is})$ is computed.

[2] Point

$$x_s^+ = Arg \max_{x \in X} \varphi(x, \xi_s) \tag{19}$$

has to be found, X is the area where the stations can be located.

[3] For modified design $\xi_{k+1,s} = (x_{1s}, \ldots, x_{ks}, x_s^+)$ matrix $M(\xi_{k+1,s})$ is computed and point

$$x_s^- = Arg \min_{x \in \xi_{k+1,s}} \varphi(x, \xi_{k+1,s}) \tag{20}$$

is deleted.

Stages 1-3 are repeated with $\xi_{k,s+1}$ coinciding with $\xi_{k+1,s}$ except the deleted point $x_{is} = x_s^-$.

The properties of iterative procedure [1]-[3] and its more sophisticated versions were frequently discussed in the statistical literature (see, for instance, the survey paper by Fedorov, 1986). The forward procedure [2] and the backward one [3] can have a variable number of steps (length of excursion). For instance, one can add on every s-th step several points $x_{1s}^+, \cdots, x_{qs}^+$ and then delete the same number of points. In extreme cases either forward or backward procedures can be used to find the best locations of new stations or to delete the least informative of existing stations. For more accurate and more detailed information about the numerical procedures of type [1]-[3] see Fedorov et al, 1987.

III. ESTIMATION AND DESIGN BASED ON THE ANALYSIS OF THE COVARIANCE STRUCTURE OF AN OBSERVED RANDOM FIELD.

In this section we shall discuss the heuristic approaches related to statistical analysis of the variance-covariance characteristics of an observed field. These approaches are all based on the assumption that this field can be described (or modelled) by some stationary random field. To make the results compatible and

comparable with the previous section we assume that the random field is generated by (1) - (3), i.e., (compare with (4)):

$$C = E[(y_j - q)(y_j - q)^T] = \sigma^2 I + F^T D_0 F \ , \ q = E[y_j] \ , \tag{21}$$

or

$$R = E[y_j y_j^T] = \sigma^2 I + F^T D_0 F + F^T \vartheta_0 \vartheta_0^T F \ , \quad j = \overline{1,J} \ . \tag{22}$$

It has to be noted that the approaches to be discussed are not using the presentation (21) or (22) and they are assumed to be model free.

Similar to Section II, let the response y be observed at points $x_1, \ldots, x_n$. To prognose $y_{0\tau}$ at the given point x_0 the linear estimators

$$\tilde{y}_{0\tau} = \lambda^T y_\tau \tag{23}$$

or

$$\tilde{y}_{0\tau} = \lambda_0 + \lambda^T y_\tau \ , \tag{24}$$

for instance, where τ could be equal to $j+1$, for instance, and $y_\tau^T = (y_{1\tau}, \ldots, y_{n\tau})$ are usually used (see Der Megreditchan, 1985; Katkovnik, 1985).

If the quadratic risk (i.e., $E[(\hat{y} - y_t)^2]$, where $\hat{y}$ is the estimator and y_t is the true value) of the prognosis has to be minimized, then in the case of (23):

$$\hat{\lambda} = Arg \min_{\lambda} E[(\lambda^T y_\tau - y_{0\tau})^2] = R_{11}^{-1} R_{10} \ ,$$

$$v_{10}^2 = E[(\hat{\lambda} y_\tau - y_{0\tau})]^2 = R_{00} - R_{01} R_{11}^{-1} R_{10} \ , \tag{25}$$

and

$$\hat{y}_{0\tau} = R_{01} R_{11}^{-1} y_\tau \ ,$$

where $R_{00} = E(y_{0\tau}^2)$, $R_{11} = E(y_\tau y_\tau^T)$ corresponds to explanatory points $x_1, \ldots, x_k$ and $R_{01}^T = R_{10} = E(y_\tau y_{0\tau})$ describes the correlation between explanatory points and point x_0, where the prognosis has to be provided.

For (24) the minimum quadratic risk estimator is defined by

$$\hat{\lambda}_0 = q_0 - C_{11}^{-1} C_{10} \ , \ \hat{\lambda} = C_{11}^{-1} C_{10} \ ,$$

$$v_{20}^2 = E[(\hat{\lambda}_0 + \hat{\lambda} y_\tau - y_{0\tau})^2] = C_{00} - C_{01} C_{11}^{-1} C_{10} \ , \tag{26}$$

and

$$\hat{y}_{0\tau} = q_0 + C_{01} C_{11}^{-1} (y_\tau - q) \ ,$$

where $q_0 = E(y_{0\tau})$, $q = E(y_\tau)$, C_{11} and $C_{10} = C_{01}^T$ are centered versions of R_{11} and R_{10} correspondingly:

$$R_{11} = C_{11} + qq^T \text{ and } R_{10} = C_{10} + q_0 q^T \ .$$

There exist several similar methods from which to choose the most informative points from the random stationary field described by its correlation function (see Der-Megreditchan, 1985, Ch.VI). We consider a method which is closely related to the field interpolation or prognosis problem, see (25), (26). The idea is clear and simple.

[1] There are k+1 points where observations can be made. The point which is best explained by all other observations, i.e., with the smallest quadratic risk v_{1i}^2 (or v_{2i}^2), has to be deleted. The subscript "i" means now that the value of the field at the i-th point is interpolated by (25) or by (26).

[2] The procedure is repeated until the total number of observation points is small enough or v_{1i}^2 (or v_{2i}^2) will not increase drastically.

The described procedure is based on the very restrictive assumptions that *the correlation matrix and the means for k+1 points are priori known.*

To some extent the procedure is similar to the deletion stage of procedure [1]-[3] from section II.

Simple matrix calculations show that if C corresponds to k+1 observation points and C^{-1} is its inverse matrix then (see for instance, Seber, 1975; Der-Megreditchan, 1985):

$$C_{ii}^{-1} = 1/v_{2i} \tag{27}$$

and therefore at every step of the above formulated procedure one has to find:

$$i^- = Arg \max_i C_{ii}^{-1} \tag{28}$$

and to delete point x_i, from the given set $x_1, \ldots, x_{k+1}$.

To include a new point (or station) one has to solve the continuous optimization problem.

$$x^+ = Arg \min_{x \in X} C_{xx}^{-1}, \tag{29}$$

where

$$C(x) = \begin{bmatrix} C & C_x \\ C_x^T & C_{xx} \end{bmatrix},$$

$$C_x = E[(y_\tau - q)(y(x) - q(x))],\ C_{xx} = E[(y(x) - q(x))^2],$$

C_{xx}^{-1} is the element of matrix $C^{-1}(x)$ with the same position as C_{xx} and

$C_{xx}^{-1} = 1/v_{2x}$, $v_{2x} = E[(\hat{\lambda}_0 + \hat{\lambda} y_\tau - y_{x\tau})^2]$.

Let C be defined by (22). Then

$$C^{-1} = (\sigma^2 I + F^T D_0 F)^{-1} = \sigma^{-2} I - \sigma^{-4} F^T (D_0^{-1} + \sigma^{-2} F F^T) F$$
$$= \sigma^{-2} I - \sigma^{-4} F^T [D_0^{-1} + M(\xi_{k+1})]^{-1} F \tag{30}$$

where

$$M(\xi_{k+1}) = \sigma^{-2} F F^T = \sigma^{-2} \sum_{i=1}^{k+1} f(x_i) f^T(x_i).$$

In (30) the following formula was used:

$$(A + E^T B E) = A^{-1} - A^{-1} E^T (E A^{-1} E^T + B^{-1}) E^T A^{-1}.$$

From (28) and (30) it immediately follows that

$$i^- = Arg \min f^T(x_i)[D_0^{-1} + M(\xi_{k+1})]^{-1} f(x_i). \tag{31}$$

and point $\bar{x} = x_{i^-}$ has to be deleted from ξ_{k+1}.

The inclusion procedure is defined (see (29) and (30)) by the following optimization problem:

$$x^+ = Arg \max_{x \in X} f^T(x)[D_0^{-1} + M(\xi_k)]^{-1} f(x) . \tag{32}$$

The substitution of D_0 by $D_0 + \vartheta_0 \vartheta_0^T$ in (31) and (32) leads to the iterative procedure for estimator (25).

IV. COMPARISON OF TWO APPROACHES

The comparison of (30) and (31) with (19),(20) and the second line of Table 1 shows that the generalized version of (28) coincides with the simplest iterative procedure for numerical construction of the D-optimal design (see Fedorov, 1986; Fedorov et al., 1987) when the criterion of optimality is based on a variance-covariance matrix of estimator (8).

Thus the complete similarity between the approaches given in Sections II and III is established.

However, here we have to remind the reader that in Section III it was assumed that the correlation structure of the random field is known, i.e., matrix C (or R, for the sake of simplicity only problem (26) will be considered) for any prescribed set $x_1, \ldots, x_n$ is known. Unfortunately, in practice this matrix has to be estimated. The necessary data are available from stations already in operation. Therefore, if one does not apply additional assumptions, only the deletion procedure (28) can be used with the natural substitution of C by some estimate $\hat{C}$. Usually (see also (21)):

$$\hat{C} = J^{-1} \sum_{j=1}^{J} (y_j - \hat{q})(y_j - \hat{q})^T ,$$

$$\hat{q} = J^{-1} \sum_{j=1}^{J} y_j .$$

It is worthwhile to emphasize that

- maximization problem (28) is very unstable under variation of C (or $\hat{C}$);
- the inverse matrix $\hat{C}^{-1}$ can be calculated only in the case when the length of the available time series (y_{ij}) is greater than the initial number of observing stations;
- the intention to delete some stations implicitly assumes that matrix $\hat{C}$ is ill-conditioned; this fact leads to significant computational difficulties;
- the initial observing network can contain a large number of stations and it causes additional computational problems in (28);
- the real quadratic risk $\hat{v}_{10}^2$ and $\hat{v}_{20}^2$ (compare with (25) and (26)) after substitution of C by $\hat{C}$ will be greater than v_{10}^2 and v_{20}^2 ;

- formulae (26) with $\hat{C}$ instead of C provide the approximate solutions of quadratic risk minimization problems, the approximation can be quite poor when the volume of the learning sample (J) has the same order as the number of observing stations (n).

Should one wish to expand the observing network, it is necessary to apply some interpolation procedure to get $\hat{C}$ including any prescribed location x.

This is the most sensitive stage (at least in theory) of the whole approach. One must invent some physically reasonable assumptions on the random field structure considered.

Usually these are assumptions on stationarity and space homogeneity of this field and assumptions on the analytical form of the covariance function $C(x,x')$, (see for instance, Der Megreditchan, 1985).

In the applied papers (most of which concern only data analysis) this set of assumptions is shadowed by technical experimental details and by discussions about the nicety of (23), (24) or similar prognoses (see for instance, Endlich et al, 1986; Clark et al, 1986, Der Megreditchan, 1985) which gives readers the hope that model free approaches are used.

In the approach considered in Section II, a practitioner faces a similar problem at the very beginning. More explicitly, he must choose the vector of basic functions f(x) (or $\eta(x,\vartheta)$ in the more general case). After overcoming this problem, rather routine and effective numerical algorithms can be used to construct an optimal observing network.

In the previous case experimental data are used to restore matrix $\hat{C}$. In the latter case they can be used to verify (or to test) hypothesis (1).

Summarizing the discussions in this section, one comes to the following conclusions:

1. From a computational point of view, the approach from Section II is more reliable. At least in the case of linear parameterization, it provides a deterministic optimal solution.
2. The approach of Section II explicitly uses the assumption on physical models forcing a practitioner to be more accurate in his physical considerations and hypothesis.

Of course, these conclusions do not cancel the usefulness of the approach discussed in Section III. Some experimental situations can be effectively described only by models based on the variance-covariance structure of an observed field. The most ideal instance (rarely realized) is when a practitioner deals with a field which is really homogeneous in space and stationary in time.

REFERENCES

Der Megreditchan, G. (1985) Methodes statistique d'analyse et d'interpolation des champs meteorologiques. Geneve, Organization Meteorlogique Mondiale, p.300.

Clark, T.L., E.C. Voldner, M.P. Olson, S.K. Seilkop, M. Alvo (1986) International Sulfur model evaluation. Report.

Endlich, R.M., B.P. Eynon, R.J. Ferek, A.D. Valdes, C. Maxwell (1986) Statistical Analysis of Precipitation Chemistry Measurements. Over the Eastern United States, UAPSP, 112, EPRI.

Fedorov, V. (1972) Theory of Optimum Experiments, Academic Press, N.Y., p.300.

Fedorov, V. (1986) Optimum Design of Experiments: Numerical Methods. WP-86-55, IIASA.

Fedorov V. and A. Uspensky (1977) Linear Estimators in the Atmospheres Sounding Inverse Problem.

Fedorov, V., S. Leonov, M. Antonovsky, S. Pitovranov (1987) The experimental design of an observation network: software and examples. WP-87-05, IIASA.

Fisk, P.R. (1967) Models of the Second Kind in Regression Analysis. JRSS, ser. 15, 29, p.266-278.

Katkovnik, V. (1985) Nonparametric identification and data smoothing. Nauka, Moscow. pp. 336.

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

Munn, R.E. (1981) The Design of Air Quality Monitoring Networks. London: McMillan Publishers Ltd.

Rao, C.R. (1973) Linear Statistical Inference and its Applications. N.Y. Wiley.

Seber, G.A.F (1977) Linear regression analysis. Wiley, N.Y. pp. 456.