VARIABILITY OF HYDRAULIC CONDUCTIVITY AND SORPTION IN A HETEROGENEOUS AQUIFER

by

GRACE HARRIET FOSTER-REID

B.S. Civil Engineering, Rice University
(1992)

Submitted to the Department of
Civil and Environmental Engineering
in Partial Fulfillment of the
Degree of

MASTER OF SCIENCE
IN CIVIL AND ENVIRONMENTAL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
March 1994

Signature of the Author _____

Department of Civil and Environmental Engineering
March 16, 1994

Certified by _____

Lynn W. Gelhar
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by _____

Professor Joseph Sussman
Chairman, Department Graduate Committee

VARIABILITY OF HYDRAULIC CONDUCTIVITY AND SORPTION IN A HETEROGENEOUS AQUIFER

by

GRACE HARRIET FOSTER-REID

Submitted to the Department of Civil and Environmental Engineering
March, 1994
In Partial Fulfillment of the Requirements for the Degree of
Master of Science in Civil and Environmental Engineering

ABSTRACT

The variability of the hydraulic conductivity (K) and the sorption coefficient (K_d) and the correlation between these two variables leads to the enhanced dispersion of contaminants. Seventy-three (73) samples, at a spacing of 0.5 m, were taken from a horizontal transect, and 26 samples, at a sampling interval of 0.15 m, were taken from a vertical transect on a vertical undisturbed face at the Handy Cranberry Bog on Cape Cod, Massachussetts. The soils at this site, Mashpee Pitted Plain Deposits, are composed of the same glaciofluvial outwash sediments as the soils at the USGS test site. The test site is near the source of an extensively studied sewage contamination plume, which extends several kilometers into a sole source aquifer that supplies drinking water for the Cape Cod area.

The hydraulic conductivity was measured with a falling head permeameter. The variance of $\ln K$ is 0.080; the correlation length is 1 m for the horizontal transect and 0.1 m for the vertical transect. The percent organic carbon (POC) was measured using a CHN Analyzer. The POC measurements and the partioning coefficient, K_{oc}, were used to calculate K_d for three hypothetical contaminants: benzene, trichloro-benzene (TCB) and pentachlorobenzene (PentCB). The soils at the bog have an average POC value of 0.018%. The variance of the POC and $\ln K_d$ for the entire sample population is 1×10^{-3} and 0.145. For a series of ten measurements of POC of the same sample, the variance is 4×10^{-6}. The correlation length of $\ln K_d$ is 1 m for the horizontal transect and 0.15 m for the vertical transect.

A linear regression of $\ln K_d$ unto $\ln K$ shows a significant correlation with a slope (b) of -0.341 and a coefficient of correlation (r) of -0.233. The correlation

between ln$K$ and lnK_d is expressed in terms of the slope and an uncorrelated residual term using a linear regression model. Several tests, based on cross-covariance and cross-spectral analyses, indicate that this simple linear regression model is an adequate representation of the correlation structure of the ln K_d and ln$K$ fields.

The longitudinal macrodispersivity for a nonreactive tracer is 0.076 m. For reactive tracers such as benzene, TCB and PentCB, the macrodispersivity is 0.081 m, 0.142 m and 0.177 m respectively.

Thesis Supervisor: Dr. Lynn W. Gelhar

Title: Professor of Civl Engineering

ACKNOWLEDGEMENT

This research was in part funded by the Parson's Fellowship from M.I.T's Civil and Environmental Engineering Department, and the Alcan Jamaica Limited Graduate Fellowship. During fieldwork, additional support was provided by the National Science Foundation (Grant No. CES-8814615) and the U.S. Geological Survey, New England District.

For his continual effort to guide, counsel and channel my energies, I thank Professor Lynn Gelhar. In his ability to impart his intuitive knowledge without stifling my thought process, Dr. Gelhar showed that he is a devoted and patient advisor. I must thank Dr. Gelhar for his effort to bridge the gaps.

I also thank the people at MIT who provided the support, advise and assistance that made this thesis a reality; namely, Pat Dixon, Dean Ike Colbert, Carmen Turner, Jesse Carty, Cynthia Stewart, Carol Solomon, Jackie Winton, Kay Thompson, Allison MacKay, Britt Holmen, Joshua Josephs and Kishan Amarasekara. I thank Dennis Le Blanc, Larry Barber and Kathy Hess at the USGS for providing insight during the developmental stages of this project, and arranging for equipment during fieldwork. Brian Handy willingly allowed sampling on his property, for that I am most thankful.

For their understanding and compromising spirit, I voice my appreciation to John Sparks and Mac Hall of Malcolm Pirnie Inc. I thank my coworkers, especially the Administrative staff, who continually provided input and stress-relief. A special note of thanks to Carla Dixie who helped with the drawings. In locating a PE

Analyzer and running the soil samples, I appreciate the assistance of Larry Wallace, Jennifer Jacob and Marie Johnson.

I exalt my family; they are the epitome of tolerance. Foremost, I thank my husband, Clifton Reid, whose undying support and encouragement allows me to soar and grow beyond my expectations. My parents, Harry and Tina, continually provided the motivation, resources and ethical background that has made my education a possibility. I thank Carl and Cina, Karen and Duane, and Ret; who showed me that it could be done. My God Parents, Brenda and P. Elliott, instinctively acted through the hearts of my family when they were not here in body.

> .... *A wise man will hear, and will increase learning; and a man of understanding will attain unto wise counsels: ... The fear of the Lord is the beginning of knowledge.*
>
> *Proverbs 1: 4-6*

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

"The last decade has seen rapid developments in theoretical research treating groundwater flow in a probabilistic framework, but actual field applications for these methods has been very limited. The practitioner wishing to apply such theory faces a bewildering maze of abstract probabilistic concepts and apparently conflicting methodologies. The natural tendency of much research in this area has been to emphasize theoretical refinements or computational methodology without developing techniques which address real field problems." (Gelhar, 1986)

Millions of dollars worth of investigative and remediation work has been performed on a contaminant plume found in the Cape Cod area. This plume resulted from the disposal of secondary-treated sewage effluent in rapid infiltration beds for more than 50 year. This 4000 m long, 1000 m wide and 30 m thick plume of groundwater plume contaminates a sole source aquifer that provides drinking water for the Cape Cod area. Since 1978, this plume has been the subject of an extensive, multidisciplinary study. Much of the work has been done at the U.S. Geological Survey's Cape Cod Toxic-Substances Hydrology Research Site, hereafter referred to as the test site, which is near the rapid-infiltration wastewater disposal facility at the Otis Air Force Base. The test site is within two kilometers (km) of the Handy Cranberry Bog, hereafter called the bog, where the samples for this study were taken from an excavated exposure. Figure 1.1 shows the location of the bog

and the test site. The soils at the bog and test site are of a similar geologic structure since they are formed of the same glacial outwash sediments. This soil type is classified as Mashpee Pitted Plain deposits and is defined as "mostly gravelly sand and pebble to cobble gravel. Locally includes boulders." (*U.S. Geological Survey Geologic Map of Cap Cod*).

1.1 BACKGROUND

In understanding the intricacies of the transport of a contaminant plume, such as that found at the test site, one must understand the porous media through which it moves. The movement of contaminants is caused by a concentration gradient that results in diffusion and dispersion, the latter of which will be a major focus in this study. Macrodispersion (A_{ij}) is a parameter which quantifies the dispersion of the contaminant. In this study, the focus will be on the dispersion in the direction of flow, which is assumed as horizontal. This longitudinal macrodispersion is denoted as A_{11}.

One factor that affects the contaminant movement is the heterogeneity of the soil as evidenced by the variability of the hydraulic conductivity (K). The effects of variability of K on the subsurface flow system has been the subject of extensive study for the past 20 years. Several researchers (*Sudicky et al, 1983; Garabedian et al.; 1988*) have found that the coefficients of the advective dispersion equation (ADE) which quantify the amount of dispersion, A_{ij}, tends to increase with residence time and travel distance. This can be explained by, the theory of enhanced mixing

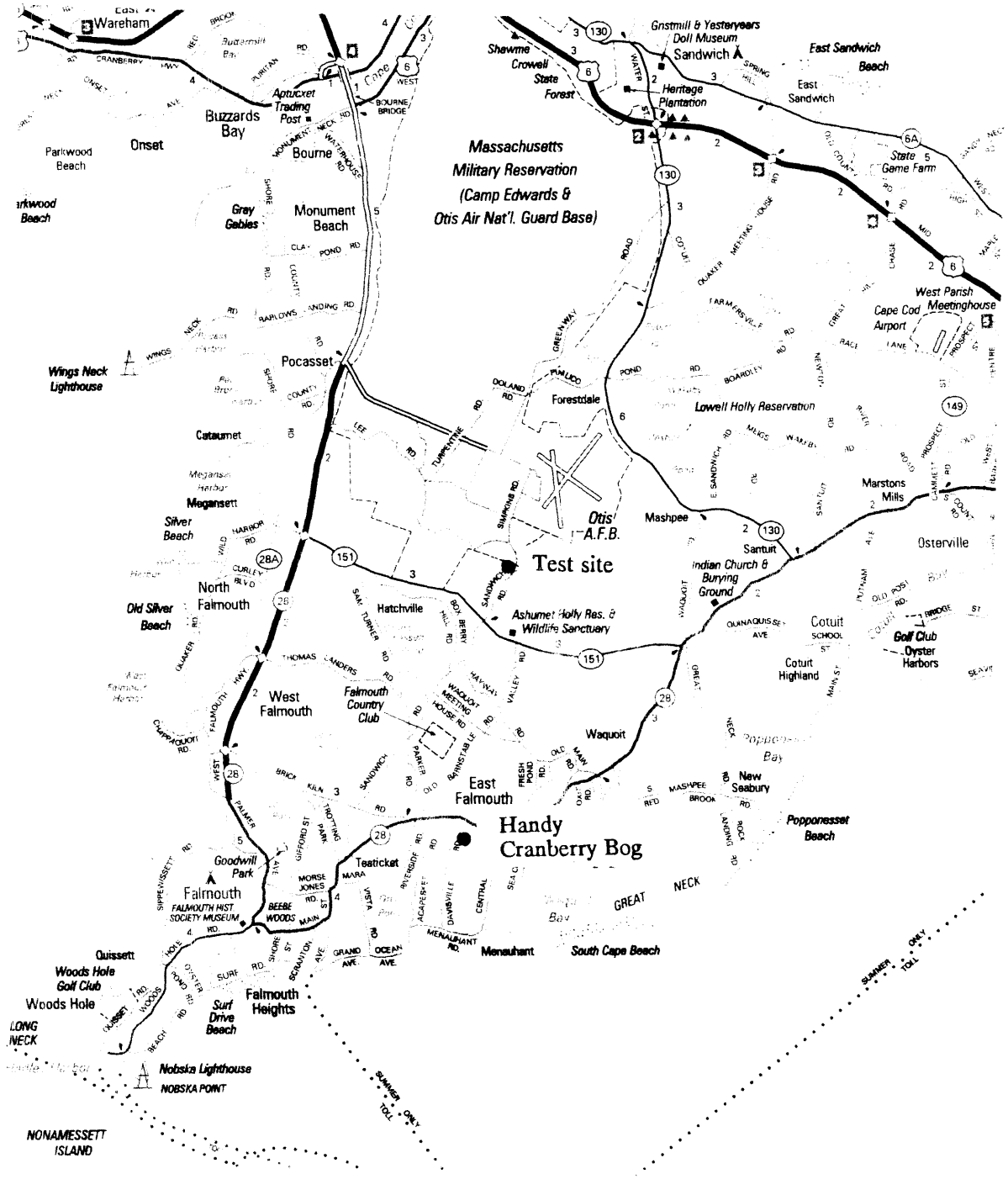

Figure 1.1 Location of the Handy Cranberry Bog and the test site, western Cape Cod, MA. Scale: 1 inch equals 2.5 miles (courtesy of AAA)

at larger scales, caused by the small-scale variation of K, which results in differential velocity as mandated by the K value (*Gelhar et al., 1979; Matheron and de Marsily, 1980; Dagan, 1982; Gelhar and Axness, 1983; Mackay et al, 1986; Neuman et al., 1987*). Researhers have developed models that account for the effect of enhanced dispersion caused by this K variation on the transport process (*Gelhar and Axness, 1983; Gelhar et al 1979*).

In recent years, the focus has broadened to include investigations of the effect of chemical heterogeneity on contaminant transport. In a situation where the velocity of the solute is retarded by sorption to the porous media the macrodispersivity is further affected. A measure of the sorptive capacity of the soil is the distribution coefficient (K_d) the ratio of the concentration of contaminant sorbed to the soil to the concentration of contaminant in solution. In particular, it has been shown via stochastic theory that the correlation between the hydraulic conductivity and sorption strongly affects the spreading of contaminants (*Garabedian et al, 1988; Valochi, 1989; Chysikopoulos et al, 1990; Robin et al., 1991*).

Field application of this stochastic theory is very limited. In *1986*, Sudicky took a total of 1279 measurements on 32 cores at a tracer test site at the Canadian Forces Bases (CFB) Borden. The results showed that the dispersion parameters for nonreactive solute found using stochastic theory and the statistics of the variability of K is consistent with that those estimated from the spread of the plume (*Freyberg, 1986).* Robin et al. (*1991*) made 1279 measurements of strontium K_d for the same cores and found a weak but significant correlation between $\ln K$ and $\ln K_d$; however, they did not evaluate the enhancement of the macrodispersivity based on their

results. By analyzing spatial moments of tracer distribution at the test site on Cape Cod, Garabedian et al. (*1988*) found that the reactive macrodispersivity for lithium was 10 times greater than the nonreactive macrodispersivity for bromide (see Figure 1.2). Mackay et al. (*1986*) studied the spatial heterogeneity of tetrachloroethelyene sorption at the Borden aquifer and found that K_d varied over an order of magnitude at a scale of a 2 meters.

1.2 MATHEMATICAL EXPRESSION OF MACRODISPERSIVITY

Using a Fickian model, as predicted by stochastic theory, the mass flux (F_D) of a contaminant is:

$$F_D = -qA_{ij}\frac{\partial c}{\partial x_j} \qquad [1.1]$$

where q is the specific discharge and $-\partial c/\partial x_j$ is the concentration gradient in the j direction. Applying the conservation of mass, the three-dimensional transport equation can be written as:

$$n\frac{\partial c}{\partial t} = \frac{\partial}{\partial x_i}\left[qA_{ij}\frac{\partial c}{\partial x_j} - cq_i\right] \qquad i,j = 1,2,3 \qquad [1.2]$$

where n is the porosity and t is the time. The two important variables in this equation are q_i and A_{ij}. The specific discharge is dependent on the ability of the medium to transmit water, otherwise called the hydraulic conductivity (K_{ij}) such that:

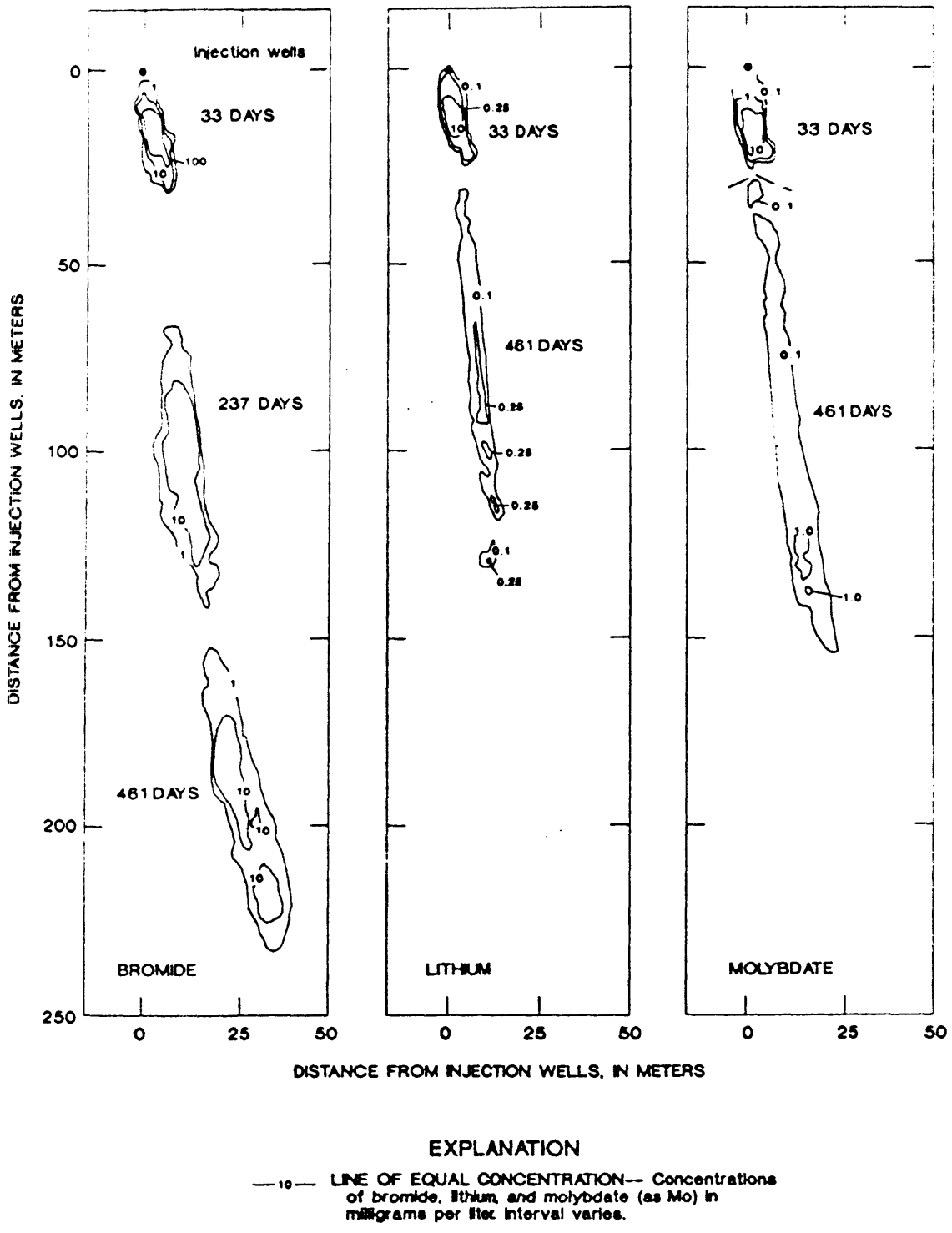

Figure 1.2 Distribution of the maximum concentrations of three tracers used at the Cape Cod site (from Le Blanc et al., 1991)

$$q_i = -K_{ij}\frac{\partial h}{\partial x_j} \qquad i,j = 1,2,3 \qquad [1.3]$$

where $-\partial h/\partial x_j$, also denoted as J_j, is the hydraulic gradient in the j direction. Based on Gelhar and Axness (*1983*), for a nonreactive tracer the longitudinal macrodispersity (A_o) is:

$$A_0 = \frac{\sigma_f^2\ \lambda}{\gamma^2} \qquad [1.4]$$

where σ_f^2 is the variance of ln$K$, λ is the correlation scale[1] in the direction of flow (horizontal) and γ is the flow factor:

$$\gamma = \frac{q}{J_l K_l} \qquad [1.5]$$

where K_l is the geometric mean K. In the case of a reactive contaminant, the relationship between ln$K$ and K_d can be expressed in the following model:

$$\rho_b K_d = b\ln K + a + \kappa \qquad [1.6]$$

where ρ_b is the bulk density of the soil, b is the slope of the linear regression curve of K_d unto ln$K$, a is the y-intercept and κ is the residual term. The longitudinal macrodispersivity, A_{11}, for the reactive contaminant is expressed as (see *Garabedian et al., 1988; Gelhar, 1993*):

$$A_{11} = \frac{\sigma_f^2\lambda}{\gamma^2}\left(1 - \frac{\gamma b}{Y}\right)^2 + \frac{1}{Y^2}(\sigma_\kappa^2\lambda_\kappa) \qquad [1.7]$$

where σ_κ^2 and λ_κ is the variance and correlation scale of the residual and Y is:

---

[1]The correlation scale is the distance over which the variation in properties of the soil is correlated

$$Y = n + \rho_b E[K_d] = nR \qquad R = 1 + \frac{\rho_b E[K_d]}{n} \qquad [1.8]$$

where R is the retardation factor.

1.3 SORPTION OF ORGANICS

Metals and organic contaminants are subject to different types of sorption. Organics experience partitioning where the compound is sorbed to the soil, preferably the soil organic matter, by molecular forces. This noncompetitive process is a function of the amount of the compound in the water and the organic content of the soil. Partitioning is represented by:

$$K_d = f_{oc} K_{oc} \qquad [1.9]$$

where K_{oc} is the organic carbon sorption coefficient and f_{oc} is the organic carbon content of the soil (*Karickhoff et al., 1979; Schwarzenbach and Westfall, 1981; Hassett et al., 1983*). Schwarzenbach and Westfall (*1981*) determined that when a soil has a *POC* value which is less than 0.1%, other processes become involved in the sorption process and the correlation between K_{oc} and f_{oc} proposed in [1.9] is no longer valid. The *POC* values for the soils at Cape Cod tend to be less than 0.1%. However, Barber (*1992*) found that even with these low *POC* values, partioning of tetrachlorobenzene and pentatchlorobenzene into organic carbon matter appears to be the primary sorption mechanism; therefore [1.9] is a suitable approximation for

the K_d values for these soils.

For this study the organic sorbents, benzene, pentachlorobenze (PentCB) and 1,2,4 trichlorobenzene (TCB), were choosen as hypothetical contaminants.

1.4 OBJECTIVES AND SCOPE

The stochastic theory outlining the effect of the variability of K and K_d and the correlation between these two variables on the transport of contaminants in the subsurface is fairly extensive. However; the field applications of these theory, especially in the area of organic sorption, has been limited.

The objective of this research is to:

i. Measure the variability of K and K_d for Cape Cod sediments. The hydraulic conductivity (K) and *POC* will be measured with a permeameter and CHN Elemental Analyzer respectively. The *POC* values will be converted to K_d values using [1.9].

ii. Evaluate the suitability of a linear regression model in the form of [1.6] to represent the relationship between K and K_d. Several tests will be performed in the covariance and spectral (or frequency) domain to evaluate the regression model.

iii. Quantify the enhanced macrodispersivity caused by the correlation between K and K_d . Using the output of the regression model and the statistics of K, K_d and the residual, the enhanced macrodispersivity will be evaluated.

CHAPTER 2 INVESTIGATIVE METHODS

2.1 SAMPLE COLLECTION

The samples for this project were collected from an outcrop at the Handy Cranberry Bog on Cape Cod, Massachusetts. The soils at this site ranged from medium to coarse grained sand. A layer of top soil was found above the outcrop. This clayey silt was approximately six inches deep at the sampling location.

A front-end loader was used to clear a vertical face until an area of undisturbed soil was accessed (see Figures 2.1, 2.2 and 2.3). Sampling was performed at 0.5 m intervals for the horizontal transect and at every 0.15 m for the vertical transect. The disturbed samples were collected with an 8 inch diameter cylindrical copper tube. The tube was forced into the sand and retrieved with the sample. The 300 g samples were placed in sealed bags and transported to the lab. Once at the lab, the samples were air dried for the permeameter test.

2.2 HYDRAULIC CONDUCTIVITY

2.2.1 Hydraulic Conductivity Experiments

The hydraulic conductivity of the sand samples was measured with a falling head permeameter. The permeameter consists of a 3-inch (3") diameter column connected to a 1" diameter transparent measuring tube with a rubber tube as shown in Figures 2.4 through 2.7. Hereafter, these tubes will be referred to as the

Scale 1" = 20 m

Figure 2.1 Photograph showing the face being cleared with a bulldozer to access an undisturbed exposure at the Handy Cranberry Bog

Scale 1" = 20 m

Figure 2.2 Photograph showing the exposure at the Handy Cranberry Bog.

Scale 1" = 1/4 m

Figure 2.3 Photograph showing the variability in the aquifer sediments at the Handy Cranberry Bog.

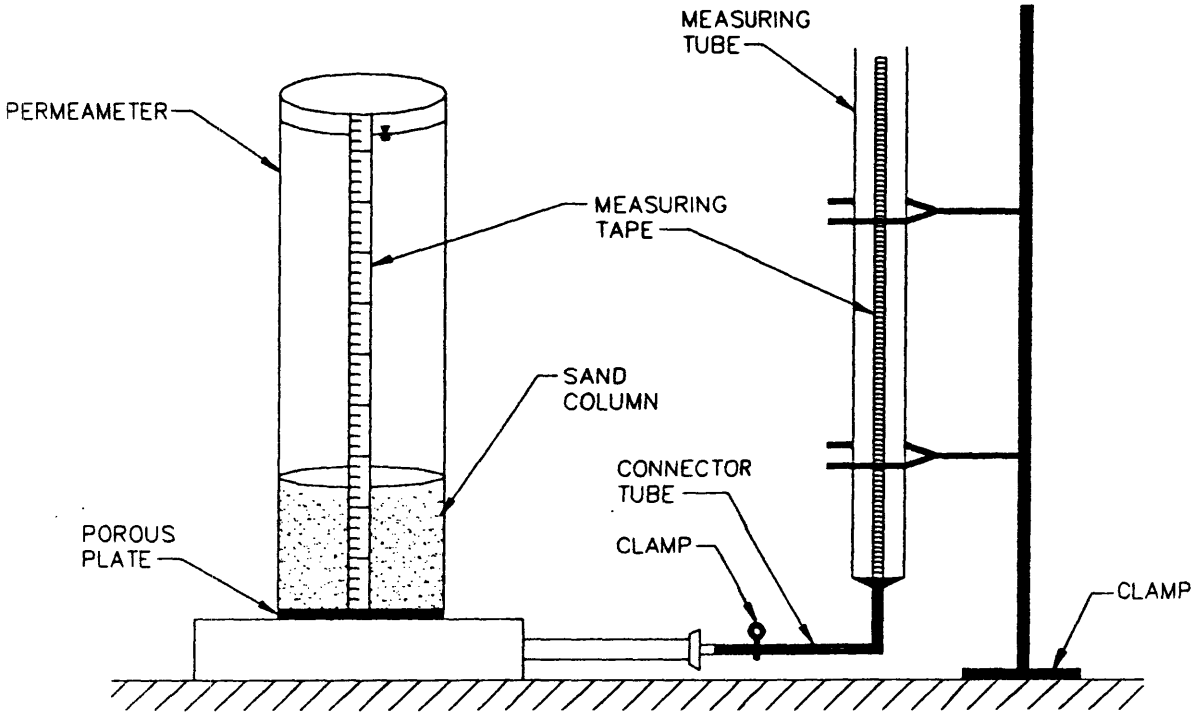

Figure 2.4 The setup of the falling head permeameter

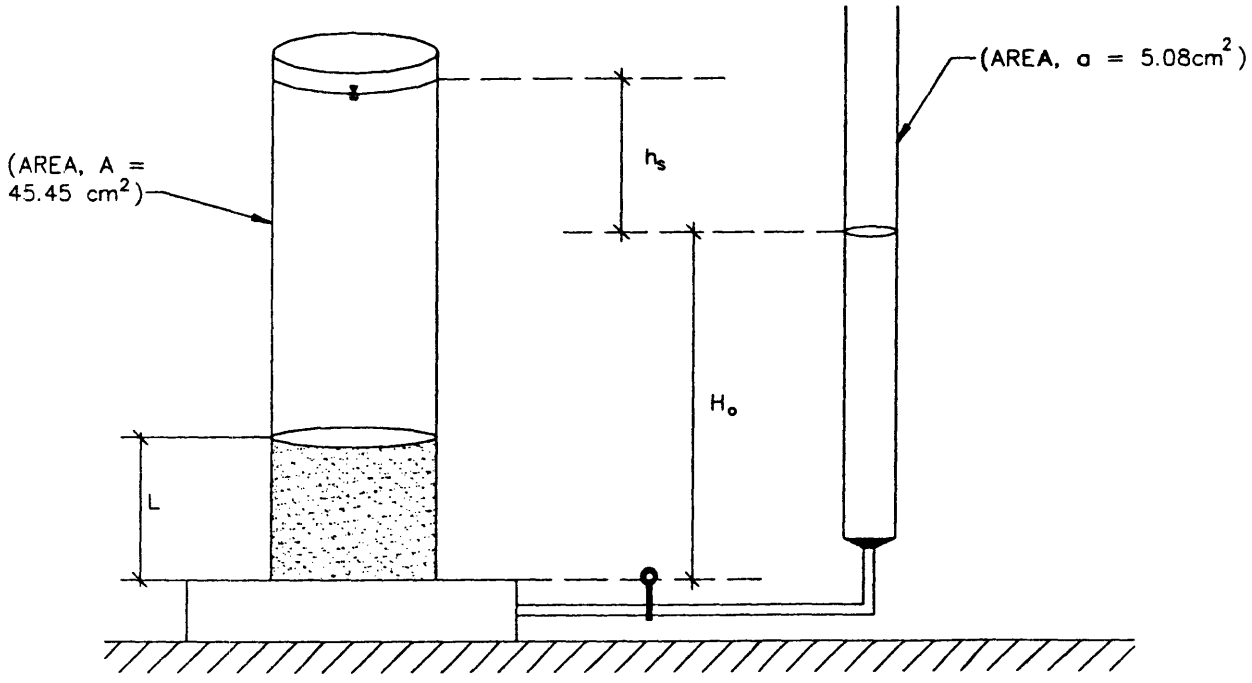

Figure 2.5 The permeameter before the test

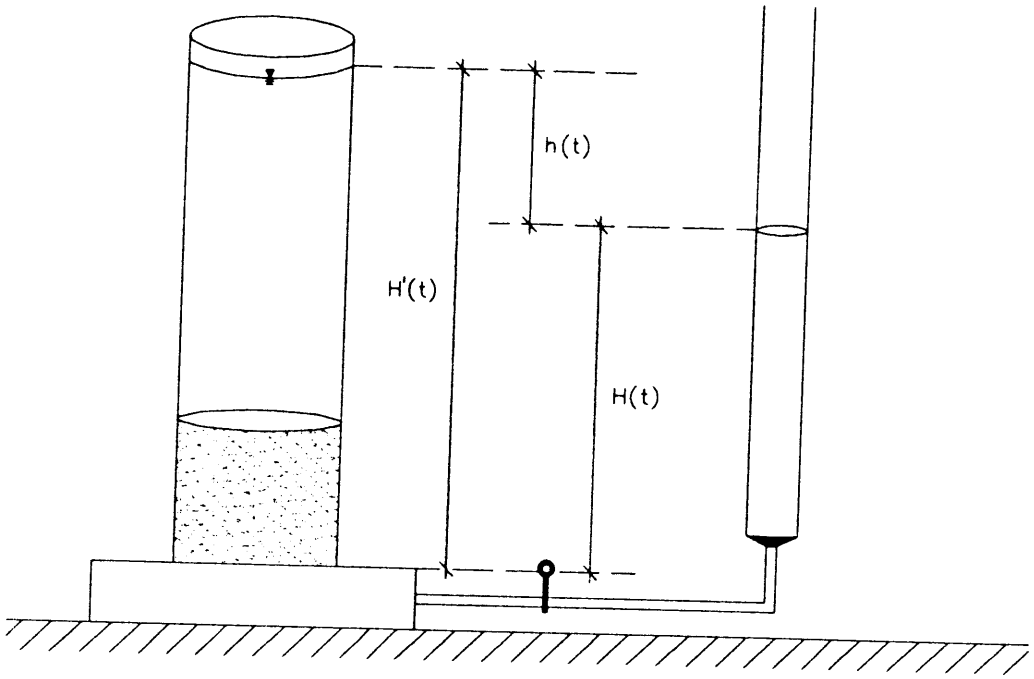

Figure 2.6 The permeameter during the test

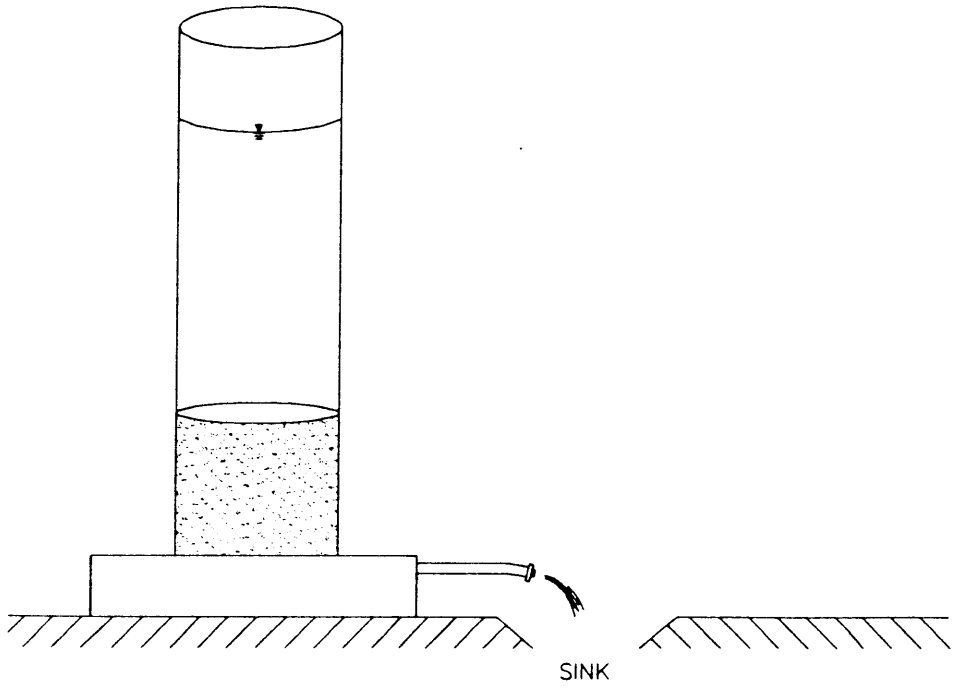

Figure 2.7 The permeameter after the test

permeameter, measuring tube and connector tube, respectively. The permeameter and measuring tube are scaled.

Prior to the permeameter test or run, the connector tube is clamped and the measuring tube is filled with water. The clamp is loosened and a volume of water equivalent to 3" of water is allowed to flow into the permeameter. The connector tube is clamped and gently tapped to remove entrapped air bubbles.

The sample is sieved using a 6.35 mm sieve. The large clasts and sand fraction is weighed, and the sand fraction placed in a plastic beaker. A target sample depth for the soil in the permeameter is calculated based on an average in-situ packing density of 5.90 cm/pound.

The sample is poured into the standing water using a funnel. This funnel is connected to a rubber tube that is long enough to span the distance from the top of the permeameter to the water surface. The sample is poured into the funnel slowly to minimize the amount of disturbance. While the sample is poured, the funnel is continuously rotated to ensure that the soil column does not develop a sloping surface. This process continues until the water level above the soil column is reduced to 1 cm. The clamp is loosened and a volume of water equivalent to 2 cm is allowed to flow into the permeameter at a rate of 5 cm/s. More soil is poured into the standing water and soil as described above. This process continues until the beaker is empty. The soil is then packed to the target depth. The clamp is then loosened and water is allowed to flow from the measuring tube to the permeameter until the water level in the permeameter is about 6 inches above the sample. The clamp is tightened and the permeameter is filled with water until the water level is approximately 8 cm below the top of the permeameter. This process is performed

at a slow rate in order to avoid disturbing the sample.

The water temperature, thickness of the sand layer, water level in the measuring tube and permeameter is measured. As a calibration method, the corresponding water levels in the permeameter and measuring tube is recorded using a carpenters level.

To start the run, the clamp is removed and a stopwatch is started. The time measurements are recorded each time the water level reaches certain marked intervals on the measuring tube.

After the test, the depth of the sample is measured. The setup is cleaned by disconnecting the rubber tubing and draining out the water. The sand and water is poured into a bucket and allowed to settle for an hour. The water is poured off and the sand is removed with a spatula and placed in an aluminum plate. The sample is air dried for approximately three days.

2.2.2 Calculation of K

In this study, the K measurements are needed for an open ground-water system. However, the permeameter is a closed, smaller-scale system with certain head losses. In using the permeameter results to represent an open system, two of these losses are:

1) h_{exp} , the loss due to the expansion in the flow area at the junction of the connector and measuring tube. This results from turbulence.
2) h_{app} , the loss due to the apparatus. This loss is caused by the resistance to the flow of water that the walls of the permeameter exhibit.

Both of these losses can be expressed in terms of q, the specific discharge. The calculated head loss (h_{loss}) is subtracted from the measured head (h) to find the absolute head h_a:

$$h_a = h - h_{loss} \quad [2.1]$$

The specific discharge is estimated by fitting a curve to the observation of $h(t)$ versus t. This curve can be represented by a polynomial:

$$h = \alpha + \beta t + \gamma t^2 + \delta t^3 + \ldots\ldots \quad [2.2]$$

Applying continuity and using $h = H' - H$, the flow is expressed as:

$$q = -\frac{d}{dt} H'(t) = -\frac{a}{A+a}\frac{dh}{dt} = \frac{a}{A}\frac{dH}{dt} \quad [2.3]$$

where:

$$\frac{dh}{dt} = \beta + 2\gamma t + 3\delta t^2 \quad [2.4]$$

The average water velocity in the connector tube (V_h) can be expressed in terms of q by applying continuity:

$$V_h = \frac{A}{a_h} q \quad [2.5]$$

where A, the area of the permeameter is 45.45 cm^2 and a_h the area of the connector tube is 0.13 cm^2. The expansion head loss, in terms of V_h is:

$$h_{exp} = K_c \frac{V_h^2}{2g} \quad [2.6]$$

K_c, the loss coefficient is proportional to the ratio of the cross-sectional flow areas

before and after expansion (*Fox and McDonald, 1985*). The area of the measuring tube is 5.08 cm², this ratio is 0.026. The corresponding K_c value is 0.97. Substituting these numerical values into [2.6]:

$$h_{exp} = \frac{K_c}{2g}\left(\frac{A}{a_h}\right)^2 q^2 = 60.425\,q^2 \qquad [2.7]$$

The apparatus head loss is modeled as a quadratic defined by equation 5.11.31 in Bear (*1972*):

$$J = \frac{h_{app}}{l_{app}} = \frac{\nu}{k_{app} g} q + \frac{1}{c g k_{app}^{1/2}} q^2 \qquad [2.8]$$

k_{app} and l_{app} is an equivalent permeability and length of the soil layer.

The total head loss is:

$$h_{loss} = \frac{\nu l_{app}}{k_{app} g} q + \left[h_{exp} + \frac{l_{app}}{c g k_{app}^{2}} \right] q^2 \qquad [2.9]$$

Dividing this equation through by q:

$$\frac{h_{loss}}{q} = \frac{\nu l_{app}}{k_{app} g} + \left[60.425 + \frac{l_{app}}{c g k_{app}^{2}} \right] q \qquad [2.10]$$

This equation will be expressed as the equation of a line:

$$y = mx + b \qquad [2.11]$$

where:

$$b = \frac{\nu l_{app}}{k_{app} g}; \qquad m = \frac{l_{app}}{c g k_{app}^2} + 60.425$$

The coefficients of [2.11] were determined using the results of ten permeameter runs performed without sand (*Thompson,* 1994). The average m and b are 10.75 s^2/cm^2 and 139.89 s^2/cm, respectively. The total head loss is:

$$h_{loss} = 10.75q + 200.32q^2 \qquad [2.12]$$

The permeameter experiment is based on the Darcy's equation:

$$q = K\frac{h_a}{l} \qquad [2.13]$$

which applies to most natural groundwater systems. This equation states that the specific discharge (q) is proportional to the hydraulic gradient (dh/dz) scaled by the K. For the permeameter, the hydraulic gradient has been replaced by $h_a(t)/l$ and the specific discharge was substituted with the velocity of the water, $v(t)$. The hydraulic conductivity is the slope of the graph of q versus h_a. Figures 2.8 and 2.9 shows the plots of q versus h_a for Samples 250 and Q.

The above analysis is based on a linear flow assumption. This assumption is valid if the Reynolds number, R_k, as defined in the following equation is greater than 1:

$$R_k = \frac{q\sqrt{k}}{\nu} \qquad [2.14]$$

where ν is the kinematic viscosity and k is the intrinsic permeability. For the highest flow calculated by equation [2.3], the Reynolds number is less than 1 and therefore the linear flow assumption is appropriate.

2.2.3 Reproducibility of Data

Three sets of duplicate soil samples were run. The purpose of these

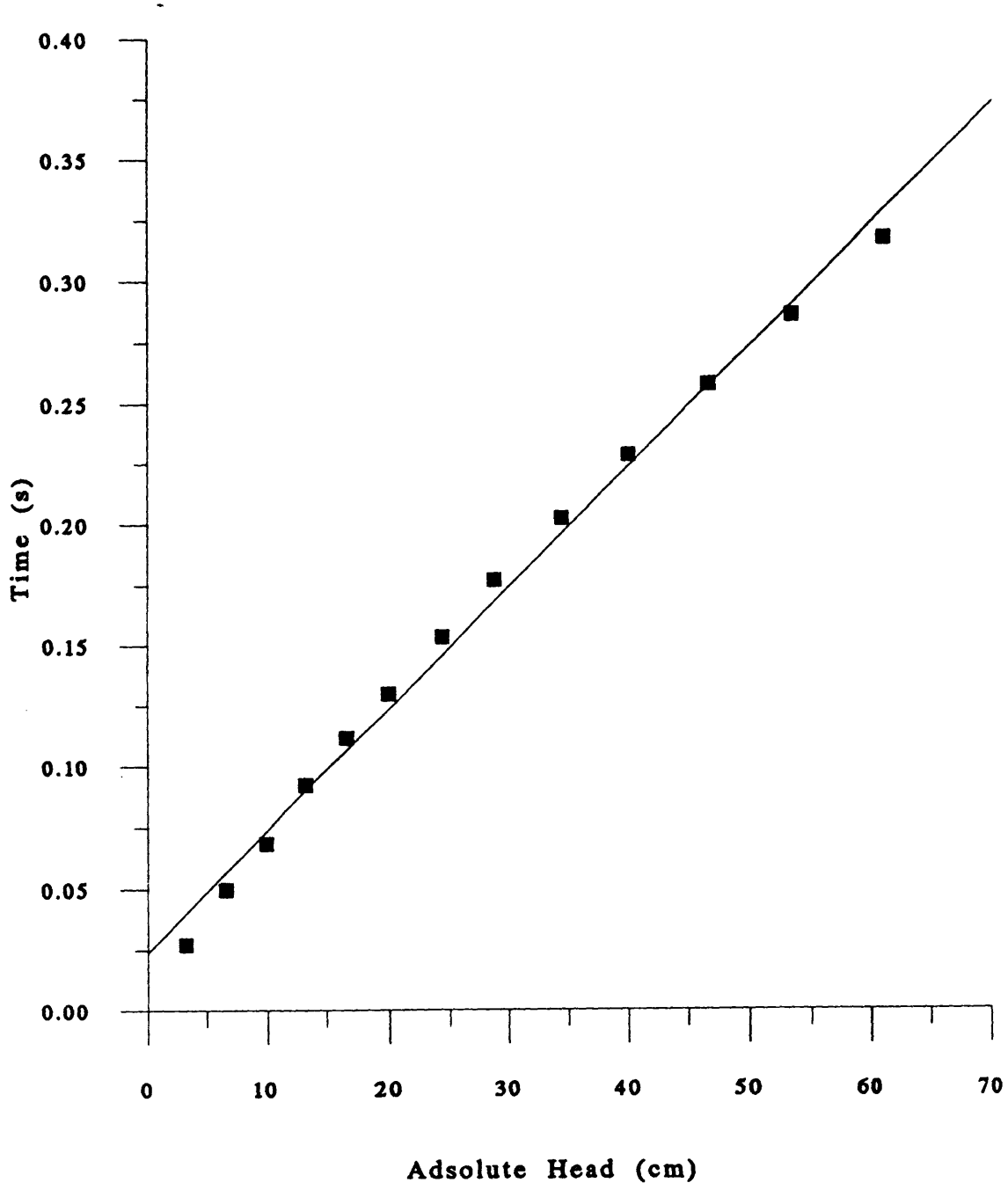

Figure 2.8 Flow as a function of the calculated head (h_a) for sample 250. The calculated head values are the measured head value less the head loss.

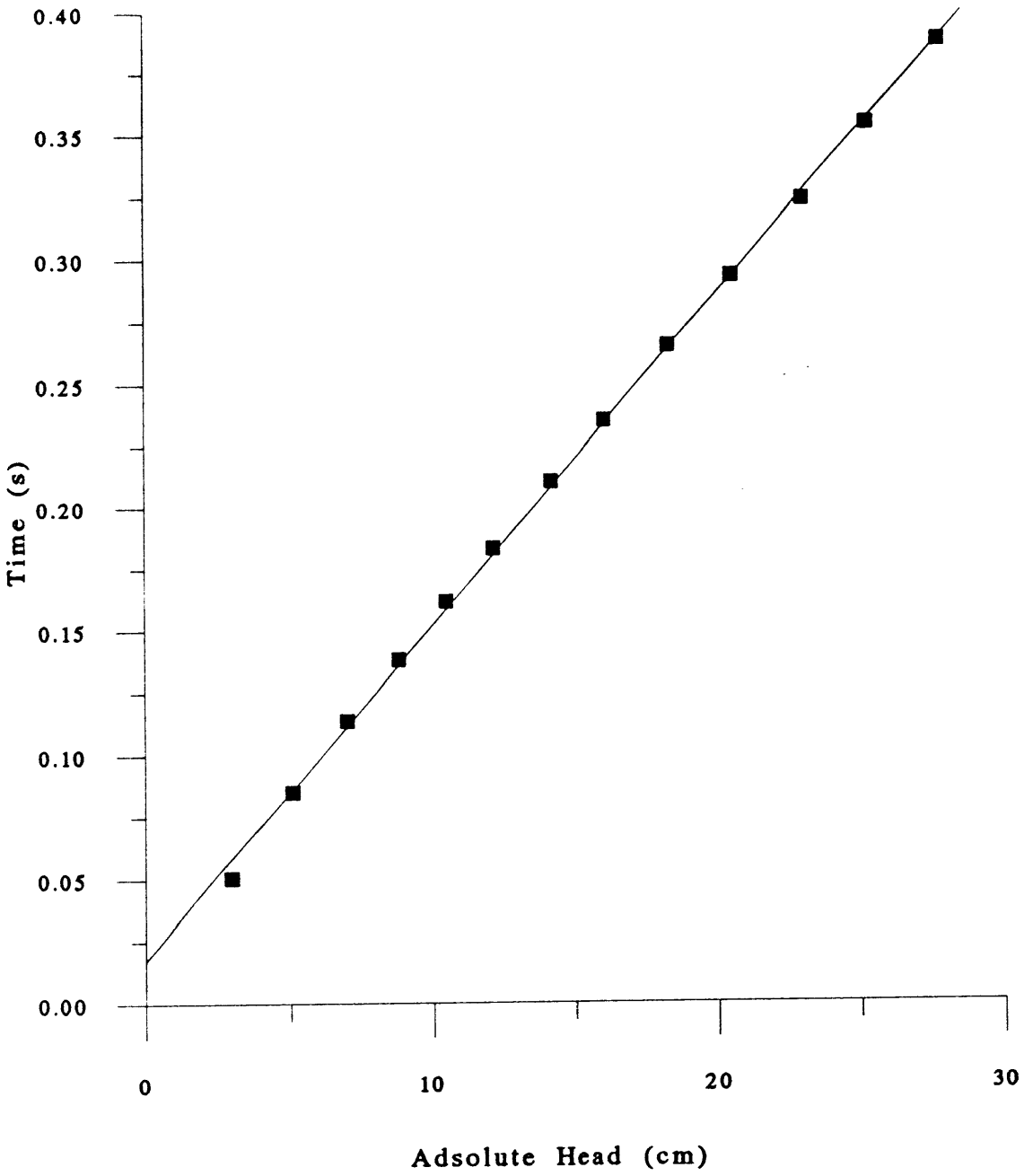

Figure 2.9 Flow as a function of the calculated head (h_a) for sample Q. The calculated head values are the measured head value less the head loss.

duplicate runs is to quantify two effects:

1) the reproducibility of the hydraulic conductivity measurements, and

2) the effect of the soil compaction on the hydraulic conductivity measurements.

Table 2.1 Reproducibility of the permeameter results

Sample ID	Depths of Sample in the Permeameter (cm)	Ln*K* (*K* in cm/s)
235	5.6	-2.62
Duplicate Sample	5.6	-2.62
Percent Difference	0.0%	0.0%
242	6.85	-2.86
Duplicate Sample	7.0	-3.07
Percent Difference	2.2%	7.4%
247	7.8	-2.55
Duplicate Sample	9.4	-3.26
Percent Difference	18.6%	24.4%

The following conclusions can be made from the above table:

1) The permeameter results are reproducible in instances where the depth of the soil column is consistent.

2) A relative increase in the sample depth above the target depth, produces roughly twice the relative decrease in the Ln*K* value.

2.3 PERCENT ORGANIC CARBON TESTING

2.3.1 Determination of a Representative Sample

Each samples obtained from the field site weighed about 300 g. Up

to 300 mg of soil was needed for the percent organic carbon analysis. It became necessary to develop a systematic method of reducing the bulk sample to a representative subsample. Several subsampling methodologies were investigated, including thief sampling, coning-and-quartering and riffling. Riffling resulted in the most representative subsamples. In 1986, the American Society of Testing and Methods (ASTM) subdivided samples by coning-and-quartering, thief sampling and riffling. The standard deviation of consecutive sixteenths of a sample which had gone through the riffling process was substantially lower than those that had been coned-and-quartered or thief sample.

Riffling is performed with a riffle-box and three receiver bins. A riffle-box consist of a metal box with sloping sides whose bottom incorporates two groups of parallel chutes of equal width. These series of chutes alternate between 60° and 300° from the horizontal. A drawing of three common types of riffle-boxes are shown in Figure 2.9.

During the riffling process, two of the receptor bins (receptors) are placed under each set of chutes. The sample is poured into the remaining bin. Care is taken to spread the sample evenly in the bin. The contents of the bin is poured into the riffle-box in an unbiased fashion. The sample is collected by the two receptors positioned under the chutes.

Riffling was performed on three samples, this resulted in subsamples of comparable masses (see Table 2.2)

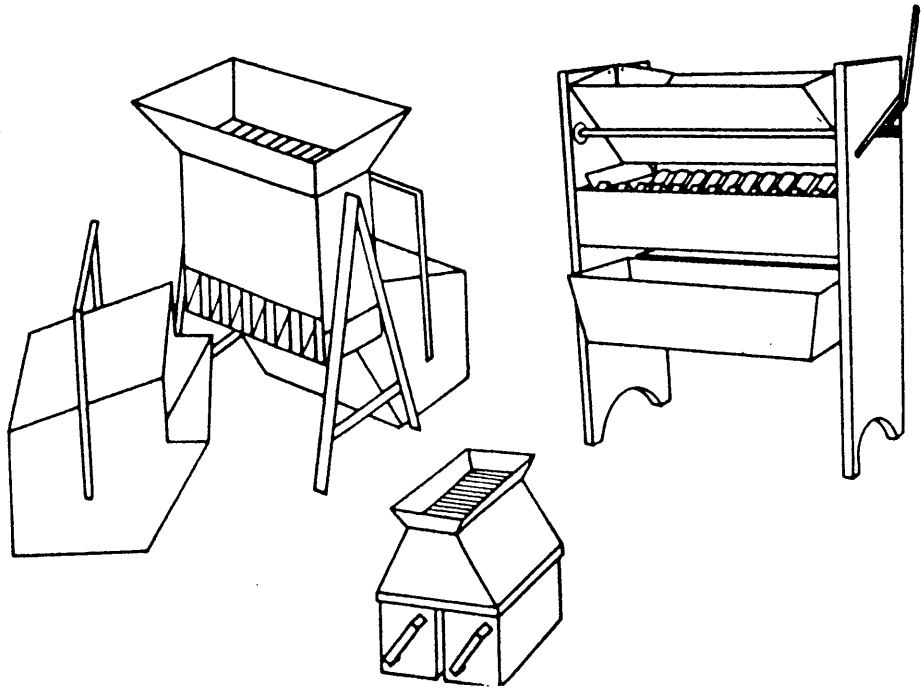

Figure 2.10 Common forms of riffle boxes (from ASTM, *1986*)

Table 2.2 Mass of subsamples divided by the riffling process

Sample ID	Original Mass (1/64th lb)	Mass after first division (1/64th lb)	Mass after 2nd division (1/64th lb)	
A	58	29	13.5	14.5
		29	15.5	13.5
B	63	31.5	16.5	14.5
		31.5	16	14.5
C	78	37.5	17	19.5
		39.5	19.5	18.5

Equal subdivision of mass does not necessarily correspond to identical subsamples (*ASTM*, 1991). The success of the riffling process is also based on the reproducibility of the results. The spread of the percent organic carbon (*POC*) values, shown in Table 2.3, also indicates that the riffling operation is effective.

About two hundred samples were riffled; it became necessary to optimize the riffling process. In designing the most successful riffling routine, two constraints were considered:

1) The amount of time spent subsampling - for the volume of sample collected and the desired subsample volume, the amount of time spent riffling becomes an issue of concern.

2) The ability of the box to handle small samples - the riffle box used is large relative to the size of the subsample required.

Two methods were investigated. In method 1, the sample was homogenized and divided into eights using the riffle box. One eight was riffled into four quarters. Each quarter was homogenized. Coning-and-quartering was used to extract about 3 mg of soil that was ground to a fine powder for *POC* analysis. Each of the subsamples were placed in glass vials.

In method 2, the sample was divided into 1/32 portions by riffling. Four of these subsamples were selected at random and the entire portion ground into a powder. The sample was homogenized and containerized in the glass vials.

The spread of the *POC* measurements for subsamples divided by these methods are shown below:

Table 2.3 Percent organic carbon values for selected subsamples

Sample ID (Method)	POC	Average POC	Standard Deviation	Deviation from the mean
D (Method 1)	0.029 0.018 0.016 0.012	0.019	0.0073	52.63% 5.26 15.79 36.84
E (Method 2)	0.014 0.013 0.012 0.009	0.012	0.0022	16.67% 8.33 0.00 25.00
F (Method 2)	0.019 0.020 0.022 0.022	0.021	0.0025	9.52% 4.76 4.76 4.76
G (Method 2)	0.020 0.018	0.019	0.0014	5.26% 5.26%

It is obvious that the second method yields better results.

2.3.2 Percent Organic Carbon Testing

The *POC* is measured with a Perkin Elmer Carbon Hydrogen Nitrogen Elemental Analyzer (CHN analyzer) oven. This oven determines the weight percent of carbon, hydrogen and nitrogen in a sample. The operation of this oven is defined in this section.

Operation of the Instrument

A schematic of the CHN analyzer is shown in Figure 2.10. The machine operates on a combustion method. There are four major zones: the pyrolysis zone, gas control zone, separation zone and detection zone. The sample is transported throughout the system by the carrier gas, helium.

At the start of each run, a known weight of sand surrounded by a tin capsule is inserted into the pyrolysis zone. In this high temperature environment, with the aid of platinized carbon reagent, this sample is converted to carbon monoxide rich gas. This gas is then passed through the scrubbing agents copper and sodium hydroxide to remove sulfur products and acid gases, respectively. The carbon monoxide mixture is transported to the gas control zone.

In the gas control zone, the gas is homogenized to ensure that the results are representative of the entire sample. This process occurs in the mixing chamber.

In the separation zone, the mixing chamber is depressurized. Frontal, not standard, chromatography is applied here. By selective retention, this method employs the steady-state stepwise signal rather than a peak signal.

In the detection zone, the carbon monoxide is passed through a thermal conductivity detector system and the weight percent of carbon, hydrogen and nitrogen determined.

These processes takes about seven minutes

Calibrating the Instrument

Before running a set of samples are run the machine must be calibrated as follows. The oven is purged with a helium flow of 15 pound per second (psi) for 200

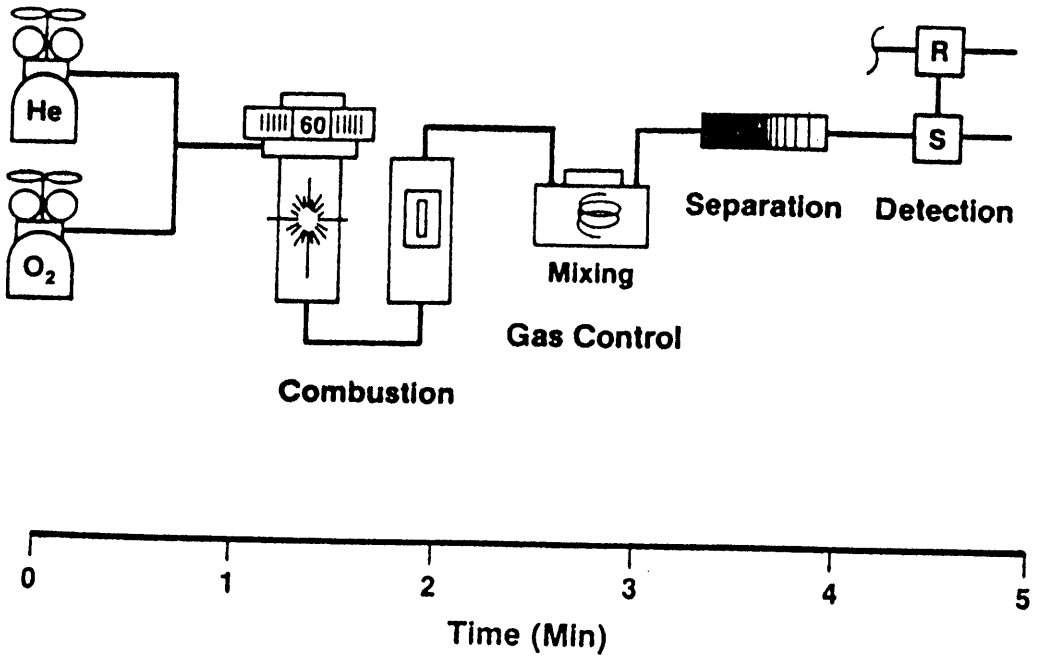

Figure 2.11 Schematic diagram of the CHN Elemental Analyzer (from Culmo, *1993*)

second duration and with an oxygen flow of 28 psi for 20 seconds. The sensors are monitored to determine the combustion temperature, reduction temperature, detector oven temperature, pressure, detector counts and signal timing.

The K-factor is a standard, acetanilide, whose elemental make up is known by the machine. A series of one K-factor, one blank, one K-factor, and one blank is run to condition the machine. Three K-factors are run to check for the reproducibility of the results. A known portion of acetanilide is run as an unknown and the result from this run is compared to the theoretical values. The K-factors and blanks are deemed acceptable if they are within the following range:

	Blanks	K-factors	True Acetanilide ± 0.3% range
C	± 30	± 0.15	71.09 (70.79 - 71.39)
H	± 100	± 3.75	6.71 (6.41 - 7.01)
N	± 16	± 0.16	10.36 (10.06 - 10.66)

Once the calibration is successful, the samples are run. If it is not successful, the series of blanks and K-factors are run as described above.

Running the Sample

At the CHN analyzer station, the tin cups are weighed on a microbalance scale. Approximately 200 mg of the sample is placed in a tin cup. The sample and cup is weighed and the weight of the sample is determined by elimination. The tin cup is folded carefully around the sample so that all of the ends are rounded; any sharp corners may get trapped in the mechanism of the CHN analyzer.

The samples are run in a sequence of 5-6 samples, 1 blank, 5-6 samples .....

Shutdown

The system is purged with helium for 20 seconds. The regulator is set to 5 psi and the oxygen supply is shut off.

2.3.3 Determination of Measurement Noise

Measurement noise is always an issue of concern. For the initial run of the machine, the *POC* values ranged from 0.009 to 0.028. For such a range, it is important to know how much measurement noise is present. The following table shows the result of ten runs on two subsamples, Samples A and B.

Table 2.4 Determination of measurement noise for the CHN Analyzer

Sample ID	POC	Mean	Percent deviation from the mean	Standard Deviation (Variance)
Sample A:	0.022		4.76%	
	0.019		9.52	
	0.019		9.52	
	0.020		4.76	
	0.020	0.021	4.76	0.002
	0.021		0.00	(4.0×10^{-6})
	0.023		9.52	
	0.024		14.29	
	0.024		14.29	
Sample B:	0.019		14.29%	
	0.020		9.52	
	0.021		4.76	
	0.021	0.022	4.76	0.001732
	0.022		0.00	(3.0×10^{-6})
	0.022		0.00	
	0.022		0.00	
	0.023		4.76	
	0.025		14.29	

The variance of the *POC* measurements for all the samples tested is 9.5×10^{-4}. The variance of *POC* for the individual samples is much less than the variance of all

the samples. Therefore, the measurement noise is significant but it is not controlling. If the measurement noise is independent of the *POC* values only, the variances would be additive and the variance of the measured *POC* would be equal to the sum of the actual variance of *POC* and the variance of the measurement noise. The variance of the noise is about 1 percent of the measured variance of the entire population and therefore it is reasonable to neglect the effect of measurement error in the subsequent statistical analysis.

The distributions of *POC* for samples A and B are shown in Figures 2.11 and 2.12, these plots closely resemble a normal distribution.

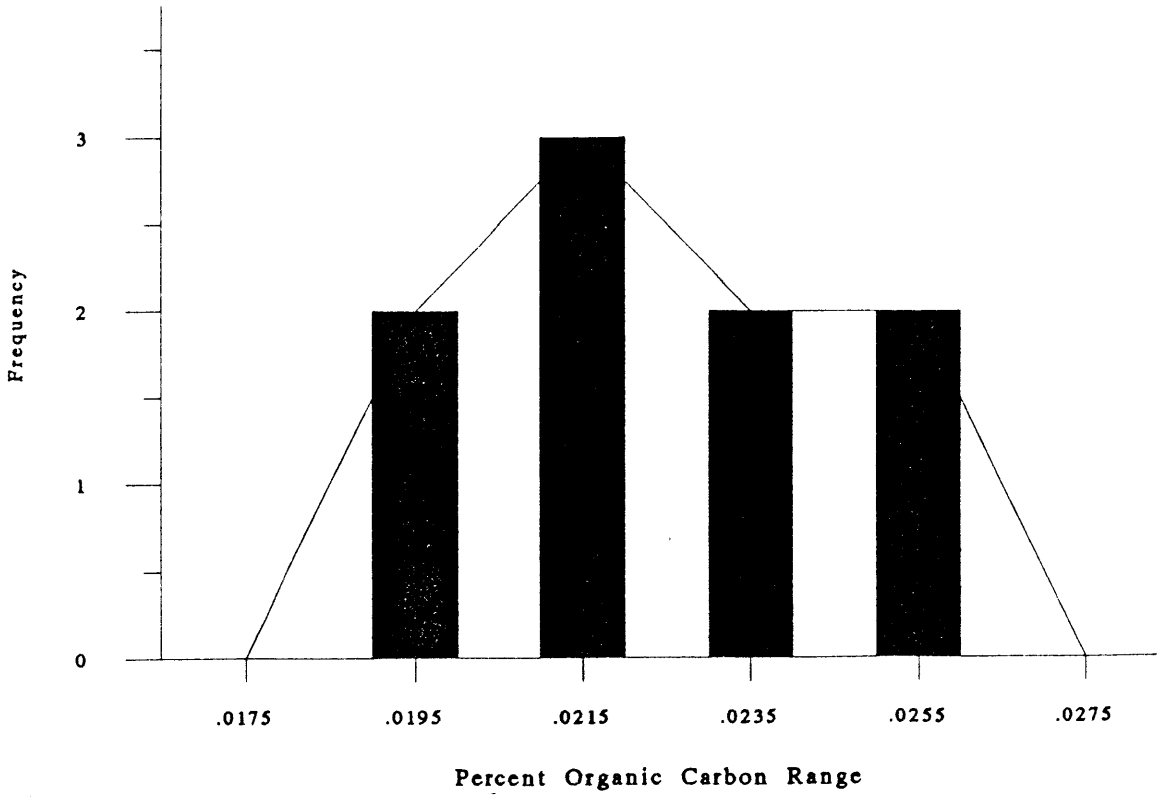

Figure 2.12 Distribution of percent organic carbon measurements for sample A.

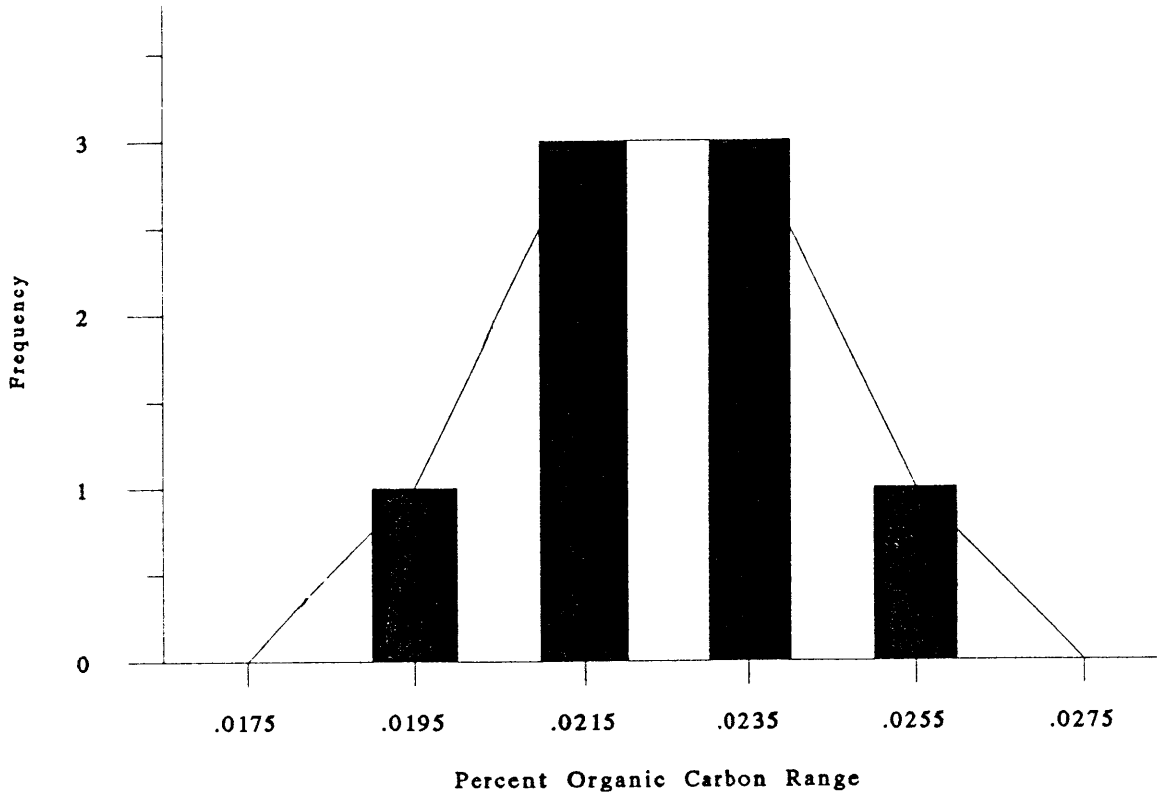

Figure 2.13 Distribution of percent organic carbon measurements for sample B.

CHAPTER 3 RESULTS AND DISCUSSION

3.1 STOCHASTIC PROCESSES AND SPATIAL SERIES[1]

"..... atmospheric variation, and variation in geologic processes, creates earth materials that have highly variable hydraulic (and chemical) properties" (*Gelhar, 1993*). This variability of aquifer properties can best be demonstrated by taking a number of measurements of physical and chemical processes over space or time. Most times, these measurements will be highly variable. Each of these processes can be defined as a stochastic or random process, since it varies significantly with space or time. A plot of this variability over space or time is called a realization. In this study, we will investigate the spatial heterogeneity of these physical and chemical processes for an aquifer material.

The cumulative probability distribution function (cdf) is used to describe a random variable, X. The cdf can be represented by the following equation:

$$F_X(x)=P[X\leq x] \qquad [3.1]$$

which represents the probability that the random variable X will be less than or equal to x. The value of the cdf ranges from 0 to 1.

The probability density function (pdf):

$$f_X(x) = \frac{dF}{dx} \qquad [3.2]$$

[1] The nomenclature used in this section is adapted from Gelhar, 1993; capital letters denote random quantities and lowercase letters denote deterministic numerical values.

The pdf can be expressed as an integral:

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\,du \qquad [3.3]$$

The total area under the pdf is unity. An example of a cdf and pdf is shown in Figure 3.1.

The moments of a random variable can be evaluated from the pdf. The mean (μ_x), or expected value, of a random variable is defined as:

$$\mu_X = E[X] = \int_{-\infty}^{\infty} u f_x(u)\,du \qquad [3.4]$$

This first moment is an indicator of the central tendency of the random variable. The mean of a random variable can be estimated from n samples of the variable as:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad [3.5]$$
$$E[\bar{X}] = \mu$$

The second moment around the mean is the variance (σ_x^2):

$$\sigma_X^2 = E[X-\mu_x^2] = \int_{-\infty}^{\infty} (X-\mu_X)^2 f(x)\,dx \qquad [3.6]$$

which for n samples is estimated from:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (X_i-\bar{X})^2 \qquad [3.7]$$
$$E[S^2] = \sigma^2$$

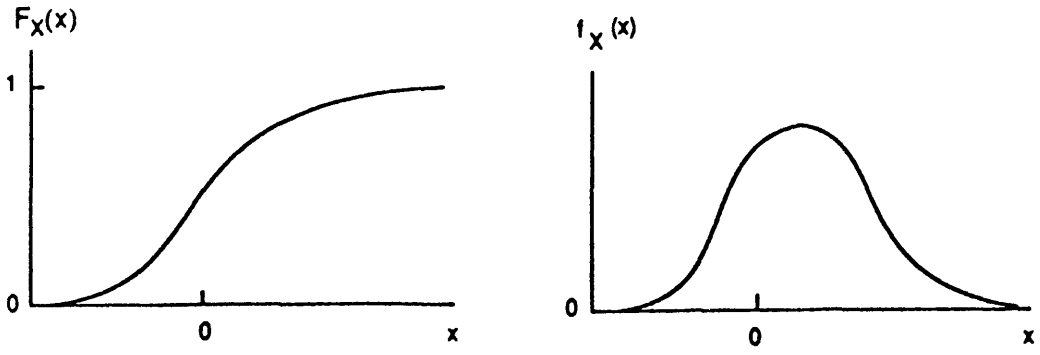

Figure 3.1 Cumulative probability distribution function and probability density function (from Gelhar, 1993)

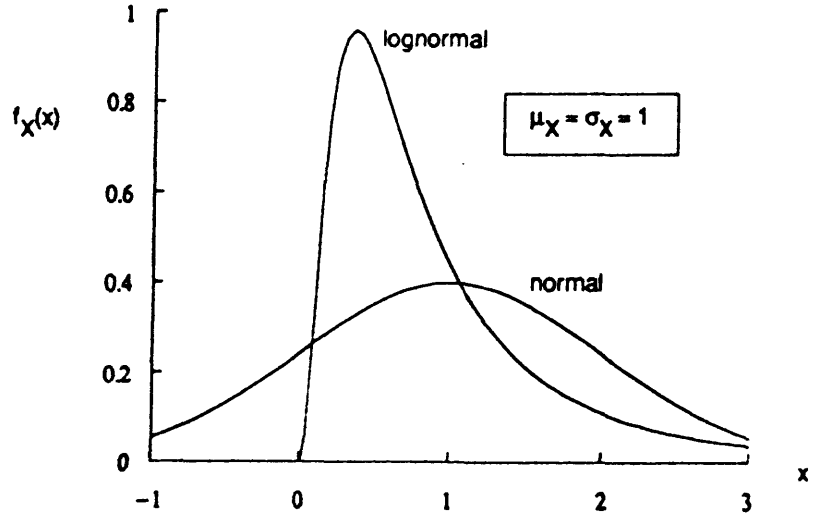

Figure 3.2 Probability density function for a normal and lognormal distribution (from Gelhar, 1993)

Both the mean and the variance of a stochastic process may change with time. The uncertainty of the estimated mean and the variance in [3.5] and [3.7] can be quantified by calculating the variance of the mean and the variance of the variance for a normal random variable, as:

$$var[\bar{X}] = E[(\bar{X}-\mu)^2] = \sigma^2/n = \sigma_X^2$$
$$var[S^2] = E[(S^2-\sigma^2)^2] = 2\sigma^4/(n-1) = \sigma_{s^2}^2 \qquad [3.8]$$

In hydrology, normal or lognormal random variables are commonly used. The normal random variable, Y, is represented by:

$$f_Y(y) = \frac{\exp[-(y-\mu)^2/2\sigma^2]}{(2\pi\sigma^2)^{1/2}} \qquad [3.9]$$

The normal, or Gaussian, distribution is completely characterisized by its mean and variance (*Gelhar, 1993*). This distribution sometimes applies to subsurface storage properties like porosity or moisture content. The pdf of the normal distribution is a bell shaped curve which may contain negative values.

A log normal distribution is one for which $Y = \ln X$ is a normal random variable. "The log normal distribution can be physically associated with phenomena that involve multiplicative effects on a large number of independent random events", (*Gelhar, 1993*). Transmissivity and hydraulic conductivity tend to have log normal distributions. For the log normal distribution, the pdf is clearly skewed with a peak smaller than the mean. The pdf for a normal and log normal distribution is shown in Figure 3.2.

The adequacy of a hypothesized probability distribution model may be evaluated using the Kolmogrov-Smirnov test. In this test, the cumulative probability

distibution of the data set is compared to a hypothesized cdf, $F_o(x)$. The test uses the difference in the maximum values of both cdfs:

$$D = \max|\hat{F}(x) - F_o(x)| \qquad [3.10]$$

This difference is compared to tabulated values for varying sample sizes and a significance level is determined.

The covariance function and autocorrelation function represent the relationship between adjacent points in the record. For a continious stationary stochastic process, $X(t)$, the covariance function, $R(\tau)$:

$$R(\tau) = E[(X(t+\tau) - E[X])(X(t) - E[X])] \qquad [3.11]$$

is dependent only on the lag (τ). The relationship between the covariance function and the autocorrelation function, $\rho(\tau)$, is:

$$\rho(\tau) = \frac{R(\tau)}{\sigma^2} = \frac{E[(X(t+\tau) - E[X])(X(t) - E[X])]}{\sigma^2} \qquad [3.12]$$

For a discretely sampled finite length realization, the autocorrelation function can be estimated using:

$$\hat{\rho}(r\Delta) = \frac{1}{(N-r)\sigma^2}\sum (X_{n=r} - \bar{X})(X_n - \bar{X}) \qquad r = 0,1,.....N-1 \qquad [3.13]$$

where N is the number of data measurements, Δ is the sampling interval, and N-r is the number of enteries in the sum. For the data points that are closer together, the data should be more correlated. Therefore, the autocorrelation function should have a larger value. This function has a value of 1 at zero lag and decreases as the lag increases.

The correlation scale (λ), is the lag corresponding to the autocorrelation function at a value of $1/e$. The correlation scale can be determined in the horizontal direction (λ_1) and the vertical direction (λ_3). In this study, we will assume that the subsurface is composed of a number of lenses with a lense length equivalent to λ_1 and a height or bed thickness equivalent to λ_3.

Realizations tend to exhibit a degree of regularity and periodicity, they can be thought of as a combination of sine and cosine waves, see Figure 3.3. The rate at which a series oscillates is represented by the angular frequency or wave number (ω), measured in radians per unit time (or length). Many sources use the frequency, ν:

$$\nu = \frac{\omega}{2\pi} \qquad [3.14]$$

measured in cycles per unit of time or length. Figure 3.4 illustrates a series of oscillations of varying frequencies and cycles. A cycle is a sine or cosine wave defined over 2π. Though visual inspection may give an indication of the number of oscillations and their contributions, there is a need for quantitative evaluation to deduce the statistical implication of these oscillations. The statistical significance of the oscillations is best gathered from spectral analysis, where the data is translated to the frequency domain with the aid of Fourier transforms.

For a stationary process, a Fourier representation can be used to represent the zero-mean[2] stochastic process. This spectral representation expressed as a function of ω is:

[2]To convert a process, X, to a zero-mean process, the mean value of X is subtracted from X

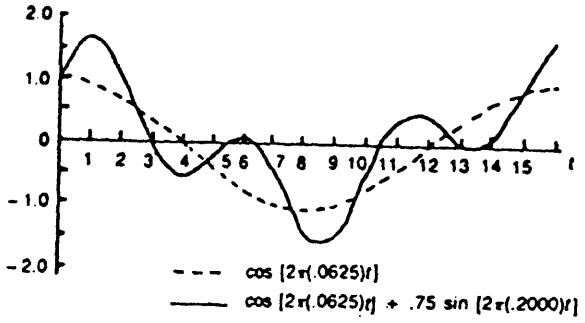

Figure 3.3 A linear combination of two periodic functions (from Shumway, 1988)

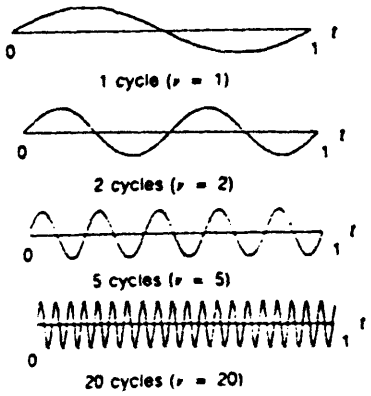

Figure 3.4 Cyclical variation using sine waves for various frequencies (from Shumway). The function plotted is $x_t = \sin(2\pi\nu t)$.

$$X(t) = \int_{-\infty}^{\infty} e^{i\omega t} dZ(\omega) \qquad [3.15]$$

For this Fourier-Stieltjes integral, Z is a stochastic process with the following properties:

$$\begin{aligned} E[dZ(\omega)] &= 0 \\ E[dZ(\omega_1) dZ^*(\omega_2)] &= 0 \end{aligned} \qquad [3.16]$$

The first condition in equation [3.16] indicates that the process is zero-mean. The second condition shows that the increments of Z [3] at varying frequencies are uncorrelated. In the case that $\omega_1 = \omega_2 = \omega$:

$$E[dZ(\omega) dZ^*(\omega)] = d\Phi(\omega) = S(\omega) d\omega \qquad [3.17]$$

where $\Phi(\omega)$ is the integrated spectrum, and $S(\omega)$ is the spectral density function or the spectrum. In [3.15], $X(t)$ is real and $e^{i\omega t}$ is complex, thus Z must be complex also. This can be thought of as a Fourier expansion in which dZ is a random amplitude. Therefore, the spectrum is the amplitude squared per frequency increment (*Gelhar, 1993*). Readable forms of the proof of the spectral representation theorem is provided in Priestly (*1981*).

The spectrum of a discrete process can be estimated using:

$$\hat{S}(\omega) = \frac{\Delta}{2\pi} \sum_{r=(-N-1)}^{N-1} \lambda(r) \hat{R}(r\Delta) e^{-i\omega r\Delta}; \qquad \hat{R}(r\Delta) = \hat{R}(-r\Delta) \qquad [3.18]$$

When $\lambda(r)$ is equal to one, this direct estimate of the spectrum is known as a pediogram. Though the spectral estimates provided by a pediogram are unbiased,

---

[3] In the text, the asterick is used to denote the complex conjugate

they tend to be erratic (*Gelhar, 1993*). The lag window ($\lambda(r)$) is used to produce consistent, but potentially biased, spectral estimates. Further discussion of the varying kinds of lag windows can be found in Priestly (*1981*). For all lag windows:

$$\lambda(r) = 0 \qquad for\,|r| \geq M < N-1 \qquad [3.19]$$

where M is an integer. This will lead to M spectral estimates at frequencies:

$$\omega_m = \frac{\pi m}{M\Delta}; \qquad m = 0,1,.....,M \qquad [3.20]$$

The lag window smooths the spectrum, possibly altering sharp peaks and depressions.

The spectrum and the covariance function shows similar information. The spectrum can be thought of as the variance of the spatial series, $X(t)$, in a frequency interval of width $d\omega$ centered at ω (*Shumway, 1988*). The covariance function also expresses the variance as a function of τ. The relationship between the spectrum and the covariance function can be expressed using Fourier transforms. The covariance function is the inverse Fourier transform of the spectrum:

$$R(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} S(\omega) d\omega \qquad [3.21]$$

thus:

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau} R(\tau) d\tau \qquad [3.22]$$

Detailed development of this equation can be found in Gelhar (*1993*). Four covariance-spectrum pairs are presented in Gelhar (*1993*), see Figure 3.5. For this analysis, the exponential and hole type covariance-spectrum pairs will be considered. The covariance function and spectrum for the exponential covariance function is:

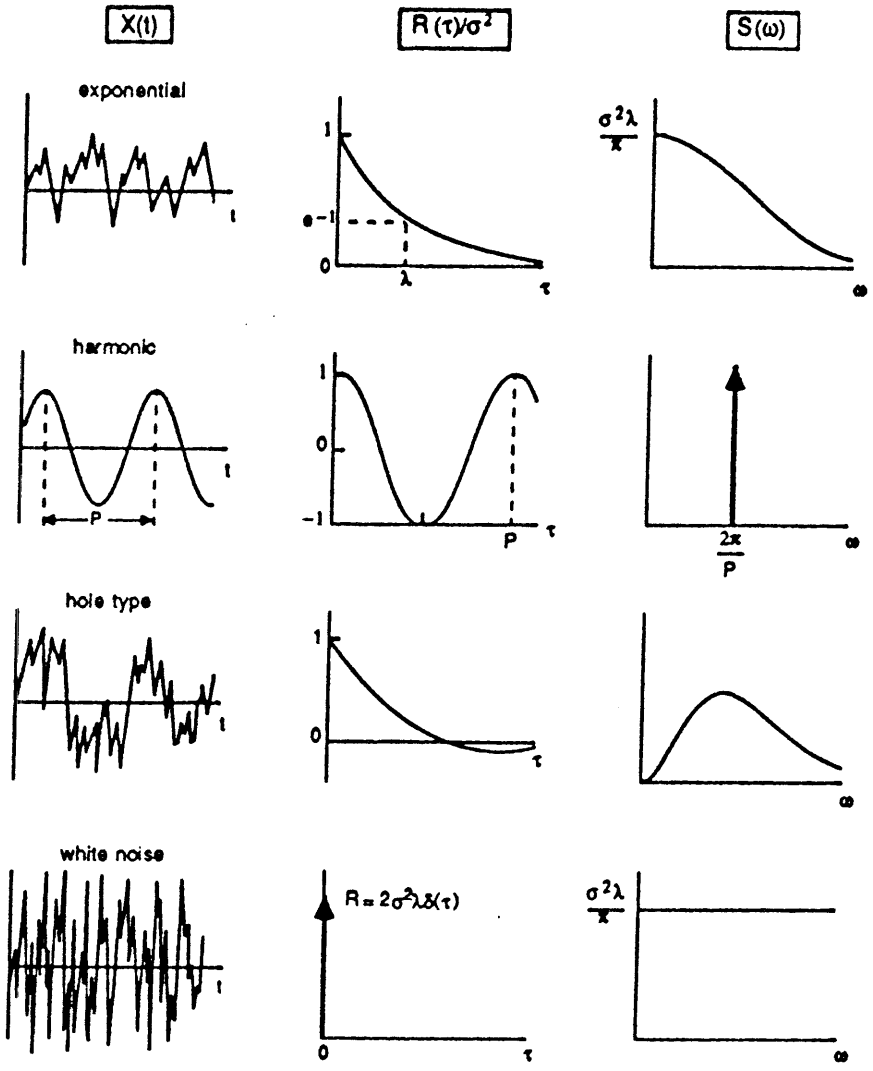

Figure 3.5 Schematic examples of covariance-spectrum pair (from Gelhar, 1993)

$$R(\tau) = \sigma^2 e^{-|\tau|/\lambda}$$
$$S(\omega) = \frac{\sigma^2 \lambda}{\pi(1+\lambda^2\omega^2)} \qquad [3.23]$$

The covariance function and spectrum for the hole type covariance function is:

$$R(\tau) = \sigma^2(1-|\tau|/l)e^{-|\tau|/l}$$
$$S(\omega) = \frac{2\sigma^2 l^3 \omega^2}{\pi(1+\lambda^2\omega^2)^2} \qquad [3.24]$$

where l is a correlation length estimated as 2.31λ. The fourth covariance-spectrum pair is white noise. This occurs when adjacent measurements are uncorrelated and as a consequence λ is equivalent to zero. The spectrum for white noise has a constant value over the frequency range.

Consider two zero-mean random processes:

$$X(t) = \int_{-\infty}^{\infty} e^{i\omega t} dZ_x(\omega) \qquad [3.25]$$

$$Y(t) = \int_{-\infty}^{\infty} e^{-i\omega t} dZ_y(\omega) \qquad [3.26]$$

The relationship between these two processes, as a function of τ, can be expressed as the cross-covariance function (R_{xy}):

$$R_{xy}(\tau) = cov[X(t+\tau), Y(t)] = E[X(t+\tau)Y(t)] \qquad [3.27]$$

or in a normalized form, by the cross correlation function:

$$\rho_{xy}(\tau) = \frac{R_{xy}(\tau)}{\sigma_x \sigma_y} \qquad [3.28]$$

where $\sigma_X^2 = E[X^2]$ and $\sigma_Y^2 = E[Y^2]$. The cross covariance function can be written in terms of the cross-spectral density function or cross-spectrum (S_{xy}) as:

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} S_{xy}(\omega) d\omega \qquad [3.29]$$

therefore:

$$S_{xy}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega\tau} R_{xy}(\tau) d\tau \qquad [3.30]$$

The cross-spectrum tells how two processes are related within various frequencies. This complex function is composed of a real and imaginary part referred to as the cospectrum (C_{xy}) and the quadrature (Q_{xy}) such that:

$$S_{xy}(\omega) = C_{xy}(\omega) - iQ_{xy}(\omega) = A_{xy}(\omega) e^{-i\theta_{xy}(\omega)} \qquad [3.31]$$

The phase spectrum (θ_{xy}) is:

$$\theta_{xy} = \tan^{-1}(Q_{xy}/C_{xy}) \qquad [3.32]$$

The amplitude of the cross-spectrum is:

$$A_{xy} = \sqrt{(C_{xy}^2 + Q_{xy}^2)} \qquad [3.33]$$

The coherency squared is:

$$w_{xy}^2(\omega) = \frac{|S_{xy}(\omega)|^2}{[S_{xx}(\omega) S_{yy}(\omega)]} \qquad [3.34]$$

3.2 RESULTS

Before the statistical analysis is performed, the distribution of K and K_d must be determined. The emperical cdf for both variables and the natural log (ln) of these variables are shown in Figures 3.6 through 3.9 in comparision with a hypothesized normal distribution calculated with the first and second moments. The Kolmogorov-Smirnov test and the visual inspection of these cdf show that the empirical cdfs of ln$K$ and lnK_d provides closer fits than the empirical cdfs of K and K_d. Consequently, ln$K$ and lnK_d were used as the random variables for all the statistical analysis performed in this study.

The statistical results for both transects are presented in the Table 3.1. Note that the table does not include a value for the mean of K_d, instead the mean of the *POC* is presented. The mean of *POC* can be converted to the mean K_d using [1.9] where the K_{oc} values for benzene, TCB and PentCB are 83 mg/l, 2040 mg/l and 13,000 mg/l, respectively.

Table 3.1: Summary of Statistics

	Horizontal Transect	Vertical Transect	Both Transect
Sample Size	73	25 (26)*	98
ln*K*: (*K* in cm/s)			
Lowest Value	-1.92	-2.01	
Highest Value	-3.45	-3.17 (-4.45)*	
Mean	-2.57	-2.56 (-2.63)*	-2.57
Variance of the mean	1.1×10^{-3}	3.5×10^{-3}	8.1×10^{-4}
Variance	0.061	0.087	0.080
Variance of the variance	1.7×10^{-4}	6.3×10^{-4}	1.3×10^{-4}
Correlation scale (λ)	1.0	0.10	
POC			
Lowest Value	0.007	0.006	
Highest Value	0.054	0.032 (0.132)*	
Mean	0.019	0.016 (0.020)*	0.018
Variance of the mean	9.7×10^{-7}	1.3×10^{-7}	6.3×10^{-7}
lnK_d: (K_d in mg/l)			
Variance	0.138	0136	0.145
Variance of the variance	5.3×10^{-4}	1.5×10^{-3}	1.5×10^{-3}
Correlation scale (λ)	1.0	0.15	
Residual			
Mean	-0.00079	0.00214	0.00118
Variance of the mean	1.9×10^{-3}	5.0×10^{-3}	1.4×10^{-3}
Variance	0.132	0.128	0.137
Variance of the variance	4.8×10^{-4}	1.3×10^{-3}	3.9×10^{-4}
Correlation scale (λ)	0.5	0.10	
Slope of lnK_d vs ln$K$	-0.339	-0.345	-0.341

* The bracketed values represent the vertical transect including the values for sample AB, the last point on the transect.

Sample AB, the last point on the vertical transect characterizes the top soil in the area. This sample has a large *POC* value, 0.132%, and consequently a large sorption coefficient[4] and small ln*K*, -4.45. This negative correlation supports the notion that low hydraulic conductivity correspondes to high sorption coefficients. However, this point is not suitable for our analysis since it is from a different population. When this sample is included in the data set, the mean changes by as much as 10%. The change in the values of ln*K* and lnK_d from sample AA (the adjacent point) to sample AB is equivalent to the range of the values of the whole transect. Appendix A shows the ln*K* and lnK_d results for both transects.

Is it suitable to treat the samples from each transect as a separate unit or should all the data be grouped together? The summary of statistics shows the mean and variance of the data for the individual transects and for both transects grouped together. Though the means and variances of *POC*, ln*K* and lnK_d for the horizontal and vertical transect are numerically different, the differences may not be statistically significant. The range within which the mean or variance is statistically equivalent can be determined by estimating the confidence interval of the mean or variance. The confidence intervals for the vertical transect will be used for this analysis since they will be wider than the confidence interval for the horizontal transect. The 95 percent confidence interval of the mean and variance can be estimated as two standard deviations normalized by the mean or variance, respectively. The relative confidence interval for the variance of lnK_d is 0.289, therefore the variance of ln*K* and lnK_d for the horizontal and vertical transect found in Table 3.1 is statistically

[4]All K_d and lnK_d values are for TCB, except where mentioned.

equivalent. Likewise, the mean of *POC* and the mean and variance of ln$K$ for the horizontal and vertical transect are statistically equivalent. Since the mean and variance of ln$K$, lnK_d and *POC* are statistically equivalent, it is reasonable to treat the data for the vertical and horizontal transect as representing a sample from the same two-dimensional stationary random field.

3.2.1 Scatter Plot

The scatter plot of ln$K$ versus lnK_d for all of the values from both transects shows a weak negative correlation between the variables (Figure 3.10). A linear regression of the data shows a negative slope of -0.341 with a coefficient of correlation, r, of -0.233.

3.2.2 Covariance Analysis

The covariance analysis includes the autocorrelation and cross-correlation functions. The spatial behavior of ln$K$ and lnK_d is observed in the realizations of ln$K$, lnK_d and the residual. The spatial behavior of the variables was presented for the horizontal and vertical directions.

Realizations and Cross Correlation Functions

The horizontal transect. Seventy-three measurements were taken at 0.5 m intervals along the horizontal transect. The realizations of ln$K$ have a distinctive high and low peak (Figure 3.11). These peaks have values of -1.92 and -3.44 and are

located at 35.5 m and 2.5 m, respectively[5]. For $\ln K_d$, there are two distinctive high peaks with values of -0.085 and 0.135 and are located at 12 m and 32 m, respectively. The lowest $\ln K_d$ value of -1.68 occurs at 3 m and 4 m. The peak values of $\ln K$ correspond to regions of low $\ln K_d$. The peaks of $\ln K_d$ correspond to regions of low $\ln K$. The smallest values of $\ln K$ does not necessarily correspond to the highest values of $\ln K_d$. The smallest values of $\ln K_d$ does not necessarily correspond to the highest values of $\ln K$. The cross correlation function ($R_{xy}/\sigma_x\sigma_y$) supports this observations (Figure 3.13). This function indicates that the maximum negative correlation occurs at a lag of -5 m. It also shows that there is a noticeable negative correlation at a lag of 0.5 m. There are regions of significant positive and negative correlation at varying lags. The realization of the residual, $\eta = \ln K_d - 0.341 \ln K$, peaks at 12.5 m and 32 m (Figure 3.14).

The vertical transect. Twenty-six measurements were taken at 0.15 m intervals along the vertical transect. The realizations of $\ln K$ and $\ln K_d$ attain a maximum value of -2.01 at 0.90 m and -0.419 at 0.15 m respectively (Figure 3.12). The maximum value of $\ln K$ does not correspond to the lowest value of $\ln K_d$. The maximum value of $\ln K_d$ does not correspond to the lowest value of $\ln K$. Low values of $\ln K$ tend to correspond to regions of high $\ln K_d$. The converse also holds true, low values of $\ln K_d$ tend to correspond to regions of high $\ln K$. The cross correlation function (Figure 3.16) shows that the maximum negative correlation occurs at a lag of -0.15 m. This function cycles from positive to negative values. The realization of the residual, η

[5]For the horizontal transect, the distances are measured from left to right. For the vertical transect, the distances are measured from the lowest point on the exposure to the highest point.

= $\ln K_d$ - 0.341ln$K$, has the highest value of 0.80 at 0.15 m; the lowest value is -0.79 at 2.55 m (Figure 3.15).

Autocorrelation Functions

Visual inspection of the realizations can yield an estimate of the lense lengths and bed thickness.

The horizontal transect. Based on the realization of ln$K$ and lnK_d, the lense length ranges from 0.5 m (the sampling interval) to 2 m. From the realization of the residual, the lense length is equal to, or less than, the sampling interval of 0.5 m. From the autocorrelation function of ln$K$, lnK_d and the residual (Figures 3.17 through 3.19) the λ_1 is 1 m, 1 m and 0.5 m for ln$K$, lnK_d and η.

The vertical transect. From the realizations of ln$K$ and lnK_d the bed thickness ranges from 0.15 m to 0.60 m. The bed thickness for the residual ranges from 0.15 m to 0.30 m. Based on the autocorrelation functions of ln$K$, lnK_d and η for the vertical transects, shown in Figures 3.20 through 3.22, λ_3 is 0.1 m, 0.15 m and 0.10 m respectively.

The autocorrelation function for both transects and both variables exhibited hole type features, see Figure 3.3. The functions decreased with increasing lag, however, they never leveled out to zero. At higher lags, where a certain degree of oscillation was evident, it became unreliable. To determine the significance of these fluctuations, the 95% confidence interval for the autocorrelation function for large lag where ($\rho \cong 0$) is computed as approximately $2/\sqrt{N_2}$, where N_2, the number of independent samples is:

$$N_2 = N\frac{1-a^2}{1+a^2} \qquad a = e^{-\Delta/\lambda} \qquad [3.35]$$

where N is the number of samples, Δ is the sampling interval, and λ is the correlation scale. All of the fluctuations fell within the 95% confidence interval, thus the fluctuations are not regarded to reflect correlation that is significantly different from zero.

In Figures 3.17 through 3.22, the autocorrelation function is presented along with an estimate of the function. This estimate is calculated as $\rho = e^{-s/\lambda}$, where s is the lag. These fits provide a reasonable smoothed estimate of the autocorrelation function.

3.2.3 Spectral Analysis

The spectra shown in Figure 3.23 through 3.30, were calculated using the function Spectrum (*X,Y,M*) in the Signal Processing Toolbox of MATLAB. The code for this program is shown in Appendix B. According to the help screen found in Matlab, this code "performs FFT (fast Fourier transform) analysis of the two sequences X and Y using the Welch method of power spectrum estimation. The X and Y sequences of N points are divided into K sections of $2M$ points each, ($2M$ must be a power of two). Using an $2M$-point FFT, successive sections are Hanning windowed, FFT'd and accumulated." Appendix B shows the codes for the program Spectrum and other related programs, Appendix C shows the spectral estimates of this program.

It is standard to present the spectral estimates on a semilog plot. The variance of the logarithm of the spectral estimate can be approximated by (see Priestley, *1981*, p470):

$$var\ \ln[\hat{S}(\omega)] = \frac{\mu M}{N} \qquad [3.36]$$

where N is the number of samples and μ is a coefficient which is dependent on the spectral window used[6]. The relative precision of this spectral estimate is proportional to $\sqrt{M/N}$. As the M value doubles, the confidence interval will increase by a factor of $\sqrt{2}$. There is a tradeoff between precision and resolution. Smaller M values leads to narrower confidence intervals but lower resolution. Larger M values leads to higher resolution coupled with less precision (larger confidence intervals).

In striking a balance between the precision and the resolution, *Gelhar (1993)* suggests that one "selects M to be relatively small compared to N, say 5 to 30 percent, and then test the sensitivity of the estimated spectra to the change in M." This test was applied to the data from the horizontal and vertical transect. For the horizontal transect, M values of 8, 16 and 32 were used. Figure 3.23 shows the spectral estimates for the varying M values. The confidence interval for the log of the spectral estimates are 0.249, 0.352 and 0.498 for M values of 8, 16 and 32 respectively. These confidence interval differ by a factor of $\sqrt{2}$. Though the spectral estimates corresponding to $M = 8$ were more precise than the other two estimates, this curve fails to highlight all the variation in the spectral density function. For the plot of $M = 32$, the significance of the fluctuations of the spectral density function

[6]For the Hanning window μ is 0.75.

cannot be determined based on the wide confidence interval. The curve for $M = 16$ shows a reasonable compromise between precision and resolution. A similar analysis was done for the spectral estimates corresponding to $M = 4$ and $M = 8$ for the vertical transect. The plot corresponding to $M = 4$ was selected.

Spectral estimates were calculated based on three spectral windows (the Hanning window, Bartlett window and Triangular window), Figure 3.24 shows the spectral estimate. It is obvious that the difference in these plots are minor compared to the 95% confidence interval and are therefore insignificant. The Hanning window was used for the all further spectral estimates. Appendix B shows the MATLAB codes that correspond to each of the respective windows.

For *the horizontal transect* the behavior of the estimated spectrums of $\ln K$, $\ln K_d$ and the η are shown in Figures 3.25, 3.26 and 3.27, these spectral estimates are referred to as S_{ff}, S_{gg} and $S_{\eta\eta}$, respectively. At the low frequency end of S_{ff}, the curve displays an increase in the spectral density function that is characteristic of the hole type spectrum. At other frequencies, the spectral estimates fluctuates within the confidence interval. For S_{gg}, the spectral estimates are higher at lowever frequencies. There is an unusual increase in the spectral density at the high frequency end of S_{ff}. This may be due to aliasing[7] or "unaccounted variances at infinitely small scales" (*Robin et al., 1991*) an example of which is measurement noise. The estimated spectrum of the residual behaves in a similar manner. The spectral results for $\ln K_d$ are shown in Figure 3.26 with the empirically estimated confidence intervals produced by MATLAB implicitly assuming that the spectral estimates are normally

[7]The "aliasing effect" occurs when "spectral contributions from higher frequencies are folded back into the calculated frequency range, and increase the estimated spectral amplitude" (Gelhar, 1993).

distributed and the standard theoretical estimates of confidence interval estimated by [3.35].

For *the vertical transect*, the spectrums of $\ln K$, $\ln K_d$ and η have quite flat slopes. Based on the short record length and the resultant low M value, the spectrum has a low resolution; it only remains to conclude that the flat slopes are caused by white noise.

The Cross Spectra and Coherency

The quadrature spectra for both transects oscillates around zero, indicating that there was no significant out of phase correlation. For the horizontal transect (Figure 3.31), the cospectrum cycles between positive and negative values, showing that the correlation is indeed scale dependent. The most significant negative correlation is at a scale of 8 m. For the vertical transect (Figure 3.32) the cospectrum cycles between negative values for smaller frequencies and positive values for larger frequencies. The coherency squared spectrum for the horizontal transect (Figure 3.33) shows that significant correlation between $\ln K_d$ and $\ln K$ occurs at a number of scales. For the coherency squared spectrum (Figure 3.34) for the vertical transect shows that significant correlation occurs at scales of less than 8 m.

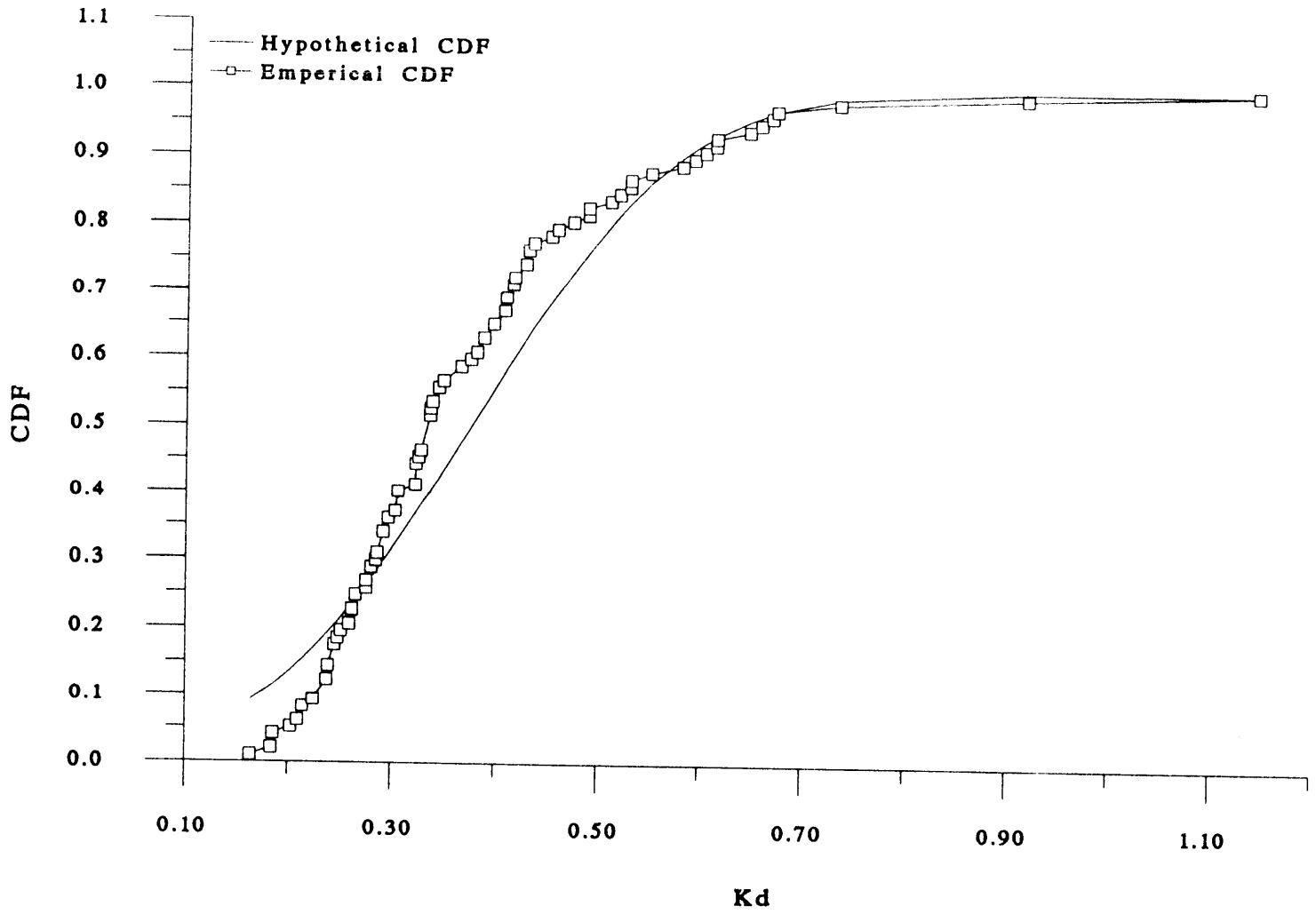

Figure 3.6 Hypothetical and empirical cumulative probability distribution function for K_d values

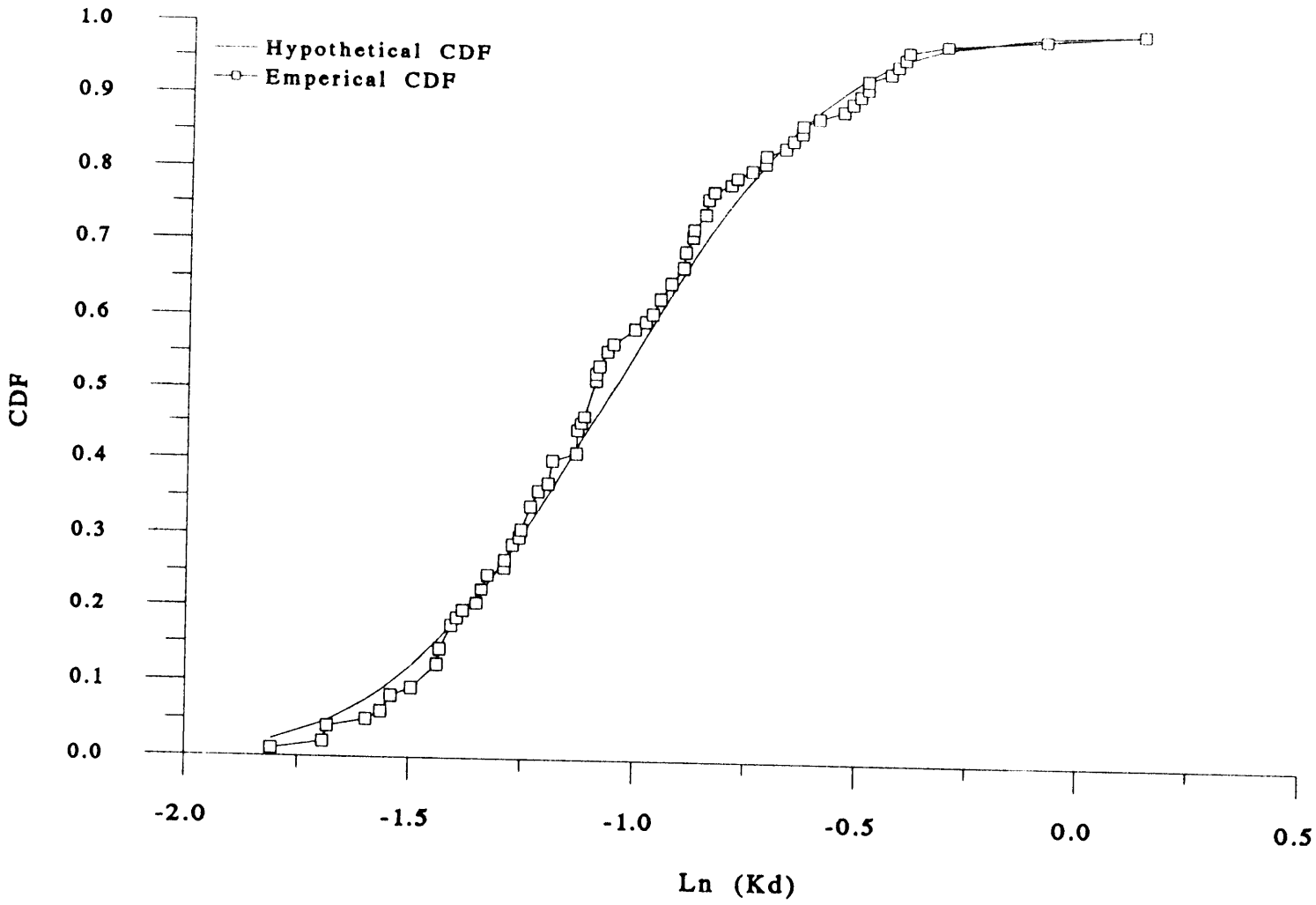

Figure 3.7 Hypothetical and empirical cumulative probability distribution function for $\ln K_d$ values

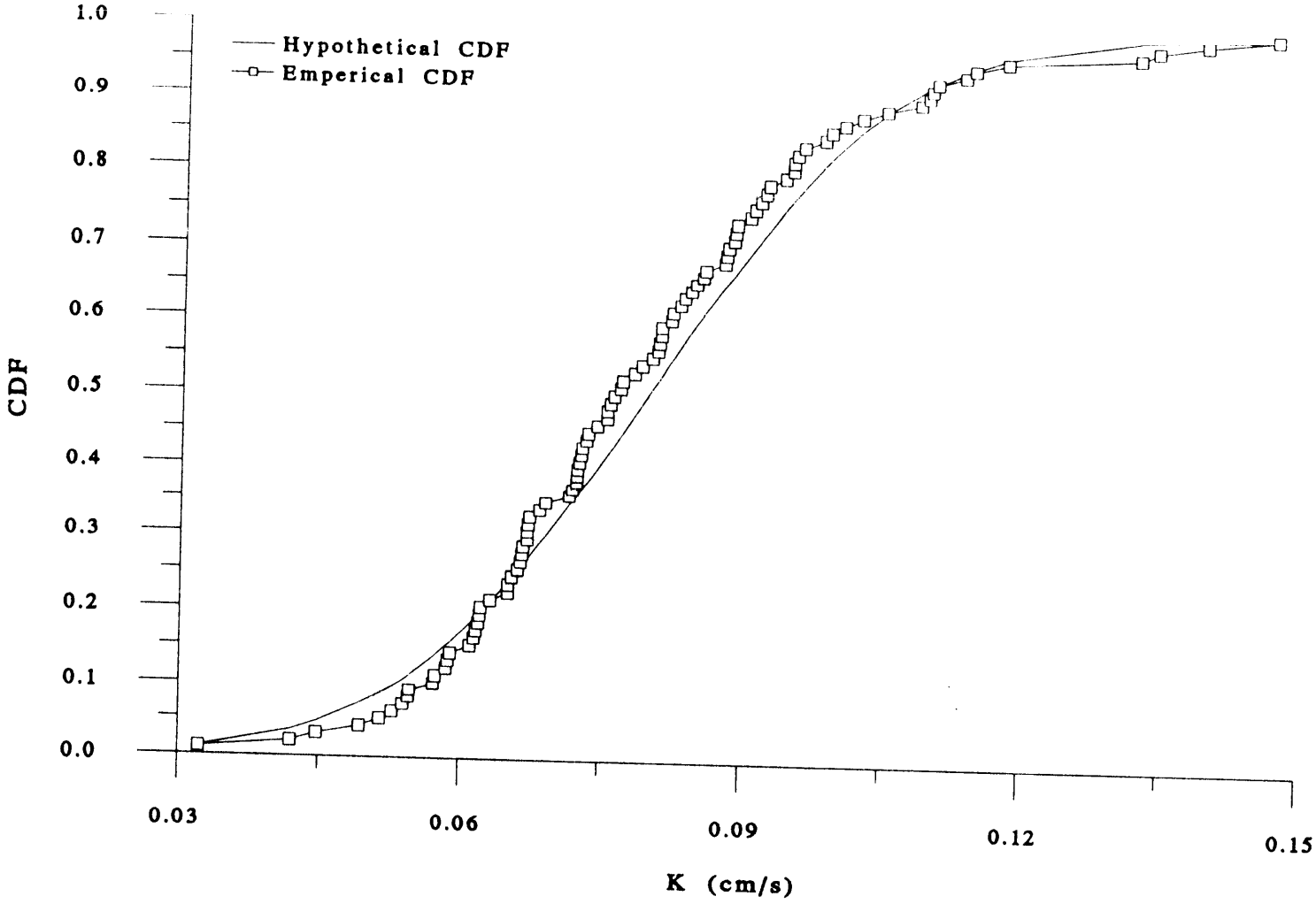

Figure 3.8 Hypothetical and empirical cumulative probability distribution function for K values

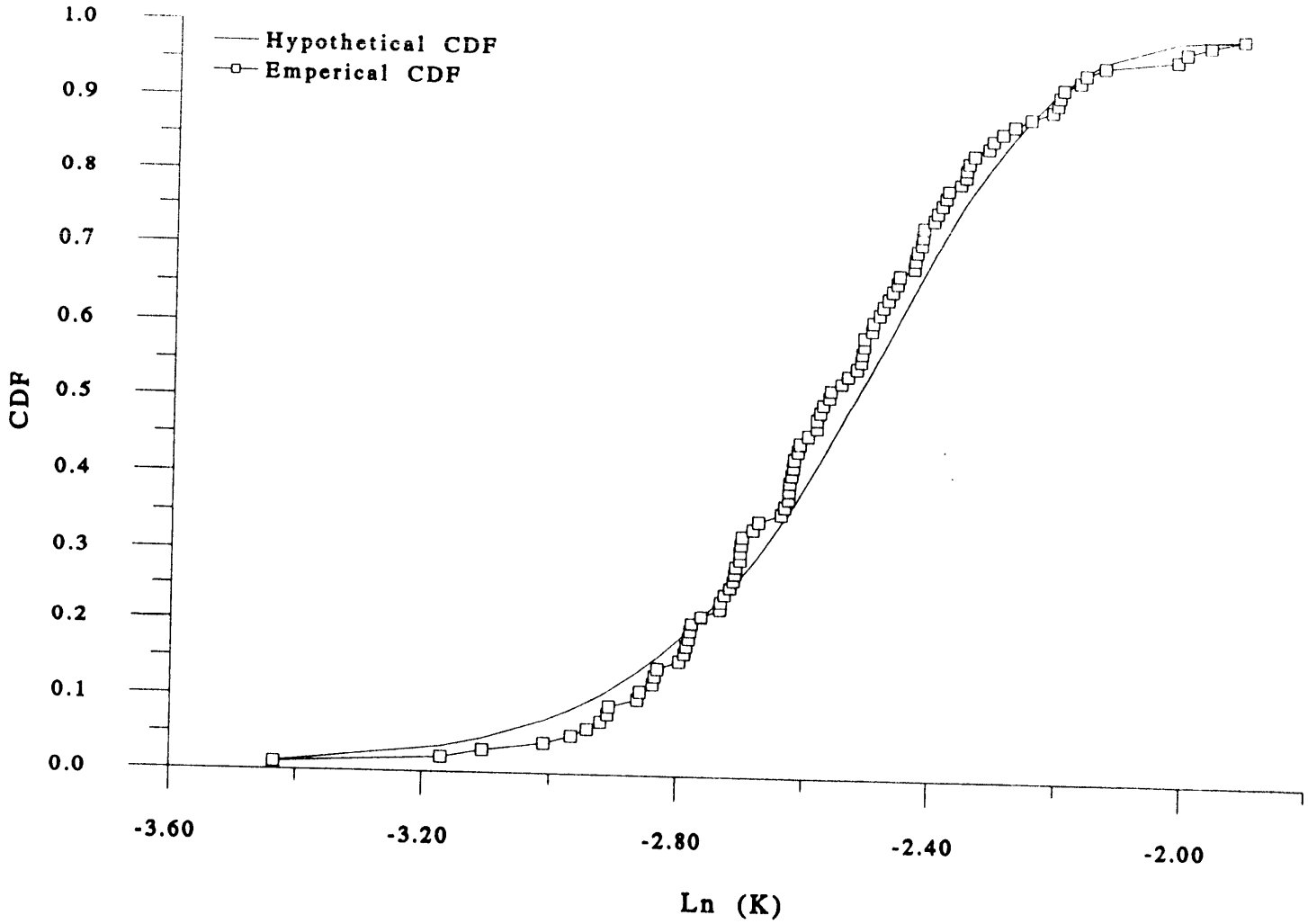

Figure 3.9 Hypothetical and empirical cumulative probability distribution function for ln*K* values

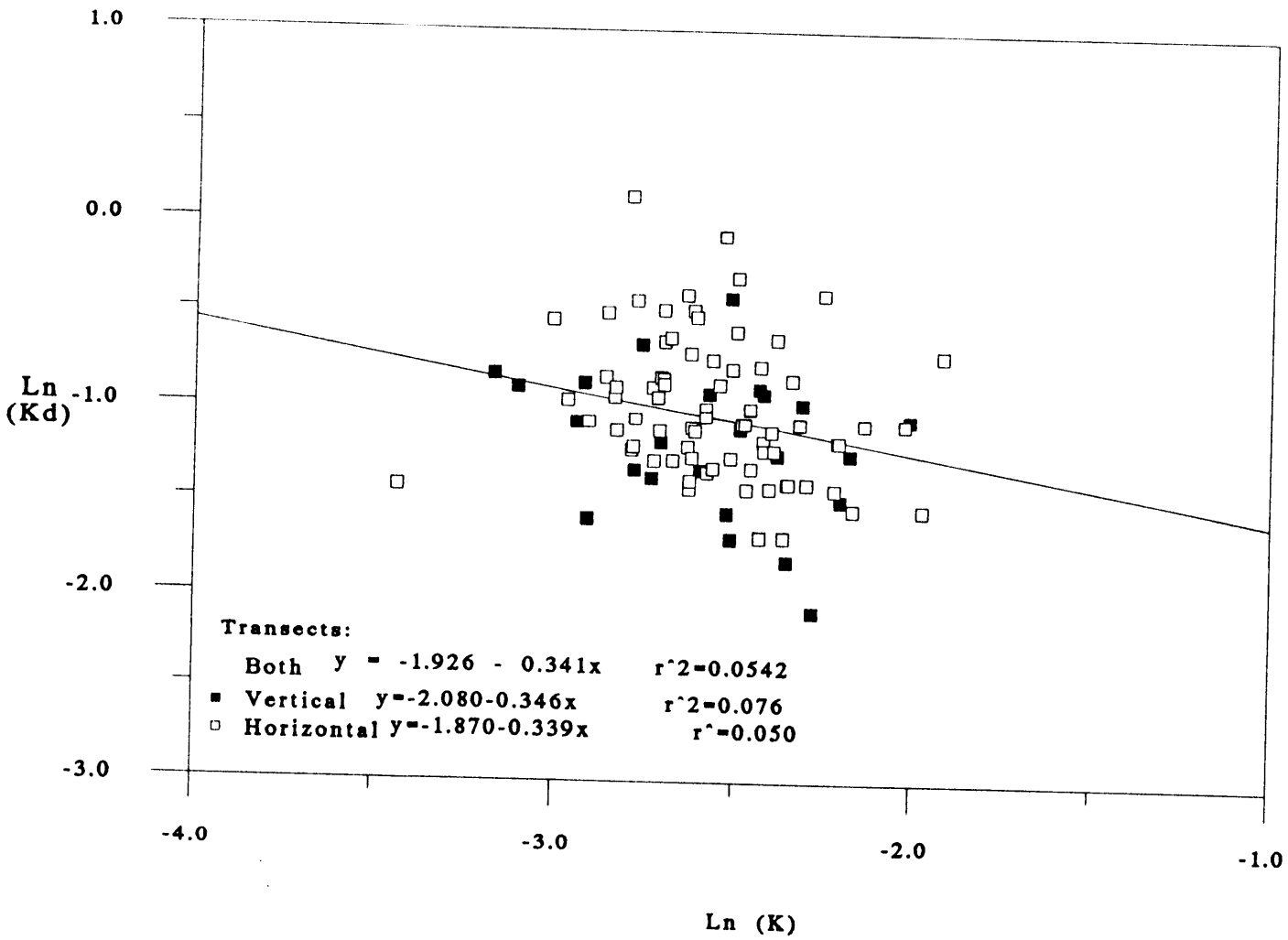

Figure 3.10 Scatter plot of ln$K$ and lnK_d for the vertical transect, horizontal transect and both transects.

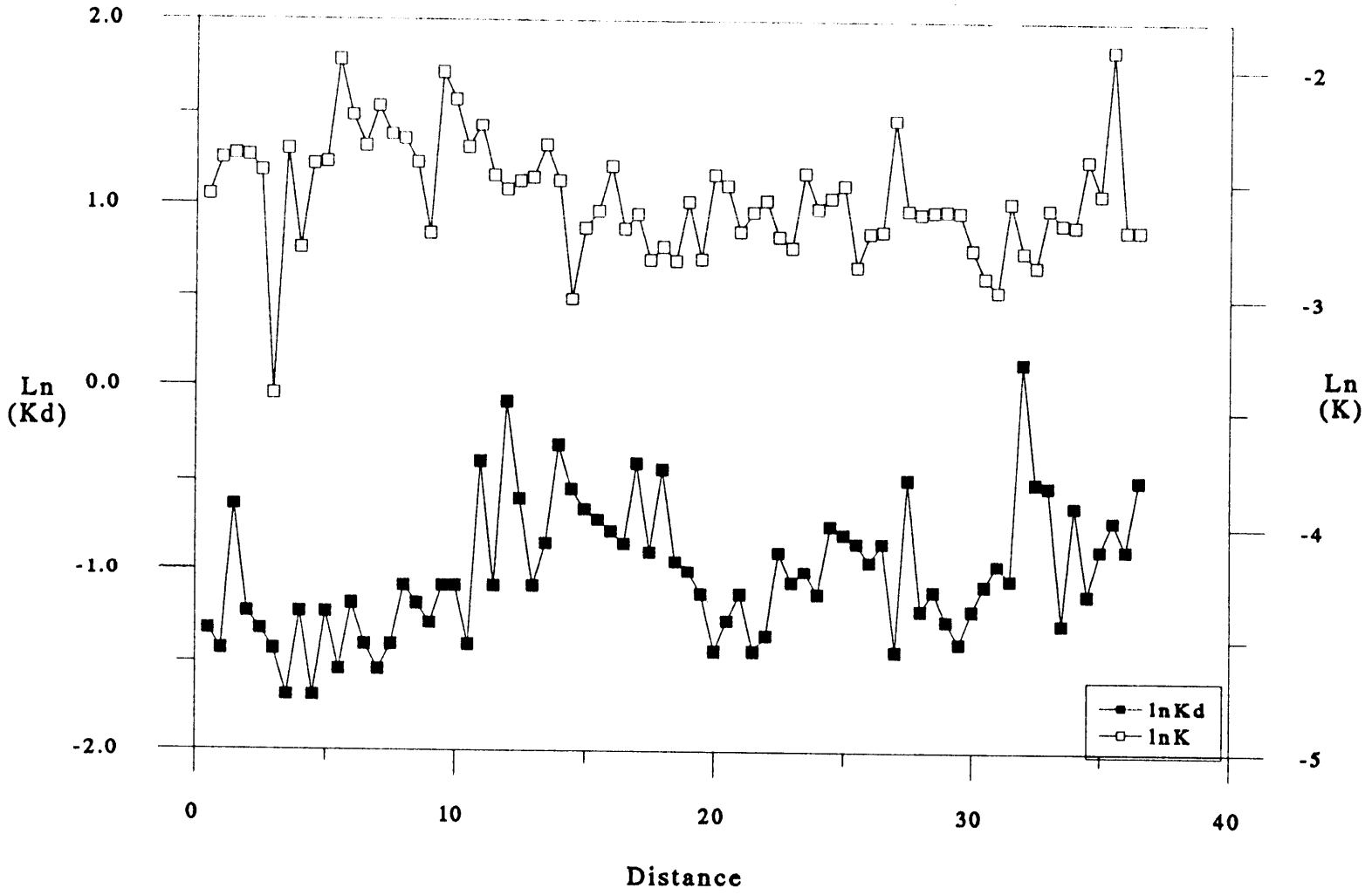

Figure 3.11 Realization of ln$K$ and lnK_d for the horizontal transect

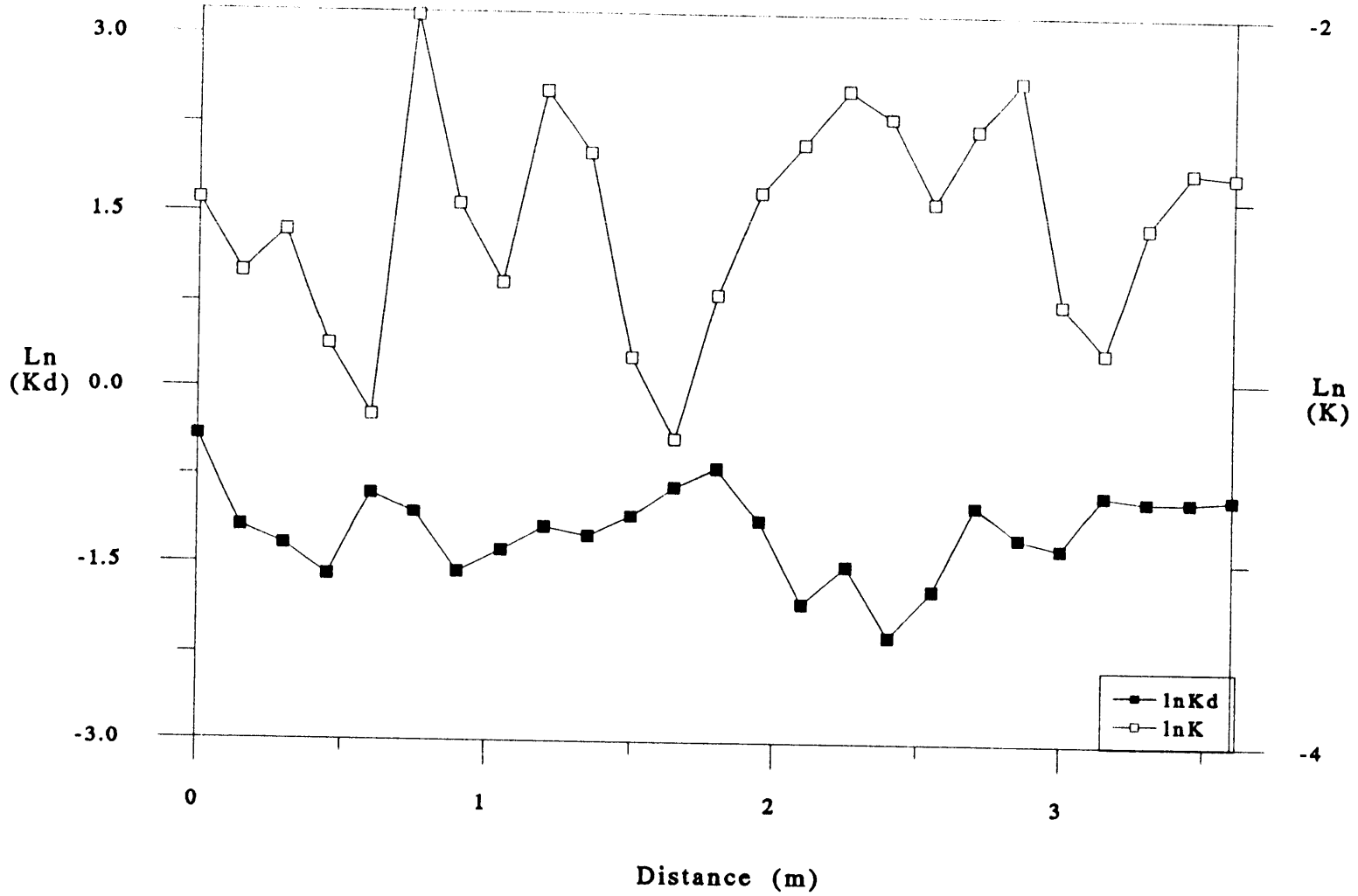

Figure 3.12 Realization of ln$K$ and lnK_d for the vertical transect

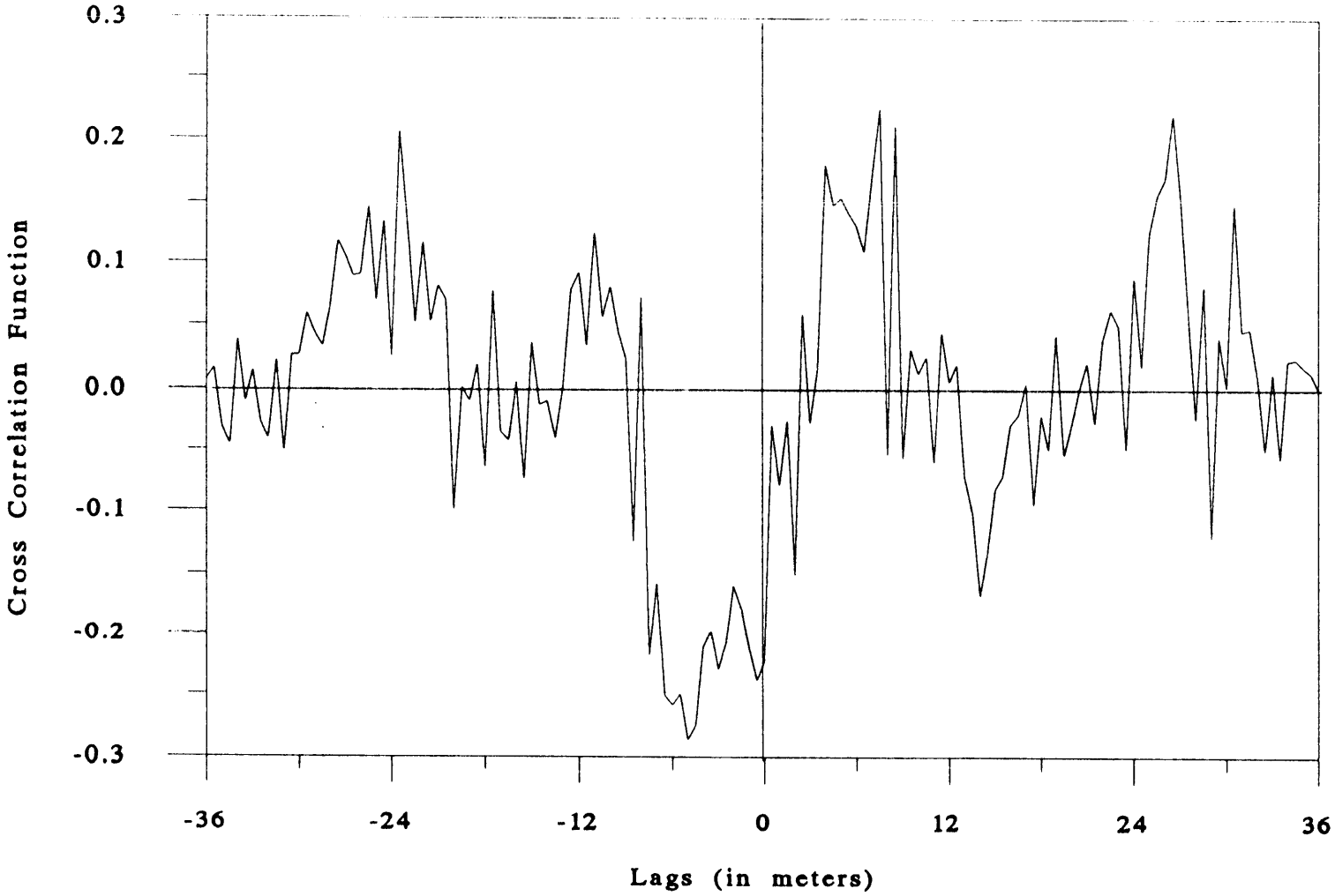

Figure 3.13 Cross-correlation function of ln$K$ and lnK_d for the horizontal transect

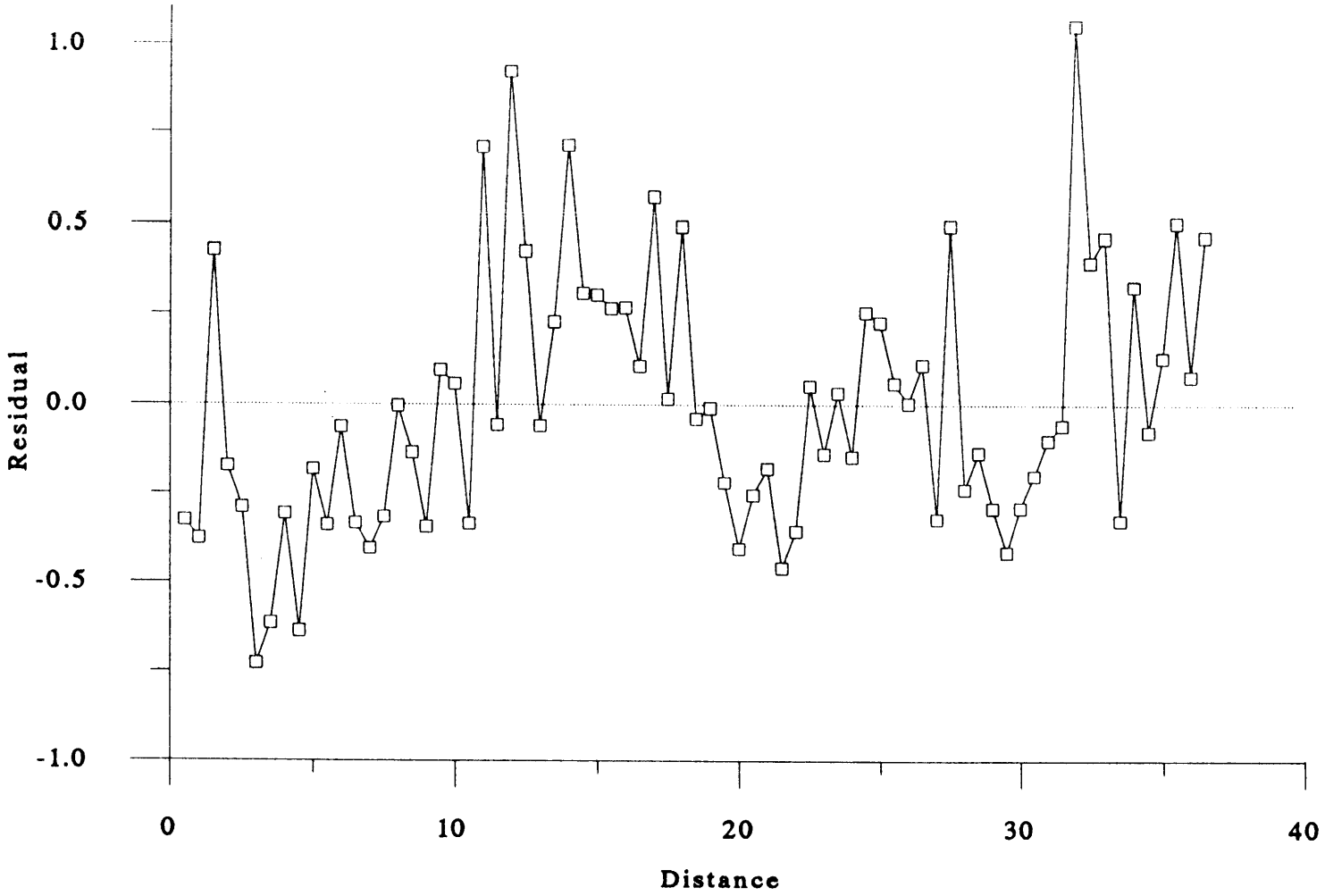

Figure 3.14 Realization of the residual for the horizontal transect

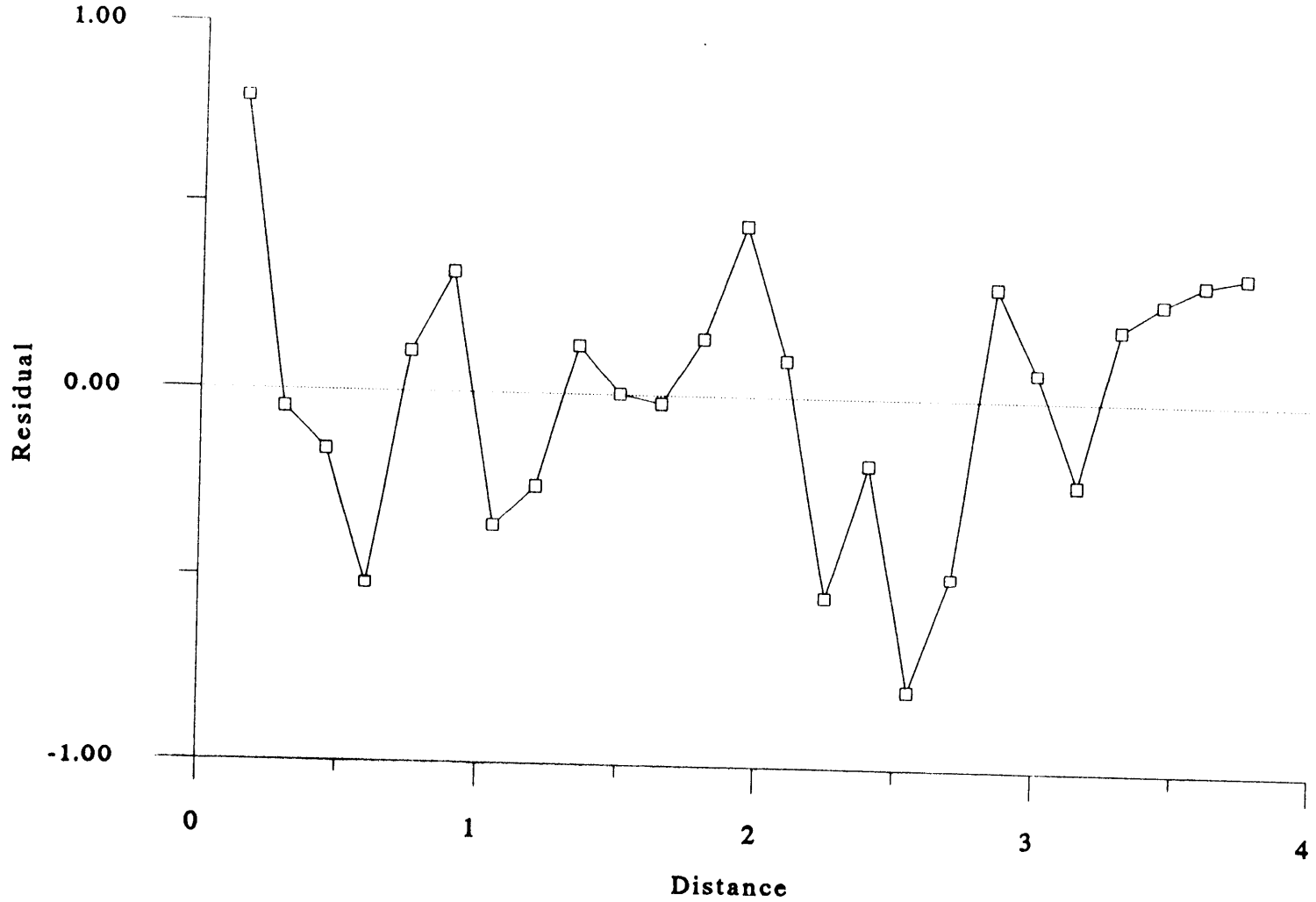

Figure 3.15 Realization of the residual for the vertical transect

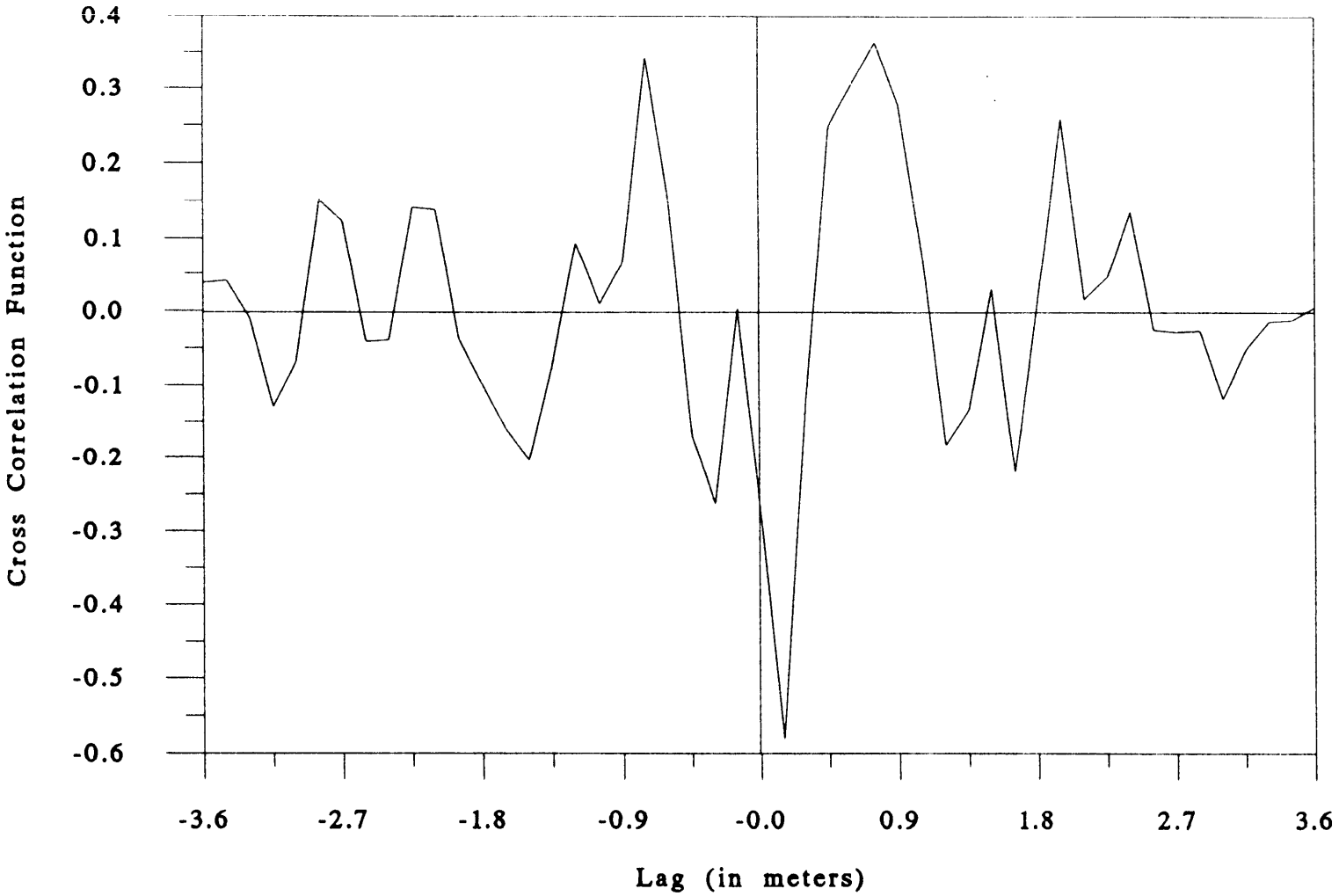

Figure 3.16 Cross-correlation function of ln$K$ and lnK_d for the vertical transect

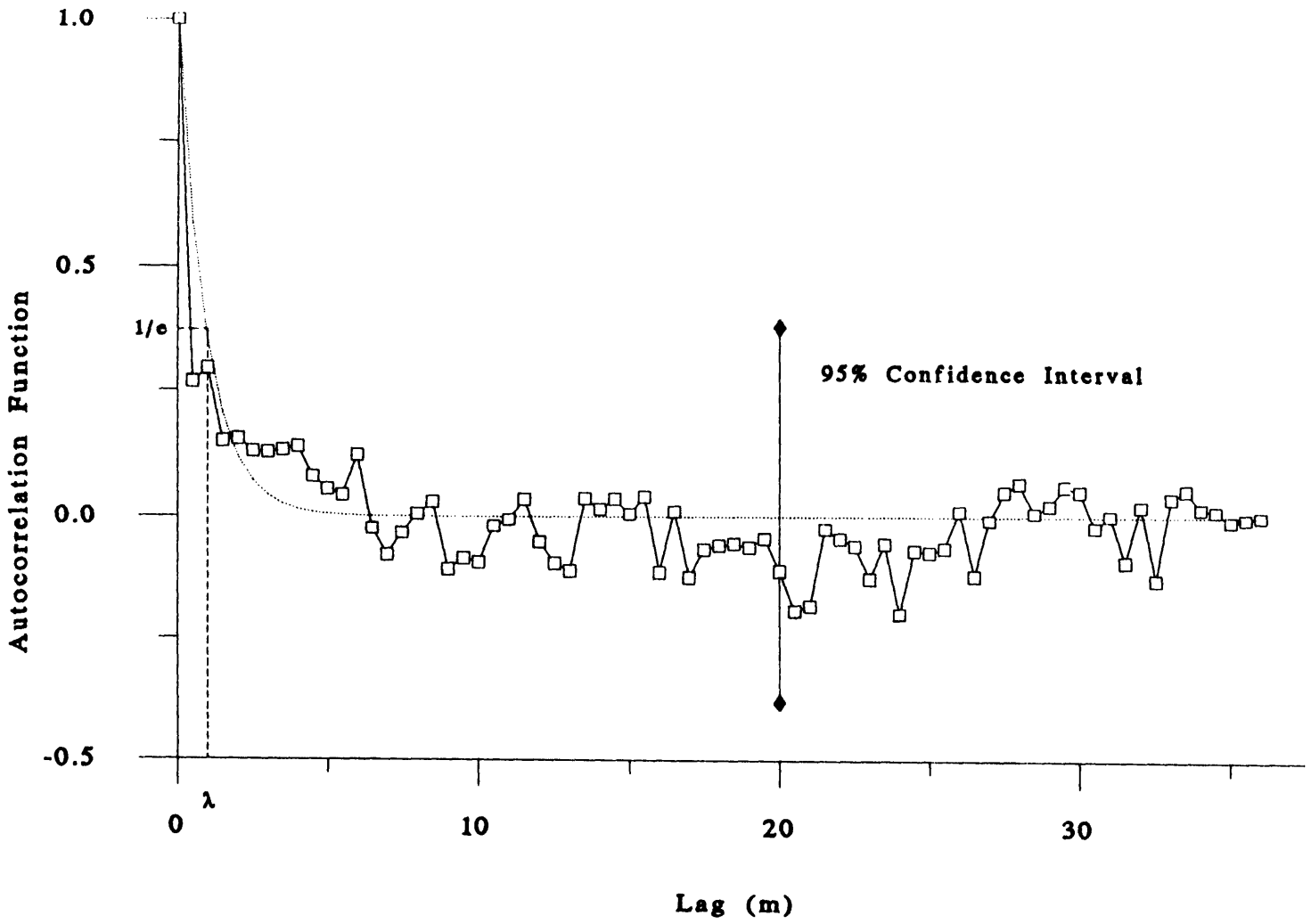

Figure 3.17 Autocorrelation function of ln*K* for the horizontal transect. The dashed line indicates the estimate of the exponential autocorrelation function, $\rho = e^{s/\lambda}$, where s is the lag.

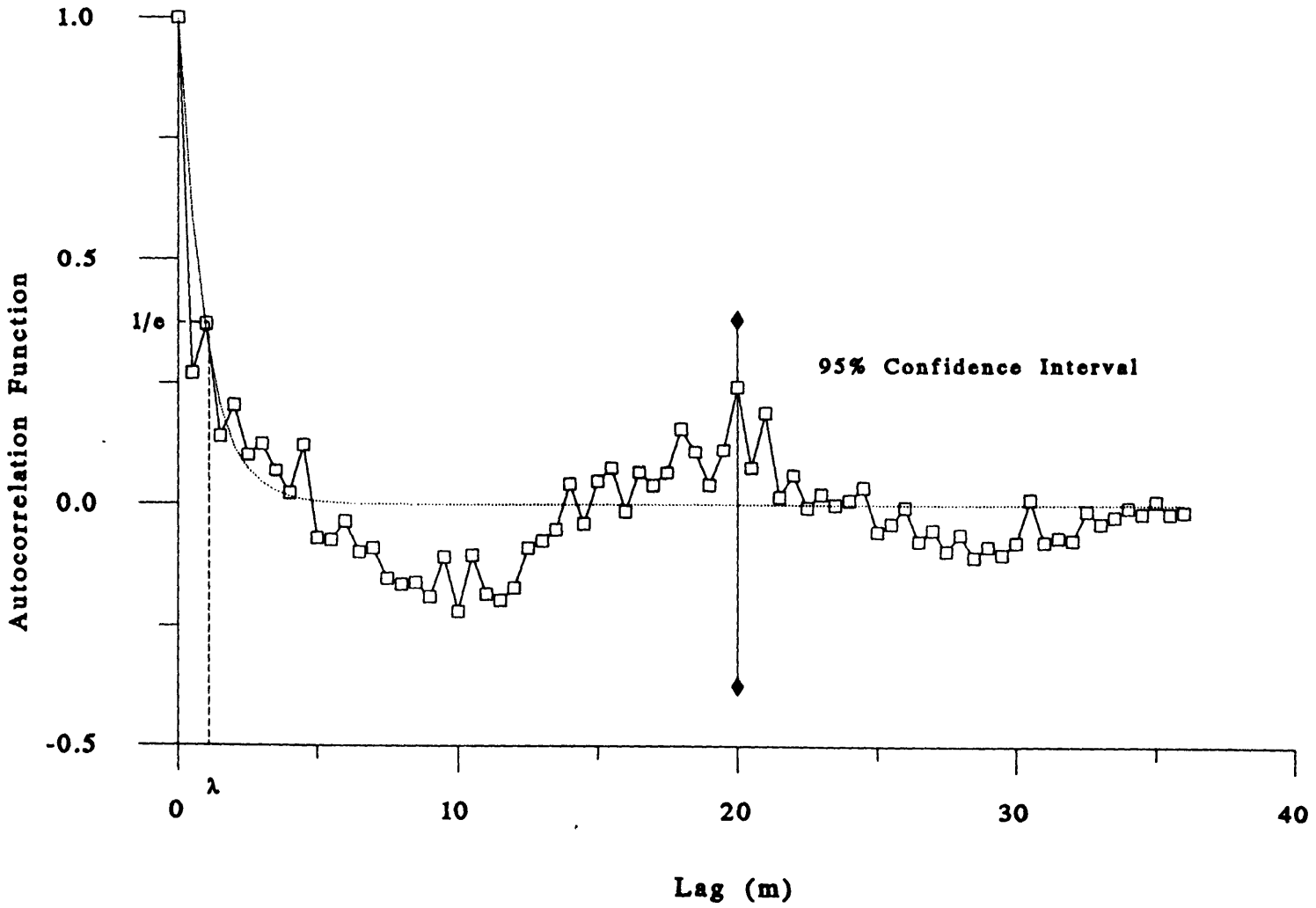

Figure 3.18 Autocorrelation function of $\ln K_d$ for the horizontal transect. The dashed line indicates the estimate of the exponential autocorrelation function $\rho = e^{s/\lambda}$, where s is the lag.

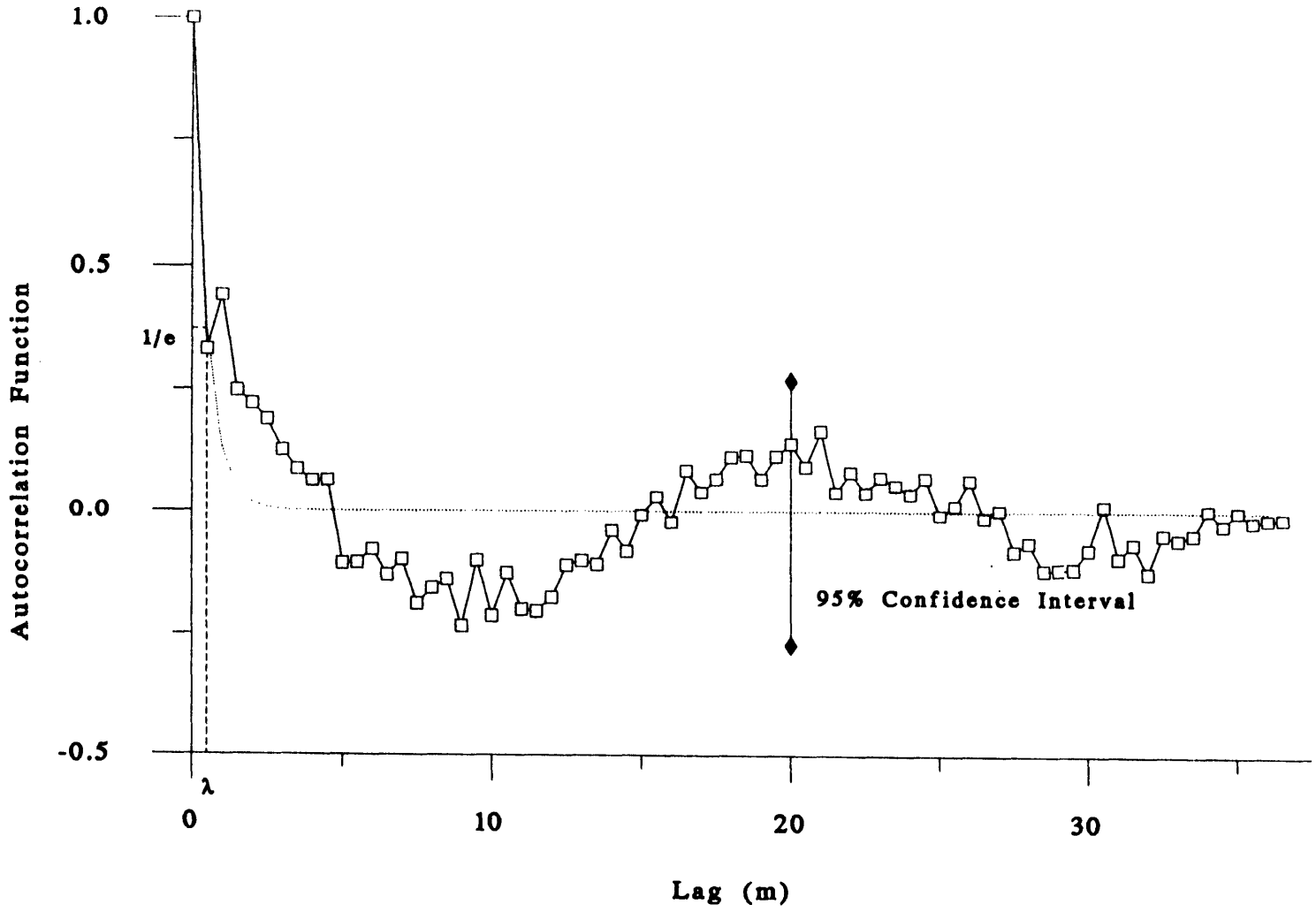

Figure 3.19 Autocorrelation function of the residual for the horizontal transect. The dashed line indicates the estimate of the exponential autocorrelation function, $\rho = e^{s/\lambda}$, where s is the lag.

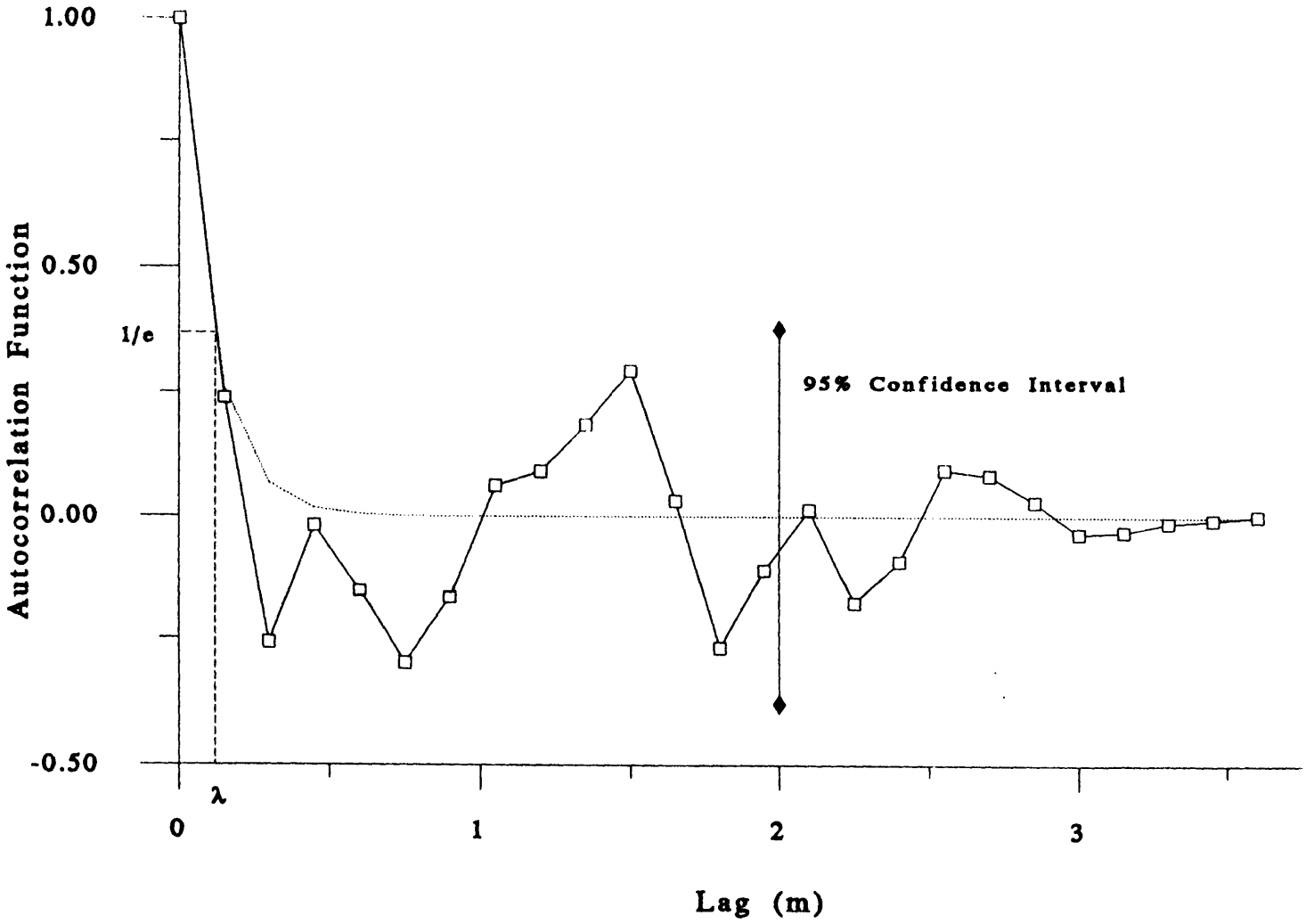

Figure 3.20 Autocorrelation function of ln$K$ for the vertical transect. The dashed line indicates the estimate of the exponential autocorrelation function, $\rho = e^{s/\lambda}$, where s is the lag.

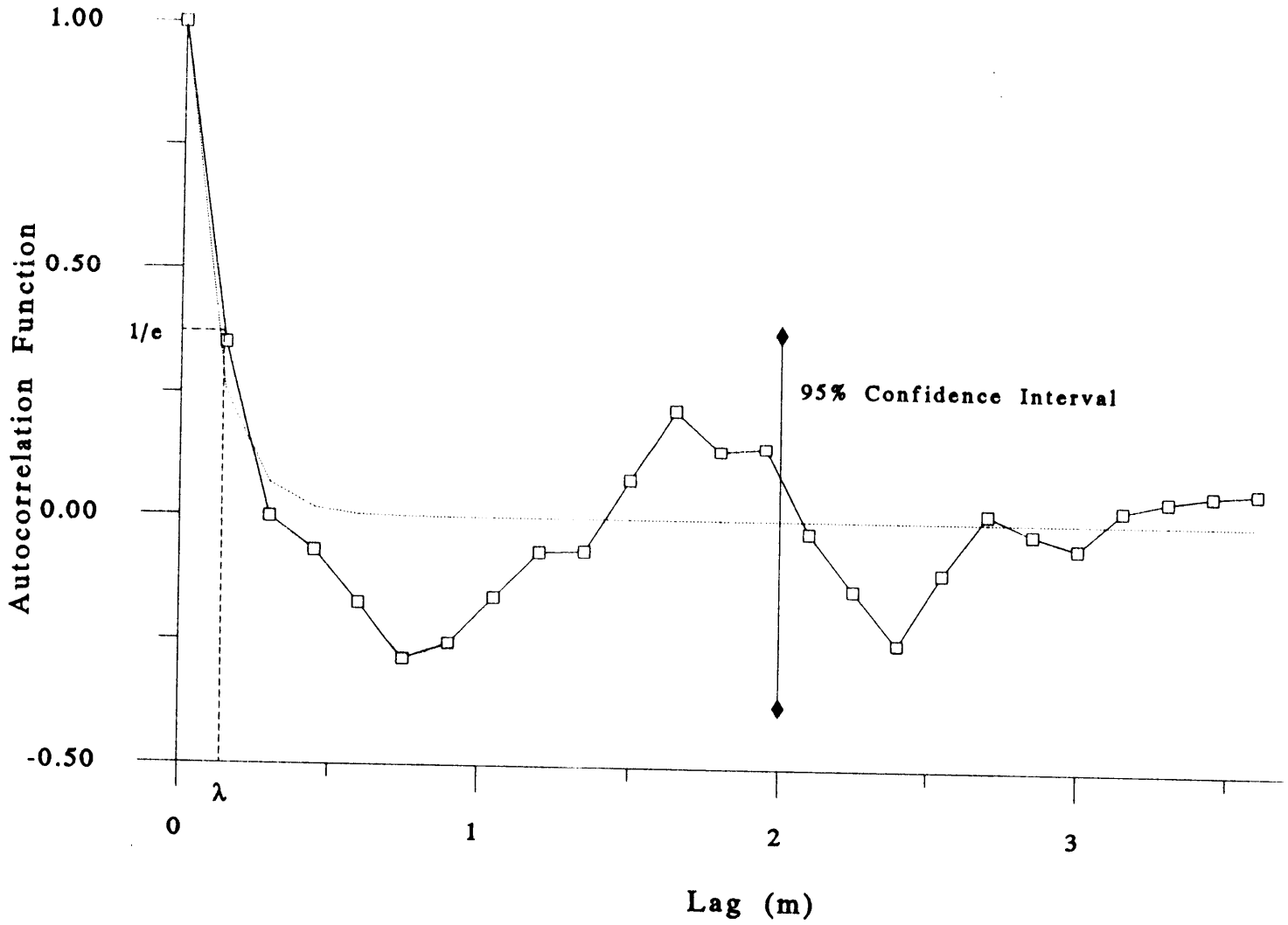

Figure 3.21 Autocorrelation function of $\ln K_d$ for the vertical transect. The dashed line indicates the estimate of the exponential autocorrelation function, $\rho = e^{s/\lambda}$, where s is the lag.

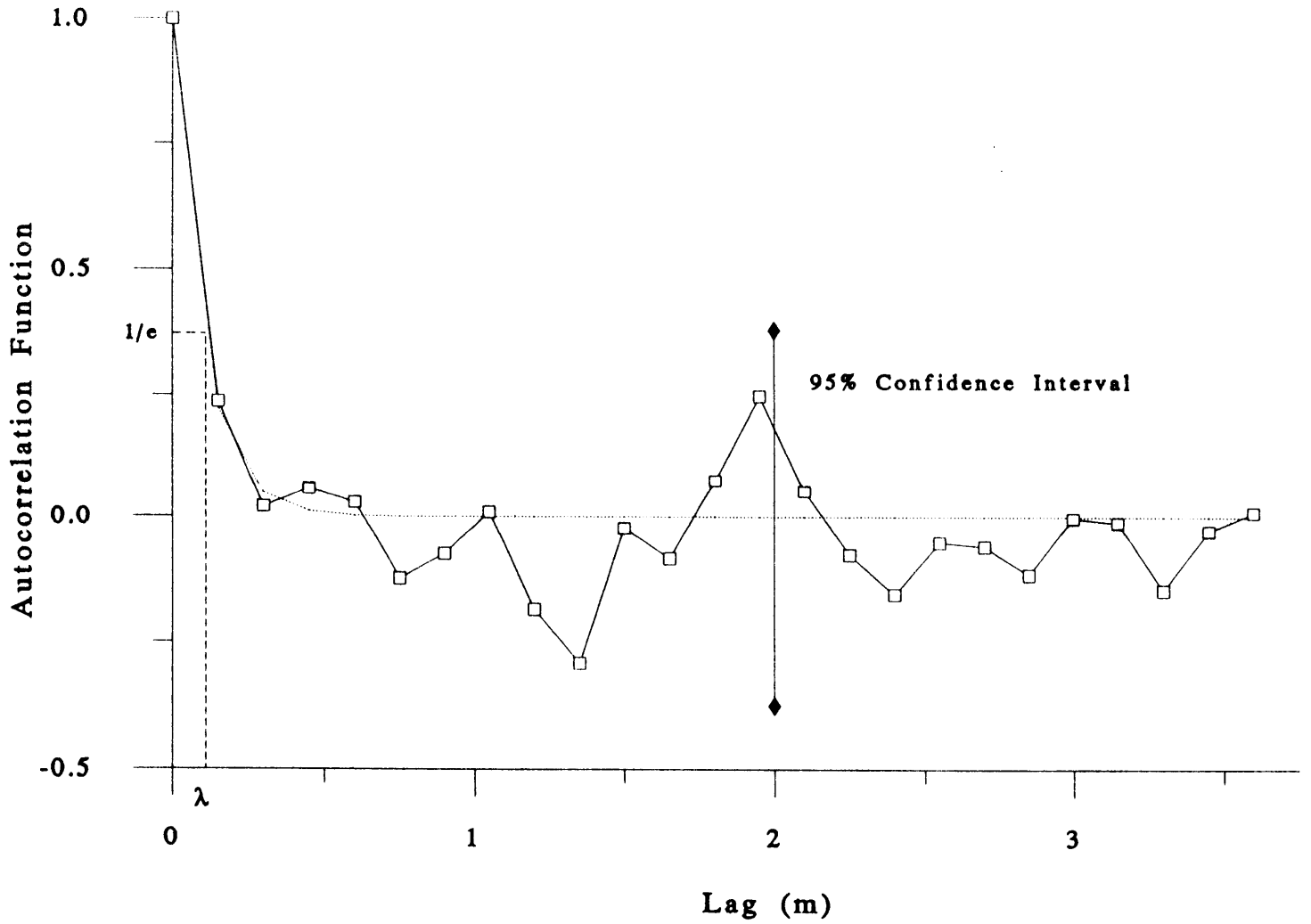

Figure 3.22 Autocorrelation function of the residual for the vertical transect. The dashed line indicates the estimate of the exponential autocorrelation function, $\rho = e^{s/\lambda}$, where s is the lag.

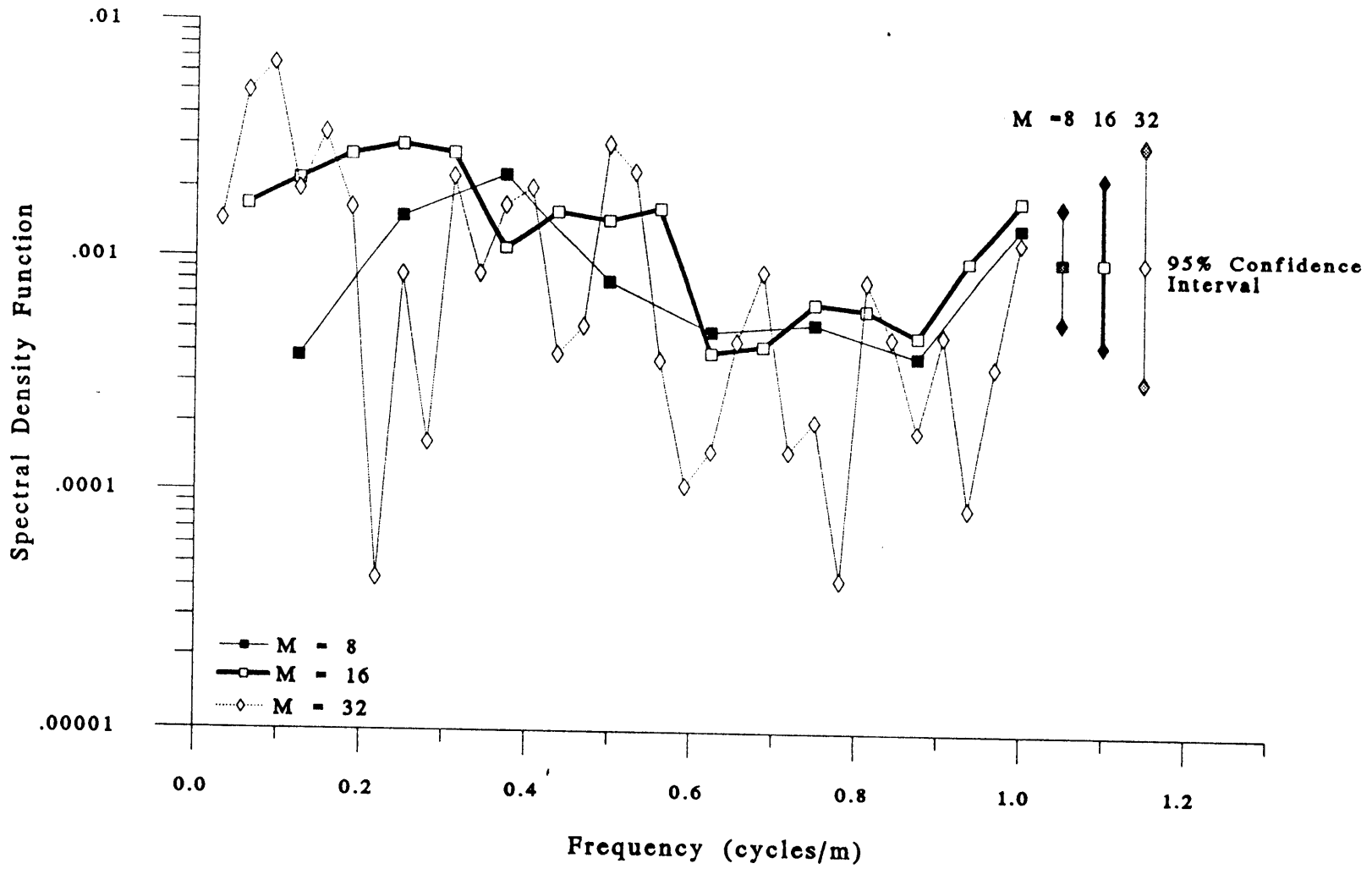

Figure 3.23 Spectral estimates using different M values. The Hanning window was used in these spectral estimates of ln$K$ for the horizontal transect.

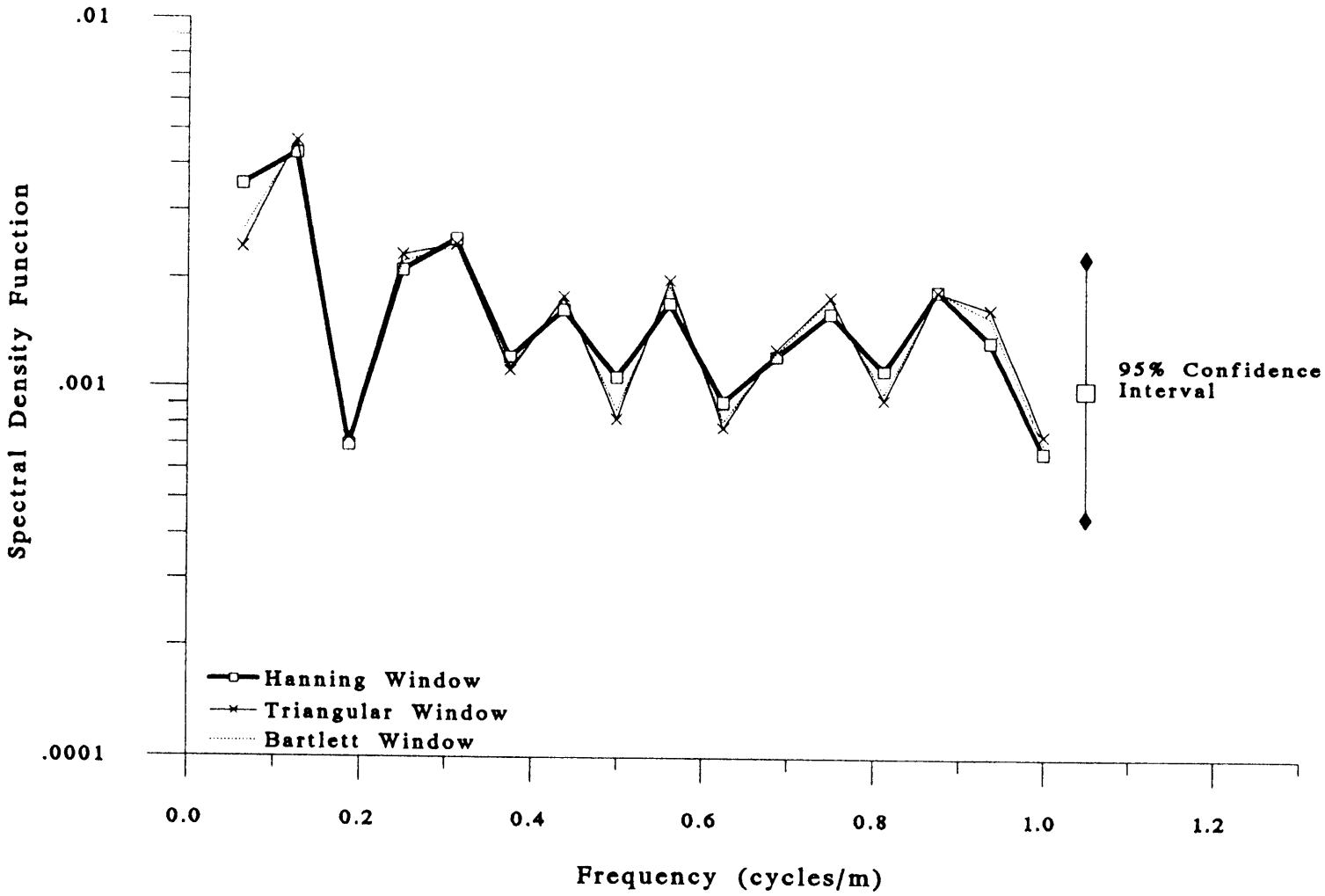

Figure 3.24 Spectral estimates using different spectral windows. This is the spectal estimate of ln$K$ for the horizontal transect with M = 16.

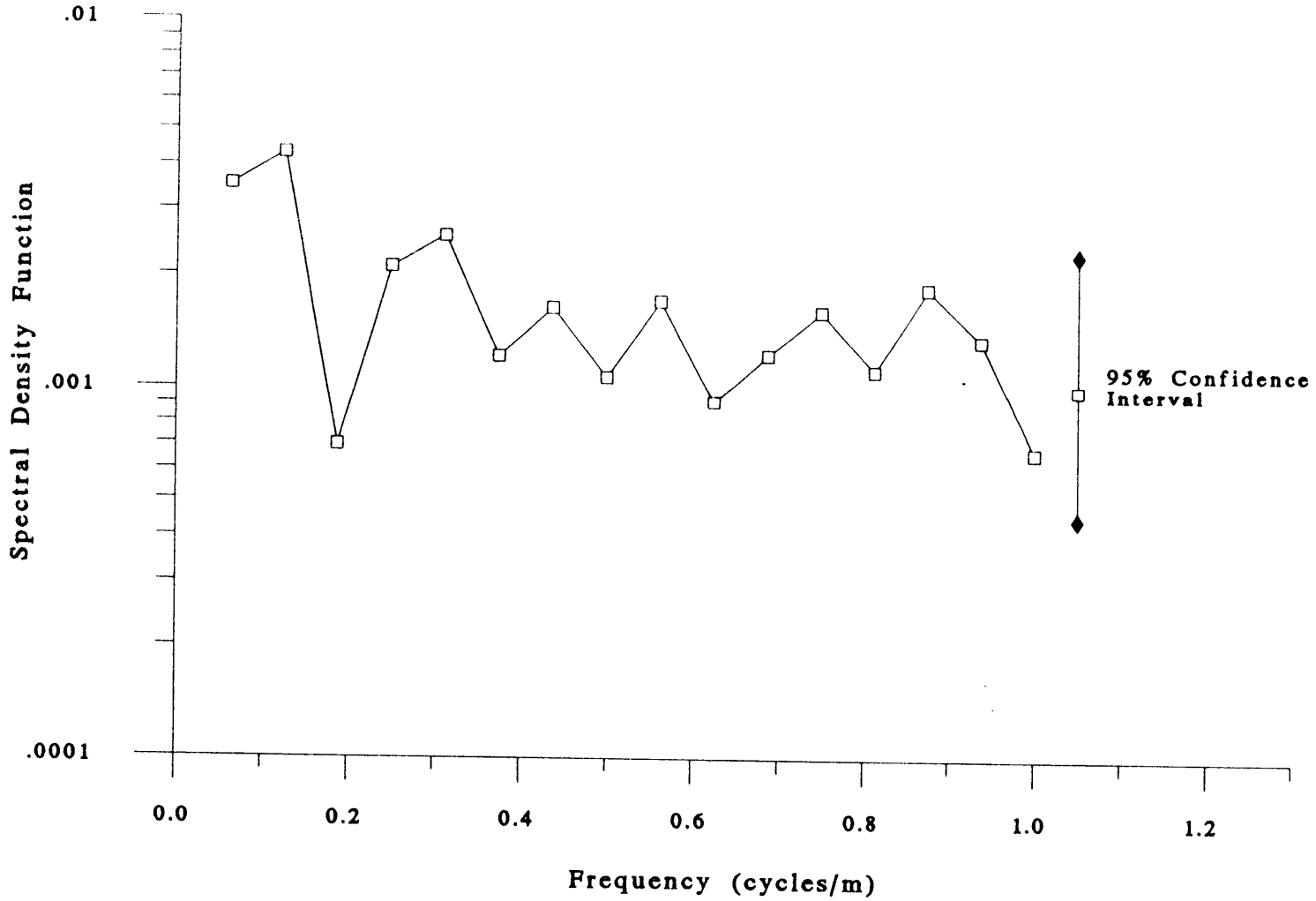

Figure 3.25 The spectral estimate of ln$K$, with M = 16, for the horizontal transect using a Hanning window. The 95% confidence interval is estimated as two times the variance [3.36].

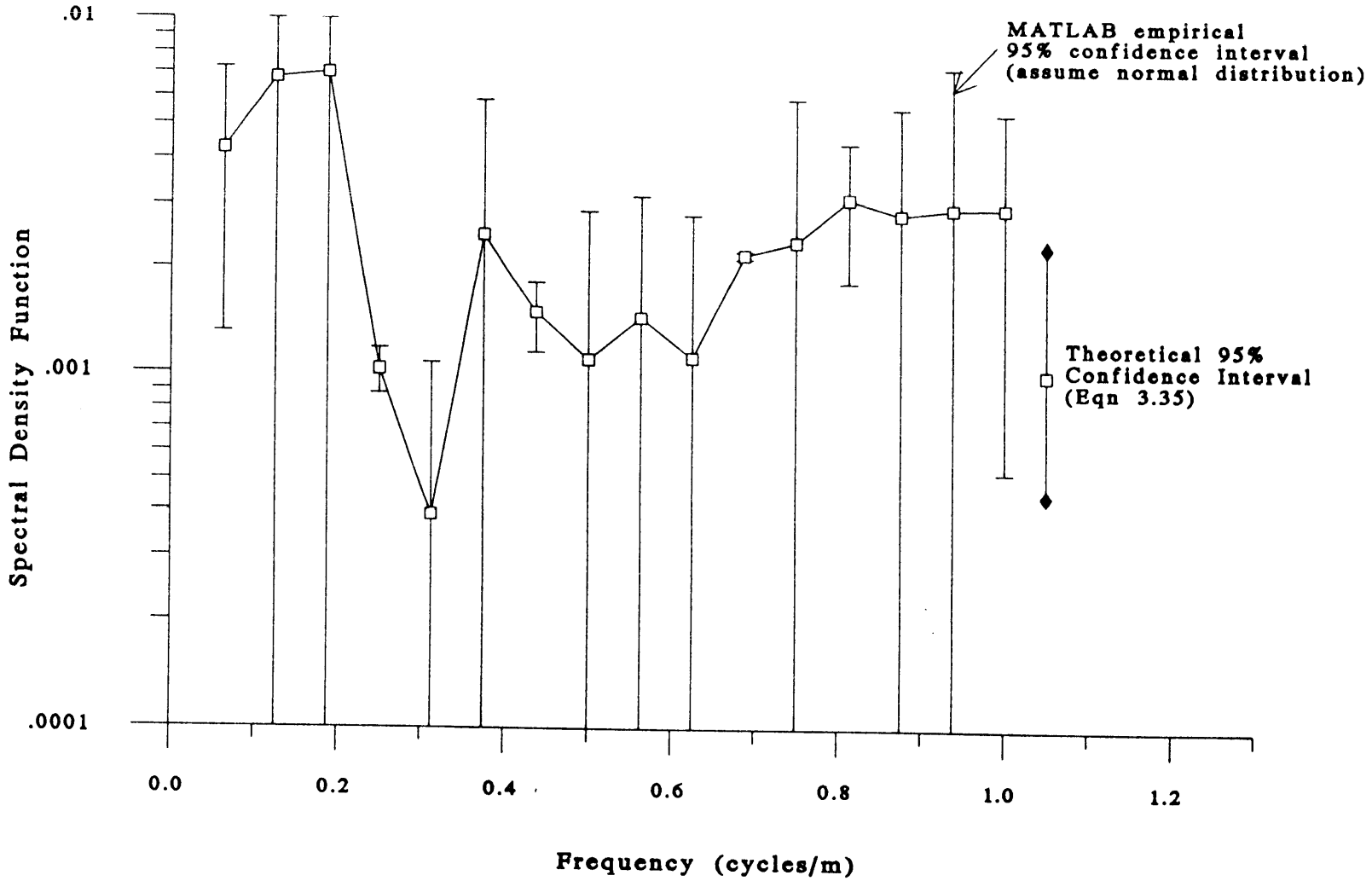

Figure 3.26 The spectral estimate of $\ln K_d$, with M=16, for the horizontal transect. The vertical bands represent the 95% confidence interval which are calculated by MATLAB. The Hanning window was used for this spectral estimate.

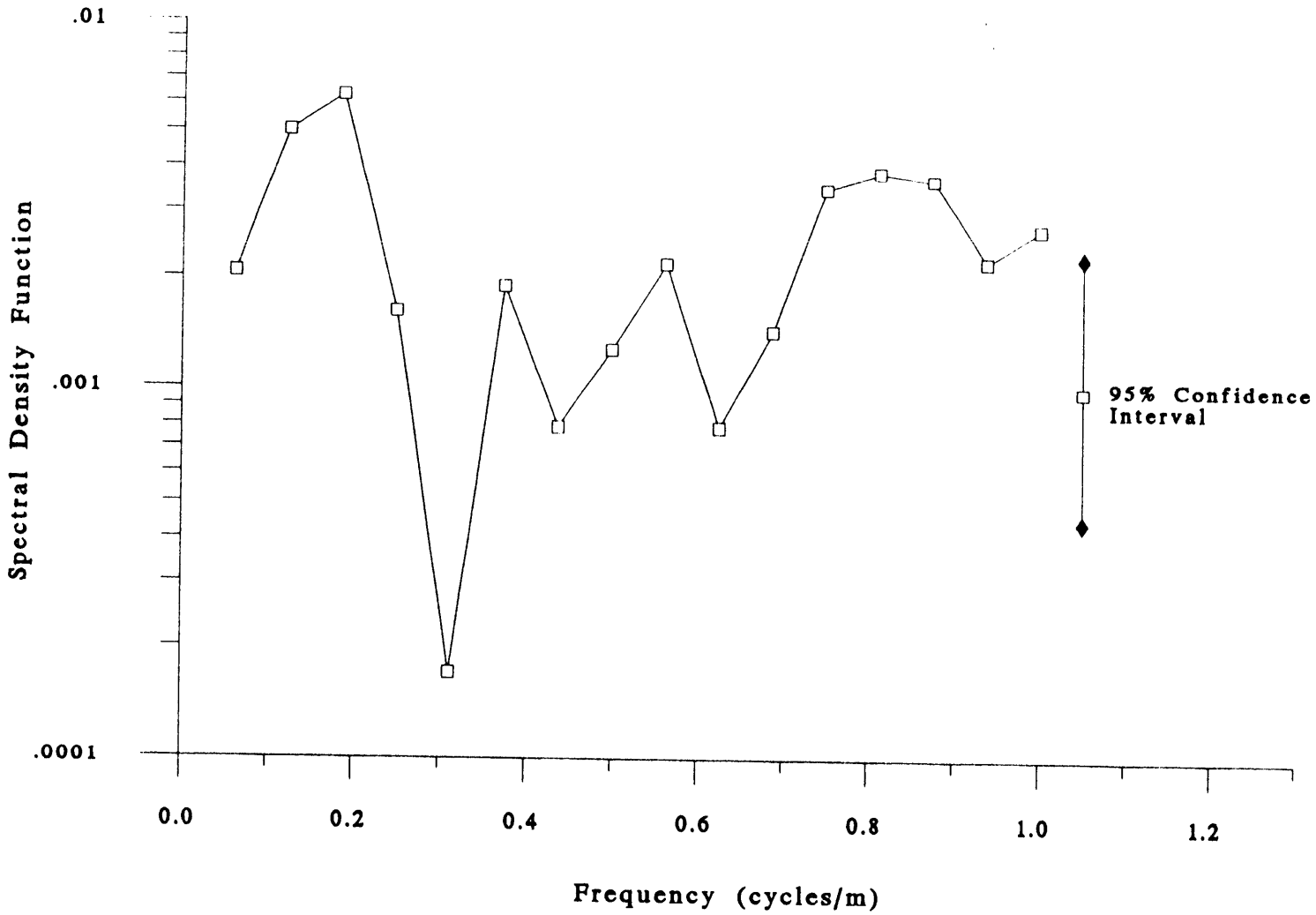

Figure 3.27 The spectral estimate for the residual with M=16, for the horizontal transect.

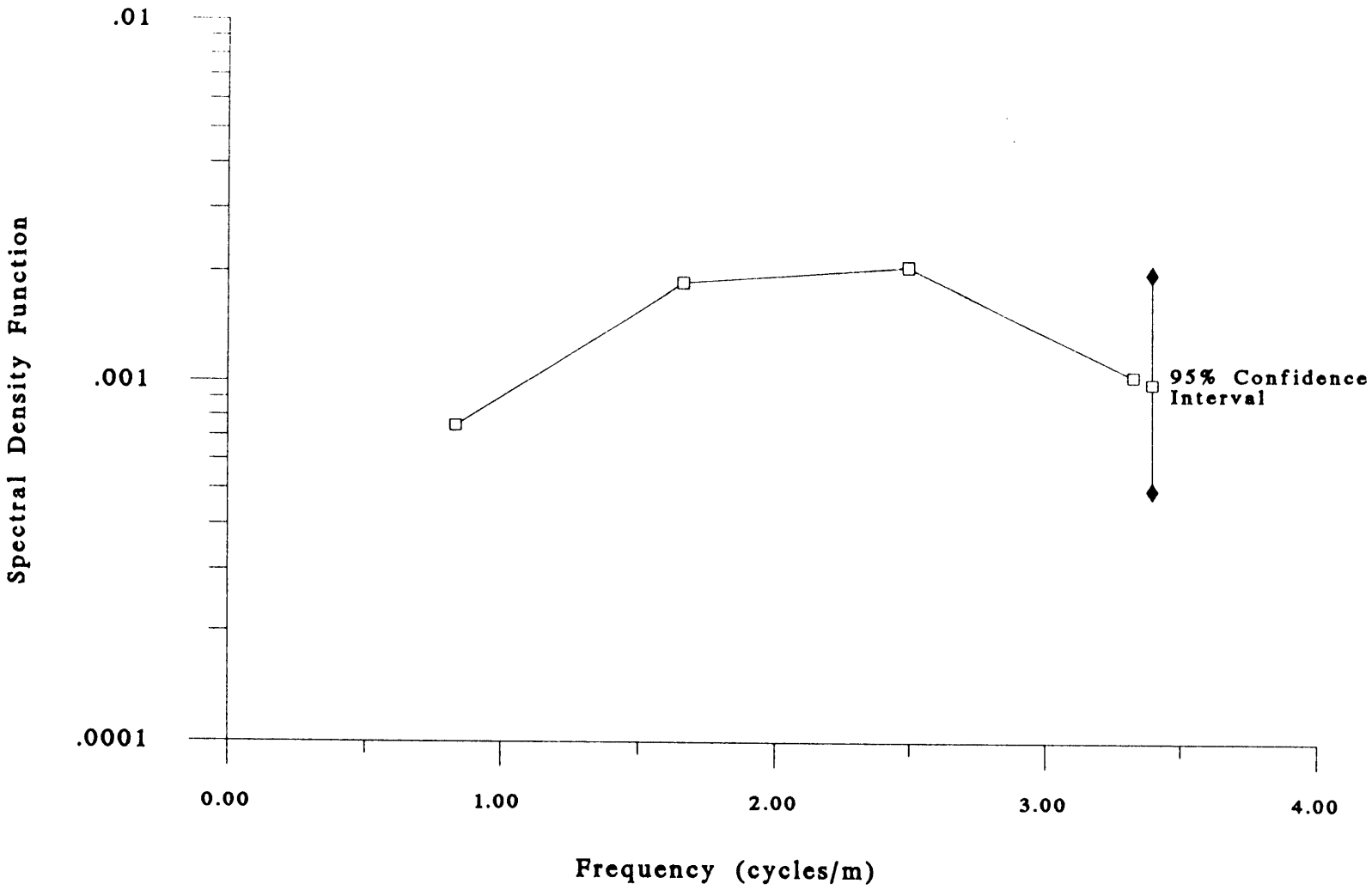

Figure 3.28 The spectral estimate of ln$K$ with M=4, for the vertical transect.

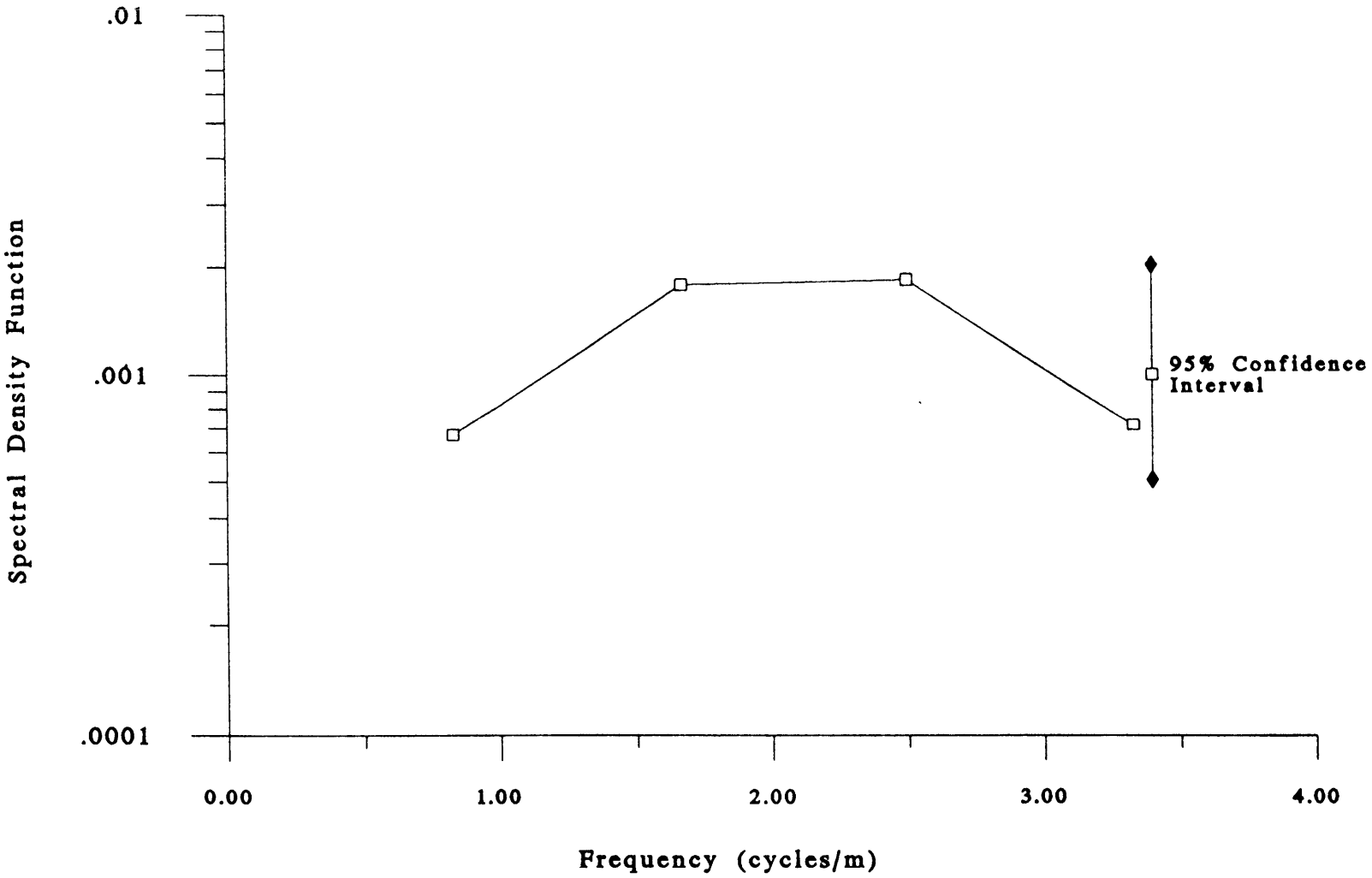

Figure 3.29 The spectral estimate of $\ln K_d$ with M=4, for the vertical transect.

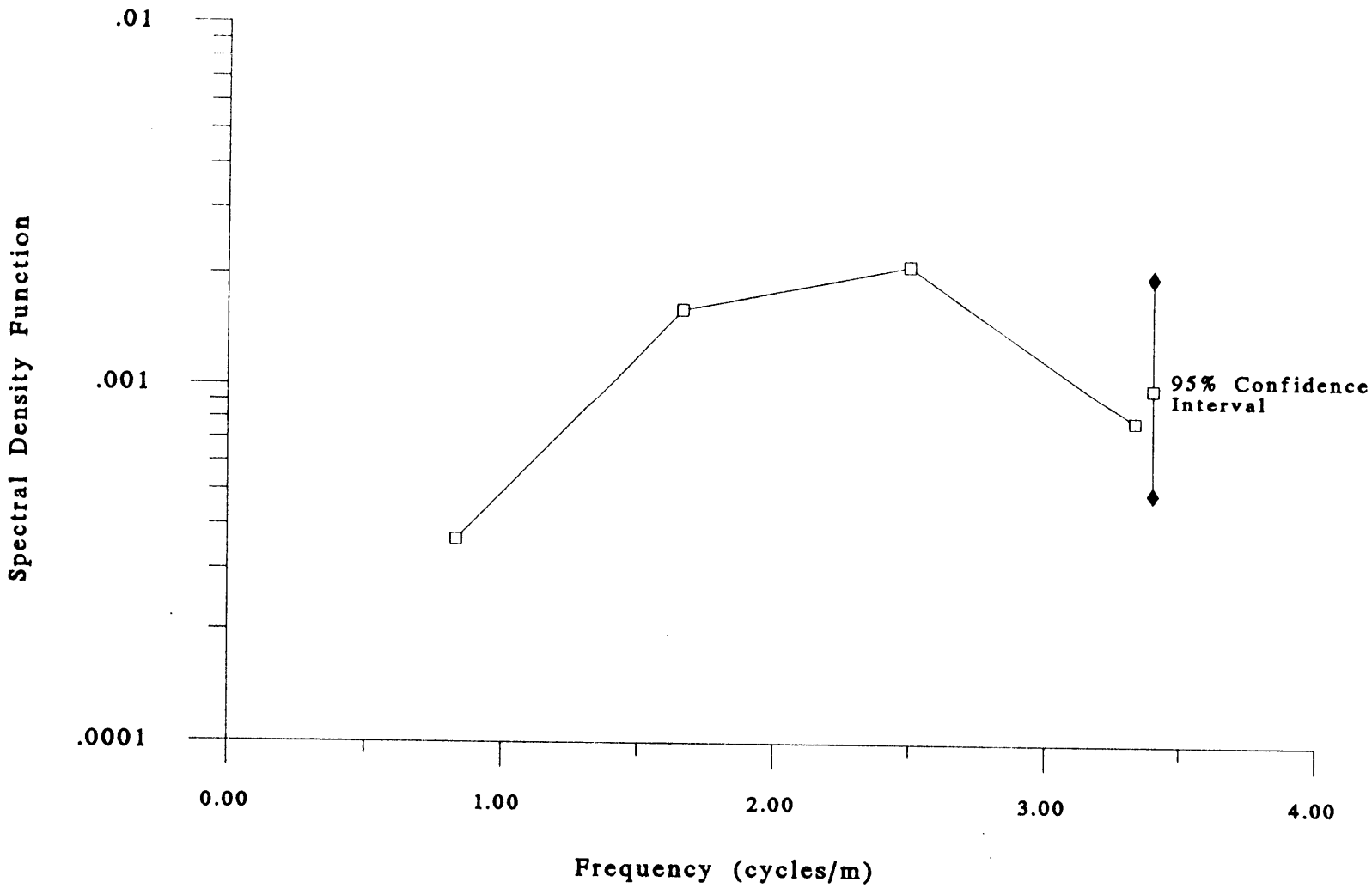

Figure 3.30 The spectral estimate for the residual ($S_{\eta\eta}$) with M = 4, for the vertical transect.

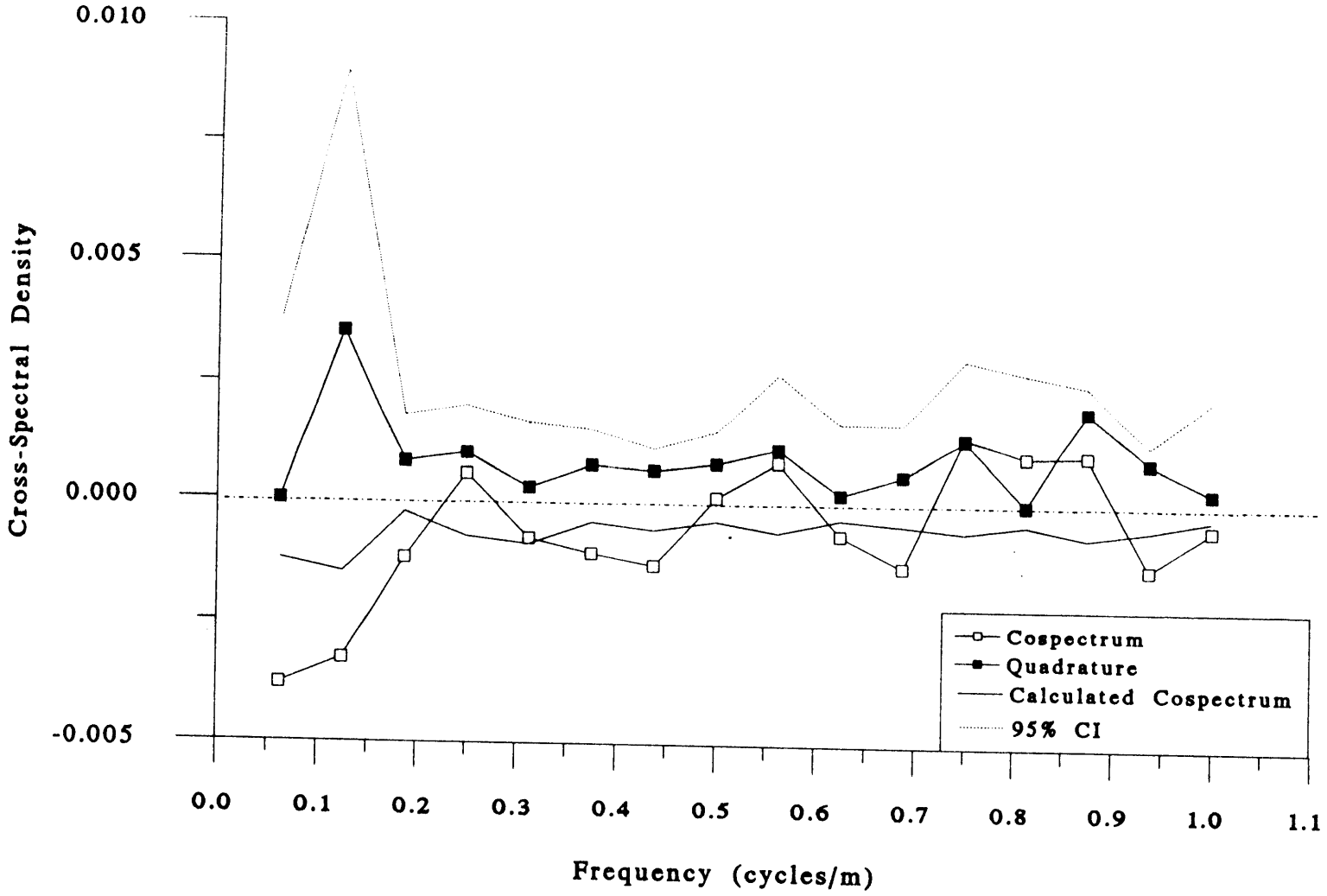

Figure 3.31 Cross-spectrum between $\ln K$ and $\ln K_d$ for the horizontal transect. The fitted cospectrum is estimated as $m\, S_{ff}$ (see Equation [3.52]).

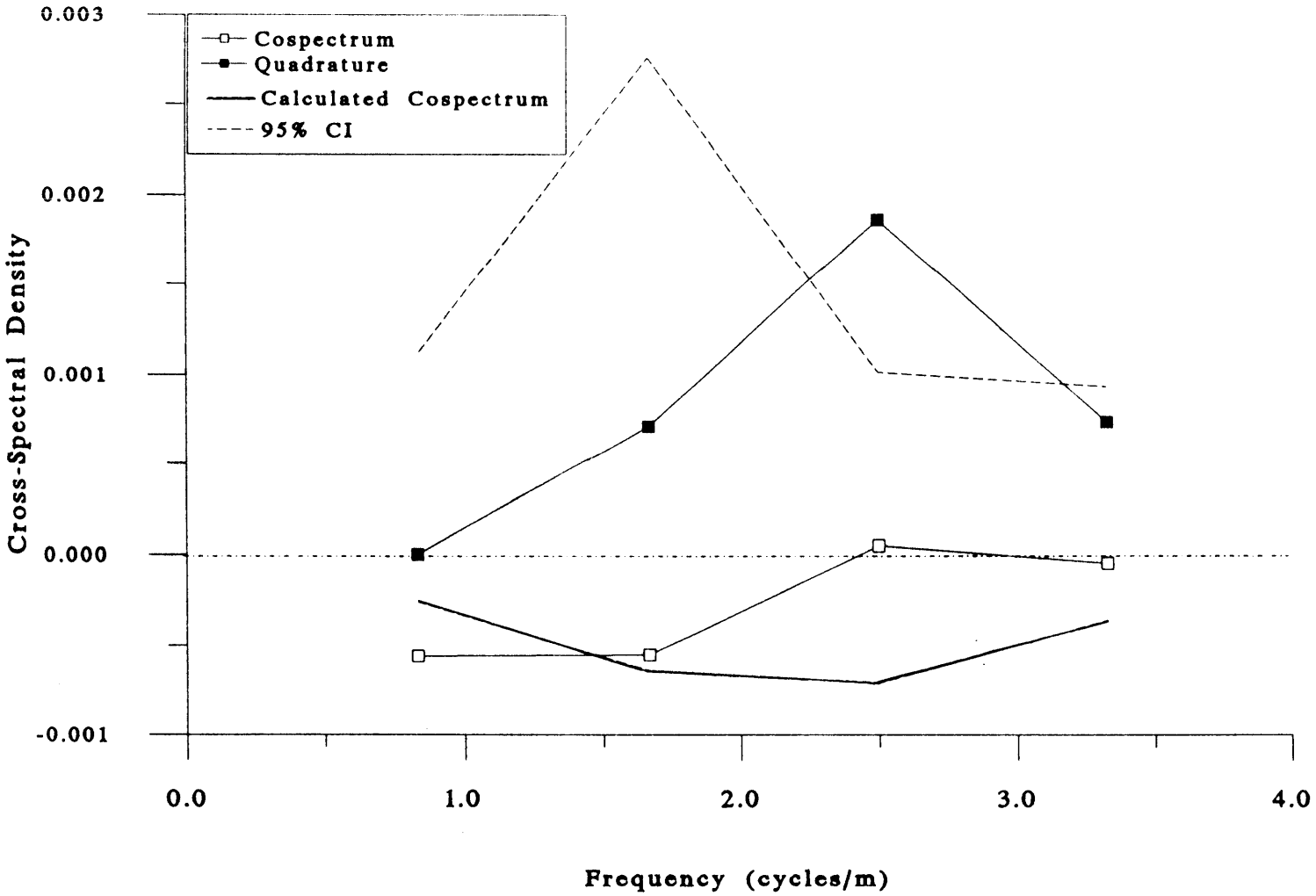

Figure 3.32 Cross-spectrum between $\ln K$ and $\ln K_d$ for the vertical transect. The fitted cospectrum is estimated as $m\, S_{ff}$ (see Equation [3.52]).

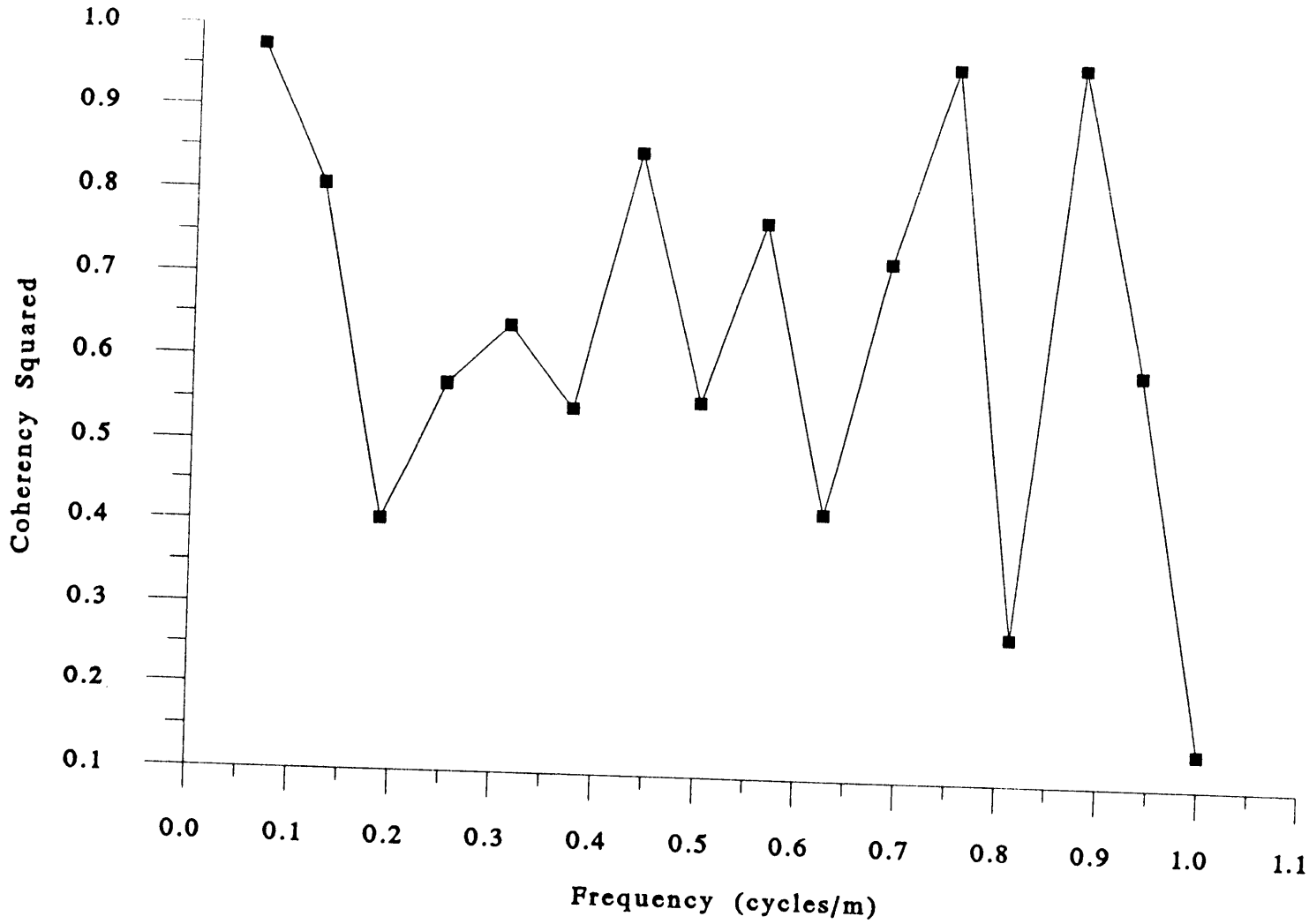

Figure 3.33 Coherency spectrum between ln$K$ and lnK_d for the horizontal transect and M = 16.

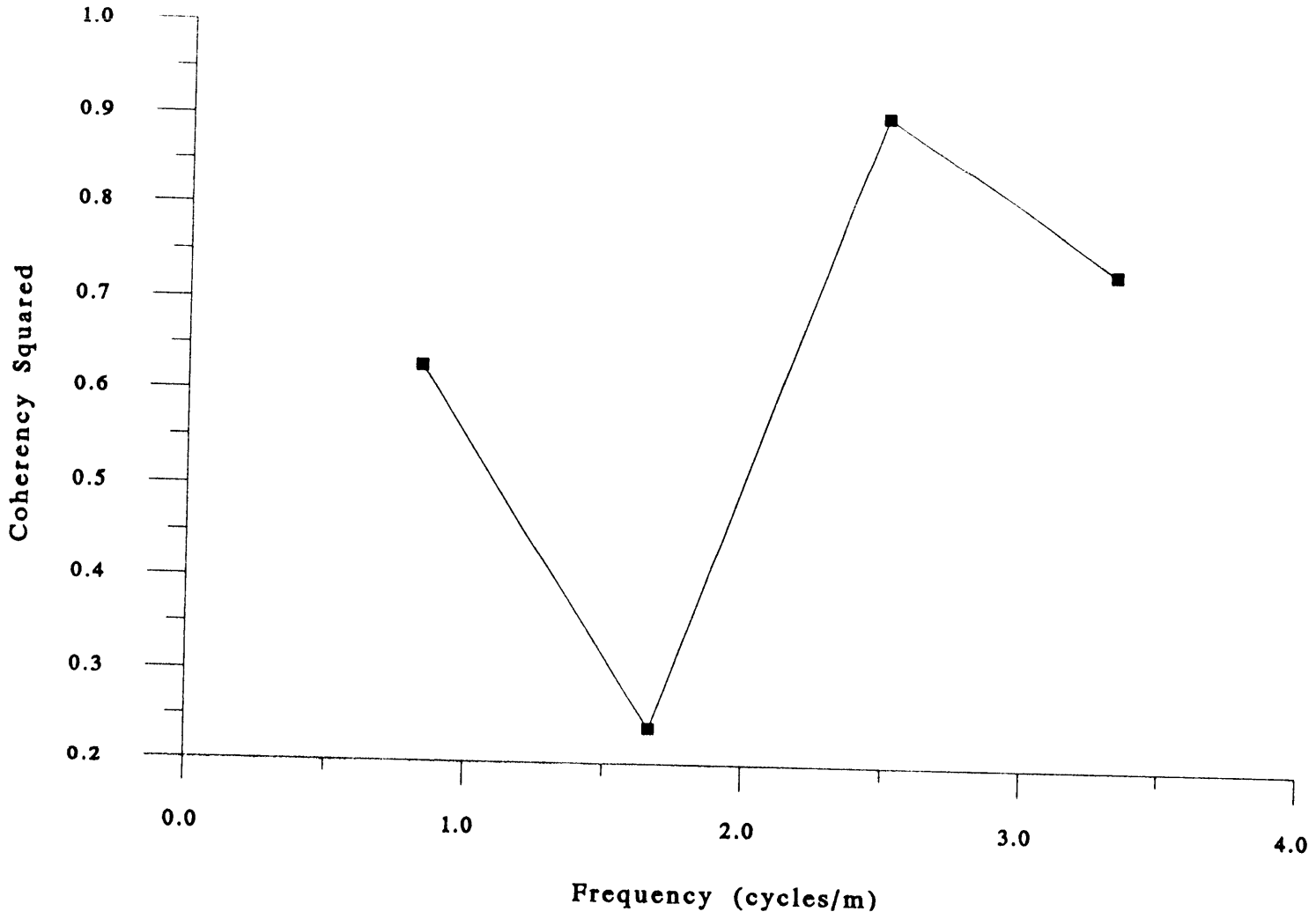

Figure 3.34 Coherency spectrum between $\ln K$ and $\ln K_d$ for the vertical transect and M = 8.

3.3 MODELS

In order to characterize the relationship between $\ln K$ and $\ln K_d$, a simple model was examined. This model is similar to the model in [1.6]. This linear regression model, termed the regression model, assumes that there is a linear relationship between $\ln K$ and $\ln K_d$. The relationship between these two variables is expressed as the equation of the scatter plot shown in Figure 3.10:

$$\ln K_d = m \ln K + c + \eta \qquad [3.37]$$

where m is the slope of the regression of the curve of $\ln K_d$ onto $\ln K$, c is the y-intercept and η is the deviation from the regression curve or the residual. In theory, the residual is a zero mean process.

The effectiveness of the model is contingent upon a number of simple assumptions. The variables, $\ln K$ and $\ln K_d$, can be expressed as zero mean processes:

$$\begin{aligned} f &= \ln K - E[\ln K] \\ g &= \ln K_d - E[\ln K_d] \end{aligned} \qquad [3.38]$$

The mean of [3.37] is:

$$E[K_d] = mE[\ln K] + c \qquad [3.39]$$

The residual is a zero mean process, thus the η term is not represented in this equation. The perbutation of [3.37] yields a form of the model:

$$g = mf + \eta \qquad [3.40]$$

It is easily shown that linear regression produces residuals which are, by

identity, uncorrelated with the dependent variable ($\ln K$). Test I of the regression model examines the correlation between f and η.

Test I: *Correlation between the f and η.* The cross-correlation functions (ccf) between f and η are shown in Figures 3.35 and 3.36 (page 102 and 103). The values of this cross-correlation function is less than the values of the cross-correlation function between $\ln K$ and $\ln K_d$ (Figures 3.13 and 3.16). The significance of the correlation can only be judged in the spectral domain where confidence intervals can be placed on the results.

The plots of the cross-spectrum (Figures 3.37 and 3.38) show that the cospectrum and quadrature spectrum tend to be less than the 95% confidence interval calculated by MATLAB for all frequencies[1]. The coherency squared spectrum (Figure 3.39 and 3.40) shows that the correlation between f and η for the horizontal and vertical transect, the coherency squared is similar to coherency squared of $\ln K$ and $\ln K_d$. Since no definitive information can be gathered from the cross-spectra and coherency squared plots, the magnitude of the cross-spectrum of f and η is examined.

The cross-spectrum between η and f is:

$$S_{\eta f}=A_{\eta f}(\omega)e^{-i\theta_{\eta f}(\omega)} \qquad [3.41]$$

where $A_{\eta f}$ is the magnitude of the cross-spectrum and $\theta_{\eta f}$ is the phase spectrum. If $A_{\eta f}$ is equivalent to zero, the f and η is not correlated. Figure 3.41 and 3.42 shows

[1] The output of the MATLAB program for the cross spectrum between f and η is found in Appendix C.

that, according to the 95% confidence interval[8], the magnitude of the cross-spectrum is may be equivalent to zero for most frequencies in the horizontal and vertical transect.

Based on the cross-correlation function, cross-spectrum, coherency squared and the magnitude of the cross-spectrum, it is difficult to ascertain a definitive conclusion about the correlation between f and η using covariance and spectral analysis.

In order to analyze the adequacy of this model, the results of the model is compared to the results from the sampled data. For this comparison, the model is converted to a covariance and spectral representation and four other tests are applied:

Test II: *Cross Covariance Method.* Multiplying [3.40] by $f(x+s)$ and taking the expected value of, we get:

$$E[g(x)f(x+s)] = mE[f(x)f(x+s)] + E[\eta(x)f(x+s)] \qquad [3.42]$$

the last term goes to zero since the η and $\ln K$ are uncorrelated. The resultant equation can be expressed in terms of the covariance function:

$$R_{fg} = mR_{ff} \qquad [3.43]$$

All of the analysis done in this report has been in terms of the autocorrelation and cross correlation function. The correlation functions and covariance functions can be related by the variance. Thus, we will use equation [3.42] in the form:

[8] The calculation of the confidence intervals are shownn in Appendix D.

$$\rho_{fg} = m\frac{\sigma_f}{\sigma_g}\rho_f \qquad [3.44]$$

Figures 3.43 and 3.44 show the cross correlation functions of the horizontal and vertical transects along with the cross correlation function calculated from [3.43] and the fitted exponential covariance [3.23] for R_{ff} as shown in Figures 3.17 and 3.20 (correlation scales from Table 3.1). This fit is representative of a smoothed version of the empirical cross correlation function.

Test III: *Autocorrelation Method.* Multiplying [3.40] by the "lagged" distribution coefficient, $g(x+s)$, we get:

$$\begin{aligned} E[g(x)g(x+s)] &= E\ [mf(x)+\eta(x)]\,[mf(x+s)+\eta(x+s)] \\ &= m^2 E[f(x)f(x+s)] + E[\eta(x)\eta(x+s)] \end{aligned} \qquad [3.45]$$

where the cross-correlation between f and η is assumed as negligible. This equation can be expressed in terms of the autocovariance function, and subsequently the autocorrelation function as:

$$\begin{aligned} R_g &= m^2 R_f + R_\eta \\ \rho_g &= m^2 \frac{\sigma_f^2}{\sigma_g^2}\rho_f + \frac{\sigma_\eta^2}{\sigma_g^2}\rho_\eta \end{aligned} \qquad [3.46]$$

Figure 3.45 and 3.46 shows the empirical and the fitted of the autocorrelation function of $\ln K_d$ for the horizontal and vertical transect; the fitted autocorrelation function calculated using [3.45]. This fit provides a smoothed estimate of the autocorrelation function with similar correlation lengths.

Test IV: *Cross-Spectrum.* The spectral representation of the sorption coefficient, hydraulic conductivity and residual can be expressed as:

$$g(x)=\int e^{i\omega x}dZ_g(\omega) \qquad [3.47]$$

$$f(x)=\int e^{i\omega x}dZ_f(\omega) \qquad [3.48]$$

$$\eta(x)=\int e^{i\omega x}dZ_\eta(\omega) \qquad [3.49]$$

Equation [3.40] can be expressed as a Fourier-Stieltjes integral:

$$dZ_g=m\,dZ_f+dZ_\eta \qquad [3.50]$$

This equation multiplied by dZ_f^* leads to an equation of the cross-spectrum:

$$E[dZ_g dZ_f^*]=S_{fg}\,d\omega=m\,E[dZ_f dZ_f^*]+E[dZ_\eta dZ_f^*] \qquad [3.51]$$

The last term goes to zero, thus:

$$S_{gf}d\omega=m\,S_{ff}(\omega)d\omega \qquad [3.52]$$

The spectrum, $S_{ff}(\omega)$ is real; according to [3.51], $S_{gf}(\omega)$, must also be real.

The complex cross spectra can also be written as:

$$S_{gf}=A_{gf}(\omega)e^{-i\theta_{gf}(\omega)}S_{ff} \qquad [3.53]$$

where A_{gf} is the magnitude of the cross-spectrum between g and f and θ_{gf} is the phase. If the cross-spectrum is real, the value of the phase, θ_{gf}, must be equal to zero. A phase of zero implies that there is no phase shift. Figures 3.47 through 3.50 show the plots of the tangent of the phase and the magnitude for the horizontal and vertical transect. For both transects, the phase fluctuates around zero. Judging from the confidence interval, most of the fluctuations are not significantly different from zero. This null phase means that there is no phase shift, therefore the fourth test is

successful. Figure 3.31 and 3.32 show the cross-spectrum calculated from [3.52] using the ln$K$ spectral estimates of Figures 3.25 and 3.28.

Test V: *Spectrum.* If [3.40] is multiplied by dZ_g^*, the following equation results:

$$E[dZ_g\ dZ_g^*]=S_{gg}dk=m^2S_{ff}(k)dk+S_{\eta\eta}(k)dk \qquad [3.54]$$

The spectrum calculated from [3.53] and the spectral estimates for f and η (Figures 3.37 and 3.38) are transposed unto the estimated spectrum in Figures 3.51 and 3.52. For both transects, the spectrum calculated by the model exhibits the same behavior as the estimated spectrum. The differences in the empirical spectral density function and that calculated by [3.18] is less than the 95% confidence interval. The regression model produces spectral estimates that are statistically equivalent to the emperical spectral estimates.

The five test indicate that the simple regression model provides a suitable first approximation of the relationship between the ln$K$ and lnK_d data.

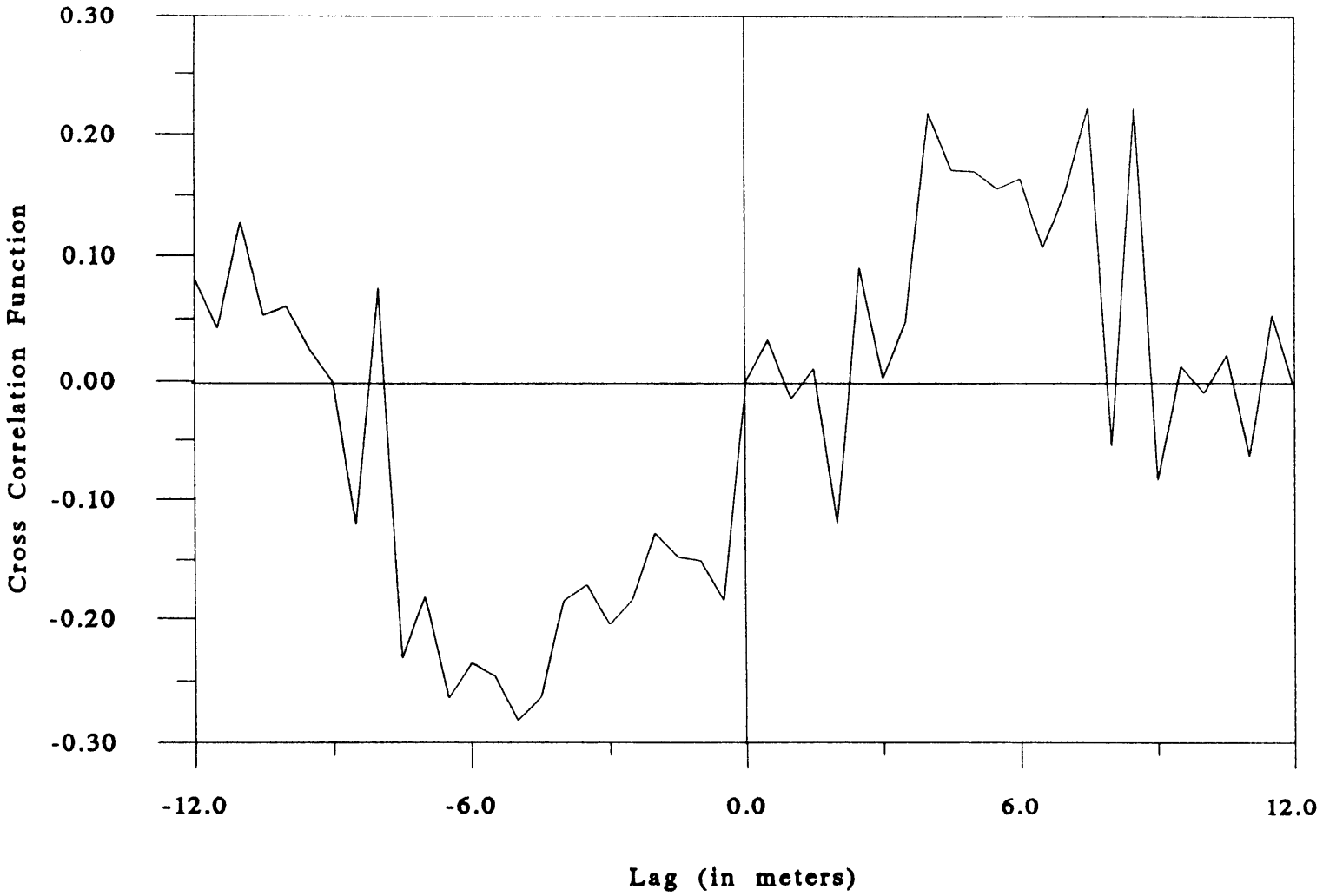

Figure 3.35 Test I of the linear regression model: Cross-correlation function of f and η for the horizontal transect.

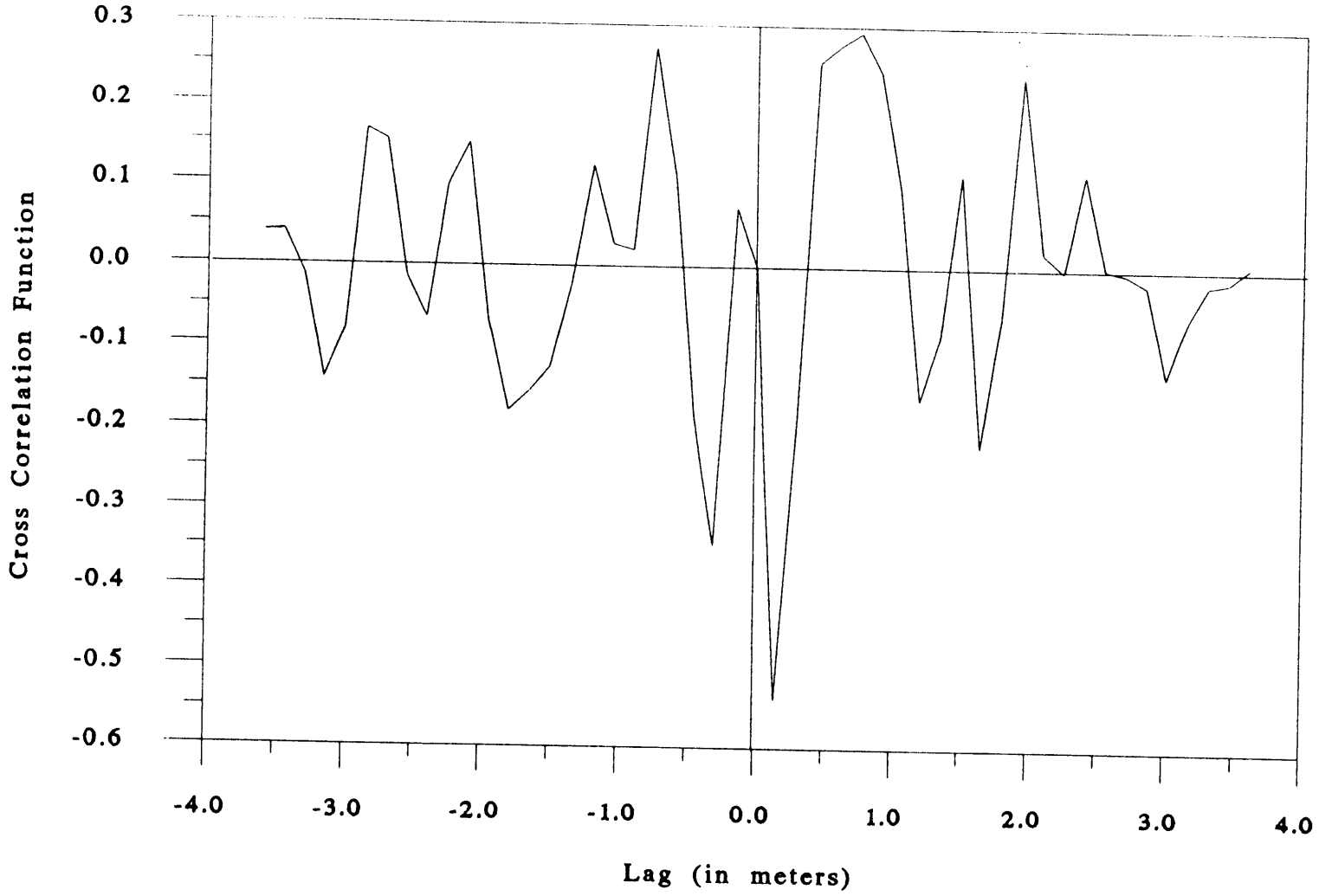

Figure 3.36 Test I of the linear regression model: Cross-correlation function of f and η for the vertical transect.

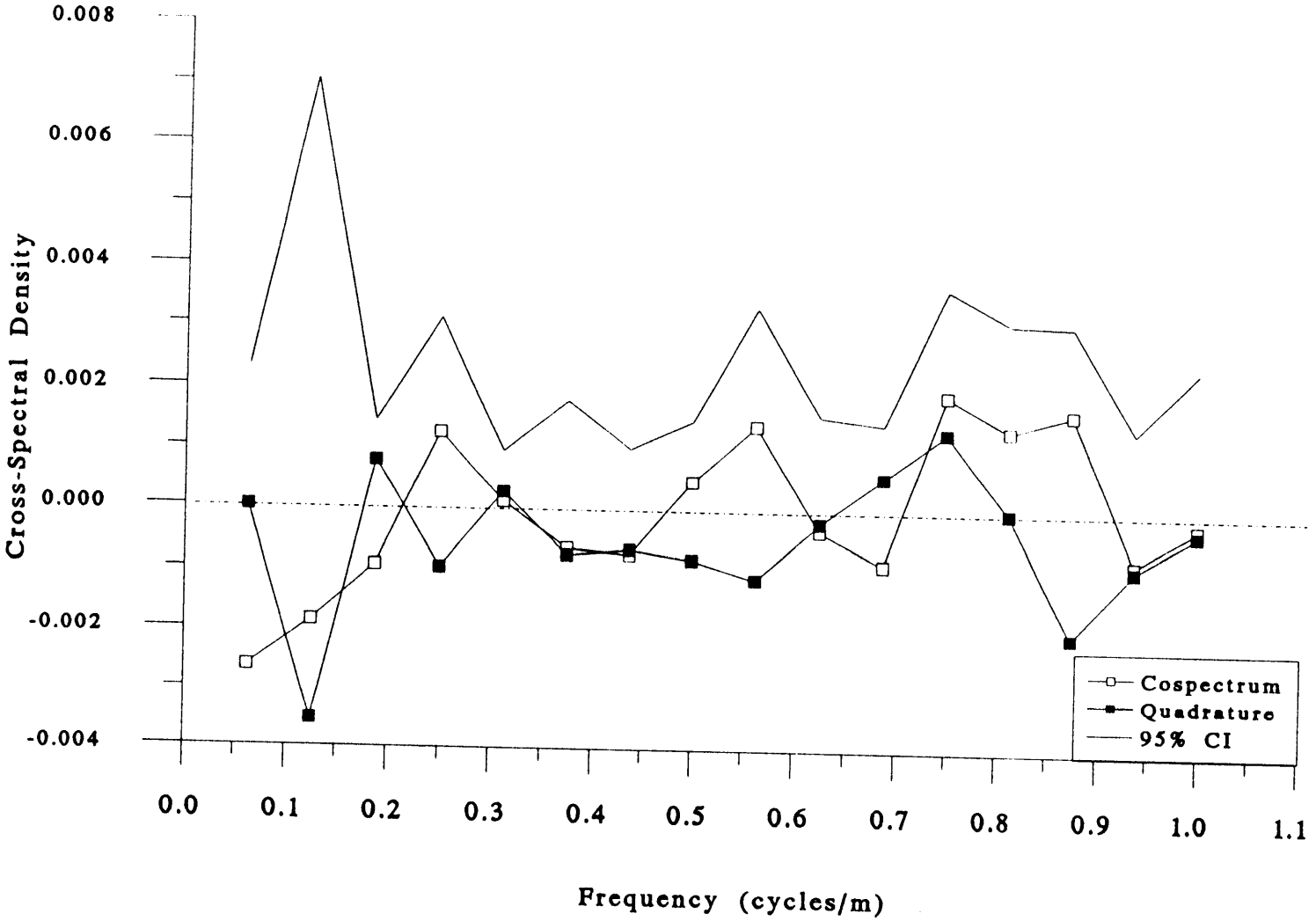

Figure 3.37 Test I of the linear regression model: Cross-spectrum between f and η for the horizontal transect with $M = 16$.

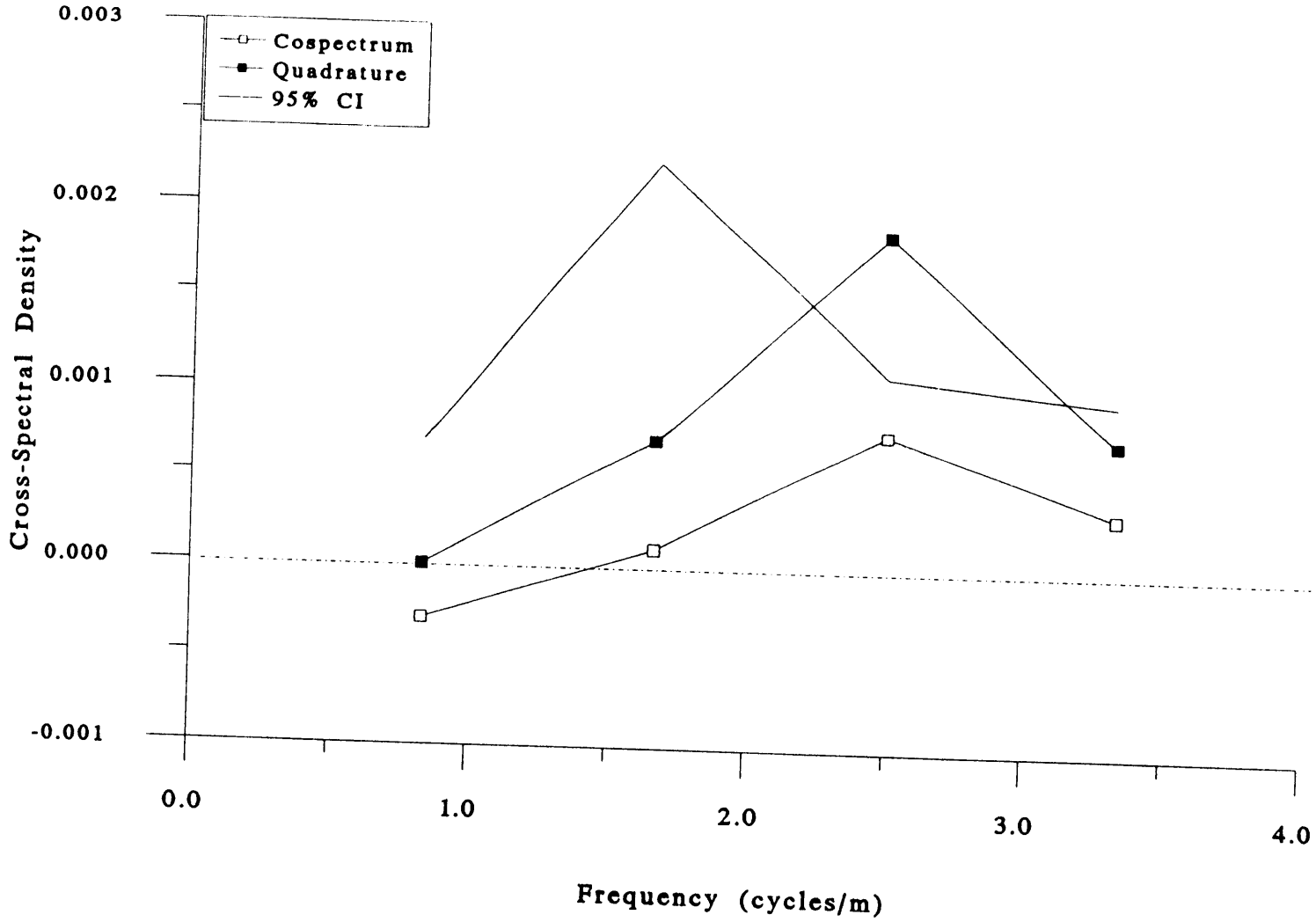

Figure 3.38 Test I of the linear regression model: Cross-spectrum between f and η for the vertical transect with $M = 4$.

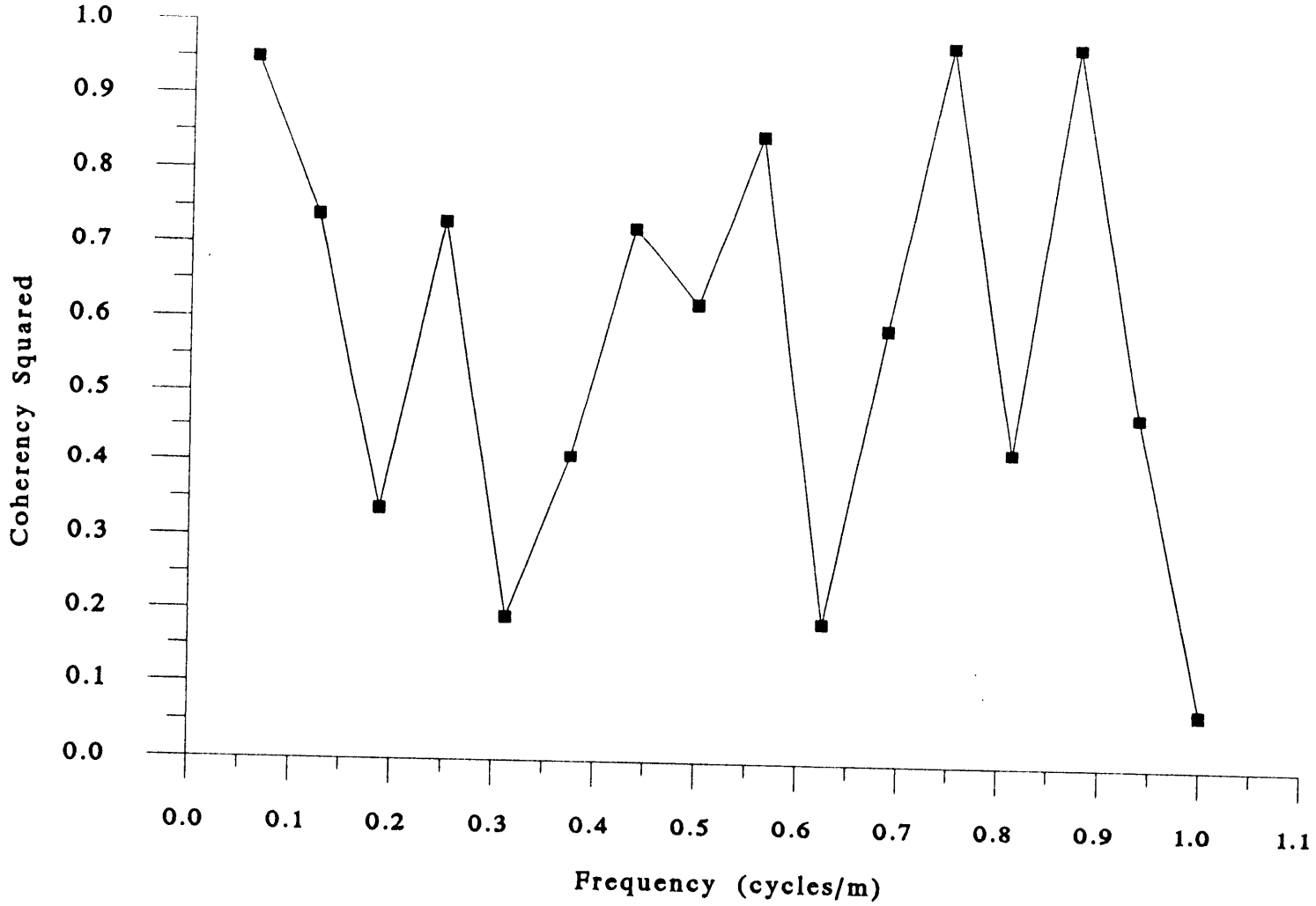

Figure 3.39 Test I of the linear regression model: Coherency squared between f and η for the horizontal transect with $M = 16$.

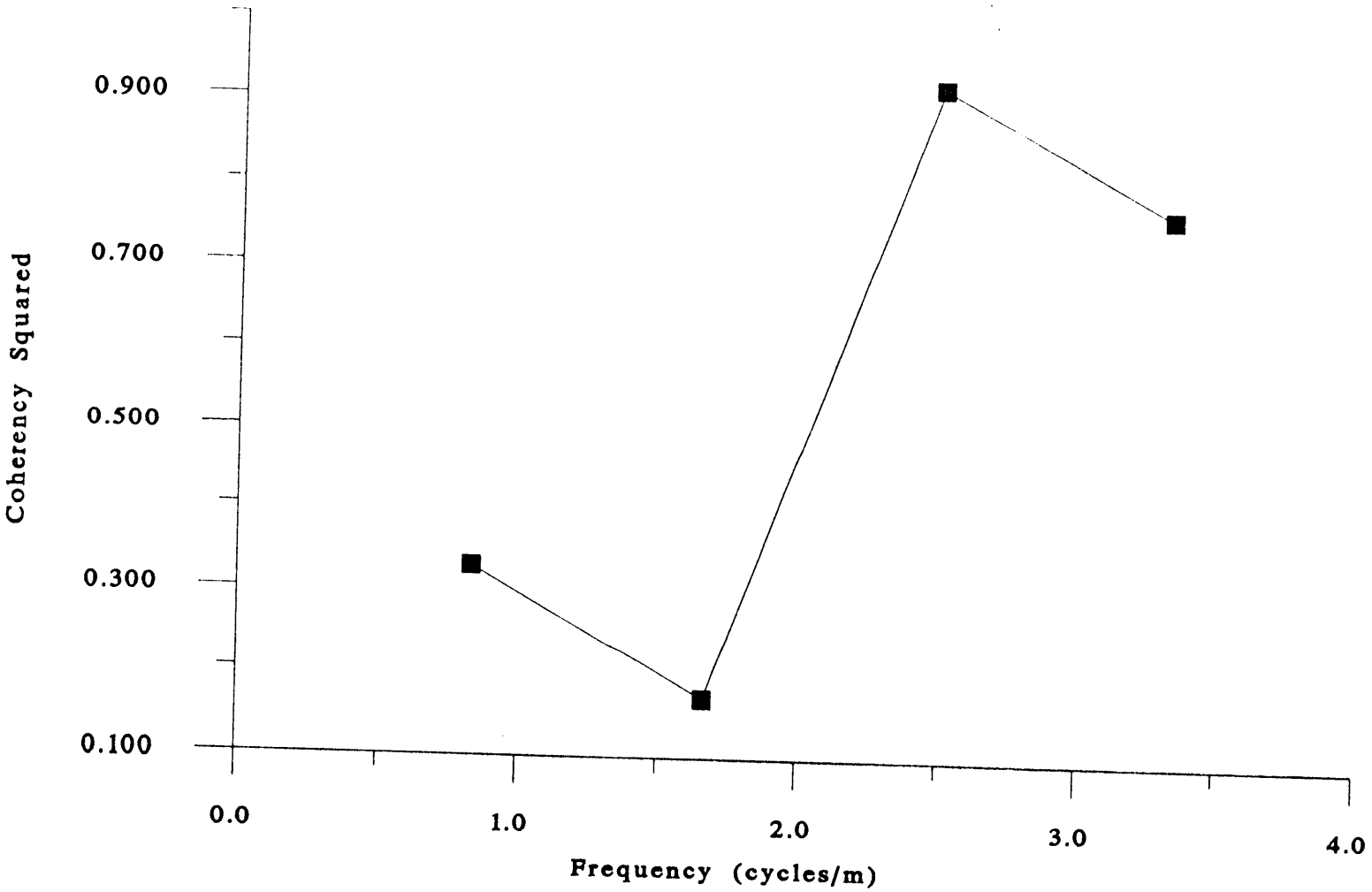

Figure 3.40 Test I of the linear regression model: Coherency squared between f and η for the vertical transect with $M = 4$.

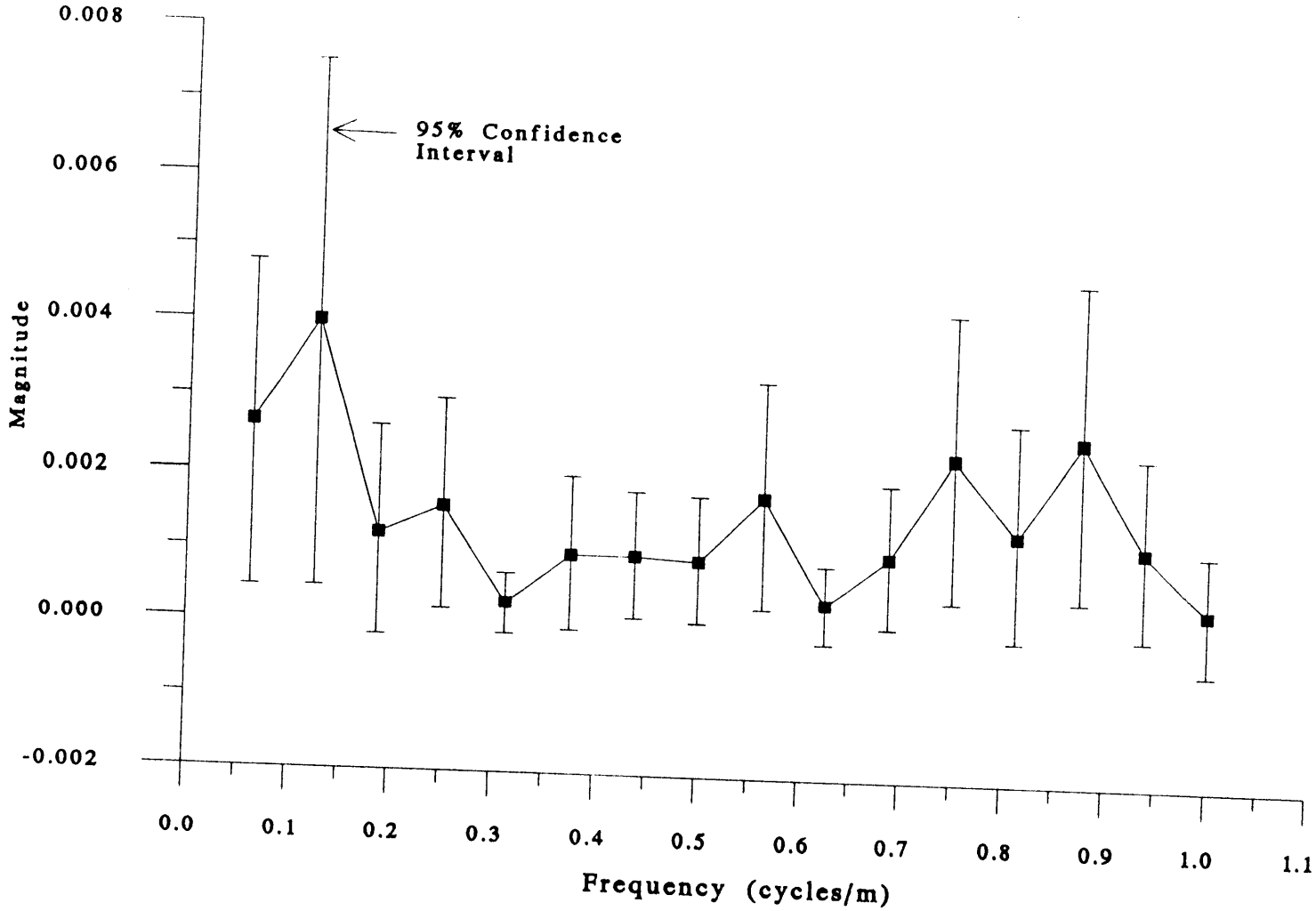

Figure 3.41 Test I of the linear regression model: Magnitude of the cross-spectrum between f and η for the horizontal transect with $M = 16$.

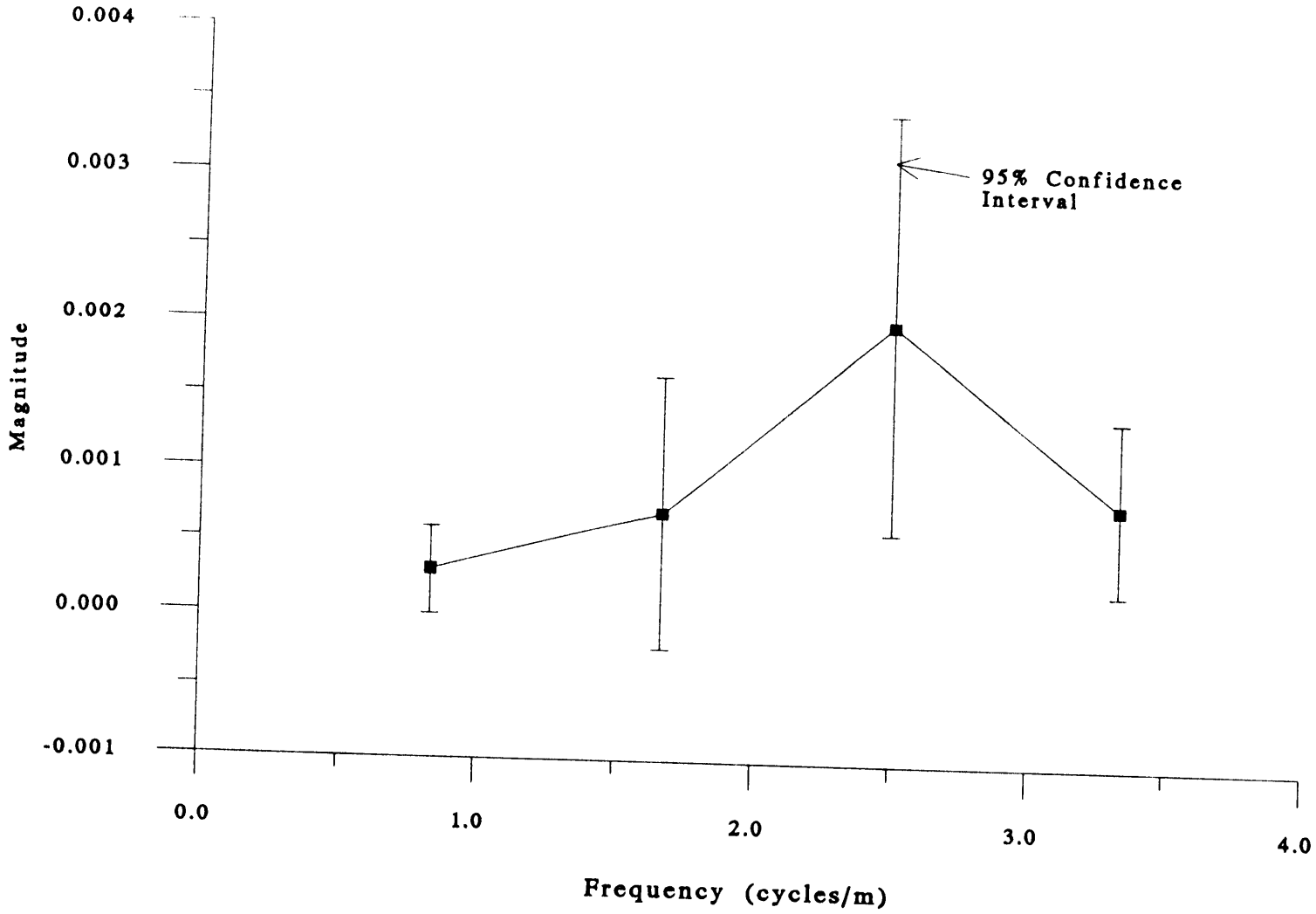

Figure 3.42 Test I of the linear regression model: Magnitude of the cross-spectrum between f and η for the vertical transect with $M = 4$.

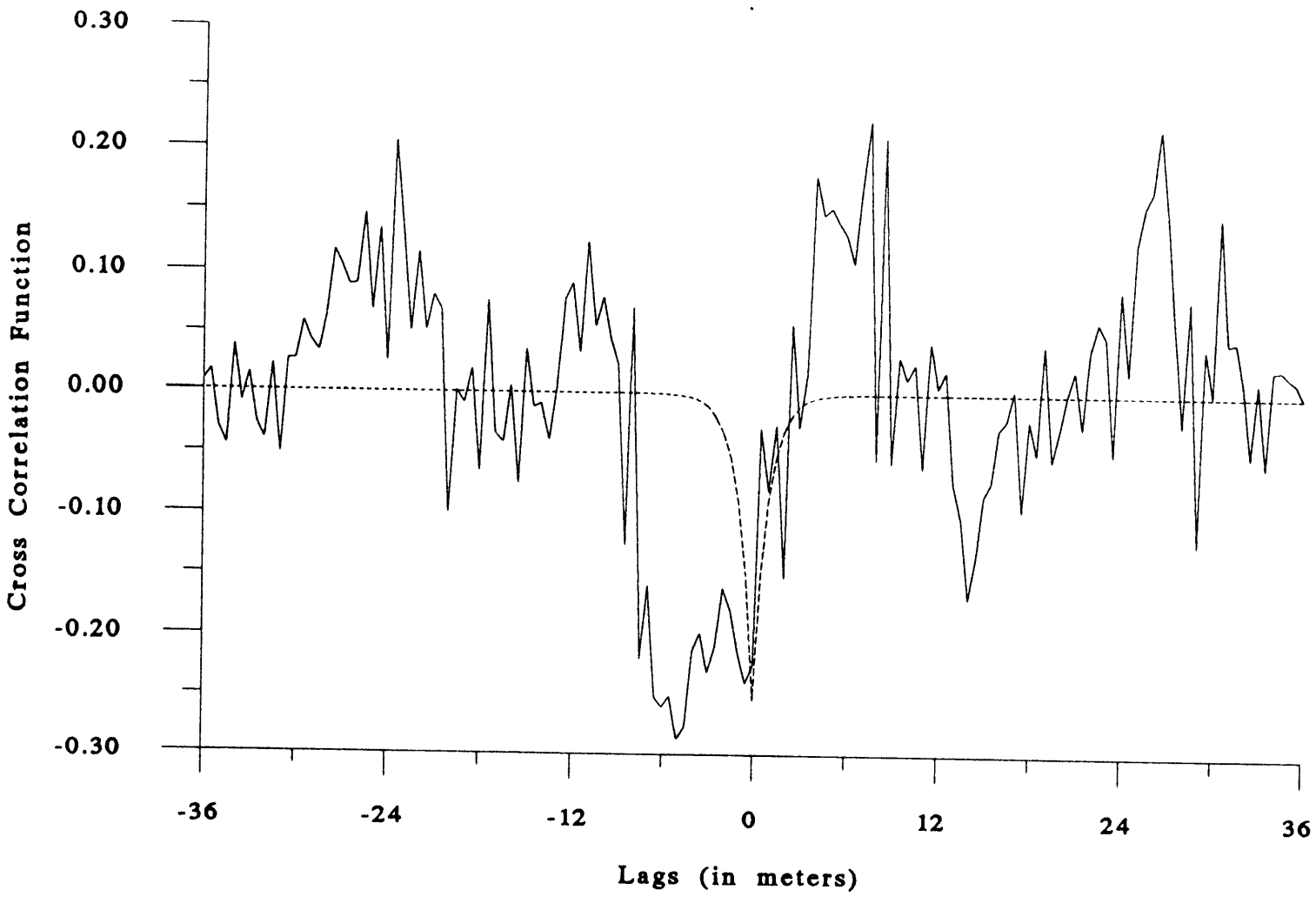

Figure 3.43 Test II the linear regression model: Calculated cross-correlation function (dashed lines) for the horizontal transect.

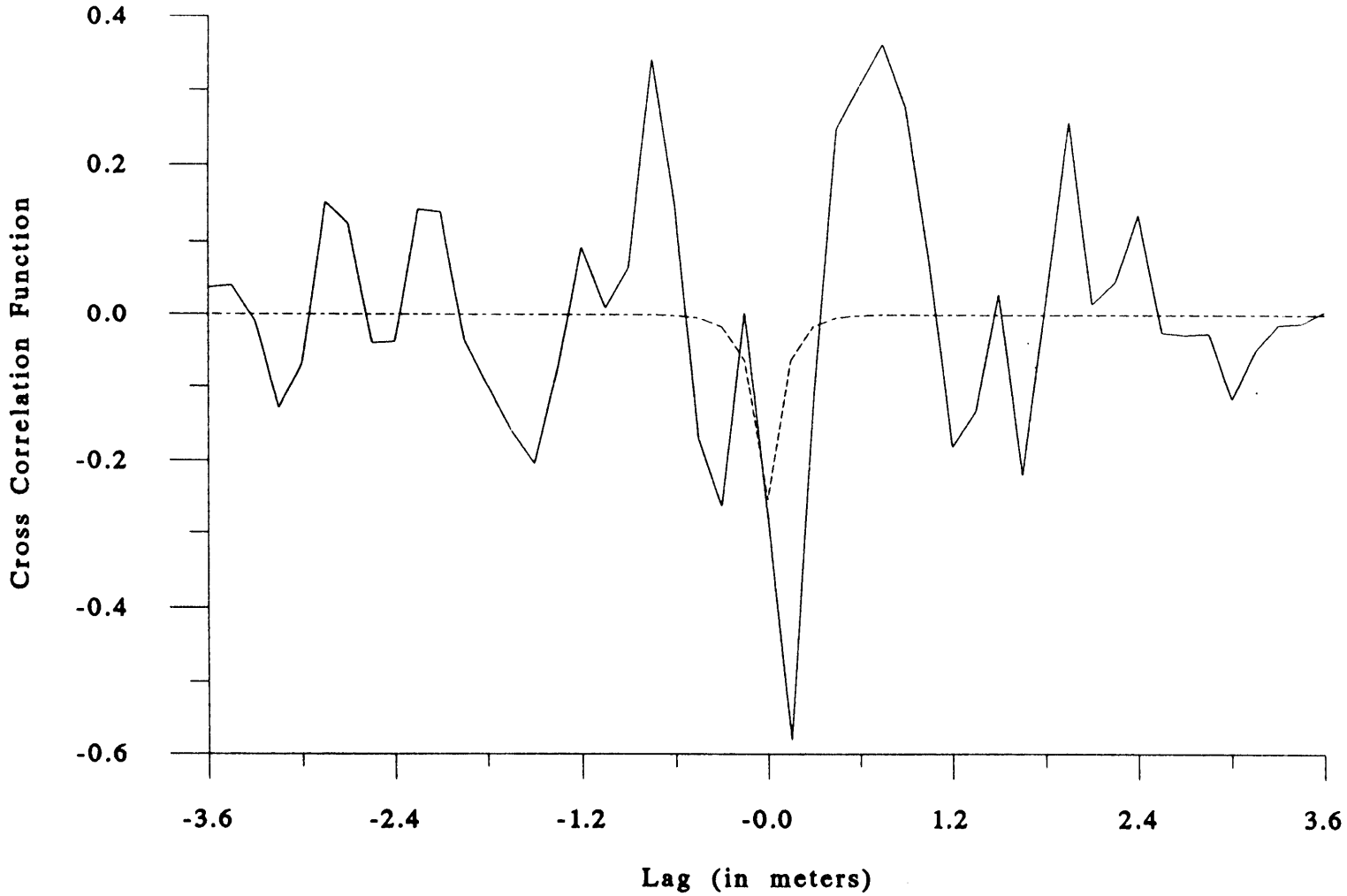

Figure 3.44 Test II of the linear regression model: Calculated cross-correlation function (dashed lines) for the vertical transect.

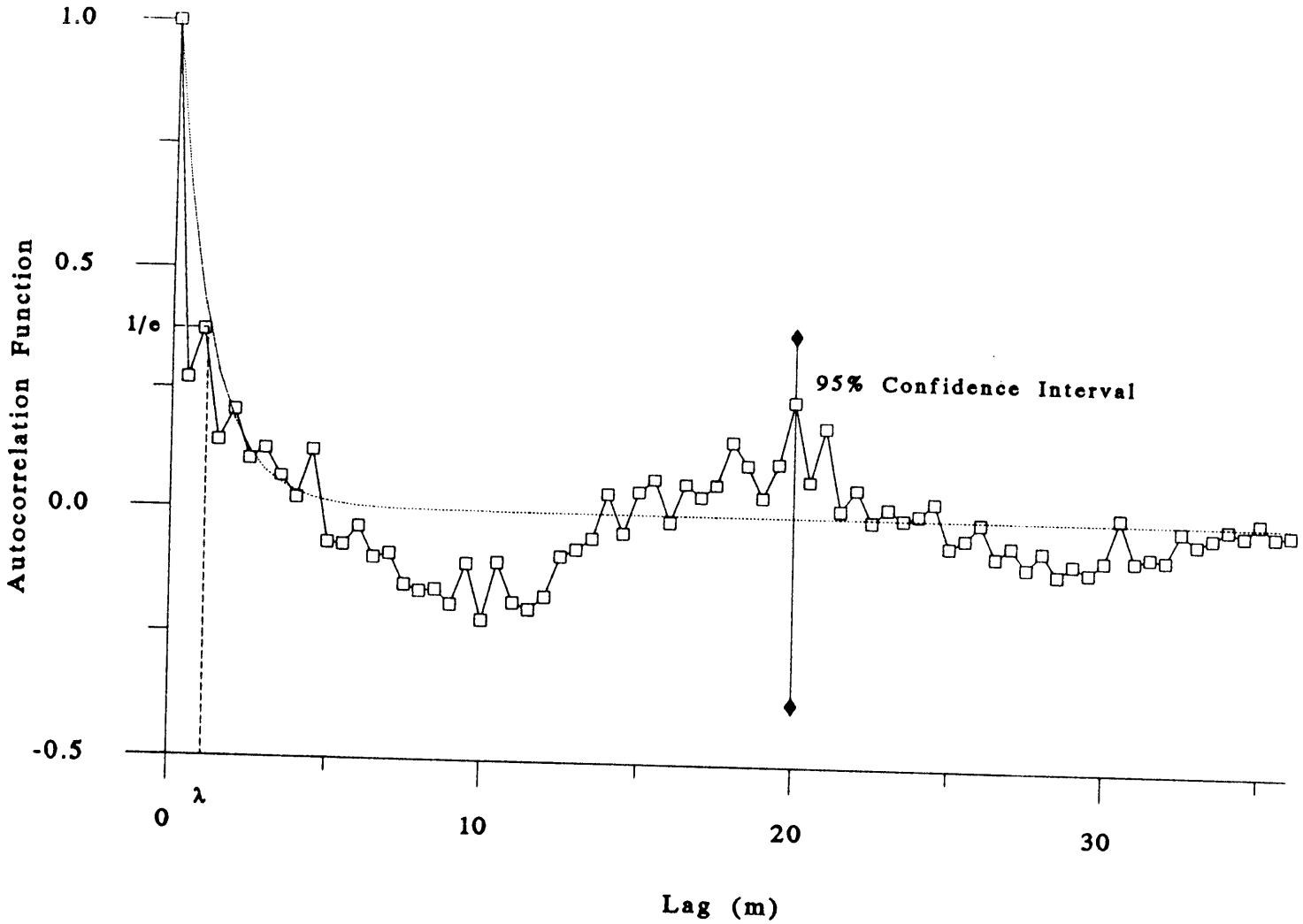

Figure 3.45 Test III of the linear regression model: Calculated autocorrelation function of $\ln K_d$ (dashed lines) for the horizontal transect.

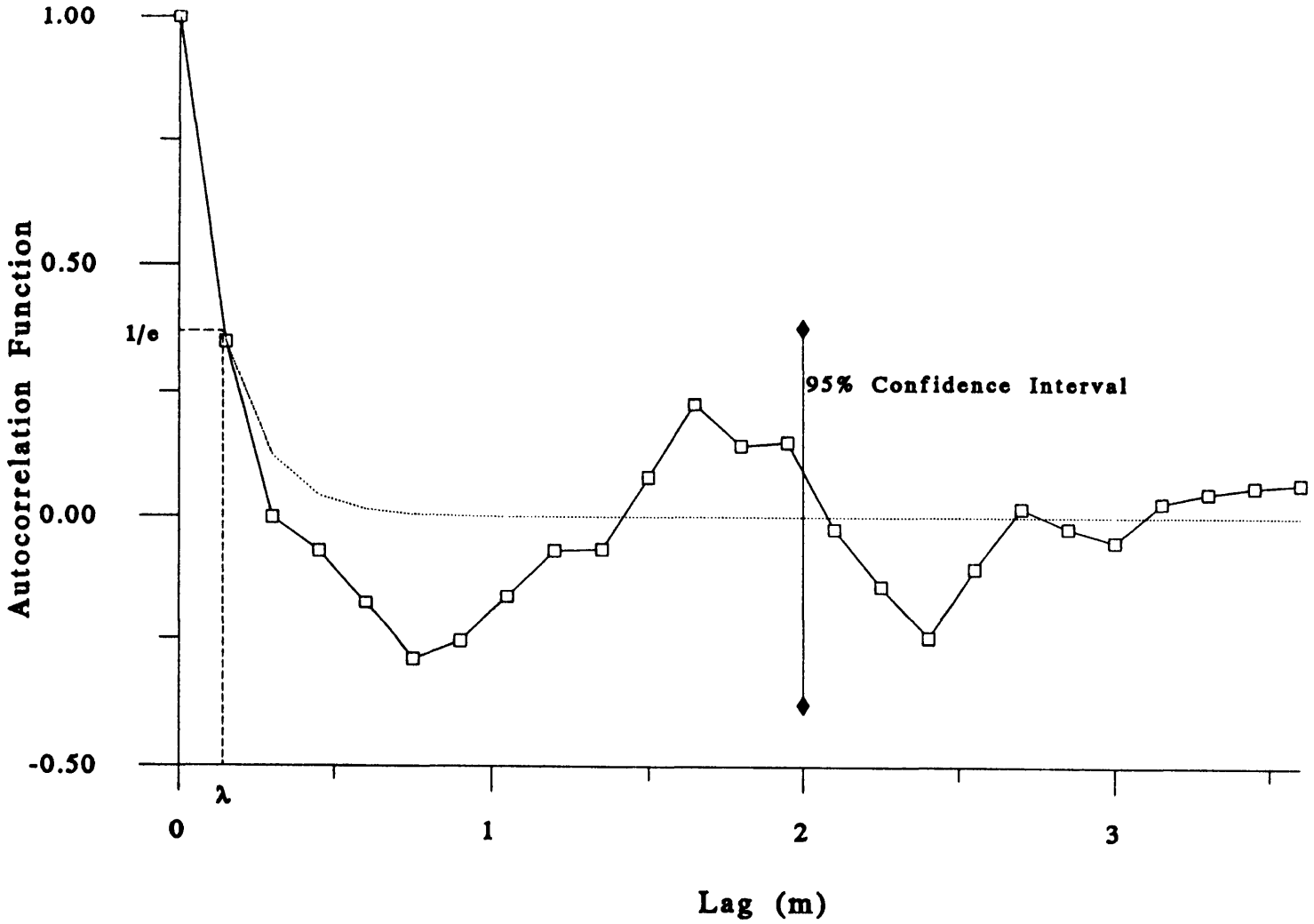

Figure 3.46 Test III of the linear regression model: Calculated autocorrelation function of $\ln K_d$ (dashed lines) for the vertical transect.

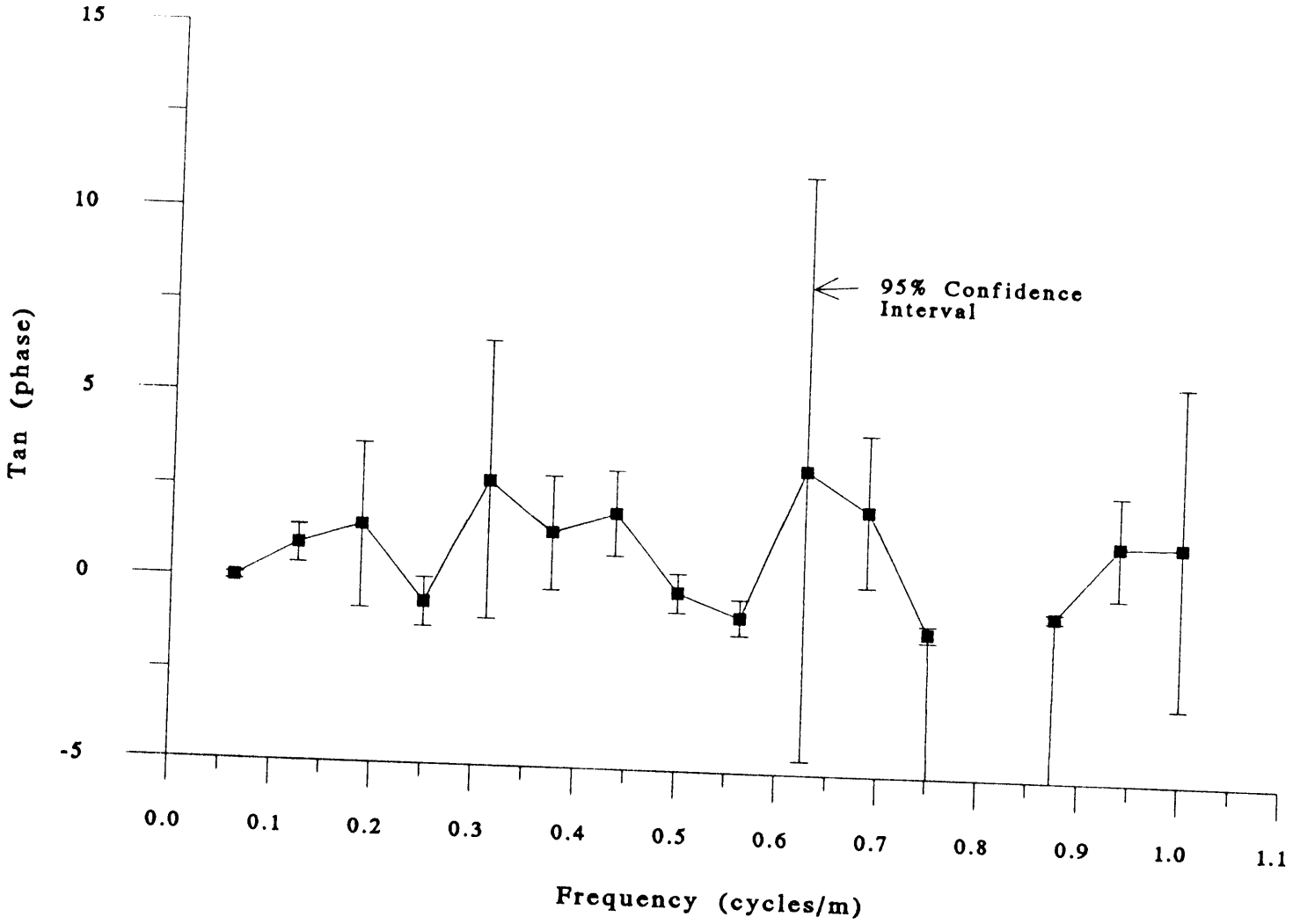

Figure 3.47 Test IV: Tangent of the phase of the cross-spectrum for the horizontal transect and M = 16.

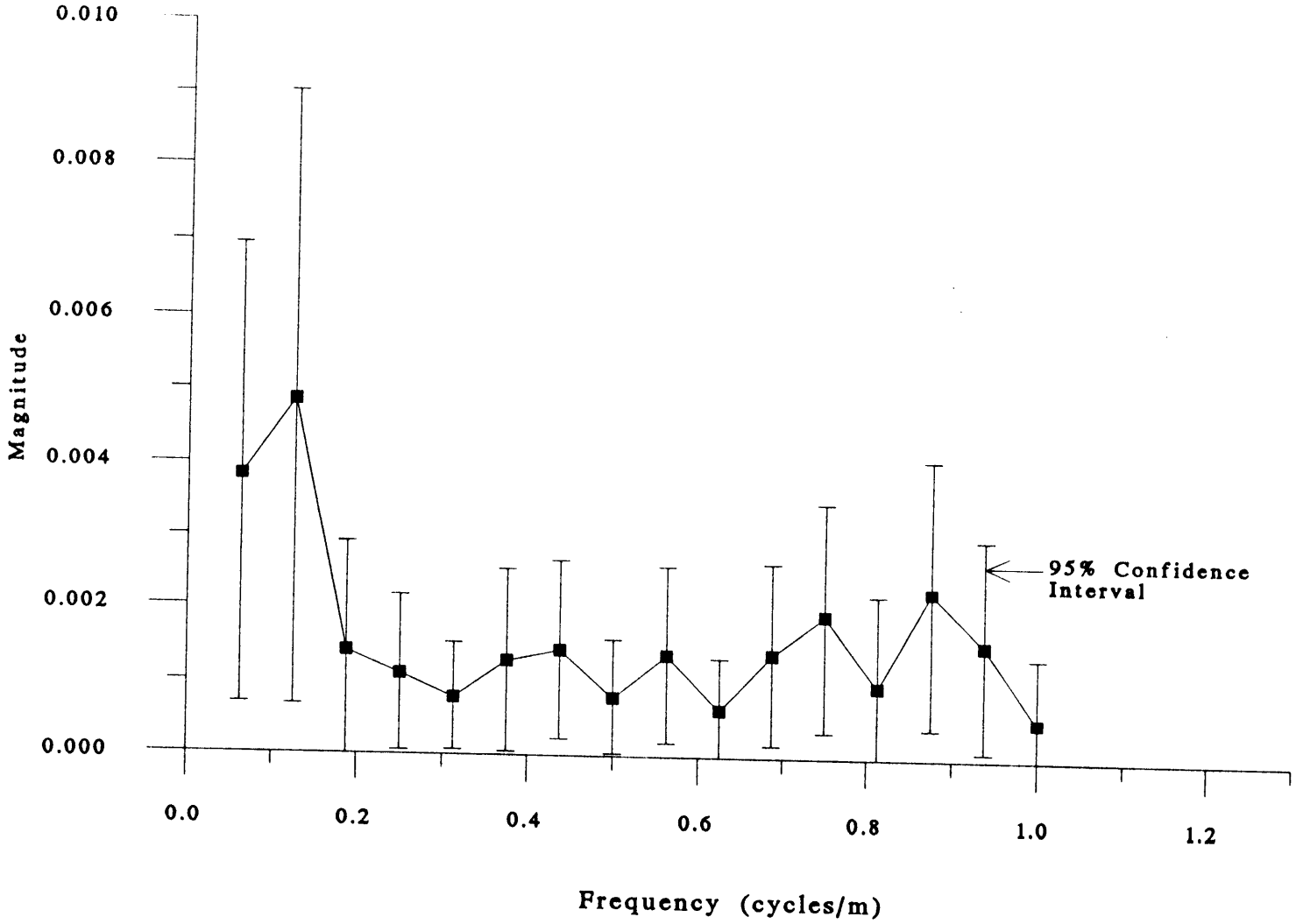

Figure 3.48 Test IV: Magnitude of the cross spectrum for the horizontal transect and $M = 16$.

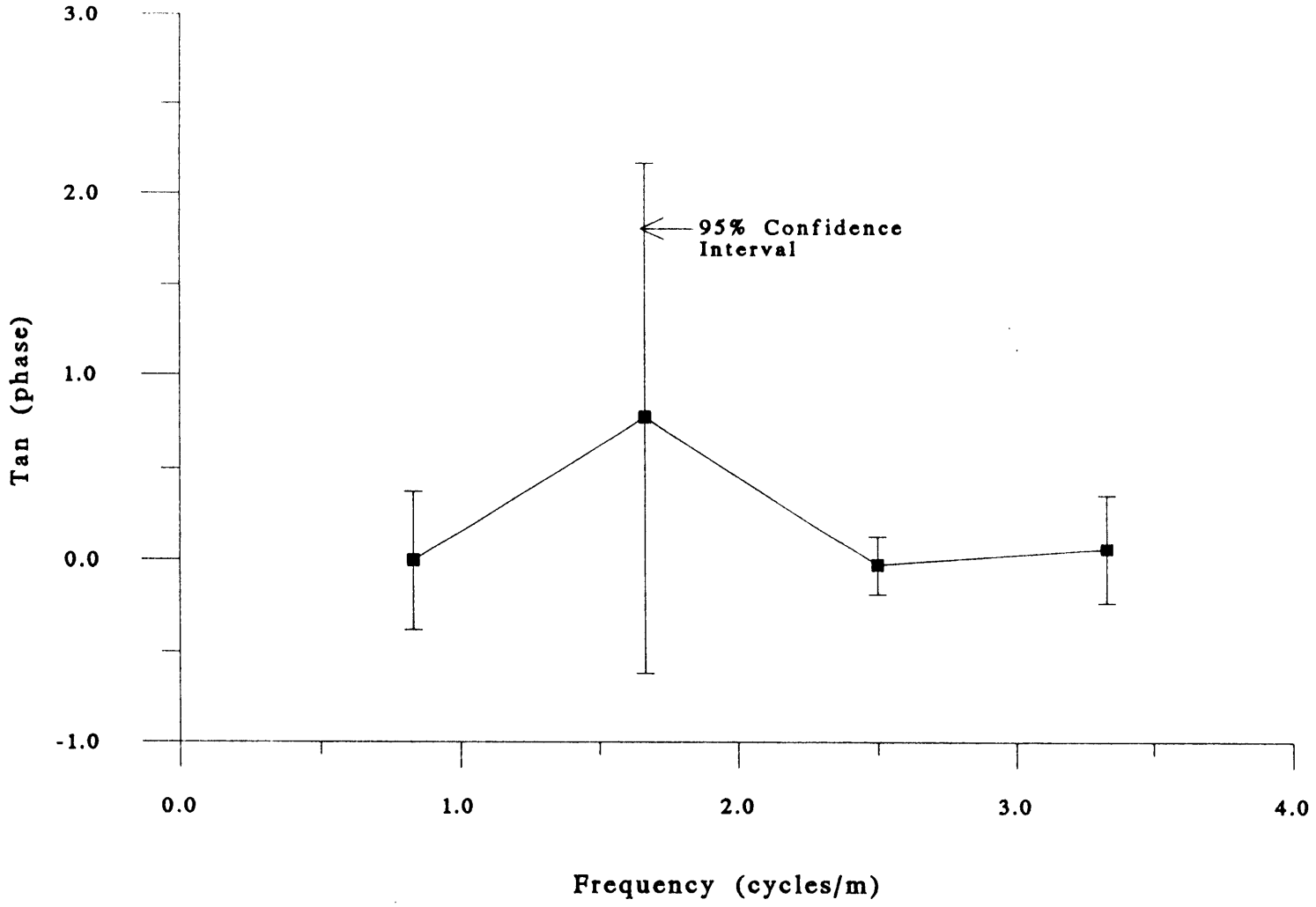

Figure 3.49 Test IV: Tangent of the phase of the cross spectrum for the vertical transect and M=4.

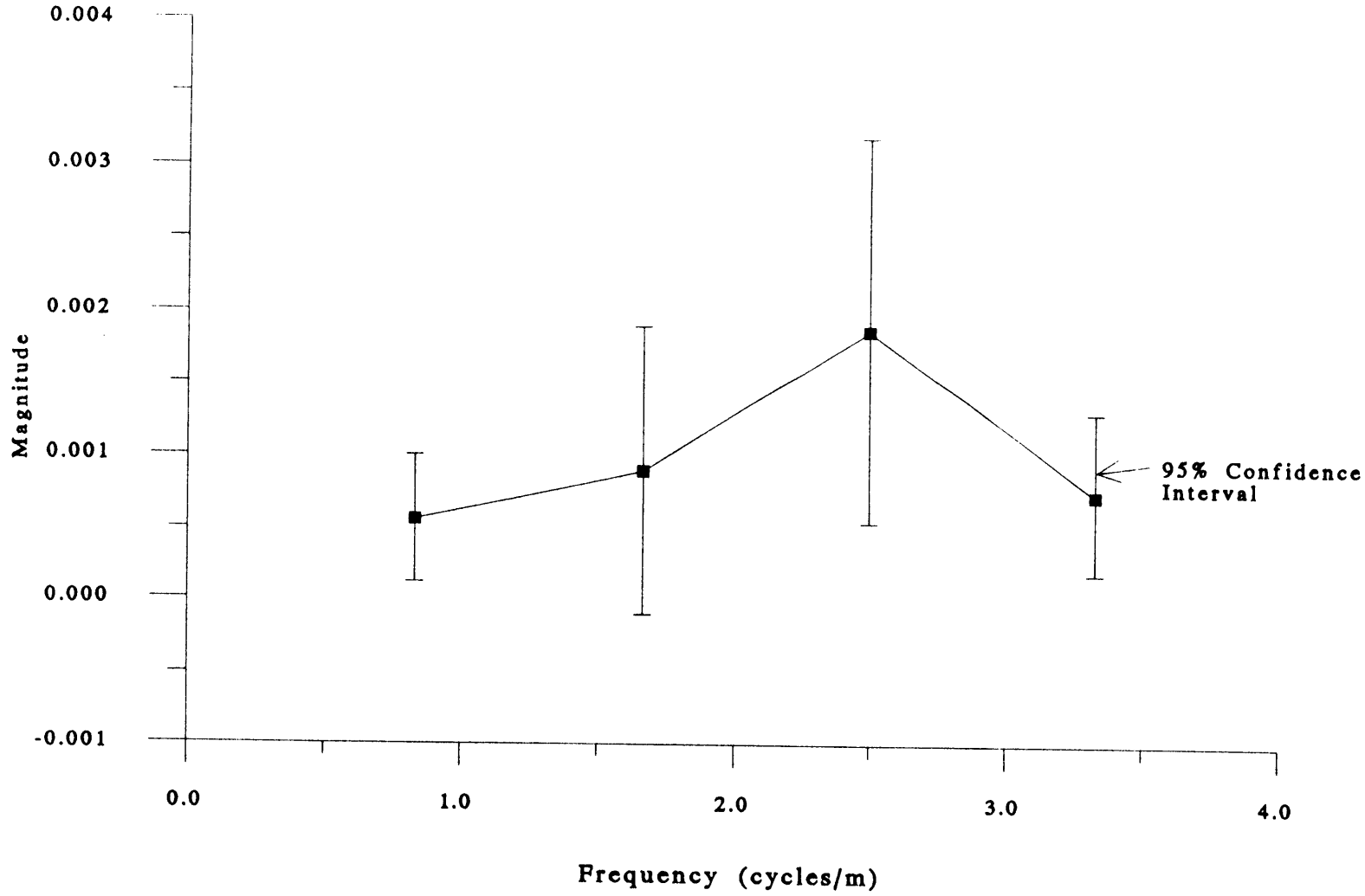

Figure 3.50 Test IV: Magnitude of the cross spectrum for the vertical transect and $M=4$.

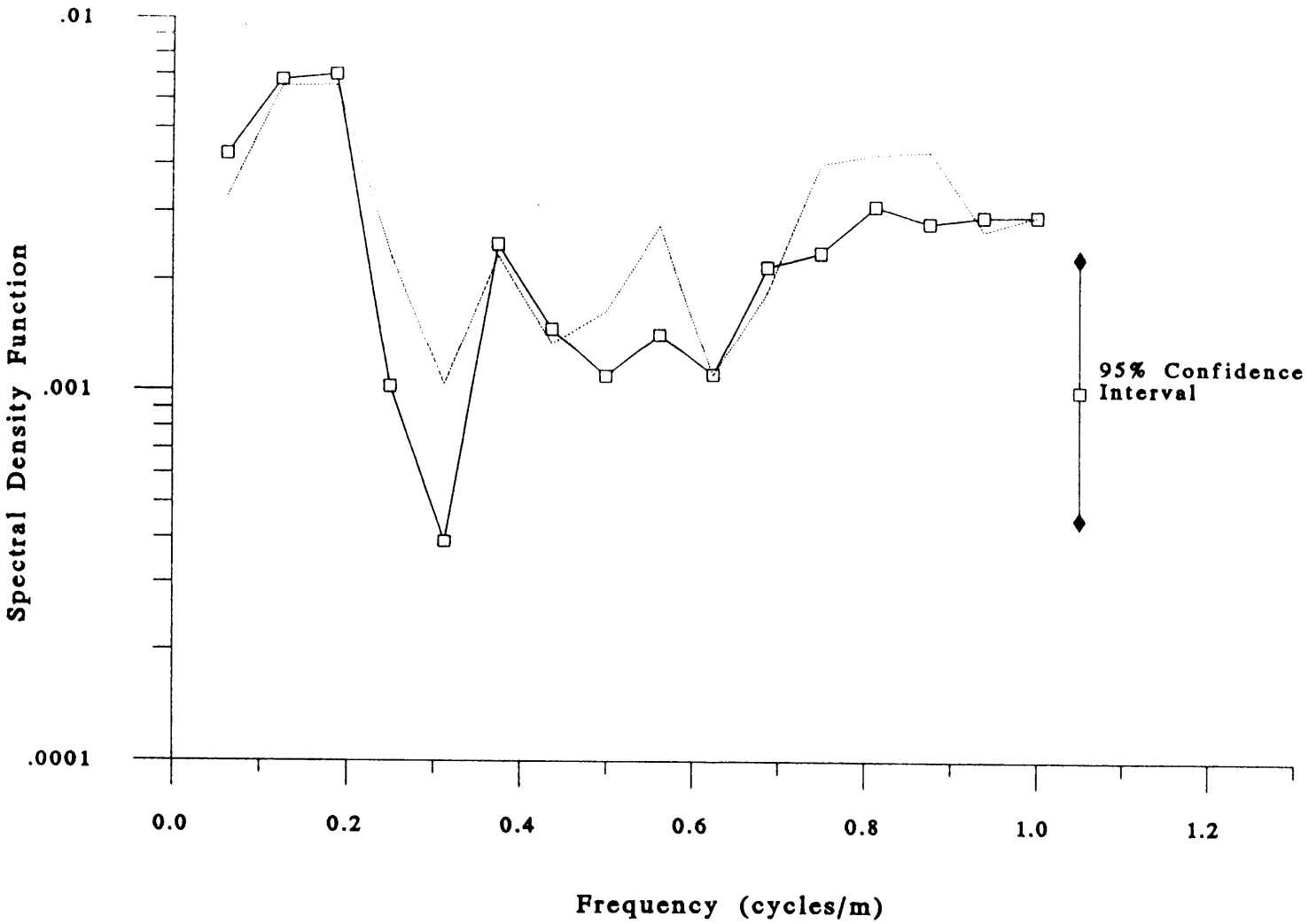

Figure 3.51 Test V of the linear regression model: Calculated spectrum of $\ln K_d$ (dashed lines) for the horizontal transect with $M = 16$.

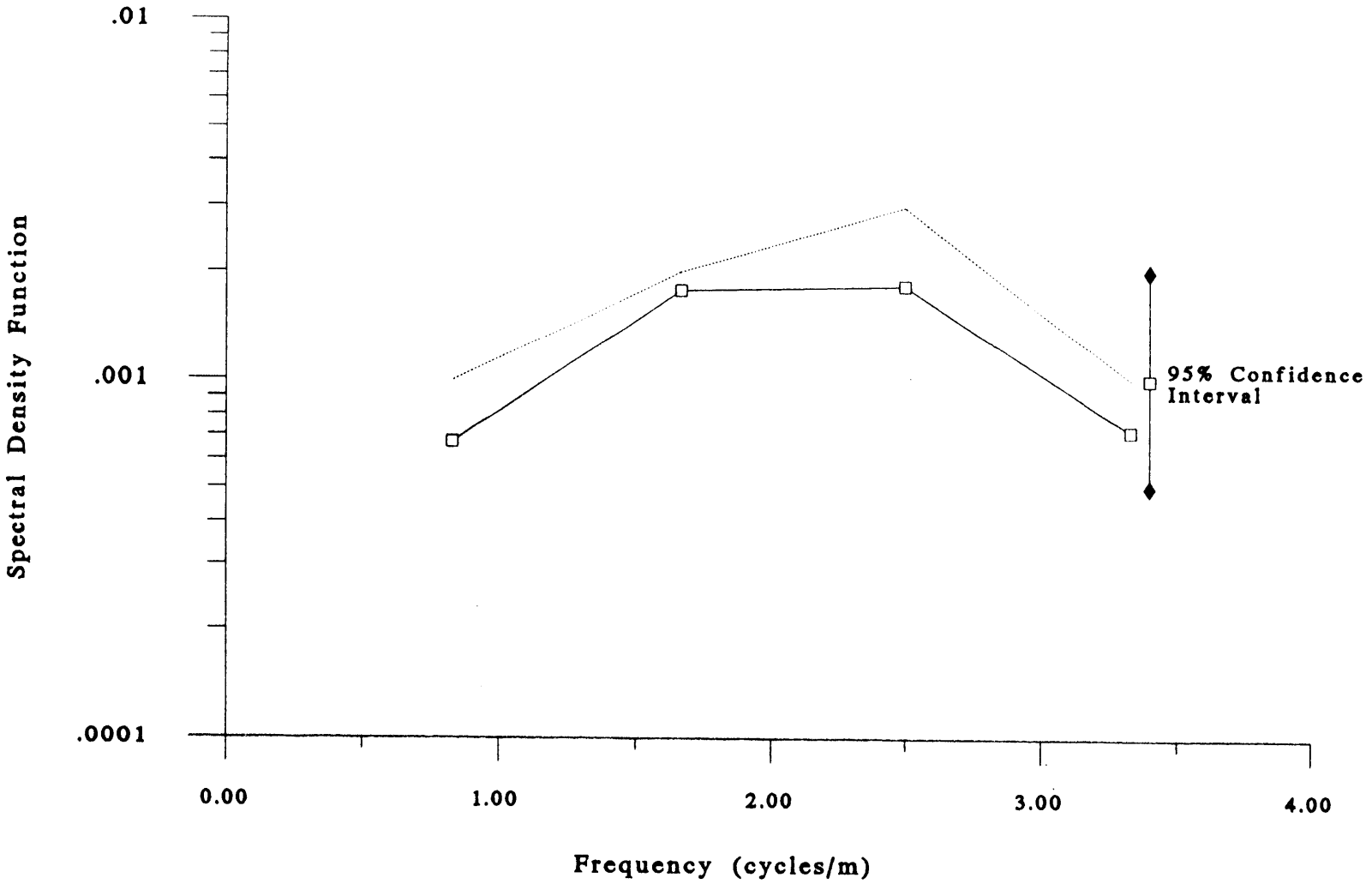

Figure 3.52 Test V of the linear regression model: Calculated spectrum of $\ln K_d$ (dashed lines) for the vertical transect with $M=4$.

3.4 LONGITUDINAL MACRODISPERSIVITY

The value of the longitudinal macrodispersivity is dependent on the parameters of the regression model in [1.6]. The model presented in section 3.3 is another form of that in [1.6], however the macrodispersivity is presented in terms of the model in [1.6]. The applicable conversions between the two models is found in Appendix E.

To calculate the macrodispersivity, the retardation factor is determined by [1.8]. The porosity is considered constant with a value of 0.30 and the density of the sand, ρ_b, is 1.855 g/cm^3. The parameter, γ, is calculated using equation (60) in Gelhar and Axness (*1983*):

$$\gamma = \exp\left[\sigma_f^2\left(\frac{1}{2} - g_{11}\right)\right] \qquad [3.55]$$

where g_{11} (0.070) is estimated using Figure 4*a* (*Gelhar and Axness, 1983*) with λ_1 = 1.0 m, λ_3 = 0.5 m. The variance of f for the horizontal transect, σ_f^2, is 0.080 thus γ is 1.035. The following table shows the macrodispersivity for benzene, TCB and PentCB:

3.2 Macrodispersivity values for Benzene, 1,2,4 Trichlorobenzene (TCB) and Pentachlorobenzene (PentCB)

Compound	K_{oc}	E[K_d]	R_f	A_0	A_{11}	A_η	A_{11}/A_0	A_η/A_0
Benzene	83	0.0158	1.10	0.076	0.0808	0.0004	1.1	0.006
TCB	2,041	0.3878	3.40	0.076	0.1415	0.0278	1.9	0.37
PentCB	13,000	2.47	16.27	0.076	0.1770	0.0492	2.3	0.65

The ratio of the macrodispersivity for a reactive tracer relative to a nonreactive tracer ranges from 1.1 to 2.3 according to the sorptive capacity of the chemical, as determined by the K_{oc}. These difference may seem modest, but they are understandable in light of the low organic carbon content of the soil and consequent low sorptive capability.

For the case of the reactive tracer, the negative correlation between ln$K$ and lnK_d leads to an additive effect in the calculation of the macrodispersivity. This addition to the macrodispersivity is the second term on the right hand side of equation [1.7], shown as A_η in Table 3.2, and ranges from 0.0004 to 0.0049. Therefore, the increase in macrodispersivity which is attributed to the correlation between ln$K$ and lnK_d leads to a 0.6% to 65% increase in the macrodispersivity and is dependent on the K_{oc} of the compound.

3.5 RELATED RESEARCH

Thompson (*1993*) used geologic information to develop a stochastic characterization on the variability of hydraulic conductivity for the glacial outwash aquifer on Cape Cod. One of the sites for which this stochastic information was gathered was the Handy Bog. Four stochastic parameters were estimated:

- the geometric mean hydraulic conductivity, K_g
- the variance of ln$K$ (var ln$K$)
- the horizontal correlation scale, λ_1
- the vertical correlation scale, λ_3.

The mean hydraulic conductivity, K_l and the variance of ln$K$ (var ln$K$) was estimated using a model which links the maximum grain size found in the outcrop to K_l and var

ln$K$. To estimate the correlation scales, the exposure was mapped and the bed lengths and thickness estimated. For the horizontal direction, λ_1 is estimated as a typical bed length. In the vertical direction, the correlation scale, λ_3, is estimated as typical bed thickness.

Table 3.3 Comparison of stochastic information estimated from geological information and stochastic information calculated from empirical data.

Stochastic Parameters	Method I: Stochastic information from sampling	Method II: Stochastic information from geologic information
Geometric Mean Hydraulic Conductivity, K_g (cm/s)	0.077	0.053
Variance of ln$K$	0.080	0.14
Horizontal Correlation Scale, λ_1 (m)	1.0 (ln$K$)	1.5
Vertical Correlation Scale, λ_3 (m)	0.10 (ln$K$)	0.12

The 95% confidence interval for the K_l and var ln$K$ is 0.003 and 2.6 x 10^{-2}. The differences in the K_l and var ln$K$, derived from method I and II, are greater than the 95% confidence interval. However, these differences are reasonable considering uncertainties in Thompson's model. This suggests that the generalized geologically-based approach of Thompson (*1994*) is quite promising as a simple, efficient and reliable method of predicting stochastic characteristics of aquifer hydraulic heterogeneity.

The concept of enhanced dispersion in the case of a retarded tracer has been studied by a number of scientist including Pickens et al. (*1981*), Garabedian et al

(*1988*) and Talbott (*1994*). The ratio of the dispersivity of the retarded solute (A_{11}) to that of the nonreactive solute (A_0) is shown in the following table:

Table 3.4 Enhanced dispersion of a retarded tracer: Comparison of four studies.

Source	A_{11}/A_0
This study: Sorption of benzene, TCB and PentCB Sampling over a 80 m record	1.1 - 2.3
Pickens et al. (*1981*): Sorption of Strontium Analysis of a tracer test	2.6
Garabedian et al. (*1988*): Sorption of lithium Interpreting a plume	10
Talbott (*1994*): Sorption of strontium Hypothetical calculation	9.7

Picken et al. (*1981*) measured the sorption of strontium using a small-scale experiment, on the order of meters, in sands at Chalk River. Garabedian et al. (*1988*) performed their tracer experiment at a much larger scale. This tracer test was performed in the same general aquifer material as that found at the Handy bog. The relative macrodispersivity of lithium (nonreactive) and bromide (reactive) was calculated based on the spatial moments of the distributions found in Figure 1.2. Talbott (*1994*) results are based on hypothetical calculations of sorption for strontium.

The ratio of the reactive to the nonreactive macrodispersivity for this study is not as large as the values calculated in the other studies. However, this may be a

result of the low organic carbon content of the soil and the resultant low sorptive capacity.

CHAPTER 4 SUMMARY AND CONCLUSIONS

4.1 OBJECTIVES

Several researchers have focused on the effects of the variability of hydraulic conductivity on the spreading of contaminants in the subsurface. It has been discovered that this heterogeneity leads to enhanced dispersion of contaminants. In a chemically heterogeneous system where the contaminant interacts with the porous media, by way of sorption, it is also plausible that this enhanced dispersion may, in part, be a result of the correlation the sorption coefficient (K_d) and the hydraulic conductivity (K).

The purpose of this research project is to:

1) measure the heterogeneity of the soil as indicated by the variability of K and K_d.
2) examine the degree of correlation between K and K_d.
3) represent this correlation using a simple model.
4) quantify the affect of 1) and 2) on the dispersion.

4.2 SUMMARY OF PROCEDURES

The sampling for this study was performed on an excavated exposure of glacial outwash at the Handy Cranberry Bog site, Cape Cod, Massachusetts. The bog is within 2 km of the USGS tracer test site, which is at southeast of the Otis Air

Force base, where an extensively studied sewage contamination plume extends to several kilometers in the sole source aquifer. The soil at the test site and the bog are composed of similar glaciofluvial outwash sediments which are classified as Mashpee Pitted Plain deposits.

At the Handy Bog, a bulldozer was used to access a vertical undisturbed face which consisted of cross-bedded sand. Seventy-three samples, at a spacing of 0.5 m, were taken from a horizontal transect and 26 samples, at a sampling interval of 0.15 m, were taken from a vertical transect. The samples weighed about 300 grams each.

The samples were analyzed for the percent organic carbon (*POC*) and the hydraulic conductivity (*K*) using the CHN Elemental Analyzer and a falling head permeameter, respectively. It became nessesary to evaluate the experimental measurement noise and the reproducibility of results in the CHN Analyzer and the permeameter. For the permeameter, the results were found to be practically reproducible in the case where the samples were packed to the same depth. Unlike the permeameter which uses the entire sample, the CHN Analyzer is equiped to handle samples which are in the order of 300 mg. Therefore, the reproducibility of the results is a function of the subsampling procedure. The bulk sample was riffled to extract 3 g subsamples. In an effort to quantify the measurement noise, 10 subsamples were analyzed, the variance of the *POC* for these subsamples is around 4×10^{-6}, the overall variance for all of the samples from the horizontal and vertical transect is about 1×10^{-3}. Therefore, the measurement noise is significant but not controlling. The *POC* measurements were used to calculate the distribution

coefficient (K_d) for three hypothesized dissolved organic contaminants; benzene, 1,2,4 Trichlorobenzene (TCB) and pentachlorobenzene (PentCB).

4.3 RESULTS AND CONCLUSIONS

For this data set, K and K_d are lognormal variables, thus ln$K$ and lnK_d were used for the statistical analysis. In the case of the Borden aquifer in Canada, Robin (*1991*) also found that these variables had lognormal distribution. Since the mean and variance of the data for both transects is statistically equivalent, the data for both transects are represented as a member of the same two-dimensional stationary random field whose mean is -2.57 and -1.003 for ln$K$ (K in cm/s) and lnK_d (K_d in ml/g), respectively.

Heterogeneity of the Soil

Though the sand appears to be homogeneous, the realization of ln$K$ and lnK_d exhibits a substantial amount of variability; the variance of ln$K$ is 0.080, that of ln K_d is 0.145. The spatial structure of this variability was quantified using the autocorrelation function which shows that, for ln$K$ and lnK_d, the horizontal correlation scales is about 1 m and the vertical correlation scales are 0.1 m and 0.15 m respectively. There is a high degree of uncertainty in the spectral estimates based on the limited number of points in the horizontal and vertical transect, 73 and 25 respectively.

Thompson (*1993*) obtained stochastic information using a generalized geologic modelling applied at several sites in Cape Cod including the Handy bog. The values of the mean hydraulic conductivity, variance of ln$K$ and the correlation scales estimated from geologic modelling are esentially equivalent to the results obtained from analyzing samples taken from Handy bog.

Correlation between lnK and lnK_d

The scatter plot between ln$K$ and lnK_d exhibits a weak negative correlation between the variables. A linear regression of lnK_d unto ln$K$ shows that this correlation is indeed significant with a slope (b) of -0.341 and a coefficient of correlation (r) of -0.233. The cross-correlation function indicate that there is a substantial negative correlation at small lags. The results of the cross-spectra is indeterminate since the cospectrum and quadrature spectrum fluctuates within the 95% confidence intervals. Based on the phase spectrum and the estimates of uncertainty, there is no out of phase correlation for ln$K$ and lnK_d.

The Linear Regression Model

Five different test were used to determine if the data can be represented by the linear regression model in which lnK_d is expressed in terms of ln$K$, an uncorrelated residual term and the slope from the best fit curve of the scatter plot of lnK_d versus ln$K$. These test include a test of the correlation between hydraulic conductivity and the residual, two test in the covariance domain and two test in the spectral domain. Based on the results of these test, the regression model is

appropriate as a first approximation representing the correlation between the two variable.

Macrodispersivity

In the aquifer material at the Handy Bog, a nonreactive tracer would have a longitudinal macrodispersivity of 0.076 m, whereas reactive tracers; such as benzene, TCB and PentCB; have a macrodispersivity of 0.081 m, 0.142 m and 0.177 m, respectively. The enhancement of dispersion due to the correlation between ln$K$ and lnK_d ranges from 0.6% to 65% and is dependent upon the K_{oc} values of the compound. This difference may seem moderate, but it is understandable in view of the low organic carbon content of the soil and the resulting low sorptive capability.

The ratio of the macrodispersivity for a retarded solute to that of a nonreactive solute ranges from 1.1 to 2.3. This ratio is low compared to value of:

- 10 obtained from a large scale tracer experiment at the test site which has the same aquifer material. Though this tracer test was performed in the same general aquifer material, this macrodispersivity represents that of metal sorption in which a lithium plume was interpreted (*Garabedian et al., 1988*),
- 9.7 obtained from a hypothetical calculation based on strontium (*Talbott, 1994*), and
- 2.6 obtained from a small scale experiment, in the order of a few meters, with strontium at sands at Chalk River (*Pickens, 1981*).

However, the difference in the ratio of the reactive to nonreactive dispersivity for this study and the others mentioned above can be attributed to the low sorptive capacity of the soil.

4.4 FUTURE RESEARCH

- The soils in this study have a low organic carbon content (the average *POC* is 0.019%). Future research should also consider the variability and correlation between $\ln K$ and $\ln K_d$ for soils with a wider range of organic carbon content such that the degree of variability would be larger than that studied in this case.
- The regression model should be tested on a variety of soils at a variety of scales to determine how the model changes with scale and variance of the K and K_d values.
- The scope of this research could be broadened to include the three-dimensional case.

REFERENCES

BALL, W.P. 1989. *Equilibrium sorption and diffusion rate studies with halogenated organic chemicals and sandy aquifer material.* Ph. D. Dissertations, Dep. Civ. Eng., Stanford Univ., Stanford, CA. 356pp.

BALL, W.P., CH BUEHLER, T.C. HARMON, D.M. MCKAY AND P.V. ROBERTS. 1990. Characterization of a Sandy Aquifer Material at the Grain Scale. *Journal of Contaminant Hydrology (5): 253-295.*

BARBER, L.B. 1992. Groundwater Contamination and Analysis at Hazardous Waste Sites. 73-119. Marcel Dekker, Inc.

BARBER,L.B., E.M. THURMAN and D.D RUNNELLS. 1992. Geochemical heterogenity in a sand and gravel aquifer: Effect of sediment mineralogy and particle size on the sorption of chlorobenzenes. *Journal of Contaminant Hydrology (9): 35-54.*

BARBER, L.B., E.M. THURMAN. M.P SCHROEDER and D.R LEBLANC. 1988. Long-Term Fate of Organic Micropollutants in Sewage-Contaminated Groundwater. *Environmental Science and Technology (22): 205-211.*

BARBER, L.B. 1990. Geochemical heterogenity in a glacial outwash aquifer: Effect of particle size and mineralogy on sorption of nonionic organic solutes: Boulder, Colo., University of Colorado, Department of Geological Sciences, unpublished Ph.D. Thesis, 238p.

BARBER, L.B. 1985. Geochemistry of organic and inorganic compounds in a sewage-contaminated aquifer, Cape Cod, Massachusetts: Boulder, Colo., University of Colorado, Department of Geological Sciences, unpublished Master's Thesis, 169p.

CULMO, R.F. Principle of Operation - The Perkin-Elmer PE 2400 CHN Elemental Analyzer. Perkin-Elmer Publication EAN-2

DAVIS, J.A., J.A. COSTON, C.C. FULLER, K.M. HESS and E. DIXON. 1993. Spatial Heterogenity of Geochemical and Hydrologic Parameters Affecting Metal Transport in Groundwater. USGS Environmental Research Brief

DOMENICO, P.A. and F.W. SCHWARTZ. 1990. *Physical and Chemical Hydrogeology.* Wiley.

ELABD, H., JURY, W.A. and M. M. CLIATH. 1986. Spatial variability of pesticide adsorption parameters. *Environmental Science Technology (20): 256-260.*

FREEZE, R.A. and J.A. CHERRY. 1979. *Groundwater.* Prentice-Hall, Inc., New York.

GARABEDIAN, S.P. 1987. Large-scale dispersive transport in aquifers: Field experiments and reactive transport theory. Department of Civil Engineering., Massachusetts Institute of Technology, Cambridge, MA. Ph.D thesis. 290pp.

GARABEDIAN, S.P., L.W. GELHAR and M.A. CELIA. 1988. Large-scale dispersive transport in an aquifer: Field experiments and reactive transport theory. Massachusetts Institute of Technology Technical Report No. 315.

GELHAR, L.W., A.L GUTJAHR and R.L NAFF. 1979. Stochastic Analysis of Macrodispersion in a Stratified Aquifer. *Water Resources Research (15): 1387-1397.*

GELHAR, L.W. and C. L AXNESS. 1983. Three-Dimensional Stochastic Analysis of Macrodispersion in Aquifers. *Water Resources Research (19): 161-180.*

GILLHAM, R.W., M.J.L. ROBIN and C.J. PTZCEK. 1990. A device for in-situ determination of geochemical transport parameters - 1. Retardation. *Ground-Water 28(5):666-672.*

HESS, K.M. 1993. Transport of Metal-Ions in Ground Water Under Spatially Varying Chemical Conditions and Aquifer Properties. Unpublished PhD Proposal. Massachusetts Institute of Technology.

KARICKHOFF, S. W., D. S. BROWN and T. A. SCOTT. 1979. Sorption of hydrophobic pollutants on natural sediments. *Water Resources Research v. 13, p. 242-248*

McGraw-Hill Dictionary of Scientific and Technical Terms Third Edition, 1984 MaGraw-Hill

MCKAY, D.M, W.P. BALL and M.G. DURANT. 1986. Variability of Aquifer Sorption Properties in a Field Experiment of Groundwater Transport of Organic Solutes: Methods and Preliminary Results. *Journal of Contaminant Hydrology (1): 119-132.*

ROBIN, M.J.L., E.A. SUDICKY, R.W. GILLHAM, and R.G. KACHANOSKI. 1991. Spatial Variability of Strontium Distribution Coefficients and Their Correlation With Hydraulic Conductivity in the Canadian Forces Base Borden Aquifer. Water Resources Research 27(10):2619-2632.

SCHWARZENBACH, R.P., and J. WESTFALL. 1981. Transport of non-polar organic compounds from surface water to ground-water - laboratory sorption studies. Environmental Science and Technology 15(11):1360-66

SHUMWAY, R.H. 1988. *Applied Statistical Time Series Analysis*. Prentice Hall.

SMITH, J.A., P.J. WITKOWSKI, and T.V. FUSILLO. 1988. Man-made organic compounds in the surface waters of the United States - a review of current understanding. U.S. Geological Survey Circular 1000:92

APPENDIX A

Hydraulic Conductivity, Percent Organic Carbon and Sorption Coefficient Measurements for the Vertical and Horizontal Transect

Horizontal Transect						
Sample ID	Distance (m)	POC	ln Kd Benzene	ln Kd TCB	ln Kd PentCB	Ln K
	Mean	0.019	-4.205	-1.003	0.849	-2.561
	St. Deviation	0.008	0.372	0.372	0.372	0.247
	Variance	6.87e-05	0.1386	0.1386	0.1386	0.0609
177	0.5	0.013	-4.529	-1.327	0.525	-2.562
178	1	0.012	-4.635	-1.432	0.419	-2.403
179	1.5	0.026	-3.836	-0.634	1.218	-2.385
180	2	0.014	-4.434	-1.231	0.620	-2.391
181	2.5	0.013	-4.529	-1.327	0.525	-2.456
182	3	0.012	-4.635	-1.432	0.419	-3.435
183	3.5	0.009	-4.886	-1.684	0.168	-2.362
184	4	0.014	-4.434	-1.231	0.620	-2.789
185	4.5	0.009	-4.886	-1.684	0.168	-2.428
186	5	0.014	-4.434	-1.231	0.620	-2.418
187	5.5	0.011	-4.743	-1.540	0.311	-1.971
188	6	0.015	-4.386	-1.184	0.668	-2.209
189	6.5	0.012	-4.609	-1.407	0.445	-2.349
190	7	0.011	-4.743	-1.540	0.311	-2.168
191	7.5	0.012	-4.609	-1.407	0.445	-2.298
192	8	0.017	-4.291	-1.088	0.763	-2.319
193	8.5	0.015	-4.386	-1.184	0.668	-2.420
194	9	0.014	-4.491	-1.289	0.562	-2.726
195	9.5	0.017	-4.291	-1.088	0.763	-2.024
196	10	0.017	-4.291	-1.088	0.763	-2.138

197	10.5	0.012	-4.609	-1.407	0.445	-2.353
198	11	0.033	-3.598	-0.395	1.456	-2.254
199	11.5	0.017	-4.291	-1.088	0.763	-2.473
201	12	0.045	-3.287	-0.085	1.766	-2.535
202	12.5	0.027	-3.798	-0.596	1.256	-2.498
203	13	0.017	-4.291	-1.088	0.763	-2.481
204	13.5	0.021	-4.050	-0.847	1.004	-2.342
205	14	0.036	-3.511	-0.308	1.543	-2.496
206	14.5	0.029	-3.744	-0.542	1.310	-3.009
207	15	0.026	-3.855	-0.653	1.198	-2.699
208	15.5	0.024	-3.916	-0.714	1.138	-2.626
209	16	0.023	-3.981	-0.778	1.073	-2.431
210	16.5	0.021	-4.050	-0.847	1.004	-2.702
211	17	0.033	-3.605	-0.402	1.449	-2.638
212	17.5	0.020	-4.094	-0.892	0.959	-2.835
213	18	0.032	-3.638	-0.435	1.416	-2.778
214	18.5	0.019	-4.148	-0.946	0.906	-2.838
215	19	0.018	-4.206	-1.003	0.848	-2.583
216	19.5	0.016	-4.331	-1.128	0.723	-2.831
217	20	0.012	-4.641	-1.439	0.413	-2.466
218	20.5	0.014	-4.474	-1.272	0.580	-2.513
219	21	0.016	-4.331	-1.128	0.723	-2.712
220	21.5	0.012	-4.641	-1.439	0.413	-2.627
221	22	0.013	-4.554	-1.351	0.500	-2.578
222	22.5	0.020	-4.094	-0.892	0.959	-2.732
223	23	0.017	-4.266	-1.064	0.788	-2.780
224	23.5	0.018	-4.206	-1.003	0.848	-2.459
225	24	0.016	-4.331	-1.128	0.723	-2.613
226	24.5	0.023	-3.948	-0.745	1.106	-2.565

227	25	0.022	-3.994	-0.792	1.060	-2.510
228	25.5	0.021	-4.043	-0.841	1.011	-2.862
229	26	0.019	-4.148	-0.946	0.906	-2.717
230	26.5	0.021	-4.043	-0.841	1.011	-2.710
231	27	0.012	-4.641	-1.439	0.413	-2.220
232	27.5	0.030	-3.689	-0.486	1.365	-2.618
233	28	0.015	-4.417	-1.215	0.637	-2.633
234	28.5	0.016	-4.315	-1.113	0.739	-2.623
235	29	0.014	-4.474	-1.272	0.580	-2.620
236	29.5	0.012	-4.597	-1.394	0.457	-2.626
237	30	0.015	-4.417	-1.215	0.637	-2.785
238	30.5	0.017	-4.284	-1.081	0.770	-2.907
239	31	0.019	-4.166	-0.963	0.888	-2.966
240	31.5	0.017	-4.253	-1.050	0.801	-2.583
241	32	0.056	-3.067	0.135	1.987	-2.796
242	32.5	0.030	-3.706	-0.504	1.348	-2.858
243	33	0.029	-3.724	-0.521	1.330	-2.610
244	33.5	0.014	-4.491	-1.289	0.563	-2.674
245	34	0.026	-3.835	-0.633	1.219	-2.682
246	34.5	0.016	-4.333	-1.131	0.721	-2.397
247	35	0.020	-4.078	-0.876	0.976	-2.546
248	35.5	0.024	-3.917	-0.714	1.137	-1.918
249	36	0.020	-4.078	-0.876	0.976	-2.701
250	36.5	0.030	-3.690	-0.487	1.364	-2.702

Vertical Transect						
Sample ID	Distance (m)	POC	ln Kd Benzene	ln Kd TCB	ln Kd PentCB	Ln K
With Sample AB						
	Mean	0.020	-4.313	-1.111	0.741	-2.633
	St. Deviation	0.023	0.554	0.554	0.554	0.464
	Variance	5.26e-04	0.3070	0.3070	0.3070	0.2153
Without Sample AB						
	Mean	0.016	-4.397	-1.194	0.657	-2.561
	St. Deviation	0.006	0.369	0.369	0.369	0.295
	Variance	3.22e-05	0.1361	0.1361	0.1361	0.0868
a	0.15	0.032	-3.621	-0.419	1.433	-2.515
b	0.30	0.015	-4.395	-1.194	0.659	-2.708
c	0.45	0.013	-4.542	-1.340	0.511	-2.597
d	0.60	0.010	-4.798	-1.596	0.256	-2.910
e	0.75	0.020	-4.098	-0.896	0.956	-3.105
f	0.90	0.017	-4.266	-1.064	0.788	-2.009
g	1.05	0.010	-4.766	-1.563	0.288	-2.521
h	1.20	0.012	-4.584	-1.382	0.470	-2.732
i	1.35	0.015	-4.386	-1.184	0.668	-2.212
j	1.50	0.014	-4.459	-1.256	0.595	-2.381
l	1.65	0.017	-4.290	-1.088	0.764	-2.941
m	1.80	0.021	-4.032	-0.829	1.022	-3.170
n	1.95	0.025	-3.873	-0.670	1.181	-2.762
o	2.1	0.016	-4.322	-1.120	0.732	-2.486
p	2.25	0.008	-5.011	-1.809	0.043	-2.354
q	2.4	0.011	-4.697	-1.495	0.356	-2.203

r	2.55	0.006	-5.283	-2.080	-0.229	-2.280
s	2.7	0.009	-4.895	-1.693	0.158	-2.511
u	2.85	0.018	-4.181	-0.978	0.873	-2.312
v	3	0.014	-4.454	-1.252	0.600	-2.177
w	3.15	0.013	-4.542	-1.340	0.511	-2.782
x	3.3	0.020	-4.076	-0.873	0.978	-2.920
y	3.45	0.019	-4.125	-0.923	0.929	-2.574
z	3.6	0.019	-4.125	-0.923	0.929	-2.422
aa	3.75	0.020	-4.098	-0.896	0.956	-2.432
ab	3.9	0.132	-2.214	0.988	2.839	-4.449

APPENDIX B
MATLAB Codes for the Spectral Estimates

```
function P = spectrum(x,y,m,noverlap)
%SPECTRUM Power spectrum estimate of one or two data sequences.
%       P = SPECTRUM(X,Y,M) performs FFT analysis of the two sequences
%       X and Y using the Welch method of power spectrum estimation.
%       The X and Y sequences of N points are divided into K sections of
%       M points each (M must be a power of two). Using an M-point FFT,
%       successive sections are Hanning windowed, FFT'd and accumulated.
%       SPECTRUM returns the M/2 by 8 array
%         P = [Pxx Pyy Pxy Txy Cxy Pxxc Pyyc Pxyc]
%       where
%         Pxx  = X-vector power spectral density
%         Pyy  = Y-vector power spectral density
%         Pxy  = Cross spectral density
%         Txy  = Complex transfer function from X to Y
%              (Use ABS and ANGLE for magnitude and phase)
%         Cxy  = Coherence function between X and Y
%         Pxxc,Pyyc,Pxyc = Confidence range (95 percent).
%
%       See SPECPLOT to plot these results.
%       P = SPECTRUM(X,Y,M,NOVERLAP) specifies that the M-point sections
%       should overlap NOVERLAP points.
%       Pxx = SPECTRUM(X,M) and SPECTRUM(X,M,NOVERLAP) return the single
%       sequence power spectrum and confidence range.
%
%       See also ETFE, SPA, and ARX in the Identification Toolbox.

%       J.N. Little 7-9-86
%       Revised 4-25-88 CRD, 12-20-88 LS, 8-31-89 JNL, 8-11-92 LS
%       Copyright (c) 1986-92 by the MathWorks, Inc.

%       The units on the power spectra Pxx and Pyy are such that, using
%       Parseval's theorem:
%
%           SUM(Pxx)/LENGTH(Pxx) = SUM(X.^2)/LENGTH(X) = COV(X)
%
%       The RMS value of the signal is the square root of this.
%       If the input signal is in Volts as a function of time, then
%       the units on Pxx are Volts^2*seconds = Volt^2/Hz.
%       To normalize Pxx so that a unit sine wave corresponds to
%       one unit of Pxx, use Pn = 2*SQRT(Pxx/LENGTH(Pxx))
%
%       Here are the covariance, RMS, and spectral amplitude values of
%       some common functions:
%         Function   Cov=SUM(Pxx)/LENGTH(Pxx)   RMS        Pxx
%         a*sin(w*t)        a^2/2            a/sqrt(2)   a^2*LENGTH(Pxx)/4
%Normal:  a*rand(t)          a^2               a          a^2
iform: a*rand(t)          a^2/12           a/sqrt(12)  a^2/12
%
%       For example, a pure sine wave with amplitude A has an RMS value
%       of A/sqrt(2), so A = SQRT(2*SUM(Pxx)/LENGTH(Pxx)).
```

```
%
%       See Page 556, A.V. Oppenheim and R.W. Schafer, Digital Signal
%       Processing, Prentice-Hall, 1975.

if (nargin == 2), m = y; noverlap = 0; end
if (nargin == 3)
      if (max(size(y)) == 1)
            noverlap = m;
            m = y;
            nargin = 2;
      else
            noverlap = 0;
      end
end

x = x(:);              % Make sure x and y are column vectors
y = y(:);
n = max(size(x));      % Number of data points
k = fix((n-noverlap)/(m-noverlap));    % Number of windows
                              % (k = fix(n/m) for noverlap=0)
index = 1:m;
w = hanning(m);        % Window specification; change this if you want:
                  % (Try HAMMING, BLACKMAN, BARTLETT, or your own)
KMU = k*norm(w)^2;     % Normalizing scale factor

if (nargin == 2)       % Single sequence case.
      Pxx = zeros(m,1); Pxx2 = zeros(m,1);
      for i=1:k
            xw = w.*detrend(x(index));
            index = index + (m - noverlap);
            Xx = abs(fft(xw)).^2;
            Pxx = Pxx + Xx;
            Pxx2 = Pxx2 + abs(Xx).^2;
      end
      % Select first half
      select = [1:m/2];
      Pxx = Pxx(select);
      Pxx2 = Pxx2(select);
      cPxx = zeros(m/2,1);
      if k > 1
            c = (k.*Pxx2-abs(Pxx).^2)./(k-1);
            c = max(c,zeros(m/2,1));
                              % (k = fix(n/m) for noverlap=0)
            cPxx = sqrt(c);
      end
      pp = 0.95; % 95 percent confidence.
      f = sqrt(2)*erfinv(pp);  % Equal-tails.
      P = [Pxx f.*cPxx]/KMU;
      return
end

Pxx = zeros(m,1); % Dual sequence case.
Pyy = Pxx; Pxy = Pxx; Pxx2 = Pxx; Pyy2 = Pxx; Pxy2 = Pxx;
```

```
for i=1:k
      xw = w.*detrend(x(index));
      yw = w.*detrend(y(index));
      index = index + (m - noverlap);
      Xx = fft(xw);
      Yy = fft(yw);
      Yy2 = abs(Yy).^2;
      Xx2 = abs(Xx).^2;
      Xy  = Yy .* conj(Xx);
      Pxx = Pxx + Xx2;
      Pyy = Pyy + Yy2;
      Pxy = Pxy + Xy;
                              % (k = fix(n/m) for noverlap=0)
      Pxx2 = Pxx2 + abs(Xx2).^2;
      Pyy2 = Pyy2 + abs(Yy2).^2;
      Pxy2 = Pxy2 + Xy .* conj(Xy);
end

% Select first half
select = [1:m/2];

Pxx = Pxx(select);
Pyy = Pyy(select);
Pxy = Pxy(select);
Pxx2 = Pxx2(select);
Pyy2 = Pyy2(select);
Pxy2 = Pxy2(select);

cPxx = zeros(m/2,1);
cPyy = cPxx;
cPxy = cPxx;
if k > 1
  c = max((k.*Pxx2-abs(Pxx).^2)./(k-1),zeros(m/2,1));
  cPxx = sqrt(c);
  c = max((k.*Pyy2-abs(Pyy).^2)./(k-1),zeros(m/2,1));
  cPyy = sqrt(c);
                              % (k = fix(n/m) for noverlap=0)
  c = max((k.*Pxy2-abs(Pxy).^2)./(k-1),zeros(m/2,1));
  cPxy = sqrt(c);
end

Txy = Pxy./Pxx;
Cxy = (abs(Pxy).^2)./(Pxx.*Pyy);

pp = 0.95; % 95 percent confidence.
f = sqrt(2)*erfinv(pp);  % Equal-tails.

P = [ [Pxx Pyy Pxy]./KMU ...
     Txy Cxy ...
     f.*[cPxx cPyy cPxy]./KMU ];                    2*SQRT(Pxx/LENGTH(Pxx))
```

SPECTRAL WINDOWS:

```
function w = hanning(n)
%HANNING HANNING(N) returns the N-point Hanning window in a column vector.
w = .5*(1 - cos(2*pi*(1:n)'/(n+1)));

function w = bartlett(n)
%BARTLETT BARTLETT(N) returns the N-point Bartlett window.
w = 2*(0:(n-1)/2)/(n-1);
if rem(n,2)
      % It's an odd length sequence
      w = [w w((n-1)/2:-1:1)]';
else
      % It's even
      w = [w w(n/2:-1:1)]';
end


function w=triang(n)
%TRIANG TRIANG(N) returns the N-point triangular window.
if rem(n,2)
      % It's an odd length sequence
      w = 2*(1:(n+1)/2)/(n+1);
      w = [w w((n-1)/2:-1:1)]';
else
      % It's even
      w = (2*(1:(n+1)/2)-1)/n;
      w = [w w(n/2:-1:1)]';
end
```

APPENDIX C
Spectral Estimates from MATLAB

The input is in the form of:
SPECTRUM(X,Y,2M)

The output in the form:
P = [Pxx Pyy Pxy Txy Cxy Pxxc Pyyc Pxyc]
where
Pxx = X-vector power spectral density
Pyy = Y-vector power spectral density
Pxy = Cross spectral density
Txy = Complex transfer function from X to Y
(Use ABS and ANGLE for magnitude and phase)
Cxy = Coherency squared function between X and Y
Pxxc,Pyyc,Pxyc = Confidence range (95 percent).

1. Horizontal Transect with an *M* value of 16, the variables are ln*K* (kh) and lnK_d (kdh)

>> spectrum(kh,kdh,32)

ans =

Columns 1 through 4

0.0888	0.1072	-0.0964	-1.0849
0.1083	0.1704	-0.0833 - 0.0893i	-0.7688 - 0.8242i
0.0175	0.1765	-0.0294 + 0.0197i	-1.6758 + 1.1240i
0.0530	0.0257	0.0137 - 0.0243i	0.2593 - 0.4584i
0.0640	0.0098	-0.0189 + 0.0068i	-0.2951 + 0.1056i
0.0308	0.0625	-0.0265 - 0.0188i	-0.8591 - 0.6103i
0.0415	0.0372	-0.0324 - 0.0163i	-0.7822 - 0.3921i
0.0270	0.0276	0.0027 - 0.0202i	0.0999 - 0.7474i
0.0433	0.0362	0.0209 - 0.0279i	0.4816 - 0.6435i
0.0231	0.0280	-0.0158 - 0.0048i	-0.6841 - 0.2086i
0.0311	0.0549	-0.0322 + 0.0145i	-1.0346 + 0.4653i
0.0407	0.0597	0.0351 + 0.0334i	0.8621 + 0.8217i
0.0283	0.0786	0.0249 - 0.0001i	0.8766 - 0.0026i
0.0470	0.0712	0.0259 - 0.0507i	0.5518 - 1.0807i
0.0342	0.0741	-0.0319 - 0.0225i	-0.9334 - 0.6578i
0.0169	0.0742	-0.0110 - 0.0076i	-0.6517 - 0.4497i

Columns 5 through 8

0.9755	0.1091	0.0746	0.0950
0.8077	0.1711	0.2165	0.2251
0.4047	0.0286	0.1755	0.0455
0.5725	0.0997	0.0036	0.0509
0.6420	0.0523	0.0171	0.0419

0.5469	0.0220	0.0857	0.0383
0.8537	0.0132	0.0085	0.0274
0.5574	0.0036	0.0448	0.0378
0.7737	0.0775	0.0441	0.0691
0.4227	0.0087	0.0429	0.0425
0.7290	0.0241	0.0013	0.0425
0.9670	0.0631	0.0901	0.0774
0.2771	0.0372	0.0330	0.0704
0.9716	0.0542	0.0695	0.0643
0.6017	0.0263	0.1091	0.0317
0.1430	0.0133	0.0611	0.0577

2. Vertical Transect with an M value of 4, the variables are ln$K$ (kv) and lnK_d (kdv)

```
>> spectrum(kv,kdv,8)

ans =

  Columns 1 through 4

   0.0629   0.0555   -0.0469            -0.7455
   0.1581   0.1489   -0.0462 + 0.0595i  -0.2925 + 0.3763i
   0.1742   0.1548    0.0039 + 0.1564i   0.0224 + 0.8975i
   0.0878   0.0593   -0.0039 + 0.0619i  -0.0440 + 0.7050i

  Columns 5 through 8

   0.6296   0.1059   0.0600   0.0937
   0.2413   0.1456   0.1372   0.2314
   0.9071   0.0572   0.0684   0.0854
   0.7389   0.1124   0.0455   0.0784
```

3. Horizontal Transect with an M value of 16, the variables are f (fh) and η (etaf)

```
>>  spectrum(kh,etah,32)

ans =

  Columns 1 through 4

   0.0888   0.0520   -0.0663            -0.7459
   0.1083   0.1264   -0.0466 - 0.0893i  -0.4298 - 0.8242i
   0.0175   0.1586   -0.0234 + 0.0197i  -1.3368 + 1.1241i
   0.0530   0.0411    0.0317 - 0.0243i   0.5983 - 0.4584i
   0.0640   0.0043    0.0028 + 0.0068i   0.0439 + 0.1056i
   0.0308   0.0481   -0.0160 - 0.0188i  -0.5201 - 0.6103i
   0.0415   0.0199   -0.0184 - 0.0163i  -0.4432 - 0.3920i
   0.0270   0.0325    0.0119 - 0.0202i   0.4389 - 0.7473i
   0.0433   0.0553    0.0356 - 0.0279i   0.8206 - 0.6435i
   0.0231   0.0199   -0.0080 - 0.0048i  -0.3451 - 0.2086i
   0.0311   0.0366   -0.0216 + 0.0145i  -0.6956 + 0.4653i
   0.0407   0.0882    0.0489 + 0.0334i   1.2011 + 0.8217i
   0.0283   0.0987    0.0345 - 0.0001i   1.2156 - 0.0026i
```

0.0470	0.0941	0.0418 - 0.0507i	0.8908 - 1.0807i
0.0342	0.0564	-0.0203 - 0.0225i	-0.5943 - 0.6578i
0.0169	0.0687	-0.0053 - 0.0076i	-0.3127 - 0.4497i

Columns 5 through 8

0.9495	0.1091	0.0227	0.0580
0.7407	0.1711	0.1041	0.1773
0.3374	0.0286	0.2015	0.0364
0.7328	0.0997	0.0391	0.0783
0.1927	0.0523	0.0053	0.0242
0.4114	0.0220	0.0594	0.0452
0.7275	0.0132	0.0165	0.0255
0.6246	0.0036	0.0348	0.0373
0.8520	0.0775	0.0728	0.0837
0.1888	0.0087	0.0157	0.0396
0.5942	0.0241	0.0184	0.0365
0.9776	0.0631	0.1299	0.0922
0.4243	0.0372	0.0015	0.0783
0.9786	0.0542	0.1041	0.0775
0.4766	0.0263	0.1020	0.0345
0.0740	0.0133	0.0364	0.0604

4. Vertical Transect with an *M* value of 4, the variables are *f* and η (eta)

> > spectrum(kv,etav,8)

ans =

Columns 1 through 4

0.0629	0.0306	-0.0252	-0.4005
0.1581	0.1358	0.0083 + 0.0595i	0.0525 + 0.3763i
0.1742	0.1782	0.0640 + 0.1564i	0.3674 + 0.8975i
0.0878	0.0671	0.0264 + 0.0619i	0.3010 + 0.7050i

Columns 5 through 8

0.3291	0.1059	0.0096	0.0583
0.1681	0.1456	0.0763	0.1886
0.9193	0.0572	0.0637	0.0913
0.7692	0.1124	0.0431	0.0805

APPENDIX D
Calculation of Confidence Intervals for the Phase and Magnitude of the Cross-Spectra

The following tables detail the calculations of the confidence intervals for the phase and magnitude of the cross-spectrum.

S_{xx} and S_{yy}

The spectral estimates from MATLAB are scaled so that the results will be in a form that is consistent with that used in Gelhar (*1993*) and Priestley (*1989*):

$$P_{xx}(f) = \frac{4\pi}{\Delta} S_{xx}(\omega)|_{\omega = 2\pi f}$$

where ω $(2\pi f)$ is the angular frequency and Δ is the sample spacing.

C_{xy} and Q_{xy}

The cospectrum and quadrature spectrum as calculated by MATLAB

$w_{xy}^2 = S_{xy}^2/(S_{xx}S_{yy})$

The coherency squared function between X and Y

Tan Phase

The tangent of the phase, ϕ_{xy}, is $-C_{xy}/Q_{xy}$.

Magnitude

The magnitude, A_{xy}, is $\sqrt{(C_{xy2} + Q_{xy}^2)}$.

Variance Magnitude and Tan Phase

According to Priestley (*1981*), the variance of the of the magnitude and the tangent of the phase is:

$$var\hat{A}_{ij}(\omega) \sim \frac{C_w}{2N} A_{ij}^2 \left\{1 + \frac{1}{|w_{ij}(\omega)|^2}\right\}$$

$$var\{-\hat{q}_{ij}(\omega)/\hat{c}_{ij}(\omega)\} \sim \{1 + \tan_{ij}^2(\omega)\}^2 \frac{C_w}{2N}\left(\frac{1}{|w_{ij}(\omega)|^2} - 1\right)$$

95% Confidence Intervals

The 95% confidence interval is estimated as plus or minus two standard deviations. The standard deviation is the square root of the variance.

1. Horizontal Transect: $x = \ln K$

$y = \ln K_d$

Frequency	S_{xx}	S_{yy}	C_{xy}	Q_{xy}
0.0625	0.003533	0.004265	-0.00384	0
0.125	0.004309	0.00678	-0.00331	0.003553
0.1875	0.000696	0.007023	-0.00117	0.000784
0.25	0.002109	0.001023	0.000545	0.000967
0.3125	0.002546	0.00039	-0.00075	0.000271
0.375	0.001225	0.002487	-0.00105	0.000748
0.4375	0.001651	0.00148	-0.00129	0.000649
0.5	0.001074	0.001098	0.000107	0.000804
0.5625	0.001723	0.00144	0.000832	0.00111
0.625	0.000919	0.001114	-0.00063	0.000191
0.6875	0.001237	0.002184	-0.00128	0.000577
0.75	0.001619	0.002375	0.001397	0.001329
0.8125	0.001126	0.003127	0.000991	0.000004
0.875	0.00187	0.002833	0.001031	0.002017
0.9375	0.001361	0.002948	-0.00127	0.000895
1	0.000672	0.002952	-0.00044	0.000302

w_{xy}^2	Tan Phase	Magnitude	Variance Magnitude	Variance Tan Phase	95% CI Magnitude	95% CI Tan Phase
0.9755	0	0.003836	2.45e-06	0.002064	0.00313	0.090869
0.8077	0.932811	0.004859	4.34e-06	0.068439	0.004168	0.523217
0.4047	1.492386	0.001408	5.66e-07	1.259178	0.001504	2.244262
0.5725	-0.56379	0.00111	2.78e-07	0.106592	0.001055	0.652968
0.642	2.779412	0.000799	1.34e-07	3.489154	0.000733	3.735856
0.5469	1.409574	0.001293	3.89e-07	0.607514	0.001247	1.558863
0.8537	1.98773	0.001443	3.72e-07	0.345275	0.001219	1.175203
0.5574	-0.13366	0.000811	1.51e-07	0.067617	0.000777	0.520064
0.7737	-0.7491	0.001387	3.63e-07	0.058591	0.001204	0.484113

0.4227	3.291667	0.000657	1.19e-07	15.72314	0.000691	7.930483
0.729	2.22069	0.001405	3.85e-07	1.074964	0.001241	2.073609
0.967	-1.0509	0.001928	6.21e-07	0.012421	0.001577	0.222902
0.2771	-249	0.000991	3.72e-07	8.24e+08	0.00122	57421
0.9716	-0.51085	0.002265	8.56e-07	0.00382	0.00185	0.123613
0.6017	1.417778	0.001553	5.28e-07	0.492968	0.001453	1.404234
0.143	1.447368	0.000532	1.86e-07	4.718017	0.000862	4.344199

2. Vertical Transect: $x = \ln K$
$y = \ln K_d$

Frequency	S_{xx}	S_{yy}	S_{xy}	Q_{xy}	w_{xy}^2
0.833	0.000751	0.000662	-0.00056	0	0.6296
1.667	0.001887	0.001777	-0.00055	0.00071	0.2413
2.5	0.002079	0.001848	0.000047	0.001867	0.9071
3.33	0.001048	0.000708	-0.00005	0.000739	0.7389

Tan Phase	Magnitude	Variance Magnitude	Variance Tan Phase	95% CI Magnitude
0	0.00056	4.87e-08	0.035299	0.000441
0.776471	0.000899	2.50e-07	0.484708	0.000999
-0.02494	0.001867	4.40e-07	0.006153	0.001327
0.063005	0.00074	7.74e-08	0.02137	0.000556

3. Horizontal Transect: **f = ln*K* - E[ln*K*]**
n = η

S_{xx}	S_{yy}	C_{xy}	Q_{xy}	w_{xy}^2	Tan Phase	Magnitude	Variance Magnitude	95% CI Magnitude
0.003501	0.002069	-0.00264	0	0.9495	0	0.002638	1.17e-06	0.002167
0.004309	0.005029	-0.00185	-0.00355	0.7407	-0.52184	0.004008	3.10e-06	0.003523
0.000696	0.00631	-0.00093	0.000784	0.3374	1.187817	0.001217	4.83e-07	0.001389
0.002109	0.001635	0.001261	-0.00097	0.7328	1.304527	0.001589	4.91e-07	0.001401
0.002546	0.000171	0.000111	0.000271	0.1927	-0.41176	0.000293	4.36e-08	0.000417
0.001225	0.001914	-0.00064	-0.00075	0.4114	-0.85106	0.000982	2.72e-07	0.001043
0.001651	0.000792	-0.00073	-0.00065	0.7275	-1.12883	0.000978	1.87e-07	0.000864
0.001074	0.001293	0.000473	-0.0008	0.6246	0.589109	0.000933	1.86e-07	0.000863
0.001723	0.0022	0.001416	-0.00111	0.852	1.275986	0.0018	5.79e-07	0.001521
0.000919	0.000792	-0.00032	-0.00019	0.1888	-1.66667	0.000371	7.13e-08	0.000534
0.001237	0.001456	-0.00086	0.000577	0.5942	1.489655	0.001035	2.36e-07	0.000972
0.001619	0.003509	0.001946	0.001329	0.9776	-1.46407	0.002356	9.23e-07	0.001922
0.001126	0.003927	0.001373	-4.0e-06	0.4243	345	0.001373	5.20e-07	0.001442
0.00187	0.003744	0.001663	-0.00202	0.9786	0.824458	0.002614	1.14e-06	0.002132
0.00136	0.002244	-0.00081	-0.0009	0.4766	-0.90222	0.001206	3.70e-07	0.001217
0.000672	0.002733	-0.00021	-0.0003	0.074	-0.69737	0.000369	1.62e-07	0.000805

4. Vertical Transect: f = ln*K* - E[ln*K*]
n = η

S_{xx}	S_{yy}	C_{xy}	Q_{xy}	w_{xy}^2	Magnitude	Variance Magnitude	95% CI Amplitude
0.000751	0.000365	-0.0003	0	0.3291	0.000301	2.19e-08	0.000296
0.001887	0.001621	0.000099	0.00071	0.1681	0.000717	2.14e-07	0.000926
0.002079	0.002127	0.000764	0.001867	0.9193	0.002017	5.10e-07	0.001428
0.001048	0.000801	0.000315	0.000739	0.7692	0.000803	8.90e-08	0.000597

APPENDIX E

Calculation of the Macrodispersivity

The calculation for the macrodispersivity proposed by Garabedian et al. (*1988*) and Gelhar (*1993*) uses the model, model 1, in [1.6]:

$$\rho_b K_d = b\ln K + a + \kappa \qquad \text{[E-1]}$$

where ρ_b is the bulk density of the soil, b is the slope of the linear regression curve of K_d unto ln$K$, a is the y-intercept and κ is the residual term. The model proposed in this study, model 2, is presented in terms of ln$K$ and lnK_d:

$$\ln K_d = m\ln K + c + \eta \qquad \text{[E-2]}$$

where m is the slope of the linear regression curve of lnK_d unto ln$K$, c is the y-intercept and η is the residual term.

The longitudinal macrodispersivity is:

$$A_{11} = \frac{\sigma_f^2 \lambda}{\gamma^2}\left(1 - \frac{\gamma b}{Y}\right)^2 + \frac{1}{Y^2}(\sigma_\kappa^2 \lambda_\kappa) \qquad \text{[E-3]}$$

where σ_κ^2 and λ_κ is the variance of the residual term in the model 1. If the parameters of model 2 is used in the calculation of macrodispersivity the following conversion must be made, where lnK_d is:

$$\ln K_d = \bar{g} + g' \qquad \text{[E-4]}$$

and the expected value of K_d is:

$$\begin{aligned} E[K_d] &= e^{\bar{g}} e^{\sigma_g^2/2} \\ &= K_{dg} e^{\sigma_g^2/2} \end{aligned} \qquad \text{[E-5]}$$

where K_{dg} is the geometeric mean of K_d. The sorption coefficient, K_d is:

$$K_d = e^{\bar{g}+g'} = e^{\bar{g}} \cdot e^{g'} \quad \text{[E-6]}$$

the pertubation K_d' expressed as a function of g' is:

$$K_d' = e^{\bar{g}}[e^{g'} - E(e^{g'})] \cong K_{dg}\left[1 + g + g'^2/2 + \ldots\ldots - e^{\sigma_g^2/2}\right] \quad \text{[E-7]}$$

$$\therefore K_d' \cong K_{dg} g' \qquad \sigma_{K_d}^2 \cong K_{dg}^2 \sigma_g^2 \quad \text{[E-8]}$$

By using the perturbation K_d', model 1 becomes:

$$\rho_b K_d' = bf + \kappa \quad \text{[E-9]}$$

Dividing through by ρ_b and substituting [E-8]:

$$K_{dg} g' = \frac{b}{\rho_b} f + \frac{\kappa}{\rho_b} \quad \text{[E-10]}$$

By substituting [E-9], K_d' can be expressed in terms of the perturbation g':

$$g' = \frac{b}{\rho_b K_{dg}} f + \frac{\kappa}{\rho_b K_{dg}}$$
$$\therefore \; b = \rho_b K_{dg} m \qquad \kappa = \rho_b K_{dg} \eta \quad \text{[E-11]}$$
$$\sigma_\kappa^2 = (\rho_b K_{dg})^2 \sigma_\eta^2$$

which is the same as the perturbation of g' which results from model 2:

$$g' = mf + \eta \quad \text{[E-12]}$$

Therefore the slope, residual value and variance of the residual of model 1, expressed as a function of model 2 is:

$$b = \rho_b K_{dg} m \quad \kappa = \rho_b K_{dg} \eta \quad \sigma_\kappa^2 = (\rho_b K_{dg})^2 \sigma_\eta^2 \quad \text{[E-13]}$$

The correlation length of the residual in model 2 (λ_η) is substituted for the correlation length of the residual in model 1 (λ_κ), this assumption is made based on Figure 4.3 in Gelhar (*1993*).