WATER RESOURCES
research center

Publication No. 106

FIELD VALIDATION OF A DISPERSION MODEL BASED ON GEOLOGICAL PARAMETERS

by

David A. Chin

Department of Civil and Architectural Engineering
University of Miami
Coral Gables, Florida 33124

UNIVERSITY OF FLORIDA

Report G1415-02

Publication No. 106

FIELD VALIDATION OF A DISPERSION MODEL BASED ON GEOLOGICAL PARAMETERS

by

David A. Chin

Department of Civil and Architectural Engineering
University of Miami
Coral Gables, Florida 33124

Technical Report No. CEN-88-1

March 1988

Subcontract between the Florida Water Resources Research Center, University of Florida, and the University of Miami under United States Department of Interior Grant Number 14-08-0001-G1415

The activities on which this report is based were financed in part by the Department of the Interior, U.S. Geological Survey, through the Florida Water Resources Research Center.

FLORIDA WATER RESOURCES RESEARCH CENTER
424 Black Hall
UNIVERSITY OF FLORIDA
Gainesville, Florida 32611

1988

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

A_{ij} = dispersivity tensor;

A_L = observed longitudinal dispersivity;

$c(x,y,z,t)$ = concentration distribution;

c_{ijkl} = concentration in grid ijk at time step l;

D_{ij} = dispersion coefficient due to head transient;

$D_{i'j'}$ = dispersion coefficient due to K variations;

e_i = error in grid i (continuous release);

$\bar{e}$ = mean error (continuous release);

erf = error function;

F_{ab} = F distribution with a and b degrees of freedom;

$G(\omega_j)$ = gain of Y(t) over X(t) at ω_j;

h_i = measured head at node i;

i = $(-1)^{-1/2}$;

$J_X(\omega_j)$ = Fourier transform of X(t);

$J_x(\omega_j)$ = Fourier transform of x(t);

$J_{\varepsilon_X}(\omega_j)$ = Fourier transform of $\varepsilon_X(t)$;

$J_Y(\omega_j)$ = Fourier transform of Y(t);

$J_y(\omega_j)$ = Fourier transform of y(t);

$J_{\varepsilon_Y}(\omega_j)$ = Fourier transform of $\varepsilon_Y(t)$;

K = hydraulic conductivity;

L = correlation length scale;

M = mass of released tracer (instantaneous release);

$\dot{M}$ = mass release rate (continuous release);

m = number of sub-records in time series;

N = number of time intervals of length Δt;

N_c = number of components with uniform power spectrum;

N_p = number of particles used to discretize source;

N_{tot} = total number of particles released;

n = porosity, also used for number of data points in time series record;

n_p = number of measurement points in Δt;

n_{ijkl} = number of particles in grid ijk at time step l;

n_{int} = number of interpolation points;

n_x = number of source particles in x direction;

n_y = number of source particles in y direction;

n_z = number of source particles in z direction;

P_{ε_X} = average power at each frequency in ε_X;

$P^{(r)}_{\varepsilon_X}$ = estimate of P_{ε_X} based on N_c components;

P_{ε_Y} = average power at each frequency in ε_Y;

$P^{(r)}_{\varepsilon_Y}$ = estimate of P_{ε_Y} based on N_c components;

p_{ijkl} = particle probability distribution;

t = time;

v = magnitude of seepage velocity;

v' = temporal velocity fluctuation;

v'' = spatial velocity fluctuation;

v_i = i component of seepage velocity;

$v_{i,q}$ = q th measurement of v_i in interval p;

$\bar{v}_{i,p}$ = average value of v_i in interval p;

$X(t)$ = time series with non-random and random components;

$x(t)$ = non-random component of $X(t)$;

$Y(t)$ = time series with non-random and random components;

$y(t)$ = non-random component of $Y(t)$;

x_o = coordinate of concentration grid origin;

x_1, x_2 = source coordinates;

y_o = coordinate of concentration grid origin;

y_1, y_2 = source coordinates;

z_o = coordinate of concentration grid origin;

z_1, z_2 = source coordinates;

α_i = i component of dispersivity;

α_L = longitudinal dispersivity;

α_T = transverse dispersivity;

β_{ij} = direction cosines between i and j axes;

$\gamma(\omega_j)$ = coherence between $x(t)$ and $y(t)$ at ω_j;

γ_o = zero coherence confidence limit;

Δt = simulation time step;

$\Delta t'$ = time interval for head measurements;

Δx = concentration grid dimension;

Δx_i = change in x_i coordinate;

Δy = concentration grid dimension;

Δz = concentration grid dimension;

$\varepsilon_X(t)$ = random component of $X(t)$;

$\varepsilon_Y(t)$ = random component of $Y(t)$;

η_i = nodal coordinate;

ξ_i = nodal coordinate; and

σ_K = variance of log-hydraulic conductivity;

σ_{11}^2 = longitudinal variance of tracer distribution;

σ_{ij}^2 = variance of tracer distribution;

$\sigma^2_{v_i}$ = variance of random fluctuations.

$\sigma^2_{\varepsilon_X}$ = variance of ε_X;

$\sigma^2_{\varepsilon_Y}$ = variance of ε_Y;

$\varphi(\omega_j)$ = phase shift of Y(t) relative to X(t);

$\chi^2_{a,b}$ = chi-square variate; and

ω_j = Fourier frequency, $2\pi j/n$.

Superscripts

* = non-dimensionalized variable, or complex conjugate.

Other:

$\langle x \rangle$, $\bar{x}$ = average of x over m sub-records.

SUMMARY

A non-empirical dispersion model has been developed which utilizes only measurable hydrogeological parameters to predict the migration of conservative contaminants in porous media. The model has been applied in an idealized formation and found to be in excellent agreement with the relevant analytic solution of the advection-dispersion equation. Since the model depends on time series measurements of hydrologic data, gaps in the data caused by instrument malfunction are an important and frequent area of concern. A technique for filling in missing data has been presented. The validity of the dispersion model has been investigated at a site where extensive measurements of the migration of a conservative tracer are available. The predicted plume dispersivity was found to vary with scale in a manner very similar to that found in most porous formations. The observed plume trajectory and mixing characteristics are found to be reproduced quite accurately by the non-empirical model. Measured concentrations at monitoring wells within the formation also show good agreement with the predicted values.

CHAPTER I

INTRODUCTION

1.1 Background

Dispersion models are primarily used for estimating the impact of potential or existing contamination sources on aquifers. It is generally assumed that dispersion is governed by the advection-dispersion equation and the focus of research has been primarily on estimating the dispersion coefficient from measurable hydrogeological parameters (e.g. Gelhar et al., 1983, Chin, 1986a, Neuman et al., 1987). In general, these theoretical formulations apply to idealized cases such as perfectly stratified formations and their ability to predict mixing in real aquifers has not been established. Furthermore, the geological parameters in these formulations are quite difficult to estimate with certainty. Geological parameters typically included in theoretical formulations are the correlation length scale and the variance of the hydraulic conductivity. Another major drawback to theoretical dispersion models is that they generally have been developed for cases where the size of the tracer cloud significantly exceeds the correlation length scales of the hydraulic conductivity. In many cases, accurate prediction of the dispersion of a cloud on smaller length scales is at least as important (Chin, 1987, Yeh, 1987). Theoretical formulations for dispersion on these smaller scales have not been developed for generalized porous formations. The more popular approaches to estimating dispersion in porous media have been either to calibrate a dispersion model using observations of tracer dispersion (Anderson, 1979) or to use dispersion parameters measured at similar

sites. The main drawbacks of calibration approaches are: (1) the calibrated parameters are rarely stationary and hence extrapolation into the future is uncertain; (2) using natural tracers, hydrogeochemical interactions can seldom be separated from natural mechanical dispersion; and (3) using man-made conservative tracers, the dispersion process can usually only be studied on relatively small length scales. In using dispersion coefficients measured at similar sites, it is generally impossible to estimate the sensitivity of the dispersion coefficient to actual site differences.

In contrasting the two main approaches to estimating dispersion in porous media, theoretical vs. calibration, practical considerations indicate that the theoretical approach would be more useful, although further refinements are undoubtedly necessary. The prediction of dispersion based only on measurable hydrogeological parameters using theoretical relations is an appealing prospect, while uncertainties, expense, and length scale limitations are implicit in the calibration approach.

1.2 Role of the Present Study

A methodology for predicting the dispersion of contaminants in porous media based only on measurable hydrogeological parameters has been developed by Chin, 1988b. In order to assess the performance of the Chin approach, the model is applied to a site where extensive hydrogeological measurements and data on the mixing of a conservative tracer are available. Since the predictions of the Chin model are not dependent on the measured tracer distribution, the comparisons between the predicted and measured movement of the tracer will provide an

unbiased measure of the adequacy of the Chin approach and also give an indication of the accuracy to be expected when applying this model. The author of this report is unable to find any previous study in which a dispersion model based only on hydrogeological data has been validated using independent field data.

1.3 Organization of the Report

This study is organized into four major chapters. Chapter II describes the formulation of the Chin model. Chapter III describes the results of the verification study in which the Chin model was applied to idealized formations and the results compared with analytic solutions of the advection-dispersion equation. The Chin model utilizes measured piezometric head data at defined nodes, and gaps in the measured data at some nodes impose limitations on the predictive capability of the model. Chapter IV describes a method that has been developed to fill in missing data, based on the inter-relationship between the measurements at different nodes. Chapter V describes the results of the validation study, where the predictions of the Chin model are compared with independent field measurements. Chapter VI provides an overall assessment of the results obtained during this study.

CHAPTER II

FORMULATION OF DISPERSION MODEL

2.1 Introduction

Dispersion of contaminants in porous media is generally assumed to be governed by the advection-dispersion equation (ADE). Classical numerical models are concerned with solving this equation in an efficient and accurate manner. Aside from numerical algorithms, the primary limitation to predicting contaminant transport in groundwater is the accuracy with which the seepage velocity and dispersivity are estimated. Seepage velocity is typically determined by solving the Darcy and continuity equations for an estimated hydraulic conductivity distribution and given boundary conditions. The dispersion coefficient may be estimated based on previous experience (e.g. Gelhar et al., 1985), or from the hydrogeological characteristics of the porous medium (e.g. Chin, 1986a, Gelhar et al., 1983, Neuman et al., 1987). Fundamental questions as to whether the actual velocity field is composed of the explicit seepage velocity plus the velocity fluctuations implicit within the dispersivity, and whether the numerical time step exceeds the Lagrangian time scale associated with the dispersivity are important requirements that are frequently overlooked. A related area of concern is that, in most field cases studied, there is significantly more hydrogeological data available than actually utilized. For example, head measurements may be available at daily intervals at several locations within the study area but are seldom analysed in detail to see how they may best be used in dispersion simulations. A diagnostic modeling approach uses these measured data

directly in the simulation instead of generating them in a numerical model. This chapter describes a diagnostic model that has been developed to efficiently utilize measured data on the spatial and temporal head fluctuations, and partition the estimated seepage velocity field into advection and diffusion components.

2.2 Formulation

2.2.1 Estimation of Flow

The traditional approach to estimating flow is to obtain a finite-difference solution of the Darcy and continuity equations for given boundary conditions. The adequacy of the assumed hydraulic conductivity field may be checked by comparing the computed heads with field measurements. Adjustment of the hydraulic conductivities to yield better agreement with observed heads is possible at this stage (i.e., the inverse problem). Uncertainty concerning boundary conditions and the non-uniqueness of the solution are areas of concern when this method is applied. In many large scale problems there are several monitoring wells within the study area and daily measurements of head are typically available. Considering that the spacing of the monitoring wells may correspond to the scale of the pump tests used to estimate the hydraulic conductivity, and assuming a fairly smooth water table in the area, then it would be reasonable to use the monitoring well locations as nodes in a finite-element grid and then estimate the seepage velocity from these heads and the measured hydraulic conductivity field. Using this method the uncertainties involved in assigning boundary conditions are eliminated.

To illustrate the method, consider the array of monitoring wells

shown in Fig.2.1. If in any element the nodal heads are h_1, h_2, and h_3 and the station coordinates are (ξ_1,η_1), (ξ_2,η_2), and (ξ_3,η_3) then the components of the seepage velocity v_i in any layer, j, within the element is given by

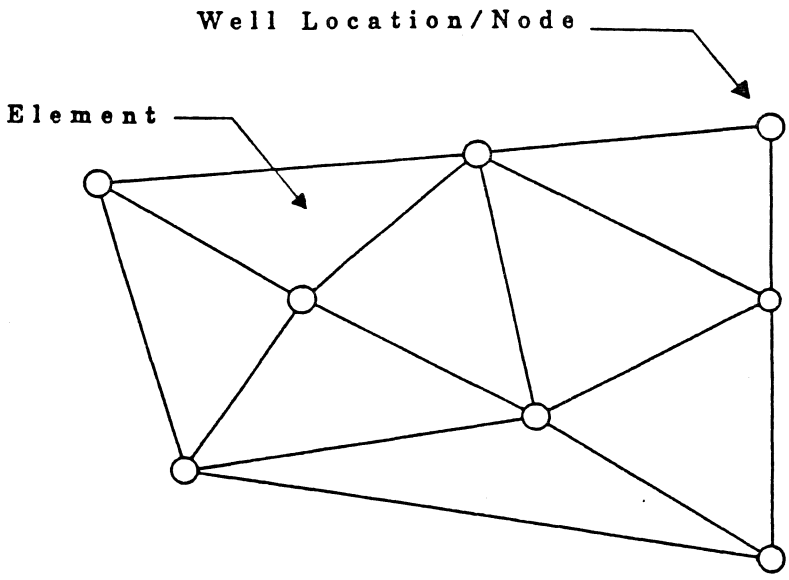
$$v_1(j) = \frac{K_j}{n\,d}\left[(\eta_3 - \eta_2)h_1 + (\eta_1 - \eta_3)h_2 + (\eta_2 - \eta_1)h_3\right] \quad \ldots (2.1a)$$

$$v_2(j) = \frac{K_j}{n\,d}\left[(\xi_2 - \xi_3)h_1 + (\xi_3 - \xi_1)h_2 + (\xi_1 - \xi_2)h_3\right] \quad \ldots (2.1b)$$

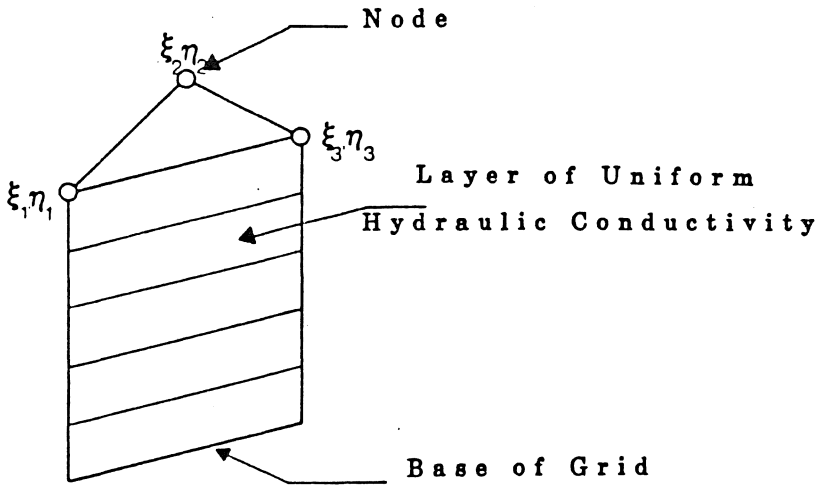
where d is defined by the relation,

$$d = \xi_1(\eta_2 - \eta_3) - \eta_1(\xi_2 - \xi_3) + (\xi_2\eta_3 - \xi_3\eta_2) \quad \ldots\ldots (2.1c)$$

where K_j is the hydraulic conductivity in layer j within the finite element, and n is the porosity. The major assumptions in this procedure are that the head variation is approximately linear in any layer and within each element, the formation is fairly homogeneous within each layer, and vertical head gradients are negligible. Although horizontal isotropy has been assumed, an equation of the same form as Eq.2.1 may be written for horizontally anisotropic media. A major concern in utilizing this method is that the time steps used in the interpolation of seepage velocity fluctuations within each element correspond to some fraction of the period of coherent nodal fluctuations. This is necessary if the space-time spectrum of the velocity field is to be accurately modeled. Combining the coherence spectrum of the nodal head fluctuations with the confidence limits for significant coherence will



(a) Plan View



(b) Elevation View

Figure 2.1 Typical Monitoring Well Array

yield the minimum frequency that is significantly coherent (e.g. Chin, 1988a). Choosing a time step, Δt, that is some fraction of the period corresponding to the minimum coherent frequency will provide a reasonably accurate description of the spatial and temporal fluctuations in the velocity field for length scales greater than the size of the grid element. Velocity fluctuations with length scales smaller than the grid element may also be interpolated using Eq.2.1. The Mean Value Theorem guarantees that these sub-grid interpolated velocities, based in interpolating incoherent head fluctuations, do occur at some location within the element. Therefore, the temporal statistics of these fluctuations within the grid element may be determined. The next section describes how these sub-grid velocities are used to determine the dispersion of contaminants.

2.2.2 Contaminant Transport Model

Based on the coherence analysis, the minimum period of head fluctuations that show significant coherence are determined. The time step, Δt, that should be used in the transport simulations is a fraction, say 0.25, of the minimum coherent period. Since the head measurements will typically be available at smaller time steps, $\Delta t'$, (1 day for USGS wells) the variance of the seepage velocities, estimated from Eq.2.1, at time scales less than Δt is given by

$$\sigma^2_{v_i} = \frac{1}{N\,m} \sum_{p=1}^{N} \sum_{q=1}^{m} (v_{i,q} - \bar{v}_{i,p})^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots (2.2)$$

where N is the number of intervals of length Δt, m is the number of measurement points within each interval, $v_{i,q}$ is the value, v_i, of the q th point within the p th interval, and $\bar{v}_{i,p}$ is the average value of

v_i in the p th interval. Since $\sigma^2_{v_i}$ is the variance of the random fluctuations whose time scale is less than Δt, a dispersion coefficient may be formed that is solely due to temporal velocity fluctuations not accounted for explicitly in the calculated advection velocity. Such a dispersion coefficient may be defined by the relation,

$$D_{ij} = \begin{cases} \frac{1}{2}\sigma^2_{v_i}\,\Delta t', & i = j \\ 0 & i \neq j \end{cases} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (2.3)$$

It should be noted that this formulation utilizes the condition that the estimated velocity fluctuations with time step $\Delta t'$ are independent and random. It is important to note that the traditional dispersion coefficient in porous media accounts for spatial variations in the seepage velocity due to random spatial variability in the hydraulic conductivity and it is an interesting problem to combine both effects. For isotropic media, the traditional dispersion coefficient is given by

$$D_{i'j'} = \begin{cases} \alpha_{i'}|v|, & i' = j' \\ 0 & i' \neq j' \end{cases} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (2.4)$$

where the primed axes are parallel and orthogonal to the mean flow direction, $\alpha_{i'}$ are the dispersivities, and $|v|$ is the magnitude of the seepage velocity. In order to combine the effects of temporal and spatial variability it is proposed that the dispersion process be simulated using a stochastic technique rather than simply solving the advection-dispersion equation (ADE). The reason for taking this

approach is that the direct simulation of the dispersion process is physically appealing, estimates of contamination migration paths may be determined with relatively little computational effort, and the negligible numerical dispersion associated with this technique. Classical approaches to the contaminant transport problem have used finite-difference (FD), finite-element (FEM), and method of characteristics (MOC) formulations to solve the ADE. These approaches have shown that FEM and MOC methods are generally more desirable than FD models, except in cases where the FD models use second order approximations for the derivatives (Anderson, 1979). The main drawbackes of FEM and higher-order FD models are associated with undesirable numerical dispersion, while the MOC approach tends to have a problem maintaining a mass balance in the system (Konikow and Bredehoeft, 1978). Using the stochastic approach, the initial tracer distribution is discretized into a large number of particles. Each particle is then advected with a deterministic velocity estimated by interpolation from the finite-element grid, and a random velocity estimated from the dispersion coefficient. The simulation process within each time interval, Δt, is composed of two steps. The first step is to interpolate the deterministic velocity, v_i, at the particle location using Eq.2.1. The next step is to determine a random velocity perturbation due to the temporal fluctuations in the velocity field. Based on Eq.2.3 we may generate a random velocity, v_i', with zero mean and variance $\sigma^2_{v_i} \Delta t'/\Delta t$. Hence due to random small scale temporal fluctuations in velocity, a particle may move with an average velocity $v_i + v_i'$ over a time interval Δt. However, because of spatial heterogeneities in hydraulic conductivity, additional dispersion will

occur and this may be simulated with an additional random velocity, v''_i, which may be generated as a random normal deviate with zero mean and variance $\sigma^2_{v_{i'}} = 2\alpha_{i'}|v + v'|/\Delta t$, where $\sigma^2_{v_{i'}}$ is the variance of the velocity components in the primed coordinate system whose axes are parallel and normal to the direction of the $v + v'$ vector. The advection procedure described here is obviously a simplification of a more complex process; however, the method is supported by the fact that it is asymptotically correct as small scale temporal fluctuations decrease ($v' \to 0$) and also as small scale hydraulic conductivity heterogeneities decrease ($\alpha_i \to 0$). The most appealing feature of the simulation technique is that the velocity field has been correctly partitioned into advection and diffusion components and the heads at the nodes are exactly equal to those measured. The particle translation procedure is summarized in the equation,

$$\Delta x_i = (v_i + v'_i + \beta_{i',i} v''_{i'}) \Delta t \quad \ldots\ldots\ldots\ldots\ldots\ldots (2.5a)$$

where

$$v'_i = N\left\{0, \sigma_{v_i} \left(\frac{\Delta t'}{\Delta t}\right)^{1/2}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.5b)$$

$$v''_{i'} = N\left\{0, \left(\frac{2\alpha_{i'}|v + v'|}{\Delta t}\right)^{1/2}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots (2.5c)$$

where $N(0,\sigma)$ indicates a random normal deviate with zero mean and standard deviation σ, and $\beta_{i',i}$ is the direction cosine between the i

and i' axes.

The tracer concentration within the porous formation is determined by superimposing a rectangular concentration grid of dimensions Δx, Δy, and Δz and calculating the probability distribution of the particles. For instantaneous releases the grid averaged concentration is given by

$$c_{ijkl} = \frac{M}{\Delta x\, \Delta y\, \Delta z}\, p_{ijkl} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.6)$$

where c_{ijkl} is the average concentration in grid i,j,k at time $l\Delta t$, M is the released mass, and p_{ijkl} is the probability that a particle located at the source at time zero will be found in grid ijk at time $l\Delta t$. It is convenient to define a non-dimensional concentration, c^*_{ijkl}, by the equation

$$c^*_{ijkl} = \frac{c_{ijkl}}{M}\, \Delta x\, \Delta y\, \Delta z \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.7)$$

Hence Eq.2.6 becomes

$$c^*_{ijkl} = p_{ijkl} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.8)$$

The quantity p_{ijkl} is calculated in the model using the relation

$$p_{ijkl} = \frac{n_{ijkl}}{N_p} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.9)$$

where n_{ijkl} is the number of particles in element ijk at time $l\Delta t$, and N_p is the total number of particles used to discretize the source. In the case of a continuous source, the mass released in any time interval

Δt is $\dot{M}\Delta t$ where $\dot{M}$ is the mass flux rate from the source. In this case the concentration is given by

$$c_{ijkl} = \frac{\int_0^{l\Delta t} \dot{M}(\tau)\,d\tau}{\Delta x\ \Delta y\ \Delta z}\ p_{ijkl} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.10)$$

In this case we may define a non-dimensional concentration by

$$c^*_{ijkl} = \frac{c_{ijkl}\ \Delta x\ \Delta y\ \Delta z}{\frac{1}{l}\int_0^{l\Delta t} \dot{M}(\tau)\,d\tau} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.11)$$

Hence Eq.2.10 becomes

$$c^*_{ijkl} = l\ p_{ijkl} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.12)$$

2.3 Simulation Parameters

Two primary parameters are associated with the simulation technique: the number of points used to discretize the source in each coordinate direction, n_x, n_y, and n_z; and the number of ensembles generated, n_e. The number of points used to discretize the source is related to the geometric definition of the source and increasing the number of these particles will improve the results for times when the concentration distribution is sensitive to the shape of the source. The number of ensembles is associated with the smoothness of the solution. A model verification study will be described subsequently in which the sensitivity of the solution to these parameters will be demonstrated.

The basic criterion governing the selection of n_x, n_y, n_z, and n_e is that the estimated concentration not be significantly changed by increasing the value of any of these parameters. This criterion has corollaries such as: it is easy to obtain an accurate solution if a large concentration grid is used, and a solution that is accurate at one time may eventually become inaccurate since the particle density may be so depleted that the selected grid size no longer yields a smooth solution. In the case of continuous releases, an additional requirement for accuracy is that the solution be unaffected by varying the time interval Δt. This criterion is related to the condition that the simulated mass release is in 'puffs', hence if the concentration grid is small relative to the 'puff' spacing then the solution will not resemble that of an instantaneous release. A remedy to this condition would be either to decrease the time step, increase the concentration grid dimensions, or linearly interpolate points from successive releases. This latter option has been found to be most economical in problems requiring a fine concentration grid, since it involves significantly less computational effort than decreasing the time step.

CHAPTER III

VERIFICATION OF THE DISPERSION MODEL

3.1 Verification Approach

The proposed model has been verified for instantaneous and continuous releases into an idealized homogeneous isotropic porous medium where analytic solutions are available. To make the configuration somewhat realistic, the location of monitoring wells within a particular area of Dade county Florida has been used in the study. This configuration is shown in Fig.3.1. The porous formation has a uniform hydraulic conductivity of 4572 m/day (15,000 ft/day), and a porosity of 0.2. The heads assigned to the well locations correspond to a linear hydraulic gradient of 0.008 percent and hence the mean seepage velocity is 1.83 m/day (6 ft/day). The longtitudinal and lateral dispersivities, α_L and α_T, were assumed to be 3.05 m (10 ft) and 0.31 m (1 ft) respectively. For the purposes of this study, the porous formation has been made sufficiently deep that a tracer released at the center of the formation does not encounter any vertical boundaries. A depth of 302 m (990 ft) was found to be more than adequate for this purpose.

3.1.1 Instantaneous Release

In the case of an instantaneous release, a parallepiped source was assumed. The dimensions of the source were 27.4 m x 27.4 m x 27.4 m (90 ft x 90 ft x 90 ft) and a mass, M, equal to 1 mass unit was released instantaneously. The analytic solution describing the resulting concentration distribution is given by (Prakash, 1984)

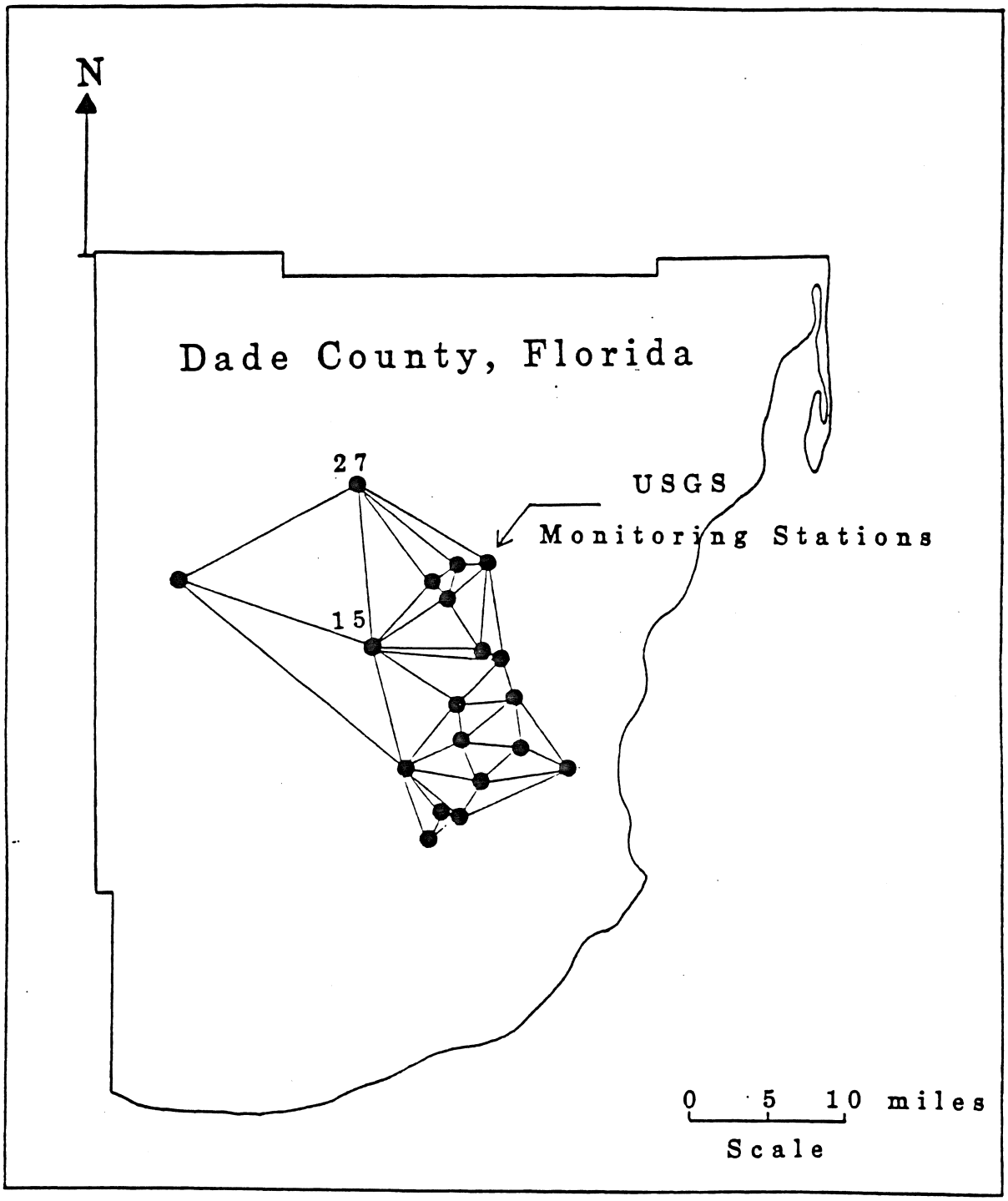


Figure 3.1: Monitoring Wells in Dade County, Florida

$$c(x,y,z,t) = \frac{M}{8(x_2-x_1)(y_2-y_1)(z_2-z_1)} \left\{ \text{erf} \frac{x-vt-x_1}{(4\alpha_L vt)^{1/2}} - \right.$$

$$\left. \text{erf} \frac{x-vt-x_2}{(4\alpha_L vt)^{1/2}} \right\} \left\{ \text{erf} \frac{y-y_1}{(4\alpha_T vt)^{1/2}} - \frac{y-y_2}{(4\alpha_T vt)^{1/2}} \right\} \left\{ \text{erf} \frac{z-z_1}{(4\alpha_T vt)^{1/2}} \right.$$

$$\left. - \text{erf} \frac{z-z_2}{(4\alpha_T vt)^{1/2}} \right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1)$$

where $c(x,y,z,t)$ is the concentration at any point (x,y,z) at time t, (x_1,x_2), (y_1,y_2), and (z_1,z_2) are the beginning and ending locations of the source in the x,y, and z coordinate directions respectively, v is the seepage velocity in the x direction, and t is time. For the purpose of comparing this solution with the numerical solution it is convenient to define the following non-dimensional variables:

$$x^* = \frac{x}{\Delta x} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.2)$$

$$y^* = \frac{y}{\Delta y} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.3)$$

$$z^* = \frac{z}{\Delta z} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.4)$$

$$t^* = \frac{t}{\Delta t} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.5)$$

$$c^* = \frac{c\ \Delta x\ \Delta y\ \Delta z}{M} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.6)$$

$$v^* = \frac{v\ \Delta t}{\Delta x} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.7)$$

$$\alpha_L^* = \frac{\alpha_L}{\Delta x} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.8)$$

$$\alpha_T^* = \frac{\alpha_T}{\Delta y} = \frac{\alpha_T}{\Delta z} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.9)$$

Combining Eqs.3.2 to 3.9 with Eq.3.1 yields

$$c^* = \frac{1}{8(x_2^*-x_1^*)(y_2^*-y_1^*)(z_2^*-z_1^*)} \left\{ \text{erf} \frac{x^*-v^*t^*-x_1^*}{(4\alpha_L^* v^* t^*)^{1/2}} - \text{erf} \frac{x^*-v^*t^*-x_2^*}{(4\alpha_L^* v^* t^*)^{1/2}} \right\} \left\{ \text{erf} \frac{y^*-y_1^*}{(4\alpha_T^* v^* t^*)^{1/2}} - \text{erf} \frac{y^*-y_2^*}{(4\alpha_T^* v^* t^*)^{1/2}} \right\}$$

$$\left\{ \text{erf} \frac{z^*-z_1^*}{(4\alpha_T^* v^* t^*)^{1/2}} - \text{erf} \frac{z^*-z_2^*}{(4\alpha_T^* v^* t^*)^{1/2}} \right\} \quad \ldots\ldots\ldots\ldots (3.10)$$

Since the numerical model yields grid averaged concentrations, we may derive the theoretical grid averaged concentration from Eq.3.10 using the relation

$$c_{ijkl}^* = \int_{x_o^*+(i-1)}^{x_o^*+i} \int_{y_o^*+(j-1)}^{y_o^*+j} \int_{z_o^*+(k-1)}^{z_o^*+k} c^*(x^*,y^*,z^*,t^*)\,dx^*\,dy^*\,dz^* \quad \ldots (3.11)$$

where (x_o^*, y_o^*, z_o^*) is the non-dimensionalized origin of the concentration grid (normalized by Δx, Δy, and Δz respectively) and it should be noted that l and t^* are equal. In the present study: $x_o^* = 160$, $y_o^* = 210$, and $z_o^* = 0$.

To verify the accuracy of the numerical method, the simulated variation of maximum grid averaged concentration with time is compared with the theoretical result in Fig.3.2 for the source discretized into 1000 particles ($n_x = n_y = n_z = 10$). Results are shown for 1, 10, and 20 ensembles (n_e). These results show that the method is capable of producing quite accurate results, even with a small number of ensembles, with accuracy increasing as the number of ensembles increases. The variation of the average error magnitude with n_e is shown in Fig.3.3. Further simulations have shown that, for each source discretization, as the number of ensembles, n_e, increases, then the accuracy of the numerical solution also increases. In fact, if the concentration distribution is not very sensitive to the shape of the source, then the accuracy has been found to depend primarily on the total number of particles used in the simulation, i.e., $n_x \times n_y \times n_z \times n_e$. This relation is shown in Fig.3.4 along with the processing time for the computer runs. The results indicate that the theoretical solution may be fairly well approximated with a relatively small amount of computational effort.

The results presented in this section generally show that the proposed numerical method for simulating the transport of instantaneous contaminant releases is capable of producing good results with fairly little computational effort and excellent results with increased effort.

3.1.2 Continuous Release

Theoretical concentration distributions for continuous releases are derived directly from instantaneous releases. The theoretical concentration distribution resulting from a steady continuous release

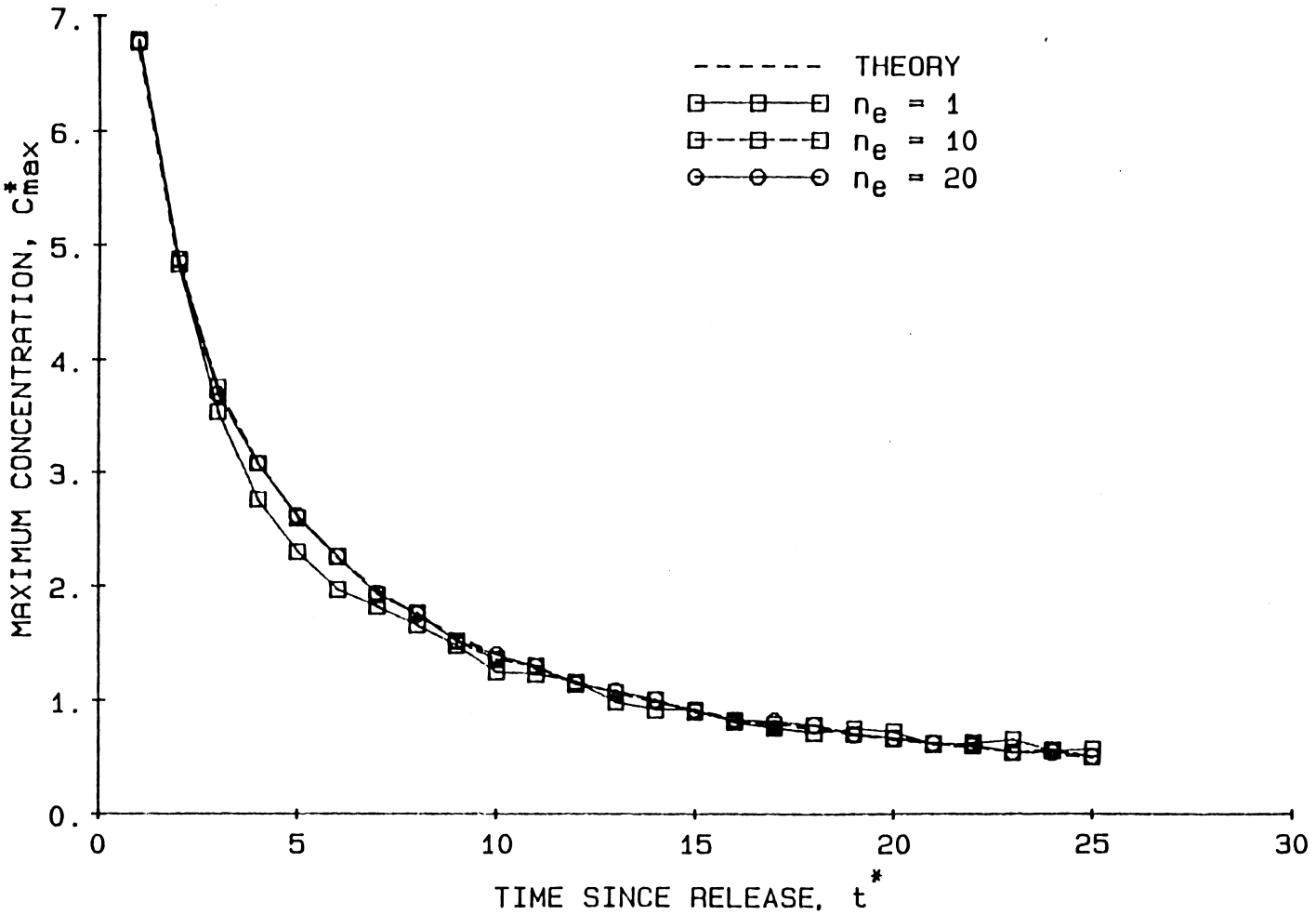


Figure 3.2: Maximum Grid Averaged Concentration vs. Time for an Instantaneous Release

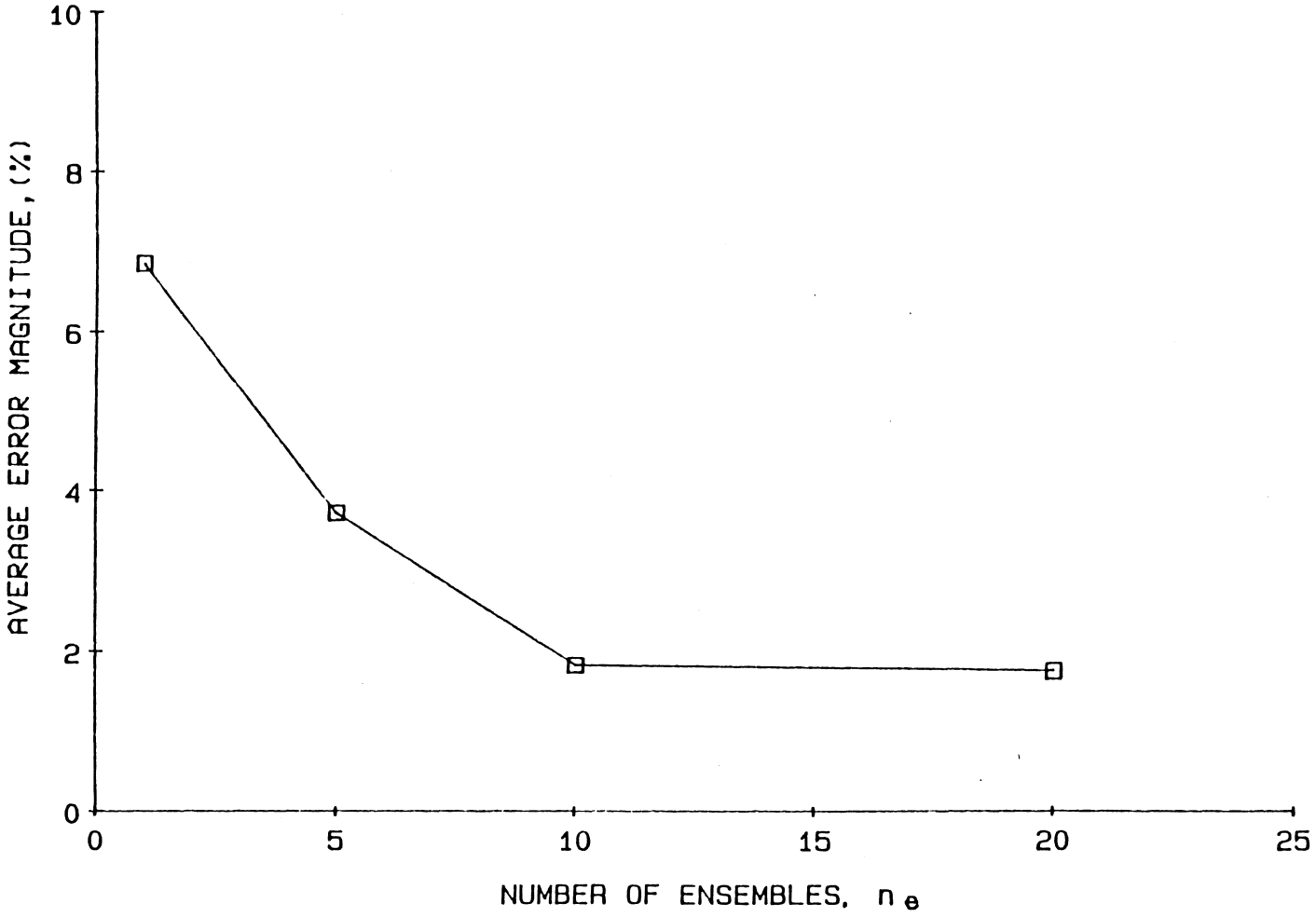


Figure 3.3: Average Error Magnitude vs. Number of Ensembles for Instantaneous Release

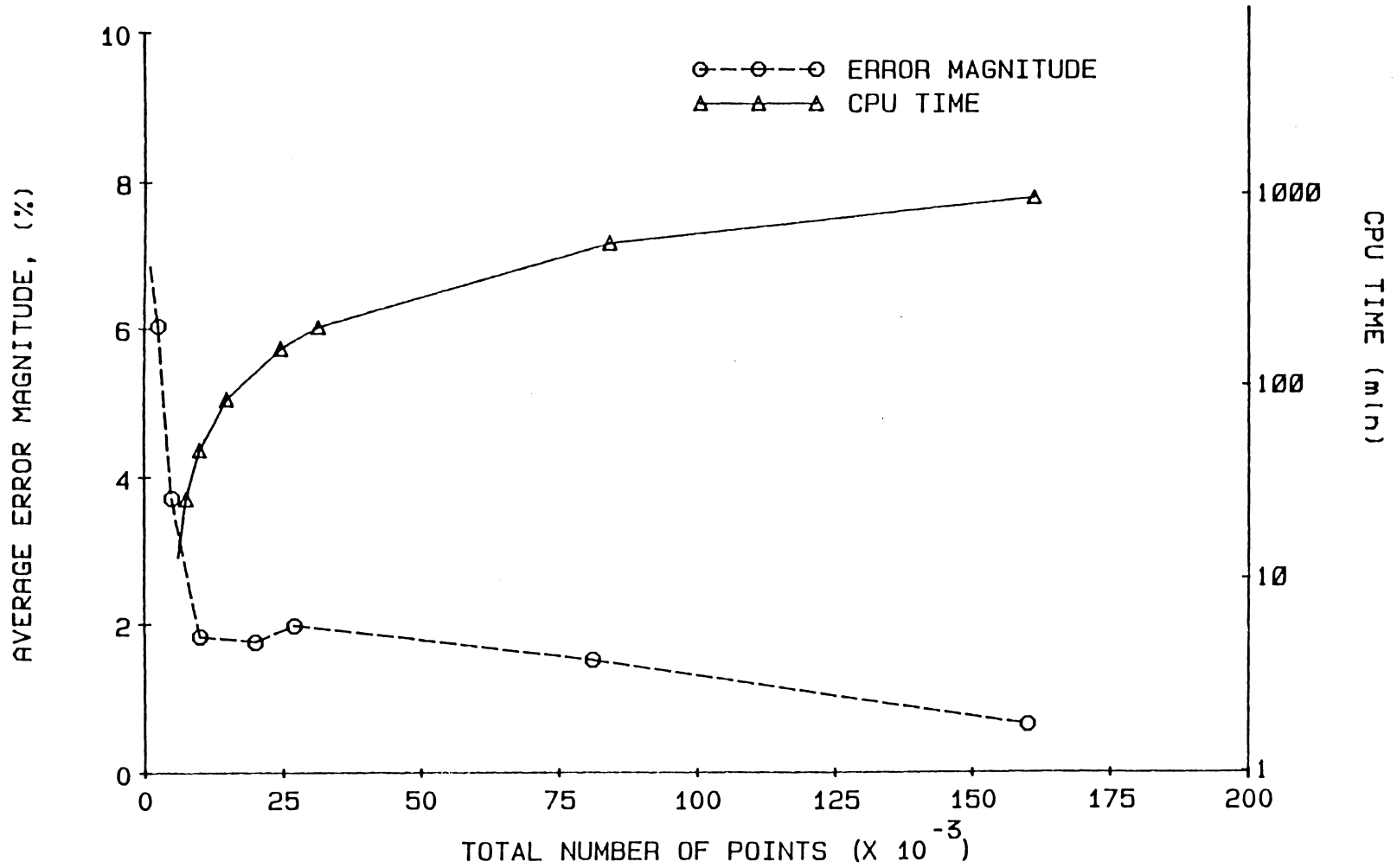


Figure 3.4: Accuracy of Simulation and CPU Time vs. Number of Particles

of contaminant from a source is given by (Prakash, 1984)

$$c^*(x^*,y^*,z^*,t^*) = \int_0^{t^*} k^*(x^*,y^*,z^*,t^*-\tau^*)\, d\tau^* \quad \ldots\ldots\ldots (3.12)$$

where $k^*(x^*,y^*,z^*,t^*)$ is equal to the concentration distribution resulting from an instantaneous release, which is the distribution given by Eq.3.10. As in the case of the instantaneous release, Eq.3.11 is used to determine the theoretical grid averaged concentration. The dimensions of the source, concentration grid dimensions, and origin, and the hydrogeological characteristics of the porous formation are the same as those for the instantaneous release. The model results are contrasted with theory by comparing the centerline concentration distributions at $t^* = 1$, 10, and 25. Discretizing the source into 27,000 particles ($n_x = n_y = n_z = 30$), using 1 ensemble ($n_e = 1$) and temporally interpolating 10 points between each 'puff' yielded results which were not sensitive to changes in these simulation parameters. The results are shown in Fig.3.5. The agreement with theory is excellent. An average error magnitude, $\bar{e}$, for each time may be computed using the formula

$$\bar{e} = \frac{\sum_{i=1}^{n_c} c^*_{ijkl}\, e_i}{\sum_{i=1}^{n_c} c^*_{ijkl}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.13)$$

where c^*_{ijkl} is the theoretical concentration in grid ijk, e_i is the magnitude of the error between the simulated and theoretical

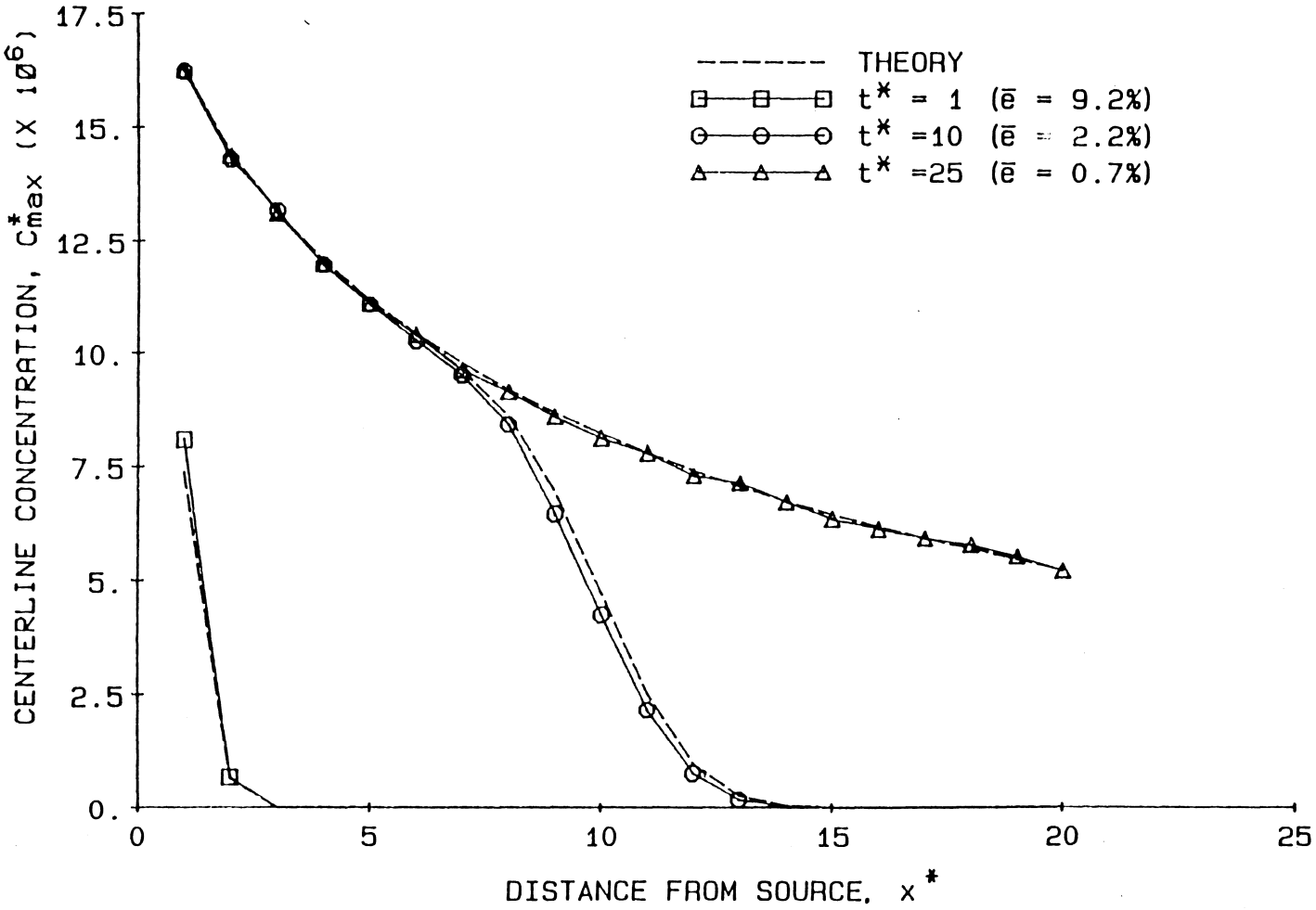


Figure 3.5: Centerline Concentration vs. Distance from Source for a Continuous Release

concentration in the centerline grid element, and n_c is the number of centerline points in the simulation. For the parameters used in this study, the errors $\bar{e}$ are shown in parentheses in Fig.3.5. The results indicate that as time increases the average error decreases and becomes quite small (0.7%) by the time $t^* = 25$. This phenomenon is a result of the fact that for small times the extent of the plume is small and the accurate definition of the source shape is very important. As time increases, downstream contamination increases and these values are not as sensitive to the source shape. The importance of interpolation points may be measured by the quantity v^* which is numerically equal to the number of grids traversed by a particle in one time step moving with the average seepage velocity. In the case of a continuous release, v^* also measures the number of grid spaces between 'puffs', and therefore indicates the need to interpolate between time steps in order to obtain a smooth solution. In the present case, $v^* = 1$, and hence some interpolation is indicated. The interpolation procedure requires linearly interpolating n_{int} points between the particle position at times t and $t + \Delta t$. Figure 3.5 shows the simulation results for $n_{int} = 5$.

3.2 Evaluation of Model Performance

A numerical approach for simulating the dispersion of contaminants in porous media has been developed. The proposed model has several features not found in classical models. The model features are: (a) uses measured head data directly; (b) uses different grids to define the porous formation and the concentration; (c) three-dimensional, unsteady model; and (d) rigorously partitions the advection and

diffusion velocity components. The model has been verified for instantaneous and continuous releases in cases where analytic solutions are available. In both cases, the agreement between the theory and numerical simulations have been excellent. The primary limitations of the model are: (a) the spatial and temporal resolution of the model results are limited by the spatial and temporal resolution of the measured head, and (b) the procedure for combining the dispersivity due to temporal variations with those due to spatial variations in the hydraulic conductivity is somewhat speculative. It is important to note, however, that the dispersivity combination procedure used in the model is asymptotically correct when dispersion is dominated by temporal fluctuations in the seepage velocity and also correct when spatial fluctuations in the hydraulic conductivity dominate.

CHAPTER IV

ESTIMATION OF MISSING DATA

4.1 Introduction

The problem of estimating hydrologic time series data at a location where only a limited record is available is an important problem in several areas of water resources engineering. In many cases, because of the non-stationary character of the time series, it is desirable to relate fluctuations at one location to those at an adjacent station where a more extensive record is available. For example, the need to know temporal variations in water table elevations at a particular location is frequently required to adequately plan several types of water resources projects. The design of water supply wells, and groundwater monitoring programs all require such information. In the present context, the Chin (1988b) model requires synoptic head measurements at several locations. Due to instrument malfunction, gaps in the data frequently exist. Such gaps severely limit the application of the Chin model. If water table records at a particular location contain gaps in the measured data and there exists a complete record several kilometers away, then it would be useful to use the synoptic portion of the record at each station to determine a relationship between the two time series and then, using this relationship, estimate the missing data at the location of interest. The problem posed falls into the category of estimating the relationship between two stochastic, non-stationary time series. Brillinger (1981) has discussed a technique that may be used to determine the relation between two stochastic stationary time series,

however, the development of data fill-in techniques for non-stationary hydrologic time series has not recieved significant attention in the literature. Bras and Rodriguez-Iturbe (1985) and Brillinger (1981) have discussed the problem of relating time series in the frequency domain, however no attention has been paid to the adequacy of such techniques for data fill-in. A weakness of classical procedures, such as those described by Bras and Rodriguez-Iturbe (1985), is that raw data are used at each station to determine the gain and phase-shift at each frequency. This procedure is not adequate for time series that have non-coherent random components which superimpose a uniform power spectrum on the coherent frequencies. In order to alleviate this deficiency, it is necessary to identify and remove the random non-coherent fluctuations from both time series before determining the gain and phase-shift between the coherent fluctuations at both stations. In this paper, a technique is presented which may be used to determine the relationship between two non-stationary time series. The procedure is verified using synthetic time series data and is validated by applying it to field water table measurements at two stations several kilometers apart.

4.2 Theory

Consider two discrete non-stationary stochastic time series X(t) and Y(t). Each may be expressed as the sum of non-random and random components, hence

$$X(t) = x(t) + \varepsilon_X(t) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.1)$$

$$Y(t) = y(t) + \varepsilon_Y(t) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.2)$$

where $x(t)$ and $y(t)$ are non-random time series and $\varepsilon_X(t)$ and $\varepsilon_Y(t)$ are random fluctuations. Suppose each series is divided into m sub-records each of length n. The discrete Fourier transform (DFT) of any sub-record of $X(t)$ is then given by the relation

$$J_X(\omega_j) = \frac{1}{n} \sum_{t=0}^{n-1} X(t) \exp(-i\omega_j t) \quad \ldots\ldots\ldots\ldots\ldots\ldots (4.3)$$

where ω_j is the Fourier frequency given by

$$\omega_j = \frac{2\pi j}{n} \, , \; j=1,(n-1)/2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.4)$$

and it has been tacitly assumed that n is odd in order to simplify the transform expressions. The DFT of $Y(t)$ may be defined similarly. Taking the Fourier transforms of Eqs.4.1 and 4.2 yield the relations

$$J_X(\omega_j) = J_x(\omega_j) + J_{\varepsilon_X}(\omega_j) \quad \ldots\ldots\ldots\ldots\ldots\ldots (4.5)$$

$$J_Y(\omega_j) = J_y(\omega_j) + J_{\varepsilon_Y}(\omega_j) \quad \ldots\ldots\ldots\ldots\ldots\ldots (4.6)$$

The random components have variances, $\sigma^2_{\varepsilon_X}$ and $\sigma^2_{\varepsilon_Y}$, and covariance, $\sigma_{\varepsilon_X \varepsilon_Y}$, given by

$$\sigma^2_{\varepsilon_X} = \sum_{j=1}^{(n-1)/2} |J_{\varepsilon_X}(\omega_j)|^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.7)$$

$$\sigma^2_{\varepsilon_Y} = \sum_{j=1}^{(n-1)/2} |J_{\varepsilon_Y}(\omega_j)|^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.8)$$

$$\sigma^2_{\varepsilon_X\varepsilon_Y} = \sum_{j=1}^{(n-1)/2} J_{\varepsilon_X}(\omega_j)\, J^*_{\varepsilon_Y}(\omega_j) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.9)$$

The quantities $J_{\varepsilon_X}(\omega_j)$ and $J_{\varepsilon_Y}(\omega_j)$ may be easily estimated if it is assumed that the random fluctuations are normally distributed. In such cases, the power spectra of the random components have distributions given by $P_{\varepsilon_X}\, \chi^2_{2,\alpha}/2$ and $P_{\varepsilon_Y}\, \chi^2_{2,\alpha}/2$ where $\chi^2_{2,\alpha}$ is the chi-square variate with 2 degrees of freedom and confidence limit α, and P_{ε_X} and P_{ε_Y} are the average power in the random component at all frequencies (Brillinger, 1981). Hence

$$P_{\varepsilon_X} = \frac{2}{n-1} \sum_{j=1}^{(n-1)/2} |J_{\varepsilon_X}(\omega_j)|^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.10)$$

$$P_{\varepsilon_Y} = \frac{2}{n-1} \sum_{j=1}^{(n-1)/2} |J_{\varepsilon_Y}(\omega_j)|^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.11)$$

Consider the power spectrum of X(t), then Eq.4.5 indicates that

$$|J_X(\omega_j)|^2 = |J_x(\omega_j)|^2 + 2\,\mathrm{Re}\left[J_x(\omega_j)\, J^*_{\varepsilon_X}(\omega_j)\right] + |J_{\varepsilon_X}(\omega_j)|^2 \quad (4.12)$$

We now assume that the non-random component of X(t), whose transform is $J_x(\omega_j)$, predominantly consists of the lower frequency components. This

condition is typical of hydrologic time series (Salas, 1980). In this case Eq.4.6 indicates that $|J_{\varepsilon_X}(\omega_j)|^2$ is the only non-zero component of the power spectrum at the high frequency end. Since the probability distribution of the spectrum of the random component is given by P_{ε_X} $\chi^2_{2,\alpha}/2$ (for an unsmoothed spectrum) then we may find the maximum value of r for which neither of the following inequalities are violated more than $0.05[(n-1)/2 - r + 1]$ times:

$$P_{\varepsilon_X}^{(r)} \frac{\chi^2_{2,0.95}}{2} > |J_{\varepsilon_X}(\omega_j)|^2, \quad j = r,(n-1)/2 \quad \ldots\ldots\ldots (4.13)$$

$$P_{\varepsilon_X}^{(r)} \frac{\chi^2_{2,0.05}}{2} < |J_{\varepsilon_X}(\omega_j)|^2, \quad j = r,(n-1)/2 \quad \ldots\ldots\ldots (4.14)$$

where

$$P_{\varepsilon_X}^{(r)} = \frac{1}{N_c} \sum_{j=r}^{(n-1)/2} |J_{\varepsilon_X}(\omega_j)|^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots (4.15)$$

and

$$N_c = \frac{n-1}{2} - r + 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.16)$$

where N_c equals the number of frequency components that satisfy the power spectrum distribution associated with a random signal. For large values of N_c we may infer that

$$P_{\varepsilon_X}^{(r)} \approx P_{\varepsilon_X} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.17)$$

If in practice N_c is found to be large, then the assumption of only low frequency non-random components is verified. Based on the above algorithm, the Fourier transform of the random component is given by

$$J_{\varepsilon_X}(\omega_j) = \begin{cases} J_X(\omega_j), & j = r,(n-1)/2 \\ P^{(r)}_{\varepsilon_X} \dfrac{J_X(\omega_j)}{|J_X(\omega_j)|}, & j = 1,r-1 \end{cases} \quad \ldots\ldots\ldots (4.18)$$

The procedure described here may be repeated for Y(t) and the random component Fourier transforms combined using Eqs.4.7 to 4.9 to determine the random component statistics. The random component statistics are determined in this manner for each sub-record. Then the average random component statistics for the time series pair are given by

$$\bar{\sigma}^2_{\varepsilon_X} = \langle \sigma^2_{\varepsilon_X} \rangle \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.19a)$$

$$\bar{\sigma}^2_{\varepsilon_Y} = \langle \sigma^2_{\varepsilon_Y} \rangle \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.19b)$$

$$\bar{\sigma}_{\varepsilon_X \varepsilon_Y} = \langle \sigma_{\varepsilon_X \varepsilon_Y} \rangle \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.19c)$$

where the notation < > refers to the average over the m sub-records.

The random component Fourier transforms may now be subtracted from the transforms of the raw data, X(t) and Y(t), to obtain the DFT of the non-random components, $J_x(\omega_j)$ and $J_y(\omega_j)$. If we define a complex quantity $A(\omega_j)$ whose magnitude equals the gain and whose argument equals the phase shift between x(t) and y(t), then

$$A(\omega_j) = \frac{J_x(\omega_j)\ J_y^*(\omega_j)}{|J_y(\omega_j)|^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.20)$$

Based on the m sub-records, we may calculate the average gain, $G(\omega_j)$ and phase shift, $\varphi(\omega_j)$, by the relations

$$G(\omega_j) = |\langle A(\omega_j)\rangle| \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.21)$$

$$\varphi(\omega_j) = \arg\langle A(\omega_j)\rangle \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.22)$$

It is important to note that relating the DFT's of x(t) and y(t) do not require that both time series be stationary, but only that their non-stationarity be significantly coherent. This type of behaviour is commonly observed in hydrologic time series such as the occurence of wet and dry years. We are justified in estimating the Fourier transform of y(t) from x(t) by a gain and phase shift only if the signals are significantly coherent at that frequency. The coherence, $\gamma(\omega_j)$, between x(t) and y(t) at frequency ω_j is defined by

$$\gamma(\omega_j) = \frac{\langle J_x(\omega_j)\ J_y^*(\omega_j)\rangle}{\langle |J_x(\omega_j)|^2\rangle^{1/2}\ \langle |J_y(\omega_j)|^2\rangle^{1/2}} \quad \ldots\ldots\ldots\ldots (4.23)$$

where the distribution of zero significance, γ_o, is given by (Koopmans, 1974, Jenkins and Watts, 1968)

$$\gamma_o = \left[1 + \frac{(m-1)}{F_{2,2(m-1)}} \right]^{1/2} \quad \ldots\ldots\ldots\ldots\ldots\ldots (4.24)$$

where $F_{a,b}$ is the F distribution and a and b are degrees of freedom. For any given confidence limit, if $\gamma(\omega_j) > \gamma_o$ then we may estimate the Fourier transform of y(t) by the relation

$$F_y(\omega_j) = \langle A(\omega_j) \rangle F_x(\omega_j) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (4.25)$$

while if $\gamma(\omega_j) < \gamma_o$ we must resort to the relation

$$F_y(\omega_j) = \langle F_y(\omega_j) \rangle \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4.26)$$

The relationship given in Eq.4.26 represents the best estimate of $F_y(\omega_j)$ for frequency components of y(t) that are not significantly coherent with those of x(t). Equation 4.26 is only approximate for non-stationary time series; however most local hydrologic time series will be significantly coherent at the dominant frequencies and spatial correlations determined from Eq.4.25 are likely to be adequate to estimate most of the non-random variance in y(t).

4.3 Verification of Procedure

The verification procedure consisted of generating several synthetic time series pairs. Then, using the above method, calculate their random components, the relationship between their non-random components, and predict a portion of the time series record that was not used in calculating the inter-relationship. The adequacy of the estimation procedure is then judged by the correlation between the simulated and known time series. Two types of time series were used in the verification process: non-stationary coherent and stationary coherent.

4.3.1 Non-Stationary Coherent Time Series

Two synthetic time series X(t) and Y(t) were generated having constituents shown in Table 4.1. The factors $\alpha(k)$ (k=1,m) modify the component amplitude for each sub- record, k, of length 365 days. A total of 12 sub-records were generated, the first 11 sub-records being used to estimate the relation between the X(t) and Y(t) series and the last sub-record used to compare the predicted and known time series. The amplitude factors $\alpha(k)$ for the 12 sub-records used in the simulations are shown in Table 4.2. The records X(t) and Y(t) are plotted in Fig.4.1. The non-stationary nature of X(t) and Y(t) is clearly evident. The relationship between X(t) and Y(t) calculated by the proposed procedure is shown in Table 4.3. The results of this simulation show that the gain and phase shift are quite accurately predicted for those components whose amplitude significantly exceeds the standard deviation of the random component. The coherence of the components with period 91 and 73 days were less than the 95% confidence limit so these components could not be estimated from X(t). In addition, these components were so small that they could not be distinguished from the random components in the time series sub-records of Y(t), hence their average value could not be calculated. A close estimation of the random component variances, Table 4.3, indicates the efficiency of the random component estimation technique. The number of frequency components used to estimate the random component statistics are shown in Table 4.4. These results show that, in 11 sub-records, on the average over 100 of the 182 frequency components had a distribution consistent with the power spectrum of a random normal deviate. This scenario indicates that significant confidence may be placed in the estimates of the random

Table 4.1.- Constituents of Non-Stationary Coherent Series

(a) Constituents of X(t)

Component	Period (days)	Amplitude (m)	Phase (radians)
1	365	1.52*α(k)	0.0
2	183	0.69*α(k)	0.0
3	122	0.52*α(k)	0.0
4	91	0.41*α(k)	0.0
5	73	0.37*α(k)	0.0

Mean, $\overline{X}$ = 1.52 m, σ_{ε_X} = 9.1 cm

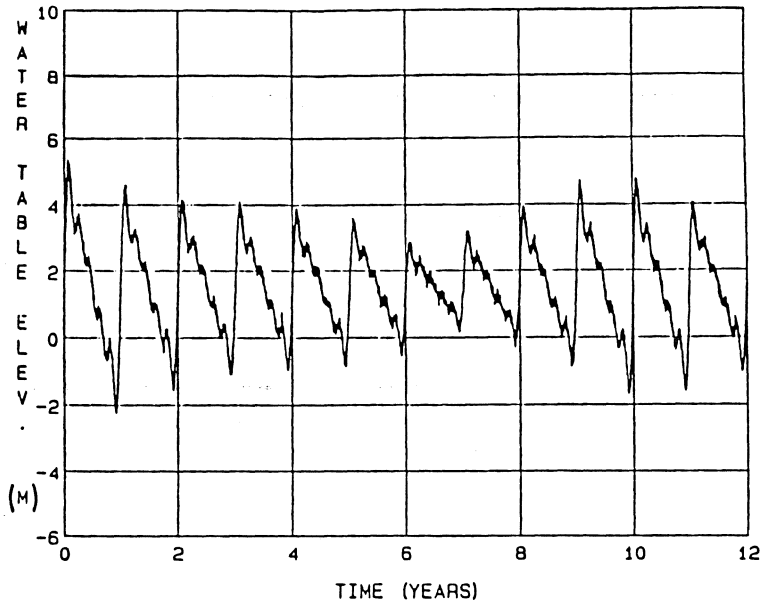
(b) Constituents of Y(t)

Component	Period (days)	Amplitude (m)	Phase (radians)
1	365	2.74*α(k)	0.300
2	183	1.17*α(k)	0.600
3	122	0.73*α(k)	-0.500
4	91	0.62*α(k)	0.700
5	73	0.48*α(k)	-0.300

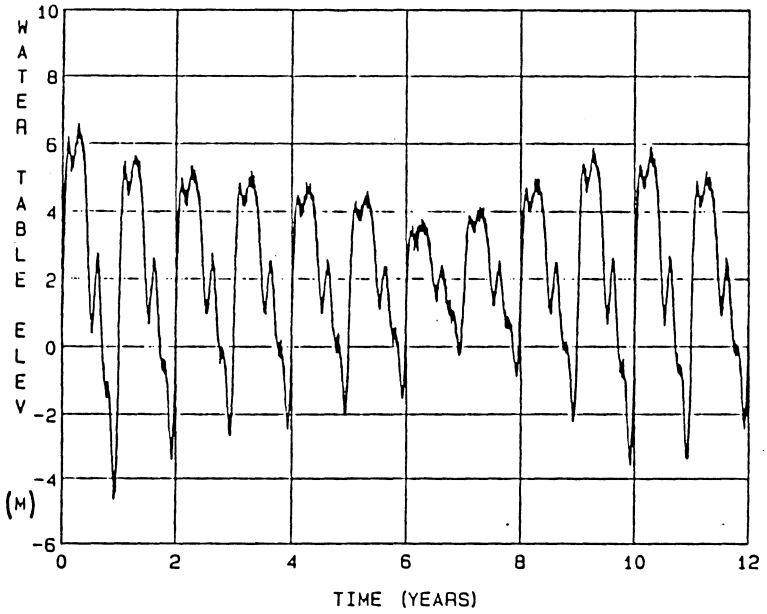
Mean, $\overline{X}$ = 2.13 m, σ_{ε_X} = 12.2 cm

Table 4.2. Amplitude Factors For Non-Stationary Coherent Series

Sub-Record, k	Factor, $\alpha(k)$
1	1.00
2	1.20
3	1.05
4	1.00
5	0.90
6	0.80
7	0.50
8	0.65
9	0.95
10	1.25
11	1.25
12	1.50



(a) Series X(t)



(b) Series Y(t)

Figure 4.1: Non-Stationary Time Series Pair

Table 4.3.- Comparison Between Estimated and Predicted Time Series Relations

(a) Non-Random Components

Period	Actual		Computed	
	Gain	Phase Shift (radians)	Gain	Phase Shift (radians)
365	1.800	0.300	1.799	0.298
183	1.699	0.600	1.711	0.611
122	1.395	-0.500	1.399	-0.494
91	1.500	0.700	-	-
73	1.295	-0.300	-	-

(b) Random Components

Actual			Computed		
$\sigma^2_{\varepsilon_X}$	$\sigma^2_{\varepsilon_Y}$	$\sigma_{\varepsilon_X\varepsilon_Y}$	$\sigma^2_{\varepsilon_X}$	$\sigma^2_{\varepsilon_Y}$	$\sigma_{\varepsilon_X\varepsilon_Y}$
(cm^2)	(cm^2)	(cm^2)	(cm^2)	(cm^2)	(cm^2)
83.6	148.6	0.0	88.4	140.7	1.4

Table 4.4.- Number of Frequencies with Random Characteristics

X(t)			Y(t)		
Maximum	Minimum	Average	Maximum	Minimum	Average
179	9	111	179	29	105

component statistics. Using the estimated time series relation shown in Table 4.3, the sub-record of Y(t) not used previously was estimated. Comparing the estimated and known fluctuations in the sub-record, a correlation coefficient of 0.95 was obtained. This high value indicates that the important features of Y(t) have been reproduced accurately.

4.3.2 Stationary Coherent Time Series

Two synthetic time series were generated having the constituents shown in Table 4.5, and are plotted in Fig.4.2. The normalized power spectra of the time series are shown in Fig.4.3. The relationship between X(t) and Y(t) calculated using the proposed procedure are compared with known quantities in Table 4.6. Eleven records each of length 365 days were used. For the non-random components the gain and phase shift are quite accurately extracted, except for components with periods of 91 and 73 days, which were found to be non-coherent with 95% confidence. The amplitude of these fluctuations were on the same order as the random fluctuations and so the procedure was unable to distinguish these components. The variances of the random components at each station were closely estimated. However the covariance of the random components is significantly overestimated. This occured because of a high covariance estimated in one of the 11 sub-records, which are all statistically the same. If this extreme value is omitted, we obtain an average covariance of 1.6 cm which is much closer to the actual value. Based on these results, we would expect the random component variances to be quite accurately estimated, while we may attach some skepticism to the estimate of covariance, although in the absence of "outliers" we would expect good results. The number of frequency components used in estimating these statistics are shown in Table 4.7.

Table 4.5.- Constituents of Coherent Series

(a) Constituents of X(t)

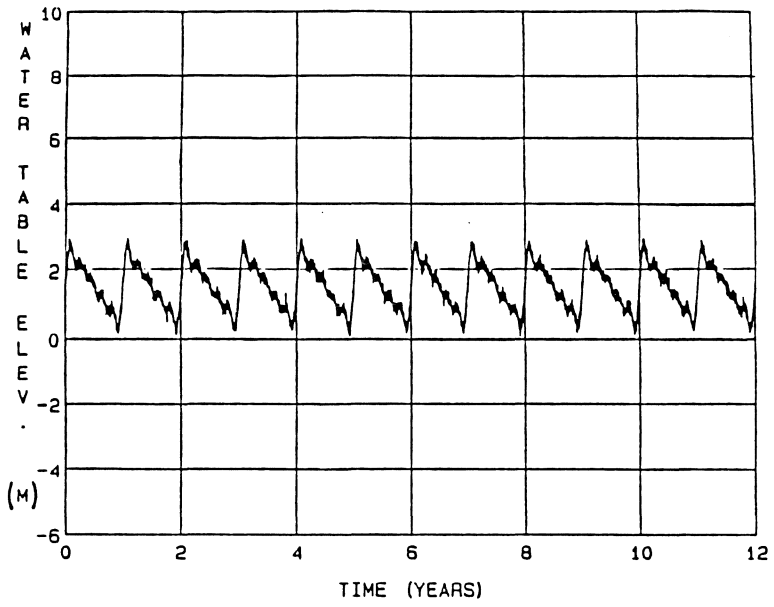
Component	Period (days)	Amplitude (m)	Phase (radians)
1	365	0.76	0.0
2	183	0.34	0.0
3	122	0.26	0.0
4	91	0.21	0.0
5	73	0.19	0.0

Mean, $\bar{X}$ = 1.52 m, σ_{ε_X} = 9.1 cm

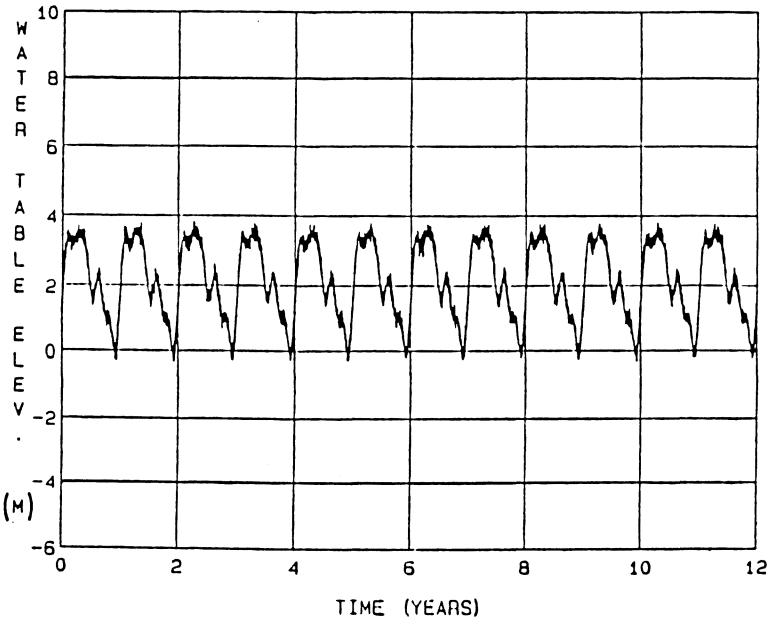
(b) Constituents of Y(t)

Component	Period (days)	Amplitude (m)	Phase (radians)
1	365	1.37	0.300
2	183	0.59	0.600
3	122	0.37	-0.500
4	91	0.31	0.700
5	73	0.24	-0.300

Mean, $\bar{X}$ = 2.13 m, σ_{ε_X} = 12.2 cm

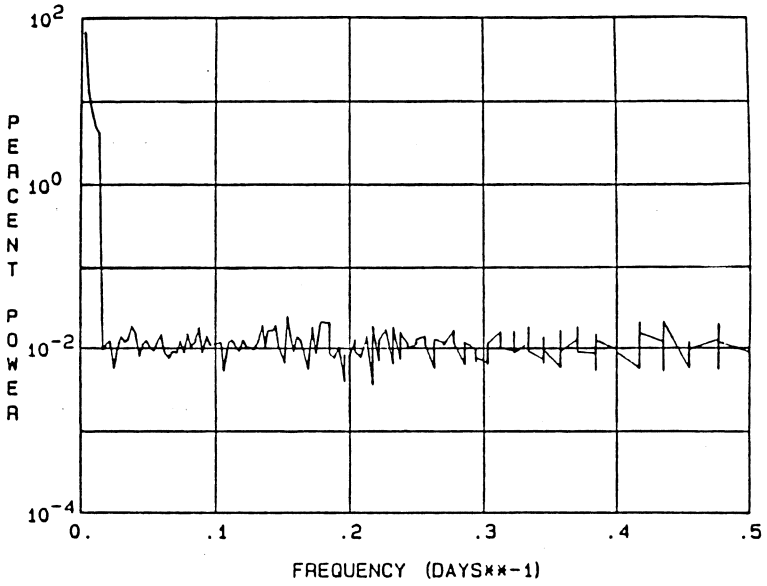


(a) Series X(t)

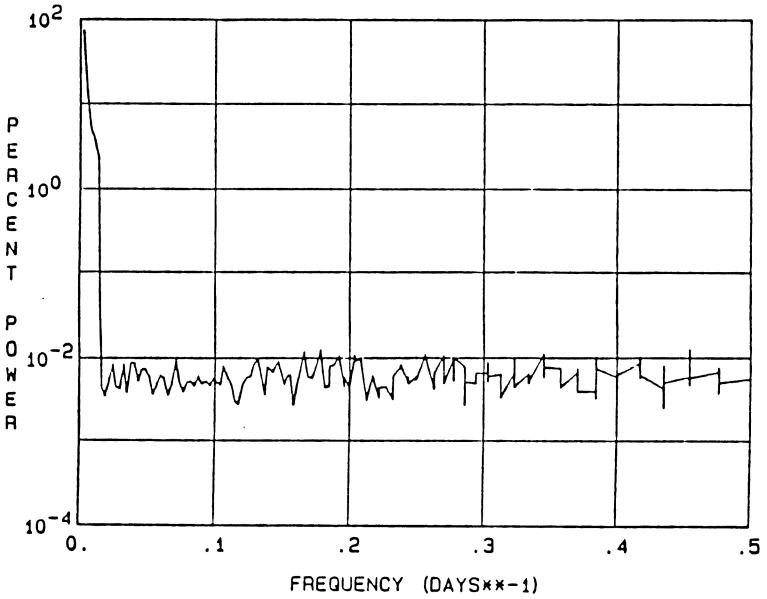


(b) Series Y(t)

Figure 4.2: Stationary Time Series Pair



(a) Series X(t)



(b) Series Y(t)

Figure 4.3: Power Spectra of Stationary Time Series

Table 4.6.- Comparison Between Estimated and Predicted Time Series Relations

(a) Non-Random Components

Period	Actual		Computed	
	Gain	Phase Shift (radians)	Gain	Phase Shift (radians)
365	1.800	0.300	1.798	0.297
183	1.699	0.600	1.717	0.617
122	1.395	-0.500	1.397	-0.491
91	1.500	0.700	-	-
73	1.295	-0.300	-	-

(b) Random Components

Actual			Computed		
$\sigma^2_{\varepsilon_X}$	$\sigma^2_{\varepsilon_Y}$	$\sigma_{\varepsilon_X\varepsilon_Y}$	$\sigma^2_{\varepsilon_X}$	$\sigma^2_{\varepsilon_Y}$	$\sigma_{\varepsilon_X\varepsilon_Y}$
(cm^2)	(cm^2)	(cm^2)	(cm^2)	(cm^2)	(cm^2)
83.6	148.6	0.0	90.9	129.7	44.8

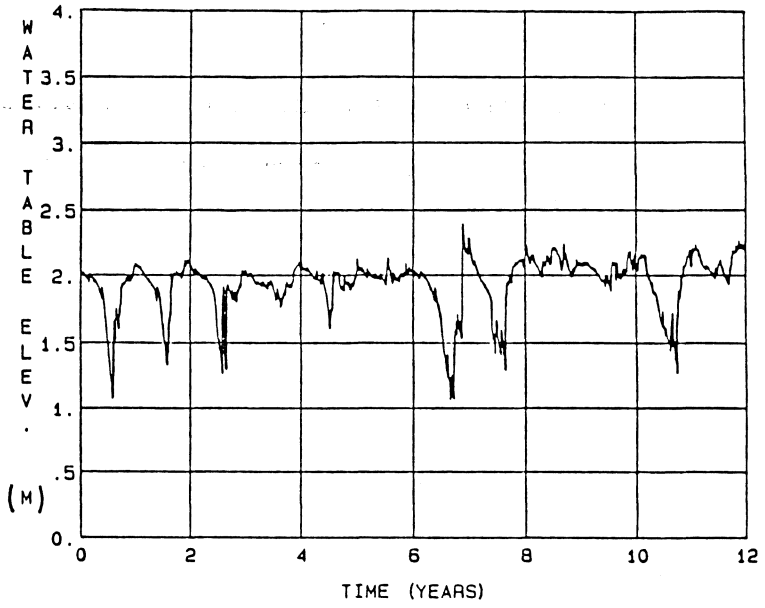
Table 4.7.- Number of Frequencies with Random Characteristics

X(t)			Y(t)		
Maximum	Minimum	Average	Maximum	Minimum	Average
179	9	111	179	4	113

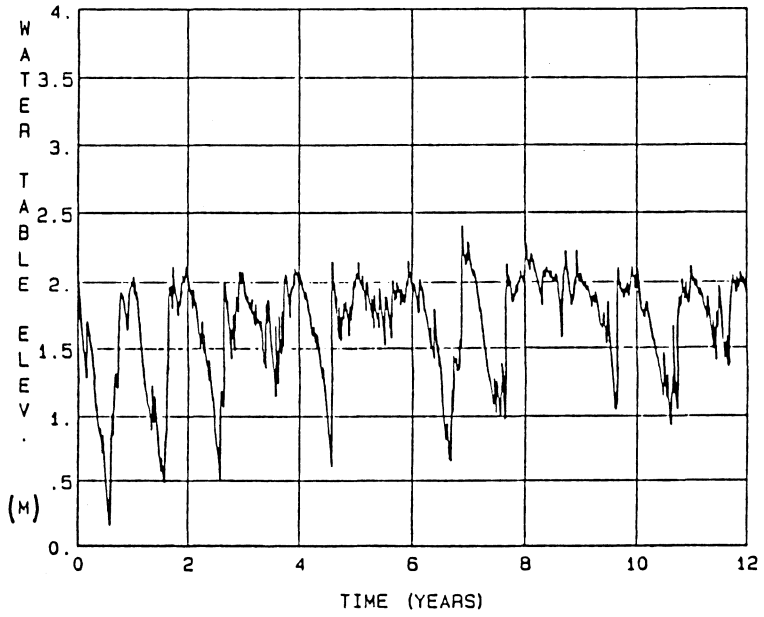
These results show that on average over 110 of the 182 components were found to have a power spectrum characteristic of a random normal deviate. This provides a sufficient number of spectral components to identify a significant portion of the random component spectrum. It may be useful to reiterate at this point that the success of the proposed technique is based on the condition that the non-random fluctuations occur at significantly fewer frequencies than are in the total spectrum and so, by the superposition property of Fourier transform, a significant portion of the spectrum is due solely to the random component spectrum, which theoretically has a uniform magnitude at all frequencies. Our ability to obtain sufficient points on which to calculate a good estimate of the magnitude of the random component spectrum is required for the technique to be successful. The station relations shown in Table 4.6 were used to predict a portion of the time series not used in calculating the station relations, the correlation coefficient was found to be 0.94. This result underscores that the proposed method is quite accurate in determining the relation between the two stationary stochastic time series X(t) and Y(t).

4.4 Validation of the Procedure

In this section the ability of the proposed technique to adequately estimate the relation between two measured time series of water table elevations is evaluated. The time series pair to be used in this validation procedure is shown in Fig.4.4. These measurements were taken in the unconfined Biscayne aquifer in Dade county, Florida, at stations that are 18.1 km apart. The non-stationary nature of these time series, along with the distinct annual cycles of wet and dry



(a) Station X



(b) Station Y

Figure 4.4: Measured Water Table Observations

seasons are the most apparent features of these time series. The procedure used was to divide the time series into 12 sub-records, each of length 365 days, then 11 of the sub-records were used to estimate the relation between the time series and this relation was used to estimate the time series at station Y from station X for the unused sub-record. The correlation coefficient between the estimated and known record then measures the validity of the proposed procedure. The correlation coefficients for each of the 12 sub-records are shown in Table 4.8. These results indicate that the method used produces good to excellent estimates of the time series for all sub-records except sub-records 6, 9, and 10. We may conclude that in 9 of 12 years the method is valid for estimating Y(t) from X(t). The validity of the random component estimation technique is demonstrated by the results shown in Table 4.9. These results indicate that in all cases a large number of frequency components, more than 44% of the total, were found to have a power spectrum characteristic of a random normal deviate and hence supports the assumption that a random component is contained within the time series. The stability in the estimated values of the variances and covariance indicate that the random component is fairly stationary. An interesting point to note is that in each sub-record estimated there was no significant difference in the correlation coefficient if only coherent frequency components were used to estimate the time series at station Y. This result demonstrates that the dominant frequency components are coherent.

4.5 Practical Applications

The results of this research are expected to be applicable

Table 4.8.- Correlation Coefficients Between Simulated and Measured Sub-Records

Sub-Record	Correlation Coefficient
1	0.90
2	0.93
3	0.83
4	0.74
5	0.75
6	0.33
7	0.89
8	0.93
9	0.41
10	0.15
11	0.90
12	0.82

Table 4.9.- Random Component Statistics at Station Y

Sub-Record	$\langle\sigma^2_{\varepsilon_X}\rangle$ (cm^2)	$\langle\sigma^2_{\varepsilon_Y}\rangle$ (cm^2)	$\langle\sigma_{\varepsilon_X\varepsilon_Y}\rangle$ (cm^2)	$\langle N_X\rangle$ (cm^2)	$\langle N_Y\rangle$ (cm^2)
1	1.3	6.7	0.3	82	90
2	1.3	7.2	0.4	74	101
3	1.1	6.5	0.2	90	95
4	1.6	6.7	0.4	90	94
5	1.5	4.2	0.2	84	93
6	1.8	7.0	0.3	84	100
7	1.4	7.1	0.4	87	92
8	1.5	6.9	0.3	84	85
9	1.5	7.3	0.3	83	95
10	1.5	6.7	0.3	96	88
11	1.5	7.0	0.3	88	93
12	1.5	6.8	0.3	82	94

Note: N_X and N_Y are the number of components in the power spectra of X(t) and Y(t) respectively that have random component characteristics.

primarily as a technique to fill-in missing data in hydrologic time series. The dominance of a few coherent low frequency components make hydrologic time series ideal for the procedure described in this research. Other applications may include studying the random component characteristics and their relation to the spreading of contaminants in the groundwater.

4.6 Summary and Conclusions

A method has been presented that may be used to estimate the relationship between two stochastic non-stationary time series. The procedure is particularly suited to hydrologic time series, which tend to be significantly coherent at a few low frequency components and include non-coherent random fluctuations. The procedure compares two scalar time series and determines (a) The random component statistics of both series; (b) The gain and phase shift between coherent frequencies; and (c) The mean Fourier coefficients for non-coherent, non-random frequency components. The procedure was verified by using synoptic portions of two synthetic time series to estimate their relationship. The estimated relationship was then compared with the known relation. Furthermore, the calculated relations were used to estimate a known portion of the time series and a correlation coefficient was used to measure the overall performance of the procedure. The method was applied to two sets of synthetic time series: non-stationary coherent and stationary coherent. In both cases the procedure performed well, with correlation coefficients between the known and estimated signals being above 0.94. It is notable, that the procedure was able to accurately estimate the random component

statistics in both cases. The validity of the procedure was demonstrated by applying it to two measured water table time series at locations 18.1 km apart. The procedure was able to accurately predict a measured time series for 9 out of the 12 years and predict fairly stationary estimates of the statistics of the random fluctuations. Practical applications of this research include fill-in of missing data and extracting random fluctuations in hydrologic time series.

CHAPTER V

VALIDATION STUDY

5.1 Introduction

The problem of predicting the movement of contaminants in aquifers where no previous contamination exists is a case frequently encountered in engineering practice. Traditional solutions assume that contaminant mixing is governed by the advection-dispersion equation, and the problem transformed to one requiring the estimation of a dispersivity. Although considerable attention has been given to developing theoretical methods of estimating dispersivity (Gelhar and Axness, 1983, Dagan, 1984, Neuman et al., 1987), it is frequently overlooked that the spatial and temporal variations in seepage velocity must be adequately accounted for in any valid dispersion model. Using the advection-dispersion equation, high frequency seepage velocity variations are typically parameterized in a dispersion coefficient while the low frequency variations are accounted for explicitly in the advection term. Because of uncertainties associated with estimating the dispersivity, it is generally desirable to describe, as much as possible, the spatial and temporal variations in seepage velocity explicitly in the advection terms. Recent field studies (Guven et al., 1984) have demonstrated that such an approach tends to produce results that are not very sensitive to the estimated dispersivity. A practical model of dispersion may be defined as one in which the parameters of the model are determined from measurable hydrogeological parameters, and therefore the model is non-empirical. Models that require calibration are impractical since the results of these studies can not

typically be extrapolated into the future with any certainty. There has not to date been any reported studies that compare the predictions of practical dispersion models to observed field measurements in order to assess the predictive capability of these models. This paper describes the results of a validation study which compares the measured dispersion of a tracer with the predictions of the Chin (1987) model, which are based only on measurable hydrogeological parameters. The details of the Chin (1987) model are described in Chapter 2. This study provides an assessment of the accuracy of the Chin approach at a site where extensive field measurements are available.

5.2 Site Description

The site used in this study has been documented extensively by Waller (1982). The source consists of a highly saline discharge from a freely flowing artesian well, Grossman well, located in Chekika Hammock State Park in Dade county, Florida. The well discharges directly into a clay lined lake which overflows into a shallow seepage basin within the Everglades marsh, Fig.5.1. The saline water then percolates directly into the highly permeable Biscayne aquifer where the water table is typically only a few feet below the land surface. The well discharge rate has varied from approximately 2000 gpm when the well was first constructed in 1944 to 1000 gpm when the well was capped in 1985. The chloride levels in the discharge have risen from approximately 900 mg/l in 1944 to 1250 mg/l in 1985. These levels significantly exceed the background chloride levels in the Biscayne aquifer of about 30 mg/l. Based on the measured data, an estimate of the discharge and chloride variations are shown in Fig.5.2. The falling discharge rate and the

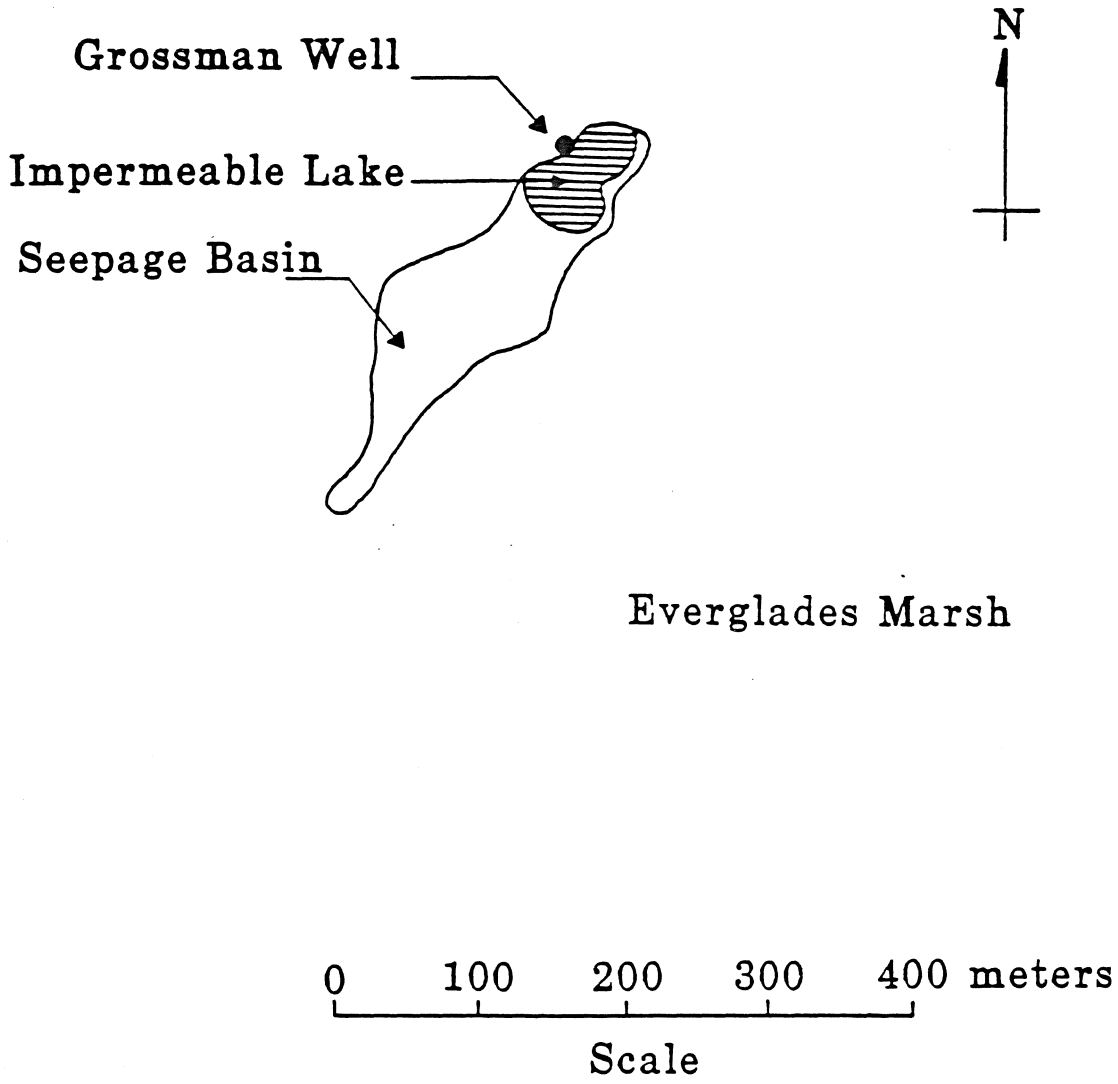


Figure 5.1: Source Configuration
(After Waller, 1982)

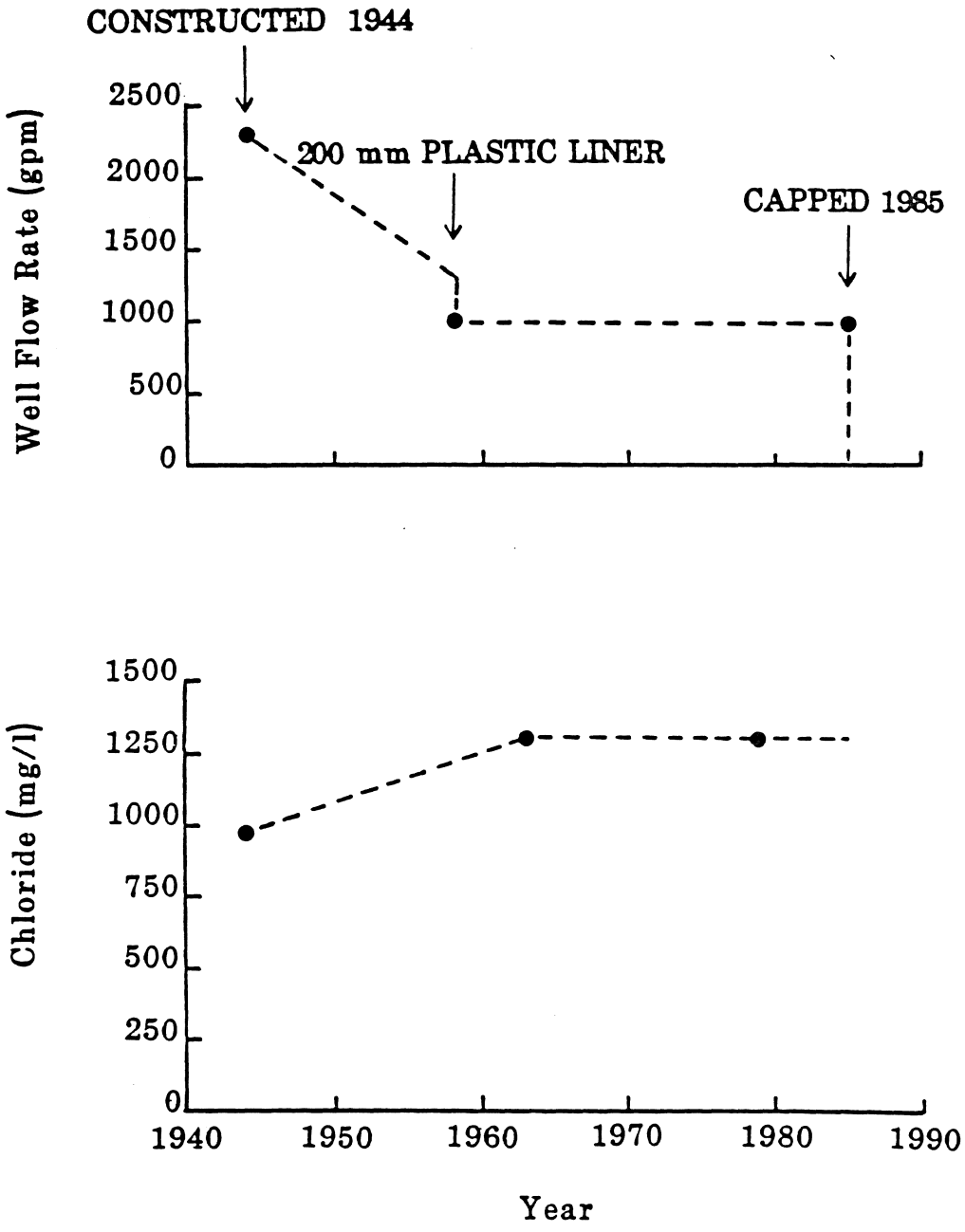


Figure 5.2: Well Discharge Characteristics

rising chloride levels yield a chloride mass flux that does not vary as much as either of the measured variables. A mass flux of 2.5 x 10^6 kg/year shows close agreement with the measured fluxes. The layout of the area surrounding Chekika is shown in Fig.5.3 and a geologic cross section XX' of the Biscayne aquifer is shown in Fig.5.4. The Biscayne aquifer is composed of regions with significantly different geological characteristics. The Miami Oolite overlays the Fort Thomson Formation which in turn overlays the Tamiami Formation. From the viewpoint of contaminant transport, the important geologic parameters are the hydraulic conductivity and the porosity. Klein and Hull (1978) have documented measured transmissivities in the Biscayne aquifer. In the middle of the study area the best estimate of the transmissivity is 3.4 x 10^5 m/day (1.1 x 10^6 ft/day). Since the Tamiami Formation is essentially an aquiclude, and the hydraulic conductivity of the Fort Thomson Formation is approximately 8 times higher than the Miami Oolite (Fish, 1987) we may deduce the hydraulic conductivities of the Miami Oolite and Fort Thomson Formation to be 1050 m/day (3440 ft/day) and 8380 m/day (27 500 ft/day) respectively. Parker et al. (1955) have estimated the specific yield of the aquifer to be typically 0.2 which, for limestone aquifers, typically corresponds to a porosity of 0.3 (Todd, 1980). In simulating transport in porous media we generally utilize the effective porosity which is adequately approximated by the specific yield. Measurements of chloride levels have been collected at several USGS operated monitoring wells. The locations of the monitoring wells are shown in Fig.5.3. The measured chloride concentrations clearly indicate that the chloride plume has caused significant elevation in chloride levels above the 30 mg/l background. A vivid

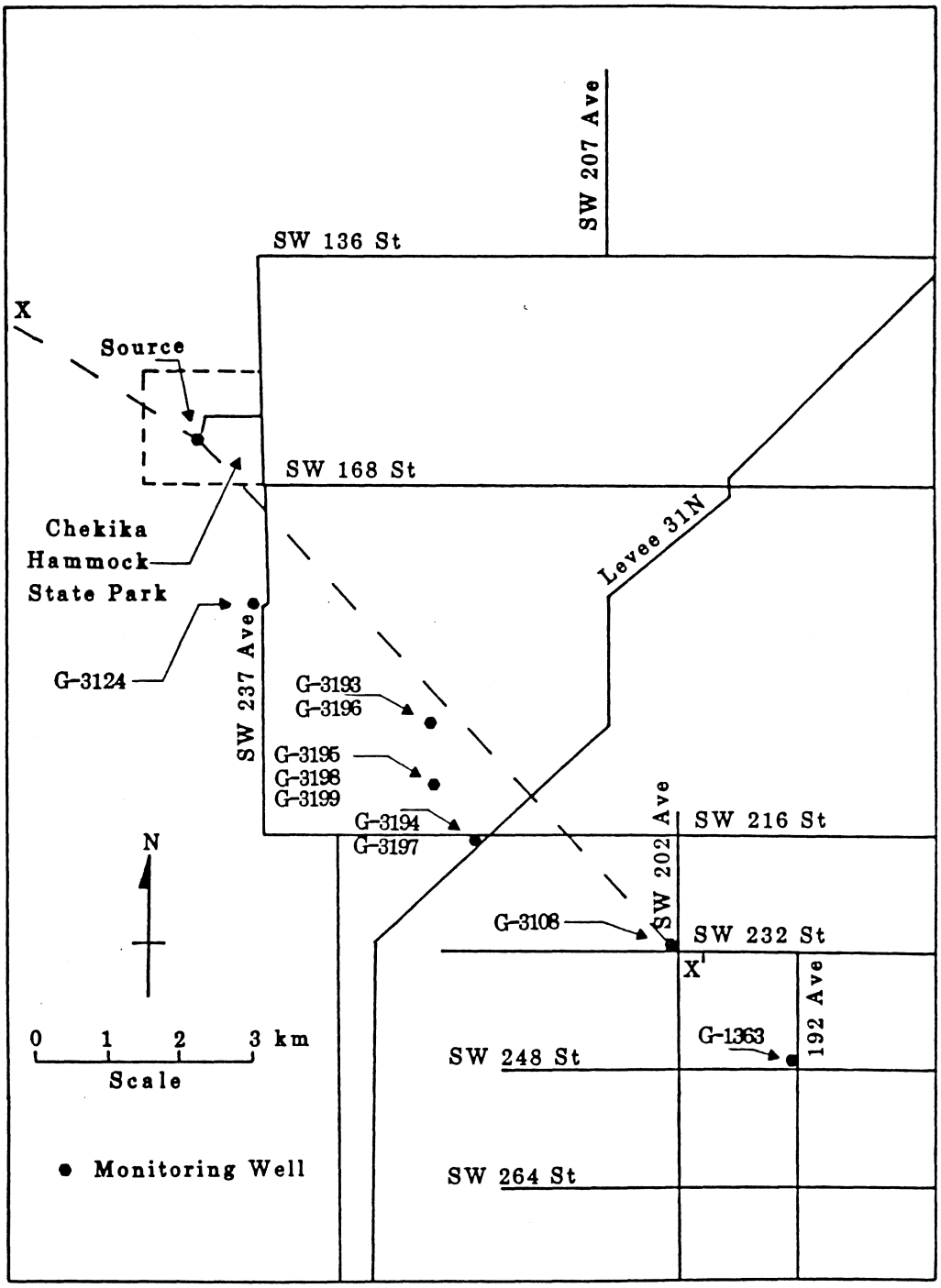


Figure 5.3: Area Surrounding Grossman Well
(After Waller, 1982)

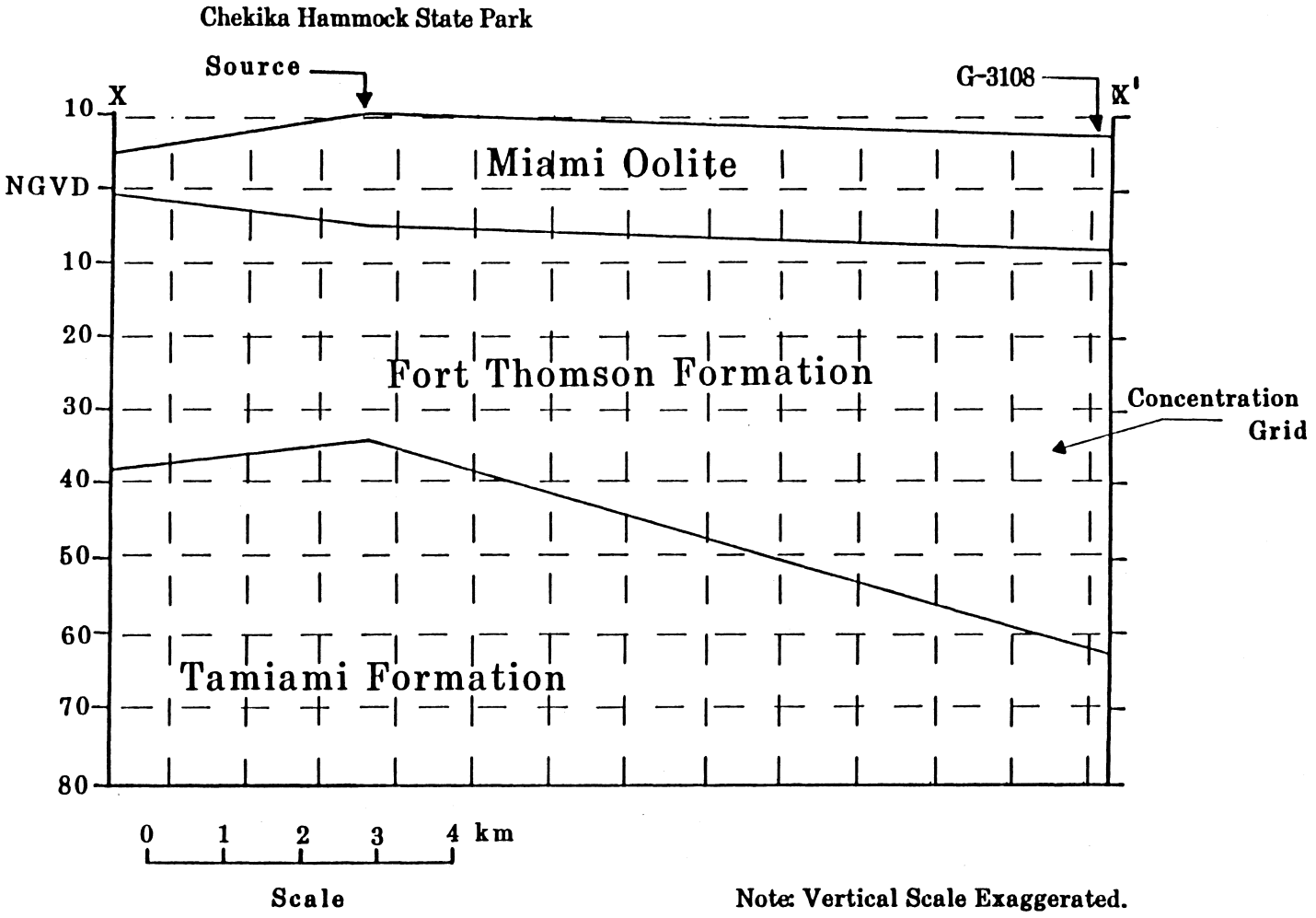


Figure 5.4: Aquifer Cross-Section (After Waller, 1982)

illustration of the chloride plume is shown in Fig.5.5, which illustrates the results of an electromagnetic conductivity survey of the area in March, 1979. The conductivity measurements are representative of the chloride levels in the upper 25 feet of the aquifer (Waller, 1982).

5.3 Model Parameters

The Chin (1987) dispersion model utilizes an array of monitoring wells as nodes in a finite element grid. The time series measurements of head at the nodes are combined with the hydraulic conductivity, effective porosity, and Darcy's law to estimate the seepage velocity fluctuations within the elements. The velocity components are then divided into diffusive and advective components. The diffusive components correspond to fluctuations on time scales less than the time step used in the model and fluctuations caused by the sub-grid scale variations in the hydraulic conductivity. The advective velocity component corresponds to seepage velocity fluctuations with time scales greater than the time step of the model.

The finite element grid used in this study is shown in Fig.5.6. The Grossman well approximately coincides with node 13. Time series measurements of head for the 13 year period from October 1973 to September 1986 were available for all nodes at a time interval of 1 day. Only sporadic measurements at all nodes were available prior to 1973. Since it was necessary to simulate the movement of the tracer for 41 years, 1944 to 1985, a 13 year cycle in the measured heads was assumed. While this assumption is essentially one of necessity, the error introduced is probably not significant since the time series

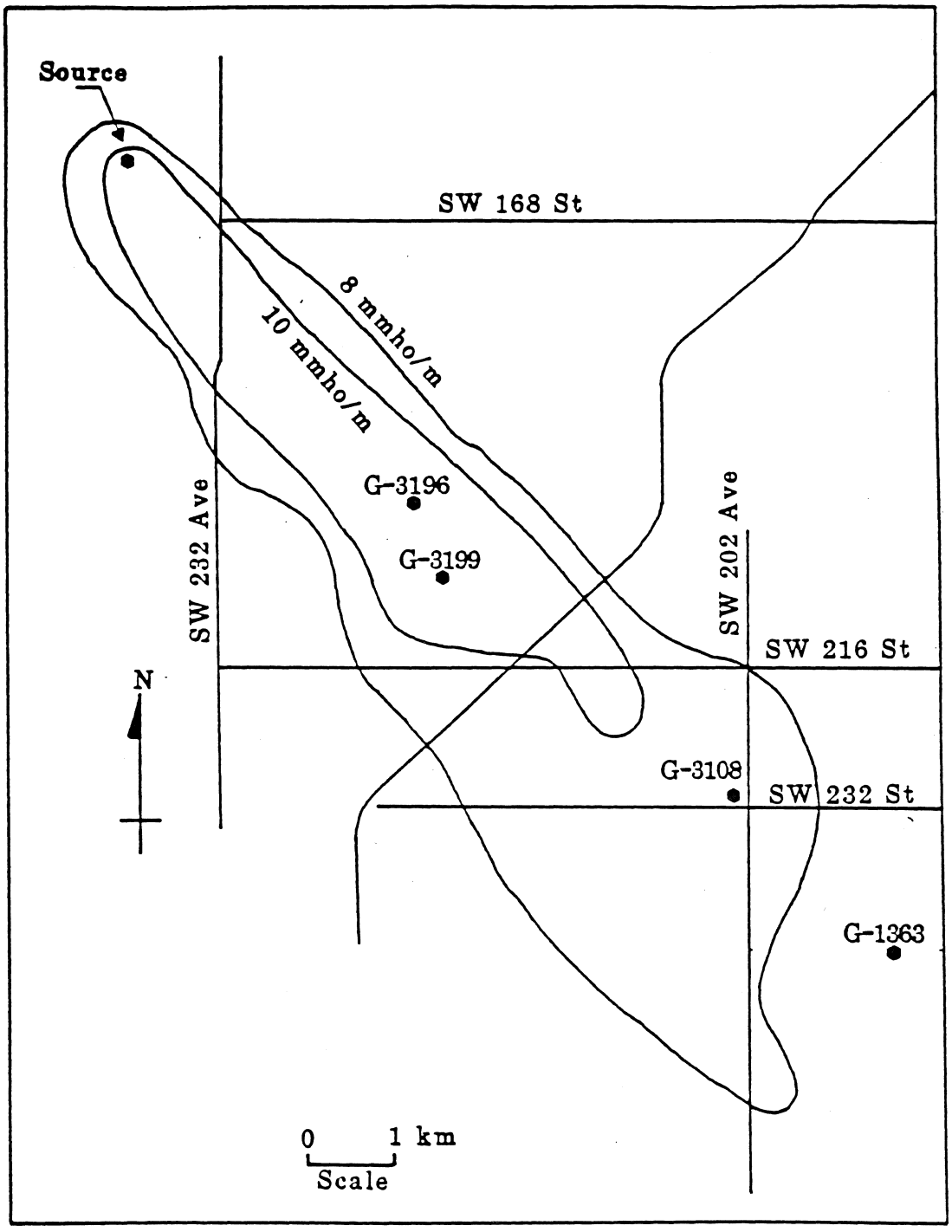


Figure 5.5: Measured Electromagnetic Contours
(After Waller, 1982)

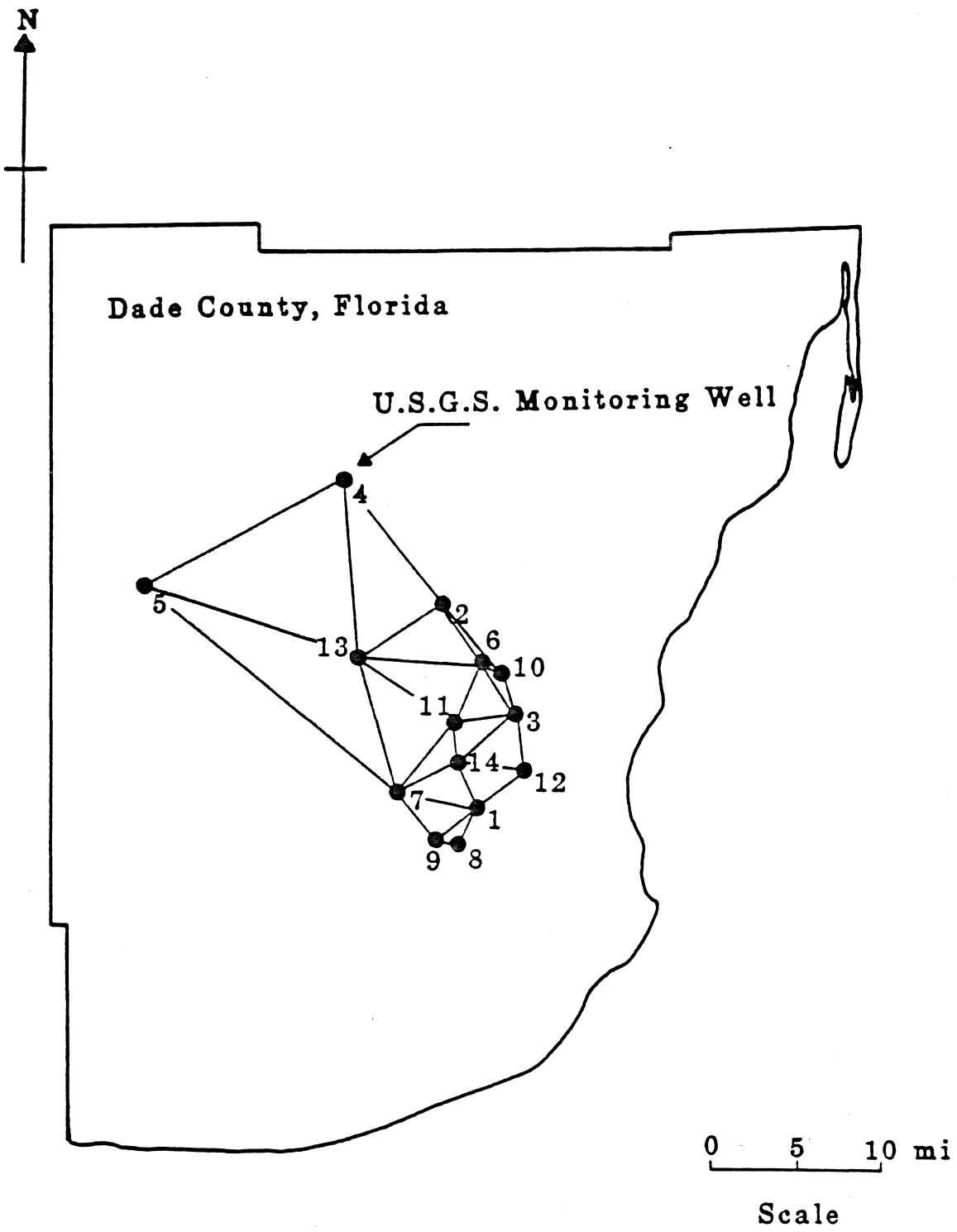


Figure 5.6: Finite Element Grid

measurements are dominated by a yearly cycle, and the time series parameters show a remarkable degree of stationarity from year to year. The properties of the time series parameters at nodes 4, 5, and 13 have been documented extensively by Chin (1988a). The hydraulic conductivity of the Miami Oolite and Fort Thomson Formation were obtained from documented measurements and, as indicated previously, are 1050 m/day (3440 ft/day) and 8380 m/day (27 500 ft/day) respectively. The effective porosity of these formations is 0.2 and the Tamiami Formation was assumed impermeable. The local dispersivity associated with sub-grid hydraulic conductivity variations was obtained using a formulation presented by Chin (1986b). Assuming the formation is isotropic and homogeneous within each grid, Chin (1986b) showed that the local longtitudinal dispersivity may be estimated by

$$\alpha_L = 0.49 \frac{\sigma_K^2}{(1 + \sigma_K^2/6)^2} L \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5.1)$$

where σ_K is the standard deviation of the Napierian logarithm of the hydraulic conductivity distribution and L is the correlation length scale of the hydraulic conductivity. For limestone formations, previous results indicate that $\sigma_K = 1.14$ is fairly typical (Chin, 1986a). We may tentatively assume that L is scaled by the size of the fractures and dissolution cavities in the limestone. The typical size of the cavities may be related directly to the magnitude of the hydraulic conductivity, by the method presented by Chin (1986b), which yields a typical cavity size of 1.5 cm for the Miami Oolite and 10 cm for the Fort Thomson Formation. Taking L to be an order of magnitude greater than this size

(Chin, 1986a) yields an estimated local dispersivity of 0.1 m for the Miami Oolite and 1 m for the Fort Thomson Formation. Taking the transverse local dispersivity, α_T, to be given by 0.1 α_L (Chin, 1986a) we obtain α_T = 0.01 m and 0.1 m for the Miami Oolite and Fort Thomson Formation respectively. It will subsequently be shown that dispersion predictions on the scale of interest are quite insensitive to the local dispersivities and are dominated by the accurately estimated spatial and temporal variations in seepage velocity.

5.4 Instantaneous Source

It is instructive to study the dispersion of an instantaneous release from the source in order to address two basic questions: (a) How does the rate of growth of dispersivity compare with the composite results of field tests (Gelhar, 1986)?; and (b) How sensitive are the predictions to the time step used in the model? To answer these questions, a mass of contaminant was released from a 0.028 m^3 (1 ft^3) cube directly below the seepage basin and in the center of the Miami Oolite. The eigenvalues of the variance of the resulting mass distribution were calculated. These principal variances are given by σ_{ii}^2 where i=1 corresponds to the principal longitudinal direction, i=2 the horizontal transverse direction, and i=3 the vertical direction. The dispersivity of the tracer cloud in the i direction, A_i, is related to the variance by the equation

$$A_i = \frac{1}{2v} \frac{d\sigma_{ii}^2}{dt} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5.2)$$

where v is the seepage velocity and t is time. The values of the

longitudinal dispersivity, A_1, associated with dispersion of a tracer released at the Grossman well site is shown in Fig.5.7 as a function of the distance of the mass centroid from the source. The results of previous tracer studies in porous media are also shown in Fig.5.7. These results demonstrate that the present results are entirely consistent with field measurements. In these simulations a time step of 30 days was used and the tracer plume was tracked for 13 years. In order to assess the sensitivity of the results to the model time step, the simulation was repeated for a time step of 1 year (365 days) and again the plume tracked for 13 years. The results are shown in Fig.5.7 and indicate that the dispersivity, A_1, converges to the value obtained using a time step of 30 days. This convergence is essentially complete within 1 km from the source. The effect of increasing the time step was further investigated by comparing the growth of the longitudinal and transverse variances for Δt = 30 and 365 days. These comparisons are shown in Fig.5.8. These results show that a time step of 365 days generally leads to an increased plume size, particularly in the horizontal dimensions. This effect is a result of parameterizing the cyclic yearly fluctuations in the seepage velocity by a statistical representation, when Δt = 365 days, compared with explicitly simulating the measured yearly variations, when Δt = 30 days. In this study, it is necessary to simulate the movement of the plume over 41 years based on only 13 years of measurements. Therefore accurate reproduction of the statistics of the measured yearly seepage velocity fluctuations is appropriate and Δt = 365 days will be used. This approach can only be justified on theoretical grounds when the time of the simulation exceeds the Lagrangian time scale of the yearly seepage velocity

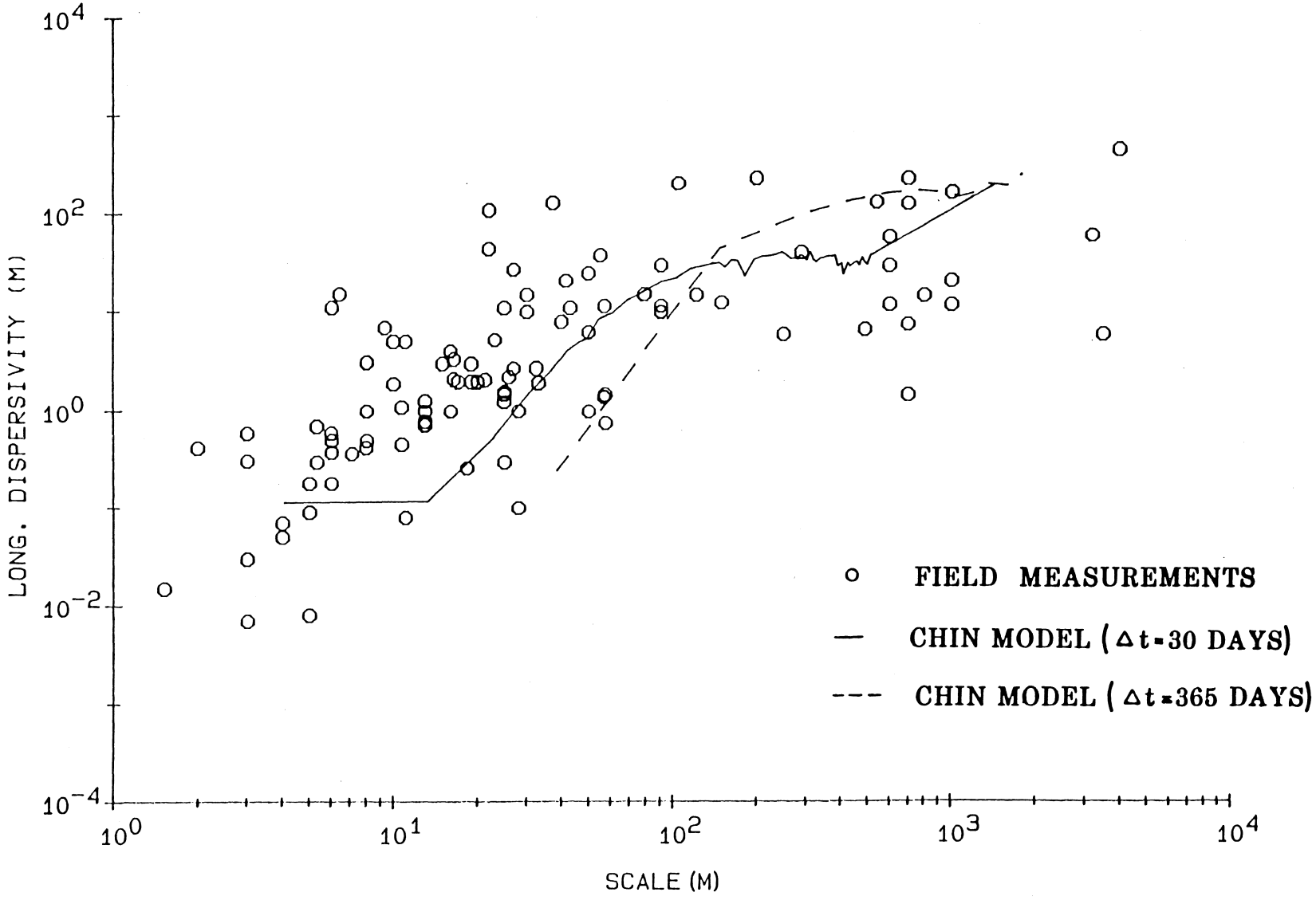


Figure 5.7: Dispersivity vs. Distance from Source

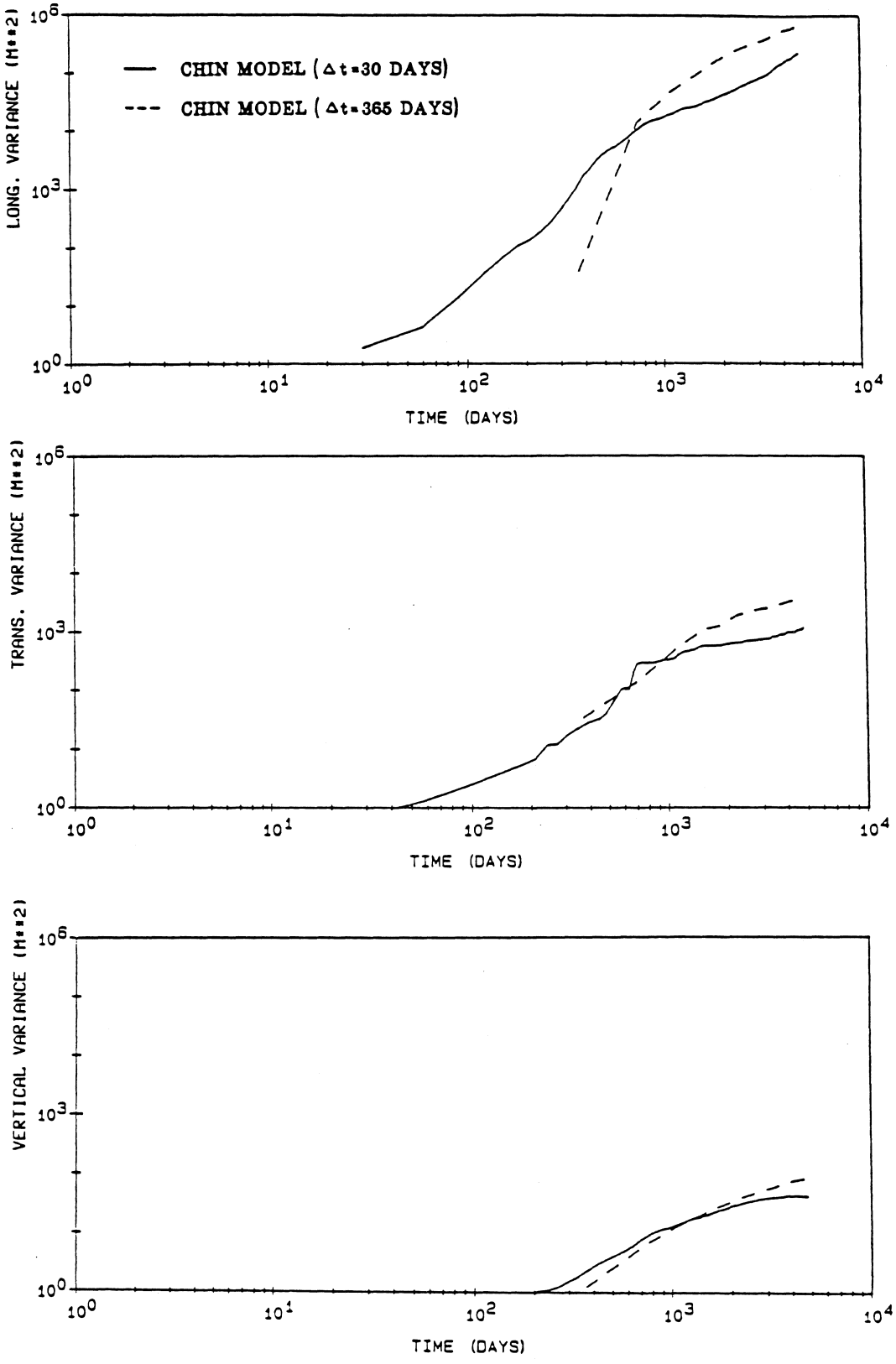


Figure 5.8: Variance vs. Time for an Instantaneous Release

fluctuations. The results presented in Fig.5.8 show that, using Δt = 30 days, the rate of growth of the variance becomes approximately constant after about 3 years (10^3 days) indicating that all Lagrangian time scales have been exceeded. Since the seepage velocity at the site is on the order of 2 m/day, we may consider the model predictions valid beyond 2 km from the source.

5.5 Continuous Source

The seepage basin adjacent to the Grossman well is a source of continuous mass flux of chloride. The dimensions of the basin, Fig.5.1, are approximately 300 m by 77 m. Combining the well discharge rate with the dimensions of the seepage basin indicates that significant vertical seepage velocities, on the order of 1 m/day may be expected in the vicinity of the seepage basin. Combining this condition with the negative buoyancy of the saline discharge indicates that vertical penetration of the plume occurs relatively rapidly in the vicinity of the source, where the total depth of the aquifer is only 14 m. This conclusion is supported by the observations of Waller (1982). In specifying the source in the Chin model, it is assumed that the plume extends directly below the seepage basin with a chloride concentration of 1200 mg/l. The chloride levels in the well discharge are given in Fig.5.2. The Chin (1987) model yields estimates of the concentration averaged over some specified volume. In general, the selected volume must be comparable to the size of the finite element grid, which controls the resolution of the model. With this in mind, the predicted concentration is averaged over a region measuring 1 km in the predominant flow direction, 0.5 km in the cross-stream direction, and 3

m in the vertical direction. An elevation view of the concentration grid is shown in Fig.5.4. Contours of predicted concentration distributions, in mg/l, are plotted by assuming the grid averaged concentration equals the concentration at the center of the grid. The growth of the plume at 10 year intervals is shown in Figs.5.9 to 5.12. This figure clearly indicates that the predominant flow direction is to the southeast, and that after 40 years the plume extends approximately 20 km (12.4 miles) from the source and is about 2.5 km (1.6 miles) wide. The model predictions may be compared directly with the conductivity measurements of Waller (1982) shown in Fig.5.5 which were taken in 1979 and show a contour of approximately background conductivity (8 mmhos/m), and a contour of the conductivity 25 % higher (10 mmhos/m). These levels are representative of the conductivity in the upper 7.6 m (25 ft) of the aquifer (Waller, 1982). These measurements may be compared with the predicted contour of concentrations slightly above background, 30.1 mg/l, and the contour of concentrations 25 % higher than background, 37.5 mg/l. Background levels were assumed to be 30 mg/l based on measurements reported by Klein and Hull, 1978. Simulated concentrations in the layer between 3 and 6 m below the surface were used for the comparison. These concentrations are representative of the average concentrations in the upper 7.6 m of the aquifer. The comparison between the predicted and measured results, 35 years after the initiation of the source is shown in Fig.5.13. The agreement between the measured and predicted levels is good and show that the Chin model has quite accurately modeled both the direction of movement and the spreading rate, especially within several kilometers of the source. The result gives support to our ability to

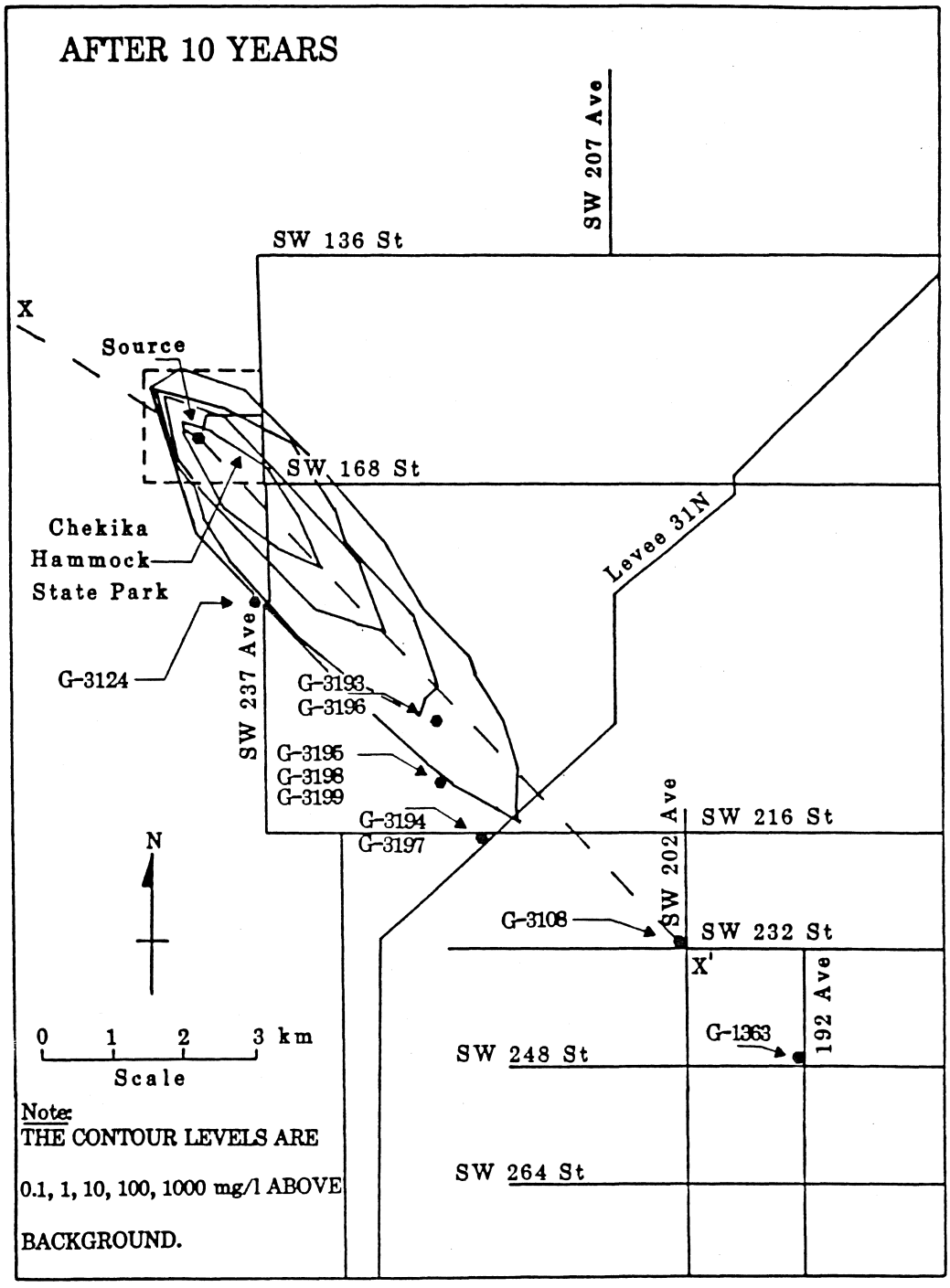


Figure 5.9: Predicted Chloride Plume After 10 Years

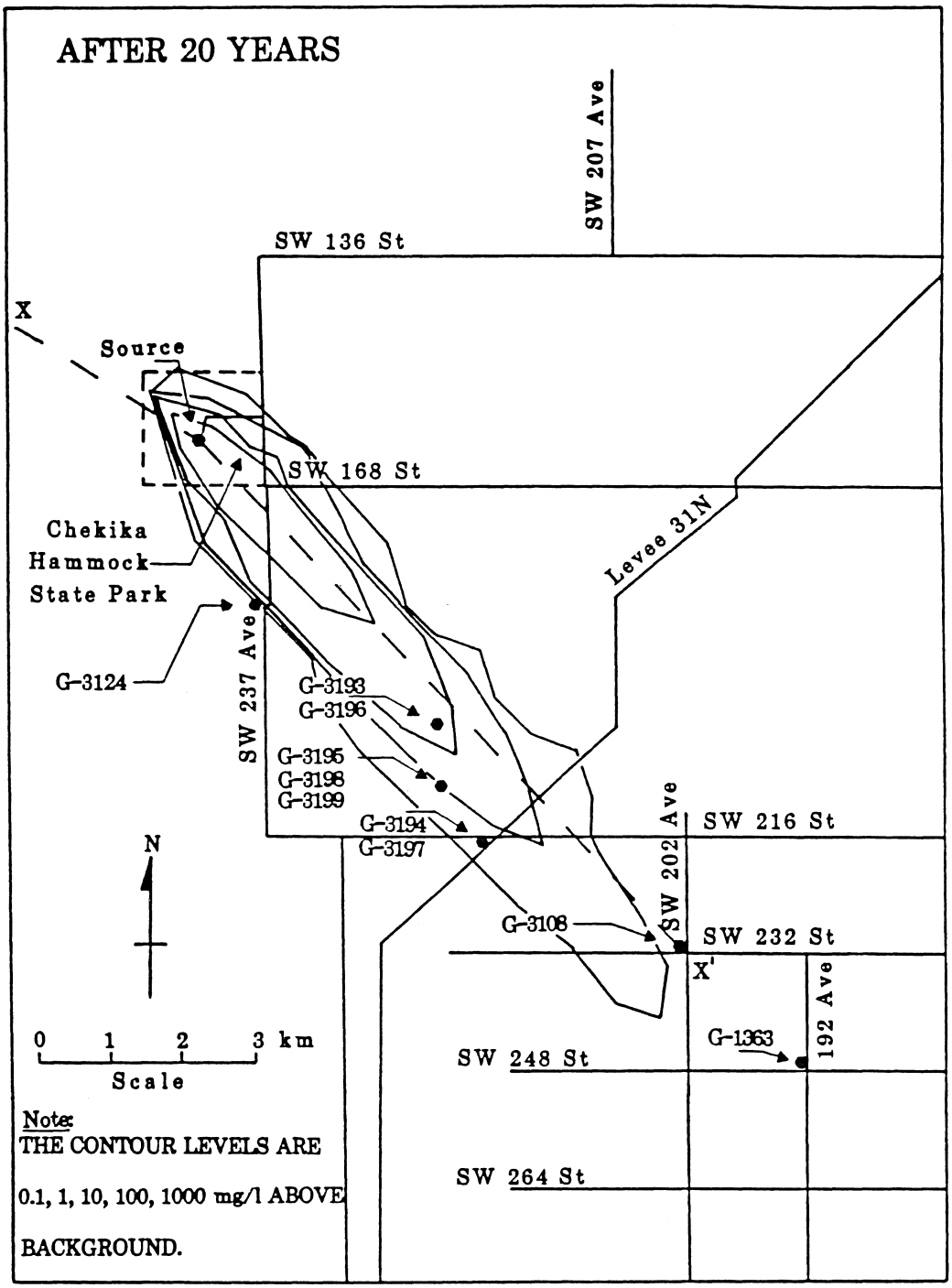


Figure 5.10: Predicted Chloride Plume After 20 Years

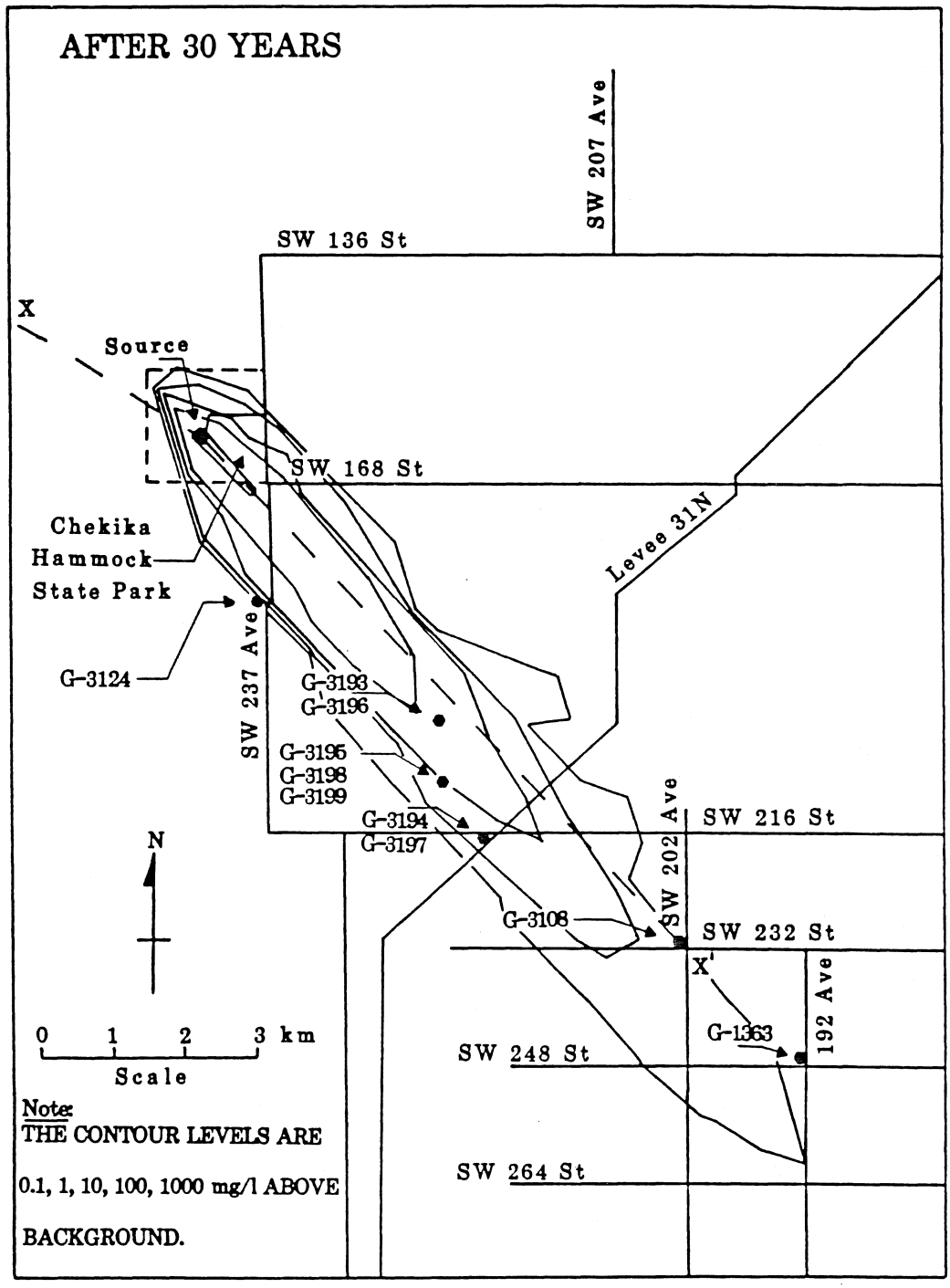


Figure 5.11: Predicted Chloride Plume After 30 Years

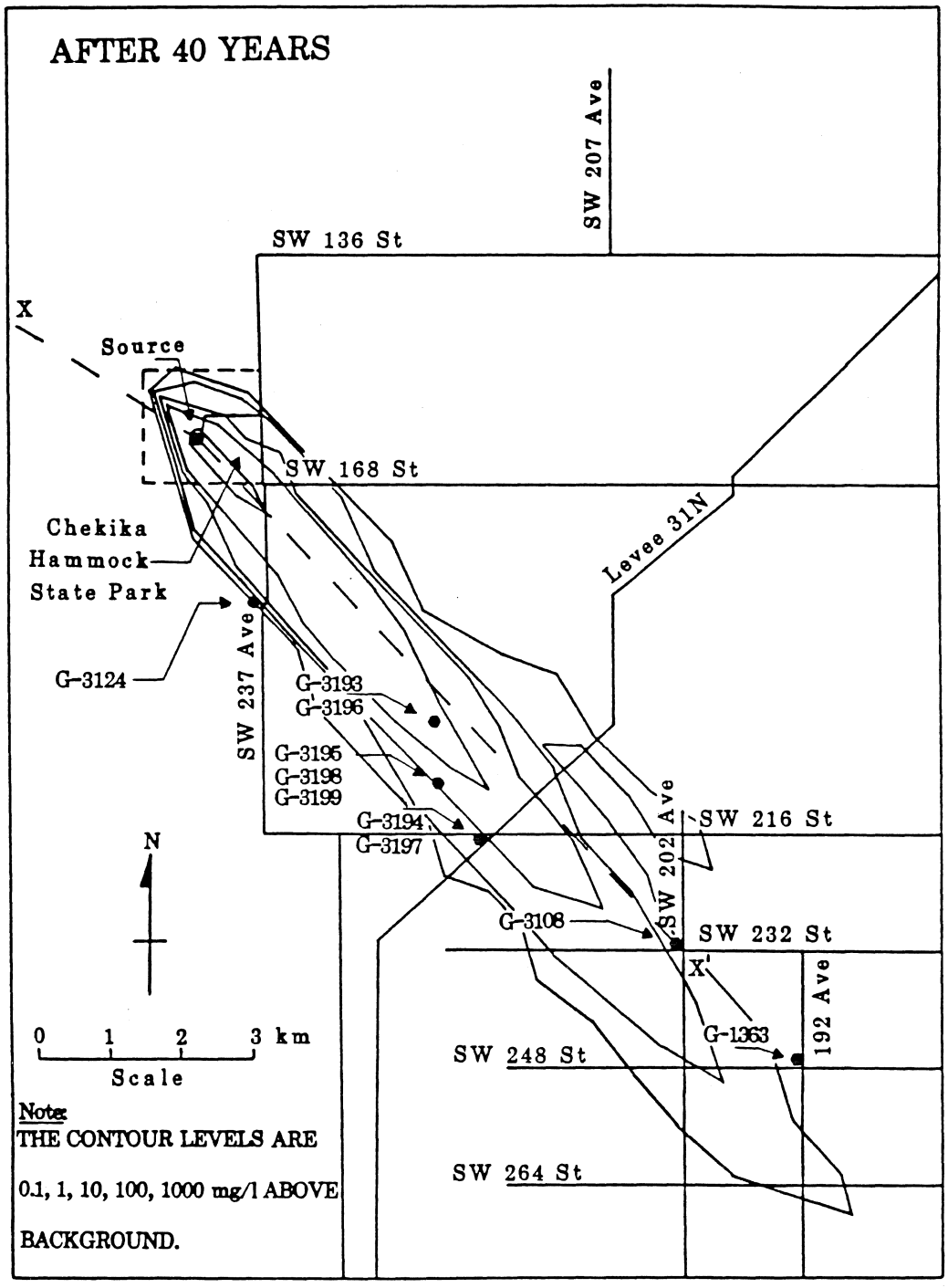


Figure 5.12: Predicted Chloride Plume After 40 Years

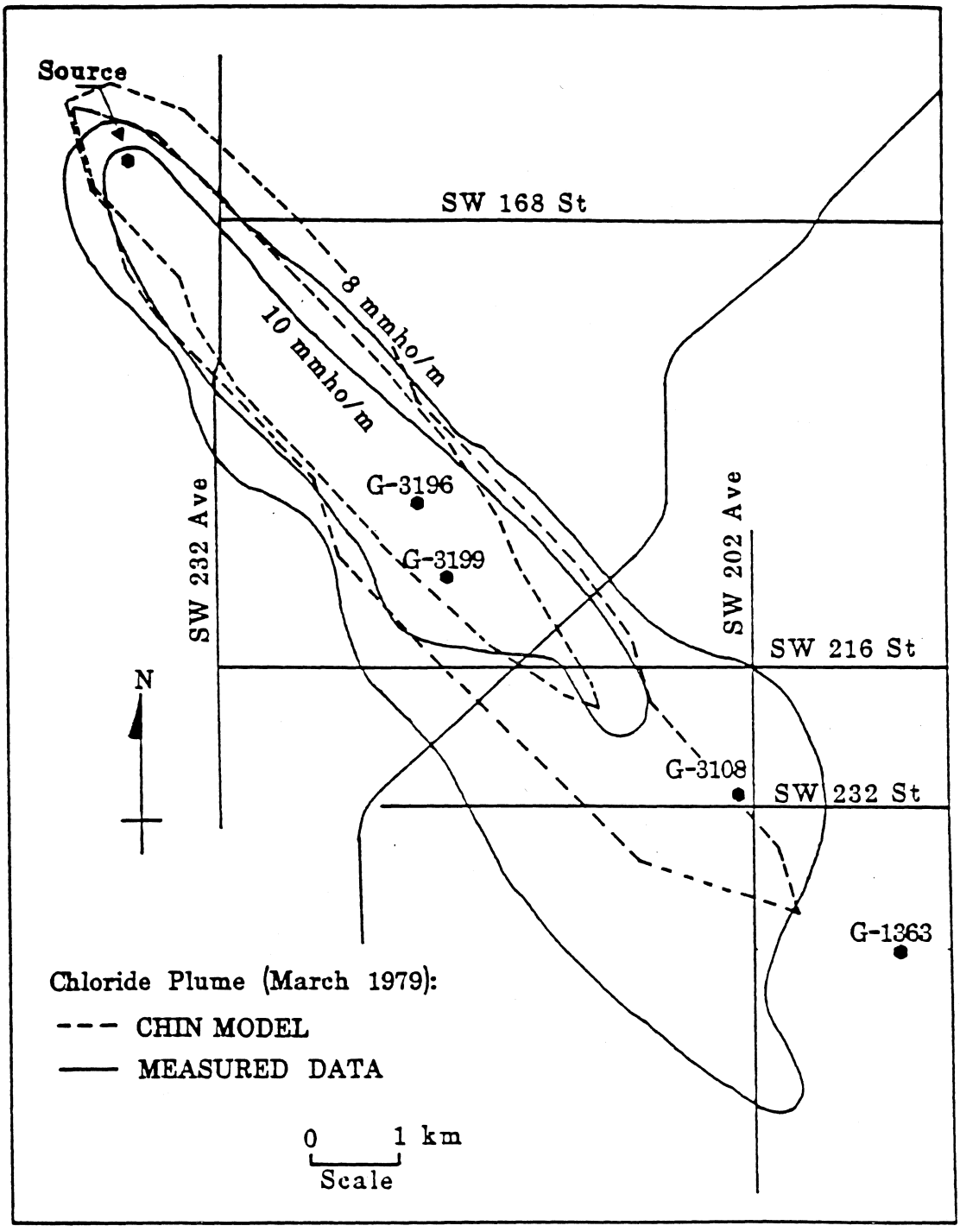


Figure 5.13: Comparison Between Predicted and Measured Plume

predict contaminant migration based only on measurable hydrogeological parameters. A more stringent test of the model is to compare the predicted concentration levels at the locations of the monitoring wells where the data on chloride levels are available. The locations of the monitoring wells in the vicinity of the source are shown in Fig.5.3. At three of the six monitoring well locations shown, the screened interval is located at several depths and multiple well designations are given. The chloride levels measured at the monitoring wells have generally indicated that the plume is well mixed over the depth of the aquifer. Since the Chin model is three-dimensional, we are able to predict the vertical variation of the chloride concentration at each monitoring well location. The measured chloride levels at the monitoring wells are compared with the predicted maximum, minimum, and average concentrations in Figs.5.14 to 5.16. These results show that, within 6 km of Grossman well, the observed chloride concentrations are predicted quite accurately by the Chin model. This is a major accomplishment of the model since the predictions are based entirely on measurable hydrogeological parameters that are independent of the field dispersion measurements. Beyond 6 km, the Chin model does predict some increase in chloride level, however the actual concentrations are underestimated.

5.6 Sensitivity Analysis

In any modeling exercise in which parameters can not be specified with certainty, a sensitivity analysis is in order to ascertain how this uncertainty affects model predictions. In the present case, the formation parameters that may vary from the 'best estimate' values used in this study are the hydraulic conductivity, porosity, and the local

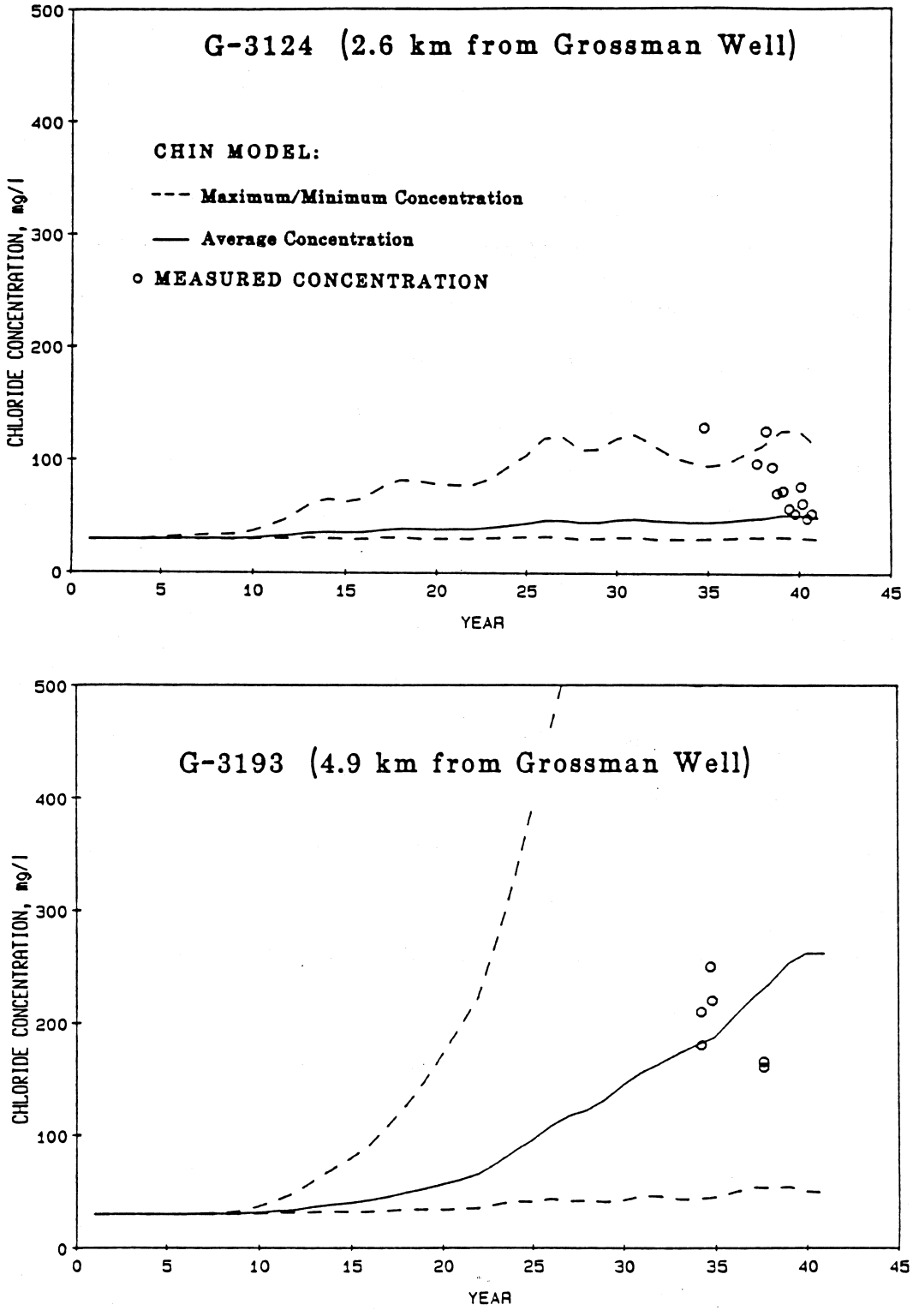


Figure 5.14: Comparison Between Predicted and Measured Concentrations At 2.6 km and 4.9 km from the Source

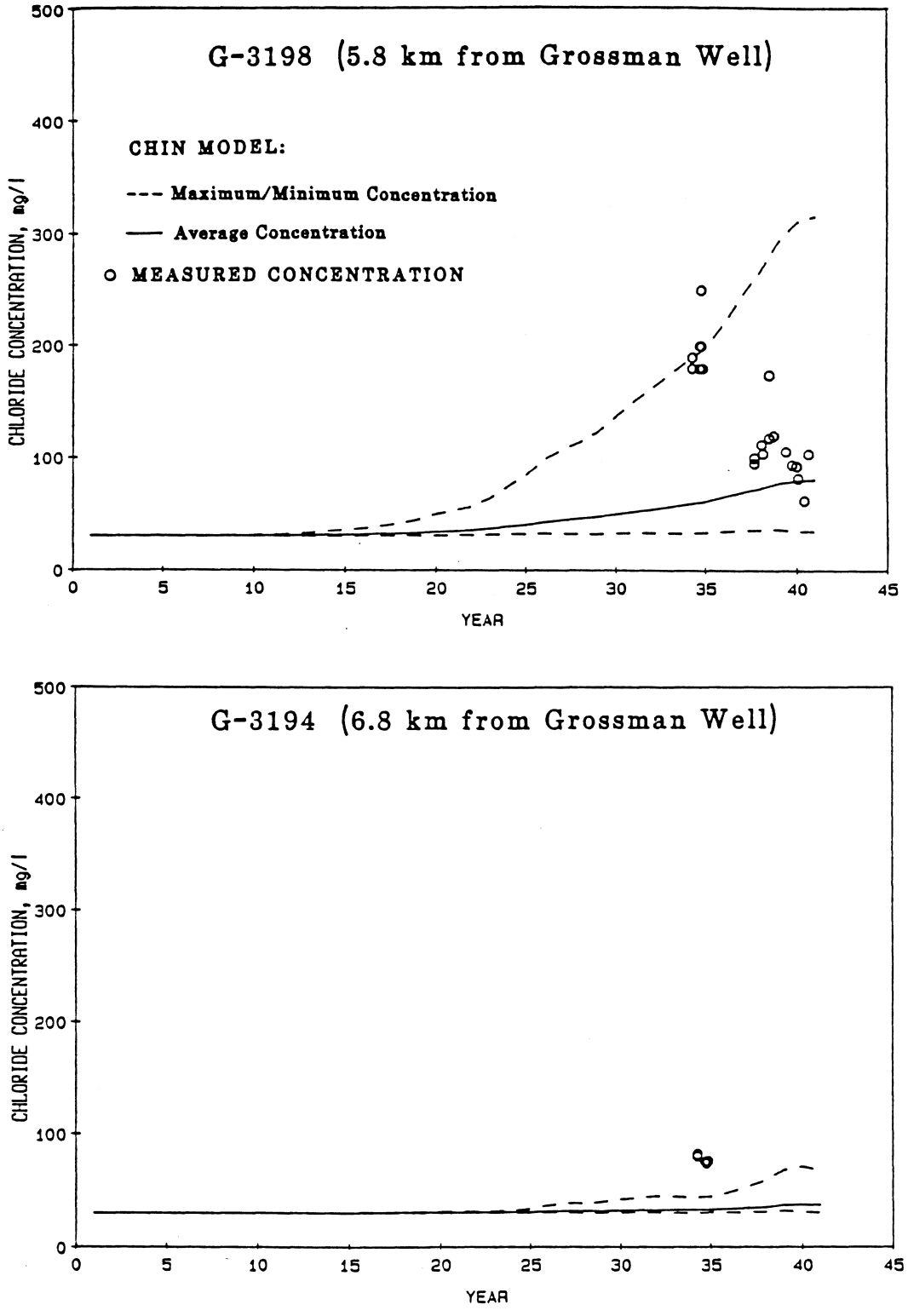


Figure 5.15: Comparison Between Predicted and Measured Concentrations At 5.8 km and 6.8 km from the Source

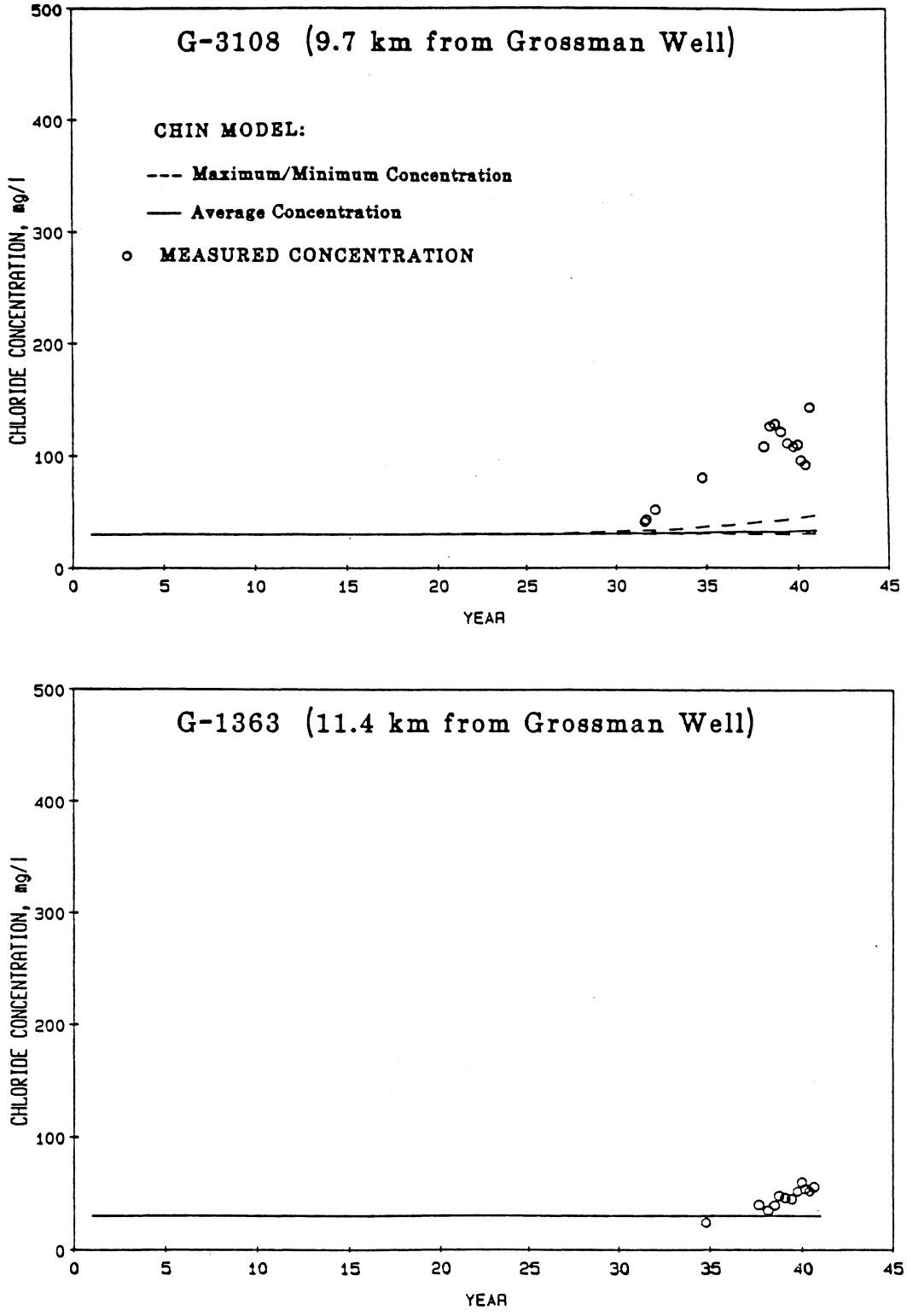


Figure 5.16: Comparison Between Predicted and Measured Concentrations At 9.7 km and 11.4 km from the Source

dispersivities. The sensitivity of the results to these parameters will be addressed in this section.

Hydraulic conductivities estimated using pump tests are generally assigned an accuracy of ± 20%. Recent pump tests in the project area have indicated that hydraulic conductivities may be somewhat higher than previously published values (Fish, 1987). Hydraulic conductivities were increased and decreased by 50% and the simulations repeated. The result is shown in Fig.5.17. These results indicate that the predicted contours are quite sensitive to such a large change in the hydraulic conductivity. As expected, the amount of mixing is increased by increasing the hydraulic conductivity.

Based on published measurements, the specific yield and porosity in the aquifer are approximately 0.2 and 0.3 respectively. The effective porosity, for calculating seepage velocities, is generally assumed to be adequately approximated by the specific yield. The impact of this assumption is reflected in Fig.5.17. The contrasted results indicate that the predictions are sensitive to the assumed porosity, with an increased porosity resulting in a shorter plume, although the maximum width of the plume is not significantly altered.

The values of dispersivity associated with sub-grid spatial variations in hydraulic conductivity are generally the focus of attention when modeling dispersion in porous media. The sensitivity of the model results to a 50% change in the dispersivities are shown in Fig.5.18. These results indicate that the model predictions are not very sensitive to this degree of variability in the dispersivity. This result is attributable to the formulation of the Chin model where an attempt is made to account explicitly for the velocity variations

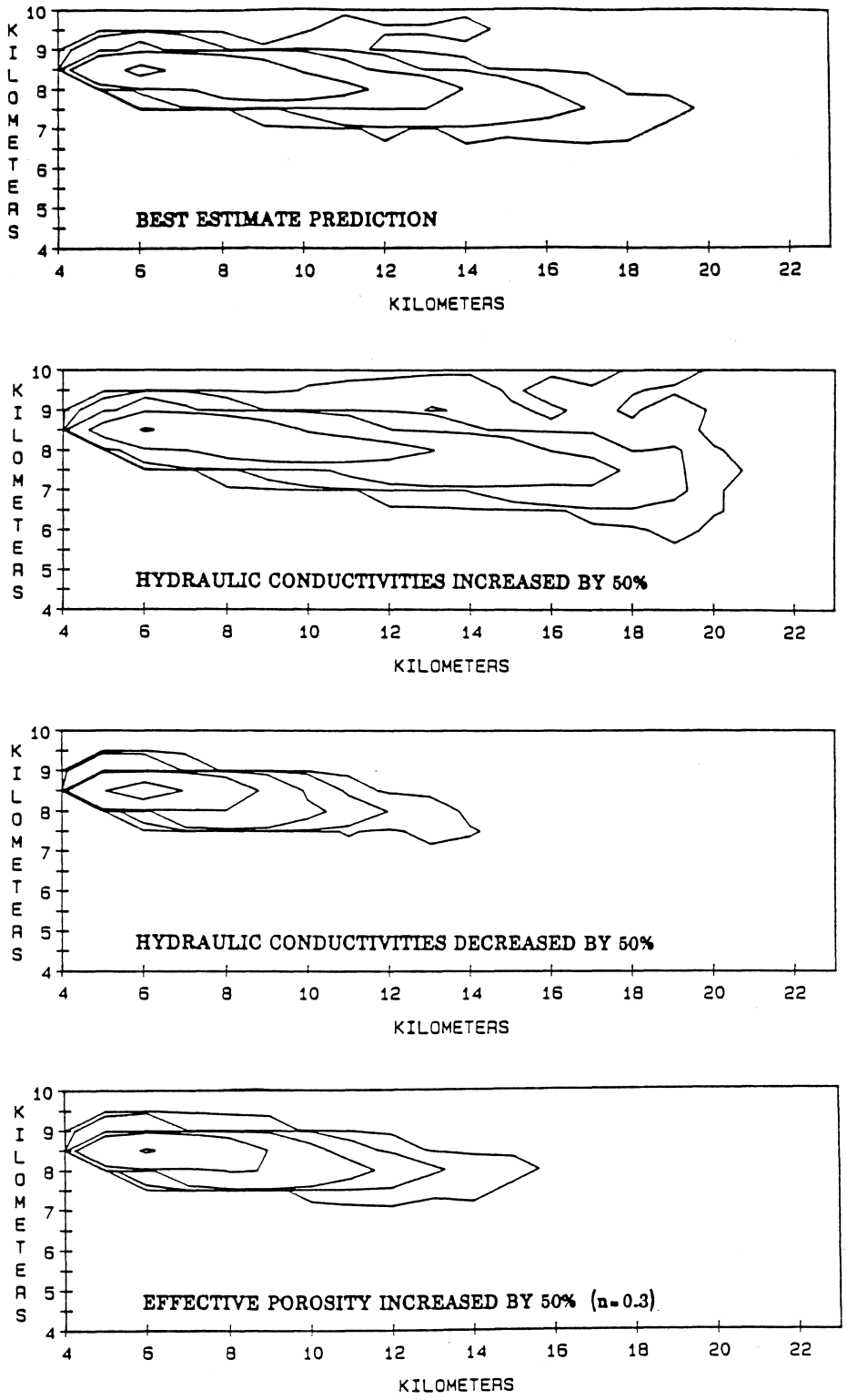


Figure 5.17: Sensitivity of Predictions to Hydraulic Conductivity and Porosity

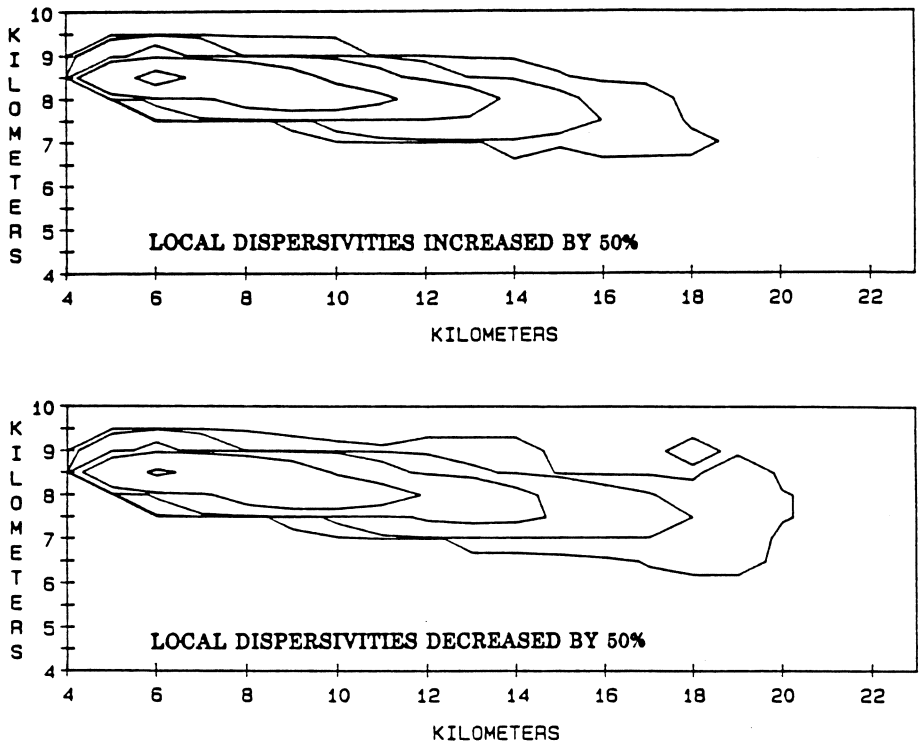


Figure 5.18: Sensitivity of Predictions to Dispersivity

dominating the mixing process. The primary effect of the dispersivity reduction is to limit the vertical mixing of the plume. This is reflected in increased plume concentrations encountered at given locations within the aquifer. It is noteworthy that the predicted results are more sensitive to the hydraulic conductivity than the local dispersivity.

5.7 Conclusions

The results of this study have demonstrated that it is possible to produce reasonable estimates of contaminant mixing in groundwater based on measurable hydrogeological parameters. The present study further indicates that accurate resolution of the spatial and temporal head variations along with the principal geologic formations and their mean hydraulic conductivities are primary requirements for accurately modeling dispersion. Of course, one should not attempt to predict mixing on scales significantly smaller than the resolution of the measurements. In the formation studied, an instantaneous release shows a growth in dispersivity that is in remarkable agreement with reported field results at several sites. This supports the basic formulation of the Chin model and clearly demonstrates that this model is superior to models which use empirical equations to describe the growth of dispersivity (e.g. Simmons, 1982), since these models would use the same empirical equations for all sites. For a continuous release from the source, the agreement between the predicted and measured plume trajectory is excellent. For the concentrations measured at certain points in the aquifer, the measured concentrations were predicted quite accurately by the Chin model, particularly within 6 km of the source.

The sensitivity analysis has shown that the model predictions are significantly less sensitive to the local dispersivity than to the hydraulic conductivity and porosity. This is significant since the dispersivity is the parameter that is estimated with the least degree of certainty. The non-empirical nature of the Chin model combined with the demonstrated validity of the model clearly shows that accurate prediction of contaminant migration in porous media over large distances is indeed possible.

CHAPTER VI

SUMMARY AND CONCLUSIONS

6.1 Summary

A model has been developed which predicts the dispersion of a contaminant in groundwater based only on measurable hydrogeological parameters. The measured data include time series of piezometric head at several locations, as well as the hydraulic conductivities and porosity within the aquifer. The model formulation has been verified by applying the dispersion model in idealized formations and comparing the predicted results with analytic solutions of the advection-dispersion equation. Both instantaneous and continuous releases were studied and, in both cases, the agreement with the Chin model was excellent. In order to apply the model in the field, time series of piezometric head measurements must be available at several locations. It is commonplace to find gaps in the measured data, frequently caused by equipment malfunction. To alleviate this problem, a technique to fill in missing head data has been developed. The technique was validated by using synoptic head measurements at a pair of locations to estimate the relationship between the measured heads and, using this calculated relationship, compare the measured head at one station with that predicted using the measured head at the other station along with the estimated relation between the measurements at the two stations. The results of this comparison demonstrated that the proposed estimation technique was capable of producing reliable estimates of the missing data. With a full complement of synoptic data at several nodes, the Chin model was applied at a site where there are detailed measurements

of the hydrogeology. In addition, fairly extensive measurements were available for a plume of contamination resulting from the continuous discharge of highly saline water from an artesian well that was in operation for about 40 years. The conservative nature of the chloride ions, the well defined plume, and the localized nature of the source, provided an excellent oppurtunity to compare the measured chloride levels with those predicted by the Chin model based only on measured hydrogeological data. The results showed that, within 6 km from the source, excellent predictions of chloride levels were obtained. Beyond 6 km, the measured chloride levels were underestimated. These results provide direct evidence that the Chin model is capable of predicting dispersion in groundwater with a fair degree of certainty, especially within the aquifer used in this study. Most field studies to date have been concerned with measuring the rate of growth of variance as a function of the distance of the source for a instantaneously released tracer. These results have shown a fair degree of consistency, especially in the longtitudinal (flow) direction. Using the Chin model, the instantaneous release of a tracer was simulated at the site and the rate of growth of variance calculated. The results were in remarkable agreement with those previously reported.

6.2 Conclusions

Based on the results of this study, we may conclude that the Chin model has been shown to provide an accurate description of dispersion at the site studied. This result demonstrates that the formulation of the model provides the capability to accurately predict dispersion in a case where analytic results are not applicable, in-situ tracers are not

used for calibration, and theoretical asymptotic dispersivities are not assumed. Because of the non-empirical nature of the Chin model, the fact that the input data consists of measurable hydrogeological parameters, and the demonstrated validity of the model, then the Chin model should prove to be a very useful tool to predict dispersion in groundwater, especially at locations where no present contamination exists.

REFERENCES

1. Anderson, M.P. (1979). "Using models to simulate the movement of contaminants through groundwater flow systems." *CRC Crit.Rev. Envir. Control*, 9(11), 97-156.

2. Bras, R.L., and Rodriguez-Iturbe, I. (1985). *Random Functions and Hydrology*. Addison-Wesley Publishing Co., Reading, Massachusetts.

3. Brillinger, D.R. (1981). *Time Series: Data Analysis and Theory*. Expanded Edition, Holden-Day Inc., San Francisco, California.

4. Chin, D.A.(1986a). "Estimation of dispersion coefficients in porous media." *Journal of Hydraulic Engineering*, ASCE, 112(7), 591-609.

5. Chin, D.A. (1986b). "Predicting contaminant movement in groundwater." *Proceedings, 1986 Annual Meeting, ASCE South Florida Section*, West Palm Beach, Florida.

6. Chin, D.A. (1987). "Macrodispersion in stratified porous media." *Journal of Hydraulic Engineering*, ASCE, 113(10), 1343-1358.

7. Chin, D.A. (1988a). "Spatial correlation of hydrologic time series," *Journal of Water Resources Planning and Management*, ASCE, in press.

8. Chin, D.A. (1988b). "Diagnostic model of dispersion in porous media." *Journal of Hydraulic Engineering*, ASCE, under review.

9. Dagan, G. (1984). "Solute transport in heterogeneous porous formations." *Journal of Fluid Mechanics*, 145, 151-177.

10. Dasgupta, D., Sengupta, S., Wong, K.V., and Nemerow, N. (1984). "Two-dimensional time dependent simulation of contaminant transport from a landfill." *Journal of Applied Mathematical Modelling*, 8, 203-210.

11. Fish, John (1987), Geologist, U.S.G.S. Miami Office, private communication.

12. Gelhar, L.W., and Axness, C.L. (1983). "Three-dimensional stochastic analysis of macrodispersion in aquifers." *Water Resources Research*, 19(1), 161-180.

13. Gelhar, L.W., Mantoglu, A., Welty, C. and Rehfeldt, K.R. (1985). "A review of field scale physical solute transport processes in saturated and unsaturated porous media." *Report No. EA-4190*, Electric Power Research Institute, Palo Alto, California.

14. Gelhar, L.W. (1986). "Stochastic subsurface hydrology from theory to applications." *Water Resources Research*, 22(9), 135S-145S.

15. Guven, O., Molz, F.J., and Melville, J.G. (1984). "An analysis of dispersion in a stratified aquifer." *Water Resources Research*, 20(10), 1337-1354.

16. Jenkins, G.M., and Watts, D.G. (1968). *Spectral Analysis and Its Applications*. Holden-Day Inc., San Francisco, California.

17. Klein, H., and Hull, J.E. (1978). "Biscayne aquifer, southeast Florida." *U.S. Geological Survey, Water Resources Investigation*, Report No.78-107.

18. Konikow, L.F., and Bredehoeft, J.D. (1978). "Computer model of two-dimensional solute transport and dispersion in ground water." *Techniques of Water-Resources Investigations of the United States Geological Survey*, Book 7, Chapter C2, USGS, Reston, Virginia.

19. Koopmans, L.H. (1974). *The Spectral Analysis of Time Series*. Academic Press, New York, New York.

20. Neuman, S.P., Winter, C.L., Newman, C.M. (1987). "Stochastic theory of field-scale Fickian dispersion in anisotropic porous media." *Water Resources Research*, 23(3), 453-466.

21. Parker, G.G., Ferguson, G.E., and Love, S.K. (1955). "Water resources of southeastern Florida." *U.S. Geological Survey, Water Supply Paper*, 1255.

22. Prakash, A. (1984). "Ground-water contamination due to transient sources of pollution." *Journal of Hydraulic Engineering*, ASCE, 112(7), 591-609.

23. Salas, J.D., Delleur, J.W., Yevjevich, V., and Lane, W.L. (1980). *Applied Modeling of Hydrologic Time Series*. Water Resources Publications, Littleton, Colorado.

24. Simmons, C.S. (1982). "A stochastic-convective ensemble method of representing dispersive transport in groundwater." *Electric Power Research Institute, Report No.CS-2558*, Project 1406-1, Palo Alto, California.

25. Todd, D.K. (1980). *Groundwater Hydrology*. John Wiley and Sons, Inc., New York, New York.

26. Waller, B.G. (1982). "Areal extent of a plume of mineralized water from a flowing artesian well in Dade county, Florida." *U.S. Geological Survey, Water Resources Investigation*, Report 82-20.

27. Yeh, T.C. (1987). "Comment on 'Modeling of scale-dependent dispersion in hydrogeologic systems' by J.F. Pickens and G.E. Grisak". *Water Resources Research*, 23(3), 522.