DEVELOPMENT OF A METHODOLOGY FOR THE DESIGN OF GROUNDWATER QUALITY SAMPLING NETWORKS

Graciela S. Herrera[1], Joseph Guarnaccia[2] & George F. Pinder[3]

Abstract - Groundwater quality sampling networks are an aid in characterizing groundwater contamination problems and in evaluating the performance of a remediation strategy. In this context the goal of a water quality sampling network is to estimate the extent and behavior of the contaminant plume. Estimating concentrations of a contaminant plume in an efficient way depends on both the location of the sampling wells and the times when the contaminant samples are taken. In response to this need two of the authors have proposed a methodology for the design of cost-effective water-quality sampling in which sampling locations and sampling times are decision variables (Herrera 1998, and Herrera and Pinder 1998).

The proposed methodology combines stochastic simulation (also known as Monte Carlo simulation) and a Kalman filter to obtain sampling network designs that minimize (we explain in the text in which sense) the uncertainty of the predicted contaminant concentration estimates. One important feature of the methodology is that the number of samples that are taken at a given time does not have to be stipulated. Rather, the method chooses the ”optimal” number of samples for each sampling time. In this paper we present an application of the approach to a hypothetical problem and a blueprint for and progress in the application of this methodology to a field problem.

Keywords - groundwater quality, sampling network design, stochastic models

1. SAMPLING NETWORK DESIGN METHODOLOGY

The groundwater-quality sampling network design methodology presented herein can be divided into two parts: 1) a method that is used to predict the uncertainty of an estimate

[1] Instituto Mexicano de Tecnología del Agua, Morelos, Mexico
[2] Ciba Specialty Chemical Corporation, Toms River, NJ

produced by data sampled at given times from a set of sampling wells; and 2) a method that uses this predicted uncertainty to choose sampling positions and sampling times.

For the first part we use a stochastic flow and transport model to compute a space-time contaminant-concentration estimate and its covariance matrix via stochastic simulation. Next, a Kalman filter is used to predict the uncertainty that the concentration estimate would exhibit if concentration data from samples taken from computed locations at different times were used to update the prior estimate. This is similar to using a space-time Kriging method to predict the uncertainty of an estimate, but here instead of using a space-time variogram obtained from an analysis of concentration data, we calculate a space-time covariance matrix from a transport model via stochastic simulation.

A function, F, of the predicted estimate uncertainty is used as a criterion to choose sampling positions and sampling times for the network. The function used depends on the objective of the design. For example, F could be the sum of the variance of the concentration estimate at locations close to a drinking water well during a given year. Different procedures, including optimization methods, can be used to select the positions and times that minimize the function F. So far, we have used a sequential procedure that selects the space-time sampling points that minimize the function F at each step, and stops when F reaches a predetermined value.

In principle any node of the model mesh can be chosen as a sampling position. Thus, if the model used is three dimensional, sampling positions at different depths can be chosen. However, in the present tests we restrict the sampling positions to a fixed, specified, depth.

1.1 KALMAN FILTER

The Kalman filter produces linear minimum-variance unbiased estimates for the state of a system given noisy data. It also establishes a way to update these estimates when a new measurement becomes available without a need to refer to old data (Jazwinski 1970).

Through stochastic-simulation tests we demonstrate that the contaminant concentration obtained from a stochastic transport model like the one we use in this work, has a strong time correlation that persists for long times (Herrera 1998). On the other hand, the Kalman filter, when applied in its traditional form to space-time problems, relies upon some assumptions that restrict the state time correlation. Thus, to be able to manage the strong time correlation of the contaminant concentration inherent in this methodology, we use a static Kalman filter applied to a space-time state.

[3] University of Vermont and Research Center for Groundwater Remediation Design, Burlington, VT

Let $C=\{c_{ijl}\}$ be a vector of concentration values at nodes of a numerical mesh, where c_{ijl} is the concentration at location (x_i,y_j) at time t_l. Let $\{z_n, n=1,2,...\}$ be a sequence of measurements of the corresponding contaminant concentrations c_{ijl}^n, where the subscript *ijl* indicates the space-time location at which the n-th sample is taken. These samples are related to the state through the linear measurement equations,

$$z_n=H_nC+v_n, \qquad (1)$$

where H_n is the n-th sampling matrix. The form of the sampling matrix depends on the relation between the data and the state. The set $\{v_n, n=1,2,...\}$ represents the measurement error. It is a white Gaussian sequence, with mean zero and covariance matrix R_n. The measurement error sequence $\{v_n\}$ and the state C are independent.

It can be shown (Jazwinski, 1970) that the minimum-variance unbiased estimate for the state variable, given the measurements $z_1, z_2, ...,z_n$, here denoted as $\hat{C}^n$, is the expected value of the state C given the data, that is, $\hat{C}^n = E\{C \mid z_1,...,z_n\}$. Note that in this notation the superscript identifies the number of measurements that are used to obtain the estimate. The covariance matrix of the estimate error is

$$P^n = E\{(C-\hat{C}^n)(C-\hat{C}^n)^T\},$$

where T denotes transpose.

Given a prior estimate of the system state, $\hat{C}^0$, and its covariance matrix, P^0, the minimum-variance linear estimate for the state and its covariance matrix can be obtained sequentially through the following formulas:

$$\hat{C}^{n+1} = \hat{C}^n + K_{n+1}(z_{n+1}-H_{n+1}\hat{C}^n),$$

$$P^{n+1} = P^n - K_{n+1}H_{n+1}P^n,$$

where

$$K_{n+1}=P^nH_{n+1}^T\,(H_{n+1}P^nH_{n+1}^T+R_{n+1})^{-1}.$$

1.2 ESTIMATION OF PRIOR MOMENTS

As was mentioned before, we use stochastic simulation to obtain the prior space-time concentration estimate and its covariance matrix. The hydraulic conductivity parameter is

represented as a spatial-random field for the stochastic model. Consequently, the velocity field, calculated through Darcy's law, also becomes a random field. Given the mean hydraulic conductivity and its variogram, conductivity realizations are obtained. The flow equation is solved numerically using each realization. The values obtained determine a velocity field that is, in turn, used to solve the contaminant transport equation and to produce a realization of the plume. Averaging the plume realizations we obtain the space-time concentration mean and its covariance matrix. We define C^0 as the mean concentration vector and P^0 as the covariance matrix associated with it. In the synthetic problem that we present, in addition to the hydraulic conductivity, the contaminant concentration at the contaminant source is also represented with a random variable but for the field application only the hydraulic conductivity is modeled as a random variable.

2. EXAMPLE PROBLEM

To illustrate our approach for the design of groundwater quality sampling networks consider the system presented in figure 1a. A contaminant source is located on the left hand side of a 0.5 by 0.5 mile region bounded on the right-hand side by a river. We want to design a contaminant-sampling program to estimate the contaminant concentrations of the moving plume during a two-year period. The concentration estimates will be associated with the nodes of what we call the *Kalman filter mesh* shown in figure 1a. Six measurements and estimation times are considered during this two-year period, that is, one every 121.7 days.

Sampling well locations and times must be selected. The criterion used to evaluate the estimates is a function of the estimate variance but other criteria can be used.

Note that the *Kalman filter mesh* is playing the role of two meshes that in general could be different, i.e. the *estimation mesh* and the *sampling mesh*. The first one would include all the nodes on which contaminant concentration estimates are sought, and the second one would include all the nodes that are possible sampling locations.

2.1 CONTAMINANT TRANSPORT SIMULATION

We use the steady-state flow equation and the conservative advection-dispersion transport equation to model the problem. The *stochastic simulation mesh* is required for the numerical solution of the flow and transport equations. The domain is divided into 40 X 40 equally sized elements. The stochastic simulation mesh is shown in figure 1b. Note that the *Kalman filter mesh* is a submesh of the *stochastic simulation mesh*. Boundary conditions for flow and transport are included in figures 1a and 1b, respectively.

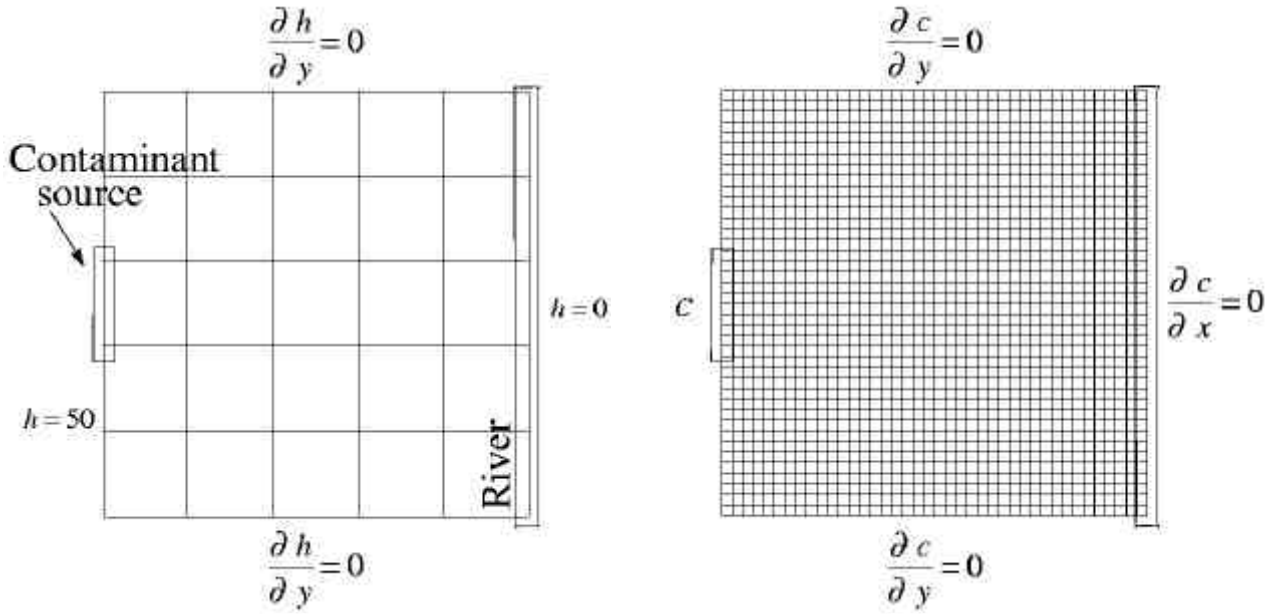


Figure 1. a) Problem set up, Kalman filter mesh, and boundary conditions for flow (*h* is in meters). b) Stochastic simulation mesh and boundary conditions for transport.

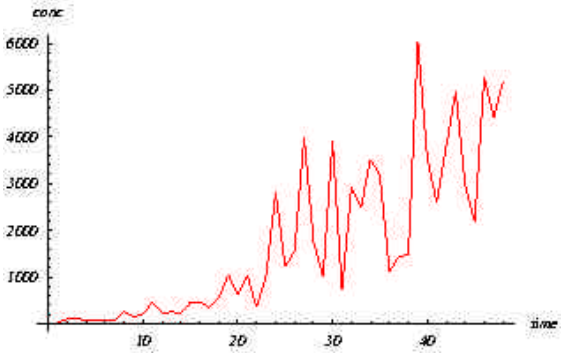


Figure 2. An example of a realization of the contaminant concentration at the source. The graph show concentraion vs time for the two year period.

Concentrations are in parts per million (ppm) and hydraulic heads are in meters. Forty-eight time-steps are used to simulate the two-year period. The contaminant source is active during all of this period. The Princeton Transport Code (PTC, Babu et al. 1993) is used in two-dimensional mode to solve these equations.

The Kalman filter mesh that we have chosen implies that there is a direct relation between the data, z, and the state, C. Each sample taken is the observation of a component of the state plus a random error (equation 1). For this reason, the n-th sampling matrix, H_n, is non zero and equal to one only at the entry of C corresponding to the position that is sampled.

2.2 HYDRAULIC CONDUCTIVITY AND SOURCE CONCENTRATION RANDOM FIELDS

We assume that the conductivity field is log-normally distributed, homogeneous, stationary, and isotropic. The mean value of the field is 3.055 and the variogram that represents the log-conductivity spatial correlation structure is

$$g(h)=0.69\left[1-\exp\left(-\frac{h}{352}\right)\right].$$

Concentration at nodes located at the source is modeled with identically distributed independent random functions. On each node, the concentration is represented as a time series, with

$$c(t) = \exp(-14 + 3t + e(t)),$$

where, *e(t)* is a zero-mean random perturbance, normally distributed and with variance 0.195. While the variance was obtained from the analysis of a time series of field data, the form of the exponent is hypothetical. For each one of the source nodes, at each simulation time-step, a different random perturbation is used. In figure 2 a realization for one of the nodes at the contaminant source is shown. The time correlation of the random perturbations is modeled with the variogram

$$g_e(t) = 0.195\left[1 - \exp\left(-\frac{t}{l_e}\right)\right],$$

where l_e is the correlation scale of e, and in this example $l_e = 11$ days.

We use a method called sequential Gaussian simulation (SGS) from the GSLIB package (Deutsch and Journel, 1992) to obtain hydraulic conductivity and contaminant concentration realizations.

A set of tests was done to check convergence of the method described above with respect to the number of realizations. It was concluded that 3,000 plume realizations were enough to obtain convergence.

2.3 SAMPLING PROGRAM

The function F used in these examples to evaluate the quality of a given estimate is the total variance, F=tvar. The total variance of the estimate obtained when samples {z_1, z_2,..., z_n} are used, is defined as the sum of the estimate variances over all locations and times. For the present example this is,

$$\text{tvar}(n) = \sum_{ijl} s^2_{ijl}(n),$$

where $s^2_{ijl}(n)$ is the variance of the estimate at the *i,j* location on the Kalman filter mesh, and *l*-th estimation time. The elements $s^2_{ijl}(n)$ are obtained from the covariance matrix P^n.

One sampling location at a time is chosen, the one selected is that which reduces tvar the most given previous sampling decisions. That is, if positions $\{\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n\}$ have been chosen, position $\mathbf{x}_{n+1}$ is chosen in the following way:

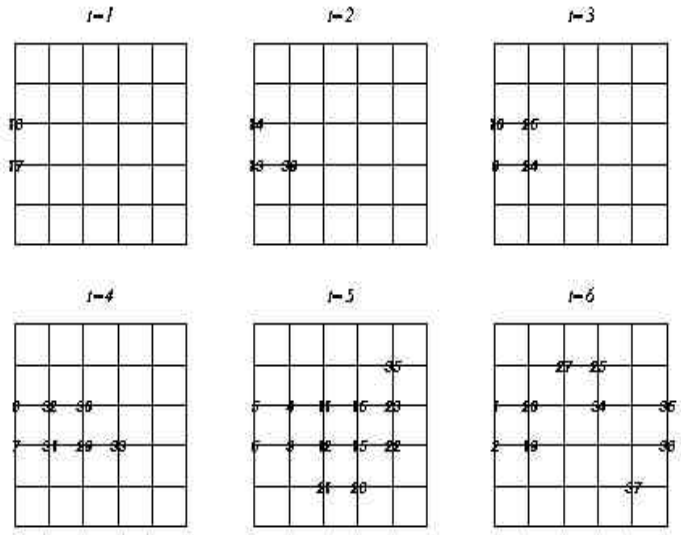
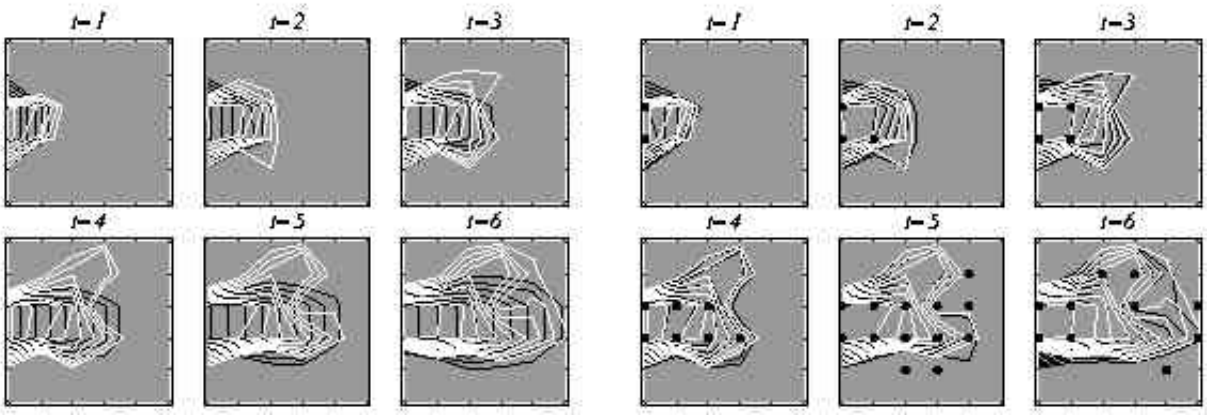


Figure 3. Sampling program for the hypothetical problem.

1) Using the Kalman filter, calculate the variance of the minimum error variance estimate obtained from data sampled at positions $\{\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n, \mathbf{X}_i\}$, for all possible sampling positions $\mathbf{X}_i$.

2) Choose the position $\mathbf{X}_i$ that minimizes the total variance.

Note that in this example we use each new sample to estimate the concentration in the whole two-year period considered. Then, a sample taken at a given time contributes to estimate concentration at both past times and future times. In the present example a sampling program with 39 samples is obtained.



a) b)

Figure 4. Comparison of the observed plume and the plume estimates (logarithmic scale). The observed plume is in white contours and the estimates are in black contours. Black dots indicate sampling locations. a) Prior estimate. b) Plume estimate for a sampling program of 39 samples.

The example analyzed assumes samples with no error (v_n=0 in equation 1). The sampling program chosen by the algorithm when minimizing the total variance is shown in figure 3. The figure includes six squares that represent the Kalman filter mesh at the six

sampling times. The numbers on the mesh indicate the order in which the samples were chosen. Number 1, for example, indicates that the first sample was chosen close to the contaminant source, at the sixth sampling time: (x_1, y_4, t_6). These numbers attest to the importance of a sample from the corresponding space-time location for reducing the total variance.

Twelve of the first twenty sampling locations are chosen close to the source. The other eight samples are located at the two central rows of the sampling mesh, either at the fifth or sixth times. This tendency to first select the samples close to the source is due to the large concentration variance at those locations.

All of the twenty first samples are chosen on the third and fourth rows of the sampling mesh, this suggests that to reduce the total variance of the concentration estimate it is important to obtain first the central tendency of the plume. The last nineteen samples seem to define the spreading of the plume; fourteen of them are located where the prior plume has its boundaries. These boundaries are shown in (figure 4).

As was mentioned before, the number of samples at each sampling time is not stipulated but is chosen by the algorithm. In figure 3 it can be seen that the number of samples augments with time until the fifth sampling time. Thus, the sampling time that gives the most information is t=5. It is interesting to note that this is not the time at which the expected plume has the largest variances; this time is t=6.

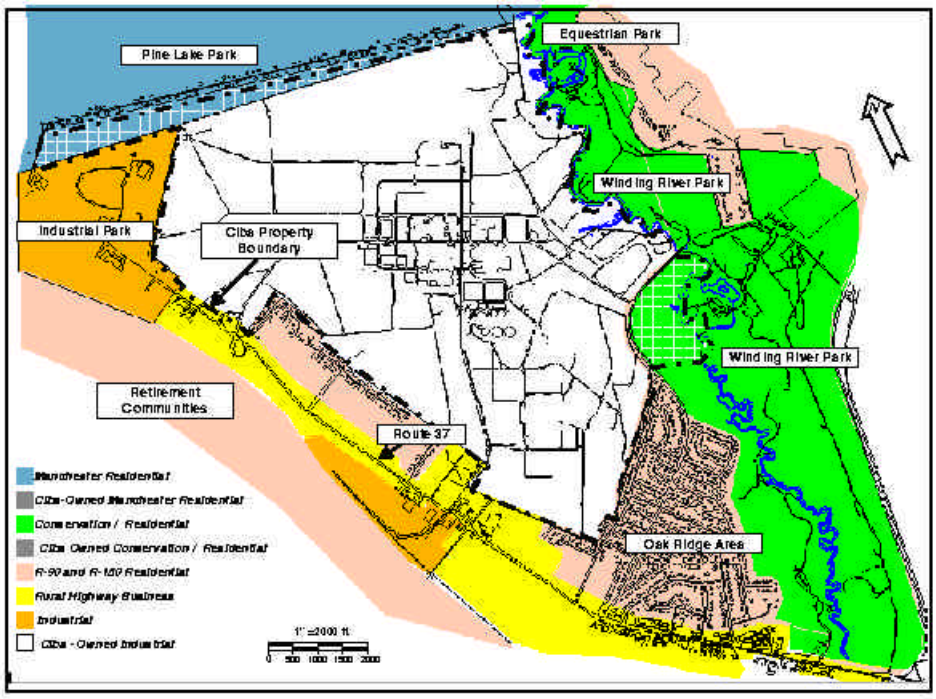


Figure 5. Study area site map. Taken from the report Ciba 1999.

2.4 OBSERVED PLUME

To investigate the quality of the plume estimates obtained by the proposed method we compare these estimates with a pre-selected plume from which samples are taken in accordance with our algorithm. The "observed" plume is selected arbitrarily from the set of realizations. In figure 4a a comparison between the observed plume and the prior estimate is presented, and in figure 4b comparison between the observed plume and the plume estimate obtained when the data from the sampling network proposed is used is presented. The observed plume is shown in white contours and the estimates in black contours. A logarithmic scale is used. As can be observed, the updated plume estimate gets very close to the observed plume.

3. FIELD PROBLEM

The objective of this part of the work is to “validate” the methodology described above by applying it to a field site. What follows outlines our progress to date.

To apply our approach, and to assess its effectiveness, we require a site at which the following are available:

• Hydraulic conductivity data

• Groundwater quality data gathered over a long period

• A flow- and contaminant-transport model calibrated for the site

We need enough hydraulic conductivity data to do a geostatistical analysis and to estimate the hydraulic conductivity mean and variogram. Groundwater quality data is needed to compare the estimate obtained with the currently available data with the one obtained from the sampling network selected by our methodology. Finally, a flow and transport model is needed for the stochastic simulation. We have chosen the aquifer underlying a superfund site in New Jersey to test the methodology because it satisfies these three requirements, and we have access to the data and a site model (see figure 5). The information about the site presented in this paper was taken from the report Ciba 1999.

3.1 DATABASE

A number of remedial investigations have been conducted at the site (Aware 1986, NUS Corporation Superfund Division 1988, CDM 1989, Eckenfelder Inc. 1991, Eckenfelder Inc. 1993, Environ Corporation 1993, Ciba 1999. As a result a database of several parameters has been compiled. For our purposes the most important data are those related to hydraulic conductivity, the temporal and spatial distribution of groundwater head, and water quality. Data related to soil are also important because they are used to characterize source areas for the transport model.

Data from 641 wells at the site are available. Most of the groundwater quality data were obtained from 1982 to date, and for some compounds data are available since the 1950s. For chlorobenzene there are 5723 measurements gathered from 1982 to 1999.

3.2 SITE HYDROLOGY AND GROUNDWATER QUALITY

The site is located in the New Jersey Coastal Plain. Within the upper 200 feet beneath the site there are nine recognized geologic members comprised of unconsolidated sands, silts, and clays. In descending order these are:

1. Upper Cohansey Member
2. Cohansey Yellow Clay Member
4. Cohansey/Kirkwood Transitional Member
5. Lower Cohansey Member
6. Upper Kirkwood Member

7. Kirkwood No. 1 Member

8. Semi-Confining Unit (Primary Kirkwood)

9. Lower Sand Aquifer (Kirkwood No. 2 Sand)

Dissolved contaminants have leached from the source areas into portions of the upper five geologic units (i.e., the Primary and Lower Cohansey Members). Units below the Lower Cohansey are uncontaminated.

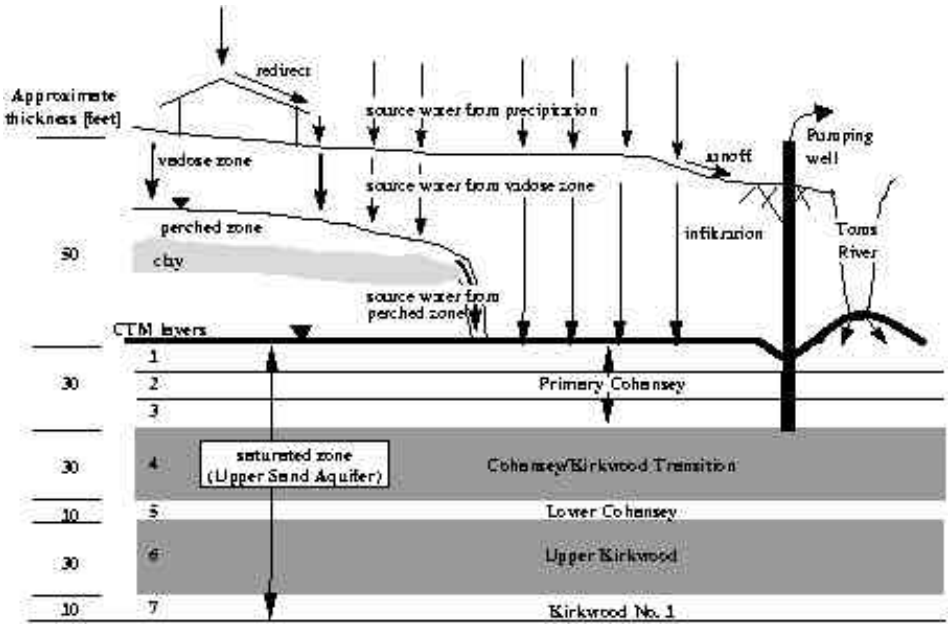


Figure 6: A conceptualized cross sectional view through the model domain. CTM stands for computational transport model. Taken from the report CIBA, 1999.

The aquifer can be divided into two hydrostratigraphic units: the Upper Sand Aquifer and the Lower Sand Aquifer. The Semi-Confining Unit separates these two subdivisions. In figure 6 a simplified view of the Upper Sand Aquifer in cross-section is shown. The gray units are conceptualized as aquitards and the white units as transmissive water bearing units.

The focus of our work is the Primary Cohansey aquifer. Groundwater flow within this aquifer is principally horizontal and it is hydraulically connected with the Toms River. Of the suite of chemicals present in the groundwater, chlorobenzene is the most ubiquitous. For this reason we chose to design the sampling network for this compound.

3.3 GROUNDWATER FLOW AND TRANSPORT MODEL

Figure 6 shows a conceptual view of the model in cross-section. The model simulates the flow and transport properties of the Upper Sand Aquifer for the years 1953 to

2028. The different geological units are represented with seven model layers. Because the Primary Cohansey is relatively thick, and because it plays an important role in defining the contaminant-transport problem, the model represents this unit as three layers. Applying the estimated-flux values to the top layer of the model incorporates the water flux entering the saturated zone from above.

Two different domains are used, one for a regional flow model and one for a local flow and transport model. The domain for the regional flow model is larger than, and contains, the local model. For the regional model, where practical, the boundary of the model is keyed into regional surface watercourses. The regional model is solved first and the resulting heads are used to define the boundary conditions for the local model.

As with the regional model, the local model simulates the Upper Sand Aquifer. The lateral boundary and the mesh used in the local model domain are shown in figure 7. The transport model includes retardation and biodegradation.

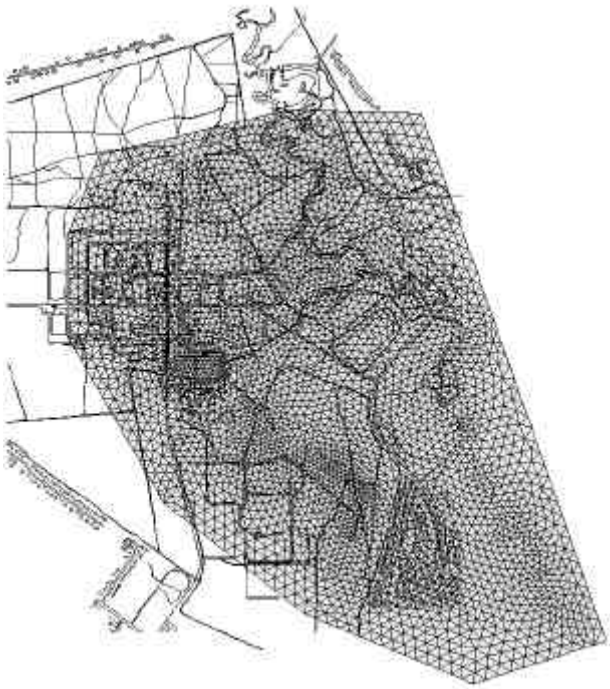


Figure 7. Domain and mesh of the local model.

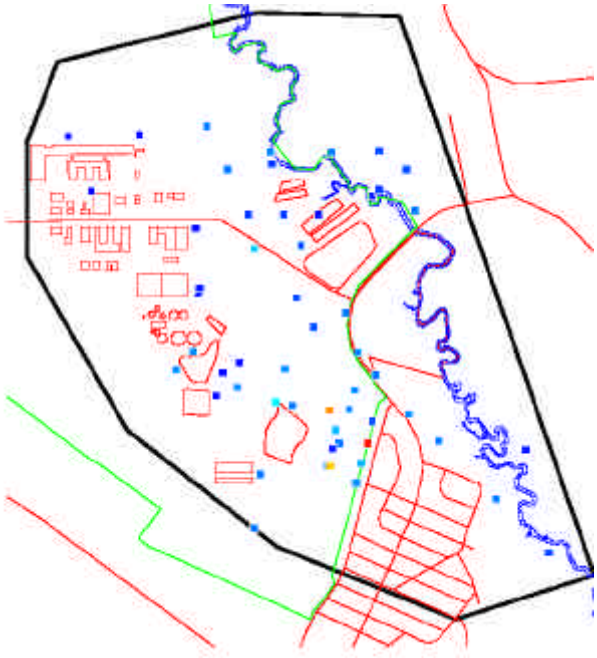


Figure 8. Positions of the wells (shown with a square) for which chlorobenzene data obtained from layer 3 during 1985 and 1986 are available. The model domain is shown with the darkest contour and the limit of the facility with a lightest one.

3.4 SAMPLING NETWORK DESIGN OBJECTIVES

As was mentioned earlier, in the network design we assume that all the samples are taken at the same depth. For this reason, the sampling network will be designed for one of the model layers. We have chosen the lower part of the Primary Cohansey aquifer, layer 3 in the model, to do the tests. Also, all the sampling wells will be in the domain of the local

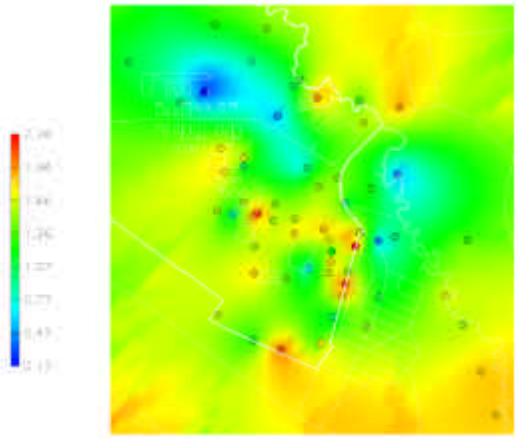


Figure 9. Hydraulic conductivity contours in logarithmic scale. The positions of the data are shown with a circle.

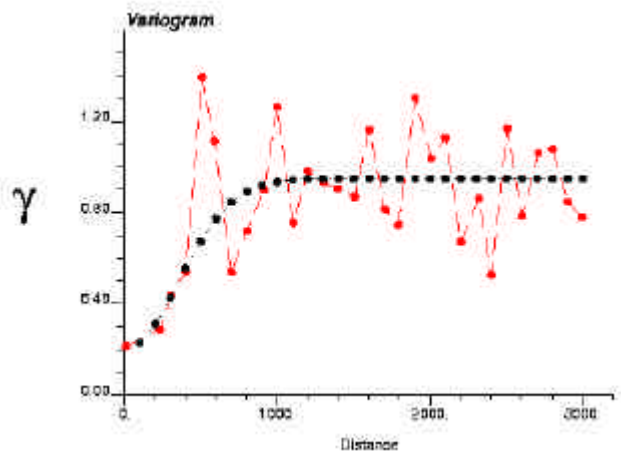


Figure 10: Comparison between the experimental variogram and the model variogram, the first is shown in gray and the second in black. Distance is in feet.

model because this is the area of most concern. As a first step, two simple tests will be performed for a period of two years:

1. optimization of the existing sampling network,

2. design of a new sampling network.

We have chosen the years 1985 and 1986 for these tests because for those years the largest number of chlorobenzene data are available at layer 3, a total of 322 data from 62 sampling wells. The wells that are located in the model domain are shown in figure 8. The wells with more data are concentrated close to the southeast limit of the facility.

In the first test from the existing samples, we will choose those sampling positions and sampling times for which we obtain a concentration estimate uncertainty equivalent to the one obtained using all samples available. For this test it will be possible to compare the resulting estimate from the sampling network suggested by our methodology and the one obtained using all the data available for the two years.

The second test consists of designing a new sampling network for the years 1985 and 1986. In the design process, well positions will be chosen from the nodes of a sub-mesh of the numerical mesh, and sampling times will be selected from the set of times defined by a frequency of half a month. The objective again will be to obtain an estimate with uncertainty equivalent to the uncertainty of the estimate obtained using all the data available.

In these tests we will assume that each sampling well location coincides with a node in the numerical mesh. Since not all the wells satisfy this assumption, we will approximate the well positions with the position of the node that is closest to it. For this model this assumption should not introduce much error because the mesh is dense and, in most cases, for every element there is only one well and it is close to one of its nodes.

We have a similar situation with the chlorobenzene data sampling times: samples were collected at different dates at each well, so it is impossible to have a simulation output for each sampling time. For this reason, we will get a simulation output every half a month and we will approximate each data sampling time with the simulation time closest to it.

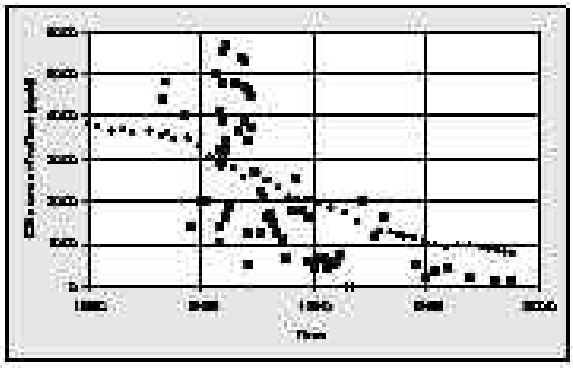


Figure 11: Chlorobenzene behavior at well 744-150 in the Primary Cohansey. The plot shows the trend in concentration (units of ppb) versus time, where the solid diamonds represent a transport model result, the squares represent an unqualified measured value, and the crosses represent 1/2 the detection limit if the result indicated not detected. Modified from CIBA, 1999

3.5 STOCHASTIC SIMULATION

In the stochastic simulation we will use only the local model. The flow boundary conditions for this model will be determined from the deterministic regional model, i.e. the regional model will be run only once with the conductivity values assigned in the calibrated model.

For the local model all the parameters will be deterministic with values equal to those of the original model, with the exception of the hydraulic conductivity in layer three. For this layer, in each model run, the hydraulic conductivity will take the values of the corresponding realization. Each hydraulic conductivity realization will be conditioned with data, that is, for the locations in which conductivity data are available each realization will take the values of the data.

3.6 HYDRAULIC CONDUCTIVITY GEOSTATISTICAL ANALYSIS

A number of slug tests have been performed at the site to estimate the hydraulic conductivity (Aware 1986, NUS Corporation Superfund Division 1988, Eckenfelder Inc. 1991, and Eckenfelder Inc. 1993). We have done a geostatistical analysis of the data obtained from these tests for the Primary Cohansey Member. Hydraulic conductivity contours in logarithmic scale for this member are shown in figure 9. The contours show that a band of low conductivity values crosses the site from northwest to southeast. This

band divides the site into two zones: the southwest and the northeast. The highest hydraulic conductivity values are found in the first of these two zones and that is the zone with more variability. In the northeast section mid-range conductivity values are found.

It has been shown that hydraulic conductivity tends to have a log-normal probability distribution (Hoeksema and Kitanidis 1985), and in our study this was confirmed. A conductivity mean value of 33.25 ft/day and a log-conductivity mean value of 1.34 were obtained using the moving window method (Isaaks and Srivastava 1989). A variogram fit for the logarithm of the data was obtained by trial and error, the best fit was obtained with the following isotropic Gaussian model:

$$g(h) = c_0 + c\left\{1 - \exp\left(-(3h/a)^2\right)\right\},$$

where c_0=0.2 , c=0.75 and a=1500 ft. The sill is equal to 0.95, the effective range is equal to 866.025 ft (=a/3), and the variogram has a nugget effect of 0.2. A comparison of the experimental variogram and the model variogram is shown in figure 10. The Gaussian model indicates that there is regularity in the site hydraulic conductivity.

4. DISCUSSION AND CONCLUSIONS

One might anticipate that with so much information about the site and the large database available, a very good match between the model results and the various field data could be obtained. However, as is often the case, the groundwater quality data are more difficult to reproduce than the groundwater head data, even with the availability of a large number of data points. Such is the case for this model. The present model correctly reproduces the slope of the groundwater head and the general trend of the contaminant concentrations, but there can be important discrepancies between individual measurements and model predictions (see figure 11 for an example). This may be explained by the following observations:

1. The model was developed for nine chemicals of concern and a compromise had to be made in the calibration between all of them;

2. Typically there is much variability in groundwater quality data and it is not easy to incorporate the sources of variability in a deterministic transport model;

3. There is a difference between the volume sampled by a monitoring point and the volume estimated by the transport model.

A stochastic transport model such as the one that we use in the proposed methodology has the advantage that it estimates the uncertainty of the predicted contaminant concentration. So, we should be able to recognize an area in which it is possible to have discrepancies like those shown in figure 11 from the uncertainty associated with it. However, it is important to recognize that the stochastic model prediction uncertainty depends on the random parameters of the model. So, it is important to evaluate what are the parameters that contribute the most to the prediction uncertainty.

For the present study we think that a second important source of uncertainty may be the variability of the contaminant concentration at the contaminant sources. In the hypothetical example considered earlier the concentration at the contaminant source was modeled as a random term. In figure 2 an example of a realization of the random contaminant concentration is shown. It can be seen that the abrupt changes in concentration are similar to the data behavior in figure 11. If modeling the contaminant sources as random variables becomes necessary, it appears that enough information exists to characterize them statistically.

The difference between the volume sampled by a monitoring point and the volume estimated by the transport model can be incorporated in the model when determining the form of the sampling matrix, $\mathrm{H_n}$.

It is interesting to note that from the geostatistical analysis a hydraulic conductivity mean of 33.25 ft/day was obtained, while in the model a best-fitting homogeneous conductivity of 90 ft/day was used for layer three. This discrepancy between the two values is not surprising, it is well known that the effective hydraulic conductivity of a random field depends not only on the conductivity mean, but also, on the conductivity variance (Gelhar 1993).

As with deterministic numerical models, when using a discrete approximation for a random field the size of the numerical mesh elements must satisfy certain conditions in order to get a good approximation. Ababou et al. (1989) proposed an empirical rule for this purpose:

$$\frac{l}{\Delta x} \geq 1 + s_Y^2, \qquad (2)$$

where $Y=\log K$, l is the integral range, Δx is a representative length of the elements of the model mesh and s_Y^2 is the log-conductivity variance.

The size of the local model mesh elements satisfies this relation. For a Gaussian model $l = a\sqrt{\Pi}/6$, which in this case yields $l = 443$ ft approximately. The area of the largest mesh elements is about 18000 ft^2, this yields an equivalent square with an area that has a side length of about 134 ft. If we take this length as representative of the mesh elements we have $l/\Delta x = 3.3$. On the other hand $1+s_Y^2 = 1.95$, so, equation 2 is satisfied.

REFERENCES

Ababou, R., McLaughlin, D., Gelhar, L. W., & Tompson, A. F. B. (1989). Numerical simulation of three-dimensional saturated flow in randomly heterogeneous porous media. *Transport in Porous Media 4,* (6), 549-565.

AWARE Incorporated (1986). Hydrological and Related Investigation, Volume IV: Miscellaneous Hydrogeologic and Geotechnical Data. Technical report, AWARE Incorporated.

Babu, D., Pinder, G. F., Niemi, A., Ahlfeld, D. P., and Stothoff, S. A., Chemical Transport by Three Dimensional Groundwater Flows, Tech. Rep. 84-WR-3, Princeton University, Department of Civil Engineering, Princeton, N.J., 1993.

CDM (1989). Final Report Ciba-Geigy Technical Enforcement. Support Document: Ground Water Modeling Report, Ciba-Geigy, Toms River, New Jersey. Technical report.

Ciba (1999). Draft Feasibility Study Report for Operable Unit 2. Technical report, Ciba Speciality Chemicals Corporation, North America, Toms River, NJ.

Deutsch, C. V., and Journel, A. G., *GSLIB, Geostatistical Software Library and User's Guide*, Oxford University Press, New York, 1992.

Eckenfelder Inc. (1991). Summary Report: Sampling, Analysis, and Monitoring Plan (SAMP) Investigation (Appendices C through G). Volume III. Technical report, Eckenfelder Inc.

Eckenfelder Inc.(1993). Summary Report of Addendum Source Area Remedial Investigation Oversight. Volume III. Appendices F through J. Technical report, Eckenfelder Inc.

Environ Corporation (1993). Revised Draft Source Control Remedial Investigation Report, Ciba-Geigy Site, Toms River, New Jersey. Appendix A: Preliminary Assesment of Potential Source Areas to the South Plume at Ciba-Geigy's Toms River Plant. Technical report, Environ Corporation.

Gelhar, L. W., *Stochastic Subsurface Hydrology*, Prentice Hall, Englewood Cliffs, New Jersey, 1993.

Herrera, G. S. (1998). *Cost Effective Groundwater Quality Sampling Network Design*. Ph. D. thesis, University of Vermont, Burlington, Vermont.

Herrera, G. S. & Pinder, G. F. (1998). Cost-effective groundwater quality sampling network design. In V. N. Burganos, G. P. Karatzas, A. C. Payatakes, C. A. Brebbia, W. G. Gray, & W. G. Pinder (Eds.), *Computational Methods in Water Resources XII*, Volume 1, Southampton, U. K., pp. 51-58. Computational Mechanics Publishers.

Hoeksema, R. J. & Kitanidis, P. K. (1985). Analysis of the spatial structure of properties of selected aquifers. *Water Resour. Res. 21* (4), 563-572.

Isaaks, E. H. & Srivastava, R. M. (1989). *Applied Geostatistics*. New York: Oxford University Press.

Jazwinski, A. (1970). *Stochastic Processes and Filtering Theory*. New York: Academic Press.

NUS Corporation Superfund Division (1988). Remedial Investigation for the Ciba-Geigy Site, Toms River, Dover Township, Ocean County, New Jersey. Technical report, NUS Corporation Superfund Division.