

UC Santa Barbara

UC Santa Barbara Previously Published Works

Title

An optimization approach for groundwater quality monitoring network design

Permalink

<https://escholarship.org/uc/item/5qk5k5f9>

Journal

Water Resources Research, 25(8)

ISSN

0043-1397

Author

Loaiciga, Hugo A

Publication Date

1989-08-01

DOI

10.1029/wr025i008p01771

Peer reviewed

An Optimization Approach for Groundwater Quality Monitoring Network Design

HUGO A. LOAICIGA

Department of Geography, University of California, Santa Barbara

The optimal sampling plan for groundwater quality monitoring is formulated as a mixed integer programming (MIP) problem. A sampling plan consists of the number and locations of sampling sites as well as the temporal sampling frequency. The MIP network problem is defined by the minimization of the variance of estimation error subject to resource and unbiasedness constraints. The mean and covariance of the spatial/temporal variable (chloride concentration measurements) are derived from the advection-dispersion equation governing mass transport. The solution for the optimal sampling proceeds in two stages: (1) parameter estimation and (2) network optimization. The MIP model was successfully tested with a network design problem in a buried valley aquifer in Butler County, Ohio. The application illustrates the role of objective function, resource constraint, mass transport processes, and hydrogeologic setting in groundwater quality monitoring network design.

1. INTRODUCTION

The sampling of spatially distributed hydrogeologic variables is of importance for aquifer characterization. Of significant current interest is the problem of sampling for subsurface pollution monitoring [Everett, 1980; Everett *et al.*, 1982; Rajagopal, 1986; U.S. Environmental Protection Agency, 1986; Gilbert, 1987; Loaiciga, 1988]. The objective of this work is to develop a methodology for the design of groundwater quality sampling networks. The design of a monitoring network consists of defining the number, locations, and sample pattern of sampling sites [Olea, 1984]. When temporal sampling is relevant, the sampling plan must specify the sampling frequency as well. Therefore in its greatest generality, a sampling plan determines where and how often to sample. The sampling plan yields information that meets specified criteria set by a decision maker. The determination of an optimal sampling plan is posed as a mixed integer programming (MIP) problem. This paper treats alternative formulations of the MIP problem corresponding to cases where there is spatial hydrogeologic variability and when spatial-temporal effects are important. The MIP approach to network design is illustrated with an application to groundwater quality monitoring at the Butler County landfill, Butler County, Ohio.

There are three approaches to the design of groundwater quality networks: (1) optimization, (2) simulation, and (3) variance reduction. Because of the complexities of monitoring subsurface pollutants, the field of network design is a developing one. There are, in fact, only two published articles that used analytical optimization in the selection of sampling sites, i.e., those by Hsueh and Rajagopal [1988] and Loaiciga [1988]. Neither one of these two studies, however, considered the boundary value problem of mass transport in the formulation of the sampling design problem, as is done in this work. Incidentally, a somewhat related approach based on statistical analysis and the one-dimensional advection equation has been presented by Knopman and Voss [1988] that has applications to the choice of sampling locations and the sampling times. Their method,

however, is of limited scope because of the simplified one-dimensional nature of the problem treated. Recent papers by Meyer and Brill [1988] and Massmann and Freeze [1987a, b] approached the problem of sampling well location for groundwater quality monitoring via simulation, in contrast to the analytical optimization formulation adapted in this paper. Whereas the analytical approach is based on a derivation of the first two moments (mean and covariance) of spatial-temporal solute concentration, the simulation approach attempts to reproduce the statistical behavior of solute distribution by repeated generation of synthetic plume realizations. The papers by Meyer and Brill [1988] and Massmann and Freeze [1987a, b] did not address the specification of (temporal) sampling frequency. In this sense, their interpretation of groundwater quality monitoring network design differed from the more general case considered herein. The variance reduction approach [Rouhani, 1985] is an iterative technique. It has not been applied to groundwater quality monitoring network design but it has potential for future applications. The variance-reduction approach adds the sampling site that contributes most to the reduction of the variance of estimation error (of the variable of interest) associated with a set of established sampling locations. Additional sampling sites are added, one at a time, until the variance of estimation cannot be further reduced, or when the additional gain in statistical accuracy is outweighed by other constraints, e.g., economic.

Various other approaches to sampling hydrological variables are available in the literature. Of particular relevance to this work are the MIP method of Hsu and Yeh [1989], and the integer programming approach by Carrera *et al.* [1984] for sampling aquifer properties. Olea [1984] presented a method to select the best pattern and sample density that would meet a specified average standard error or maximum standard error of estimation of spatial variables. The method was based on a nearest-neighbor analysis of two dimensional point distributions of spatial variables. Some other interesting applications of geostatistical methods to hydrologic network design can be found in the works by Bastin *et al.* [1984], Fiering [1965], and Bras and Rodriguez-Iturbe [1985].

Copyright 1989 by the American Geophysical Union.

Paper number 89WR00769.
0027-1397/89/89WR-00769\$05.00

2. SOME CONSIDERATIONS ON INDEPENDENT AND DEPENDENT OBSERVATIONS

2.1. A Classical Sampling Result

The classical approach to statistical sampling [Cochran, 1977] considers random variables that are independent and identically distributed (iid). If, in addition, the variables are normally distributed, a fundamental result specifying the optimal sample size n^* required to estimate the population mean from the sample mean is given by (see details in the appendix)

$$n^* = (z_{\alpha/2})^2 \sigma^2 / E^2 \quad (1)$$

(n^* is rounded to the nearest, larger, integer) in which $z_{\alpha/2}$ is a standard normal variable such that $\alpha/2$ of the probability mass follows to the right of $z_{\alpha/2}$; α is the significance level (usually $\alpha = 0.05$ in which case $z_{\alpha/2} = 1.96$); σ^2 is the population variance of the variable being sampled; and E is the specified tolerable error incurred in estimating the population mean from the sample mean. For α specified, (1) indicates that (1) for fixed tolerable error E , an increase in the variance σ^2 implies a larger sample size, and (2) for fixed variance, a smaller tolerable error requires a larger sample size. The effect of an increase in uncertainty (represented by the variance) and in the accuracy of results (specified by the tolerable error) is to increase the sample size. The relevance of (1) to MIP formulations is discussed in detail in section 4.

In spite of the usefulness of (1) in pointing out basic tendencies of the sample size, it is of limited value when the random phenomenon under consideration has a statistical dependence structure, e.g., when sampled observations are positively correlated, as is the case with hydrogeological variables of interest. Another factor of critical importance not considered in (1) concerns resource availability. Resource constraints and spatial (and possibly temporal) correlation must be incorporated in the selection of a sampling plan as shown below.

2.2. The Effect of Spatial Correlation

Suppose that the concentration of a chemical species in groundwater solution is to be measured at a number n^* of monitoring wells. Assume that the concentrations at the monitoring wells are iid so that the n^* well locations can be randomly selected (see subsection 3.1 for a discussion on potential sample patterns). If, in addition, resources are not limiting, the number of observations n^* is obtained from (1). Let us introduce the following proposition: n^* is equal to or less than the number, n' , of observations required to meet a specified tolerable error when the observations are positively correlated, i.e.,

$$n^* \leq n' \quad (2)$$

(a proof is presented in the appendix). Equation (2) indicates that using (1) when it is not applicable would lead us to err on the unsafe side, since (1) calls for less sampling sites than what is actually required to meet a specified tolerable error of estimation. Equation (1) provides a lower bound to the number of required sampling sites when the observations are statistically dependent and positively correlated.

3. OPTIMAL SAMPLING PLAN: TIME INDEPENDENT CASE

3.1. General Considerations

Suppose that a set of N locations is specified as the maximum number of potential sampling sites from which to draw a sampling plan. In groundwater monitoring studies some of the specified locations could be already existing wells. The possible locations of the N sampling sites are determined by accessibility, cost considerations, environmental factors, hydrogeological setting, or by theoretical considerations relevant to experimental design [Everett, 1980; Everett *et al.*, 1982; Barcelona *et al.*, 1983, 1985; Olea, 1984; Canter and Knox, 1985]. For example, when monitoring a groundwater contaminant plume by wells arranged in a regular grid, the geometrical configuration of the well locations and the plume's shape determine the adequacy of the grid configuration. Gilbert [1987] has tabulated the probability that the plume is missed by the wells as a function of grid and plume geometry. As a second example relevant to the definition of the N potential sites layout consider a data collection program for hydraulic conductivities. Suppose that the correlation length scale of hydraulic conductivity is l meters (i.e., the distance beyond which observations of hydraulic conductivity become uncorrelated is equal to l). If the potential sampling sites are all separated from each other by a distance exceeding l , the observations will be uncorrelated. If, in addition, there are no resource constraints and hydraulic conductivity (or, more likely, its logarithm) has a normal distribution over the section of aquifer being sampled, then, the optimal sample mean is given by (1). The distribution of the n^* sampling sites amongst the N possible locations (there are $N!/(n^*!(N - n^*)!)$ possible well distributions) can be randomly selected in this case. It is assumed in the following that the N potential well locations are specified prior to the formulation of the MIP problem for sampling plan selection. An example is given in section 7.

3.2. The Spatial Model

The spatially distributed hydrogeologic variable, $z(\mathbf{x})$, is expressed as the sum of a mean, $m(\mathbf{x})$, and a zero mean, second-order stationary, error term $v(\mathbf{x})$, i.e.,

$$z(\mathbf{x}) = m(\mathbf{x}) + v(\mathbf{x}) \quad (3)$$

in which $\mathbf{x}^T = (x_1, x_2, x_3)$ is a (row) location vector of coordinates x_r , $r = 1, 2, 3$. For the sake of argument, suppose that the mean at a location $\mathbf{x}_i$, denoted by m_i , is given by the polynomial expression

$$m_i = \sum_{k=1}^L f_k^i \beta_k \quad (4)$$

in which the β_k are parameters and the f_k^i are known functions of the spatial coordinates of the i th sampling site. In general, the mean can be constant (i.e., $L = 1$, $f_k^i = m$ (constant), and $\beta_1 = 1$, in (4), nonlinear in the parameters, or governed by a spatial deterministic process as shown in section 6. The polynomial model of (4) is oftenly used in universal kriging of spatial variables [Loaiciga *et al.*, 1988].

In addition to the mean, the spatial model specification includes a covariance structure (the covariance between any two locations $\mathbf{x}_i$ and $\mathbf{x}_j$ is denoted by σ_{ij} , when $i \neq j$, and by σ_i^2 when $i = j$). When modeling a physical process within a

stochastic framework, the covariance structure is determined by the governing equations relevant to the process of interest. For example, when modeling the spatial distribution of groundwater solutes, as is the case in this study, the covariance structure σ_{ij} is determined by the advection-dispersion equation governing mass transport. Details about the mean (m_i) and covariance (σ_{ij}) determination are deferred to section 6, where model specification is addressed at length.

In the analysis of spatial variables it is sometimes convenient to use the variogram instead of the covariance to describe the statistical structure of the spatial variable. The variogram is simply the variance of the difference ($z(\mathbf{x}_i) - z(\mathbf{x}_j)$) between the observations at locations i and j . In this work the covariance will be used. This is justified, since hydrogeologic variables have finite second moments, making the covariance and variogram representation of spatial statistical dependence equivalent. The models presented below (see section 7) can be readily modified in terms of the variogram according to well-established statistical relationships (see, for example, *de Marsily* [1986]).

The spatial model of (3) is suitable for hydrogeologic variables that are time independent. An important example is hydraulic conductivity that can be assumed time independent in most field conditions where groundwater physico-chemical properties do not vary significantly with time.

3.3. Optimization Criterion and Constraints

Let y_i ($i = 1, 2, \dots, N$) be a binary variable that takes the value 1 or 0 depending on whether or not an observation is made at the i th sampling site. Let λ_i ($i = 1, 2, \dots, N$) be a spatial interpolating weight assigned to the i th observation, z_i , when constructing an estimator of the hydrogeological variable $z(\mathbf{x}_0)$ at location $\mathbf{x}_0$. The estimator $\hat{z}(\mathbf{x}_0)$ is a linear combination of the observations taken at the developed sites, i.e.,

$$\hat{z}(\mathbf{x}_0) = \sum_{i=1}^N \lambda_i y_i z_i \quad (5)$$

Note that while y_i is binary ($y_i = 0$ or 1), λ_i is real and continuous, hence the denomination mixed-integer problem adopted in this work; $\hat{z}_0$ (and the observation z_i) can be either a point value or a spatial average. For example, if a pump test is conducted in an homogeneous and isotropic aquifer under radial flow conditions, the averaging domain for the calculated transmissivity is a cylindrical section of aquifer of radius equal to the (finite) zone of influence of pumping and a height equal to the fully penetrating depth of the well. The process of averaging is called regularization and it is well described in various works. In the context of hydrogeologic sampling the paper by *Carrera et al.* [1984] presented a clear exposition of this theme.

The objective of the MIP problem is to minimize the variance of the estimation error $\hat{z}_0 - z_0$; i.e.,

$$\text{minimize}_{\lambda_i, y_i} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j \sigma_{ij} - 2 \sum_{i=1}^N \lambda_i y_i \sigma_{i0} + \sigma_0^2 \quad (6)$$

in which σ_0^2 is the variance of z_0 and σ_{i0} denotes the covariance between z_i and z_0 ($i = 1, 2, \dots, N$).

In addition to having minimum variance of estimation error the estimator $\hat{z}_0$ is required to be unbiased, i.e., $E(\hat{z}_0) = m_0$, where m_0 is the mean or expected value of z_0 ($z_0 = z(\mathbf{x}_0)$). If the mean is modeled by the polynomial mean in (4) the unbiasedness condition becomes

$$\sum_{i=1}^N \lambda_i y_i f_k^i = f_k^0 \quad k = 1, 2, \dots, L \quad (7a)$$

Equation (7) simplifies to

$$\sum_{i=1}^N \lambda_i y_i = 1 \quad (7b)$$

when the mean of the hydrogeologic variable is a constant over the spatial domain of interest.

There is a resource constraint that limits the number of sampling sites. Suppose that the available budget is R and that the cost of sampling the i th site is C_i ($i = 1, 2, \dots, N$). The resource constraint becomes

$$\sum_{i=1}^N y_i C_i \leq R \quad (8)$$

The variable y_i is binary, i.e.,

$$y_i = (0 \text{ or } 1) \quad i = 1, 2, \dots, N \quad (9)$$

Equations (6)–(9) constitute the MIP problem for determining a sampling site for spatially distributed variables.

4. SIMPLIFIED FORMULATIONS OF THE MIP PROBLEM

4.1. The Effect of Lack of Correlation

Suppose that the mean of the hydrogeologic variable is constant and that the observations at different locations are uncorrelated. Lack of correlation occurs when the distances separating the sampling sites exceed the correlation length-scale of the hydrogeologic variable. The MIP problem simplifies to

$$\text{minimize}_{\lambda_i, y_i} \sum_{i=1}^N \lambda_i^2 y_i \sigma_i^2 \quad (10)$$

(note that $y_i^2 = y_i$) subject to the unbiasedness condition

$$\sum_{i=1}^N \lambda_i y_i = 1 \quad (11)$$

and to the resource constraint

$$\sum_{i=1}^N y_i C_i \leq R \quad (12)$$

The solution of the problem embodied by (10)–(12) can be obtained by inspection and it consists of sampling the j th sampling site associated with the smallest variance σ_j^2 . Therefore $\lambda_j = y_j = 1$ and all other λ and y are equal to zero (it is assumed that $C_i \leq R$, $i = 1, 2, \dots, N$). If the variances are all equal to each other, i.e., $\sigma_i^2 = \sigma^2$, $i = 1, 2, \dots, N$, the solution to (10)–(12) is obtained by choosing at random one

of the sampling sites, say, the i th one, and setting $\lambda_i = y_i = 1$. All other λ and y are equal to zero.

4.2. Estimating the Population Mean

Assume that the observations z_i are iid with mean μ and variance σ^2 . The estimator of μ is the sample mean

$$\bar{z} = \sum_{i=1}^N \lambda_i y_i z_i \quad (13)$$

where

$$\lambda_i = 1 / \left(\sum_{i=1}^N y_i \right) \quad i = 1, 2, \dots, N \quad (14)$$

The objective is to minimize the variance of $\bar{z}$ subject to the unbiasedness (equation (11)) and resource (equation (12)) constraints. The variance of $\bar{z}$ is given by

$$V(\bar{z}) = \sigma^2 / \left(\sum_{i=1}^N y_i \right) \quad (15)$$

The minimization of (15) (subject to the unbiasedness and resource constraints) is obtained by inspection. Define a permutation i_k ($k = 1, 2, \dots, N$) of the N location indices for the sampling sites so that the cost of sampling associated with each of the locations are arranged from smallest to largest, i.e., $C_{i_1} \leq C_{i_2} \leq \dots \leq C_{i_N}$. The optimal sampling plan is to choose n_1 locations where n_1 is the largest positive integer such that

$$\sum_{k=1}^{n_1} C_{i_k} \leq R \quad n_1 \leq N \quad (16)$$

In other words, the set of n_1 locations associated with the n_1 lowest sampling costs constitute the optimal solution. If the sampling cost is constant ($C_i = C$, $i = 1, 2, \dots, N$) the n_1 locations can be randomly selected subject to the constraint $n_1 \cdot C \leq R$, provided that n_1 does not exceed N .

The minimization of the variance of the sample mean (equation (15)) subject to the unbiasedness (equation (11)) and resource (equation (12)) constraints when the variances σ_i^2 , $i = 1, 2, \dots, N$, are unequal introduces some complications. Note that the variance of the sample mean becomes

$$V(\bar{z}) = \left(1 / \sum_{i=1}^N y_i \right)^2 \sum_{i=1}^N y_i \sigma_i^2 \quad (17)$$

Define a permutation i_k ($k = 1, 2, \dots, N$) of the N location indices so that the variances of the observations are arranged from smallest to largest, i.e., $\sigma_{i_1}^2 \leq \sigma_{i_2}^2 \leq \dots \leq \sigma_{i_N}^2$. Define the function

$$f(n) = \sum_{k=1}^n \sigma_{i_k}^2 / n^2 \quad (18)$$

within the interval $n = 1, 2, \dots, P$, where P is the largest positive integer satisfying

$$\sum_{k=1}^P C_{i_k} \leq R \quad P \leq N \quad (19)$$

The value, say, n_2 , that minimizes the function $f(n)$ in (18) defines the optimal set of sampling locations that correspond to the n_2 lowest variances.

4.3. The Effect of Changing the Objective Function

It has been shown in the models of subsection 4.2 that there is a tradeoff between the objective of minimizing the estimation variance (that calls for enlarging the sample size) and the need to keep the sampling plan within allowable resource availability. Next, let us examine the effect of changing the objective function to be the minimization of the sampling cost,

$$\text{minimize } \sum_{i=1}^N C_i y_i \quad (20)$$

subject to the unbiasedness condition (11) and to a restriction on the magnitude of the tolerable error E incurred in estimating the population mean by the sample mean (equation (1)), i.e.,

$$\sum_{i=1}^N y_i \geq z_{\alpha/2}^2 \sigma^2 / E^2 \quad (21)$$

The resource constraint (12) is also imposed to keep the sampling plan within budget.

Let

$$Q = (z_{\alpha/2})^2 \sigma^2 / E^2 \quad Q \leq N \quad (22)$$

(Q is rounded to the nearest, larger, integer). The solution to (20) subject to the aforementioned constraints (equations (11), (12), and (21)) is found by inspection. Recall the permutation i_k ($k = 1, 2, \dots, N$) of location indices for which the sampling costs at the N locations are arranged from smallest to largest, i.e., $C_{i_1} \leq C_{i_2} \leq \dots \leq C_{i_N}$. The optimal sampling plan is given by the set of locations whose indices are $i_1, i_2, \dots, i_Q$, i.e., by the Q locations associated with the Q lowest sampling costs. It is evident that by minimizing the cost of the sampling plan the smallest possible number Q , of sampling locations (as dictated by the specified tolerable, E) is chosen. In contrast, when the criterion is to minimize the estimation variance, the largest possible number of sampling locations (as dictated by the resource constraint) is called for. In either case, the sampling cost and/or the variance at each location determine which locations are included in the sampling plan.

5. OPTIMAL SAMPLING PLAN: TIME-DEPENDENT CASE

5.1. Time-Dependent Spatial Model

There are hydrogeological variables that vary over time, such as the concentration of a chemical species dissolved in groundwater. A sampling plan for time-dependent variables consists of the locations and times at which observations of the hydrogeological variable are to be taken. The observation taken at the i th location ($i = 1, 2, \dots, N$) and at the t th time ($t = 1, \dots, T$) will be denoted by z_{it} . The purpose of the sampling plan is to gather data in order to estimate and make statistical inferences about the value of the hydrogeological variable at some location x_0 and at time T , z_{0T} . In general, the observations z_{it} and the variable z_{0T} can be spatial averages rather than a point value.

The estimator of z_{0T} , denoted by $\hat{z}_{0T}$, is defined by the following weighted sum of observations z_{it} :

$$\hat{z}_{0T} = \sum_{i=1}^N \sum_{t=1}^T \lambda_{it} y_{it} z_{it} \quad (23)$$

in which y_{it} is a binary decision variable equal to 1 or 0. It takes the value 1 when an observation is taken at the i th sampling site at time t . Otherwise, it takes the value 0; λ_{it} is a weighting coefficient assigned to the observation z_{it} . While y_{it} is binary, λ_{it} is a continuous, real variable.

The mean of z_{it} is denoted by m_{it} and it represents the spatial trend of the hydrogeologic variable at time t . In addition, the complete spatial-temporal model specification requires a covariance structure. When the regionalized variable is governed by a physical process then the mean m_{it} and covariance σ_{ij}^{rs} between observations z_{it} and z_{js} , depend upon the process in question. For the case of interest in this work, i.e., mass transport, the mean and covariance are determined by the advection-dispersion equation as shown in section 6.

5.2. Time-Dependent MIP Problem

The time-dependent MIP problem for sampling plan selection represents a generalization of the time-independent formulation discussed in subsection 3.3. The objective is to minimize the variance of estimation error $\hat{z}_{0T} - z_{0T}$ subject to unbiasedness (i.e., $E(\hat{z}_{0T}) = m_{0T}$) and resource constraints. The formulation of the time-dependent MIP problem parallels that given in subsection 3.3 and it is readily shown to be given by the following expressions:

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^N \sum_{j=1}^N \sum_{r=1}^T \sum_{s=1}^T \lambda_{ir} \lambda_{js} y_{ir} y_{js} \sigma_{ij}^{rs} \\ & - 2 \sum_{i=1}^N \sum_{t=1}^T \lambda_{it} y_{it} \sigma_{i0}^{tT} + \sigma_{00}^2 \end{aligned} \quad (24)$$

subject to

$$\sum_{i=1}^N \sum_{t=1}^T \lambda_{it} y_{it} m_{it} = m_{0T} \quad (25)$$

$$\sum_{i=1}^N \sum_{t=1}^T y_{it} C_{it} \leq R \quad (26)$$

in which σ_{i0}^{tT} represents the covariance between z_{it} and z_{0T} , σ_{00}^2 is the variance of z_{0T} , and C_{it} is the cost of sampling at the i th location at time T . An application of the general MIP problem given in (24)–(26) is presented in section 7.

6. MODEL SPECIFICATION IN NETWORK DESIGN

The time-independent and time-dependent MIP formulations were presented in sections 3 and 5 respectively. It was assumed there that the model parameters, i.e., the means and covariances of the hydrogeologic variable were known. In practice, those parameters need to be estimated prior to or simultaneous to the network design problem solution. The procedure for parameter specification is illustrated in this

section with the estimation of dispersivities and covariance parameters in a groundwater quality problem. Subsequently, the spatial and temporal distribution of mean chemical concentration can be readily derived and input to the MIP formulations for network design. The reader is referred to *Loaiciga and Marino* [1987a, b] for an example of time-dependent groundwater flow parameter identification (an extensive review is available in the work by *Willis and Yeh* [1987]).

6.1. The Mean Concentration Equation of Miscible Displacement

Let us start with the equation describing transport of a nonreactive and conservative tracer (i.e., neglecting adsorption, chemical reactions, radioactive decay, etc.; see, for example, *Bear* [1972])

$$\partial c / \partial t = \partial / \partial x_i (d_{ij} \partial c / \partial x_j) - v_i \partial c / \partial x_i \quad (27)$$

in which c denotes solute concentrations, v_i is the seepage velocity in the i th Cartesian direction ($i = 1, 2, 3$), and d_{ij} is the local coefficient of hydrodynamic dispersion. The index notation is used in (27). We seek to derive an equation for the mean concentration. For this purpose, the concentration and velocity are decomposed into the sum of a mean and a (zero mean) stationary random component, i.e., $c = C + c'$ and $v_i = V_i + v'_i$, where the primed variables are the random components. To derive the mean concentration equation it is assumed that (1) d_{ij} depends on the mean seepage velocities through the equation [*Scheidegger*, 1961] $d_{ij} = a_{ijkl} V_k V_l / V$, where V is the magnitude of the seepage velocity and a_{ijkl} is the dispersivity tensor and (2) the dispersive flux is given by the expression [*Gelhar and Axness*, 1983]

$$\overline{v'_i c'} = -V A_{ij} (\partial C / \partial x_j) \quad (28)$$

where the overbar represents the expectation operation and A_{ij} is the macrodispersivity tensor. It is emphasized that the *Scheidegger* theory of dispersion is applicable to statistically homogeneous media (see, for example, the equivalent definition of nonuniform homogeneity by *Freeze* [1975]) and in general the mean seepage velocities may vary from point to point but the dispersivity tensor is assumed constant. The same interpretation is taken in (28), i.e., although the mean seepage velocity may vary in space (due to, say, effect of boundary conditions and flow domain geometry) the macrodispersivity tensor is constant. Using the mean perturbation decomposition of the concentration and seepage velocities as well as assumptions (1) and (2) stated above in (27) lead, after taking expectations, to the mean equation of tracer concentration in terms of macroscopic variables

$$\partial C / \partial t = \partial / \partial x_i (D_{ij} \partial C / \partial x_j) - V_i \partial C / \partial x_i \quad (29)$$

in which the macroscopic hydrodynamic dispersion coefficient D_{ij} can be readily shown to be

$$D_{ij} = V A_{ij} + d_{ij} \quad (30)$$

It is implicit in (30) that the sum of the macrodispersivity (A_{ij}) and local dispersivity (d_{ij}) tensors must be over the same coordinate axes. *Gelhar and Axness* [1983] provided a detailed analysis of A_{ij} in (30). *Neuman et al.* [1987] exam-

ined the relative magnitudes of A_{ij} and d_{ij} using the Peclet number as a basis for the analysis.

When the porous medium is isotropic the components of the macrodispersivity tensor are [Gelhar and Axness, 1983]

$$A_{11} = \sigma_f^2 l / \gamma^2 \quad (31)$$

$$A_{22} = A_{33} = (\sigma_f^2 a_L / 15 \gamma^2) (1 + 4a_T / a_L) \quad (32)$$

in which σ_f^2 is the variance of log hydraulic conductivity, l represents the correlation length scale of loghydraulic conductivity, a_L and a_T are the longitudinal and transversal local dispersivities, respectively, and $\gamma = 1 + \sigma_f^2/6$. The components of the local dispersivity tensor d_{ij} are $d_{11} = a_L V$; $d_{22} = d_{33} = a_T V$. It is being implied that the main principal axis of d_{ij} (along the $i = 1$ direction) coincides with the direction of the mean seepage velocity and with the orientation of the main principal direction of the macrodispersivity tensor A_{ij} . The former expressions for the macrodispersivity (A_{ij}) and the local dispersivity (d_{ij}) tensor define the hydrodynamic dispersion coefficient D_{ij} in (30) when the medium is isotropic (which is the case of interest in this study for the purpose of application). The reader may consult Gelhar and Axness [1983] or Neuman *et al.* [1987] for the relevant macrodispersivity expressions in anisotropic media.

Equation (30) is a partial differential equation describing the spatial-temporal behavior of mean concentration. Once the parameters a_L , a_T , σ_f^2 , and l are estimated, (30) can be numerically discretized and the mean or trend of concentration is readily obtained by simulation of the numerical model.

6.2. Parameter Estimation and Model Specification

The estimation of parameters in this work follows the maximum likelihood (ML) approach by Loaiciga and Mariño [1986]. For completeness, their approach is briefly summarized herein. The distribution of concentrations is modelled according to (27). The first step of ML estimation to discretize (27) numerically to yield the following stochastic first-order linear system [Loaiciga and Marino, 1986]:

$$A_c = Bc_{t-1} + Dx_t + w_t \quad (33)$$

in which c , x , and w are vectors of concentrations, boundary conditions, and modeling errors, respectively. A , B , and D are suitable dimensioned matrices whose elements are functions of the longitudinal (a_L) and transversal (a_T) dispersivities. The error vector w_t is assumed zero mean with constant covariance W and uncorrelated over time. The covariance W is assigned a specific structure, e.g.,

$$w_{ij} = \theta \exp(-h_{ij}/v) \quad (34)$$

where w_{ij} is the ij th element of W , θ is an unknown parameter, v is an unknown correlation scale, and h_{ij} is the distance between the i th and j th locations. The exponential model for the error term covariance is widely used in hydrogeology [de Marsily, 1986]. The uncorrelated temporal structure of the error term does not constitute a loss generality, since it is straightforward to show [Loaiciga and Marino, 1987a, b] that by a simple reparameterization a correlated, autoregressive, error term is easily redefined as an uncorrelated random vector. Concentration observations must be available at specific locations. These measurements are taken at existing wells that are part of the optimized

monitoring network. The vector of measurements, z , is related to the concentration vector, c , by the equation

$$z_t = Gc_t \quad (35)$$

in which G is a matrix defined by the geometric configuration of the numerical discretization grid associated with (33) and by the location of the measurement points. The dimension of the measurement vector z is less than, or at most equal to, the dimension of c . The linear system (33) and the measurement or observation process (35) define the model upon which ML estimation is based, once the additional assumption that the error term w is Gaussian is introduced.

The parameter set to be estimated by maximum likelihood consists of the longitudinal and transversal dispersivities (a_L and a_T , respectively), and the parameters θ and v associated with the covariance of the error term w (see equation (34)). The variance σ_f^2 of log hydraulic conductivity (see equations (31) and (32)) can be estimated from field measurements of hydraulic conductivity, after transforming these values logarithmically and applying the moment estimator for the variance (P. F. Hudak and H. A. Loaiciga, Contaminant transport simulation and parameter estimation in a buried valley aquifer in southwest Ohio, submitted to *Groundwater*, 1988). The correlation length scale l can also be inferred from log hydraulic conductivity data as shown by Hudak and Loaiciga (1988). The reader is referred to Loaiciga and Marino [1986] for details of the ML parameter estimation procedure involving the dynamic model (33) and (35) (an application is found in section 7). Note that the covariance of concentration measurements z_{it} and $z_{js}(\sigma_{ij}^s)$ follows from (35), i.e., σ_{ij}^s is the ij th element of the following covariance (G' denotes the transpose of G),

$$\text{Cov}(z_r, z_s) = G \text{Cov}(c_r, c_s) G' \quad (36)$$

where $\text{Cov}(c_r, c_s)$ is readily obtained from the linear first-order system (33) (see details in the work by Anderson [1978] and Loaiciga and Mariño [1985]).

Upon the estimation of the parameters a_L , a_T , l , and σ_f^2 , the means of the concentration observations can be obtained from the simulation of the discretized form of (29). Such means, m_{it} , for all i , t are given by the elements of the following mean observation vector m_t :

$$m_t = GC_t \quad (37)$$

where C_t is the mean concentration vector at time t obtained from the simulation of the numerically discretized form of (29). Once the means (m_{it}) and covariances (σ_{ij}^s) of concentration measurements are derived, the time-dependent MIP problem for optimal sampling plan (equations (24)–(26)) is completely specified. The next section presents an application of the MIP formulation for sampling plan optimization.

7. APPLICATION

The time-dependent MIP formulation (24)–(26) was used to determine the optimal sampling for chloride plume distribution at the Butler County landfill, located 10 miles north-east of the City of Hamilton, Ohio, on the floodplain of the Great Miami River (Figure 1). The landfill served as a disposal site for municipal solid wastes mainly. During the period of landfilling, development proceeded in two phases of operation. The phase 1 portion of the landfill (see area 1 in Figure 1) was active from January 6, 1971 to April 3, 1975.

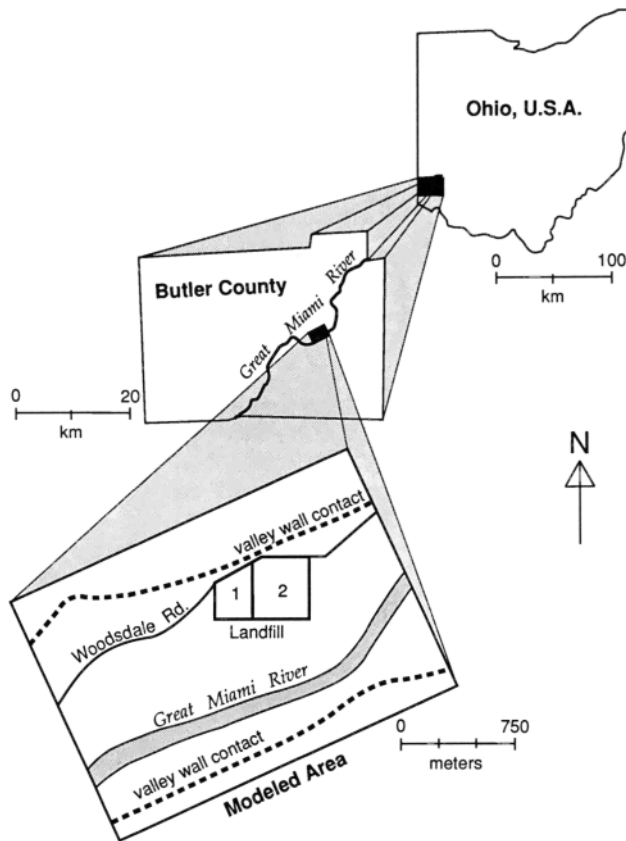


Fig. 1. Location of study area.

Refuse was placed in the phase 2 area (Figure 1) from April 3, 1975 to January 1, 1985 when the landfill operation was terminated.

7.1. Hydrogeologic Setting

The aquifer consists of glaciofluvial deposits comprised predominantly of sand and gravel, which form a highly productive buried valley underlying the landfill. Silt and clay deposits are present in lesser amounts in thin, laterally discontinuous lenses. The unconsolidated deposits range from 6 to 12 m in thickness and overlie highly impermeable

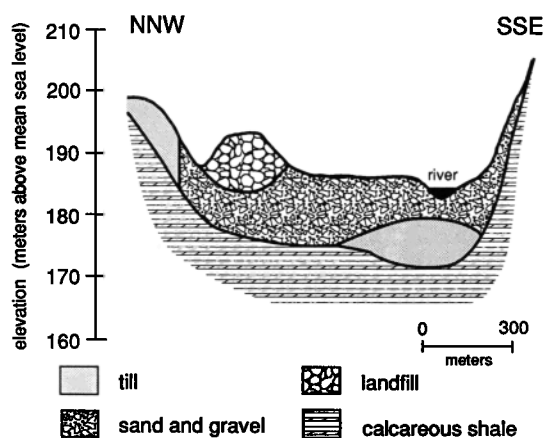


Fig. 2. Generalized NNW-SSE geologic cross section through modeled area.

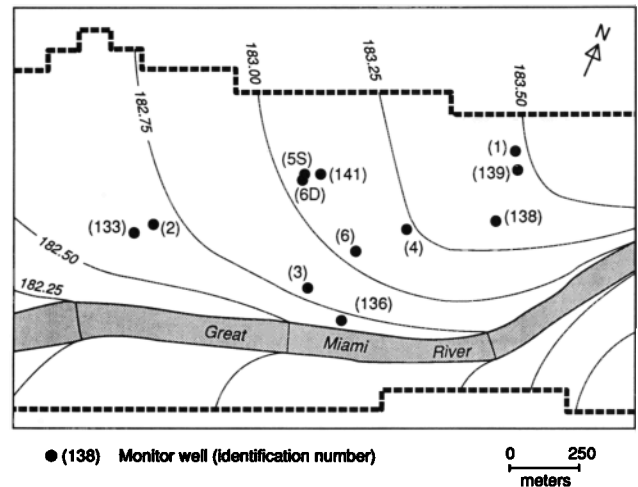


Fig. 3. Monitor well locations and steady state head distribution (solid contours in meters above mean sea level).

bedrock. Figure 2 depicts a generalized cross section of the study area. The areas to the immediate north-northwest (NNW) and south-southeast (SSE) of the study area are prominent upland surfaces underlain by relatively impermeable glacial till, ranging from 1 to 12 m in thickness which is underlain by shale bedrock. The NNW and SSE boundaries of the study area were adequately modeled as no-flow boundary conditions. On the contrary, no natural geologic boundaries are present transverse to the axis of the buried valley at the east-northeast (ENE) or west-southwest (WSW) margins of the modeled region. The Great Miami River flowing in the vicinity of the landfill has a strong (effluent) hydraulic connection with the groundwater flow regime. Under flood conditions the high river stage produces bank storage and reverses the direction of groundwater flow in the adjacent aquifer. Hudak and Loaiciga (1988) provided a detailed hydrogeologic description of the site under study.

7.2. Data Acquisition and Flow System Modeling

The experimental work conducted at the Butler County landfill has been reported by Hudak and Loaiciga (1988). A summary of their findings is presented herein. Figure 3 shows the existing set of 12 monitoring wells at the study site. Piezometric heads and groundwater samples were measured or collected biweekly from mid-June to late-November, 1987. In addition, falling head piezometer tests as well as analysis of grain size distribution were conducted to obtain experimental information on the hydraulic properties of the flow system at the study site.

The field and laboratory measurements of hydraulic conductivity and the measured piezometric levels were instrumental in deriving a steady state piezometric head distribution (shown in Figure 3) representative of the predominant flow regime at the site. The flow model calibration led to estimates of recharge rate (i.e., percolation to the water table) and hydraulic conductivity of 46 cm/year and 1.3×10^{-3} m/s, respectively. The hydraulic conductivity measurements resulted in an estimate of the loghydraulic conductivity variance of σ_f^2 equal to 1.51 and a correlation scale $l = 2.0$ m. The calibrated flow model was tested and found to be adequate for the modeling of groundwater flow velocities responsible for advective and dispersive transport pro-

TABLE 1. Parameter Estimation Results

Method	a_L , m	a_T , m	θ	v , m
Maximum likelihood	30.1 (8.9)	2.8 (0.72)	7.3 (1.9)	5.4 (2.0)
Calibration	53.4	5.3	NA	NA

Values in parentheses are the standard deviations. NA, value not available.

cesses. The flow system was represented as being essentially two dimensional (in the horizontal plane), in a homogeneous and isotropic medium (Hudak and Loaiciga, 1988).

7.3. Groundwater Quality and Parameter Estimation

Groundwater samples collected at the study site were analyzed for chloride concentration. A background concentration (i.e., concentration of uncontaminated groundwater) of approximately 20 ppm was inferred on the basis of chemical analysis. Chloride concentrations at the study site were relatively high and suitable for accurate measurement. In addition, chloride is a nonreactive tracer and hence adequate for modeling according to (27).

Table 1 shows the values of estimated parameters by the approach of Loaiciga and Marino [1986]; it also shows the values of dispersivities derived by Hudak and Loaiciga (1988) by calibrating a mass transport model. The ML estimates yielded dispersivity values that were within the same order of magnitude as those obtained by calibration, although the latter were approximately twice as large as the ML values. Since the ML is a statistical estimation method it provided the standard deviations of parameter estimators, that are shown within parentheses in Table 1. Once the estimated dispersivities and the hydraulic parameters, σ_f^2 and l , were determined, the mean concentration equation (29) was fully specified in terms of its parameters and discretized for simulation. A typical (i.e., quasi steady state)

distribution of mean chloride concentration is depicted in Figure 4.

The parameter estimates a_T , a_L , θ , v , σ_f^2 , and l enabled the complete specification of the observation means, m_{it} , and covariances, σ_{ij}^{rs} , required for the solution of the monitoring network design problem embodied in the MIP formulation of (24)–(26). This issue is addressed next.

7.4. Solution to the MIP Network Design Problem

Figure 5 shows the grid containing the possible well locations (excluding the existing wells shown in Figure 3). The potential well sites were chosen to be at the center of each of the active cells within the boundaries of the modeled area (an active cell is an internal cell which is neither a river node (R) or a constant head (H) and excluding those located within the landfill). Overall there are 203 active cells and hence 203 possible monitoring sites, in addition to the 12 existing wells shown in Figure 3.

The MIP model (see equations (24)–(26)) was tested by developing an optimal sampling plan that would minimize the variance of concentration estimation error along the cells bordering the northern margin of the river (shown as a cross-hatched region in Figure 6). The estimation region comprehends the area where groundwater flow lines intersect the river and hence constitutes the critical portion of the aquifer with respect to monitoring. The optimal sampling plan consists of selecting the location and the timing at which concentration measurements z_{it} must be collected in order to minimize the variance of concentration estimation error, at some time T , over the river margin shown in Figure 6. The sampling horizon is divided into biweekly time intervals $t = 1, 2, \dots, T'$, and concentration measurements can be taken at those discrete times only. For this application, the MIP network problem was solved for a duration of the sampling horizon equal to $T' = 26$ two-week periods (i.e., approximately one year). The time at which the concentration was estimated was set equal to $T = T'$, i.e., coincident with the last potential sampling time.

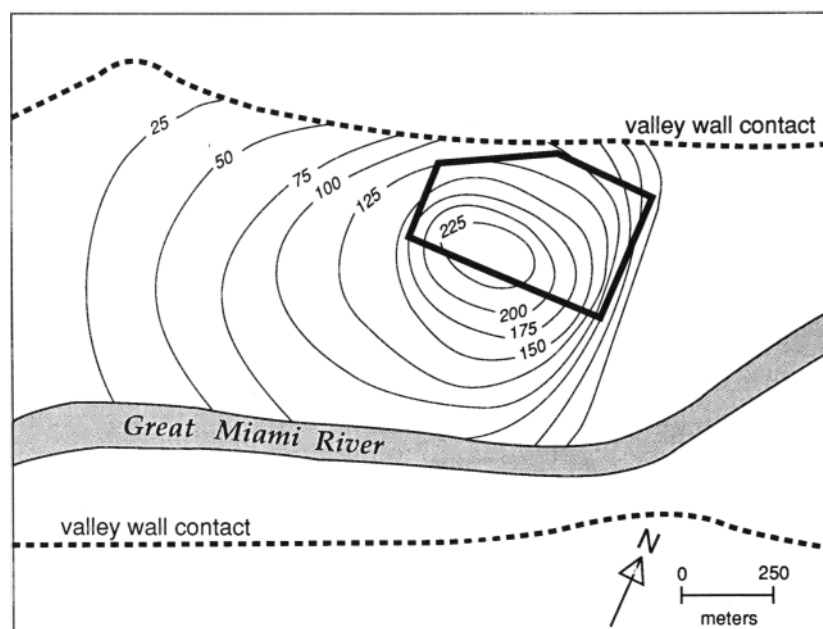


Fig. 4. Quasi steady state mean chloride concentration distribution (contours in ppm).

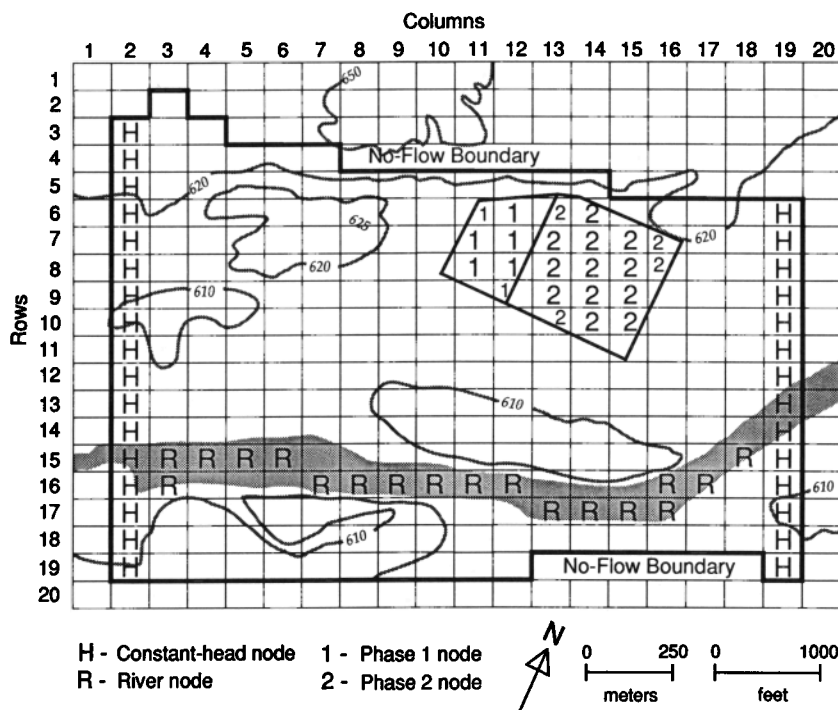


Fig. 5. Model representation.

Besides the means (m_{it}) and covariances (σ_{ij}^{rs}), the sampling costs (C_{it}) and budget (R) are required prior to the MIP problem (see equations (24)–(26)) solution. From an analysis of well installation, sampling and chemical analysis costs it was determined that the average sampling cost C_{it} could be assumed constant for all locations and sampling times, with $C_{it} = \$50/\text{sample}$, for all i, t . The total sampling budget was set equal to $R = \$100,000$. Because of the unconsolidated nature of the deposits which consist mainly of sand and fine gravel, it is inexpensive to install monitoring wells by exerting pressure on 2" drive points. The field sampling costs were due to personnel expenses to run a peristaltic pump for sample retrieval. The \$50/sample figure is a prorated value. Note that if initial well installation costs were relatively high,

then by a trivial modification of the resource constraint (26) new wells could be assigned a higher cost.

The solution of the MIP problem for the estimated means (m_{it}), covariances (σ_{ij}^{rs}), specified sampling cost (C_{it}), budget (R), and sampling horizon T' was obtained by the branch-and-bound algorithm [Fletcher, 1987], which has been applied previously by Carrera *et al.* [1984] to network design for groundwater flow monitoring. The details of the mixed integer solution by the branch-and-bound method are outlined in the work by Fletcher [1987] and will not be repeated herein. Suffice it to say, however, that the selection of sampling locations and times for the time-dependent MIP problem of section 5.2 is by no means trivial nor can it be deduced by simple inspection. The estimated average concentration over the area of interest (shown in Figure 6) was equal to 65 ppm chloride with a standard deviation of estimation error equal to 9 ppm. Figure 7 displays the optimized network (excluding the original, existing, wells) corresponding to $T' = 26$. There are 64 monitoring sites in Figure 7 (the total sampling plan includes also the twelve original wells) and their locations correspond to the cell center. Overall, the total number of sampling sites is equal to 76, with the original 12 being shown in Figure 3 and the other 64 wells in Figure 7. The optimized sampling plan yielded that each chosen monitoring site is to be sampled during the entire sampling period at times $t = 1, 2, \dots, T'$ (in this application the initial sampling time, $t = 1$, corresponded to the last field data collection date, i.e., to September 25, 1987). This uninterrupted sampling schedule is attributed to (1) the constant sampling cost at all sites and times and (2) the concentration distributions are quasi-steady (see Figure 4) with minor spatial and temporal changes due mainly to seasonal variations of groundwater recharge. The total sampling cost was equal to \$98,800 ($= 76 \times 26 \times 50$), barely short of the \$100,000 budget. This seems to indicate that the

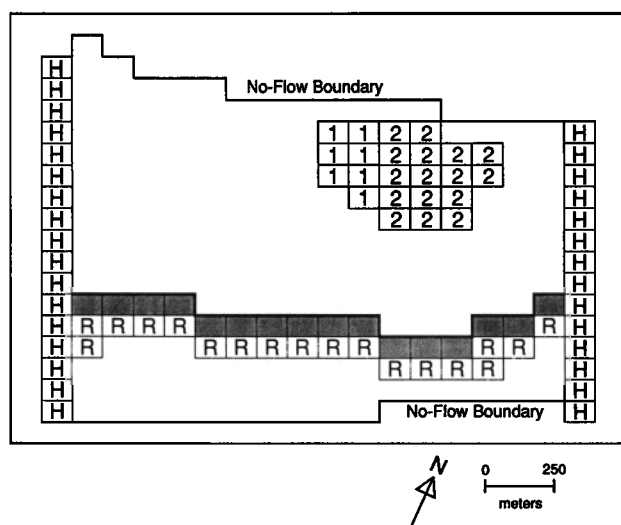


Fig. 6. Model representation and the estimation region (cross-hatched).

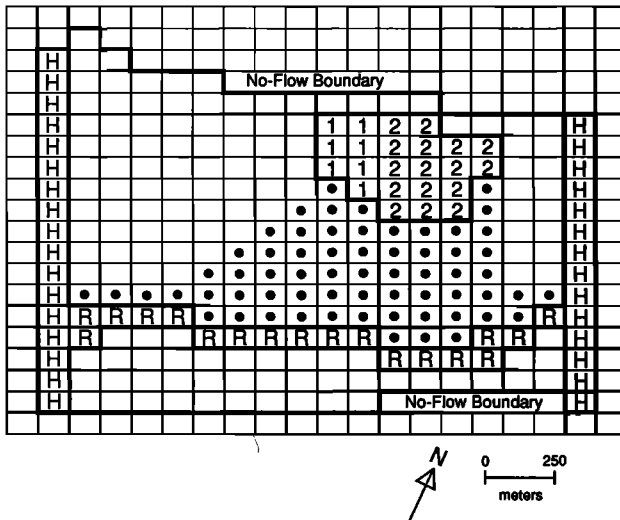


Fig. 7. Optimal monitoring sites (dots) distribution for sampling horizon $T^* = 26$.

optimal solution includes as many sampling locations and sampling times as the budget constraint permits, seeking to minimize the estimation variance, while satisfying the unbiasedness constraint (25).

The spatial distribution of wells is not surprising given that the mean advection-dispersion equation (29) constrains the network design problem. Notice, however, that the final well locations and sampling frequency needed to obtain the best statistical estimate of concentrations cannot be deduced by inspection a priori unless the mathematical programming problem (24)–(26) is fully solved. By doing so, the hydrogeologic characteristics of mass transport, statistical accuracy, and optimal resource allocation are all combined to yield the optimized sampling plan.

It is worth pointing out that the network design problem is dependent upon parameter estimates of hydrogeologic variables and the grid configuration chosen to conduct the numerical flow and concentration simulations. An analysis of the effect of hydrogeologic uncertainty in groundwater management modeling can be found in the work by *Loaiciga and Marino* [1987]. For the problem solved in this paper, the grid configuration is already optimized from hydrogeologic considerations studied in the work by *Hudak and Loaiciga* (1988).

8. CONCLUSIONS

In summary, a new approach to groundwater quality monitoring has been developed and successfully applied to a real hydrogeologic setting. This paper introduced time-independent and time-dependent MIP formulations for optimal hydrogeologic sampling. It was shown that the geostatistical approach can be coupled with statistical parameter estimation and field information and subject to the mass transport equations to provide a realistic representation of the spatial/temporal statistical structure of the field variable. The theoretical analysis and application of this work lead to the following conclusions.

1. Network design depends upon the hydrogeologic and statistical properties of the processes governing mass transport. It is necessary to explicitly introduce the governing advection-dispersion equations as constraints in the statisti-

cal design for proper specification of the first two moments of the regionalized variable, which show spatial/temporal variability.

2. Network design involves a two-stage process: (1) parameter estimation and (2) network optimization. It is important to have adequate field data to characterize the flow properties of the system and make a sound definition of the location of estimation area.

3. Resource constraints determine the optimal number of sampling locations and times, but the actual spatial distribution and temporal frequency is governed by the mean and covariance structure of the stochastic variable.

4. The groundwater quality network design problem indicates that the correlation of a stochastic variable governed by a partial differential equation must be developed from the governing equation itself. This will determine the spatial and temporal evolution of correlation. It is not possible to specify, extraneously, spatial/temporal covariance models for the concentration measurements.

5. The MIP approach yields optimal sampling plans (at least locally optimal), that satisfy proper statistical conditions of estimates (unbiasedness, minimum variance) as well as hydrogeologic and resource constraints. Rather than using an incremental, one step at a time search for a set of sampling locations, the MIP scans over the entire set of possible locations in search for a global (or in the worst case, for a local) optimal solution. Notice that the variance reduction approach to network design always yields suboptimal solutions because it scans for sampling sites from amongst subsets of the entire decision space.

6. Groundwater quality monitoring network design presents a variety of interesting optimization possibilities including time-dependent and time independent formulations. Correlation structure, objective function, statistical conditions, resource constraints, and hydrogeologic characteristics are shown to play important roles in the determination of an optimal sampling plan.

APPENDIX : PROOF OF EQUATION (1)

Suppose that n observations $z_1, z_2, \dots, z_n$ are independent and identically distributed observations drawn from a normal distribution with mean μ and variance σ^2 . Then the sample mean $\bar{z} (= \sum_{i=1}^n z_i/n)$ is normally distributed with expected value μ and variance σ^2/n . By normalizing $\bar{z}$ one obtains

$$Z = \frac{\bar{z} - \mu}{\sigma/\sqrt{n}} \quad (A1)$$

where Z is a standard normal variable with zero mean and unit variance. From (A1) the following probability statement holds:

$$P\left(-z_{\alpha/2} \leq \frac{\bar{z} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right) = 1 - \alpha \quad (A2)$$

from which it follows that

$$P(\bar{z} - z_{\alpha/2}(\sigma/\sqrt{n}) \leq \mu \leq \bar{z} + z_{\alpha/2}(\sigma/\sqrt{n})) = 1 - \alpha \quad (A3)$$

indicating that a $100(1 - \alpha)\%$ confidence interval for the mean μ is given by $\bar{z} \pm z_{\alpha/2}(\sigma/\sqrt{n})$ where the error, E_r , in estimating μ by $\bar{z}$ is

$$E = z_{\alpha/2}(\sigma/\sqrt{n}) \quad (A4)$$

Equation (1) follows at once from (A4) after solving for n .

Next, consider n observations $z_1, z_2, \dots, z_n$ that are normally distributed with mean μ and variance σ^2 and statistically dependent (hence correlated). This situation could be envisioned in a homogenous soil where the logarithm of hydraulic conductivity is a normal variable with constant mean and variance over some spatial domain. The sample mean $\bar{z}$ has expected value μ and variance $V(\bar{z})$:

$$V(\bar{z}) = (1/n^2) \left(n\sigma^2 + 2 \sum_{i=1}^n \sum_{j>1}^n \sigma_{ij} \right) \quad (\text{A5})$$

where σ_{ij} is the covariance between z_i and z_j . The covariance σ_{ij} can be rewritten as follows:

$$\sigma_{ij} = \sigma^2 \rho_{ij} \quad (\text{A6})$$

in which ρ_{ij} is the correlation coefficient. In the analysis of hydrogeological variables, which are generally positively correlated, we are interested in values of ρ_{ij} in the interval $0 \leq \rho \leq 1$, although, mathematically, $-1 \leq \rho \leq 1$. The covariance in (A5) can be rewritten as follows:

$$V(\bar{z}) = (\sigma^2/n^2) \left(n + 2 \sum_{i=1}^n \sum_{j>1}^n \rho_{ij} \right) \quad (\text{A7})$$

The sample mean can be standardized using the mean μ and the square root of $V(\bar{z})$ to yield the tolerable error

$$E = z_{\alpha/2}(V(\bar{z}))^{1/2} \quad (\text{A8})$$

Let ε be a positive and arbitrarily small real number such that $\rho_{ij} \geq \varepsilon$, for all i, j ($i \neq j$). Substituting the value ε for ρ_{ij} in (A7) leads to the following inequality:

$$V(\bar{z}) \geq (\sigma^2/n)(1 + (n-1)\varepsilon) \quad (\text{A9})$$

Squaring both sides of (A8) and using (A9) leads to the expression

$$E^2 \geq (z_{\alpha/2})^2 (\sigma^2/n)(1 + (n-1)\varepsilon) \quad (\text{A10})$$

solving for n in (A10) and recalling the expression for n^* in (1) yields

$$n \geq n^*(1 - \varepsilon)/(1 - n^*\varepsilon) \quad (\text{A11})$$

which is valid provided that $\varepsilon < (1/n^*)$. Letting $\varepsilon = 0$ (i.e., $\rho_{ij} \geq 0$) in (A11) and denoting n by n' (to indicate that n' represents the sample size when correlation is considered) results in

$$n' \geq n^* \quad (\text{A12})$$

with strict equality holding only if $\rho_{ij} = 0$ (for all $i, j, i \neq j$), in which case the problem reduces to the statistically independent case already presented in (1). Equation (A12) proves the result in (1).

Incidentally, if one assumes that the correlation coefficient is constant, i.e., $\rho_{ij} = \rho$ ($i, j = 1, 2, \dots, N$), in (A7) it is straightforward to show that

$$n^* = n'/[1 + (n' - 1)\rho] \quad (\text{A13})$$

which is the classical Yule's equation. Equation (A13) means that the equivalent number of independent observations (n^*) is only a fraction of the total number (n') of correlated observations when the correlation coefficient is equal to ρ .

REFERENCES

- Anderson, T. W., Repeated measurements of autoregressive processes, *J. Am. Stat. Assoc.*, 73(362), 371-378, 1978.
- Barcelona, M. J., J. P. Gibb, and R. A. Miller, A guide to the selection of materials for monitoring well construction and ground-water sampling, technical report, Ill. State Water Surv., Champaign, August 1983.
- Barcelona, M. J., J. P. Gibb, J. A. Helfrich, and E. E. Garske, Practical guide for groundwater sampling, *Ill. State Water Surv. Contract Rep. 372*, Champaign, 1985.
- Bastin, G., B. Lorent, C. Duque, and M. Gevers, Optimal estimation of the average areal rainfall and optimal selection of rain gauge locations, *Water Resour. Res.*, 20(4), 463-470, 1984.
- Bras, R. L., and I. Rodriguez-Iturbe, *Random Functions in Hydrology*, Addison-Wesley, Reading, Mass., 1985.
- Canter, L. W., and R. C. Knox, *Groundwater Pollution Control*, Lewis Publishers, Chelsea, Mich., 1985.
- Carrera, J., E. Usunoff, and F. Szidarovsky, A method for optimal observation network design for groundwater management, *J. Hydrol.*, 73, 147-163, 1984.
- Cochran, W. G., *Sampling Techniques*, 3rd ed., John Wiley and Sons, New York, 1977.
- deMarsily, G., *Quantitative Hydrogeology*, Academic, San Diego, Calif., 1986.
- Everett, L., *Groundwater Monitoring*, General Electric Company, Shenectady, N.Y., 1980.
- Everett, L. G., L. G. Wilson, and L. G. McMillon, Vadose zone monitoring concepts for hazardous waste sites, *Ground Water*, 20(3), 312-324, 1982.
- Fiering, M., An optimization scheme for gaging, *Water Resour. Res.*, 1(4), 463-470, 1965.
- Fletcher, R., *Practical Methods of Optimization*, John Wiley and Sons, New York, 1987.
- Freeze, R. A., A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media, *Water Resour. Res.*, 11(5), 725-741, 1975.
- Gelhar, L., and C. L. Axness, Three dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.*, 19(1), 161-180, 1983.
- Gilbert, R. O., *Statistical Methods for Environmental Pollution Monitoring*, van Nostrand-Reinhold, New York, 1987.
- Hsu, N. S., and W. Yeh, Optimum experimental design for parameter identification in groundwater hydrology, *Water Resour. Res.*, 25(5), 1025-1040, 1989.
- Hsueh, Y. W., and R. Rajagopal, Modeling groundwater quality sampling decisions, *Ground Water Monitor. Rev.*, 121-134, 1988.
- Knopman, D. S., and C. I. Voss, Discrimination among some one-dimensional models of solute transport in porous media: Implications for sampling design, *Water Resour. Res.*, 24(11), 1859-1876, 1988.
- Loaiciga, H. A., Groundwater monitoring network design, in *Proc. VII Intl. Conf. Computational Methods in Water Resources*, vol. 2, edited by M. A. Celia, L. A. Ferrand, C. A. Brebia, N. G. Gray, and G. F. Pinder, pp. 371-376, Computational Mechanics Publications/Elsevier, New York, 1988.
- Loaiciga, H. A., and M. A. Mariño, An approach to parameter estimation and stochastic control in water resources with an application to reservoir operation, *Water Resour. Res.*, 21(11), 1575-1584, 1985.
- Loaiciga, H. A., and M. A. Mariño, Estimation and inference in the inverse problem, paper presented at Proc. Water Forum '86: World Water Issues in Evolution, Am. Soc. Civ. Eng., Long Beach, Calif., Aug. 4-6, 1986.
- Loaiciga, H. A., and M. A. Mariño, The inverse problem in confined aquifer: Identification and estimation with extensions, *Water Resour. Res.*, 23(1), 92-104, 1987a.
- Loaiciga, H. A., and M. A. Mariño, Parameter estimation in groundwater: Classical, Bayesian and deterministic assumptions and their impact on management policies, *Water Resour. Res.*, 23(6), 1027-1035, 1987b.
- Loaiciga, H. A., R. H. Shumway, and W. Yeh, Linear spatial interpolation analysis with an application to the San Joaquin Valley, California, *Stochastic Hydrol. Hydraul.*, 2(2), 113-136, 1988.
- Massmann, J., and R. A. Freeze, Groundwater contamination from waste management sites: The interaction between risk-based

- engineering design and regulatory policy, 1, Methodology, *Water Resour. Res.*, 23(2), 351–367, 1987a.
- Massmann, J., and R. A. Freeze, Groundwater contamination from waste management sites: The interaction between risk-based engineering design and regulatory policy, 2, Results, *Water Resour. Res.*, 23(2), 368–380, 1987b.
- Meyer, P. D., and E. D. Brill, A method for locating wells in a groundwater monitoring network under conditions of uncertainty, *Water Resour. Res.*, 24(8), 1277–1282, 1988.
- Neuman, S. P., C. L. Winter, and C. M. Newman, Stochastic theory of field-scale Fickian dispersion in anisotropic porous media, *Water Resour. Res.*, 23(3), 453–456, 1987.
- Olea, R. A., Sampling design optimization for spatial functions, *Math. Geol.*, 16(4), 365–391, 1984.
- Rajagopal, R., Conceptual design for a groundwater quality monitoring strategy, *Environ. Professional*, 8(3), 244–264, 1986.
- Rouhani, S., Variance reduction analysis, *Water Resour. Res.*, 21(6), 837–846, 1985.
- Scheidegger, A. E., General theory of dispersion in porous media, *J. Geophys. Res.*, 66(10), 3273–3278, 1961.
- U.S. Environmental Protection Agency, RCRA groundwater monitoring technical enforcement guidance document, U.S. Environ. Protect., Washington, D. C., 1986.
- Willis, R., and W. Yeh, *Groundwater Systems Planning and Management*, Prentice-Hall, Englewood Cliffs, N.J., 1987.
- H. A. Loaiciga, Department of Geography, University of California, Santa Barbara, CA 93106.

(Received June 7, 1988;
revised March 28, 1989;
accepted March 28, 1989.)