Q19941207 0002

NETWORK DESIGN AND ACCURATE PREDICTION OF INFILTRATION AT THE LAS CRUCES TRENCH SITE

by
Gordon W. Wittmeyer (CNWRA), Budhi Sagar (CNWRA),
and Michael Cruse (Computation Research Professionals)

Abstract

Soil hydraulic data from the Las Cruces trench experiment are used to define material zone properties for a unsaturated flow model using several sampling configurations. Predictions of water content obtained from the flow model are compared to *in situ* water content measurements from a 310 day infiltration experiment. A single conceptual model of the flow system is selected and material properties are assigned to each zone using both purely random, and stratified-random sampling configurations. For each sampling configuration, 10 infiltration simulations are performed to obtain estimates of the mean prediction error. Two measures are used to compare the predicted and observed water contents: (i) the sum of squared errors of the predicted water contents, and (ii) the first and second moments of the infiltrated moisture plume. Both measures indicate that there is a general improvement in the accuracy of the model predictions as the number of samples is increased, although the marginal increase in accuracy decreases rapidly as more samples are taken. Based on the mean sum of squared errors measure, the purely random sampling configuration produces the most accurate predictions. However, comparison of the moisture plume moments lead to equivocal results. The generally superior accuracy of the purely random sampling configurations suggests that spatial correlation of the most important soil hydraulic parameters is rather weak within each material zone. Semivariogram analyses or the soil hydraulic data used in model construction indicates that both the saturated hydraulic conductivity and the pore size distribution parameter are spatially uncorrelated. Because the soils at the trench exhibit definite lateral structure based on color and texture, it is suggested that site characterization measurements should have a support size sufficient to conform to the observed stratigraphy.

Introduction

Regulatory agencies responsible for radiological safety require that site characterization efforts at a proposed high-level or low-level radioactive waste repository be sufficient to support a demonstration of its capability to safely isolate waste. However, both budget considerations and concerns regarding disturbance of the site require that drilling boreholes, taking samples, and conducting laboratory and field tests be minimized. Thus, constructors of disposal facilities are eventually confronted with the question of how much site characterization data is enough. Whenever possible, this question is posed quantitatively in terms of an error tolerance or an acceptable risk of failure to reject the site when the site is unable to meet the required performance criteria. The specified error tolerance or acceptable probability of failure is typically very small in the case where the site is investigated for disposal of hazardous chemical waste, low-level radioactive waste, or highly radioactive spent fuel.

Prior to site characterization, quantitative measures of the natural variability of the geology, hydrology and climatology of the site are unknown. However, reconnaissance level studies at the proposed site often provide data that may be used to produce a qualitative description of the sites's natural heterogeneity. Based on such preliminary data, decisions are made to drill boreholes and take observations at a few locations. With the additional quantitative data obtained from these boreholes, further locations for sampling are selected, and the process is repeated until it becomes apparent that

enough data is available to make the requisite site decisions. A great amount of subjective judgment is involved at each step of this iterative process. For that reason, Hudak and Loaiciga (1993) refer to this iterative process, in which formal decision theory is not used, as the qualitative approach. In the qualitative approach the sampling configuration is selected based on the investigator's: (i) conceptual model of the hydrostratigraphy, (ii) interpretation of how the spatial variability of its hydraulic properties influence the subsurface flow regime, and (iii) intuition as to what may constitute an optimal sampling network.

More quantitative methods, referred to as the simulation approach by Hudak and Loaiciga (1993), have also been developed. In the quantitative approach, one must first develop a premise about the worth of data in order to decide whether sufficient data have been gathered. Generally speaking, the marginal utility of data decreases as the size of database increases. According to James and Freeze (1993), hydrogeologic data have worth only if they can aid in making decisions between alternative courses of action. Early studies pertaining to design of sampling networks and data worth [Baecher, (1972); Maddock, (1973); and Gates and Kisiel, (1974)] ignored the spatial correlation that usually exists in natural geologic media at one or more scales. If the data are spatially correlated, gathering additional data from locations near existing measurement locations provides less information than gathering data that are located farther away. Of course an implicit assumption in this discussion is that the correlation structure is somehow known before the sampling plan is implemented which is not always the case. Recent studies [James and Freeze, (1993); McKinny and Loucks, (1992); Feinerman et al., (1985); Rouhani, (1985); Grosser and Goodman, (1985); Marin et al., (1989); and Graham and McLaughlin, (1989)] incorporate the effects of spatial correlation through the use of kriging. When estimation procedures which account for spatial correlation are coupled with formal decision theories that are based on well-defined loss/gain functions, formal mathematical procedures for determining optimal sampling networks can be developed [Bogardi et al., (1985); Massman et al., (1991).]

In this paper, a numerical investigation is presented which examines the worth of additional data and the design of sampling patterns using data from a field experiment. The data are from a carefully planned field experiment funded by the Nuclear Regulatory Commission (NRC) for the express purpose of validating mathematical models of flow and transport in partially saturated porous media [Wierenga et al. (1986), Wierenga et al. (1989), Nicholson et al. (1989), Nicholson et al. (1994).] The NRC's interest in such experiments is to develop a knowledge base so that technically sound decisions regarding the appropriateness of models used by constructors of disposal facilities for nuclear waste could be formed. Similarly, the NRC is also interested in determining adequacy of site characterization data which provides the motivation for this study. Because the original objective of this field experiment was to validate models, measurements of flow and transport properties were collected on a dense sampling network. The iterative characterization method described above was not followed in characterizing this experimental site; instead, the dense sampling network was implemented to ensure that no important site characteristic was overlooked. Such dense measurement networks are impractical at any real disposal site of any appreciable extent for two reasons: (i) the excessive cost associated with the dense network, and (ii) the possibility that an excessive number of boreholes and shafts may reduce the site's ability to isolate waste. Nevertheless, this experiment presents an excellent data set to investigate the issue of data sufficiency since the ultimate "true" characteristics of the site are known *a priori*. A variety of data sampling patterns are employed to determine how the number and location of sample points affect the accuracy of predictions obtained from mathematical models. Although generally applicable methods cannot be derived from this study, these analyses nevertheless provide useful insight into the issue of data sufficiency.

Quantitative Network Design Methods

Much of the literature on sampling network design is devoted to developing formal mathematical procedures for selecting additional measurement locations so as to minimize the error in estimating a spatially distributed variable over its entire domain using limited measurements. For example, given an array of existing boreholes at which the transmissivity of the aquifer has been estimated, a hydrogeologist may need to determine where additional boreholes should be drilled in order to minimize the error associated with estimating the transmissivity of the entire aquifer without exceeding a fixed budget. Typically, the transmissivity of the aquifer between borehole locations is estimated using statistically based interpolation procedures such as kriging, or, purely deterministic methods like inverse distance weighting. If kriging or a similar technique such as Gaussian conditional mean estimation is used, the estimation or kriging variance can be used as a surrogate measure for the error of estimation. One may then arrive at the optimum sampling plan by configuring the network of boreholes so that, for a given number of boreholes, the mean estimation variance is minimized. However, as noted by Christakos (1992), optimal sampling plans for a fixed sample size, even when they exist, are, in practice, quite difficult to determine because the optimization problems are computationally untenable.

Christakos (1992) addresses the problem of sampling or network design using concepts from the theory of spatial random fields. For the case where the primary concern is the mean areal performance of the sampling network, relatively straightforward measures of the average areal efficiency can be formulated. Christakos (1992) suggests minimizing the average estimation error

$$\sigma_A^2 = \frac{1}{K}\sum_{k=1}^{K} \sigma_Y^2(s_k), \tag{1}$$

where K is the number of points at which Y, the logarithm of transmissivity, is estimated, and the $\sigma_Y^2(s_k)$ are the point estimation errors at the additional sample locations s_k. For a linear estimator, like ordinary kriging, the point estimation errors are given by

$$\sigma_Y^2(s_k) = C_Y(0) - \sum_{j=1}^{m} \zeta_j C_Y(s_k - s_j) - \mu, \tag{2}$$

where m is the number of existing sample locations, μ is a Lagrange multiplier, $C_Y(h)$ is the spatial covariance of the log-transmissivity for two locations separated by h, and ζ_j are the kriging weights. If the covariance of the spatial random function is known *a priori*, an optimal sampling network may be constructed in the absence of observations by using Eqs. (1) and (2).

Since the number of possible sampling patterns is large, even for a moderate sample size or sampling density, one usually selects the pattern from a collection of basic sampling plans. The primary sampling patterns can be divided into three groups: (i) systematic or regular sampling patterns, (ii) stratified sampling patterns, and (iii) random sampling patterns (Christakos, 1992). With systematic sampling plans, like those shown in Figures 1a and 1b, the sampling points are located at the centroids of regular polygons used to tessellate the 2D domain. In stratified sampling patterns, such as shown in Figures 1c and 1d, the 2D domain is divided into regular or irregular, nonoverlapping regions, with one sampling location randomly selected from each region. For purely random sampling patterns, as shown in Figure 1e, the arrangement of sample locations exhibits no structure. Depositional processes, such as

alluviation, form geologic fabrics which exhibit different correlation structures at different scales. Large-scale changes in geologic structure, like the transition from a clay to a coarse sand layer, may be described deterministically (Gelhar et al., 1994). Small-scale variations in properties, such as porosity or hydraulic conductivity changes within a clay layer, may be described probabilistically (Gelhar et al., 1994). Gelhar et al. (1994) recommend using a clustered sampling pattern (shown in Figure 1f) to determine the soil property means and variances, as well as, identify the large-scale trends.

Real world sampling design problems are constrained both by fixed budgets which limit the number of boreholes drilled and the number of samples taken, and by the practical impossibility of considering all sample patterns when seeking the mathematically optimal solution. Moreover, to properly employ these formal mathematical procedures, one must already have obtained enough site characterization data to estimate the covariance structure of the parameter of interest. The network design problem is made even more difficult if the goal is to use the measurements obtained from the sampling network to predict the values of another, related variable. For example, one may wish to design a sampling network for transmissivity such that the mean prediction error for hydraulic head field is minimized. In this case, the head prediction error may be approximated by the estimated head variance derived either by numerically propagating the uncertainty in Y through the flow equation, or by analytically solving the small perturbation expansion of the flow equation.

Description of Experimental Observations

The experimental site is located northeast of Las Cruces, New Mexico on the New Mexico State University College Ranch near the north end of the Doña Ana Mountains. The climate at the experiment site is semi-arid. Class A pan evaporation is 239 cm per year, while average annual precipitation is only 23 cm, over half of which falls during the summer monsoon season. As shown in Figure 2, a trench 26.4 m long, 4.8 m wide, and 6 m deep was excavated to facilitate instrumentation, site characterization, and visual inspection of the infiltration experiments. Plot 2, located adjacent to the north face of the trench, was irrigated by an array of 80 drip lines aligned parallel to the trench face and covering a surface strip 12 m long and 1.2 m wide. For the Plot 2b experiment, which is the focus of this study, water was applied to the strip at a rate of 1.82 cm/day for 70 days. Chromium, boron, and pentafluorobenzoic acid were applied to the strip during the first 15 days, and tritium, bromide, and 2,6-difluorobenzoic acid were applied during days 29 through 44. Water content was monitored by neutron probes through the network of access tubes shown in Figure 2. Matric potential was monitored by tensiometers installed in the face of the trench. Solute samples were obtained using suction lysimeters installed in the trench face and destructive sampling of small diameter soil cores at selected locations to the north of the trench.

Approximately, 600 soil cores and 600 disturbed soil samples were taken during the construction of the trench. As described in Wierenga et al. (1989), 50 samples of each type, spaced at 50 cm intervals, were taken along nine horizontal transects corresponding to the nine soil layers identified in the north face of the trench. In addition, samples were taken along three vertical transects at a spacing of 13 cm. Water retention data were obtained from the soil cores by determining water content at suctions of 10, 20, 40, 80, 120, 200, and 300 cm H_2O. The high suction end for each of the water retention curves was determined from crushed and sieved portions of the disturbed soil samples which were placed in pressure plate extractors and subjected to pressures of 1, 5, and 15 bar.

The van Genuchten model (van Genuchten, 1980) was used to describe the water retention or moisture characteristic curve. This model is given by

$$\theta = \theta_r + \frac{(\theta_s - \theta_r)}{[1 + (\alpha h)^n]^m}, \tag{3}$$

where θ_r is the residual water content, θ_s is the saturated water content, h is suction head, and α, m, and n are unknown fitting parameters. It is further assumed that $m = 1 - 1/n$, which facilitates the derivation of a closed-form expression for the unsaturated hydraulic conductivity as a function of water content or pressure head based on Mualem's (1976) predictive model for the unsaturated hydraulic conductivity. The van Genuchten-Mualem predictive model for unsaturated hydraulic conductivity is

$$K(h) = K_s \frac{[1 - (\alpha h)^{n-1}[1 + (\alpha h)^n]^{-m}]^2}{[1 + (\alpha h)^n]^{m/2}}, \tag{4}$$

where K_s is the saturated hydraulic conductivity.

In a previous modeling study of the Plot 2a experiment conducted by Wittmeyer and Sagar (1993), difficulty was encountered in determining a suitable set of initial pressures, since many of the water content values measured by neutron probes were less than the estimated residual water contents (Wierenga at al., 1989). It was assumed that the soil hydraulic parameters, which were determined from cores taken in a vertical cross-section parallel and immediately adjacent to the north face of the trench, could be directly projected normal to the north trench face and made coincident with the planar arrays of neutron probe access tubes at 2, 6, and 10 m from the face (Figure 2). Any node, i, in the computational mesh that resides in the intersection of the rectangular areas of influence for soil sample location, l, and the neutron probe location, j, was assigned the initial pressure head

$$h_i = \frac{1}{\alpha_l}\left[\left(\frac{\theta_j - \theta_{r_l}}{\theta_{s_l} - \theta_{r_l}}\right)^{\frac{n_l}{1-n_l}} - 1\right]^{n_l}. \tag{5}$$

When $\theta_j < \theta_{rl}$, Eq. (5) implies negative saturation, and the pressure head estimated by Eq. (5) is meaningless. To resolve this problem the pressure head at any node for which this condition occurred was set to 15 bar, the pressure at which the residual water contents were defined (Wierenga et al., 1989). Consequently, the estimated initial water content in many portions of the model domain exceeded the measured water content.

To eliminate this difficulty θ_r, α and n were re-estimated for each of the measured water retention curves. An estimation procedure suggested by Mualem (1976) whereby θ_r, α and n are determined subject to the constraint that the computed water retention curve passes through the point corresponding to the highest suction and lowest water content was implemented using constrained nonlinear least squares. Approximately 100 soil samples for whose retention curves the minimum water content did not coincide with the highest applied suction were removed from consideration. Unfortunately there remained some locations in the model domain where the re-estimated residual water content still exceeded the measured initial water content. The problem was resolved by including an additional point to each set of water retention data consisting of the lowest water content measured by neutron probe and the highest suction measured by thermocouple psychrometer, followed by re-estimation of θ_r, α and n.

Network Design for Predicting Infiltration

The goal of the network design problem addressed in this paper is to determine the number and configuration of sample locations where the unsaturated soil hydraulic properties should be measured in order to minimize the error in predicting the propagation of a moisture plume. This problem, while similar to the network design problem of choosing the measurement locations for Y that minimize the prediction error in hydraulic head, is greatly complicated by the nonlinearity of the unsaturated flow equation, and the large number of parameters needed to adequately describe the unsaturated hydraulic conductivity and soil moisture characteristic functions. At the time this study was performed, all site characterization activities and infiltration experiments had already been completed at the Las Cruces field experiment site. Our strategy for this study was to use a limited amount of the available qualitative and quantitative data to construct the conceptual flow model, select soil hydraulic parameters from the existing database, and compare the predicted water contents along the y=2m transect to the measured water content values.

Soil hydraulic data may be used both to design the basic conceptual model and to assign parameter values to the numerical model's zones. However, since the object of this study is to estimate the impact of different sampling schemes on the accuracy of model predictions, we chose a single conceptual model and varied the parameters of that model based on the sampling scheme. For example, a conceptual model composed of two parameter zones, within each of which a single sample was collected to obtain parameter values, cannot be compared to a model composed of a single parameter zone, in which two samples were collected to assign effective, average properties to the zone. Clearly two soil samples were used in each case, but because the conceptual models employed are fundamentally different (one zone versus two) it would be difficult to draw conclusions about the superiority of one sampling pattern over the other. Figure 3 depicts the factors considered in conducting this study. In Figure 3, the x_1 axis represents the number of parameter zones in the conceptual model, the x_2 axis represents the number of sampling strata per model zone, and the x_3 axis represents the number of randomly located samples collected per stratum. For this study a single conceptual model was used which consisted of 9 horizontal layers or parameter zones coincident with the nine soil horizons identified on the north face of the trench. By selecting a single conceptual model the study could be focused on the experimental factors most relevant to site characterization: (i) how many samples should be taken (axis x_3), and (ii) what sampling pattern should be used (axis x_2).

Figure 4 shows the model domain, the definition of the nine model parameter (or material) zones, and the locations of the actual soil samples taken in this portion of the trench face. The model domain extends 10 m laterally to the east from the northwest corner of the trench, and 7 m vertically from the strip source to a location one meter below the bottom of the trench. The computational mesh used for the numerical model, as shown in Figure 5, consists of 81 nodes in the horizontal and 97 nodes in the vertical direction. The mesh is increasingly refined towards the center and top of the domain to accommodate the large pressure head gradients expected to occur at early time directly below the strip source. Boundary conditions were everywhere defined to be no-flow, except at the bottom of the model and at the 1.2 m horizontal portion at the top where the strip infiltration source is located. To simulate gravity drainage, a negative pressure head gradient of unit magnitude was specified at the bottom of the domain. At the strip source, the boundary condition was a prescribed flux of 1.82 cm/day during the first 70 days of simulation and zero thereafter. The redistribution of the moisture plume was monitored for 240 days after infiltration ceased.

All simulations described in this paper were performed using the computer code PORFLOW, Version 1.2 (Sagar and Runchal, 1990), which is similar to Version 1.0 except that the newer version employs the mass-conserving mixed-formulation for the storage term, and allows the specification of a prescribed gradient boundary condition. PORFLOW uses the method of finite volumes to discretize the unsaturated flow equation in space and fully implicit backward differencing for discretization in time. Pre-conditioned conjugate gradient method is used to solve the set of discrete equations resolving the nonlinearities by employing Piccard iterations.

Two basic sampling patterns were employed in this study: (i) random sampling, and (ii) stratified sampling. For the purely random sampling plan, one or more soil samples were randomly selected from the existing sample locations in each of the nine model zones. When a single random sample was taken, the soil hydraulic properties corresponding to that sample were assigned to the entire zone. When two or more samples were taken, the arithmetic averages for θ_s, θ_r and n, and the geometric averages for K_s and α, were assigned to the entire model zone. Here we have assumed that θ_s, θ_r and n are normally distributed, and K_s and α are lognormally distributed. For the stratified sampling plan, each of the nine model zones was divided into two or more sampling strata as shown in Figure 6. One or more soil samples were then randomly selected from each stratum, and all samples from the model zone were appropriately averaged to determine the effective model parameters.

Truly random sampling could not be performed since we were constrained to choose from the existing sampling network described earlier (Figure 4). Spatial random sampling was approximated by: (i) mapping each sampling stratum to a unit square, (ii) generating a pair of uniform $[0,1]$ random numbers corresponding to the Cartesian coordinates of a point in the unit square, and (iii) selecting the sample location within the stratum closest to the randomly generated point. Because the horizontal sampling transects were intentionally aligned with the nine soil horizons, sample location randomness was effectively constrained to the horizontal axis, except where a stratum intersected one of the vertical transects.

Two metrics were used to determine which sampling plan produced the most accurate estimates of the propagation of the moisture plume. The first metric was based on the direct point-to-point comparison of the measured and predicted water contents at the neutron probe measurement locations in the y=2m transect collected at 55 measurement times, where $t \in [0,310]$. This measure, called the cumulative sum of squared errors, is given by

$$\mathrm{SSE} = \sum_{j=1}^{N_T} \sum_{i=1}^{N_{OBS}} \left(\theta_{ij}^{*} - \hat{\theta}_{ij}\right)^2, \tag{6}$$

where N_{OBS} is the number of neutron probe measurement locations, N_T is the number of measurement times, θ^* are the measured water contents, and $\hat{\theta}$ are the predicted water contents. The second metric was based on the first and second moments of the water content difference plumes. The moment of order l in the x-direction and k in the z-direction for the observed water content difference plume is

$$M^{*}_{lk}(t) = \iint_{Z\,X} \left[\theta^{*}(x,z,t_0) - \theta^{*}(x,z,t)\right] x^{l} z^{k} dx dz, \tag{7}$$

where $\theta^{*}(x,z,t_0)$ and $\theta^{*}(x,z,t_0)$ are the measured water contents at point (x,z) and times t_0 and t, respectively. Similarly, the moment of order l in the x-direction and k in the z-direction for the predicted water content difference plume is

$$\hat{M}_{lk}(t) = \iint_{Z\,X} \left[\hat{\theta}(x,z,t_0) - \hat{\theta}(x,z,t)\right] x^{l} z^{k} dx dz. \tag{8}$$

The x and z coordinates of the plume centroid for either the measured or predicted water content differences are generically given by

$$X_c(t) = M_{10}(t) \,/\, M_{00}(t), \tag{9}$$

$$Z_c(t) = M_{01}(t) \,/\, M_{00}(t). \tag{10}$$

The centroid is used to represent the mean position of the plume, while the second moment about the centroid is used to represent the extent of the spread of the plume about its center. Equations for the second moments are generically given by

$$S_{xx}(t) = M_{20}(t) \,/\, M_{00}(t) - X_c^2(t), \tag{11}$$

$$S_{zz}(t) = M_{02}(t) \,/\, M_{00}(t) - Z_c^2(t). \tag{12}$$

The square roots of the second moments have units of length and are interpreted as measures of plume spread.

Network design configurations tested in this study are shown in Table 1. Because each model simulation took between 1.5 and 5 CPU hours on a Sun SPARCstation 10/40, the total number of sampling patterns that could be investigated was restricted to a small number. In addition, the maximum number of sample strata per zone was limited to six by the horizontal spacing of the existing sample locations. As shown in Table 1, five stratified sampling patterns in which one measurement was randomly selected from each stratum (9-2-1, 9-3-1, 9-4-1, 9-5-1, 9-6-1), three purely random sampling patterns in which one, two and four measurements were selected (9-1-1, 9-1-2, 9-1-4), and two sampling patterns in which each model zone was divided into two strata with two and four measurements randomly selected from each stratum (9-2-2, 9-2-4) were tested.

Because the actual sample measurement locations were selected randomly from each sample stratum, each sampling pattern or configuration has many possible realizations. As an example, consider the sampling pattern 9-3-1, which is shown for the second soil layer or material property zone in Figure 6. For layer 2 alone there are 180 possible realizations, and when all 9 material zones are considered, there are over 4.27×10^{20} possible realizations for configuration 9-3-1. It is clear that one cannot select the optimum sampling pattern from Table 1 using only a single realization to evaluate the

measures described by Eqs. (6) through (12). In order to account for the effects of randomness, ten realizations were taken for each network design configuration shown in Table 1. Thus for each of the sampling patterns shown in Table 1, ten values of the measures given by Eqs. (6) through (12) were obtained. The median, mean, and standard deviation of these measures over the space of realizations were then used to compare the various sampling configurations.

In order to assess whether ten realizations were sufficient to obtain sample statistics that were sufficiently accurate for selecting one sampling configuration over another, the standard deviation of the mean of the SSE measure described by Eq. (6) was analyzed for sampling configuration 9-3-1. For this particular configuration, 25 realizations of the sampling scheme were obtained, the corresponding 25 flow simulations conducted, and 25 values of the SSE measure obtained. A statistical experiment was then performed with these 25 values of the SSE. Two thousand samples of size 2 ($N_s=2$) were drawn from the population of 25 without replacement. (Imagine that the 25 balls are labelled with the SSE estimates and placed in an urn. Two balls are randomly drawn from the urn one at a time without replacing the first ball, the SSE values of each ball then recorded and the balls returned to the urn. This process is performed 2000 times.) The mean of each sample of two SSEs (2000 in all) was calculated. The mean and standard deviation of the 2000 mean SSEs were then calculated from the values so obtained. This statistical experiment was repeated for sample sizes $N_s=3,\ldots,24$, and for each sample size 2000 estimates of the SSE were obtained. The results of this statistical experiment are shown in Figure 7 where the mean SSE (dashed line) and its plus and minus one standard deviation limits (error bars) are plotted as a function of the number of realizations (2 to 24). The decrease of the standard deviation of the mean SSE with increasing number of realizations is obvious (the standard deviation decreased from 0.55 for $N_s=10$ to 0.28 for $N_s=20$). While, the uncertainty (or standard deviation) of the mean SSE for $N_s=10$ may not be small enough to always distinguish between sampling configurations, the added computational burden makes performing a greater number of simulations impractical.

Results of Study

Plots of the mean SSE versus time for sample patterns 9-1-1, 9-2-1, 9-3-1, 9-4-1, 9-5-1, and 9-6-1 are shown in Figure 8. During the wetting portion of the infiltration experiment, which extends from day 0 to day 70, sampling configuration 9-1-1 produces a somewhat smaller mean SSE than do the other sampling configurations. However, once the moisture redistribution period begins after day 70, sample patterns 9-4-1, 9-5-1, and 9-6-1 produce the smallest mean SSE. Figure 9 shows the mean SSE at day 310 for these six stratified sampling patterns. It is clear that dividing each model zone into two strata leads to a significant reduction in the mean SSE. Further increases in the number of sampling strata per zone lead to a general decrease in the mean SSE, although it is also apparent that the percentage decrease is not as great as that from 9-1-1 to 9-2-1. In Figure 10, the median SSE at day 310 is plotted for sample patterns 9-1-1, 9-2-1, 9-3-1, 9-4-1, 9-5-1, and 9-6-1. Because the median is less affected by outliers than the mean, the graph in Figure 10 is somewhat smoother and better depicts the general decrease in the marginal value of adding another stratum and sample in each model zone. Note that in Figure 10, sampling configuration 9-4-1 actually produces a lower median SSE estimate than do 9-5-1 and 9-6-1, although the standard deviation of the 10 runs used to define each of the three points exceeds the maximum difference between the medians (Table 2). Note that the decrease in the median SSE using 2 strata instead of 1 is 14 percent, while the median SSE decreases by 11 percent moving from 2 strata to 4 strata.

In Figure 11, the median SSE at day 310 is plotted for sample patterns 9-1-1, 9-1-2, and 9-1-4. As stated in Table 1, these sampling configurations correspond to using a single stratum per model zone,

with one, two, and four random samples selected per stratum. Figure 11 clearly shows that there is a marked decrease in the median SSE as the number of random samples increases from one to four. The median SSE is reduced 21 percent by increasing from 1 to 2 samples, but only 10 percent by increasing from 2 to 4 samples. The decrease in the marginal value of additional samples is similar to that observed for the stratified sampling patterns, although the rate of decrease in the marginal value is greater.

In Figure 12, the median SSE at day 310 is plotted for sample patterns 9-2-1, 9-2-2, and 9-2-4. These sample patterns correspond to delineating two sample strata per model zone, and taking one, two, and four random samples in each of the two strata. Figure 12 is similar to Figure 11 in that there is marked decrease in the median as the number of random samples increases from one to four. The median SSE is reduced 28 percent by increasing from 1 to 2 samples per stratum, but only by 7.6 percent by increasing from 2 to 4 samples per stratum. Once again, the now familiar decrease in the marginal value of additional samples is observed, however, in this case the decrease is greater.

Analyses of the means of the first [Eqs. (8) and (9)] and second moments [Eqs. (10) and (11)] lead to somewhat different, and somewhat equivocal, sets of observations about the marginal value of additional measurements. Figures 13a, 13b, and 13c show the evolution of the vertical coordinate of the centroid of the water content difference plume, the horizontal spread of the plume, and the vertical spread of the plume, respectively, for sample patterns 9-1-1, 9-2-1, 9-3-1, 9-4-1, 9-5-1, and 9-6-1. From Figure 13a it may be concluded that increasing the number of sampling strata leads to a general improvement in the ability of the model to predict the depth of penetration of the water content difference plume. It is also apparent from Figure 13a that improvement in predicting the Z-coordinate of the centroid of the plume is far greater during the moisture redistribution period. Figure 13b indicates that using fewer strata improves the mean accuracy of predicting the horizontal spreading of the plume during the infiltration phase, while using more strata increases the mean prediction accuracy during the redistribution phase. Conversely, from Figure 13c, it appears that using fewer strata decreases the mean accuracy of predicting the vertical spread of the water content difference plume during infiltration, while using more strata decreases the mean prediction accuracy during redistribution.

While it is disconcerting to discover that in some cases the moment-based measure of model performance suggests that using more samples may lead to poorer predictions, the root cause of these inconsistent results is most likely the selection of an incorrect conceptual model. In previous work, the authors noted that their models were unable to replicate a distinct vertical bifurcation in the water content distribution that was observed at a depth of 3 m (Wittmeyer and Sagar, 1993). Spatial analyses of the grain-size distribution data indicated that there was a thin, but distinct, coarse sandy layer at 3 m depth. Soil samples collected from this horizon for determination of the soil hydraulic properties clearly failed to capture the capability of this sandy layer to transmit relatively large water fluxes without a measurable increase in the water content (Wittmeyer and Sagar, 1993). In this study we were again limited to using the original soil moisture characteristic data, and thus the model should be expected to produce biased predictions. For this reason, one must not conclude from the moment comparisons that using additional samples necessarily reduces the accuracy of model predictions.

If the soil hydraulic model parameters are homogeneous within a model zone, one soil sample per zone is sufficient for characterization. If the correlation length for a particular parameter is less than the largest dimension of the model zone, stratified sampling should be used to obtain an effective parameter value. Finally, if the model parameters exhibit no spatial correlation, a number of purely random samples should be taken within a model zone to obtain the effective parameter value used for that zone. There are two sample configurations for which two samples were selected from each zone (9-2-1,

9-1-2), and three sample patterns for which four samples were selected from each material zone (9-4-1, 9-1-4, 9-2-2). The median SSE for pattern 9-2-1 is 18.17, while that for sample pattern 9-1-2 is only 16.70. Similarly, the median SSE for pattern 9-4-1 is 16.25, while that for sample pattern 9-1-4 is 14.95. In both instances, the purely random pattern produces more accurate model predictions than does the stratified sampling pattern, which suggests that the soil hydraulic parameters to which the flow model is most sensitive are spatially uncorrelated. We may reject the possibility that each zone is homogeneous, both on the basis of visual inspection of the trench face, and because using two random samples per stratum instead of one produces more accurate predictions in terms of the SSE measure [median SSE for: (i) 9-1-1 equals 21.25 and 9-1-2 equals 16.70, and (ii) 9-2-1 equals 18.17 and 9-2-2 equals 15.36].

Inasmuch as the definitions of the nine soil horizons were largely based on the observed lateral structure, and the general continuity of soil color and texture, the apparent superiority of purely random sampling over stratified sampling within each soil horizon is difficult to explain. If, however, the color and textural continuity observed within each soil layer does not imply spatial correlation of the soil hydraulic parameters obtained from the soil cores, this result may be explained. To test this hypothesis, horizontal sample semivariograms for $\ln(\alpha)$, n, θ_r, θ_s, and $\ln(K_s)$ were constructed, and are shown in Figures 14a-e, respectively. Each sample semivariogram was computed from the 160 soil samples in the model domain using a constant lag of 0.55m, which slightly exceeds the 0.5m horizontal sampling interval, and a window of 60 degrees. Although semivariogram models were not fitted to any of these sample semivariograms, visual inspection indicates that those for $\ln(\alpha)$, θ_r, and θ_s could be represented by spherical models with approximate ranges of 6m, 7m, and 7m, respectively. The sample semivariogram for n suggests that this parameter is mildly spatially correlated, although a large portion of its sill may be due to a nugget effect. However, the horizontal sample semivariogram for $\ln(K_s)$ suggests that this parameter is not spatially correlated, and would be best represented by a pure nugget model. If the water content predictions made with the flow model are most sensitive to the values of saturated hydraulic conductivity, the lack of spatial correlation implied by its sample semivariogram supports the observation that a purely random sampling pattern is superior, on average, to a stratified sampling pattern. If instead the model is most sensitive to α, θ_r, and θ_s, the observed superiority of purely random sampling to stratified sampling cannot be explained at this time.

Conclusions

The efficiency of several sampling configurations was studied by utilizing data from a detailed field experiment in Las Cruces, New Mexico, on flow and transport in unsaturated soils. The sampling configurations studied were of two general types: (i) purely random sampling in which the location of the sampling points is spatially random, and (ii) stratified sampling in which the domain of interest is divided into strata and sampling points are located randomly in each strata. Our interest centered on investigating sampling configurations that will provide the "best" prediction of the transient propagation of the moisture plume. Thus, we did not adopt the traditional approach where the goodness of the sampling network is based on minimizing the error in obtaining an effective value of the property being measured. The large amount of soil characterization data at the experimental site provided us with an opportunity to test the sampling configurations with actual measurements of the moisture plume.

The results indicate that the marginal value of additional samples decreases rapidly as the sample size increases. Thus the greatest improvement in model predictions was obtained when the sample size was increased from one sample per material zone to two samples per material zone. Thereafter, the rate

of improvement decreased significantly. Of the two measures of "goodness" employed (i) the median cumulative squared error, and (ii) the moment-based measures of the centroid and spread of the moisture plume, the first measure appeared to provide more consistent results.

Note that a common conceptual model was used in evaluating various sampling schemes. While this is a reasonable strategy for this particular study, it is to be emphasized that any new data can alter not only the parameter values but also the conceptual model. For example, if the new sample indicates the existence (or absence) of a soil structure that was not indicated by the previous data, then a change in the conceptual model must be made. With the new conceptual model, the interative process of determining the worth of further data will start anew.

Overall, for this particular site and experiment, it was observed that random, non-stratified, sampling provided more accurate simulation results. This appears to imply that spatial correlation of the most important soil hydraulic parameters within each material zone is rather weak. Through semivariogram analysis, it was found that $\ln(K_s)$ and van Genuchten's n were indeed laterally uncorrelated within the model area. However, θ_s, θ_r and α were found to be correlated. These results appear to indicate that accurate prediction of the propagation of the moisture plume is most sensitive to $\ln(K_s)$ and van Genuchten's n. However, contrary to the sample semivariogram shown in Fig 14e, a sample semivariogram constructed for *in situ* measurements of saturated hydraulic conductivity obtained with Guelph permeameters shows a very strong lateral, spatial correlation. Inasmuch as the support size of the permeameter measurements is greater than that of the soil cores from which the saturated hydraulic conductivity data used in this study were obtained, one should expect the statistical structure of the fields to be quite different. This observation underscores the importance of obtaining measurements of hydraulic parameters on a scale that conforms to the observed stratigraphic structure of the site geology.

It appears that the sufficiency of data can only be determined in terms of the modeling results and the measures selected for determining worth of data. In addition, whether additional data will modify a decision can only be judged after the additional data is collected. All existing procedures for judging the worth of data before its collection are based on assumptions that imply statistical extrapolation of prior data. The numerical experiment described above indicates that an iterative coupling between data collection and modeling will be advantageous for deciding data sufficiency.

Acknowledgments

The authors wish to thank Amvrossios C. Bagtzoglou, Robert G. Baca, and Wesley C. Patrick who provided technical and programmatic reviews of this paper. Their comments greatly improved the clarity and overall quality of this paper. This report was prepared to document work performed by the Center for Nuclear Waste Regulatory Analyses (CNWRA) for the Nuclear Regulatory Commission under Contract No. NRC-02-93-005. The activities reported here were performed on behalf of the NRC Office of Nuclear Regulatory Research, Division of Regulatory Applications. The report is an independent product of the CNWRA and does not necessarily reflect the views or regulatory position of the NRC.

References

Baecher, G. B. 1972. Site Exploration: A Probabilistic Approach, Ph.D. Thesis, Mass. Inst. of Technol., Cambridge.

Bogardi, I., A. Bardossy, and L. Duckstein. 1985. Multicriterion network design using geostatistics. Water Resour. Res., v. 21, pp. 199-208.

Christakos, G. 1992. Random Field Models in Earth Sciences. Academic Press, San Diego, California.

Feinerman, E., E. Bresler, and G. Dagan. 1985. Optimization of a spatially variable resource: an illustration for irrigated crops. Water Resour. Res. v.21, pp. 793-800.

Freeze, R. A., B. R. James, J. Massman, T. Sperling, and J. L. Smith. 1992. Hydrogeological decision analysis, 4, the concept of data worth and its use in development of site investigation strategies. Ground Water, v. 30 pp. 574-588.

Gates, J. S., and C. C. Kisiel. 1974. Worth of additional data to a digital computer model of a groundwater basin. Water Resour. Res., v. 10, pp. 1031-1038.

Gelhar, L. W., M. A. Celia, and D. McLaughlin. 1994. Modeling Field Scale Unsaturated Flow and Transport Processes. NUREG/CR-5965. Nuclear Regulatory Commission. Washington, DC.

Graham, W., and D. McLaughlin. 1989. Stochastic analysis of nonstationary subsurface solute transport, 1. Unconditional moments. Water Resour. Res., v. 25, pp. 215-232.

Grosser, P. W., and A. S. Goodman. 1985. Determination of groundwater sampling frequencies through Bayesian decision theory, Civ. Eng. Syst., v. 2, pp. 186-194.

Hudak, P. F., and H. A. Loaiciga. 1993. An optimization method for monitoring network design in multilayered groundwater flow systems. Water Resour. Res. v. 29, pp. 2835-2845.

James, B. R., and R. A. Freeze. 1993. The worth of data in predicting aquitard continuity in hydrogeological design. Water Resour. Res. v. 29, pp. 2049-2065.

McKinney, D. C., and D. P. Loucks. 1992. Network design for predicting groundwater contamination. Water Resour. Res., v. 28, pp. 133-147.

Maddock, T. III. 1973. Management model as a tool for studying the worth of data, Water Resour. Res., v. 9, pp. 270-280.

Marin, C. M., M. A. Medina, Jr., and J. B. Butcher. 1989. Monte Carlo analysis and Bayesian decision theory for assessing the effects of waste sites on groundwater, I, Theory. J. Contam. Hydrol, v. 5, pp. 1-13.

Massman, J., R. A. Freeze, L. Smith, T. Sperling, and B. James. 1991. Hydrogeological decision analysis, 2. Applications to groundwater contamination. Ground Water, v. 29, pp. 536-548.

Mualem, Y. 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. Water Resour. Res. 12(3): 513-522.

Nicholson, T. J. 1990. Recent accomplishments in the INTRAVAL project - a status report on validation efforts, proc. Geoval 90 symposium on validation of geosphere flow and transport models, Stockholm, Sweden, May-14-17.

Nicholson, T. J., R. Hills, M. Rockhold, G. Wittmeyer, T. C. Rasmussen, and A. Guzman. 1994. Conclusions from working group 1: Partially-saturated porous and fractured media test cases Las Cruces trench and Apache Leap tuff studies. Presented and distributed at the OECD/NEA symposium on the Validation of Geosphere FloW and Transport Models: GEOVAL-94.

Sagar, B., and A. K. Runchal. 1990. PORFLO-3: A Mathematical Model for Fluid Flow, Heat and Mass Transport in Variably Saturated Geologic Media - Theory and Numerical Methods, Version 1.0. WHC-EP-0042. Richland, Washington: Westinghouse Hanford Company.

Rouhani, S. 1985. Variance reduction analysis. Water Resour. Res., v. 21, pp. 837-846.

van Genuchten, M. Th. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 44: 892-898.

Wierenga, P. J., L. W. Gelhar, C. S. Simmons, G. W. Gee, and T. J. Nicholson. 1986. Validation of Stochastic Flow and Transport Models for Unsaturated Soils: A Comprehensive Field Study, NUREG/CR-4622, Nuclear Regulatory Commission, Washington, DC.

Wierenga, P. J., D. Hudson, J. Vinson, M. Nash and A. Toorman. 1989. Soil Physical Properties at the Las Cruces Trench Site, Soil and Water Science Report No. 89-002, University of Arizona, Tucson: AZ.

Wittmeyer, G., and B. Sagar. 1993. Model complexity and model validity: application to the Las Cruces trench experiment, INTRAVAL Test Case 10. Proceedings Site Characterization and Model Validation, Focus '93. American Nuclear Society. La Grange Park, IL. 209-216.

Table 1. Sampling configuration designation and description.

Sampling Configuration Designation	Number of Model Zones	Number of Sample Strata per Model Zone	Number of Samples per Stratum
9-1-1	9	1	1
9-2-1	9	2	1
9-3-1	9	3	1
9-4-1	9	4	1
9-5-1	9	5	1
9-6-1	9	6	1
9-1-2	9	1	2
9-1-4	9	1	4
9-2-2	9	2	2
9-2-4	9	2	4

Table 2. Mean SSE, median SSE, and standard deviation of SSE over 10 realizations conducted for each sampling configuration.

Sampling Configuration Designation	Mean Cumulative Squared Error	Standard Deviation of Cumulative Squared Error	Median Cumulative Squared Error
9-1-1	22.42	5.74	21.25
9-2-1	17.13	2.48	18.17
9-3-1	17.85	2.29	18.03
9-4-1	16.15	1.15	16.25
9-5-1	16.70	1.90	16.45
9-6-1	16.08	1.25	16.48
9-1-2	16.90	2.91	16.70
9-1-4	15.66	2.30	14.95
9-2-2	15.58	1.63	15.36
9-4-2	14.09	0.865	14.19

List of Figures

Figure 1a-1f. Sampling patterns: (a) Systematic sampling pattern on a regular, rectangular grid, (b) Systematic sampling pattern on a regular, hexagonal grid, (c) Stratified, random sampling pattern on a regular, rectangular grid, (d) Stratified, random sampling pattern on a regular, hexagonal grid, (e) Purely random sampling pattern, (f) Random, clustered sampling pattern [from Christakos, (1992)].

Figure 2. Plan view of the Las Cruces trench site.

Figure 3. Design factors for conceptual model selection and data sampling. Rectangular zone with dashed perimeter delineates the factor plane in (x_2,x_3) considered here.

Figure 4. Location of soil samples given by solid triangles. Horizontal lines delineate the nine material property zones or soil layers.

Figure 5. Computational mesh.

Figure 6. Selection of sample locations for stratified-random sampling. Horizontal strip represents one material zone or layer used in model. Vertical dashed lines divide material zone into three sampling strata. Stars represent randomly selected sample locations in each stratum. Circled triangle represents the actual soil sample location closest to the randomly selected point.

Figure 7. Mean cumulative sum of squared error as a function of sample size given by the dashed horizontal line. Vertical bars represent one standard deviation limits.

Figure 8. Mean cumulative sum of squared errors as a function of time for the six stratified sampling configurations.

Figure 9. Mean cumulative sum of squared errors at day 310 for the six stratified sampling configurations.

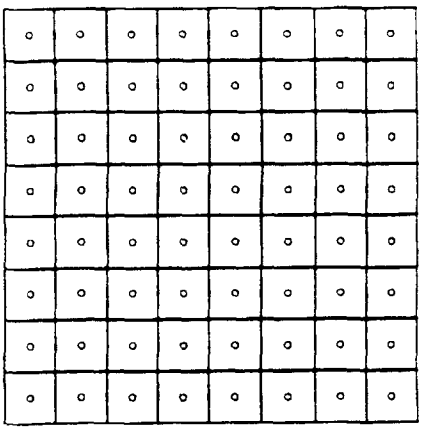
Figure 10. Median cumulative sum of squared errors at day 310 for the six stratified sampling configurations.

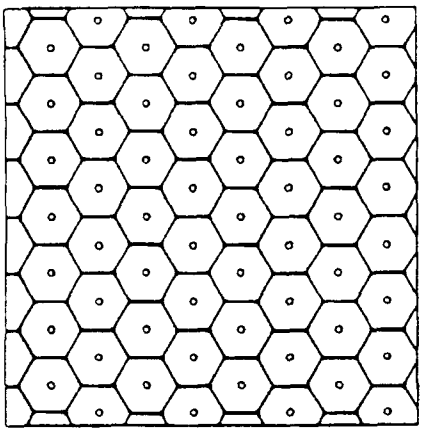
Figure 11. Median cumulative sum of squared errors at day 310 for sample patterns 9-1-1, 9-1-2, and 9-1-4.

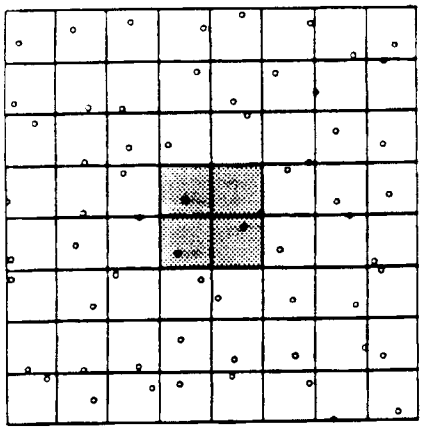
Figure 12. Median cumulative sum of squared errors at day 310 for sample patterns 9-2-1, 9-2-2, and 9-2-4.

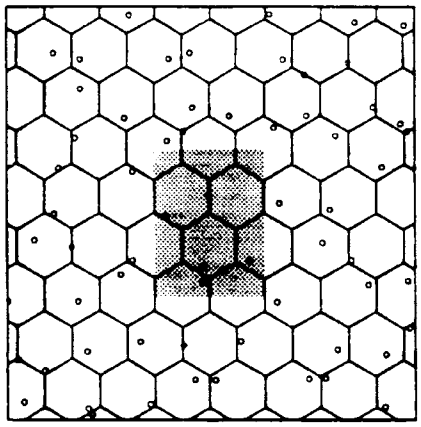
Figure 13a-13c. (a) Mean vertical position of water content difference plume centroid as a function of time for six stratified sampling patterns, (b) Mean horizontal spread of water content difference plume as a function of time for the six stratified sampling patterns, (c) Mean vertical spread of water content difference plume as a function of time for six stratified sampling patterns.

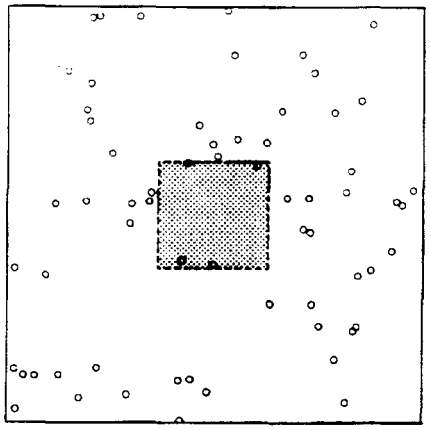
Figure 14a-14b. Sample semivariograms for : (a) $\ln(\alpha)$, (b) n, (c) θ_r, (d) θ_s, (e) $\ln(K_s)$.

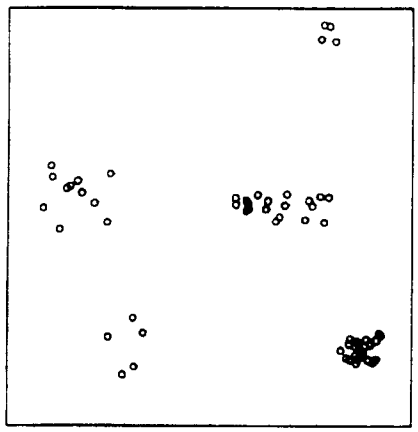












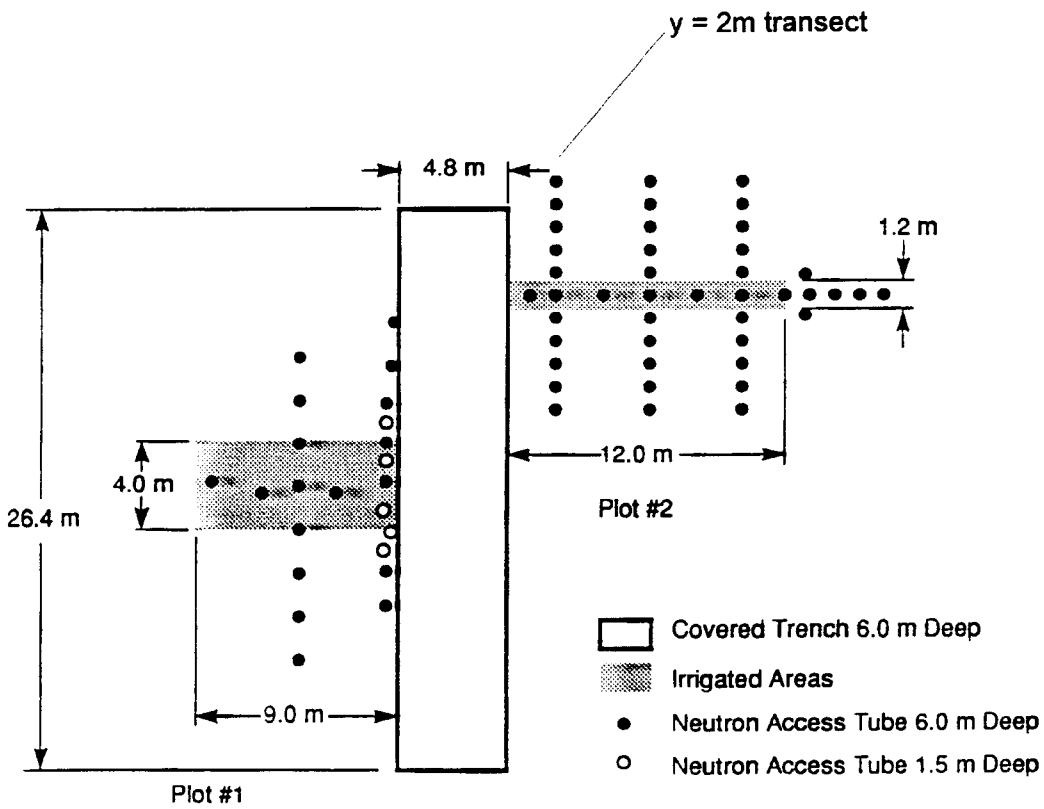
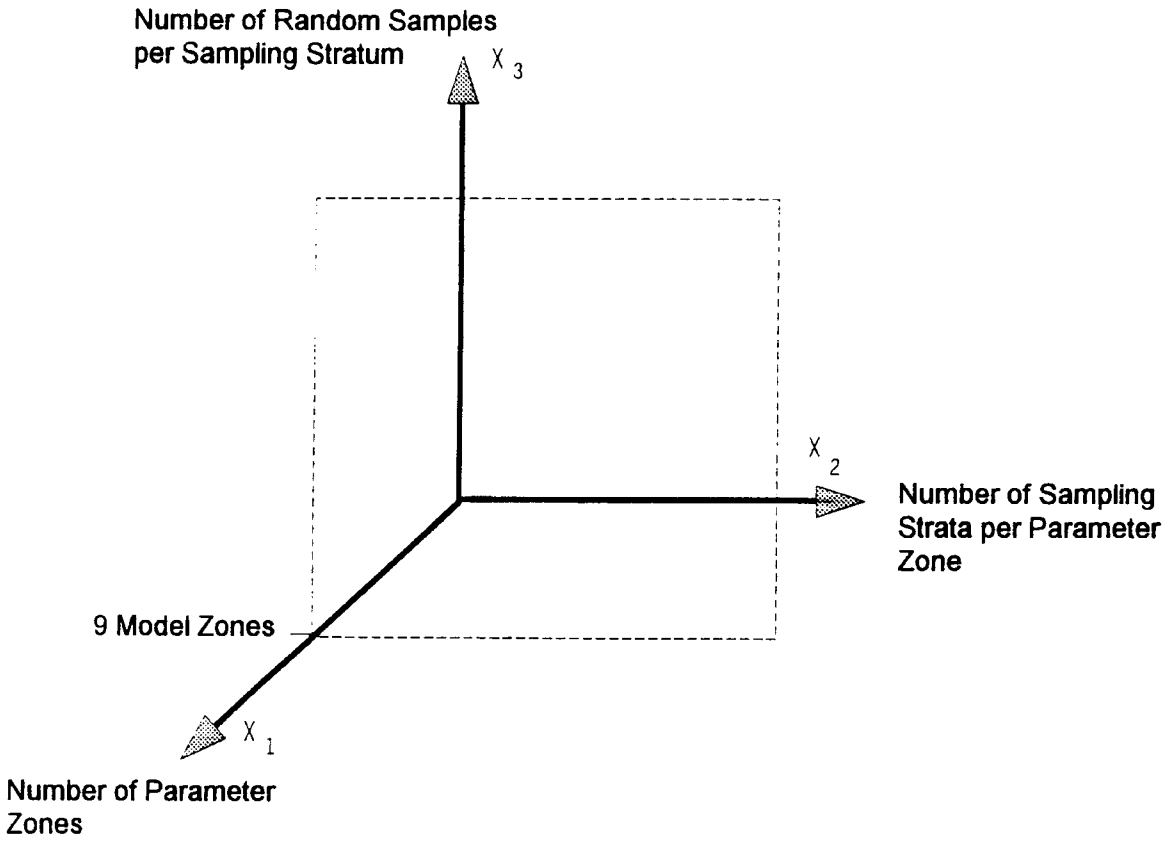
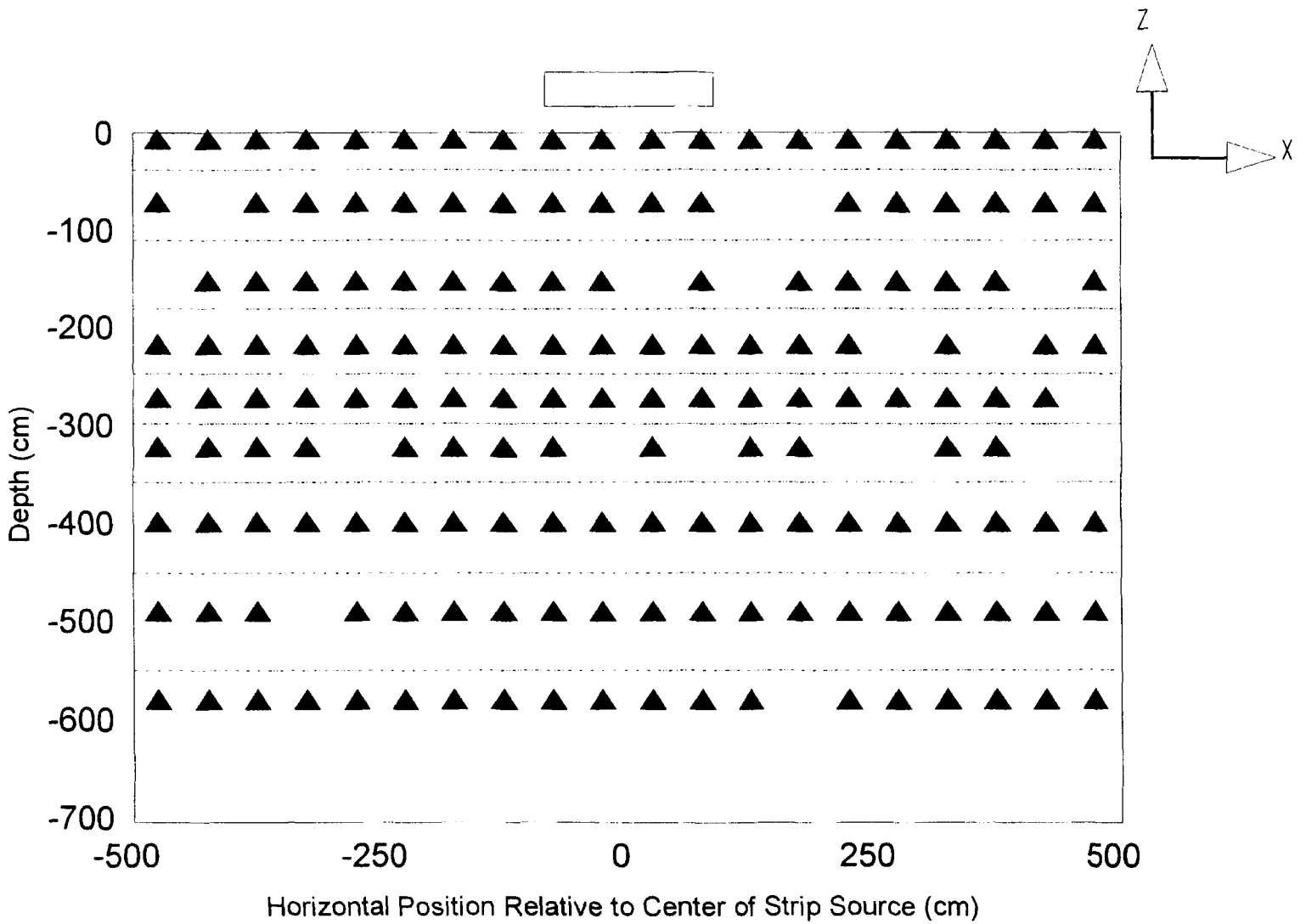
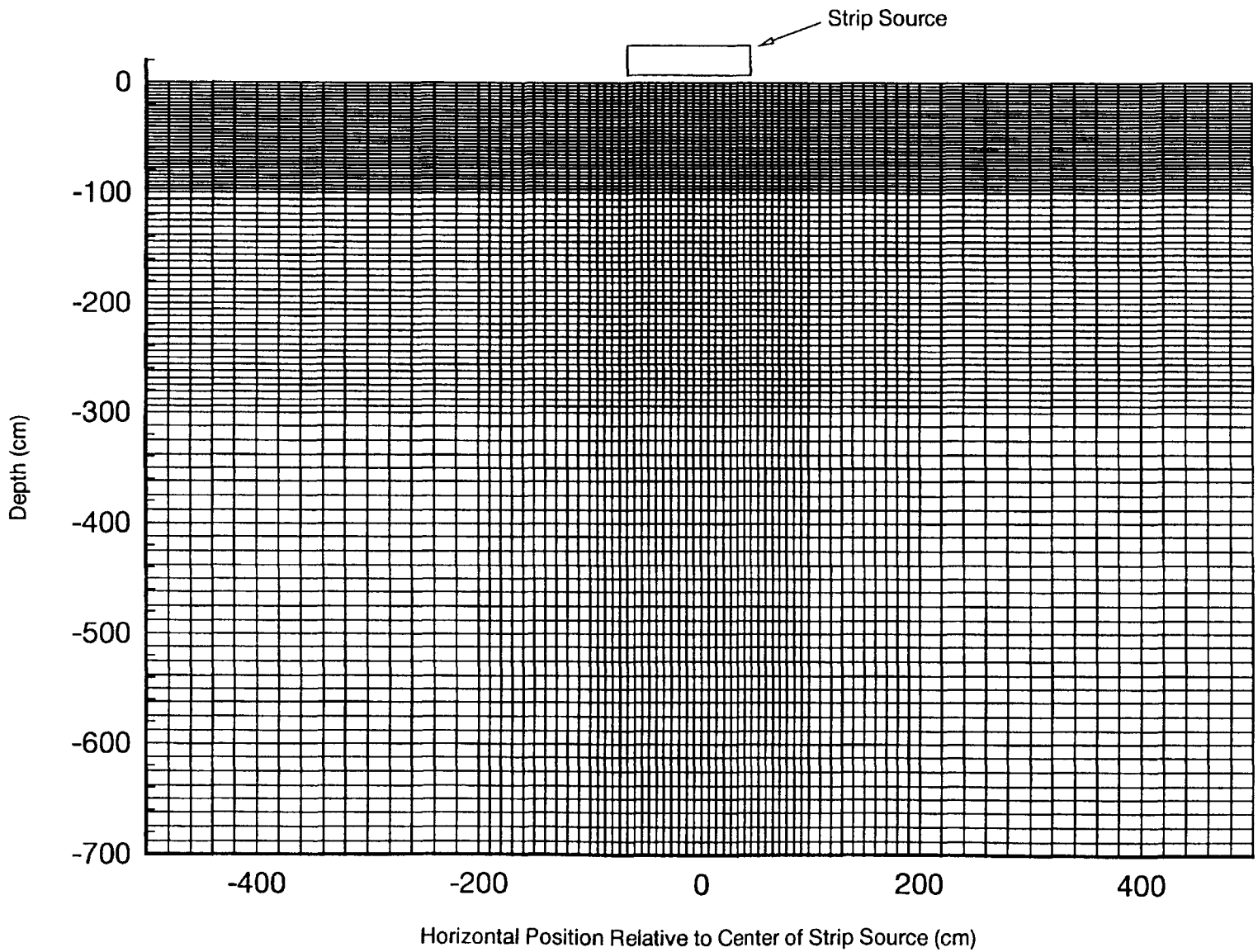

North
y = 2m transect
4.8 m
1.2 m
12.0 m
Plot #2
4.0 m
26.4 m
9.0 m
Plot #1
Covered Trench 6.0 m Deep
Irrigated Areas
Neutron Access Tube 6.0 m Deep
Neutron Access Tube 1.5 m Deep

Fig 3 26/43


Number of Random Samples per Sampling Stratum
X 3
X 2
Number of Sampling Strata per Parameter Zone
9 Model Zones
X 1
Number of Parameter Zones


Z
X
0
-100
-200
-300
-400
-500
-600
-700
Depth (cm)
-500
-250
0
250
500
Horizontal Position Relative to Center of Strip Source (cm)


Strip Source
0
-100
-200
-300
-400
-500
-600
-700
Depth (cm)
-400
-200
0
200
400
Horizontal Position Relative to Center of Strip Source (cm)

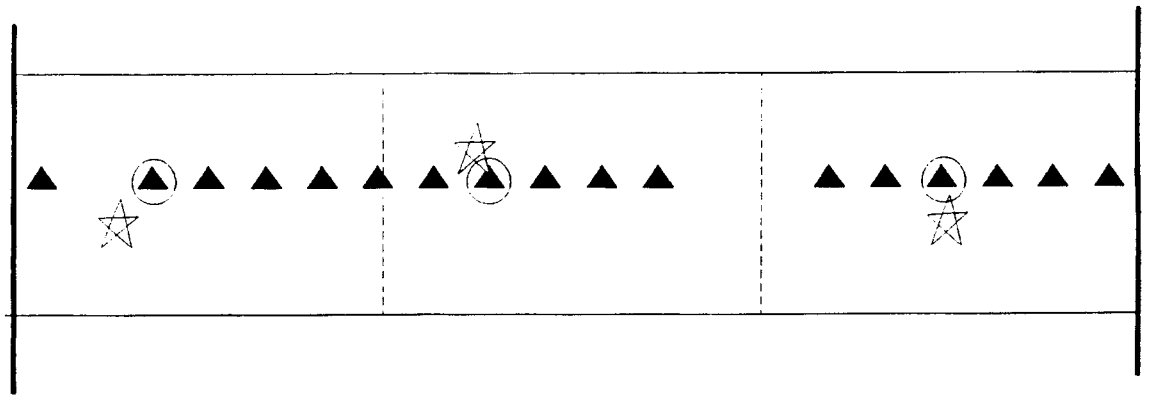
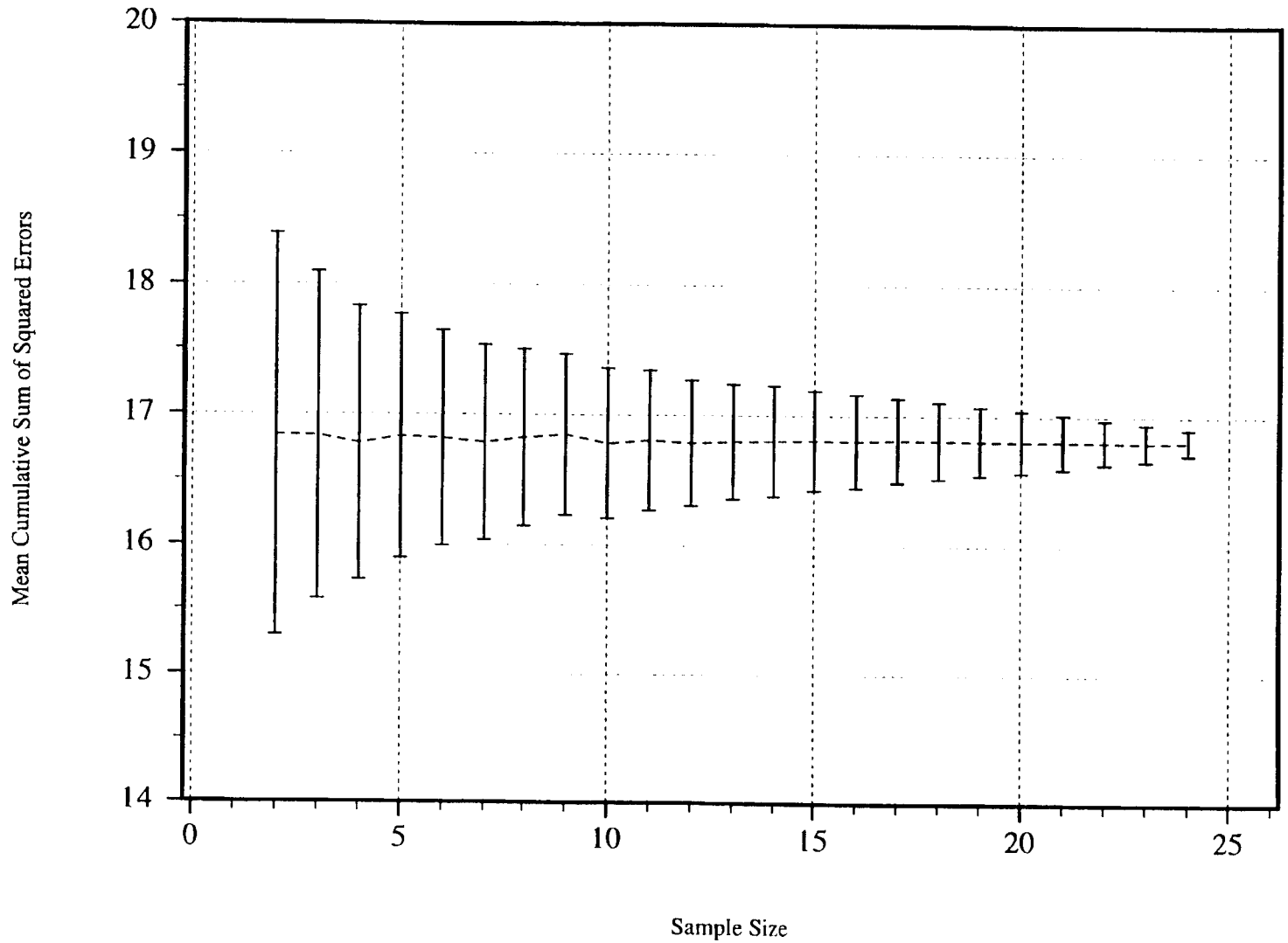


Fig 2


20
19
18
17
16
15
14
0
5
10
15
20
25
Mean Cumulative Sum of Squared Errors
Sample Size

31/43

Fig 5

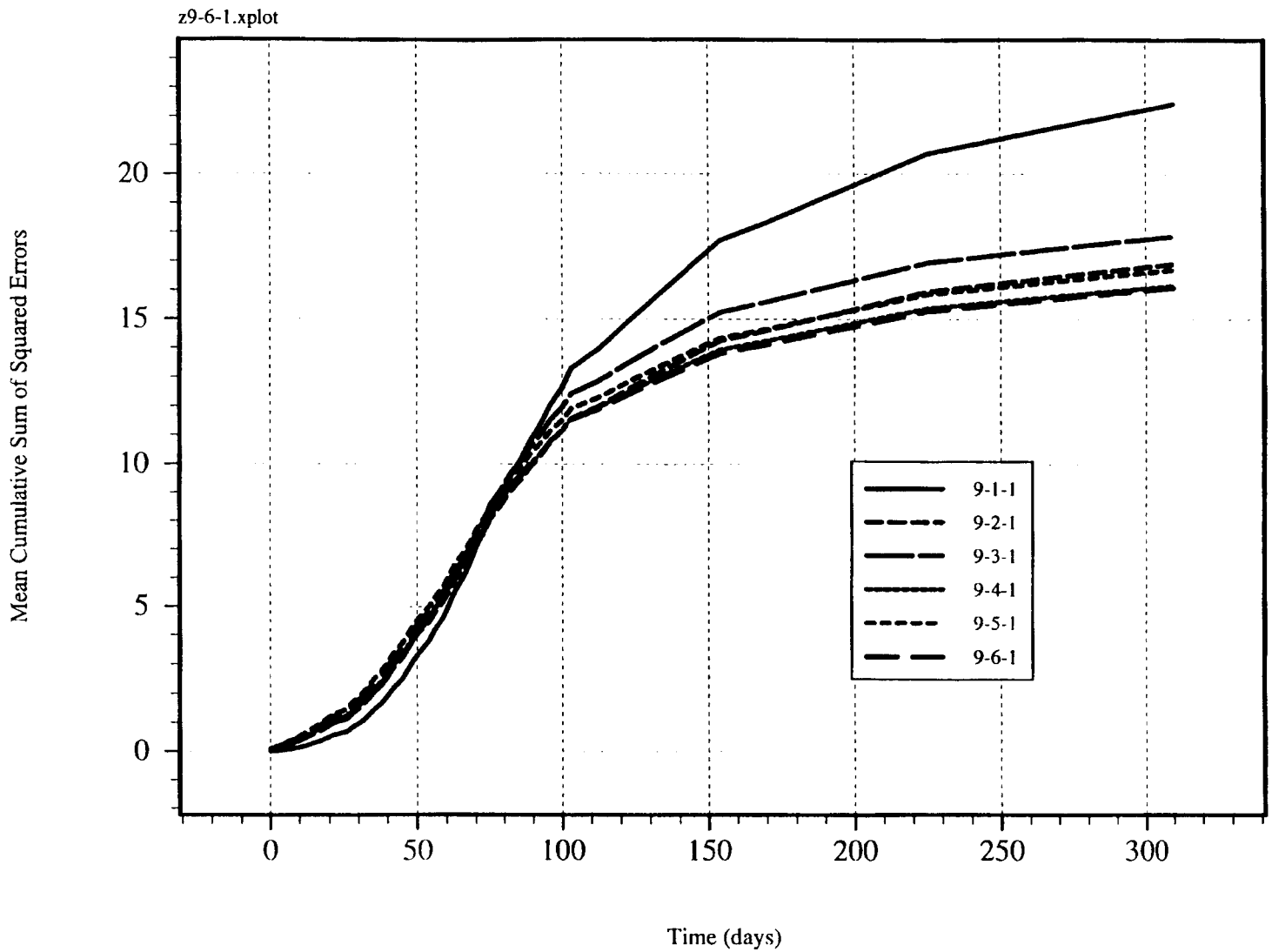

z9-6-1.xplot
Mean Cumulative Sum of Squared Errors
20
15
10
5
0
0
50
100
150
200
250
300
Time (days)
9-1-1
9-2-1
9-3-1
9-4-1
9-5-1
9-6-1

Fig 9

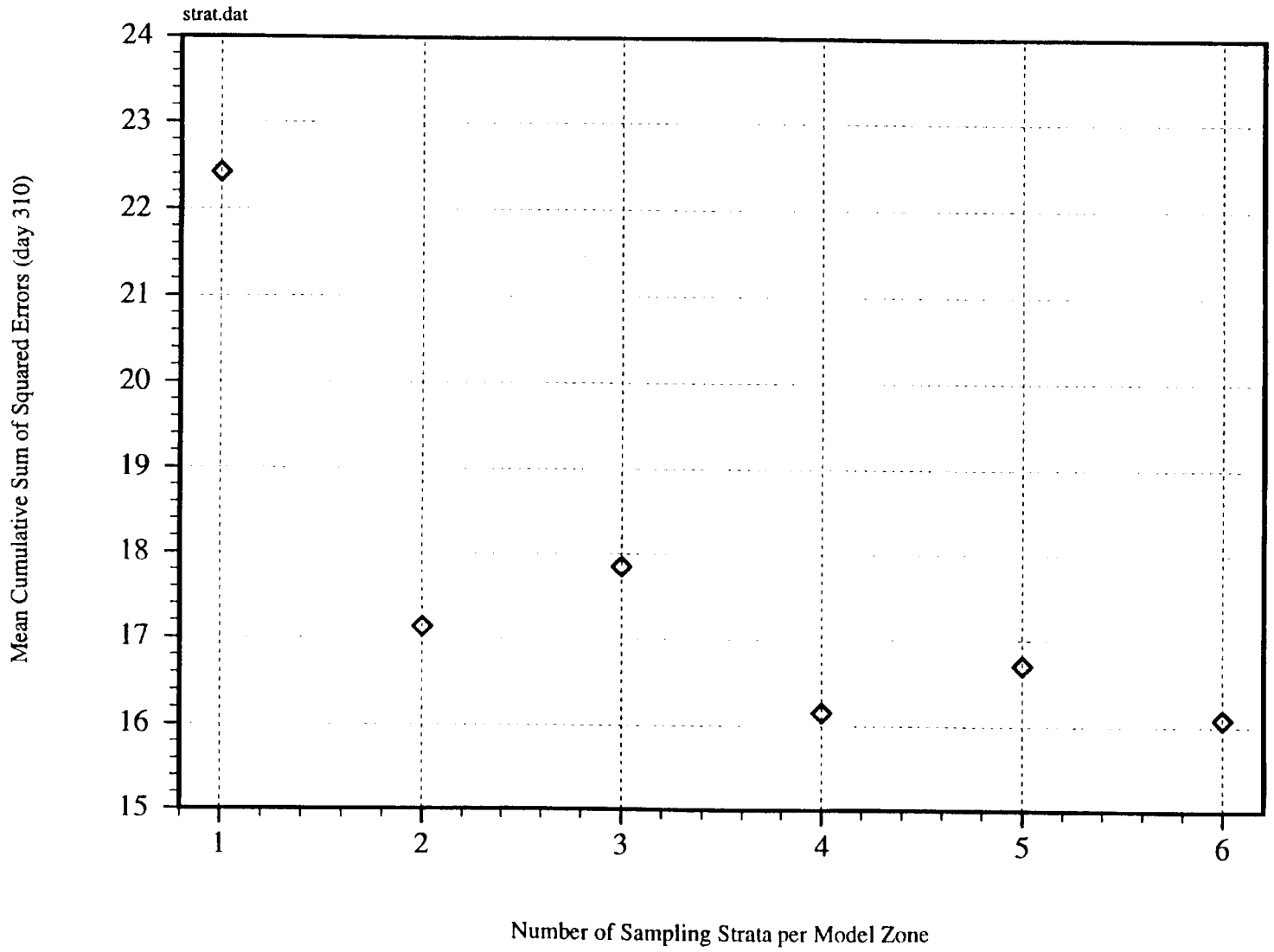

strat.dat
24
23
22
21
20
19
18
17
16
15
1
2
3
4
5
6
Mean Cumulative Sum of Squared Errors (day 310)
Number of Sampling Strata per Model Zone

Fig 10

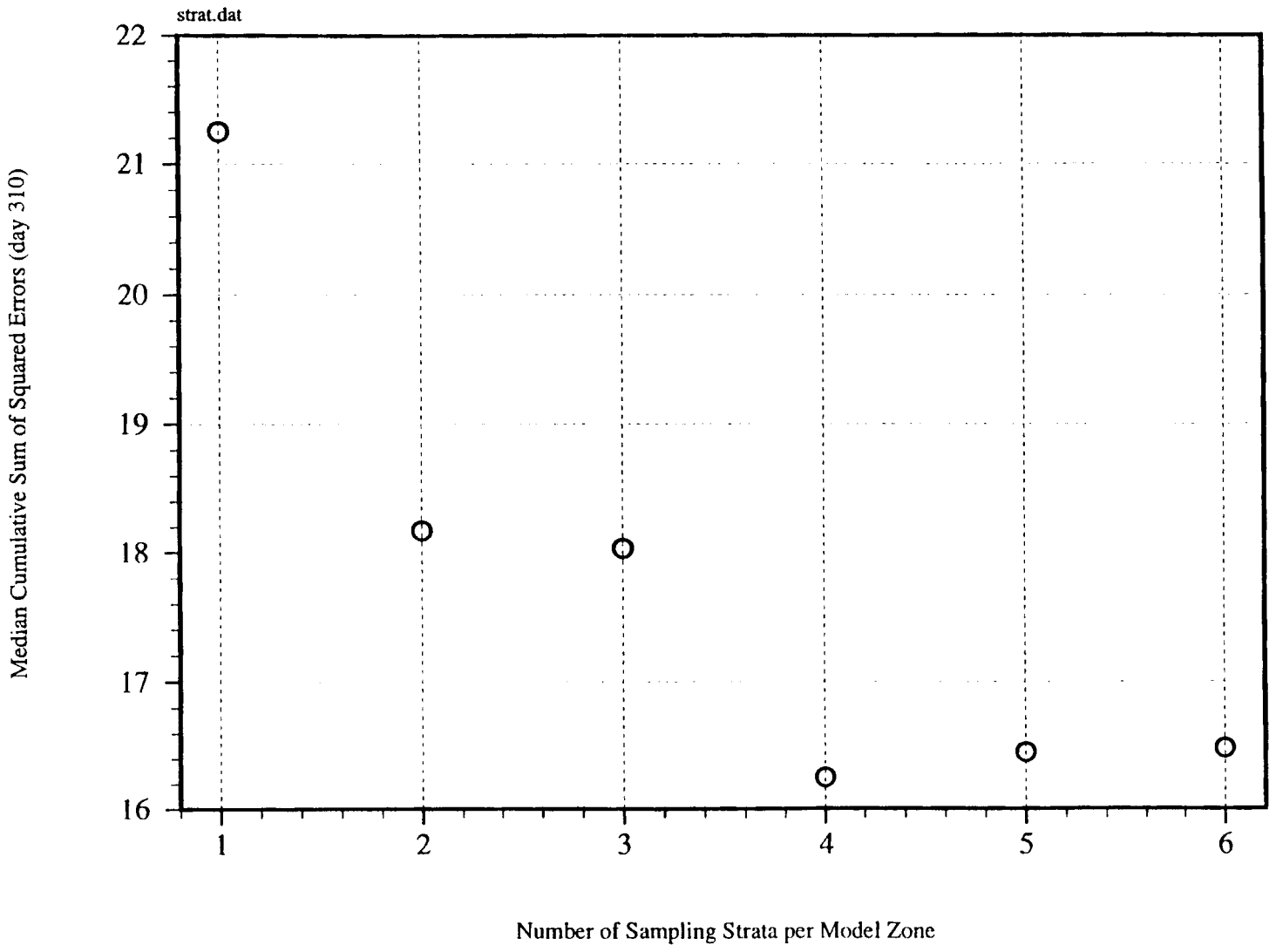

strat.dat
22
21
20
19
18
17
16
1
2
3
4
5
6
Median Cumulative Sum of Squared Errors (day 310)
Number of Sampling Strata per Model Zone

Fig 11

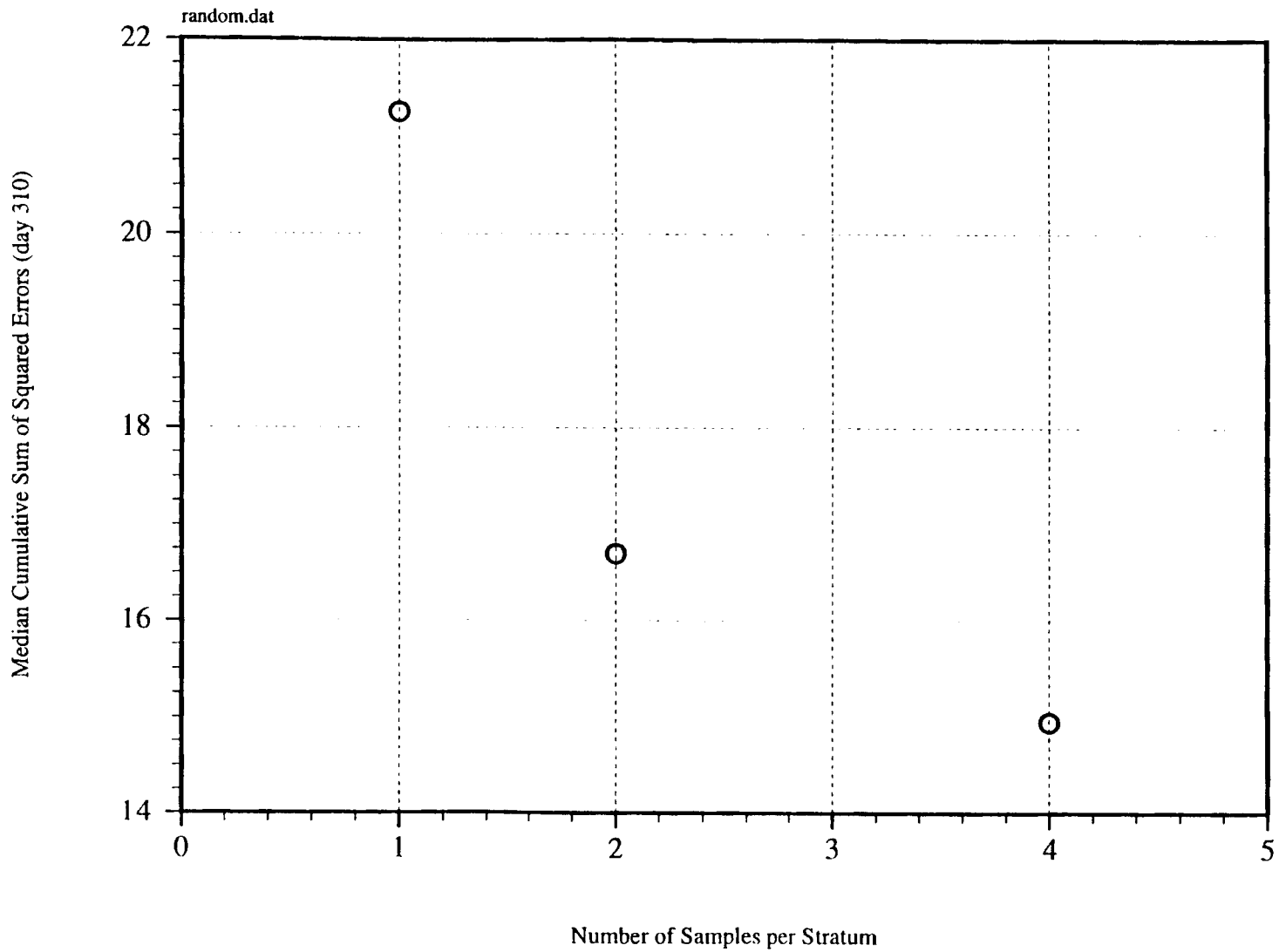
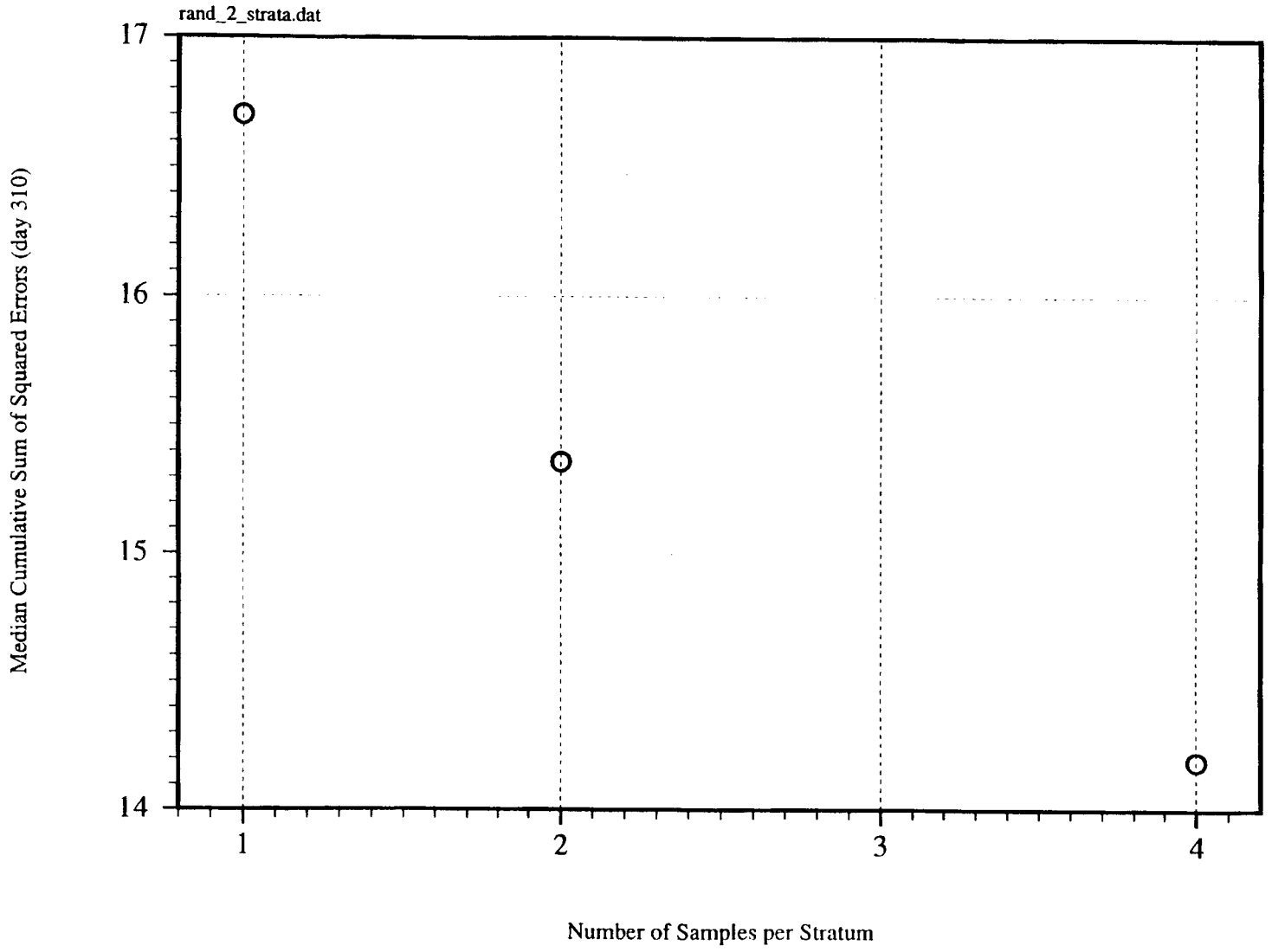
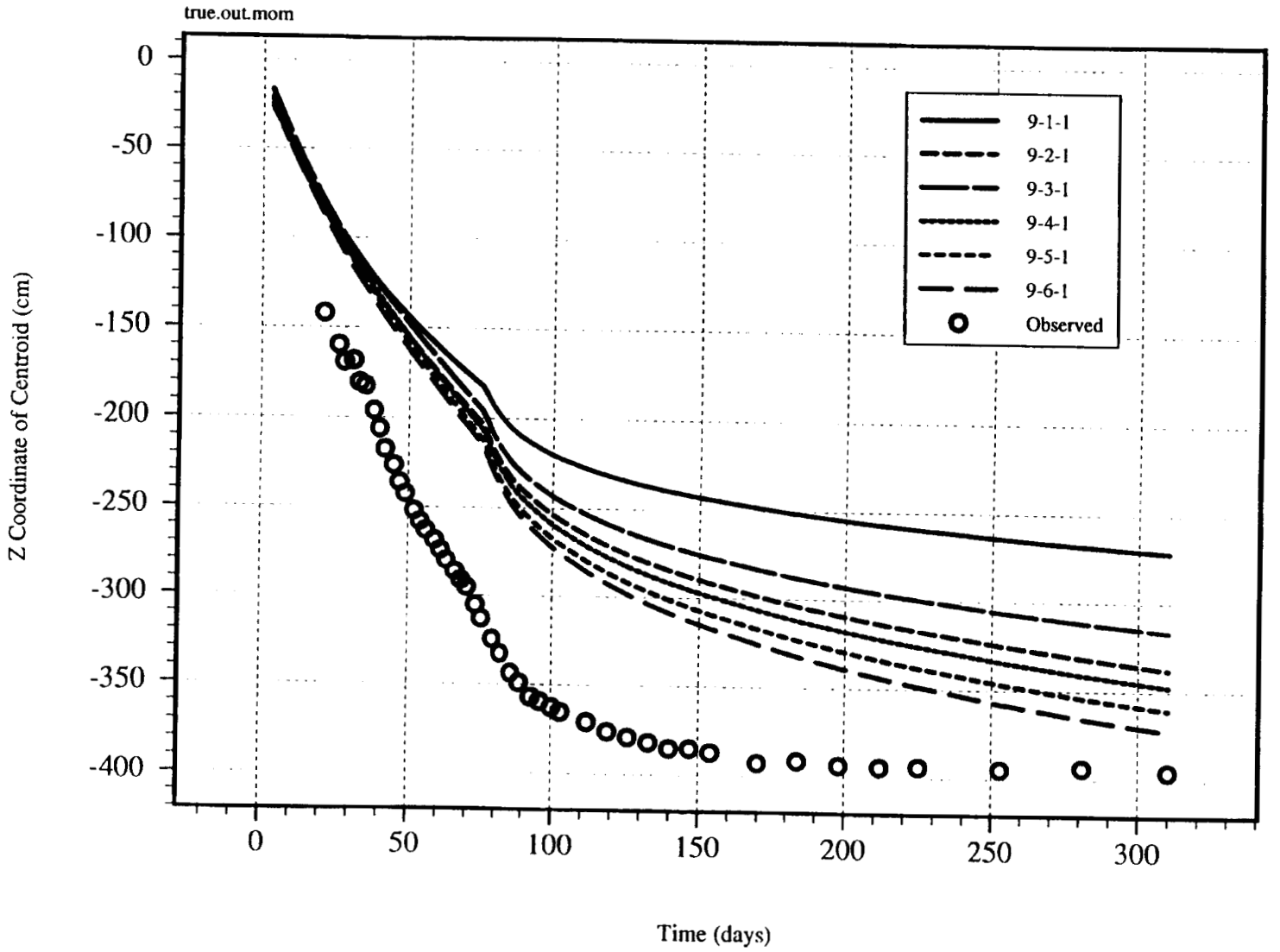

random.dat
Median Cumulative Sum of Squared Errors (day 310)
22
20
18
16
14
0
1
2
3
4
5
Number of Samples per Stratum

Fig 12


rand_2_strata.dat
17
16
15
14
1
2
3
4
Median Cumulative Sum of Squared Errors (day 310)
Number of Samples per Stratum

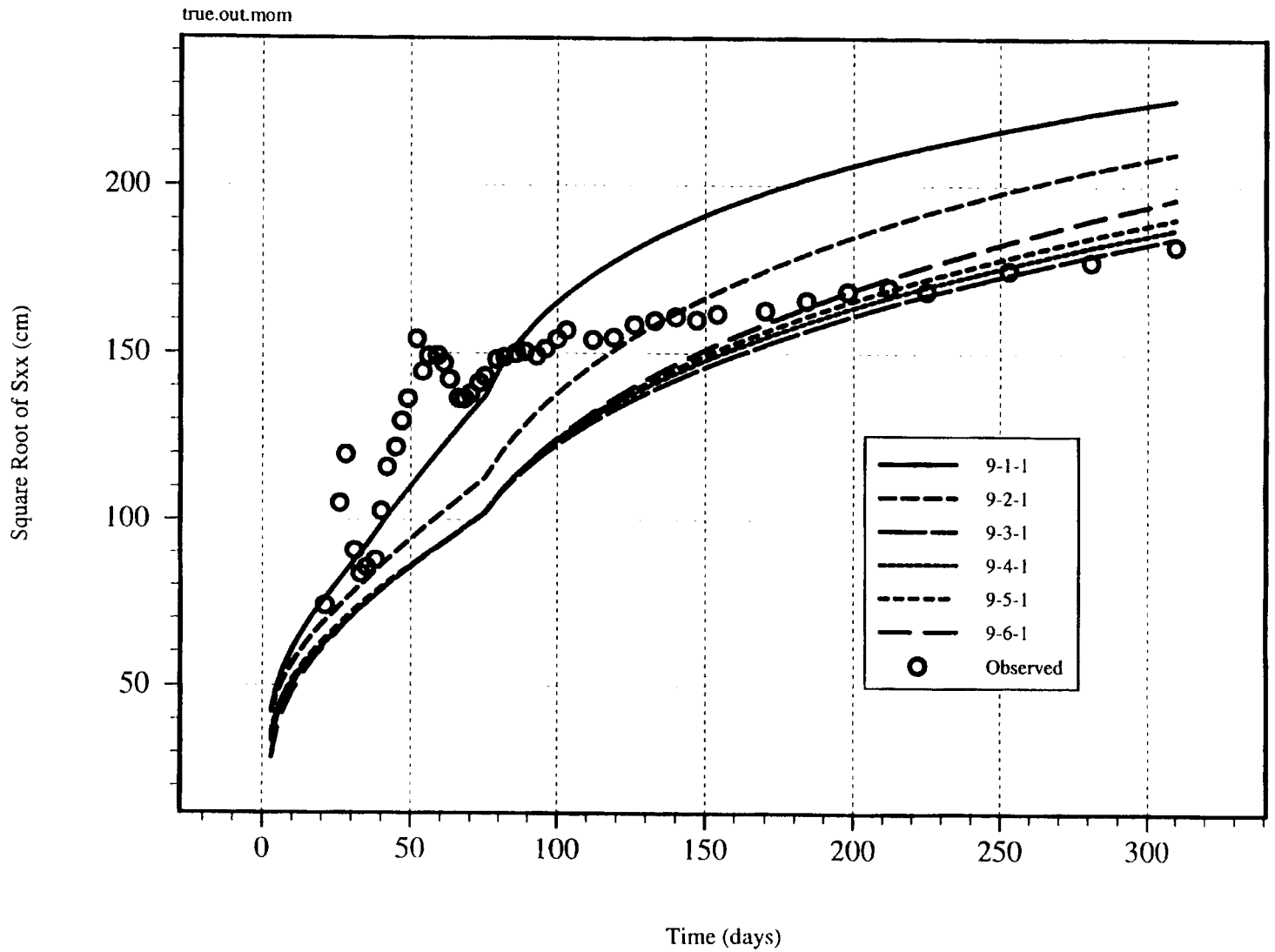
36/43

Fig 12a


true.out.mom
Z Coordinate of Centroid (cm)
0
-50
-100
-150
-200
-250
-300
-350
-400
0
50
100
150
200
250
300
Time (days)
9-1-1
9-2-1
9-3-1
9-4-1
9-5-1
9-6-1
Observed

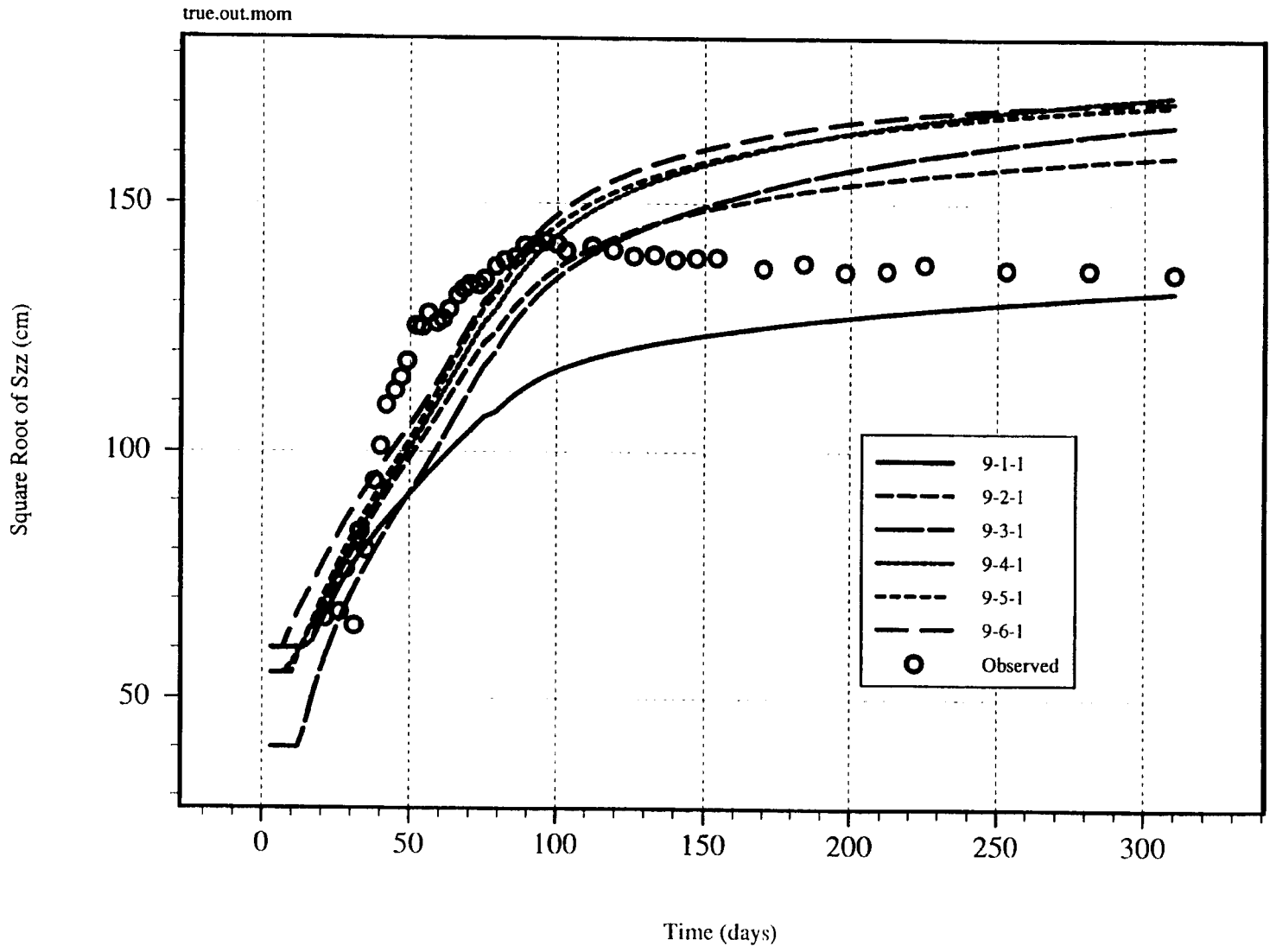
37/43

Fig 13b

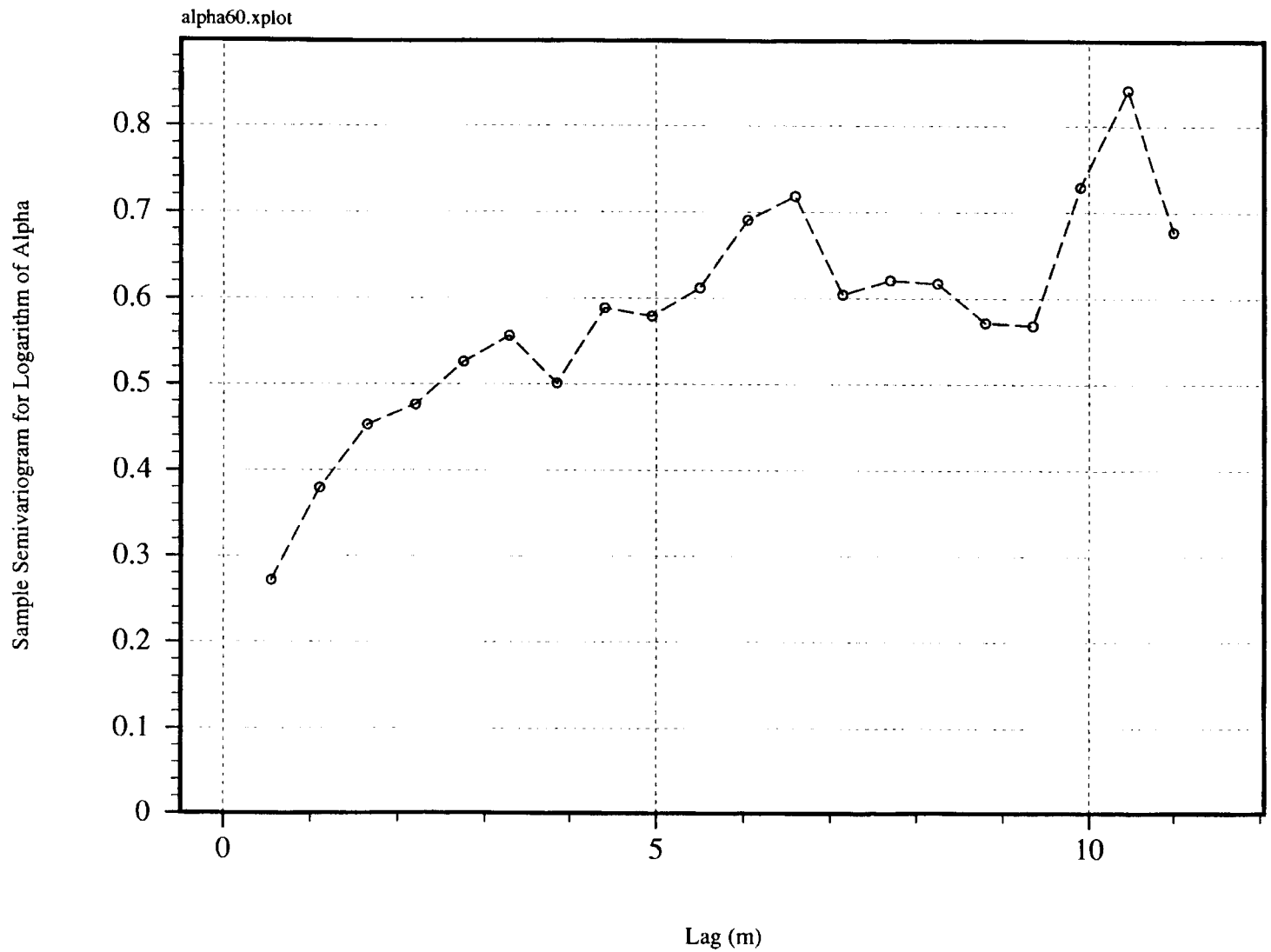

true.out.mom
Square Root of Sxx (cm)
200
150
100
50
0
50
100
150
200
250
300
Time (days)
9-1-1
9-2-1
9-3-1
9-4-1
9-5-1
9-6-1
Observed

38/43

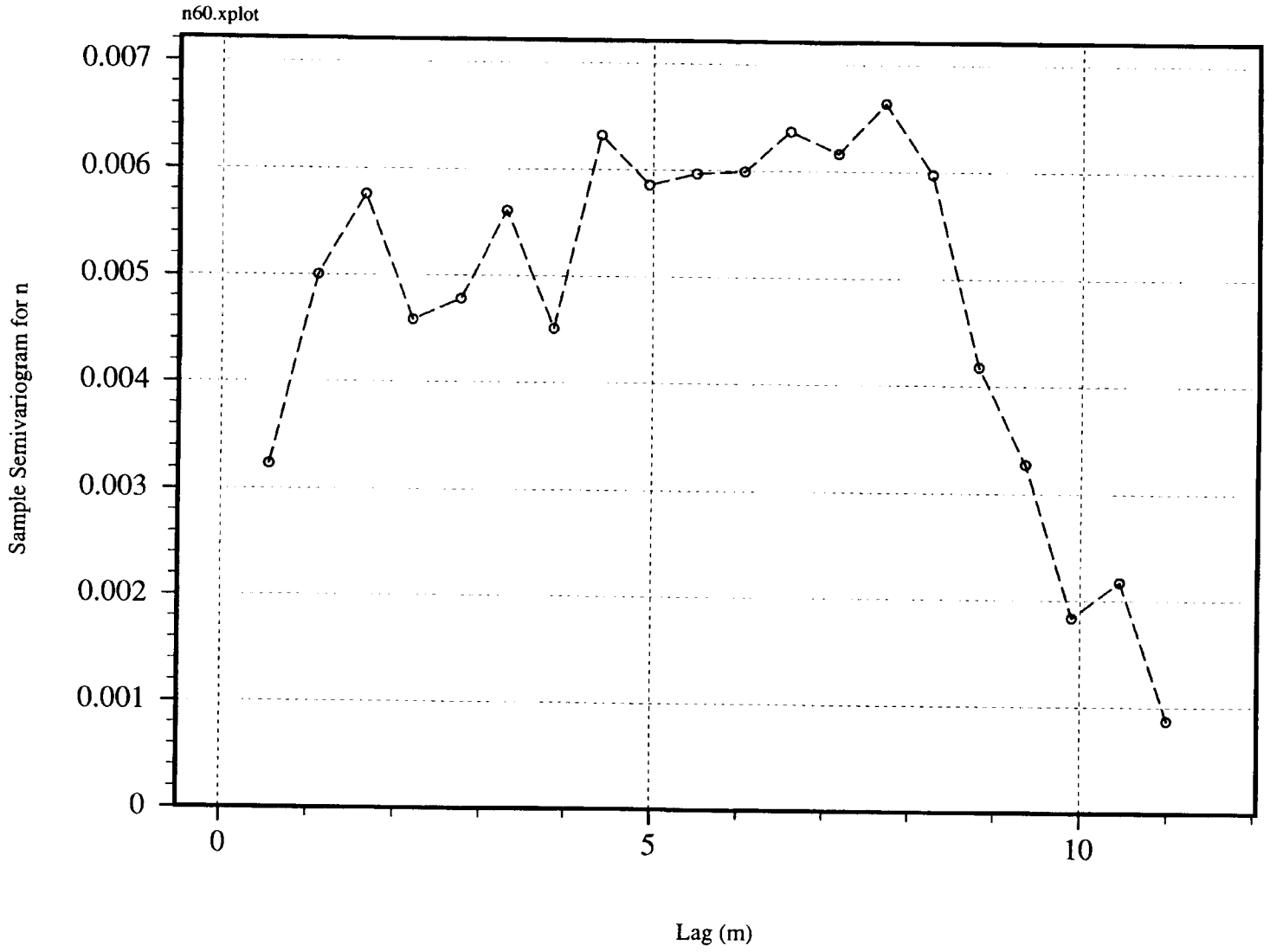
Fig 13c


true.out.mom
Square Root of Szz (cm)
150
100
50
0
50
100
150
200
250
300
Time (days)
9-1-1
9-2-1
9-3-1
9-4-1
9-5-1
9-6-1
Observed

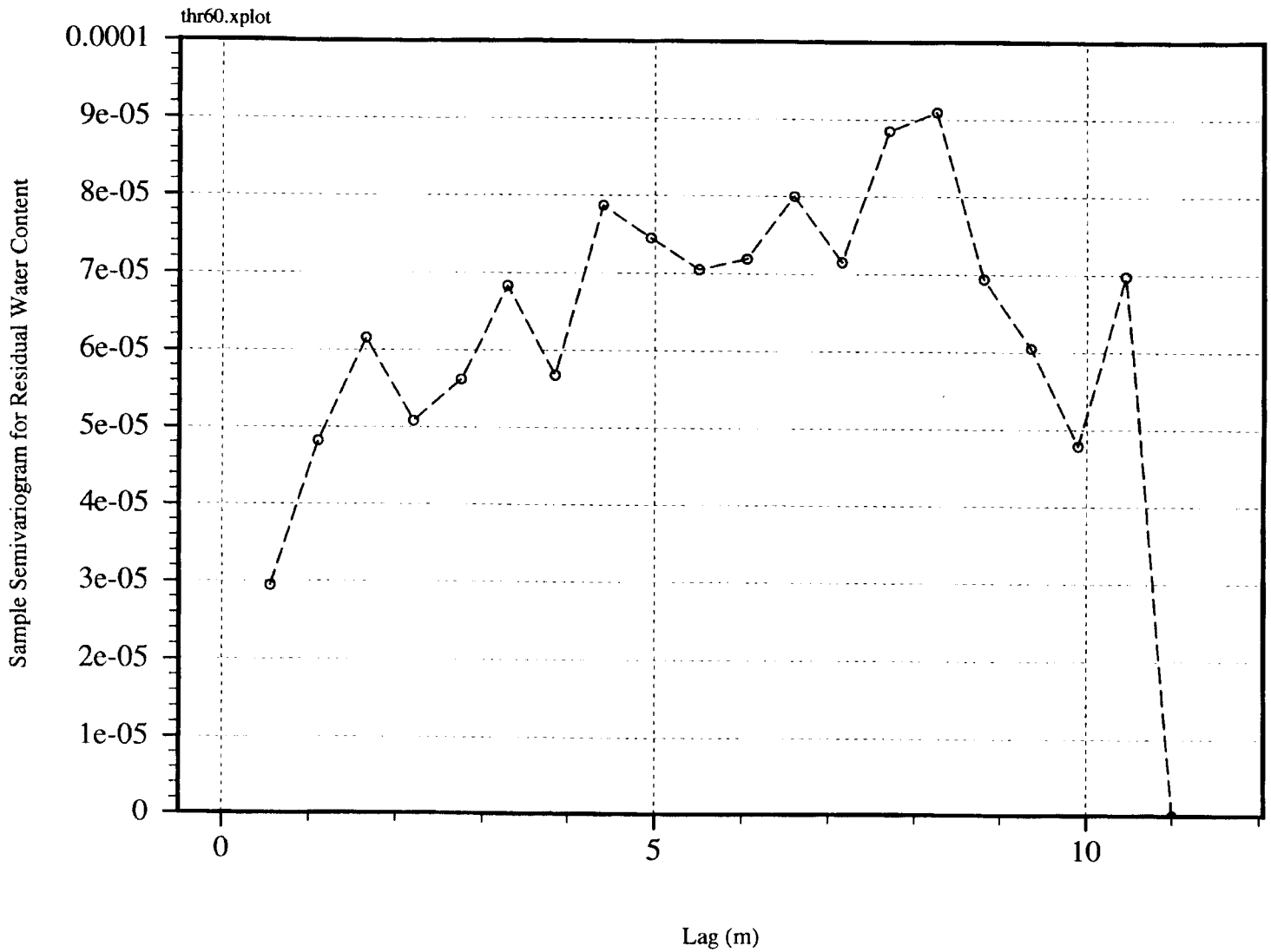
39/43


alpha60.xplot
Sample Semivariogram for Logarithm of Alpha
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
5
10
Lag (m)

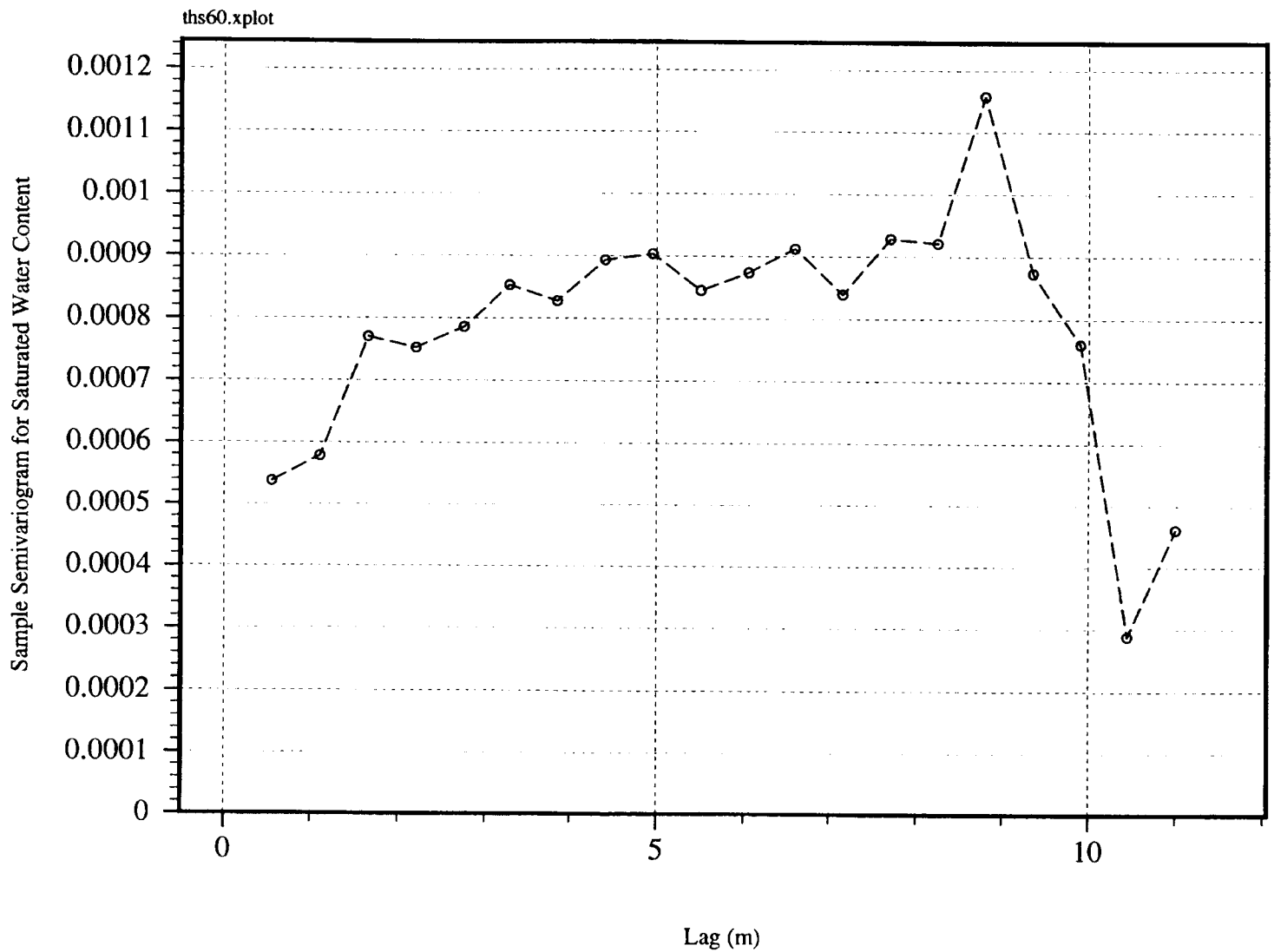
40/43

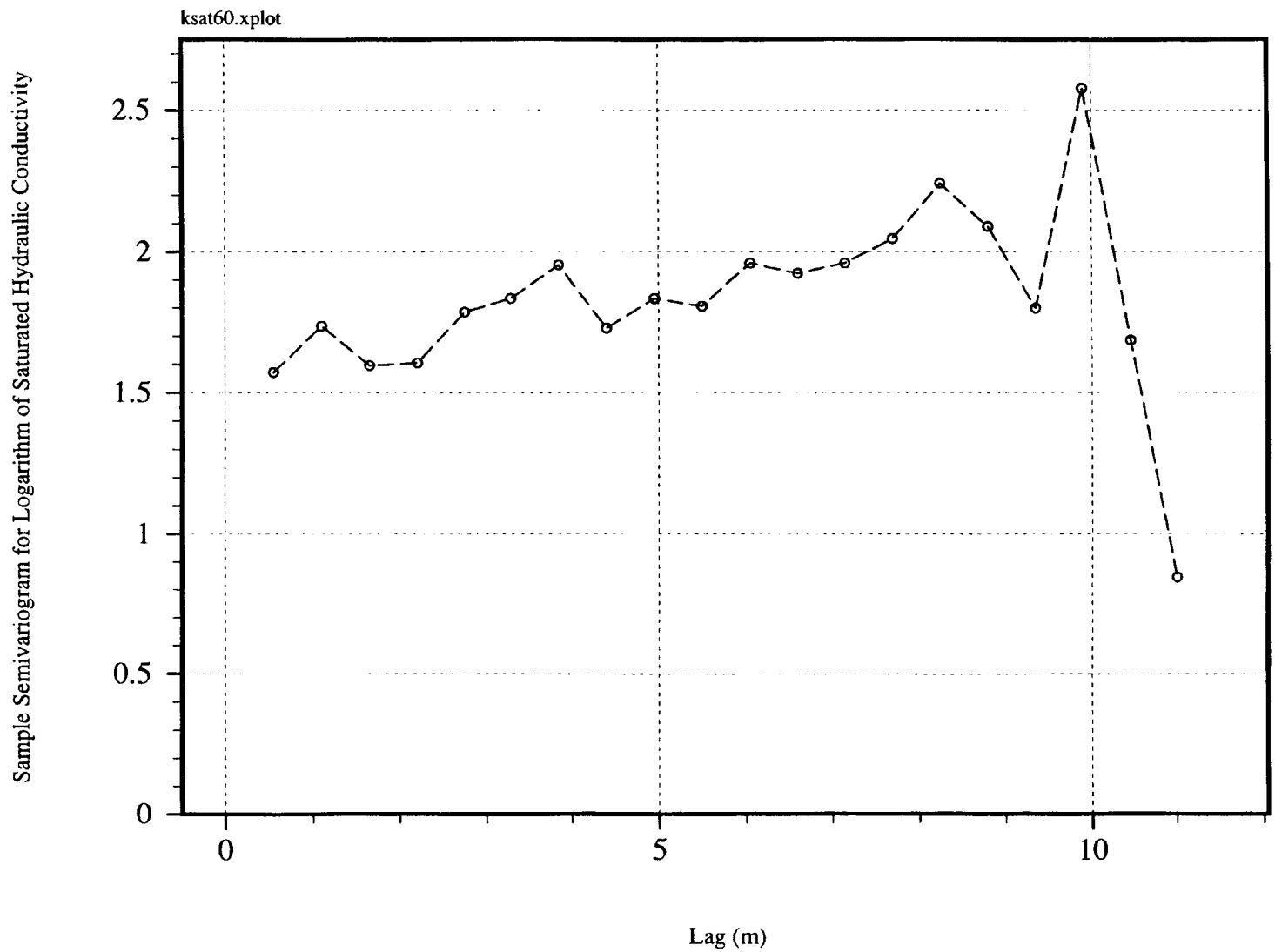

n60.xplot
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0
Sample Semivariogram for n
0
5
10
Lag (m)

41/43


thr60.xplot
0.0001
9e-05
8e-05
7e-05
6e-05
5e-05
4e-05
3e-05
2e-05
1e-05
0
Sample Semivariogram for Residual Water Content
0
5
10
Lag (m)

42/43


ths60.xplot
Sample Semivariogram for Saturated Water Content
0.0012
0.0011
0.001
0.0009
0.0008
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0
0
5
10
Lag (m)


ksat60.xplot
Sample Semivariogram for Logarithm of Saturated Hydraulic Conductivity
2.5
2
1.5
1
0.5
0
0
5
10
Lag (m)