The 3rd International Conference on Water Resources and Arid Environments (2008)
and the 1st Arab Water Forum

Optimizing Groundwater Monitoring Network in Highly Heterogeneous Aquifers Using Geostatistics: Technique Development with a Few Case Studies

Shakeel Ahmed[1], A. Nabi[1], D. Kumar[1], P.D. Sreedevi[1] and J. Perrin[2]
1. National Geophysical Research Institute, Indo-French Centre for Groundwater Research, Hyderabad, India
2. BRGM, Water Division, Resource Assessment, Discontinuous Aquifers Unit, Indo-French Centre for Groundwater Research, Hyderabad, India

Abstract

A few new procedures have been developed using geostatistical technique such that the number of monitoring wells was reduced without loosing the monitoring benefits. The objectives of the geostatistical optimization of the monitoring network has been that the monitored parameter should (a) represent the true variability of the parameter and (b) provide its estimate on unmeasured locations with a desired accuracy in the form of the variance of the estimation error.

A small watershed with a semi-confined aquifer in fractured granite rocks of a semi-arid region having an area of 60 km2 have been studied. The flow and the fluoride contamination in the area have to be simulated. A number of bore wells (55 to 60) based on geomorphology, land use and other variations were decided to have a first hand network for measurement of water-levels and water quality. The monitoring network consisted of 55 bore-wells out of which 25 were specially drilled for monitoring the water levels and 30 were selected from the existing irrigation wells. The values of water levels (amsl) of the aquifer from the given network of monitoring wells have been geostatistically analyzed to study the variability of the head data. The analysis of the variance of the estimation error obtained using kriging technique on a suitably finer grid in the entire area has been the basis to discard some of the private irrigation wells. The wells were discarded such a way that the variance of the estimation error does not exceed a pre-decided value. Thus after a few iterations, a network of 40 monitoring wells have been evolved consisting of 25 specially drilled observation wells and 15 private irrigation wells keeping the variance of the estimation error within the desired limit. In another study a simple method has been developed using the cross-validation technique of geostatistical estimation to analyse and optimize an existing network for monitoring fluoride. In the same watershed, water samples have been collected from 60 wells and analyzed for water chemistry and fluoride. The fluoride values were analyzed geostatistically and the variogram calculated was validated through cross-validation test. Using the result of the cross-validation, it has been possible to assign a priority

index to all these measurement points (ranking the measuring wells) and depending on the constraints such as financial and logistic, a network could be reduced without loosing the outcome by selecting the wells on the basis of the priority index. A number of monitoring networks with less number of wells have thus been prepared and compared. It was found that monitoring fluoride from a network of only 30 wells selected using the priority index could provide almost same variability that obtained from monitoring all the 60 wells. It was also observed that with the use of the priority index we had still to monitor some cluster of wells that one could have easily discarded by a visual decision without any scientific analysis.

Introduction

Appropriate and adequate data are essential for the success of any scientific study. Scarcity of data and their collection on isolated location mainly in the field of hydrogeology makes it necessary to adopt special procedures or an estimation technique for bridging the gap between field measurements and data requirements. Numerical simulation of flow and transport processes in an aquifer necessitates, dividing and discretizing the natural heterogeneous system into a number of small volumes called mesh which are assumed to be uniform with almost no variation of the aquifer properties over it. To satisfy this condition, it is necessary to discretize the system into much finer and hence more number of grids. Although with the availability of more powerful computers, computation with large number of grids/mesh is not a difficulty but the data preparation that is to assign the aquifer parameters to each grid/mesh becomes cumbersome. Thus an appropriate estimation procedure is required to provide an unbiased, minimum variance and with unique value over the entire area of the mesh.

Geostatistical techniques in the form of "Theory of Regionalized Variables" were initially developed to apply to mining problems (Matheron, 1963). But soon, hydrogeologists have realized its applications to the groundwater hydrology and a number of studies have been carried out on hydrogeological problems. Works of Delhomme (1978), Mizell (1980) Aboufirassi and Marino (1983), Neuman (1984), Ahmed and Marsily (1987, 1993), Ahmed et al. (1988), Roth et al (1996) etc. have shown more applications of Geostatistics in groundwater hydrology. However, multivariate and non-stationary Geostatistics have found comparatively more applications in groundwater hydrology. Also some of them have to be suitably modified as well as some special procedures developed for a meaningful application of Geostatistics in this field. Delhomme (1976) has developed the method of Kriging with Linear Regression, Kriging using erroneous data, Kriging in the presence of a fault etc. Conditional simulation has also been applied in aquifer modeling (Delhomme, 1979). Galli and Meunier (1987) and Ahmed (1987) have worked on Kriging with an External Drift. Ahmed and Marsily (1987) have compared a number of multivariate Geostatistical methods in estimating transmissivity using data on transmissivity and specific capacity. Also Ahmed (1987) has developed a special antisymmetric and anisotropic cross-covariance between residuals of hydraulic head and transmissivity based on the work of Mizell (1980) and used coherent nature of various covariances to cokrige transmissivity and hydraulic head in solving an Inverse Problem (Ahmed and Marsily, 1993).

Geostatistical estimation variance reduction, cross-validation techniques etc. are a few procedures that could study adequacy of a given monitoring network and could evolve an optimal monitoring network with some given constraints. The advantage of the geostatistical estimation technique is that the variance of the estimation error could be calculated at any point without having the actual measurement on that point (well). Thus the benefits to be accrued from an additional measurement could be studied prior to its measurement.

Monitoring through an optimal network can reduce the amount of data required while providing the desired accuracy and a few simple procedures have been developed for a geostatistical optimization of the monitoring network in this study.

Geostatistics: The Theory of Regionalized Variables

Geostatistics, a statistical technique based on the theory of regionalized variables, is used to describe spatial relationships existing within or among several data sets. The application of kriging (various techniques of Geostatistics) to groundwater hydrology was initiated by Delhomme (1976,1978,1979), followed by a number of authors, viz., Delfiner and Delhomme (1983), Marsily et al (1984), Marsily (1986), Aboufirassi and Marino (1983, 1984), Gambolti and Volpi (1979a, b) and many others. In geostatistics, the geological phenomenon is considered as a function of space and/or time. Let z(x) represents any random function of the spatial coordinates z with measured values at "N" locations in space $z(x_i)$ i = 1,2,3,…N, and suppose that the value of the function z has to be estimated at the point x_o. Each measured value z_1, z_2, z_3,….z_N contributes in part to the estimation of the unknown value z_o at location x_o. Taking this into account, and assuming a linear relation between z_o and the z_i, the expected value of z_o, which in the "kriging estimate" z_o*, can be defined as:

$$z^*(x_o) = \lambda_1 z_1 + \lambda_2 z_2 + \lambda_3 z_3 + \lambda_4 z_4 + \ldots\ldots\ldots \lambda_N z_N$$

$$\text{or }: z^*(x_o) = \sum_{i=1}^{N} \lambda_i \, z(x_i) \qquad (1)$$

where $z^*(x_o)$ is the estimation of function z(x) at the point x_o and λ_i are the weighting factors. For making this predictor unbiased and optimal, the following two conditions need to be satisfied.

i. On an average, the difference between the estimated value and the true (unknown) value, i.e. the expected value of the estimation error should be zero. This un-biasness condition is also sometimes called the universality condition.

$$E\,[z^*(x_o) - z(x_o)\,] = 0 \qquad (2)$$

ii. The condition of optimality means that the variance of the estimation error should be minimum.

$$\sigma_K^2\,(x_o) = \text{var}\,[(z^*(x_o) - z(x_o)] \quad \text{is minimum.} \qquad (3)$$

Using equations (1) and (2), we get

$$\sum_{i=1}^{N} \lambda_i = 1 \tag{4}$$

Expanding equation (3), we obtain

$$\sigma_K^2(x_o) = \sum_{i=1}^{N} \lambda_i \sum_{j=1}^{N} \lambda_j \; E\,[z\,(x_i)\,z\,(x_j)] + E\,[z^*(x_o)^2\,] - 2\sum_{i=1}^{N} \lambda_i \; E\,[z\,(x_i)\,z^*(x_o)] \tag{5}$$

The best unbiased linear estimator is the one which minimizes $\sigma_K^2(x_o)$ under the constraint of equation (4). Introducing the Langrange multipliers and adding the term -2 $\mu\,(\sum_{i=1}^{N} \lambda_i$ -1), we obtain

$$\begin{aligned} Q \quad &= \sigma_K^2(x_o) - 2\,\mu\,(\sum_{i=1}^{N} \lambda_i - 1) \\ &= \sum_{i=1}^{N} \lambda_i \sum_{j=1}^{N} \lambda_j \; C\,(x_i\,,\,x_j) + C\,(0) - \\ &\qquad 2\sum_{i=1}^{N} \lambda_i\, C\,(x_i\,,\,x_o) - 2\mu \sum_{i=1}^{N} \lambda_i + 2\mu \end{aligned} \tag{6}$$

To minimize the above equation, making partial differentiation of Q w.r.t. λ_i and μ and equating them to zero, we obtain the following kriging equations:

$$\sum_{j=1}^{N} \lambda_j\, C\,(x_i\,,\,x_j) - \mu = C\,(x_i\,,\,x_o) \tag{7}$$

i = 1,2,3,…..N

$$\sum_{j=1}^{N} \lambda_j = 1 \tag{8}$$

where C $(x_i\,,\,x_j)$ is the covariance between points x_i and x_j. Substituting equation (7) into equation (5), we obtain the variance of the estimation error.

$$\sigma_K^2(x_o) = C\,(0) - \sum_{i=1}^{N} \lambda_i \; C\,(x_i\,,\,x_o) + \mu \tag{9}$$

The square root of this equation gives us the standard deviation $\sigma_K(x_o)$, which means that with 95% confidence, the true value will be within $z^*(x_o) \pm 2\,\sigma_K(x_o)$.

In case the covariance cannot be defined, we can derive the following kriging equations:

$$\sum_{j=1}^{N} \lambda_j \gamma (x_i , x_j) + \mu = \gamma (x_i , x_o) \qquad (10)$$

i =1,2,3,…..N

$$\sum_{j=1}^{N} \lambda_j = 1 \qquad (11)$$

where $\gamma(x_i , x_j)$ is called the variogram between points x_i and x_j (see below), and the variance of the estimation error becomes:

$$\sigma_k^2 = \sum_{i=1}^{N} \lambda_i \ \gamma (x_i, x_o) + \mu \qquad (12)$$

Equations (10) and (11) are a set of (N+1) linear equations with (N+1) unknowns and on solving them we obtain the value of λ_i, which are used to calculate the equation(1) and (9) or (12).

Kriging is the best linear unbaised estimator (BLUE) as this estimator is a linear function of the data with weights calculated according to the specifications of unbaisedness and minimum variance. The weights are determined by solving a system of linear equations with coefficients that depends only on the variogram, which describes the spatial structure (correlation in space) of the function z. The weights are not selected on the basis of some arbitrary rule, but depend in fact on how the function varies in space. A major advantage of kriging is that it provides the way to estimate the magnitude of the estimation error, which is the rational measure of the reliability of the estimate. Since the variance of the estimation error depends on the variogram and on the measurements location, therefore, before deciding on selecting a new location for measurement, or deleting an existing measurement point, the variance of the new estimation error can be calculated and a better measurement network can be designed based on the minimum variance even prior to any new measurement.

The Variogram

It has been observed that all the important hydrogeological properties and parameters such as piezometric head, transmissivity or hydraulic conductivity, storage coefficient, yield, thickness of aquifer, hydrochemical parameters, etc. are all functions of space. According to Marsily (1986) these variables (known as regionalized variables) are not purely random, and there is some kind of correlation in the spatial distribution of their magnitudes. The spatial correlation of such variables is called the structure, and is normally described by the variogram. The experimental variogram measures the average dissimilarity between data separated by a vector h (Goovaerts, 1997). It is calculated according to the following formula.

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left[z(x+h) - z(x) \right]^2$$

h = separation distance between two points, also called the lag distance. A generalized formula to calculate the experimental variogram from a set of anisotropic scattered data can be written as follows (Ahmed, 1995):

$$\gamma(\underline{d},\underline{\theta})=\frac{1}{2N_d}\sum_{i=1}^{N_d}[z(x_i+\hat{d},\hat{\theta})-z(x_i,\hat{\theta})]^2 \qquad (13)$$

where,

$$d-\Delta d\leq\hat{d}\leq d+\Delta d\ ,\theta-\Delta\theta\leq\hat{\theta}\leq\theta+\Delta\theta \qquad (14)$$

with $$\underline{d}=\frac{1}{N_d}\sum_{i=1}^{N_d}\hat{d}_i\ ,\ \underline{\theta}=\frac{1}{N_d}\sum_{i=1}^{N_d}\hat{\theta}_i \qquad (15)$$

where d and θ are the initially chosen lag and direction of the variogram with Δd and $\Delta\theta$ as tolerance on lag and direction respectively. $\underline{d}$ and $\underline{\theta}$ are actual lag and direction for the corresponding calculated variogram. N_d is the number of pairs for a particular lag and direction.

Equation (15) avoids the rounding-off of the errors of pre-decided lags and directions (only multiples of the initial lag and fixed values of θ are taken in conventional cases). It is very important to account for every term carefully while calculating variograms. If the data are collected on a regular grid, Δd and $\Delta\theta$ can be taken as zero and $\underline{d}$ and $\underline{\theta}$ become d and θ, respectively. Often, hydrogeological parameters exhibit anisotropy and hence variograms should be calculated at least in 2 to 4 directions to ensure the existence or absence of anisotropy. Of course, a large number of samples are required in that case.

The variogram model is the principal input for both interpolation and simulation schemes. However, modeling the variogram is not a unique process. Various studies tried to correlate the mathematical expressions normally used to describe variogram models to the physical characteristics of the parameters.

Variograms obtained by modeling an experimental variogram are often not unique. It is therefore necessary to validate it. Cross-validation is performed by estimating the random function z at the points where measurements are available (i.e. at data points) and comparing the estimates with the data (e.g. Ahmed and Gupta, 1989). In this exercise, measured values are removed from the data set one by one when the data point is estimated, and the same is repeated for the entire data set. Thus, for all the measurement points, the true measured value (z), the estimated value (z*) and the variance of the estimation error (σ^2) become available. This leads to computing the following statistics, which need to satisfy:

$$\frac{1}{N}\sum_{i=1}^{N}(z_i-z_i^*)\approx 0.0 \qquad (16)$$

$$\frac{1}{N}\sum_{i=1}^{N}(z_i-z_i^*)^2\approx \min \qquad (17)$$

$$\frac{1}{N}\sum_{i=1}^{N}(z_i - z_i^*)^2 / \sigma_i^2 = 1 \qquad (18)$$

$$\frac{\left| z_i - z_i^* \right|}{\sigma_i} \leq 2.0 \qquad \forall i \qquad (19)$$

Various parameters of the variogram model are gradually modified to obtain satisfactory values of equations 16 to 19. During the cross-validation, many important tasks are accomplished, such as:

- Testing the validity of the structural model (i.e. the selected variogram).
- Deciding the optimum neighbourhood for estimation.
- Selecting suitable combinations of additional information, particularly in case of multivariate estimation.
- Sorting out unreliable data.

Description of The Study Area

The Maheshwaram watershed, an experimental research site of about 53 km² in the Ranga Reddy district (Fig. 1) of Andhra Pradesh, India, is presented. This watershed is underlain by granitic rocks and is representative of Southern India catchments in terms of overexploitation of its weathered hard-rock aquifer (more than 700 borewells in use), its cropping pattern, rural socio-economy (based mainly on traditional agriculture), agricultural practices and semi-arid climate. The granite outcrops in and around Maheshwaram form part of the largest of all granite bodies recorded in Peninsular India. Alcaline intrusions, aplite, pegmatite, epidote, quartz veins and dolerite dykes traverse the granite. There are three types of fracture patterns in the area, viz. (i) mineralised fractures, (ii) fractures traversed by dykes, and (iii) late-stage fractures represented by joints. The vertical fracture pattern is partly responsible for the development of the weathered zone and the horizontal fractures are the result of the weathering, as shown by Maréchal et al (2003). Hydrogeologically, the aquifer occurs both in the weathered zone and in the underlying weathered-fractured zone. However, due to deep drilling and heavy groundwater withdrawal, the weathered zone has now become dry. About 150 dug wells were examined and the nature of the weathering was studied. The weathered-zone profiles range in thickness from 1 to 5 m below ground level (bgl). They are followed by semi-weathered and fractured zones that reach down to 20 m bgl. The weathering of the granite has occurred in different phases and the granitic batholith appears to be a composite body that has emerged in different places and not as a single body. One set of pegmatite veins displacing another set of pegmatitic veins has been observed in some well sections. Joints are well developed in the main directions – N 0^0 – 15^0E, NE-SW, and NW-SE that vary slightly from place to place.

The groundwater flow system is local, i.e. with its recharge area at a topographic high and its discharge area at a topographic low adjacent to each other. Intermediate and regional groundwater flow systems also exist since a significant hydraulic conductivity exists at depth. Aquifers occur in the permeable saprolite (weathered) layer, as well as in the weathered-fractured zone of the bedrock and the quartz pegmatite intrusive veins when they are jointed and fractured. Thus only the development of the saprolite zone and the fracturing and interconnectivity between the various fractures allow a potential aquifer to develop, provided that a recharge zone is connected to the groundwater system.

The mean annual precipitation (P) is about 750 mm, of which more than 90% falls during the monsoon season. The mean annual temperature is about 26 °C, although in summer ("Rabi" season from March to May), the maximum temperature can reach 45 °C. The resulting potential evaporation from the soil plus transpiration by plants (PET) is 1,800 mm/year. Therefore, the aridity index (AI = P/PET = 0.42) is within the range 0.2-0.5, typical of semiarid areas (UNEP 1992). Although the annual recharge is around 10-15 % of rainfall, the water levels have been lowered by about 10 m during the last two decades due to intensive exploitation. The transmissivity of the fractured aquifer was measured by seven pumping tests in wells and it varies considerably from about 1.7 x 10^{-5} m^2/s to about 1.7 x 10^{-3} m^2/s, and a low storage coefficient (0.6 %) indicates a weak storage potential of the aquifer (Maréchal et al., 2004). The withdrawal which increases year by year has to be controlled to allow recharge by rainfall to maintain or restore the productive capacity of the depleted aquifer.

Our data show that the weathered-fractured layer is conductive mainly from the surface down to a depth of 35 m, the range within which conductive fracture zones with transmissivities greater than 5 x 10^{-6} m^2/s are observed (Maréchal et al. 2004). The lower limit corresponds to the top of the unweathered basement, which contains few or poorly conductive fractures (T < 5 x 10^{-6} m^2/s) and where only local deep tectonic fractures are assumed to be significantly conductive at great depths, as observed by studies in Sweden (Talbot and Sirat, 2001), Finland (Elo, 1992) and the United States (Stuckless and Dudley, 2002). An unpublished study of the same area has statistically confirmed this result by the analyses of airlift flow rates in 288 boreholes, 10 to 90 m deep, where the cumulative airlift flow rate ranges from 1.5 m^3/h to 45 m^3/h, and the mean value increases drastically in the weathered-fractured layer at depths between 20 and 30 m. Below 30 m, the flow rate is constant and does not increase with depth. In practice, drilling deeper than the bottom of the weathered-fractured layer (30 – 35 m) does not increase the probability of improving the well discharge. The data confirm that the weathered-fractured layer is the most productive part of the hard-rock aquifer, as already shown elsewhere by other authors (e.g. Houston and Lewis 1988, Taylor and Howard 2000).

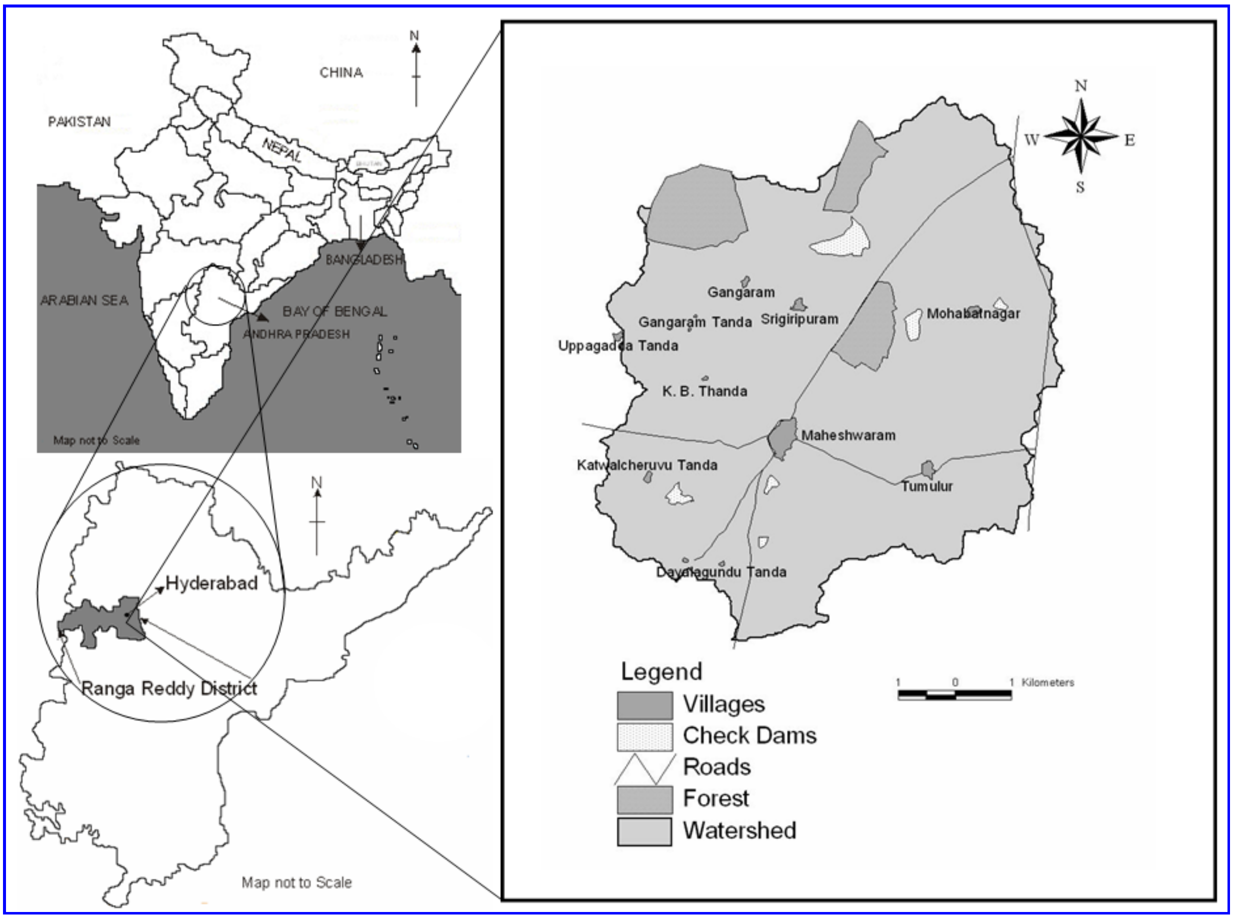

Figure 1. Geographical location of the study area and the watershed

Two different scales of fracture networks have been identified and characterized by hydraulic tests (Maréchal et al., 2004): a primary fracture network (PFN) which affects the matrix at the decimetre scale, and a secondary fracture network (SFN) affecting the blocks at the borehole scale. The latter is the first one described below.

The secondary network is composed of two conductive fractures sets - a horizontal one (HSFN) and a sub-vertical one (VSFN) as observed on outcrops. They are the main contributors to the permeability of the weathered-fractured layer. The average vertical density of the horizontal conductive set ranges from 0.15 m^{-1} to 0.24 m^{-1}, with a fracture length of a few tens of meters (10 - 30 m in diameter for the only available data). This corresponds to a mean vertical thickness of the blocks ranging from ≈4 m to ≈7 m. The strong dependence of the permeability on the density of the conductive fractures indicates that individual fractures contribute more or less equally to the bulk horizontal conductivity ($K_r = 10^{-5}$ m/s) of the aquifer. No strong heterogeneity is detected in the distribution of the hydraulic conductivities of the fractures, and therefore no scale effect was inferred at the borehole scale. The sub-vertical conductive fracture set connects the horizontal network, ensuring a vertical permeability ($K_z = 10^{-6}$ m/s) and a good connectivity in the aquifer. Nevertheless, the sub-vertical set of fractures is less permeable than the horizontal one, introducing a horizontal-to-vertical anisotropy ratio for the permeability close to 10 due to the preponderance of horizontal fractures.

As discussed earlier, the horizontal fracture set is due to the weathering processes, through the expansion of the micaceous minerals, which induces cracks in the rock. These fractures are mostly sub-parallel to the contemporaneous weathering surface, as in the flat Maheshwaram watershed where they are mostly horizontal (Maréchal et al., 2003). In the field, the bore wells drilled with a fairly homogeneous spacing throughout the watershed confirm this conclusion: of 288 wells, 257 (89%) were drilled deeper than 20 meters and 98% of these are productive. In any case, the probability of a vertical well crossing a horizontal fracture induced by such wide-scale weathering is very high.

The primary fracture network (PFN), operating at the block scale, increases the original matrix permeability of $K_m = 10^{-14} – 10^{-9}$ m/s to $K_b = 4 \times 10^{-8}$ m/s. Regarding the matrix storage, it should also contribute to the storage coefficient in the blocks of $S_b = 5.7 \times 10^{-3}$. The storage in the blocks represents 91% of the total specific yield ($S_y = 6.3 \times 10^{-3}$) of the aquifer; storage in the secondary fracture network accounts for the rest. This high storage in the blocks (in both the matrix and the PFN) and the generation of the PFN would result from the first stage of the weathering process itself. The development of the secondary fracture network (SFN) is the second stage in the weathering process: that is why the words “primary” and “secondary” have been chosen to qualify the different levels of fracture networks. The obtained storage values are compatible with those of typical unconfined aquifers in low-permeability sedimentary layers.

The universal character of granite weathering and its worldwide distribution underline the importance of understanding its impact on the hydrodynamic properties of the aquifers in these environments.

Constrained Optimization of Groundwater Levels

Groundwater is mainly found in a coupled system of weathered and fractured granitic rocks in the study area of Maheshwaram. The two zones could be assumed to form a single and often semi-confined aquifer. However, due to over exploitation and successive reduction in the rainfall recharge, the water table has declined and the saturated flow is mainly confined to aquifer consisting of highly fractured rocks only. The general elevation of the area ranges from 600 m (amsl) to 660 m (amsl). The water levels are being monitored through a network of about 55 bore wells. About 30 wells (indicated as IFW) out of the 600 irrigation wells existing in the area have been selected for water level measurement based on the drainage pattern present, variation in rock formation covering the study area (Fig. 2). Later 25 wells (indicated as IFP) taping the fractured aquifer have been drilled up to a depth of 45m to monitor the water level fluctuation. These wells have been drilled based on the recommendation from geophysical investigations. The water level measurements have been carried out on monthly basis for a period of almost 2-3 hydrological cycles. Thus it was decided to reduce the number of IFW wells such that:

- all the wells are monitored in a shortest possible time say one single day,
- discard some of the irrigation wells fitted with pumps as it was difficult to monitor static levels in these wells and
- reduce the cost of monitoring also without loosing the monitoring benefits.

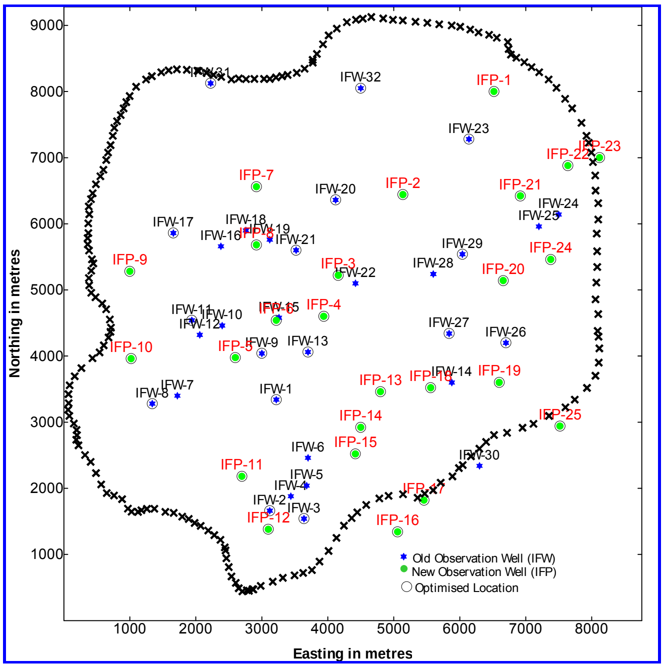

Figure 2. Location of wells indicating all and optimized wells.

Thus to obtain an optimal monitoring network having 25 IFP wells and minimize the IFW wells such that the kriging estimation of water levels provide standard deviation of the estimation error not more than 8 m (against the average standard deviation of 12 m of the water level data) in the entire area. Through a special procedure the IFW wells were removed one by one and the impact with the above constraints were analyzed by comparing the sum of squared difference of the estimation error from a desired value. Finally a network with 25 IFP wells and 15 IFW wells have been evolved for monitoring the water levels every month. The contour map of the standard deviation of the estimation errors (σ_K) from a network of 55 wells as well as from a network of 40 wells are drawn using a suitable kriging method (Fig. 2). It is very clear that using the optimized monitoring network it is still possible to maintain the same magnitudes of σ_K (Fig. 3 & Fig. 4).

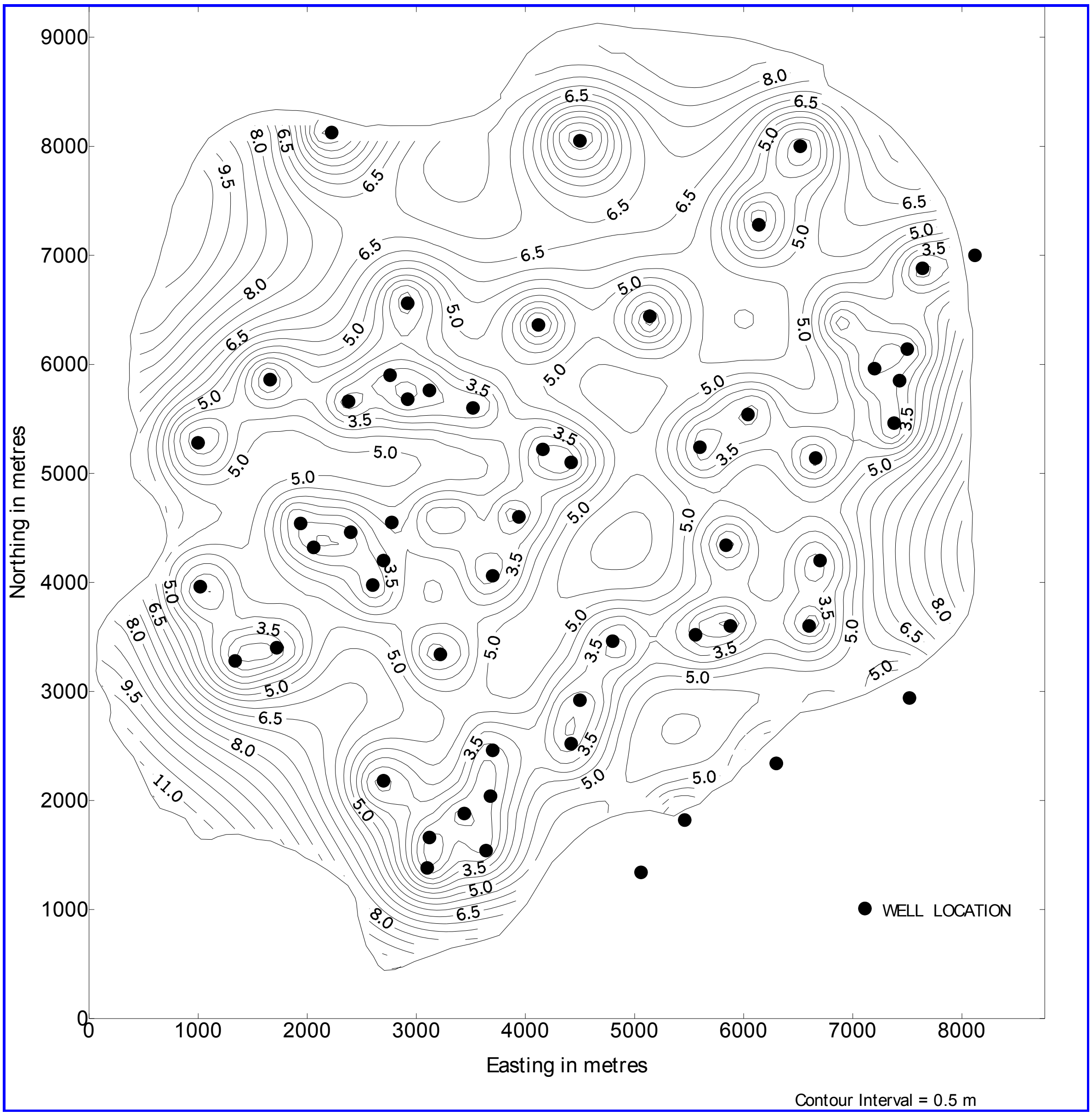

Figure 3. Variation of σ (k) using data from 55 wells

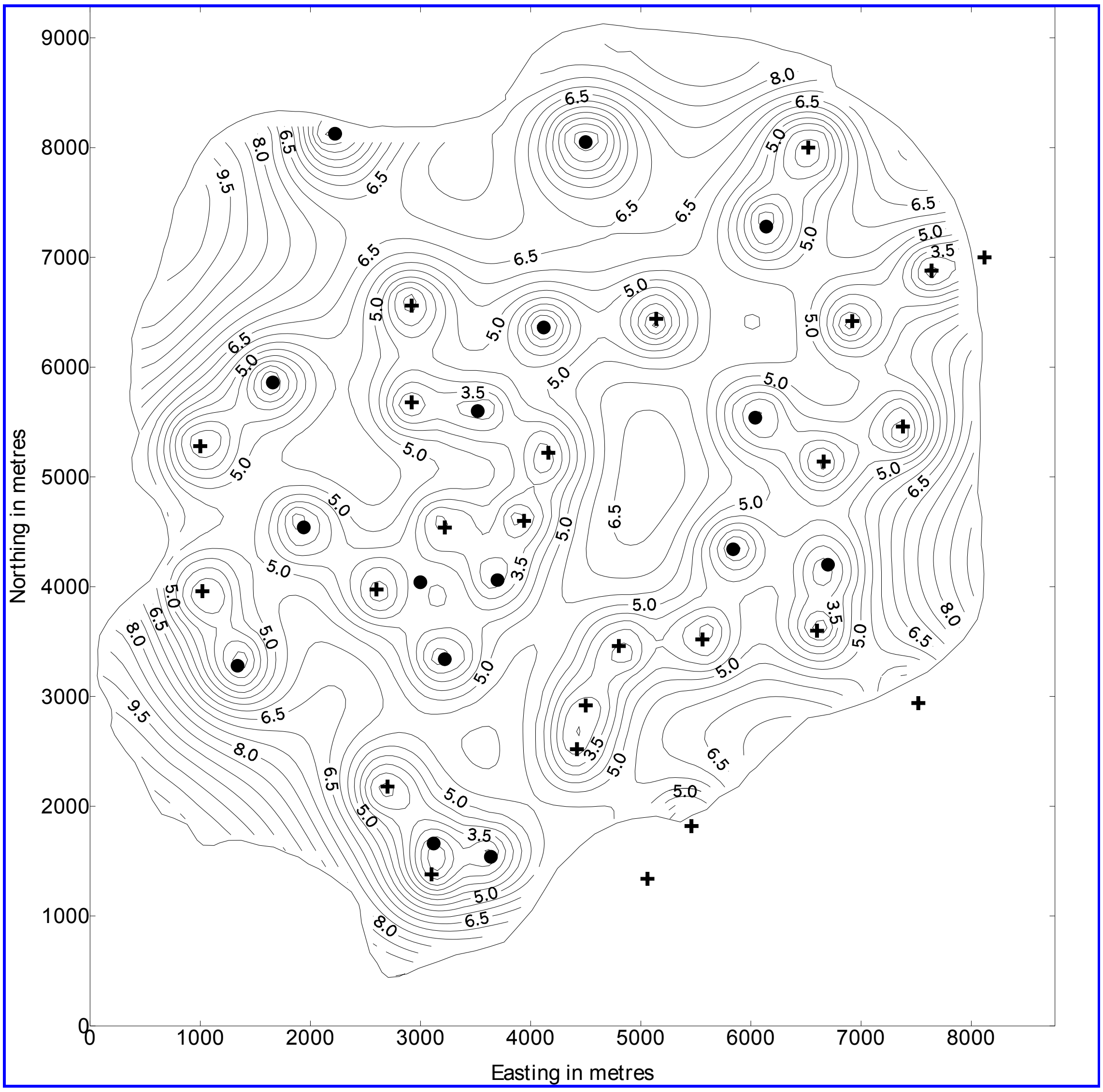

Figure 4. Variation of σ (k) using data from 40 wells.

Assignment of Priority Index To Wells Monitoring Fluoride In Groundwater

Although the area is not much affected with the fluoride problems but the situation is alarming as the mean itself is exceeding 1.5 mg/l, the WHO upper limit. During July 2001, groundwater samples for the analysis of Fluoride were collected from 60 wells fairly distributed in the area. The fluoride content in the area varies from 0.5 mg/l to 2.97 mg/l with a mean value of 1.6 mg/l. The variogram from the values of Fluoride has been calculated and fitted with theoretical model. The variogram parameters are as reported below.

γ (d) = 0.1 + 0.25 Sph (2000) ……….(20)

A cross-validation test (Ahmed and Gupta, 1989) to validate the variogram was thus performed and values given by equations 16 and 19 mainly were analysed for the final variogram decided by the Cross-validation tests. The values of equation 16 indicates the difference of the estimated value from the measured values and a low value of equation 16 suggests that the parameter could be well estimated at this point and need not be measured. Thus, based on the values of equation 16, a priority index could be assigned to the measurement points starting from the highest value of equation 16.

The number of wells to define desired size of the monitoring network could be decided based on the resources available including the man-power and the analyses facilities etc. Then a network could be prepared from Table 1 by picking the wells with decreasing priority. Table 2 shows the statistics from the measured values for comparing the various monitoring networks. It is very clear that the minimum and maximum values measured are present in all the cases. The mean and the variance are increasing consistently as the number of measurement points is decreasing. However, the change in the mean value is almost negligible. Fig. 5 and Fig. 6 show the iso-lines of fluoride drawn based on the 60 measurement points and 30 measurement points respectively coming from the network decided on the basis of the priority index. The two contours show more or less similar distribution and the regionalized picture is almost identical.

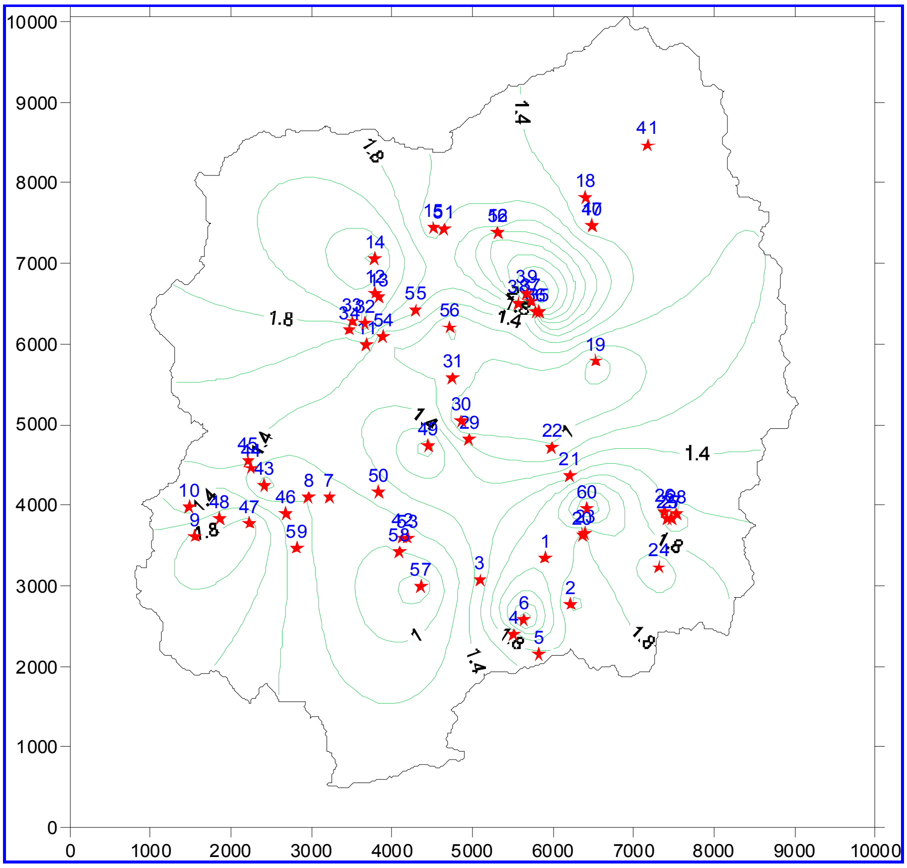

Figure 5: Iso-values of Fluoride content with location of 60 observation wells.

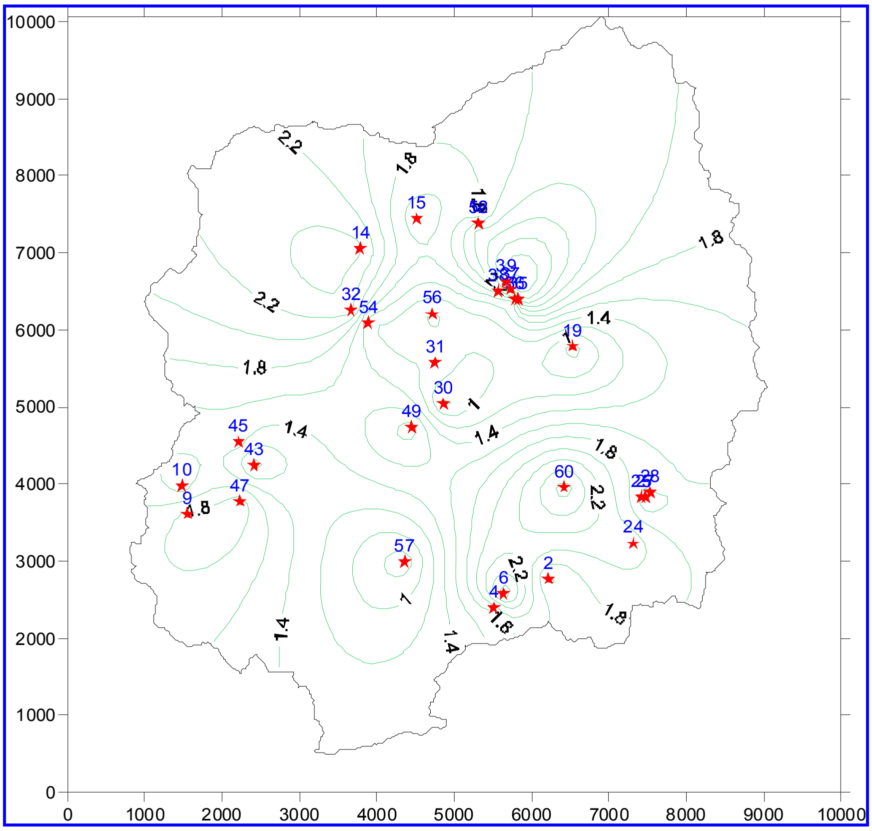

Figure 6: Iso-values of Fluoride content with location of 30 observation (Optimized) wells.

Conclusions

The geostatistical estimation (particularly the block estimates) has the advantage of providing uniform value over the mesh of the aquifer and also over the isolated and variable size of the meshes The most important contribution of the geostatistical estimates is with the fact that it provides the values of the estimation variance and with this magnitude the aquifer model could be calibrated at all the meshes in an unbiased way. Of course, there are many advantages of these methods particularly in Groundwater modeling using them, the model is calibrated at all the meshes and would be free from the biasness of the aquifer modeler. The model parameters should then be changed on the meshes where the above criteria are not satisfied. A faster and unbiased model calibration thus could be achieved.

The constrained optimization of the monitoring network with only 40 wells will ensure that all the wells are measured in the shortest possible time every month. Also that the revised network consists of all the 25 wells without pumping and one has to be only careful for monitoring the 15 private wells fitted with pumps for irrigation. The revised network will provide almost the same accuracy as that obtained from the network of 55 measurement wells.

It is finally recommended that the method of estimating kriging variance on a block/area may be used in data collection network design. Of course, the better way of analysing and designing a network is to discretize the area into a number of blocks and design a network by reducing the estimation variance on an average basis. This procedure could be repeated by reducing the size of discretized blocks until there is no change in the average statistics of the estimation variance.

Using the result of the geostatistical cross-validation, it has been possible to assign a priority index to all these measurement points and depending on the constraints such as financial and logistic, a network could be reduced without loosing the outcome. A number of monitoring networks with less number of wells have been prepared and compared. The reduced network using the reduced monitoring wells still have some clusters that could have been discarded if it was reduced without any scientific study such as the present one.

Acknowledgments

Most of the work reported here are the outcomes of one of the project funded by CEFIPRA, New Delhi. Director, NGRI is thankfully acknowledged to permit the publication of the same. My fellow scientists from NGRI and BRGM, France have provided useful input and discussions at times. The authors are grateful to the organizers of the seminar for their kind invitation to present this work.

Table 1. Priority index for selecting wells for the measurement of Fluoride.

Priority Index	Well No.	X in m	Y in m	F measured	F Estimated	Difference	Σ	Diff/σ
51	1	5902.212	3344.458	1.97	1.897	-0.0726	0.4507	-0.1611
8	2	6219.405	2769.202	1.54	2.284	0.744	0.4627	1.6079
31	3	5094.45	3066.748	1.47	1.726	0.2561	0.4688	0.5463
10	4	5508.345	2391.547	1.81	2.372	0.5619	0.4083	1.376
43	5	5822.104	2147.406	1.73	1.854	0.1245	0.4536	0.2744
1	6	5636.137	2572.363	2.82	1.731	-1.089	0.4052	-2.6872
53	7	3222.197	4091.376	1.14	1.208	0.0678	0.415	0.1633
50	8	2961.081	4091.376	1.2	1.117	-0.0833	0.4031	-0.2067
26	9	1552.695	3613.776	1.89	1.527	-0.3631	0.4376	-0.8296
20	10	1477.736	3974.646	1.27	1.726	0.4561	0.4408	1.0347
58	11	3680.533	5991.858	1.51	1.531	0.0209	0.3994	0.0524
36	12	3794.402	6630.438	2.06	2.238	0.1775	0.368	0.4824
40	13	3830.45	6580.084	2.18	2.038	-0.1419	0.3675	-0.3862
9	14	3783.339	7058.447	2.46	1.878	-0.5825	0.4571	-1.2742
28	15	4516.714	7441.443	1.43	1.716	0.2856	0.3906	0.7312
16	16	5311.125	7380.408	1.76	2.248	0.4876	0.3474	1.4037
38	17	6480.521	7471.96	1.16	1.312	0.1524	0.3499	0.4356
47	18	6400.03	7816.809	1.26	1.148	-0.1121	0.4474	-0.2505
6	19	6530.493	5791.205	0.7	1.615	0.9152	0.5374	1.703
52	20	6366.652	3625.22	1.76	1.828	0.0681	0.3592	0.1896
33	21	6205.481	4360.693	1.54	1.774	0.234	0.4273	0.5476
44	22	5983.275	4710.883	1.02	1.138	0.1179	0.4537	0.2597
55	23	6394.499	3644.293	1.81	1.84	0.0295	0.359	0.0822
5	24	7313.841	3227.728	2.14	1.213	-0.9271	0.4917	-1.8854
30	25	7416.647	3830.45	1.6	1.329	-0.2712	0.3624	-0.7483
41	26	7383.269	3916.663	1.51	1.651	0.1407	0.3776	0.3726
24	27	7472.151	3835.791	1.21	1.627	0.4173	0.3602	1.1585
23	28	7533.377	3888.434	1.65	1.233	-0.4175	0.3755	-1.1118
57	29	4955.595	4810.828	1.15	1.129	-0.0212	0.4182	-0.0506
19	30	4858.321	5038.946	0.73	1.193	0.4625	0.4083	1.1328
14	31	4749.983	5569.189	1.21	0.706	-0.5043	0.4514	-1.1173
11	32	3663.939	6252.783	2.32	1.765	-0.5554	0.3823	-1.4527
60	33	3508.299	6283.301	1.97	1.975	0.0048	0.3762	0.0129
32	34	3469.389	6177.252	1.65	1.897	0.2472	0.3814	0.6482
3	35	5827.826	6402.319	2.69	1.657	-1.0325	0.3624	-2.8494
2	36	5786.055	6393.927	1.54	2.574	1.034	0.3608	2.8655
29	37	5722.159	6530.493	2.74	2.457	-0.2828	0.3695	-0.7654
13	38	5566.71	6499.976	1.81	2.335	0.5246	0.3894	24
22	39	5669.325	6618.994	2.97	2.543	-0.4274	0.3776	-1.1319
42	40	6488.913	7469.672	1.3	1.162	-0.1375	0.35	-0.393
56	41	7174.033	8463.782	1.26	1.289	0.0291	0.5541	0.0526
45	42	4127.806	3596.991	1.02	1.137	0.1167	0.3674	0.3177
25	43	2411.002	4243.964	0.9	1.301	0.401	0.4061	0.9875
35	44	2247.16	4458.35	1.16	1.35	0.1895	0.5065	20
27	45	2213.782	4546.851	1.49	1.159	-0.331	0.382	-0.8665
54	46	2680.51	3891.486	1.31	1.28	-0.0301	0.4085	24
21	47	2230.566	3774.756	1.92	1.467	-0.4532	0.4196	-1.08
59	48	1858.252	3833.502	1.69	1.706	0.0156	0.4114	21
7	49	4449.957	4736.06	1.92	1.102	-0.818	0.454	22

Priority Index	Well No.	X in m	Y in m	F measured	F Estimated	Difference	Σ	Diff/σ
46	50	3833.312	4163.855	1.27	1.383	0.1131	0.4537	0.2494
39	51	4642.981	7420.081	1.65	1.504	-0.1458	0.3883	-0.3755
15	52	5310.935	7378.119	2.25	1.762	-0.4878	0.3474	-1.4043
48	53	4190.939	3590.125	1.15	1.04	-0.11	0.3693	-0.2978
12	54	3888.052	6091.803	1.22	1.769	0.5488	0.4002	1.3712
34	55	4290.884	6413	1.64	1.437	-0.2026	0.4283	20
18	56	4720.038	6204.718	0.96	1.425	0.4646	0.4439	1.0468
17	57	4354.971	2993.506	0.67	1.14	0.4696	0.462	1.0166
49	58	4087.942	3417.7	1.01	0.92	-0.0898	0.3984	27
37	59	2814.978	3468.054	1.28	1.444	0.1641	0.4628	0.3546
4	60	6417.959	3960.913	2.72	1.728	-0.9917	0.4167	-2.3797

Table 2. Statistics of measured values of F for different networks

Network size	Statistics on the parameter values in mg/L			
	Minimum	Maximum	Mean	Variance
60 (july 2001)	0.67	2.97	1.60	0.286
55	0.67	2.97	1.61	0.304
50	0.67	2.97	1.61	0.324
45	0.67	2.97	1.66	0.336
40	0.67	2.97	1.70	0.352
35	0.67	2.97	1.71	0.378
30	0.67	2.97	1.74	0.426

References

Aboufirassi, M., and Marino, M.A., (1984) Cokriging of aquifer transmissivities from field measurements of transmissivity and specific capacity, J. Math. Geol., 16(1), 19-35.

Aboufirrasi, M. and M.A. Marino (1983): Kriging of water levels in Souss aquifer, Morocco. Math. Geol., 15(4), 537-551.

Ahmed, S. and G. de Marsily (1993): Cokriged estimation of aquifer transmissivity as an indirect solution of inverse problem - A practical approach. Water Resour. Res., 29(2), 521-530.

Ahmed, S. & Gupta, C.P. (1989) Stochastic spatial prediction of hydrogeologic parameters: role of cross-validation in krigings. International Workshop on “Appropriate Methodologies for Development and Management of Groundwater Resources in Developing Countries” February 28-March 4, 1989, India, Oxford & IBH Publishing Co. Pvt. Ltd., IGW, Vol. III, 77-90.

Ahmed, S., G. de Marsily and Alain Talbot (1988) Combined use of hydraulic and electrical properties of an aquifer in a geostatistical estimation of transmissivity, Ground Water, 26(1):78-86.

Ahmed, S. (1987): Estimation des transmissivités des aquifères par methodes Géostatistique Multivariables et resolution du Problème Inverse. Doctoral thesis, Paris School of Mines, France, 148pp

Ahmed, S., (1995) An interactive software for computing and modeling a variograms. In Mousavi and Karamooz (eds.) Proc. of a conference on "Water Resources Management (WRM'95)", August 28-30. Isfahan University of Technology, Iran, 797-808.

Ahmed, S. (2004) Geostatistical Estimation Variance approach to optimizing an air temperature monitoring network, Water, Air and Soil Pollution, 158 (1): 387-399.

Ahmed, S. and G. de Marsily (1987): Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. Water Resour. Res., 23(9), 1717-1737.

Das, S., et al., Manual: Hydrograph Network Stations, Central Groundwater Board, 1995.

Delfiner, P. and Delhomme, J.P. (1983) Optimum Interpolation by kriging, in “Display and Analysis of Spatial Data”, J.C. Davis and M.J. McCullough (eds.), 96-114, Wiley and Sons, London.

Delhomme, J.P. (1976): Application de la théorie des variables régionalisées dans les sciences de l'eau. Doctoral thesis, Paris School of Mines, France, 130pp.

Delhomme, J.P. (1978): Kriging in Hyrosciences. Adv. in Water Resour., 1(5), 251-266.

Delhomme, J.P. (1979): Spatial variability and uncertainty in groundwater flow parameters: a Geostatistical approach. Water Resour. Res., 15(2), 269-280.

Elo, S. (1992) Geophysical indications of deep fractures in the Narankavaara-Syote and Kandalaksha-Puolanka zones, Geological Survey of Finland, 13: 43-50.

Galli A. and G. Meunier (1987): Study of a gas reservoir using the external drift method. In Matheron and Armstrong (eds.) Geostatistical case studies, D. Reidel Hingham, 105-120.

Gambolati, G. and Volpi, G. (1979a) Groundwater contour mapping in Venice by Stochastic Interpolators, 1. Theory, Water Resour. Res., 15(2), 281-290.

Gambolati, G., and Volpi, G. (1979b) A conceptual deterministic analysis of the kriging technique in hydrology. Water Resour. Res., Vol.15, no.3, 625-629.

Goovaerts, P. (1997) Geostatistics for Natural Resources Evaluation. New York: Oxford University Press.

Houston, J.F.T. and Lewis, R.T. (1988) The Victoria Province drought relief project, II. Borehole yield relationships, Groundwater, 26(4): 418-426.

Marsily, G de, Lavedan, G., Boucher, M., and Fasanini, G. (1984) Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model, in Geostatistics for Natural Resources Characterization, part 2, edited by G. Verly et al., 831-849, D. Reidel, Hingham, Mass.

Marsily, G de. (1986) Quantitative Hydrogeology: Groundwater Hydrology for Engineers: Academic Press, Inc., London.

Matheron, G. (1963): Traité de Géostatistique appliquée. Mémoires du BRGM, No. 14, Edition Technip.

Mizell, S.A.(1980): Stochastic analysis of spatial variability in two -dimentional Groundwater flow with implications for observation-well network design. Doctoral thesis, New maxico institute of mining and technology, U.S.A.,133 pp.

Neuman, S.P. (1984): Role of Geostatistics in subsurface hydrology. In Verly et al. (eds.) Geostatistics for natural resources characterization, Proc. NATO-ASI, D. Reidel Pub. Co., 787-816.

Roth,C.,J.P. Chiles and C.de Fouquet (1996): Adapting geostatistical transmissivity simulations to finite differece flow simulators. Water Resour. Res. Vol. 32 (10), 3237-3242.

United Nations Environmental Programme (1992) World atlas of desertification. Edward Arnold, Sevenoaks, UK.