Minimizing Error Variance in Estimates by Optimum Placement of Samples – A Comparison of Optimization Techniques

Brandon Wilde

Many optimization algorithms and objective functions have been applied to the problem of field measurement design. This paper compares six optimization algorithms. The objective function is the minimization of the estimation variance at all locations. The algorithms compared are: genetic, random search, modified random search, gradient, Nelder-Mead simplex, and Hooke-Jeeves pattern search. Hooke-Jeeves and modified random search are found to be the most efficient algorithms for this problem.

Introduction

Problems that require sampling are common in all types of engineering, whether it be stream water sampling, ore deposit sampling, ground water sampling, or petroleum reservoir sampling (Ayyub, et.al, 1990). The ground water hydrologist takes samples from wells to determine the extent of a plume of toxic waste that has entered a ground water aquifer. The mining engineer takes ore samples to estimate the expected yield of a vein. The reservoir engineer drills wells to determine the productivity of a petroleum reservoir. In all cases, the goal is to develop a field measurement design (FMD) that consists of the optimal number of measurements and the optimal locations of these measurements.

Many techniques have been created to determine optimal FMD, primarily in the area of groundwater sampling and monitoring. A number of these techniques use stochastic models and consider a number of possible locations to optimize some objective function, usually involving sampling cost (Meyer and Brill, 1988; Meyer et.al, 1994; Storck, et.al, 1997). Another technique is to locate the samples in such a manner that the estimation variance is minimized (Fiereing, 1965; Carerra, et.al, 1984; Virdee and Kottegoda, 1984; Rouhani, 1985; Rouhani and Hall, 1988; Loaiciga, 1989). Other methods include the minimization of the coefficient of variation (Criminisi et.al, 1997), maximization of coverage (Hudak and Loaiciga, 1992) and minimization of cost (Andricevic, 1990; Zhang, 2005).

Some of the optimization techniques applied to FMD include non-linear integer programming (Fiering, 1965, Carrera et.al, 1984), mixed integer programming (Loaiciga, 1989), integer programming (Andricevic, 1990; Hudak and Loaiciga, 1992,1993), linear programming, (Meyer and Brill, 1988), simulated annealing (Meyer et.al, 1994; Storck, et.al, 1997), and genetic algorithms (Zhang and Pinder, 2005; Catania and Paladino, 2008).

Common among all of the methods previously applied is the assumption of stationarity, that is, the assumption that the model statistics are constant in the modeling domain. Each method uses this assumption. For instance, the assumption of stationarity is necessary for the generation of stochastic models of the geology. It is also a necessary assumption to minimize the estimation variance. This assumption is not always justified. Natural phenomena show complex patterns of spatial variation that may not be represented well by stationary parameters. Common departures from stationarity include: different directions of continuity; differing behaviours of high and low valued areas; abrupt changes in the variogram across different rock types; or smooth variations in the variogram direction within the modeling domain (Boisvert, et.al, 2009).

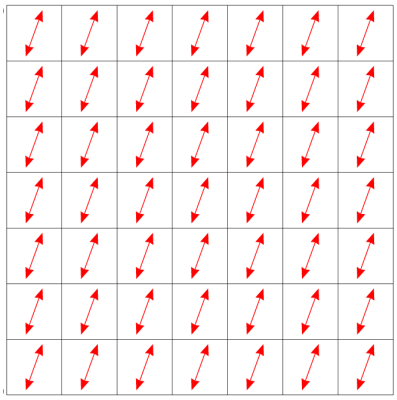

Figure 1: Field with a constant direction of continuity.

It is interesting to consider FMD in the presence of such non-stationarity, particularly as affected by different directions of continuity. All previous work has considered a constant direction of continuity (Figure 1) or no direction of continuity as part of the assumption of stationarity. In order to consider determining the optimal FMD in the presence of varying directions of continuity (Figure 2), it is important to first determine which optimization method is best suited to the problem. That is the aim of this study. This study does not consider FMD in the presence of locally varying

directions of continuity. This study is directed at determining which optimization method is most applicable for FMD where direction of continuity is constant. This is done by applying various optimization techniques to optimal FMD for a field with a constant (or no) direction of continuity. These techniques are compared to see which method gives the best results. That method can then be applied to the variable direction field to determine the optimal FMD. The method that works best for a field with constant direction of continuity will also work best for a field with varying direction of continuity. A small change to the calculation of the objective function is the only difference. To simplify further discussion, fields with no continuity direction (isotropic) are considered to have constant continuity direction.

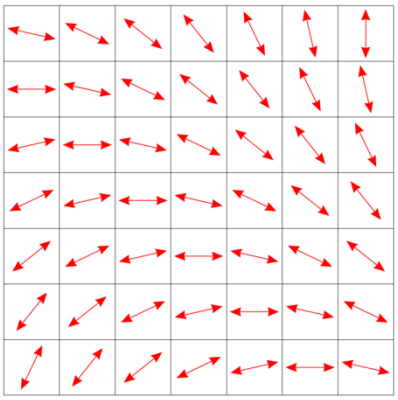

Figure 2: Field with varying directions of continuity.

This study is somewhat unconventional in the fact that the optimum locations of samples for a field with constant continuity direction are already known. McBratney et.al (1981) showed that for a field with no continuity direction and monotonic increasing semivariogram, the optimal sample locations are located on an equilateral triangular sampling grid. This is based on the fact that estimation variance is proportional to distance. As distance to a sampled location from an unsampled location increases, so too does the estimation variance at that unsampled location. The maximum distance between an interpolated point and its nearest sampling point can be minimized by sampling on an equilateral triangular grid where the maximum distance to a sample is 0.63*d* where *d* is the distance between samples. A square grid is also close to optimal with a maximum distance to a sample of 0.71*d.*

These configurations are optimal only where there is no direction of continuity. Where a field has a direction of continuity (anisotropy), McBratney et.al (1981) shows that the optimum sample locations fall on an isosceles triangular grid which is aligned with the direction of continuity and stretched in proportion to the anisotropy ratio along this direction.

Six new optimization techniques are applied to the problem. The optimization techniques considered are: gradient, genetic algorithm, Nelder-Mead simplex, Hooke Jeeves pattern search, randomized search, and a modified version of randomized search. The gradient method requires the determination of a gradient and step size for every location at every iteration while the other five methods are based solely on evaluations of the objective function. The genetic algorithm is the only one of these techniques which has been applied before (Zhang and Pinder, 2005; Catania and Paladino, 2008). It is reapplied here in order to compare to the other methods.

Comparison Example

In order to compare these optimization methods, I will consider an area that is 1000m x 1000m in size. This area is discretized into a grid with 400 cells, each with a side length of 50m. The estimation variance will be calculated at the center of each of these 400 cells and the sum of the 400 estimation variances will be taken as the objective function. The locations of 25 samples will be optimized.

Objective Function

The estimate variance can be calculated at any location and does not depend on the value of the estimate at that location. Most interpolation methods are a weighted average and have the general form shown in equation (1.1) where *u* is a location vector, $Z(u)^*$ is an estimate at location *u*, $Z(u_i)$, $i=1,\ldots,n$ are *n* data values and λ_i refer to weights. The optimal estimate at a location depends on the choice of the weights. In choosing the weights we could consider the closeness of the data to the location being estimated, the redundancy between data values, anisotropic continuity (direction), and the magnitude of the anisotropic continuity. Anisotropic continuity, or anisotropy, refers to the condition of a geological variable being more continuous in one direction than another.

$$Z^*(u) = \sum_{i=1}^{n} \lambda_i \cdot Z(u_i) \qquad (1.1)$$

A common method for choosing the weights is called kriging and is based on the idea of minimizing the error between the estimated value at a location and the unknown true value at that location. This error is defined in equation (1.2) where *Z(u)** is the estimate at a location and *Z(u)* is the true value at that location. Expanding this out and replacing *Z(u)** with the relation from (1.1) we get the equation shown in (1.3). Note that $E\{[Z(u)]^2\}$ is now equal to $C(0)$ based on the relation $C(h) = E\{Z(u) \bullet Z(u+h)\}$ which is the definition of covariance. Replacing other terms with their equivalent covariance notations we get equation (1.4). In order to minimize this equation we take the derivative with respect to the weights, λ_i, and setting this equal to zero we get (1.5). This system of *n* equations with *n* unknown weights is known as the simple kriging system. These weights, when used to generate an estimate of the variable at the unsampled location (1.1), give the minimized error variance. The corresponding minimized estimation variance (which is what we are interested in for this project) is defined in (1.6) (Deutsch, 1998).

$$E\left\{[Z^*(u) - Z(u)]^2\right\} \tag{1.2}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j E\{Z(u_i)\cdot Z(u_j)\} - 2\cdot\sum_{i=1}^{n}\lambda_i E\{Z(u)\cdot Z(u_i)\} + C(0) \tag{1.3}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j C(u_i,u_j) - 2\cdot\sum_{i=1}^{n}\lambda_i C(u,u_i) + C(0) \tag{1.4}$$

$$\sum_{j=1}^{n}\lambda_j C(u_i,u_j) = C(u,u_i), \quad i = 1,...,n \tag{1.5}$$

$$\sigma^2{}_{SK}(u) = C(0) - \sum_{i=1}^{n}\lambda_i \bullet C(u,u_i) \tag{1.6}$$

The equation given in (1.6) defines the estimation variance at one unsampled location. For this project, we are interested in minimizing the sum of the estimation variances at every unsampled location:

$$ObjFun = \min\sum_{\alpha}^{nloc}\sigma^2{}_{SK}(u_\alpha) = \min\sum_{\alpha}^{nloc}\left[C(0) - \sum_{i=1}^{ndat}\lambda_{i,\alpha}\bullet C(u_\alpha,u_{i,\alpha})\right]$$

Where *nloc* is the number of unsampled locations (in this case, 400) and *ndat* is the number of data considered at each unsampled location (for this example, 25). The *C(0)* term is a constant and can be removed yielding the pseudo-variance:

$$ObjFun = \min\sum_{\alpha}^{nloc}\left[-\sum_{i=1}^{ndat}\lambda_{i,\alpha}\bullet C(u_\alpha,u_{i,\alpha})\right]$$
$$= \max\sum_{\alpha}^{nloc}\left[\sum_{i=1}^{ndat}\lambda_{i,\alpha}\bullet C(u_\alpha,u_{i,\alpha})\right]$$

Instead of minimizing a negative objective function, a positive function can be maximized with the same result. For this study, the author will maximize the positive function.

Base Case

In order to compare the different optimization techniques, a base case is established based on knowledge of the optimum sample locations. The base case is for a field with no direction of continuity (isotropic). An equilateral sample grid and a square sample grid are considered for the base case. The objective function value for the square sample grid (Figure 3a) is 290.6 while the objective function value for the equilateral sample grid (Figure 3b) is 288.1. This conflicts with the findings of McBratney et.al (1981) which demonstrated that the equilateral sample grid would provide the optimal solution. The reason for this conflict is an artifact of the model, namely the edge effects. If more samples were taken such that the samples extended past the grid domain, the equilateral sample grid would provide the more optimal solution. In any case, the values are very similar and a base case has been established against which the various optimization methods may be compared.

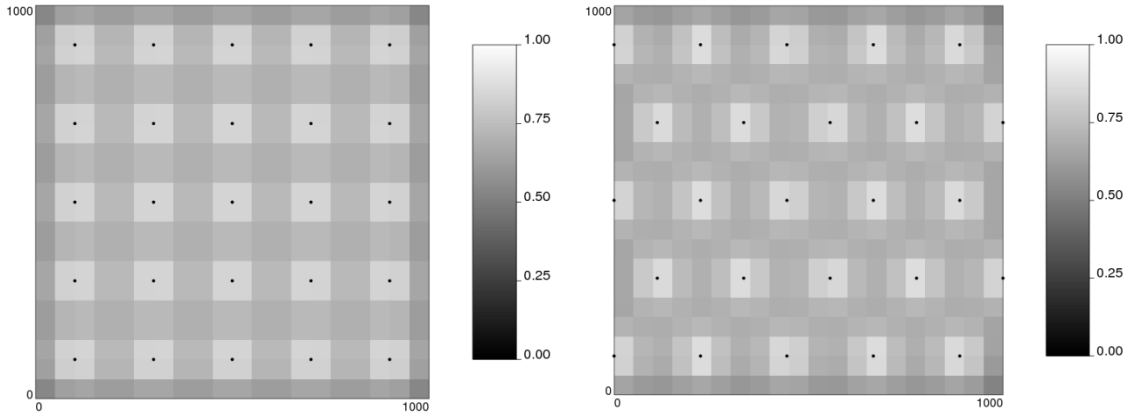

Figure 3: Pseudo-variance base case values for the a) square sample grid and b) equilateral sample grid. Note the edge effects along the top and bottom of plot b.

Optimization Algorithms

It is now useful to examine the details of how each optimization algorithm is applied. Two metrics are used to evaluate the performance of each method. One is the improvement in the objective function for a given number of iterations of the algorithm and the other is the number of objective function evaluations required to achieve a set value for the objective function. These values are different since one iteration of the algorithm may require multiple evaluations of the objective function.

Genetic Algorithm

The genetic algorithm uses binary representations of the decision variables concatenated together into one long string called a chromosome. Cross-over and mutation operations are performed on the chromosomes in an effort to create a chromosome with increased fitness. The decision variables in this case are the coordinates of the sample points. In order to represent the coordinates in binary, a decision had to be made regarding the precision of the coordinates. A trade-off must be made between computational efficiency and decision variable precision. Increased precision in the decision variables results in longer chromosomes and vice versa. For this study, a precision on the order of ~5m is deemed appropriate. This enables each sample coordinate to be represented with 16 bits; 8 bits for the x-coordinate and 8 bits for the y-coordinate. An 8-bit binary number is capable of representing 256 separate values, from 0 to 255. Therefore, the x- and y-coordinates of the sample points will have a precision of 1000/256 = ~3.9m. This is within the required 5m precision. This is the equivalent of having 512 possible sample locations from which the optimal 25 are chosen.

This algorithm is implemented by randomly generating 50 coordinates (x- and y-coordinates for the 25 sample locations) each for 10 different 'individuals' in the population. These coordinates are converted to their binary representations, or chromosomes. The chromosomes have a length, *L*, of 400 bits (50 coordinates, 8 bits each). The fitness of each of these 10 individuals is evaluated. The roulette-wheel scheme is used to select 10 parents from the population. These 10 parents are then randomly paired to create 5 pairs. The cross-over probability of 0.9 is applied to determine whether the cross-over operation is performed for each pair. The mutation probability of 0.05 is applied to determine whether the mutation operation is performed for each individual. The fitness of these 10 individuals is evaluated and the process repeats.

As the algorithm proceeds, a record of the most fit chromosome is kept as the best-so-far. Each new population is checked to see whether the best-so-far is surpassed by any of the population. If there is no improvement within the population, that is, if none of the individuals in the new population are more fit than the best-so-far, the least fit individual is replaced with the best-so-far. This process of 'elitism' aids the algorithm in the determination of the best configuration. The algorithm proceeds for a predetermined number of iterations or until the objective function experiences no further improvement.

The result of applying the genetic algorithm with a population of 10 individuals is shown in Figure 4b. The objective function was increased from an initial best value of 265.8 to 287.5 in 1000 iterations, very near the base case optimum of 290. The objective function was evaluated 3621 times to reach a value of 285.

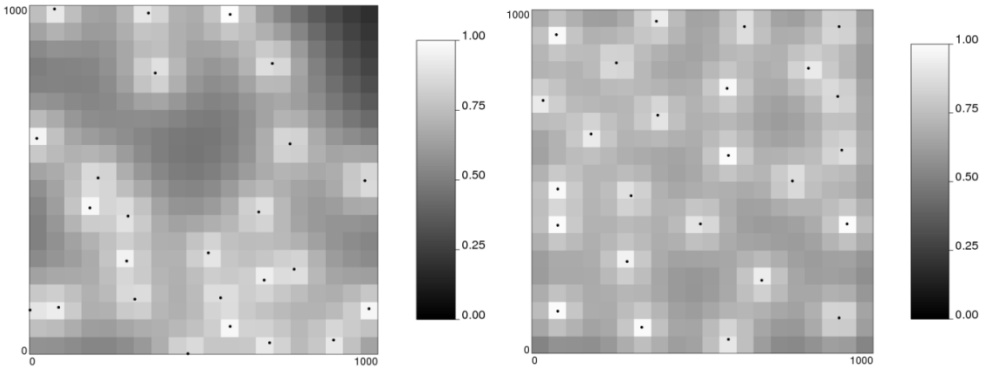

Figure 4: a) Initial best and b) final result produced by the genetic algorithm with a population of 10.

The same process was repeated using a population of 20 individuals. This modification allowed the objective function to be increased even further to 289.8 in 1000 iterations. The objective function was evaluated 5661 times to reach a value of 285.

Nelder-Mead Simplex

This optimization method is a derivative free method which uses only the objective function evaluation. It uses the concept of a simplex which is a geometric object determined by an assembly of *n+1* points, $\boldsymbol{p}$, in the *n* dimensional space (Chong et.al, 2008). As applied to this problem, each point, $\boldsymbol{p}_i$, consists of the *x,y* coordinates of the *n* data. There will, therefore, be *2n+1* initial configurations, or points. These points will be generated randomly within the field domain. The objective function is evaluated for the *2n+1* configurations. The centroid of the *2n* best points is calculated and the reflection move performed. The objective function is evaluated for the new point determined by the reflection move. This value is compared to the other objective function values. This comparison determines whether the expansion, outside contraction, or inside contraction move is performed. If the contraction moves both fail, the shrinkage move is performed (Chong & Zak, 2008).

For this study, 51 different configurations of the 50 coordinates are randomly generated. The objective function for each of these configurations is determined. The centroid of the 50 best configurations is calculated and the reflection move performed. The objective function of the configuration created by the reflection move is evaluated. If this value is better than any of the other values, the expansion move is performed. If it is better than the second worst configuration and worse than the best configuration, this configuration is kept and replaces the worst. If it is better than the worst and worse than the second worst, the outside contraction move is used. If it is worse than the worst, the inside contraction move is performed. If both contraction moves are unable to generate an improved configuration, the shrinkage move is used where all coordinates are moved half the distance towards the best configuration. The results of applying the Nelder-Mead Simplex method are shown in Figure 5. The objective function value is increased from an initial best of 273.6 to 288.9 in 1000 iterations. The objective function was evaluated 643 times to reach a value of 285.

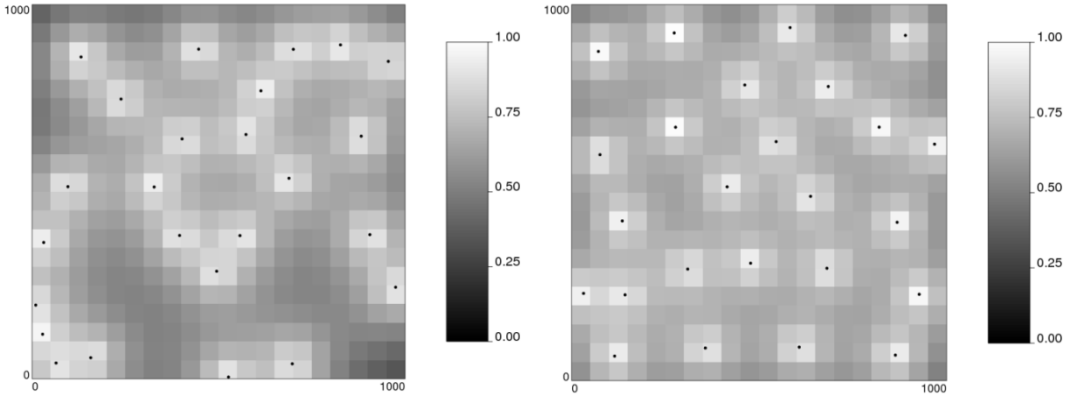

Figure 5: a) Initial best and b) final configuration for the Nelder-Mead Simplex method.

Hooke Jeeves Pattern Search

This optimization method is also a derivative free method. It uses a variety of moves to locate the optimal point. It is similar to exploring unknown territory in the fact that it moves to a base point, explores the surrounding area, then moves to a new base point to explore that area. The algorithm starts with a randomly generated base point. The exploratory move is performed first where each decision variable is visited in turn, increasing and decreasing the value of that variable by a certain step size and keeping the value that gives the best objective function value. Once each decision variable is visited, a new base point is determined based on the current and previous base point. This is called the pattern move. The exploratory move is then repeated at the new base point. If the exploratory move is unable to locate a better point, the step size is reduced and the exploratory move performed again. This procedure is repeated until the step size is sufficiently small.

The algorithm is applied to this problem by randomly generating 25 x- and y-coordinates. Each of these 50 coordinates is then visited in turn and adjusted up and down by a step size of 25m. This exploratory move locates a new base point. This base point and the starting point are used to locate the third base point. The exploratory move repeats at this base point and the algorithm proceeds, reducing the step size, until the step size is less than 0.5m.

The results of applying the Hooke Jeeves method are shown in Figure 6. The objective function value increases from 253.3 to 290.8 in 1000 iterations. Interestingly, this value is achieved after only 10 iterations of the algorithm. To reach a value of 285, 222 evaluations of the objective function are performed.

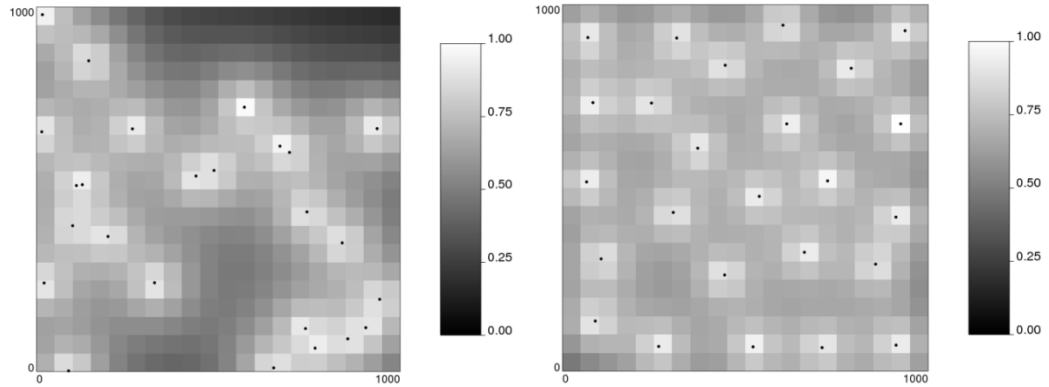

Figure 6: a) Initial and b) final configurations for the Hooke Jeeves method.

Randomized Search

This optimization method is also a derivative free method which uses only the objective function evaluation. It proceeds by considering random sample locations. A random initial configuration is created and the objective function evaluated for that configuration. The next configuration is then randomly generated and its objective function value determined. If the objective function value of the second arrangement is an improvement over that of the first, the second configuration is kept. If the objective function value of the second configuration is not an improvement, it is rejected and a new configuration is chosen. This process proceeds until no further improvements can be made or until a preset number of iterations is reached.

For this study, 25 sample locations are randomly generated within the 1000 x 1000m field. The value of the objective function is determined for this configuration. 25 different sample locations are then randomly generated and the objective function value determined for this configuration. The algorithm proceeds in this manner, keeping track of the best-so-far configuration, until the preset number of iterations is reached. The best-so-far at the completion of the algorithm is the assumed optimal solution.

The objective function increases from 253.3 to 276.3 in 1000 iterations. After 100,000 objective function evaluations, the value was only 279.4. The results of applying this method are shown in Figure 7.

Modified Randomized Search

Consider the step in the randomized search algorithm where a new point is generated. For an *n* dimensional problem, *n* new values are generated which define the point. This part of the algorithm will be modified where instead of generating *n* new values to create a new point, only 1 of the *n* dimensions

will be changed from the previous point. The dimension to be changed will be randomly generated as well as the value for this dimension. The other *n-1* dimensions will be held constant and the objective function evaluated for the new point. If the objective function is improved the point is kept, otherwise it is rejected and the process repeated.

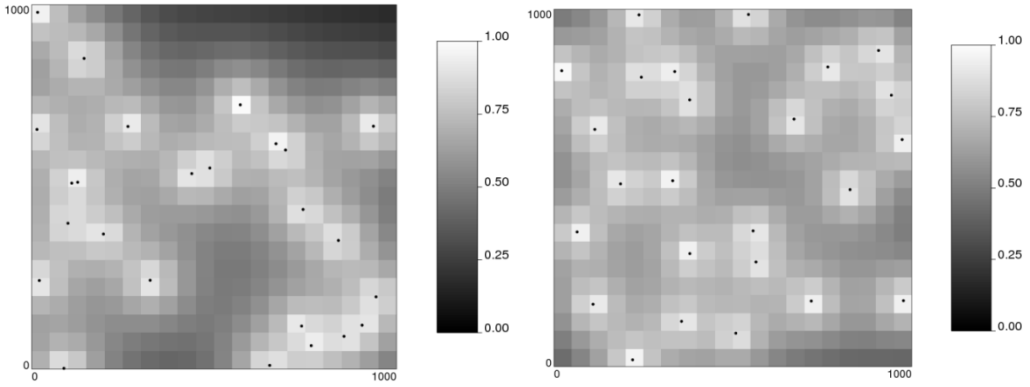

Figure 7: a) Initial and b) final configurations for the randomized search method.

For this study, 25 sample locations are randomly generated within the 1000 x 1000m field. The objective function value for this configuration is evaluated. One of the 50 coordinates is randomly selected and a new value randomly generated. The objective function for this new configuration is evaluated. If it is improved, the change is kept, otherwise the change is rejected. The process repeats until the specified number of iterations is reached, keeping track of the best-so-far configuration. The best-so-far at the completion of the algorithm is the assumed optimal solution.

The objective function increases from 253.3 to 288.6 in 1000 iterations. It takes 128 objective function evaluations to reach a value of 285. The results of applying this method are shown in Figure 8.

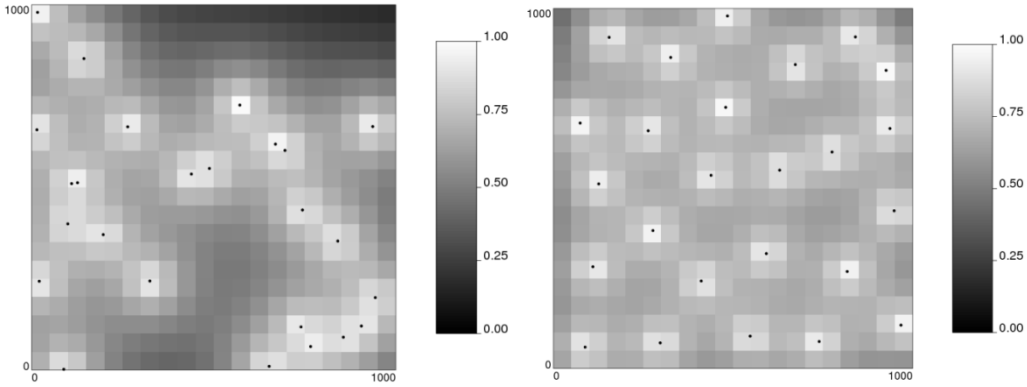

Figure 8: a) Initial and b) final configurations for the modified randomized search method.

Gradient

The gradient method is an optimization method which depends on the function and gradient values at each iteration (Chong et.al, 2008). As such, it is necessary to find the gradient, or first derivative, of the objective function with respect to the data locations. For this method, I have decided to limit the number of data considered to two. This will greatly simplify the determination of the gradient of the objective function. The objective function now becomes:

$$ObjFun = \max \sum_{i}^{nloc} \left[\lambda_\alpha C(u_\alpha, u_i) + \lambda_\beta C(u_\beta, u_i) \right]$$

Note the notation change where *i* now represents the unsampled locations, α represents the closest data to the location, and [illegible] represents the next closest data to the location.

In order to derive the gradient, we must first examine how the covariance is defined. As mentioned previously, the definition of covariance is $C(h) = E\{Z(u) \bullet Z(u+h)\}$. It is common to define

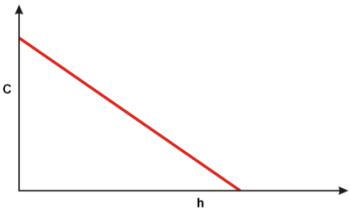

Figure 9: Covariance structure for the gradient method.

covariance, *C,* as a function of separation distance, *h.* In order to keep the calculation of the objective function simple, I have chosen to employ a linear relationship between *C* and h^2 as shown in Figure 9. This will eliminate the need for the calculation of the square root of the anisotropic distance, h^2 as defined in (1.7) where *ax, ay,* and *az* represent anisotropic distances. For this study, I will only be considering 2-D data placement, we can therefore drop the z component of (1.7). We can convert this calculation for h^2 into matrix notation as shown in (1.8) where *V* represents the vector between points $P_{\square}$ and $P_{\square\square}$ and *R* represents the anisotropic rotation matrix taking into account both the magnitude and direction of anisotropic continuity. Covariance between any two points can then be calculated quickly using th (1.9) where *m* represents the slope of the *C* vs. h^2 plot and *b* represents the *C*-axis intercept.

$$h^2 = \left(\frac{x}{ax}\right)^2 + \left(\frac{y}{ay}\right)^2 + \left(\frac{z}{az}\right)^2 \tag{1.7}$$

$$h^2 = V^T R^T R V \tag{1.8}$$

$$C = \max\left(mV^T R^T RV + b, 0\right) \tag{1.9}$$

It is also important to examine how the weights, $\square$, are defined. We already saw in (1.5) how to solve for the weights. We must solve for *n* weights when *n* data are considered. In order to simplify calculations, I have decided to limit the number of data considered to two. This allows me to solve for the weights ahead of time based on the relations shown in (1.10)-(1.12) where $C_{12}=C_{21}$ and $C_{11}=C_{22}=1$.

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} C_{1,i} \\ C_{2,i} \end{bmatrix} \tag{1.10}$$

$$\lambda_1 = C_{1,i} - C_{12}\lambda_2 \tag{1.11}$$

$$\lambda_2 = \frac{C_{2,i} - C_{21}C_{1,i}}{1 - C_{12}^2} \tag{1.12}$$

Now that the weights and covariance functions are defined, we can take the derivative of the objective function to obtain the gradient. We take the derivative with respect to the data locations, *P*. If we let *O* represent the objective function, then the derivative of the objective function is defined as:

$$\frac{\partial O}{\partial P_k} = \lambda_\alpha \frac{\partial C_{\alpha i}}{\partial P_k} + \frac{\partial \lambda_\alpha}{\partial P_k} C_{\alpha i} + \lambda_\beta \frac{\partial C_{\beta i}}{\partial P_k} + \frac{\partial \lambda_\beta}{\partial P_k} C_{\beta i} \tag{1.13}$$

Where:

$$\frac{\partial C_{\alpha\beta}}{\partial P_\alpha} = 2mR^T RV \tag{1.14}$$

$$\frac{\partial C_{\alpha\beta}}{\partial P_\beta} = -2mR^T RV \tag{1.15}$$

$$\frac{\partial \lambda_\beta}{\partial P_k} = \frac{\frac{\partial}{\partial P_k}\left\{C_{\beta i} - C_{\alpha\beta}C_{\alpha i}\right\}\bullet\left(1 - C_{\alpha\beta}^2\right) + \frac{\partial}{\partial P_k}\left\{1 - C_{\alpha\beta}^2\right\}\bullet\left(C_{\beta i} - C_{\alpha\beta}C_{\alpha i}\right)}{\left(1 - C_{\alpha\beta}^2\right)^2} \tag{1.16}$$

$$\frac{\partial \lambda_\alpha}{\partial P_k} = \frac{\partial C_{\alpha i}}{\partial P_k} - C_{\alpha\beta}\frac{\partial \lambda_\beta}{\partial P_k} - \frac{\partial C_{\alpha\beta}}{\partial P_k}\lambda_\beta \tag{1.17}$$

With the gradient defined, we have everything we need to perform the gradient algorithm. It will proceed as follows:

1. An initial point or configuration is chosen
2. The gradient at each point is calculated, the search direction is taken to be the negative of the gradient. If the gradient is zero, we have found the optimal configuration and the algorithm terminates.
3. The step size at each point is determined using the golden section 1-D search
4. The next point is determined by adding the product of the step size and direction to the previous point

5. The iteration number is increased and we return to step 2

As mentioned, the objective function is calculated differently for this method than the others. As such, direct comparison of objective function values with those from the other methods is not possible. A new base case was established for this method based on an equilateral grid. The objective function value for this configuration is 272. The results of applying this method are shown in Figure 11. The objective function value increased from an initial value of 204.3 to 266.3 with 17 iterations of the algorithm. The objective function stopped increasing at this. The

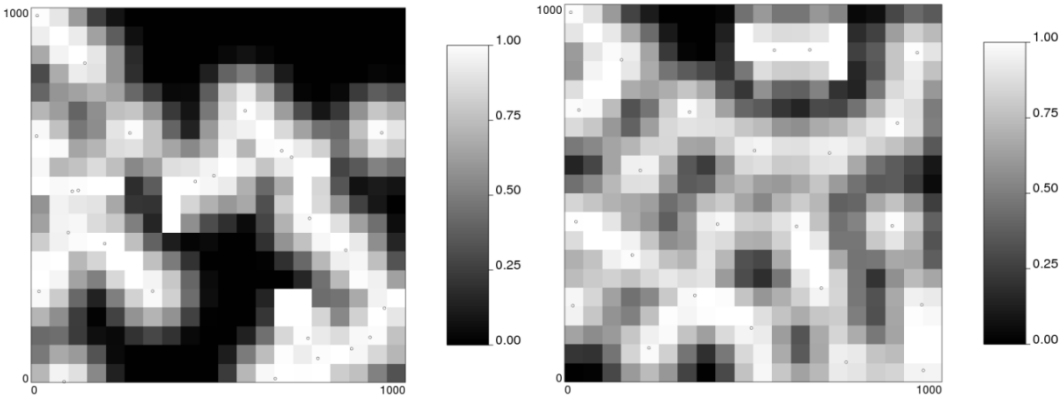

Figure 11: Results of applying the gradient method. a) Initial and b) final configurations.

objective function was evaluated 9550 times during these 17 iterations. The line search to determine the step size requires many evaluations of the objective function.

Discussion

The optimization results for the five derivative-free methods are summarized in Table 1. The randomized search method is the only truly poor method among them. Even after many thousands of iterations, the algorithm was unable to approach an optimal solution. This is likely due to the fact that no neighborhoods were used in the generation of new points. Each configuration was completely random with no consideration of previous configurations. The other four methods performed comparably well at generating optimal configurations.

Table 1: Optimization results for the five derivative-free methods.

Method		Objective Function value after 1000 iterations	# of objective function evaluations to reach 285
Hooke Jeeves		290.8	222
Nelder-Mead Simplex		288.9	643
Randomized Search		276.3	>100,000
Modified Randomized Search		288.6	128
Genetic Algorithm	Pop = 10	287.5	3621
	Pop = 20	289.8	5661

The genetic algorithm was the next poorest method after randomized search. It required many times more iterations than the best three methods. Another drawback of this method is the lack of precision. A given string of binary digits is capable of representing only a limited number of values. The trade-off between efficiency and precision would require reduced efficiency to gain increased precision.

The Hooke Jeeves, Nelder-Mead and modified randomized search methods performed comparably. Each generated an optimal solution with great efficiency. The Hooke Jeeves was determined to be the most efficient method. It generated the most optimal solution with a minimal number of objective function evaluations. Although the modified randomized search was able to reach the threshold of 285 with less function evaluations, the Hooke Jeeves method is deemed better due to its ability to find a more optimal solution. The modified randomized search method required more than 3200 iterations to reach a solution as optimal as that created by the Hooke Jeeves method in 10 iterations. The Nelder-Mead method was unable to reach an equivalently optimal solution after 10000 iterations. This disparity is reduced by realizing that each iteration of the Hooke Jeeves method required approximately 75

objective function evaluations compared to the 1.06 evaluations per iteration for the Nelder-Mead method.

The gradient method performed well in that the objective function increased a great deal with few iterations of the algorithm. It has two main disadvantages: the optimal value is not approached before the objective function levels off; a large number of objective function evaluations are required to determine the step size.

Conclusions

Field measurement design is an important aspect of many engineering studies. Many different optimization methods have been applied to the solution of numerous objective functions. A number of optimization methods had not been applied to this problem. This study was performed to determine the suitability of the methods that had not been tried (Nelder-Mead, Hooke Jeeves, randomized search, modified randomized search, gradient) and to compare to a method that had (genetic algorithm). The Hooke Jeeves pattern search method was determined to be the most appropriate for optimizing the locations of the samples. The modified randomized search and Nelder-Mead simplex methods also performed well. It is recommended that the Hooke Jeeves method be applied to the determination of the optimal FMD in the presence of varying continuity directions.

References

Andricevic, R. 1990. Cost-effective network design for groundwater flow monitoring. Stochastic Hydrol. Hydraul.

Ayyub, B.M., McCuen, R.H. 1990, *Optimum Sampling for Structural Strength Evaluation.* Journal of Structural Engineering. V.116. No.2.

Boisvert, J.B., Manchuk, J.G., Deutsch, C.V. 2009. *Kriging in the presence of locally varying anisotropy using non-Euclidean distances.* Mathematical Geosciences. In press.

Carerra, J., Usunoff, E., Szidarovszky, F. 1984. *A method for optimal observation network design for groundwater management.* Journal of Hydrology. 73.

Catania, F., Paladino, O. 2008. *Optimal sampling for the estimation of dispersion parameters in soil columns using an Iterative Genetic Algorithm.* Environmental Modeling and Software. 24.

Chong, E.K.P., Zak, S.H. 2008. *An Introduction to Optimization.* 3rd Ed. John Wiley & Sons. New Jersey.

Criminisi, A., Tucciarelli, T., Karatzas, G.P. 1997. *A methodology to determine optimal transmissivity measurement locations in groundwater quality management models with scarce field information.* Water Resources Research. V.33. No.6.

Deutsch, C.V., Journel, A.G. 1998. *GSLIB: Geostatistical Software Library and User's Guide.* Oxford University Press. New York.

Fiering, M.B. 1965. *An optimization scheme for gaging.* Water Resources Research. V.1. No.4.

Hudak, P.F., Loaiciga, H.A. 1992. *A location modeling approach for groundwater monitoring network augmentation.* Water Resources Research. V.28. No.3.

Hudak, P.F., Loaiciga, H.A. 1993. *An optimization method for monitoring network design in multilayered groundwater flow systems.* Water Resources Research. V.29. No.8.

Loaiciga, H.A. 1989. *An optimization approach for groundwater quality monitoring network design.* Water Resources Research. V.25 No.8

McBratney, A.B., Webster, R., Burgess, T.M. 1981. *The design of optimal sampling schemes for local estimation and mapping of regionalized variables – I.* Computers & Geosciences. V.7. No.4.

Meyer, P.D., Brill, E.D. 1988. *A Method for Locating Wells in a Groundwater Monitoring Network Under Conditions of Uncertainty.* Water Resources Research. V.24. No.8.

Meyer, P.D., Valocchi, A.J., Eheart, J.W. 1994. *Monitoring network design to provide initial detection of groundwater detection.* Water Resources Research. V.30. No.9.

Rouhani, S. 1985. *Variance reduction analysis.* Water Resources Research. V.21. No.6.

Rouhani, S., Hall, T.J. 1988. *Geostatistical schemes for groundwater sampling.* Journal of Hydrology. 103.

Storck, P., Eheart, J.W., Valocchi, A.J. 1997. *A method for the optimal location of monitoring wells for detection of groundwater contamination in three-dimensional heterogeneous aquifers.* Water Resources Research. V.33. No.9

Virdee, T.S., Kottegoda, N.T. 1984. *A brief review of kriging and its application to optimal interpolation and observation well selection.* Hydrological Sciences Journal. V.29. No.4.

Zhang, Y., Pinder, G.F. 2005. *Least Cost Design of Groundwater Quality Monitoring Networks.* Water Resources Research. V.41.