Università degli Studi di Parma
Facoltà di Ingegneria

Dottorato di Ricerca in Ingegneria Civile - XVIII Ciclo
Curriculum: Protezione Idraulica del Territorio (ICAR/02)

Andrea Zanini

Approccio Geostatistico per la Soluzione dei Problemi Inversi nelle Acque Sotterranee: Applicazioni e Sviluppi

Dissertazione per il conseguimento del titolo di Dottore di Ricerca

Tutore: Prof. Ing. Maria Giovanna Tanda
Co-Tutore: Prof. Ing. Peter K. Kitanidis
Coordinatore del Dottorato: Prof. Ing. Paolo Mignosa

Parma, Gennaio 2006

Università degli Studi di Parma
Facoltà di Ingegneria

Dottorato di Ricerca in Ingegneria Civile - XVIII Ciclo
Curriculum: Protezione Idraulica del Territorio (ICAR/02)

Andrea Zanini

Geostatistical Approach for Solving Inverse Problems in Groundwater: Applications and Improvements

Dissertation for the Degree of Doctor of Philosophy

Advisor: Prof. Eng. Maria Giovanna Tanda
Co-Advisor: Prof. Eng. Peter K. Kitanidis
Phd Coordinator: Prof. Eng. Paolo Mignosa

Parma, January 2006

All endings are also beginnings.
We just don't know it at the time...

M. Albom

Contents

List of Figures

List of Tables

Preface

This work is the result of the research activity performed by the writer during the Phd course (XIII Cycle) regarding the inverse problems in groundwater. The work consists of two parts, preceded by a short summary on the hydraulics of groundwater. The first part regards a general introduction on the inverse problems, explaining the main inverse methods and focusing on the quasi-linear geostatistical procedure; the second part verifies this methodology using synthetic cases and proposes new improvements to the two studied problems: the recovery of the release history of pollutants and the aquifer parameter estimation. At last the main codes developed for the study are shown in appendices.

First of all I would like to thank my advisor Prof. Maria Giovanna Tanda, who gave me the great opportunity of this Phd; she helped me to solve problems and sent me abroad to increase my knowledge. I thank Prof. Peter K. Kitanidis, my advisor at the Stanford University, who has allowed an intercontinental cooperation. Both the advisors are great volcanoes of advice. Ilaria Butera, for all of her suggestions. Mike Fienen and Matthew Sylvester for the review of this thesis. I thank heartily all the people, graduate and undergraduate students, with whom I shared the office in the last three years, to bear me. I thank Fabiana and all of my friends. Last but not least I thank my parents, to whom I dedicate this work; without them I could have never reached this goal.

Parma, January 2006

Introduction

The great interest in environmental issues has drawn the community to an attention to the quality of groundwater. Scientific efforts in groundwater flow studies have primarily focused on the flow and transport behavior and on the identification of the corresponding parameters.

After the nineties, increasing attention has been focused on the problem of recovery the release history of a pollutant. The knowledge of the pollution injection function provides information about the future pollution spread and allows better planning of remediation actions [*Liu and Ball,* 1999; *Snodgrass and Kitanidis,* 1997; *Skaggs and Kabala,* 1994; *Butera and Tanda,* 2003]. Moreover, from a legal and regulatory point of view, it is also important to determine the release time and duration and the highest values of concentration of the injected solution: an available release history can be a useful tool for sharing the costs of remediation of a polluted area among the responsible parties.

The Mathematical modeling is the basis of the studies of the evolution of the pollution. Setting up a model is a complex and a difficult task, because the main problem is the evaluation of the aquifer parameters. Usually the scientists have few field data, and with that they have to model wide areas; this implies the introduction of errors (due to the large approximations) into the modeling. These parameters, usually, are estimated from scarce data because they are difficult and costly to obtain. The accuracy of the estimate depends on the number of the measurements, their locations in the studied area, the observation error and the sensitivity of the observed quantity to the real field.

Both of the problems - release history and parameters identification - are represented by ill-posed problems especially inverse problems. The literature regarding the inverse problem is wide and regards several branches of sciences and mathematics. During the last 40 years several methods were developed to solve inverse problems, for instance the Tikhonov regularization, the minimum relative entropy theory, the adjoint state method and the geostatistical method.

The methodology applied in this work is the quasi-linear geostatistical approach proposed by *Kitanidis* [1995]. This approach was chosen because it is a statistical method so that it is possible to evaluate the unknown function and the related uncertainty to it at the same time. It has been widely applied during the last 10 years by several authors with good results [*Snodgrass and Kitanidis,* 1997; *Michalak and Kitanidis,* 2002, 2003; *Butera and Tanda,* 2003; *Boano et al.*, 2005]. This work presents applications and improvements of the quasi-linear geostatistical approach:

The first application concerns the **recovery of the release history** of the pollutants; it consists in the evaluation of the release function of a pollutant starting from concentration measurements. A brief literature review on this topic is presented and the geostatistical approach proposed by *Snodgrass and Kitanidis* [1997] and subsequent developments are summarized. A new improvement (evaluation of the transfer function) about the possibility to apply the methodology to non uniform flow cases (pumping well, heterogeneous hydraulic conductivity fields, etc.) is

described. This new improvement enables the evaluation of the transfer function using a numerical model; this allows extension of the geostatistical approach to any case without using rough simplifications as the use of a 1-D or 2-D homogeneous model.

The second application of the geostatistical approach presented in this work is the **estimation of the hydraulic parameters**. Starting from field measurements, for instance of the transmissivity, it is possible using an interpolator to evaluate the whole transmissivity field of the study area. However, these kind of measurements are expensive and with few monitoring points the resulting transmissivity field is not reliable. Therefore, head measurements are frequently used, because they are easier and cheapless expensive to evaluate. So in the last 30 years several methods [for instance *Kitanidis and Vomovoris,* 1983; *Rubin and Dagan*, 1987a; *Giudici et al.,* 1988; etc] regarding parameter identification were developed. In this work the quasi-linear methodology is applied to parameter estimation following the work presented by *Kitanidis* [1995]. This method is a very efficient procedure, but for strongly nonlinear cases it requires some add ons. It is based on heads measured in specific points of the study area, then a forward problem is performed with an initial value of transmissivity. The following step is to correct the initial transmissivity field as far as the forward problem represents correctly the heads measured. The first proposed improvement is the updating of unknowns from an iteration to the next one. This procedure allows to choose the correct parameter in the Gauss-Newton iterations and to speed up the process. Then the theory regarding the conditional realizations, proposed in *Kitanidis* [1995], is tested to reproduce the highly nonlinear transmissivity field. Moreover considering the possibility to apply the estimation of hydraulic parameters to a very well defined grid it has been decided to summarize and test the methodology proposed by *Nowak et al.* [2003] that allows to speed up the matrix multiplication in order to decrease the computation time.

The work is structured in two part; the first presents a general introduction on inverse problems and describes the quasi-linear geostatistical approach. The second one proposes few improvements to the methodology and analyzes several cases.

Chapter 1

Fundamentals of Hydraulics of Groundwater

The basic governing equations[1] (for fully saturated soil) of flow and transport in groundwater are summarized in this section to better clarify the following chapters.

1.1 Flow

Darcy's experimental results concluded that the rate of flow Q [L^3T^{-1}] is proportional to a cross-sectional area A [L^2] and to the difference of head $(h_1 - h_2)$ [L] and inversely proportional to the length L [L]:

$$Q = \frac{K \cdot A(h_1 - h_2)}{L} \tag{1.1}$$

where K [LT^{-1}] is the hydraulic conductivity that represents the ability of the aquifer material to transport water under hydraulic gradients. Considering infinitesimal length the relationship $\frac{(h_1 - h_2)}{L}$ tends to $i = -\frac{dh}{dx}$ that represents the hydraulic gradient (the minus sign means that the flow goes from high head to low head). So equation (1.1) can be written as $Q = K \cdot i \cdot A$ and $U = Q/A = K \cdot i$ represents the velocity of the fluid. This expression is valid for one dimensional motion; considering three dimensional flux, the velocity is represented by $\mathbf{U} = -\mathbf{K} \cdot \nabla h$, so the Darcy's law becomes:

$$Q = -\mathbf{K} \cdot \nabla h \cdot A$$

where $\mathbf{K}$ is the symmetric tensor of the hydraulic conductivity

$$\mathbf{K} = \begin{bmatrix} k_{xx} & k_{xy} & k_{xz} \\ & k_{yy} & k_{yz} \\ & & k_{zz} \end{bmatrix}$$

[1] This section is based on *Bear* [1972], *Bear and Bachmat* [1990] and *Domenico and Schwartz* [1998]

The velocity along the three main Cartesian directions results:

$$\begin{aligned} U_x &= -k_{xx}\frac{\partial h}{\partial x} - k_{xy}\frac{\partial h}{\partial y} - k_{xz}\frac{\partial h}{\partial z} \\ U_y &= -k_{yx}\frac{\partial h}{\partial x} - k_{yy}\frac{\partial h}{\partial y} - k_{yz}\frac{\partial h}{\partial z} \\ U_z &= -k_{zx}\frac{\partial h}{\partial x} - k_{zy}\frac{\partial h}{\partial y} - k_{zz}\frac{\partial h}{\partial z} \end{aligned}$$

i.e.

$$\mathbf{U} = -\mathbf{K}\nabla h \tag{1.2}$$

The continuity equation (conservation of mass) for a porous medium saturated with water is the following:

$$div\,(\rho\mathbf{U}) + \frac{\partial n\rho}{\partial t} \pm \rho q = 0 \tag{1.3}$$

where $\mathbf{U}$ [LT^{-1}] represents the Darcy's velocity, ρ [ML^{-3}] is the density of the fluid, n [-] is the porosity and q [T^{-1}] is a term of source/sink per unit of volume. Substituting the Darcy velocity in the equation (1.3) it results:

$$div\,[\rho\,(-\mathbf{K}\nabla h)] + \frac{\partial n\rho}{\partial t} \pm \rho q = 0$$

that is the diffusion equation. Considering that the density ρ is constant in the space but variable in time with few approximation

$$\frac{1}{\rho}\frac{\partial n\rho}{\partial t} \simeq S_s\frac{\partial h}{\partial t}$$

where S_s [L^{-1}] represents the specific storativity. Equation (1.3) can be rewritten as:

$$div\,(\mathbf{K}\nabla h) = S_s\frac{\partial h}{\partial t} \pm q$$

In steady state condition $\dfrac{\partial h}{\partial t} = 0$ and without source or sinks the equation (1.3) collapses to $div\,(\mathbf{K}\nabla h) = 0$. Expanding this equation it results:

$$\frac{\partial}{\partial x}\left(K_{xx}\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_{yy}\frac{\partial h}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_{zz}\frac{\partial h}{\partial z}\right) = 0$$

in a homogeneous domain it is reduced to

$$K_{xx}\frac{\partial^2 h}{\partial x^2} + K_{yy}\frac{\partial^2 h}{\partial y^2} + K_{zz}\frac{\partial^2 h}{\partial z^2} = 0 \tag{1.4}$$

In isotropic medium $K_{xx} = K_{yy} = K_{zz}$ so the equation (1.4) becomes

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} + \frac{\partial^2 h}{\partial z^2} = 0$$

that is the Laplace's equation: $\nabla^2 h = 0$. If the flow can be considered horizontal the transmissivity (T [$\mathrm{L}^2\mathrm{T}^{-1}$]) indicates the ability of the aquifer to transmit water through its entire thickness, and

it is equal to the integral (over a vertical line) of the product of the hydraulic conductivity and the aquifer's thickness dz:

$$T = \int_0^b k_{xx} dz \tag{1.5}$$

The aquifer's storativity S [-] is defined as the volume of water released from (or added to) the aquifer per unit horizontal area of aquifer and per unit decline (or rise) of the average (over the vertical) piezometric head in the aquifer.

$$S = \int_0^b S_s dz \tag{1.6}$$

1.2 Transport of Contaminants

The partial differential equation describing the fate and transport of contaminants of species r in three dimensional, transient groundwater systems can be written as follows:

$$\frac{\partial (nC^r)}{\partial t} = \frac{\partial}{\partial x_i}\left(nD_{ij}\frac{\partial C^r}{\partial x_j}\right) - \frac{\partial}{\partial x_i}(nv_i C^r) + q_s C_s^r + \sum R_n \tag{1.7}$$

where C^r [ML^{-3}] represents the dissolved concentration of species r; t [T] is time; x_i [L] is the distance along the respective Cartesian coordinate axis; D_{ij} [$\mathrm{L}^2\mathrm{T}^{-1}$] is the hydrodynamic dispersion coefficient tensor; v_i [LT^{-1}] is the effective velocity (it is related to the Darcy velocity through the relationship $v_i = U_i/n$); q_s [T^{-1}] is the volumetric flow rate per unit volume of aquifer representing fluid sources and sinks; C_s^r [ML^{-3}] is the concentration of the source or sink flux for species r, and $\sum R_n$ [ML^{-3}] represents the reaction term. $\frac{\partial}{\partial x_i}(nv_i C^r)$ represents the advection term and describes the transport of miscible contaminants at the same velocity as the groundwater.

For many field-scale contaminant transport problems, the advection term dominates over other terms. To measure the degree of advection dominance, a dimensionless Peclet number is usually used; it is defined as

$$P_e = \frac{|\mathbf{v}|\, L}{D} \tag{1.8}$$

where $|\mathbf{v}|$ [LT^{-1}] is the magnitude of the effective velocity vector, L [L] is the characteristic length and D [$\mathrm{L}^2\mathrm{T}^{-1}$] is dispersion coefficient. For pure advection problems, the Peclet number approaches infinity.

Dispersion in porous media refers to the spreading of contaminants over a region greater than the one predicted from the advection process. Dispersion is caused by mechanical dispersion, a result of deviations of actual velocity on a microscale from the average groundwater velocity and by molecular diffusion driven by concentration gradients. Molecular diffusion is generally secondary and negligible, compared with the effects of mechanical dispersion, and only becomes important when groundwater velocity is very low. The sum of mechanical dispersion and molecular diffusion is termed hydrodynamic dispersion. The hydrodynamic dispersion tensor (D_{ij}) for a porous medium,

is defined [*Scheidegger*, 1961] in the following component form:

$$\begin{aligned}
D_{xx} &= \alpha_L \frac{v_x^2}{|v|} + \alpha_T \frac{v_y^2}{|v|} + \alpha_T \frac{v_z^2}{|v|} + D^* \\
D_{xy} &= (\alpha_L - \alpha_T) \frac{v_x v_y}{|v|} + D^* = D_{yx} \\
D_{xz} &= (\alpha_L - \alpha_T) \frac{v_x v_z}{|v|} + D^* = D_{zx} \\
D_{yy} &= \alpha_L \frac{v_y^2}{|v|} + \alpha_T \frac{v_x^2}{|v|} + \alpha_T \frac{v_z^2}{|v|} + D^* \\
D_{yz} &= (\alpha_L - \alpha_T) \frac{v_y v_z}{|v|} + D^* = D_{zy} \\
D_{zz} &= \alpha_L \frac{v_z^2}{|v|} + \alpha_T \frac{v_x^2}{|v|} + \alpha_T \frac{v_y^2}{|v|} + D^*
\end{aligned} \tag{1.9}$$

where D_{xx}, D_{yy}, D_{zz}, [L^2T^{-1}] are the principal component of the dispersion tensor; α_L [L] is the longitudinal dispersivity; α_T [L] is the transverse dispersivity; D^* [L^2T^{-1}] represents the effective molecular diffusion coefficient; v_x, v_y, v_z [LT^{-1}] are the components of the effective velocity vector v, $|v| = \sqrt{v_x^2 + v_y^2 + v_z^2}$.

The fluid sink/source term of the governing equation, $q_s C_s$, represents solute mass entering the model domain through sources or leaving the model domain through sinks. Sinks or sources may be classified as areally distributed sinks or sources or as point sinks or sources. The areally distributed sinks or sources include recharge and evapotranspiration. The point sinks or sources include wells, drains, and rivers.

The reactive terms, can be described, in general, as:

$$\sum R_n = -\rho_b \frac{\partial \bar{C}^k}{\partial t} - \lambda_1 n C^k - \lambda_2 \rho_b \bar{C}^k$$

where ρ_b [ML^{-1}] is the bulk density of the subsurface medium; $\bar{C}^k$ [MM^{-1}] is the concentration of species k sorbed on the subsurfaces solid; λ_1 [T^{-1}] is the first order reaction rate for the dissolved phase; and λ_2 [T^{-1}] is the first order reaction rate for the sorbed (solid) phase. In the applications of the present thesis only non reactive pollutants have been considered such that the term $\sum R_n = 0$.

Part I

Inverse Problems - General Overview

Chapter 2

Introduction

The literature on inverse and ill-posed problem has grown substantially over the last 50 years in many branches of science and mathematics. Typically ill-posed problems include numerical differentiation on noisy data, the inverse heat conduction problem, interpretation of geophysical data and the inverse problems of groundwater hydrology, i.e. the recovery of the release history of pollutants and the evaluations of the hydraulic parameters of aquifers.

A well-posed mathematical-physical problem, according to *Hadamard* [1902], must satisfy three requirements:

- Existence: there exists a function which satisfies governing equations and subsidiary conditions.
- Uniqueness: there is only one solution of the problem.
- Stability: the variation of solution can be arbitrarily small, provided the variations of input data are sufficiently small.

If anyone of these requirements is not satisfied, the problem is ill-posed.

Analyzing Figure 2.1 it is possible to better clarify the concept of ill-posed problem. A well-posed problem is defined by a cause **C** and a model **M** (see Figure 2.1). The model **M** is assumed continuous and well-defined, so for each cause **C** only one effect **E** is obtained, and for small perturbation of **C** there are small perturbations in the effect **E**. So the problem is well-posed, as meeting the conditions mentioned above, honours the conditions of existence, uniqueness and stability. Otherwise the problems that do not respect these condition are ill-posed.

A particular ill-posed problem is the inverse one problem. For instance, the determination of the cause **C** as a function of the effect **E** and the model **M**, or the determination of the model **M** as a function of **C** and **E**. Intuitively, *existence* of an inverse solution appears to be no problem

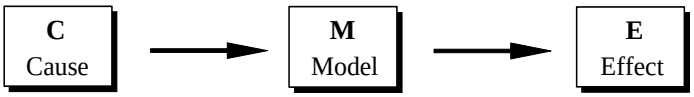

Figure 2.1: Sketch of problem

at all, since the physical reality must be the solution. In practice the observation error of the variables cannot be avoided. As a result, an accurate solution of an inverse problem may not exist. Different combination of hydrogeologic conditions may lead to similar observations of water level and solute concentration. It is thus impossible to uniquely determine the particularities of an aquifer by only observing the variables, i.e. the *non-uniqueness* of inverse solutions is often encountered. A solution, although it is existent and unique, can not be accepted, if it does not continuously depend upon the input data. It is known that forward solutions in groundwater modeling are always stable, (for example when hydraulic parameters and/or boundary conditions change slightly, the water level should be slightly affected). Unfortunately, inverse solutions in groundwater modeling are often *unstable*.

During the last 50 years several methods were developed to solve inverse problems. *Sun* [1994] summarized the methods to solve the inverse problems that deals with the groundwater in four categories: indirect methods; direct methods; adjoint state methods and stochastic methods.

As example of the indirect method it is relevant to list the Gauss-Newton optimizations algorithm, the optimization method which it is specially designed for minimizing the objective function which has the form of the sum square functions with regularization. In many cases there are difficulties associated with instability and nonuniqueness so several methods were developed to find the "best possible" solution of the problem. The strategy of these methods is to avoid solving the ill-posed problem directly and to solve a related well-posed problem whose solution is closed to the solution of the real problem. Solutions found by these methods have been termed regularized solutions or quasi solutions; an example of this is Tikhonov regularization.

The direct methods are not widely used, they are based on a general form of equation error criteria. The inverse solution is obtained by solving a system of superdeterministic equations. The ill-posedness of inverse problems in this case it is manifested by the ill-conditioning of coefficient matrix. Data processing, adding constraints and parameterization are always necessary (using the direct methods) for improving the stability of inverse solutions.

The adjoint state method is based on the variational theory; its applications include not only parameter identification, but also sensitivity analysis, reliability estimates and observation design. With the adjoint state method, the forward problem, for instance, the concentration as the dependent variable, is replaced by the adjoint equation, with the adjoint state as the dependent variable. The adjoint state is a function that describes the marginal change in the performance measure due to, for instance, a unit injection of mass. The application of the adjoint state method to groundwater modeling has been extensive in recent years [*Sun and Yeh*, 1990a, 1990b; *Sun*, 1994; *Neuaper and Wilson,* 1999; *Michalak and Kitanidis*, 2004].

The stochastic method is another important field of solutions of inverse problems. To solve inverse problems it is necessary to know prior information, field measurements and a conceptual model relating the observation to the unknown parameters. There are three main advantage of using the stochastic methods: first the prior information is easy to be incorporated into the statement of the inverse problem; second the unknown parameters and their uncertainty can be estimated at the same time; third the over-parameterization problem may be avoided, because unknown parameters are regarded as a stochastic fields and described by only a few statistical parameters.

2.1 Example

A problem is well-posed if it is solvable in a unique way and its solution depends with continuity from the data. For instance the integral of a continuous function is a well-posed problem because

it has a unique solution, while the derivative of a function is a ill-posed problem. Derivatives are often encountered in inverse problems (the derivative is the inverse of the integral).

Considering the linear problem:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{y} \tag{2.1}$$

if $\mathbf{A}$ is a square matrix and $det\mathbf{A} \neq 0$, equation (2.1) is a well-posed problem and its solution is $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{y}$. Otherwise if $\mathbf{y} \notin \mathrm{Im}(\mathbf{A})$, equation (2.1) is not solvable so an approximate solution $\mathbf{x}_0$, that allows: $\|\mathbf{A}\mathbf{x}_0 - \mathbf{y}\| = \min\{\|\mathbf{A}\mathbf{x} - \mathbf{y}\|\}$, can be chosen as a regularized solution.

Different methods to solve the problem described by the equation (2.1) are present in literature, in the following two methods are shown. The first is the Tikhonov regularization [*Tikhonov and Arsenin*, 1977] briefly described and the second is the statistical method (that is described in the following chapter and it is applied to several cases).

2.2 Tikhonov Regularization

The objective is to solve the equation (2.1); if $\ker \mathbf{A} \neq \mathbf{0}$ there are infinite solutions, and the method looks for the one with minimum norm. Tikhonov regularization [*Tikhonov and Arsenin*, 1977] replaces the ill-posed problem with a well-posed minimization problem. The solution of the equation (2.1) is reachable using the Lagrange multipliers, minimizing the following:

$$\|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2 + \alpha \|\mathbf{x}\|^2 \tag{2.2}$$

where α is the Lagrange multiplier and represents the Tikhonov's regularization parameter. The solution $\mathbf{x}$ is considered an approximation of the true solution, because the problem in equation (2.1) has only one solution. The solution of the equation (2.1) is an $\mathbf{x}_0$ vector that satisfies $\mathbf{y} = \mathbf{A} \cdot \mathbf{x}_0 + \boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is the error. Using Tikhonov's regularization method the solution is

$$\mathbf{x}_\alpha = \left(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I}\right)^{-1} \mathbf{A}^T\mathbf{y} \tag{2.3}$$

The main problem of the method is the choice of the regularization parameter α that has to be not too small or too large; several methodologies were developed with the aim of determine an approximate α.

Applying the Tikhonov regularization to recover the release history of pollutants in groundwater *Skaggs and Kabala* [1994] proposed the following methodology. The transport process in groundwater can be represented by an integral [see *Jury and Roth*, 1990]. So given the linear Fredholm equation of the first kind

$$y(\omega) = \int_a^b K(\lambda, \omega) s(\lambda) d\lambda \tag{2.4}$$

where $K(\lambda, \omega)$ and $y(\omega)$ are known and $s(\lambda)$ is the unknown function. Solving the (2.4) for s is an ill-posed problem, and an infinite set of functions s^* that satisfy the following (2.5) exists:

$$\left\| y(\omega) - \int_a^b K(\lambda, \omega) s^*(\lambda) d\lambda \right\| \leq \varepsilon \tag{2.5}$$

where ε represents the noise level data. The Tikhonov regularization [*Tikhonov and Arsenin*, 1977] replaces the ill-posed problem with a well-posed minimization problem. The function to be

minimized is:

$$V(\alpha) = \left\| y(\omega) - \int_a^b K(\lambda,\omega)s(\lambda)d\lambda \right\|^2 + \alpha^2 \|Ls\|^2 \tag{2.6}$$

where L is the regularization operator and α is the regularization parameter which determines the relative weight of L. The most common structure of regularization operator is

$$\|Ls\|^2 = \int_a^b \left(\frac{d^n s}{d\lambda^n}\right)^2 d\lambda \tag{2.7}$$

where $\frac{d^n s}{d\lambda^n}$ is the nth derivative of s with the zeroth derivative being equal to s. The accuracy of the solution depends on finding a good value for α as mentioned above several method were proposed to find an optimal value of α. *Skaggs and Kabala* [1994] used the method proposed by *Provencher* [1982] and *Obenchain* [1977].

Let's define:

$$F = \frac{(V(\alpha) - V(0))(N_y - N_0)}{V(0)N_0} \tag{2.8}$$

where $V(0)$ is the residual computed from the equation (2.6), with α close to zero, N_0 is the number of degrees of freedom associated with $V(0)$, N_y is the number of observations of y and $V(\alpha)$ is the residual obtained from the equation (2.6) for a particular value of α. The ratio in the equation (2.8) can be evaluated using the Fisher distribution (F), $P(F, v_1, v_2)$, where v_1 and v_2 are the number of degrees of freedom associated with the numerator and denominator, respectively, of F.

A good choice of α is a value that results in equation (2.8) being close to $F_{0,5}$, where $F_{0,5}$ is the value of F such that half of the area under the $Fisher$ distribution lies to the left of $F_{0,5}$ and half to the right. *Provencher* [1982] found that a solution obtained with this criterion can be interpreted as being a parsimonious solution: the simplest solution that fits the data. This solution may not contain all the details of the true solution but all the components that are required by the data and exclude numerical artifacts.

2.3 The Statistical Method

Starting from the problem described by the equation (2.1) and considering a measurement $\mathbf{y}$ affected by a noise $\boldsymbol{\eta}$ the problem becomes:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{y} = \mathbf{y_0} + \boldsymbol{\eta} \tag{2.9}$$

assuming that the noise is characterized by a Gaussian distribution with zero mean and covariance known, $\mathbf{x}$ is also a random variable; the two variables $\boldsymbol{\eta}$ and $\mathbf{x}$ are independent so that $Cov(\mathbf{x}, \boldsymbol{\eta}) = 0$. Three variables are involved: $\mathbf{x}$, $\mathbf{y}$, and $\boldsymbol{\eta}$. Let's assume to solve the problem by means of a linear estimation of $\mathbf{x}$:

$$\mathbf{x} = \mathbf{L}\mathbf{y} = \mathbf{L}(\mathbf{y}_0 + \boldsymbol{\eta})$$

and the best choice of $\mathbf{L}$ is the one that minimizes the following equation:

$$J(\mathbf{L}) = tr\left\{Cov(\mathbf{L}\mathbf{y} - \mathbf{x})\right\} = tr\left\{E\left[(\mathbf{L}\mathbf{y} - \mathbf{x})^T(\mathbf{L}\mathbf{y} - \mathbf{x})\right]\right\} \tag{2.10}$$

$\mathbf{x}$ and its realizations are unknown, and also $\mathbf{y}$ is unknown but the minimization problem with respect to $\mathbf{L}$ is solvable. $\mathbf{L}\mathbf{y}_0 = \mathbf{L}\mathbf{A}\mathbf{x} - \mathbf{L}\boldsymbol{\eta}$, so the minimization problem becomes:

$$\begin{aligned} & (\mathbf{Ly}-\mathbf{x})^T(\mathbf{Ly}-\mathbf{x}) \\ = \; & (\mathbf{LAx}-\mathbf{L}\boldsymbol{\eta}-\mathbf{x})^T(\mathbf{LAx}-\mathbf{L}\boldsymbol{\eta}-\mathbf{x}) = \\ = \; & tr\left\{(\mathbf{LAx}-\mathbf{L}\boldsymbol{\eta}-\mathbf{x})(\mathbf{LAx}-\mathbf{L}\boldsymbol{\eta}-\mathbf{x})^T\right\} = \\ = \; & tr\left\{\mathbf{L}\left[\mathbf{Axx}^T\mathbf{A}^T+\boldsymbol{\eta}\boldsymbol{\eta}^T\right]\mathbf{L}^T-\mathbf{L}(\mathbf{Ax}-\boldsymbol{\eta})\mathbf{x}^T-\mathbf{x}\left(\mathbf{x}^T\mathbf{A}^T-\boldsymbol{\eta}^T)\mathbf{L}^T+\mathbf{xx}^T\right)\right\} \end{aligned} \quad (2.11)$$

Integrating the (2.11) with respect to $\mathbf{x}$ and $\boldsymbol{\eta}$ and considering that $\mathbf{x}$ and $\boldsymbol{\eta}$ are independent the (2.10) becomes:

$$J(\mathbf{L}) = tr\left\{\mathbf{L}(\mathbf{ARA}^T+\mathbf{V})\mathbf{L}^T-\mathbf{LAR}-\mathbf{RA}^T\mathbf{L}^T+\mathbf{R}\right\}$$

where $\mathbf{R}=\mathbf{xx}^T$ and $\mathbf{V}=\boldsymbol{\eta}\boldsymbol{\eta}^T$, $(\mathbf{V}>0)$. In order to minimize J it is sufficient to obtain the minimization of

$$tr\left\{\mathbf{LPL}^T-\mathbf{LAR}-\mathbf{RA}^T\mathbf{L}^T\right\} \quad (2.12)$$

where $\mathbf{P}=\mathbf{ARA}^T+\mathbf{V}>0$ solving

$$tr\left\{\left[\mathbf{LP}^{1/2}-\mathbf{RA}^T\mathbf{P}^{-1/2}\right]\left[\mathbf{LP}^{1/2}-\mathbf{RA}^T\mathbf{P}^{-1/2}\right]^T-\mathbf{RAA}^T\mathbf{R}^T\right\}$$

so it is enough to minimize

$$tr\left\{\left[\mathbf{LP}^{1/2}-\mathbf{RA}^T\mathbf{P}^{-1/2}\right]\left[\mathbf{LP}^{1/2}-\mathbf{RA}^T\mathbf{P}^{-1/2}\right]^T\right\}=\left\|\mathbf{LP}^{1/2}-\mathbf{RA}^T\mathbf{P}^{-1/2}\right\|^2$$

the minimum is reached for $\mathbf{L}_0=\mathbf{RA}^T\mathbf{P}^{-1}$ so the best estimate of $\mathbf{x}$ is

$$\hat{\mathbf{x}}=\mathbf{L}_0\mathbf{y}=\mathbf{RA}^T\mathbf{P}^{-1}\mathbf{y}=\mathbf{RA}^T\left(\mathbf{ARA}^T+\mathbf{V}\right)^{-1}\mathbf{y} \quad (2.13)$$

In many cases the elements η_i of the noise $\boldsymbol{\eta}$ are independent so that its covariance is

$$cov(\boldsymbol{\eta}) = diag\left[v_1, v_2, ..., v_n\right], \text{ with } v_i > 0$$

Frequently the statistics of the errors are constant, so $v_i=\alpha_i$ with $i=1,...,m$, in this case the $\boldsymbol{\eta}$ is called white noise and the estimate becomes:

$$\hat{\mathbf{x}}=\mathbf{RA}^T\left(\mathbf{ARA}^T+\alpha\mathbf{I}\right)^{-1}\mathbf{y}=\left(\mathbf{A}^T\mathbf{AR}+\alpha\mathbf{I}\right)^{-1}\mathbf{RA}^T\mathbf{y} \quad (2.14)$$

and if $\mathbf{R}=\mathbf{I}$ the equation (2.14) becomes: $\hat{\mathbf{x}}=\left(\mathbf{A}^T\mathbf{A}+\alpha\mathbf{I}\right)^{-1}\mathbf{A}^T\mathbf{y}$ that is the same form of the Tikhonov's regularization. That means that the method proposed by Tikhonov is a particular case of the statistical one in practice the Tikhonov regularization gives the same weight to all the information.

The statistical method proposes another way to evaluate the parameter α. It is possible to apply this method when the error of the estimate satisfies the following equations:

$$\begin{aligned} E(\boldsymbol{\eta}) &= 0 \\ E(\boldsymbol{\eta}\boldsymbol{\eta}^T) &= Cov(\boldsymbol{\eta}) = \sigma^2\mathbf{I} \end{aligned} \quad (2.15)$$

where $\mathbf{I}$ is a $n \times n$ matrix. The second line of (2.15) implies that $tr\left[E\left(\boldsymbol{\eta}\boldsymbol{\eta}^T\right)\right] = n\sigma^2$ that represents the variance of the error. Considering $\delta = \sqrt{n}\sigma$ and $E\left(\left\|\mathbf{x}_{\alpha,\delta} - \mathbf{x}_\alpha\right\|^2\right)$ with

$$\mathbf{x}_\alpha = \left(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I}\right)^{-1}\mathbf{A}^T\mathbf{y}, \mathbf{x}_{\alpha,\delta} = \left(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I}\right)^{-1}\mathbf{A}^T(\mathbf{y} + \boldsymbol{\eta})$$

So $E\left(\left\|\mathbf{x}_{\alpha,\delta} - \mathbf{x}_\alpha\right\|^2\right) = n\sigma^2 tr\left[\mathbf{A}\left(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I}\right)^{-1}\left(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I}\right)^{-1}\mathbf{A}\right]$. α is chosen in order to minimize the residual above the difference between $\mathbf{x}_{\alpha,\delta}$ and $\mathbf{x}_\alpha$:

$$\frac{\left\|(\mathbf{y} + \boldsymbol{\eta}) - \mathbf{A}\mathbf{x}_{\alpha,\delta}\right\|^2}{\left\|tr\left[\mathbf{A}\left(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I}\right)^{-1}\left(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I}\right)^{-1}\mathbf{A}\right]\right\|}$$

Chapter 3

Geostatistical Method

3.1 Introduction

The history of the geostatistical approach to the inverse problem starts with the basic works of *Matheron* [1963, 1971, 1973] and then the methodology was developed by several authors as for instance: *Kitanidis and Vomvoris* [1983], *Dagan* [1985], *Hoeksema and Kitanidis* [1984, 1985, 1989], *Rubin and Dagan* [1987a, 1987b] and *Wagner and Gorelick* [1989].

This work is along the line of the work proposed by *Kitanidis and Vomovoris* [1983] and *Kitanidis* [1995]; a list of the main papers relating to this methodology, is presented in the following. *Kitanidis and Vomvoris* [1983] proposed the geostatistical inverse methodology to estimate hydrogeological parameters and applied it to a one dimensional case. *Hoeksema and Kitanidis* [1984] extended the methodology to a 2-D aquifer. Then in 1989 the same authors applied the methodology to the prediction of the transmissivities, heads and velocity [*Hoeksema and Kitanidis,* 1989]. *Kitanidis* [1995] improved the quasi-linear geostatistical theory; in the paper the author explained and applied the method to a nonlinear case. In the last ten years several enhancements were proposed. *Snodgrass and Kitanidis* [1997] applied the methodology to the recovery of pollutants history in 1-D case, and *Michalak and Kitanidis* [2002, 2003, 2004a] made several improvements to this procedure. In 2002 they applied the methodology to a real case [*Michalak and Kitanidis,* 2002] (this is the first application of the geostatistical approach to contaminant source identification in a multi dimensional domain, in a non-point source, and with both spatially and temporally distributed data). In the meanwhile *Butera and Tanda* [2001] applied the method to a 2-D aquifer, then in 2002 and 2003 *Butera and Tanda* [2002, 2003] studied the case of an aquifer with multiple point sources and with areal sources. In the work of *Michalak and Kitanidis* [2003] the authors proposed a new method to enforce parameter non negativity (to avoid negative concentration data); in 2004 [*Michalak and Kitanidis*] used the adjoint state method applied to the geostatistical procedure to identify the source release, this work allows to use the methodology with a multidimensional heterogeneous media. In the work of *Butera and Tanda* [2004] the authors considered the case of temporally distributed data and the possibility to apply the methodology to a heterogeneous aquifer, considering the theory of the stochastic transport of pollutants. Further contributions to the methodology come from *Kitanidis and Lane* [1985] who applied the Gauss-Newton method for maximum likelihood. *Kitanidis* [1996] made a comparison between the geostatistical approach and the maximum a posteriori probability method. *Kitanidis and Shen* [1996] proposed a methodology for the estimate of solute concentration contour maps and volume averages and presented a

method for the optimization of a parameter (see the following chapter) that constrains concentrations to positive values. *Kitanidis* [1997] presented a method to obtain a stable and a reasonable estimate with the minimum possible suppositions about the unknown function or its structure. In the work of *Kitanidis* [1999] the author examined which generalized covariance function produces the flattest possible estimate of an unknown function that is consistent with the data. *Nowak et al.* [2003] proposed a spectral method regarding the matrix multiplications to speed up the procedure. *Nowak and Cirpka* [2004] developed a modified Levenberg-Marquardt algorithm for the geostatistical approach in order to increase the stability in cases with strong nonlinearity.

3.2 Quasi-Linear Geostatistical Approach

This section gives a brief description of the quasi-linear geostatistical approach applied to the recovery of the conductivity field, a detailed one can be find in the work of *Kitanidis* [1995]. The approach consists of two steps:

the first step (**structural analysis**) in geostatistics is to characterize the random field by functions of a few parameters that describe the statistic field. $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are the structural parameters of the process, $\mathbf{s}$, that is a spatially variable parameter (such as the log conductivity). $\mathbf{s}$ is the $n \times 1$ vector of discretized values with expected value $E[\mathbf{s}] = \mathbf{X}\boldsymbol{\beta}$ where $\mathbf{X}$ is a known $n \times p$ matrix, $\boldsymbol{\beta}$ are p unknown coefficients and covariance matrix: $E\left[(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})^T\right] = \mathbf{Q}(\boldsymbol{\theta})$. The covariance matrix $\mathbf{Q}$ is considered a known function of the parameters $\boldsymbol{\theta}$.

The observations $\mathbf{z}$ ($m \times 1$) are related to the unknown spatial process and the other parameters through $\mathbf{z} = \mathbf{h}(\mathbf{s}, \mathbf{r}) + \mathbf{v}$, where $\mathbf{r}$ are other unknown parameters (such as boundary conditions, sources and sinks, ...) and $\mathbf{v}$ is the observation error with normal distribution, zero mean, and covariance matrix $\mathbf{R} = E\left[\mathbf{v}\mathbf{v}^T\right]$, that is fixed or a known function (assuming that the errors are taken to be independent and identically distributed with variance σ_R^2, the covariance matrix becomes $\mathbf{R} = \sigma_R^2 \cdot \mathbf{I}$, where $\mathbf{I}$ is the identity matrix). The standard deviations of the measurements errors (which are the square root of the diagonal elements of $\mathbf{R}$) define how closely the observations should be reproduced. The vector $\mathbf{z}$ is a random vector being function of $\mathbf{s}$ and $\mathbf{v}$, that are random vectors. The probability density function (pdf) of $\mathbf{z}$ depends on the distribution of $\mathbf{s}$ and $\mathbf{v}$ and also on the function $\mathbf{h}$ and it is generally complex to evaluate.

From the Bayes theorem the posterior pdf of a state vector $\mathbf{s}$ given an observation vector $\mathbf{z}$ is proportional to the likelihood of the data given the state, times the prior pdf of the state:

$$p''(\mathbf{s}) = \frac{p(\mathbf{z}|\mathbf{s})p'(\mathbf{s})}{p(\mathbf{z})} \tag{3.1}$$

where $p(\mathbf{z}) = \int_s p(\mathbf{z}|\mathbf{s})p'(\mathbf{s})ds = \int_s p(\mathbf{z}, \mathbf{s})ds$, the prior, $p'(\mathbf{s})$ = probability of the events (that represents the assumed structure of the unknown function) is Gaussian and can be defined by the following equations:

$$p'(\mathbf{s}|\boldsymbol{\beta}, \boldsymbol{\theta}, \mathbf{r}) \propto \exp\left(-\frac{1}{2}(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})^T \mathbf{Q}^{-1}(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})\right) \tag{3.2}$$

The likelihood function is also a Gaussian function represented by:

$$p(\mathbf{z}|\mathbf{s}, \boldsymbol{\beta}, \boldsymbol{\theta}, \mathbf{r}) \propto \exp\left(-\frac{1}{2}(\mathbf{z} - \mathbf{h}(\mathbf{s}, \mathbf{r}))^T \mathbf{R}^{-1}(\mathbf{z} - \mathbf{h}(\mathbf{s}, \mathbf{r}))\right) \tag{3.3}$$

so the posterior is

$$p''(\mathbf{s}|\boldsymbol{\beta},\boldsymbol{\theta},\mathbf{r}) \propto \exp\left(-\frac{1}{2}\left(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\right)^T\mathbf{Q}^{-1}\left(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\right)-\frac{1}{2}\left(\mathbf{z}-\mathbf{h}(\mathbf{s},\mathbf{r})\right)^T\mathbf{R}^{-1}\left(\mathbf{z}-\mathbf{h}(\mathbf{s},\mathbf{r})\right)\right) \tag{3.4}$$

which is Gaussian.

$$\begin{aligned} p(\mathbf{z},\mathbf{s}|\boldsymbol{\beta},\boldsymbol{\theta},\mathbf{r}) &\propto \exp\left(-\frac{1}{2}\left(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\right)^T\mathbf{Q}^{-1}\left(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\right)-\frac{1}{2}\left(\mathbf{z}-\mathbf{h}(\mathbf{s},\mathbf{r})\right)^T\mathbf{R}^{-1}\left(\mathbf{z}-\mathbf{h}(\mathbf{s},\mathbf{r})\right)\right) \\ p(\mathbf{z}|\boldsymbol{\beta},\boldsymbol{\theta},\mathbf{r}) &\propto \int_s \exp\left(-\frac{1}{2}\left(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\right)^T\mathbf{Q}^{-1}\left(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\right)-\frac{1}{2}\left(\mathbf{z}-\mathbf{h}(\mathbf{s},\mathbf{r})\right)^T\mathbf{R}^{-1}\left(\mathbf{z}-\mathbf{h}(\mathbf{s},\mathbf{r})\right)\right) ds \end{aligned} \tag{3.5}$$

The elimination of $\boldsymbol{\beta}$ can be achieved by working with the restricted likelihood, obtained through averaging over all values of the drift coefficient $\boldsymbol{\beta}$, so 3.5 becomes:

$$\begin{aligned} p(\mathbf{z}|\boldsymbol{\theta},\mathbf{r}) &= \int_s p(\mathbf{z}|\boldsymbol{\beta},\boldsymbol{\theta},\mathbf{r})d\boldsymbol{\beta} \\ p(\mathbf{z}|\boldsymbol{\theta},\mathbf{r}) &\propto |\mathbf{R}|^{-1/2}|\mathbf{Q}|^{-1/2}\left|\mathbf{X}^T\mathbf{Q}^{-1}\mathbf{X}\right|^{-1/2}\cdot \\ &\quad \int_s \exp\left(\frac{1}{2}\left(\mathbf{z}-\mathbf{h}(\mathbf{s},\mathbf{r})\right)^T\mathbf{R}^{-1}\left(\mathbf{z}-\mathbf{h}(\mathbf{s},\mathbf{r})\right)+\frac{1}{2}\mathbf{s}^T\mathbf{G}\mathbf{s}\right) ds \end{aligned} \tag{3.6}$$

where

$$\mathbf{G} = \mathbf{Q}^{-1} - \mathbf{Q}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{Q}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Q}^{-1}$$

The following step is the application of the restricted maximum likelihood estimation, that consists of finding the values of the parameter $\boldsymbol{\theta}$ and $\mathbf{r}$ that maximize the expression $p(\mathbf{z}|\boldsymbol{\theta},\mathbf{r})$, in equation (3.6). Minimizing the log of the posterior pdf (3.6) is equal to maximizing the posterior probability but easier to handle, so that the negative log of the posterior pdf is

$$L = -\ln p''\left(\mathbf{s}|\boldsymbol{\beta},\boldsymbol{\theta},\mathbf{r}\right) \tag{3.7}$$

Assuming that $\mathbf{Q}$ and $\mathbf{R}$ are already been assigned, the observation equation is $\mathbf{z} = \mathbf{h}(\mathbf{s}) + \mathbf{v}$. The negative log-likelihood function (3.7) that is represented by the following

$$\frac{1}{2}\left(\mathbf{z}-\mathbf{h}(\mathbf{s},\mathbf{r})\right)^T\mathbf{R}^{-1}\left(\mathbf{z}-\mathbf{h}(\mathbf{s},\mathbf{r})\right)+\frac{1}{2}\left(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\right)^T\mathbf{Q}^{-1}\left(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\right) \tag{3.8}$$

must be minimized with respect to $\mathbf{s}$ and $\boldsymbol{\beta}$ so (3.8) becomes:

$$\frac{1}{2}\left(\mathbf{z}-\mathbf{h}(\mathbf{s},\mathbf{r})\right)^T\mathbf{R}^{-1}\left(\mathbf{z}-\mathbf{h}(\mathbf{s},\mathbf{r})\right)+\frac{1}{2}\mathbf{s}^T\mathbf{G}\mathbf{s} = \min \tag{3.9}$$

The appropriate iterative procedure is the Gauss-Newton method, obtained by linearizing $\mathbf{h}$ about the most recent estimate and solving the optimization problem starting with an estimate $\widetilde{\mathbf{s}}_0 = \overset{\wedge}{\mathbf{s}}_k$. The procedure starts finding the derivative of $\mathbf{h}$ about $\mathbf{s}$ at $\widetilde{\mathbf{s}}$: $\widetilde{\mathbf{H}} = \left.\dfrac{\partial \mathbf{h}}{\partial \mathbf{s}}\right|_{\mathbf{s}=\widetilde{\mathbf{s}}}$, this sensitivity matrix is found using the adjoint state method. Assuming that the actual $\overset{\wedge}{\mathbf{s}}$ is close to $\widetilde{\mathbf{s}}$, approximate $\mathbf{h}\left(\overset{\wedge}{\mathbf{s}}\right) = \mathbf{h}\left(\widetilde{\mathbf{s}}\right) + \widetilde{\mathbf{H}}\left(\overset{\wedge}{\mathbf{s}} - \widetilde{\mathbf{s}}\right)$ and defining:

$$\widetilde{\boldsymbol{\Sigma}} = \widetilde{\mathbf{H}}\mathbf{Q}\widetilde{\mathbf{H}}^T + \mathbf{R} \tag{3.10}$$

In the parameter estimation step, $\mathbf{z}_0 = \mathbf{z} - \mathbf{h}\left(\widetilde{\mathbf{s}}\right) + \widetilde{\mathbf{H}}\widetilde{\mathbf{s}}$ and $\mathbf{H}$ are treated as constant. Carrying out the integration and simplifying, the $p(\mathbf{z}|\boldsymbol{\theta}, \mathbf{r})$ can be written:

$$\begin{aligned} p(\mathbf{z}|\boldsymbol{\theta}, \mathbf{r}) \quad \propto \quad & |\boldsymbol{\Sigma}|^{-1/2} \left|\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Sigma}^{-1}\mathbf{H}\mathbf{X}\right|^{-1/2} \\ & \exp\left[-\frac{1}{2}\mathbf{z}_0^T\left(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{H}\mathbf{X}\left(\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Sigma}^{-1}\mathbf{H}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Sigma}^{-1}\right)\mathbf{z}_0\right] \end{aligned} \tag{3.11}$$

The problem of maximum likelihood estimation is equivalent to minimize the negative logarithm of the pdf

$$\begin{aligned} L \quad = \quad & \frac{1}{2}\ln|\boldsymbol{\Sigma}| + \frac{1}{2}\ln\left|\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Sigma}^{-1}\mathbf{H}\mathbf{X}\right| + \\ & + \frac{1}{2}\mathbf{z}_0^T\left(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{H}\mathbf{X}\left(\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Sigma}^{-1}\mathbf{H}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Sigma}^{-1}\right)\mathbf{z}_0 \end{aligned} \tag{3.12}$$

minimizing with respect to θ_i and considering

$$\boldsymbol{\Xi} = \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{H}\mathbf{X}\left(\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Sigma}^{-1}\mathbf{H}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Sigma}^{-1}$$

$$g_i = \frac{\partial L}{\partial\theta_i} = \frac{1}{2}Tr\left[\boldsymbol{\Xi}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right] - \frac{1}{2}\mathbf{z}_0^T\left(\boldsymbol{\Xi}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\boldsymbol{\Xi}\right)\mathbf{z}_0$$

Considering the Fisher information matrix $\mathbf{F}$,

$$\mathbf{F}_{ij} = \mathbf{E}\left[\frac{\partial^2 L}{\partial\theta_i\partial\theta_j}\right] = \frac{1}{2}Tr\left[\boldsymbol{\Xi}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\boldsymbol{\Xi}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\right] \tag{3.13}$$

it is possible to apply the Gauss-Newton method to solve the problem [see for more details *Kitanidis and Lane*, 1985].

$$\boldsymbol{\theta}_{l+1} = \boldsymbol{\theta}_l - \mathbf{F}^{-1}\mathbf{g} \tag{3.14}$$

For a complex case it may be helpful to add the Marquardt modification replacing $\mathbf{F}^{-1}$ with $(\mathbf{F}+\lambda\mathbf{I})^{-1}$, where λ is a positive parameter.

Once the iterations have converged it is possible to form and solve the following cokriging system (second step: **estimate of the spatial function**):

$$\begin{bmatrix} \widetilde{\boldsymbol{\Sigma}} & \widetilde{\mathbf{H}}\mathbf{X} \\ \left(\widetilde{\mathbf{H}}\mathbf{X}\right)^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Lambda}^T \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{H}}\mathbf{Q} \\ \mathbf{X}^T \end{bmatrix} \tag{3.15}$$

where $\boldsymbol{\Lambda}^T$ is an $m \times n$ matrix of coefficients and $\mathbf{M}$ is a $p \times n$ matrix of multipliers. Then the cokriging estimate ($\boldsymbol{\Lambda}$ form) is

$$\hat{\mathbf{s}} = \boldsymbol{\Lambda}\mathbf{z_0} \tag{3.16}$$

If $\mathbf{s}$ is practically equal to $\widetilde{\mathbf{s}}$, the algorithm has converged and the covariance matrix of estimation is:

$$\mathbf{V} = -\mathbf{X}\mathbf{M} + \mathbf{Q} - \mathbf{Q}\widetilde{\mathbf{H}}^T\boldsymbol{\Lambda}^T \tag{3.17}$$

otherwise $\tilde{\mathbf{s}}$ is set equal to $\hat{\mathbf{s}}$ and the procedure is repeated.

This procedure is analogous to the one used in kriging processes. Another approach ($\boldsymbol{\xi}$ form) is analogous to the procedure for function estimation. The solution is:

$$\hat{\mathbf{s}} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{Q}\tilde{\mathbf{H}}^T\boldsymbol{\xi} \tag{3.18}$$

where the $\hat{\boldsymbol{\beta}}$ and $\boldsymbol{\xi}$ coefficients are found by solving a single linear system of $m + p$ equations:

$$\begin{bmatrix} \tilde{\boldsymbol{\Sigma}} & \tilde{\mathbf{H}}\mathbf{X} \\ \left(\tilde{\mathbf{H}}\mathbf{X}\right)^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi} \\ \hat{\boldsymbol{\beta}} \end{bmatrix} = \begin{bmatrix} \mathbf{z} - \mathbf{h}\left(\tilde{\mathbf{s}}\right) + \tilde{\mathbf{H}}\tilde{\mathbf{s}} \\ \mathbf{0} \end{bmatrix} \tag{3.19}$$

When the iterative procedure has converged, the posterior covariance matrix can be computed, through the linearization approximation, as follows:

$$\begin{bmatrix} \mathbf{P}_{yy} & \mathbf{P}_{yb} \\ \mathbf{P}_{yb}^T & \mathbf{P}_{bb} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma} & \tilde{\mathbf{H}}\mathbf{X} \\ \left(\tilde{\mathbf{H}}\mathbf{X}\right)^T & \mathbf{0} \end{bmatrix}^{-1} \tag{3.20}$$

where $\mathbf{P}_{yy}$ is $(m \times m)$, $\mathbf{P}_{yb}$ is $(m \times p)$, and $\mathbf{P}_{bb}$ is $(p \times p)$. Then the covariance matrix of estimation is

$$\mathbf{V} = \mathbf{Q} - \mathbf{Q}\mathbf{H}^T\mathbf{P}_{yy}\mathbf{H}\mathbf{Q} - \mathbf{X}\mathbf{P}_{bb}\mathbf{X}^T - \mathbf{X}\mathbf{P}_{yb}^T\mathbf{H}\mathbf{Q} - \mathbf{Q}\mathbf{H}^T\mathbf{P}_{yb}\mathbf{X}^T$$

The methodology is summarized in the flow chart represented in Figure 3.1.

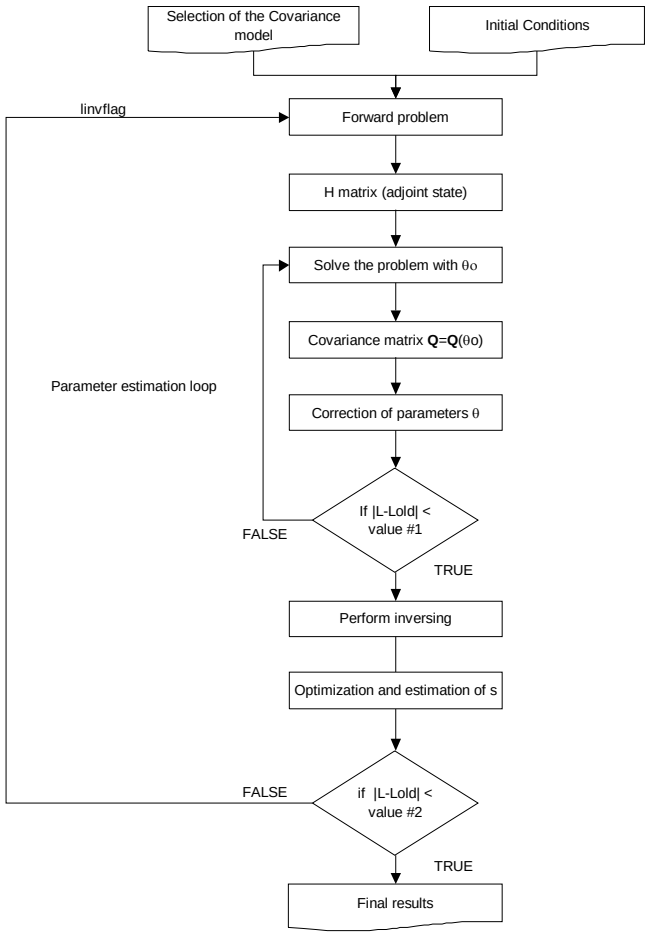

Figure 3.1: Flow chart of the methodology.

Part II

Applications

Chapter 4

Application to the Recovery of Pollutants Release History

4.1 Literature Review

In the last 50 years one of the main problems in environmental issues is the groundwater contamination caused by industrialization, chemical fertilizer, waste disposal etc. Increasing environmental consciousness has developed several tools for the reclamation of groundwater but who has to pay? So, it was stated the principle: "The polluter pays", but who is the polluter? Several techniques were developed to find the cause of the pollution and in the last 30 years several methods were developed to study the release time history of the pollutants. One of the first methods was proposed by *Gorelick et al.* [1983]. This approach combines numerical simulation methods and linear programming with estimation techniques and multiple regression aimed at identifying the groundwater source locations and magnitudes; it is applicable to well defined groundwater systems. *Wagner* [1992] developed a deterministic approach to the simultaneous estimation of model parameters and the solute source characteristics. *Skaggs and Kabala* [1994] applied Tikhonov regularization to find the release history of the pollution source. Through this method the inverse problem is reduced to a minimization problem with a unique solution; this method, however, is affected by the plume measurement errors and by the accuracy of the transport parameter estimates. Moreover it presents an internal parameter whose estimation is quite difficult and arbitrary. *Skaggs and Kabala* [1995] proposed a quasi reversibility method that it is easier to implement, but it is less accurate than Tikhonov regularization. *Woodbury et al.* [1996, 1998] proposed the Minimum Relative Entropy approach to recover the release and the evolution histories of a plume. *Aral and Guan* [1996] used genetic algorithms and response matrix technique to identify sources of groundwater pollution. *Mahar and Datta* [1997] applied nonlinear optimization models for the identification of the unknown groundwater pollution sources by means of embedding technique and for the estimation of the aquifer parameters [*Mahar and Datta,* 2000, 2001]; they specifically addressed the problem of designing an optimal monitoring network for efficient source identification [*Mahar and Datta,* 1997]. *Snodgrass and Kitanidis* [1997] and *Michalak and Kitanidis* [2002, 2003, 2004b, 2004c] developed and applied the geostatistical methodology to recover the release history. *Butera and Tanda* [2001, 2002] improved the geostatistical methodology proposed by *Snodgrass and Kitanidis* [1997] for a 2-D case, with multiple sources and in a heterogeneous field [*Butera and Tanda,* 2003]. *Liu and Ball* [1999] applied the Tikhonov regularization, proposed by *Skaggs*

Method	Reference	Estimated Function	Source Type	Domain	Dimensions	Site
		Deterministic Approaches				
Tikhonov regularization	*Skaggs and Kabala* [1994]	release history	point	homogeneous	1-D	hypothetical
	Skaggs and Kabala [1998]	release history	point	homogeneous	1-D	hypothetical
	Liu and Ball [1999]	concentration history	interface	nonuniform	1-D	DAFB[a]
	Neupauer et al. [2000]	release history	point	homogeneous	1-D	hypothetical
Quasi-reversibility	*Skaggs and Kabala* [1995]	release history	point	homogeneous	1-D	hypothetical
	Bagtzoglou and Atmadja [2003]	historical distribution	point	heterogeneous	1-D	hypothetical
Spectral analysis	*Birchwood* [1999]	location and time of rectangular pulse	point	homogeneous	1-D	hypothetical
Nonregularized nonlinear least squares	*Alapati and Kabala* [2000]	release history	point	homogeneous	1-D	hypothetical
Progressive genetic algorithm methods	*Aral et al.* [2001]	release history and source location	point	heterogeneous	3-D	hypothetical
Marching-jury backward beam equation method	*Atmadja and Bagtzoglou* [2001]	release history	point	nonuniform	1-D	hypothetical
	Bagtzoglou and Atmadja [2003]	historical distribution	point	heterogeneous	1-D	hypothetical
		Stochastic Approaches				
Minimum relative entropy	*Woodbury and Ulrych* [1996]	release history	point	homogeneous	1-D	hypothetical
	Woodbury et al. [1998]	release history	patch or point	homogeneous	3-D	GL[b] and hypothetical
	Neupauer et al. [2000]	release history	point	homogeneous	1-D	hypothetical
Geostatistically based methods	*Snodgrass and Kitanidis* [1997]	release history	point	homogeneous	1-D	hypothetical
	Michalak and Kitanidis [2004]	concentration history	interface	nonuniform	1-D	DAFB[a]
	Michalak and Kitanidis [2002]	release history	patch	homogeneous	3-D	GL[b]
	Michalak and Kitanidis [2003]	concentration history	interface	nonuniform	1-D	DAFB[a]
	present work	historical distribution	any	heterogeneous	3-D	hypothetical

[a]Dover Air Force Base, Delaware.
[b]Gloucester Landfill, Ottawa, Ontario, Canada.

Figure 4.1: Summary of the methodologies, from *Michalak and Kitanidis* [2004b].

and Kabala [1994] to a real site. *Neuaper and Wilson* [1999] showed that the backward model probabilities (probability of where the particle was located at some prior time) are adjoint states of resident concentration and provided a methodology for obtaining the governing equations of these probabilities. *Birchwood* [1999] used a Fourier based inverse technique to recover the source location and the release history of the groundwater contaminant plume from breakthrough curve data obtained at a single monitoring well. *Alapati and Kabala* [2000] used a nonlinear least squares method without regularization to recover the release history of a groundwater contaminant plume given its measured spatial distribution. *Atmadja and Bagtzoglou* [2000, 2001] improved the Backward Beam Equation (Marching-Jury Backward Beam Equation-MJBBE) to solve the Advection Dispersion Equation within a contaminant source identification context. *Bagtzoglou and Atmadja* [2003] made a comparison between the Quasi Reversibility method and the MJBBE to recover conservative contaminant plume spatial distribution. *Michalak and Kitanidis* [2004b] applied the adjoint state method to the geostatistical procedure with the aim at recovering of the precedent distribution of contaminant at a given point back in time. The described methodology is able to manage heterogeneous fields as it can use numerical codes for solving the flow and transport problems. *Neuaper and Wilson* [2005] improved the backward probability model [*Neuaper and Wilson*, 1999] adding the possibility to sample at multiple locations and times and applied the methodology to a real case to identify the source location.

More specific literature review and several comparisons were carried out by *Neuaper et al.* [2000], *Atmadja and Bagtzoglou* [2001] and *Michalak and Kitanidis* [2004b], see Figure 4.1.

4.2 Methodology

The methodology proposed by *Kitanidis* [1995] and *Snodgrass and Kitanidis* [1997] has been studied and applied because it resulted the most promising after the literature review. The methodology is briefly explained in the following; for more details see *Kitanidis* [1995] and *Snodgrass and Kitanidis* [1997]. The value of the concentration may be estimated by the following expression:

$$\mathbf{z} = \mathbf{h}(\mathbf{s}, \mathbf{r}) + \mathbf{v} \tag{4.1}$$

where $\mathbf{z}$ is an $m \times 1$ vector of observations, $\mathbf{s}$ is the unknown release function discretized in n time intervals, $\mathbf{r}$ is a vector that contains other parameters needed by the model function $\mathbf{h}(\mathbf{s}, \mathbf{r})$ (for instance the aquifer parameters). In this work the parameters $\mathbf{r}$ are assumed known so the equation (4.1) is reduced to $\mathbf{z} = \mathbf{h}(\mathbf{s}) + \mathbf{v}$. $\mathbf{h}(\mathbf{s})$ is an $m \times 1$ vector that represents the transport process. $\mathbf{v}$ is an $m \times 1$ vector of measurements errors assumed with zero mean and known covariance matrix $\mathbf{R}$. The unknown function $s(t)$ is analyzed as a stochastic random process characterized by the mean $E[\mathbf{s}] = \mathbf{X}\boldsymbol{\beta}$ and covariance matrix $\mathbf{Q}(\boldsymbol{\theta}) = E\left[(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})^T\right]$, where $\mathbf{X}$ is a known $n \times p$ matrix and $\boldsymbol{\beta}$ are $p \times 1$ unknown drift coefficients; $\boldsymbol{\theta}$ are the unknown parameters of the covariance matrix $\mathbf{Q}$. Taking for instance the Gaussian distribution, the parameters of $\mathbf{Q}$ are the variance (σ^2) and the characteristic time length (l). For conservative solute transport the relationship between the observed concentration and the solute input is linear so the equation (4.1) can be simplified to

$$\mathbf{z} = \mathbf{H} \cdot \mathbf{s} + \mathbf{v} \tag{4.2}$$

where $\mathbf{H}$ is a known matrix $m \times n$ and represents the matrix of the transfer functions. The estimation procedure proposed by *Kitanidis* [1995], see section 3.2, is divided into two parts; first the **structural parameters** $\boldsymbol{\theta}$ of the chosen distribution are found then the unknown release function is **estimated**. The parameters $\boldsymbol{\theta}$ are estimated maximizing the probability that the process represents the observation $\mathbf{z}$, through the following equation:

$$p(\mathbf{z}|\boldsymbol{\theta}) \propto |\boldsymbol{\Sigma}|^{-1/2} \left|\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Sigma}^{-\mathbf{1}}\mathbf{H}\mathbf{X}\right|^{-1/2} \exp\left[-\frac{1}{2}\mathbf{z}^T\boldsymbol{\Xi}^{-1}\mathbf{z}\right] \tag{4.3}$$

where

$$\boldsymbol{\Sigma} = \mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R} \tag{4.4}$$

($\boldsymbol{\Sigma}$ is a $m \times m$ matrix) and

$$\boldsymbol{\Xi} = \boldsymbol{\Sigma}^{-\mathbf{1}} - \boldsymbol{\Sigma}^{-\mathbf{1}}\mathbf{H}\mathbf{X}(\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Sigma}^{-\mathbf{1}}\mathbf{H}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Sigma}^{-\mathbf{1}} \tag{4.5}$$

($\boldsymbol{\Xi}$ is a $m \times m$ matrix). Maximizing the (4.3) is equivalent to minimizing the negative logarithm of $p(\mathbf{z}|\boldsymbol{\theta})$:

$$L(\boldsymbol{\theta}) = \frac{1}{2}\ln|\boldsymbol{\Sigma}| + \frac{1}{2}\ln\left|\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Sigma}^{-\mathbf{1}}\mathbf{H}\mathbf{X}\right| + \frac{1}{2}\mathbf{z}^T\boldsymbol{\Xi}^{-1}\mathbf{z} \tag{4.6}$$

The $\boldsymbol{\beta}$ parameters are eliminated by the integration described in equation (3.6) (see the section 3.2) this process removes the bias caused by the unknown coefficients $\boldsymbol{\beta}$. The minimization of equation (4.6) can be achieved by taking derivatives of $L(\boldsymbol{\theta})$ respect to $\boldsymbol{\theta}$ and setting them to zero. This process can be defined by:

$$g_i = \frac{\partial L}{\partial \theta_i} = \frac{1}{2}Tr\left[\boldsymbol{\Xi}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right] - \frac{1}{2}\mathbf{z}^T\left(\boldsymbol{\Xi}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\boldsymbol{\Xi}\right)\mathbf{z}$$

and this equation can be solves numerically using the Gauss-Newton iterative method:

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - \mathbf{F}^{-1}\mathbf{g}$$

where $\mathbf{F}$ is the Fisher information matrix:

$$\mathbf{F}_{ij} = \frac{1}{2} Tr \left[\boldsymbol{\Xi} \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i} \boldsymbol{\Xi} \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j} \right]$$

Sometimes it is necessary to apply the Marquardt modification by replacing $\mathbf{F}^{-1}$ with $(\mathbf{F} + \lambda \mathbf{I})^{-1}$, where $\mathbf{I}$ is the identity matrix and λ is a positive parameter.

Once that the iterations have converged and the $\boldsymbol{\theta}$ are evaluated it is possible to obtain the best estimate of the release function $s(t)$ by solving the universal Kriging:

$$\hat{\mathbf{s}} = \boldsymbol{\Lambda} \cdot \mathbf{z}$$

The Kriging method evaluates the $\boldsymbol{\Lambda}$ coefficients $(n \times m)$ with the constrains of unbiased mean and minimum variance of the error, solving the following system:

$$\begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{HX} \\ (\mathbf{HX})^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Lambda}^T \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{HQ} \\ \mathbf{X}^T \end{bmatrix} \tag{4.7}$$

where $\mathbf{M}$ is a matrix of multiplier $(p \times n)$. The covariance matrix $(n \times n)$ of the error of the estimate is

$$\mathbf{V} = -\mathbf{XM} + \mathbf{Q} - \mathbf{QH}^T\boldsymbol{\Lambda}^T \tag{4.8}$$

The transfer matrix $\mathbf{H}$, discretized in n regular time intervals Δt, results the following:

$$\mathbf{H} = \Delta t \cdot \begin{bmatrix} f(x_1, T - t_1) & \dots & f(x_1, T - t_n) \\ f(x_2, T - t_1) & \dots & f(x_2, T - t_n) \\ \dots & \dots & \dots \\ f(x_m, T - t_1) & \dots & f(x_m, T - t_n) \end{bmatrix} \tag{4.9}$$

The function $f(x_i, T - t_j)$ represents the transfer functions (TFs) located in the point i at time $T - t_j$. In simple cases such as 1-D or 2-D uniform flow, an analytical solution for the TF is available (for instance equation (5.11)). This methodology (**unconstrained case**) is functional and fast but does not enforce the non-negativity of the concentration (see Figures 5.7 and 5.17); then, with the aim at avoiding this kind of problem, a transformation of the unknown variable $\mathbf{s}$ has been proposed [*Kitanidis and Shen*, 1996; *Snodgrass and Kitanidis*, 1997]. The new unknown becomes:

$$\widetilde{\mathbf{s}} = \alpha \left(\mathbf{s}^{1/\alpha} - 1 \right) \tag{4.10}$$

where α is a positive number and chosen as small as possible while ensuring that $\widetilde{\mathbf{s}} > -\alpha$; (in this work $\alpha = 2$). The values of $\mathbf{s}$ are then constrained to be positive (**constrained case**) and they are physically compatible.

The equation (4.1) in the transformed spaces becomes:

$$\mathbf{z} = \widetilde{\mathbf{h}}\left(\widetilde{\mathbf{s}}\right) + \mathbf{v} = \mathbf{h}\left[\left(\frac{\widetilde{\mathbf{s}} + \alpha}{\alpha} \right)^{\alpha} \right] + \mathbf{v} \tag{4.11}$$

In this case the transfer function $\widetilde{\mathbf{h}}\left(\widetilde{\mathbf{s}}\right)$ is not linear with respect to the unknown $\widetilde{\mathbf{s}}$ and the solution is reached iteratively [see *Kitanidis*, 1995]. The procedure starts with making an initial estimate of the unknown $\widetilde{\mathbf{s}}_l$ then the derivative of $\widetilde{\mathbf{h}}$ with respect to $\widetilde{\mathbf{s}}$ at $\widetilde{\mathbf{s}}_l$ is found:

$$\widetilde{\mathbf{H}}_l = \left.\frac{\partial \widetilde{\mathbf{h}}}{\partial \widetilde{\mathbf{s}}}\right|_{\widetilde{\mathbf{s}}=\widetilde{\mathbf{s}}_l} \tag{4.12}$$

Then it is possible to find the new estimate: $\hat{\mathbf{s}}_{l\_1} = \mathbf{\Lambda} \cdot \mathbf{z}_{0l}$, where $\mathbf{z}_{0l} = \mathbf{z} - \widetilde{\mathbf{h}}\left(\widetilde{\mathbf{s}}_l\right) + \widetilde{\mathbf{H}}_l \widetilde{\mathbf{s}}_l$, and $\widetilde{\mathbf{h}}\left(\widetilde{\mathbf{s}}_0\right) = \int_0^t \left(\frac{\widetilde{\mathbf{s}}_0 + \alpha}{\alpha}\right)^{\alpha} f\left(x, t-\tau\right) d\tau$, where $\widetilde{\mathbf{s}}_0$ is the initial guess. The $\mathbf{\Lambda}$ $(n \times m)$ matrix of coefficients is evaluated from the solution of the following system:

$$\begin{bmatrix} \mathbf{\Sigma} & \widetilde{\mathbf{H}}_l \mathbf{X} \\ \left(\widetilde{\mathbf{H}}_l \mathbf{X}\right)^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}^T \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{H}}_l \mathbf{Q} \\ \mathbf{X}^T \end{bmatrix}$$

where $\mathbf{\Sigma} = \widetilde{\mathbf{H}}_l \mathbf{Q} \widetilde{\mathbf{H}}_l^T + \mathbf{R}$ and $\mathbf{M}$ is a matrix of multiplier $(p \times n)$. The transfer function matrix $\widetilde{\mathbf{H}}_l$ for the transformed variable $\widetilde{\mathbf{s}}_l$ becomes:

$$\widetilde{\mathbf{H}}_l = \Delta t \cdot \begin{bmatrix} \left(\frac{\widetilde{\mathbf{s}}_0(t_1) + \alpha}{\alpha}\right)^{\alpha-1} f\left(x_1, T - t_1\right) & \dots & \left(\frac{\widetilde{\mathbf{s}}_0(t_n) + \alpha}{\alpha}\right)^{\alpha-1} f\left(x_1, T - t_n\right) \\ \left(\frac{\widetilde{\mathbf{s}}_0(t_1) + \alpha}{\alpha}\right)^{\alpha-1} f\left(x_2, T - t_1\right) & \dots & \left(\frac{\widetilde{\mathbf{s}}_0(t_n) + \alpha}{\alpha}\right)^{\alpha-1} f\left(x_2, T - t_n\right) \\ \dots & \dots & \dots \\ \left(\frac{\widetilde{\mathbf{s}}_0(t_1) + \alpha}{\alpha}\right)^{\alpha-1} f\left(x_m, T - t_1\right) & \dots & \left(\frac{\widetilde{\mathbf{s}}_0(t_n) + \alpha}{\alpha}\right)^{\alpha-1} f\left(x_m, T - t_n\right) \end{bmatrix}$$

The structural parameters $\boldsymbol{\theta}$ are evaluated at each iteration minimizing the equation:

$$L\left(\boldsymbol{\theta}\right) = \frac{1}{2} \ln\left|\mathbf{\Sigma}\right| + \frac{1}{2} \ln\left|\mathbf{X}^T \widetilde{\mathbf{H}}_l^T \mathbf{\Sigma}^{-1} \widetilde{\mathbf{H}}_l \mathbf{X}\right| + \frac{1}{2} \mathbf{z}_{0l}^T \mathbf{\Xi}^{-1} \mathbf{z}_{0l}$$

where $\mathbf{\Xi} = \mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1} \widetilde{\mathbf{H}}_l \mathbf{X} \left(\mathbf{X}^T \widetilde{\mathbf{H}}_l^T \mathbf{\Sigma}^{-1} \widetilde{\mathbf{H}}_l \mathbf{X}\right)^{-1} \mathbf{X}^T \widetilde{\mathbf{H}}_l^T \mathbf{\Sigma}^{-1}$.

In the parameter estimation step $\mathbf{z}_{0l}$ and $\widetilde{\mathbf{H}}_l$ are treated as constants and the minimization is obtained using the Gauss-Newton method (see equation (3.14)). Once $\boldsymbol{\theta}$ is found it is used to find and update the value of $\widetilde{\mathbf{s}}_l$. This iterative process is continued until $\widetilde{\mathbf{s}}_l$ and $\boldsymbol{\theta}$ converge. The final value of $\widetilde{\mathbf{s}}_l$ is the best estimate of the transformed function and its covariance function is given by the equation $\mathbf{V} = -\mathbf{X}\mathbf{M} + \mathbf{Q} - \mathbf{Q}\widetilde{\mathbf{H}}^T \mathbf{\Lambda}^T$.

At the end of the process it is necessary to apply the inverse transformation of equation (4.10) to calculate the release function $s(t)$; the best estimate results:

$$\hat{\mathbf{s}} = \left(\frac{\widetilde{\mathbf{s}}_l + \alpha}{\alpha}\right)^{\alpha} \tag{4.13}$$

In the last ten years several improvements of the methodology were proposed. *Butera and Tanda* [2001] applied the methodology proposed by *Snodgrass and Kitanidis* [1997] to a 2-D domain (see the following sections) and they made an analysis on the impact of measurements concentration errors, on the impact of errors in aquifer parameters estimate and erroneous identification of the hydraulic gradient direction. *Michalak and Kitanidis* [2002] applied the methodology to a real case and enhanced the methodology with the spatially and temporally distributed data. So the sensitivity matrix becomes:

$$\mathbf{H} = \begin{bmatrix} f(\mathbf{x}_1, T_1 - t_1) & \dots & f(\mathbf{x}_1, T_1 - t_m) \\ f(\mathbf{x}_2, T_2 - t_1) & \dots & f(\mathbf{x}_2, T_2 - t_m) \\ \dots & \dots & \dots \\ f(\mathbf{x}_n, T_n - t_1) & \dots & f(\mathbf{x}_n, T_n - t_m) \end{bmatrix} \tag{4.14}$$

Butera and Tanda [2002, 2003] applied the methodology both to sources of finite dimensions and point sources and to single and multiple sources. For instance, for two sources mutually independent, the $\mathbf{H}$ matrix becomes:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 \end{bmatrix} \tag{4.15}$$

and $\mathbf{s} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix}$ where 1 means the first source and 2 the second one. The covariance matrix becomes:

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_1 & 0 \\ 0 & \mathbf{Q}_2 \end{bmatrix} \tag{4.16}$$

Michalak and Kitanidis [2003] developed a new method to enforce parameter non negativity, that uses the Markov Chain Monte Carlo theory. This new improvement was also tested to a real case [*Michalak and Kitanidis*, 2004a] previously analyzed by *Liu and Ball* [1999]. *Butera and Tanda* [2003] applied the methodology with temporally distributed data to a heterogeneous stochastic field. The objective of the paper of *Michalak and Kitanidis* [2004b] doesn't deal with the recovery of the release function but with the determination of the antecedent distribution of contaminant at a given point back in time. That is a new point of view on the inverse problem of transport that needs only a few hypotheses on the boundary condition (i.e. no location of extension of the source is required) and seems to give more information in space. But since the unknown functions are several, it needs a large amount of computational efforts for the backward recovering or eventually for the forward forecasting of the pollutant transport. With the recovering of the pollutant release history, described in the present paper, both the backward and forward procedures are easy and very computing time saving.

In the following chapters new developments of the methodology proposed by *Snodgrass and Kitanidis* [1997] are shown.

Chapter 5

New Developments in the Recovering the Pollutants Release History: Transfer Function

Some approaches developed in the literature in order to solve the second type of inverse problem (for instance geostatistical approach and Tikhonov regularization method [*Skaggs and Kabala*, 1994]) require the computation of the function that describes the effect in time, in a certain location of the aquifer, of an impulsive release of pollutant in the source. This function, named transfer or Kernel function (TF) can be analytically determined if the problem has a simple geometry and boundary conditions. In many cases the characteristics of the groundwater flow field do not allow an analytical transfer function formulation; this is the case, for instance, of non uniform in the mean flow field due to complex boundary conditions, or to the existence of pumping wells, or to high heterogeneity of conductivity fields [*Sudicky*, 1986]. With the available procedures the technician has to reduce the real problem to a very simplified scheme to which the analytical transfer function can be applied. As a consequence a rough approximation of results can be expected.

In this work a new methodology for the evaluation of the transfer function is presented; it allows the application of the approach to real cases without using heavy simplification. This methodology is based on the analogy between the TF and the Instantaneous Unit Hydrograph (IUH) in surface hydrology. It is possible to evaluate the TF using a numerical model that represents the study area, so that for each monitoring point it is possible to get a transfer function. The numerical model has to be set up previously, starting from real field data. Then using a step input function for each source it is possible to evaluate the breakthrough curve. The transfer function is evaluated making the time derivative of the breakthrough curve. Two analytical examples are compared to test the methodology, then several synthetic cases are shown to highlight the great potential of the methodology.

5.1 Mathematical Formulation

As shown in the literature review proposed by *Michalak and Kitanidis* [2004b] (see Figure 4.1) only a few methodologies proposed in the last years are flexible for application to multiple sources, heterogeneity, multi dimensional case and non uniform flow, and none of them has all those char-

acteristics (see section 4.1). In this paper a new methodology (preliminary results are reported in *Butera et al.* [2004], *Tanda et al.* [2005] and *Zanini et al.* [2005]) developed by the writer, with the aim at overcoming the limitations caused by the non uniformity of flow field, is described in the following.

The equation (5.1) represents the transport process considering a non reactive solute and a medium with constant porosity.

$$\frac{\partial (nC)}{\partial t} = -\nabla (nC \cdot \mathbf{v}) + \nabla (\mathbf{D} \cdot n \nabla C) + C_{in}(t)\, \delta (\mathbf{x} - \mathbf{x}_0) \tag{5.1}$$

$\mathbf{v}$ is the effective velocity, n the porosity, $\mathbf{D}$ the dispersion coefficients and $C(\mathbf{x}, t)$ the concentration at the location $\mathbf{x}$ at time t. Equation (5.1) is a linear differential equation; considering as boundary and initial conditions: $C(\mathbf{x}, 0) = 0$; $C(\infty, t) = 0$, $C(\mathbf{x}_0, t) = C_{in}(t)$, where $C_{in}(t)$ represents the source release, the solution of equation (5.1) is given by the following integral [*Jury and Roth*, 1990]:

$$C(\mathbf{x}, t) = \int_o^t C_{in}(\tau) \cdot f(\mathbf{x}, t - \tau)\, d\tau \tag{5.2}$$

where $f(\mathbf{x}, t - \tau)$ is the Kernel or transfer function (TF). Considering the variable transformation $y = t - \tau$, so $dy = -d\tau$ and for $\tau = 0$ $y = t$, for $\tau = t$ $y = 0$, the equation (5.2) becomes:

$$C(\mathbf{x}, t) = \int_o^t C_{in}(t - y) \cdot f(\mathbf{x}, y)\, dy \tag{5.3}$$

turning y in τ the (5.3) results:

$$C(\mathbf{x}, t) = \int_o^t C_{in}(t - \tau) \cdot f(\mathbf{x}, \tau)\, d\tau \tag{5.4}$$

The Heaviside function (called also unit step function) is considered as the input function ($C_{in}(t) \equiv H(t)$); $H(t)$ is represented by the following, see also Figure 5.1:

$$H(t) = \begin{cases} 0 : t < 0 \\ 1 : t > 0 \end{cases} \tag{5.5}$$

The integral (5.4) due to the Heaviside function for $t > 0$ becomes

$$C_H(\mathbf{x}, t) = \int_o^t 1 \cdot f(\mathbf{x}, \tau) d\tau \tag{5.6}$$

The time derivative[1] of (5.6) is

$$\frac{\partial C_H(\mathbf{x}, t)}{\partial t} = \int_0^t \frac{\partial f(\mathbf{x}, t)}{\partial t} d\tau + f(\mathbf{x}, t) \cdot 1 - f(\mathbf{x}, 0) \cdot 0 \tag{5.7}$$

[1] Differentiation under the integral sign [*Amerio*, 1974]:

$$\frac{\partial}{\partial y} \int_{\alpha(y)}^{\beta(y)} f(x, y) dx = \int_{\alpha(y)}^{\beta(y)} \frac{\partial f(x, y)}{\partial y} dx + f(\beta(y), y) \frac{\partial \beta(y)}{\partial y} - f(\alpha(y), y) \frac{\partial \alpha(y)}{\partial y}$$

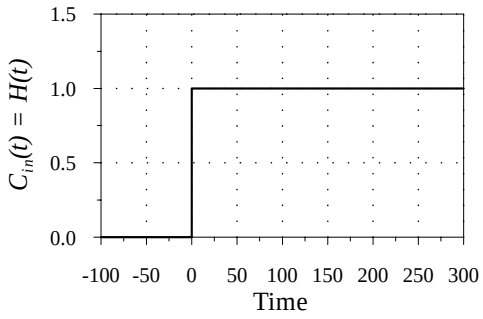

Figure 5.1: Heaviside function, equation (5.5), used as input function.

where the first ($\int_0^t \frac{\partial f(\mathbf{x},t)}{\partial t} d\tau$) and the third terms ($f(\mathbf{x},0)\cdot 0$) are null; then the equation (5.7) results

$$\frac{\partial C_H(\mathbf{x},t)}{\partial t} = f(\mathbf{x},t) \tag{5.8}$$

Equation (5.8), obtained for a stepwise input function, allows to outline a methodology for the identification of the TF also in complex cases (for instance in heterogeneous media, or in the presence of wells, recharge, etc.) where the TF is not always analytically known. The new methodology proposed in this work takes advantage of equation (5.8) so that it is possible to apply the available procedures, as e.g. the Tikhonov regularization [*Skaggs and Kabala*, 1994] and the geostatistical procedure [*Snodgrass and Kitanidis*, 1997] for a generic flow field.

It is possible to recognize an analogy with a well known hydrological problem: the computation of the runoff from a basin after a rain event. In fact equation (5.2) can be compared to the convolution integral that joins the net rainfall to the outflow in a given cross section of a river. As known, the impulse response function of the basin is called Instantaneous Unit Hydrograph (IUH) [*Chow*, 1964; *Wilson*, 1990]. In general it is possible to recover the IUH of a basin if the net rainfall and the outflow in the closing section of a basin are known. The determination of the IUH is very easy if the net rainfall hyetograph is very long and constant in time (indefinite step condition): the runoff hydrograph becomes as S curve and its derivative coincides with the IUH function. These concepts can be extended to the present problem: the equation (5.2), at a given location with coordinates $\mathbf{x}$, is comparable to the surface hydrology convolution integral, where the IUH plays the role of the TF and the net rainfall the role of the release function.

Unfortunately it is not possible to register the response of an aquifer due to an indefinite step injection, because such test can requires too long a time; however, a numerical model of flow and transport of the studied aquifer is often available for the prediction of the pollutant fate or for planning remediation actions; this model can also be used to simulate the response of the aquifer to an appropriate injection function. A step injection (Heaviside function, see equation (5.5) and Figure 5.1) can be applied using a numerical model of flow and transport of the studied area; the model has to be already calibrated and the source location also known [*Snodgrass and Kitanidis*, 1997]. In each node of the computation mesh it is possible to register a response function (breakthrough curve) in terms of concentration versus time with, generally, the shape of a S curve (see Figure 5.2a). The time derivative of this function (see Figure 5.2b) is the TF that describes the response due to an impulsive injection of pollution into the aquifer.

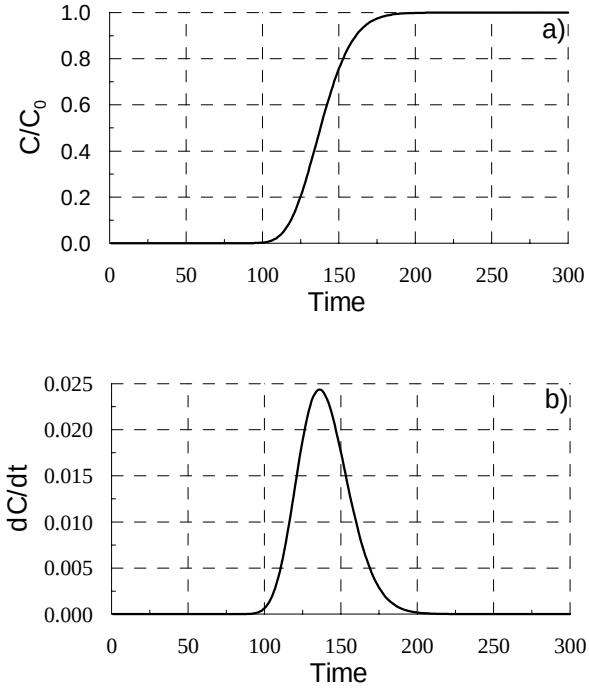

Figure 5.2: a) Breakthrough curve b) Derivative of breakthrough curve.

The choice of using an input step function and then deriving numerically the breakthrough curve rather than computing directly the aquifer response due to an impulsive injection, can be explained by the need of reducing numerical errors caused by the discrete time representation of the process, that are unavoidable in the numerical computations. The numerical description of an impulse necessarily have to develop in a small but finite time (not infinitesimal) so the aquifer response is not correct because it is due to a finite injection and not to an impulsive one. Besides, the responses of areas far from the source and less exposed to the contamination, can be underestimated. The study of the transport of a step input of infinite time length is more accurate and not affected by such kind of errors.

For a given measurement point, the breakthrough curve is an output of the numerical model and it consists in a series of concentration values recorded in time. The shorter is the temporal interval between two computations, the more accurate is the computation of the numerical derivative of the TF curve; by definition, the derivative can be calculated as the rate $\Delta C/\Delta t$, with $\Delta t \to 0$.

Once the TFs are obtained, it could be possible to forecast the evolution of the pollution in the aquifer by means of equation (5.2), if the release history of the source is known, but this application is useless since the numerical model of the aquifer is available. Instead, the TFs, obtained by the numerical derivative, are very useful to solve the inverse problem of the second kind object of this work. Toward this end the methodology proposed by *Snodgrass and Kitanidis* [1997] can be adopted.

5.2 Mathematical Statement of the Problem

In the following, 1-D and 2-D flow cases are considered for the application of the geostatistical approach making use of the considerations reported in section 5.1. In all the examples the procedure is applied in the constrained case.

5.2.1 Mathematical Statement of the 1-D Problem

Concerning the transport process of a conservative solute in a 1-D steady flow the advection dispersion equation (5.1) collapses in the following:

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x}\left(D(x)\frac{\partial C}{\partial x}\right) + \frac{\partial}{\partial x}(vC) \tag{5.9}$$

v is the effective velocity, D the longitudinal dispersion and $C(x,t)$ the concentration at the distance x from the source at time t. Equation (5.9) is a linear differential equation, with boundary and initial condition: $C(x,0) = 0$; $C(\infty,t) = 0$, $C(0,t) = C_{in}(t)$, where $C_{in}(t)$ represents the source release; the solution of equation (5.9) is given by the following integral:

$$C(x,t) = \int_o^t C_{in}(\tau) \cdot f(x, t-\tau)\, d\tau \tag{5.10}$$

where $f(x,t-\tau)$ represents the kernel function called also transfer function (TF). $f(x,t-\tau)$ describes the effect due to an impulsive release of the source at time $t-\tau$ on the measured concentration at time t at point x. The solution of the ADE formulated in equation (5.10), that joins the input function (contaminant release) to the output function (distribution of solute in groundwater), holds also for complex cases, also for those not analytically solvable, provided that the process preserve the linearity [*Jury and Roth*, 1990]. Previous studies [*Jury et al.*, 1982; *Rinaldo et al.*, 1989] considered the transport processes in a statistical framework and the TF becomes the probability density function (pdf) that means the probability of the solute released in x_0 at time $t-\tau$ of being in x at time t.

The TF, in the 1-D case with uniform flow and constant dispersion coefficient, can be analytically represented by the following [*Jury and Roth*, 1990]:

$$f(x,t) = \frac{x}{2\sqrt{\pi D t^3}} \exp\left(-\frac{(x-vt)^2}{4Dt}\right) \tag{5.11}$$

5.2.2 Homogeneous 1-D Flow and Transport: Analytical Versus Numerical TF

The methodology outlined has been tested through a numerical example proposed in the literature [*Snodgrass and Kitanidis*, 1997]; it deals with the 1-D transport by advection and dispersion of a conservative solute through a homogeneous porous media. The release function adopted $C_{in}(t)$, shown in Figure 5.3, is the one proposed by *Skaggs and Kabala* [1994]:

$$s(t) = \exp\left(\frac{-(t-130)^2}{50}\right) + 0.3\exp\left(\frac{-(t-150)^2}{200}\right) + 0.5\exp\left(\frac{-(t-190)^2}{98}\right) \tag{5.12}$$

Notice that, in agreement with the symbols adopted in the previous section $C_{in}(t)$ is replaced by $s(t)$ (i.e. $C_{in}(t) \equiv s(t)$).

Then the 1-D problem stated in equation (5.9), plus the boundary conditions given in the above section 5.2.1, is solved through the integral (5.10) using the TF of equation (5.11); all variables are made dimensionless. At time $T = 300$, some concentration data are sampled at given locations of the flow field (Figure 5.4); the data collected are stored in the $\mathbf{z}$ measurement vector.

The release function is recovered by the geostatistical methodology (see section 4.2) adopting the expression (5.11) as TF. A Gaussian covariance function is assumed for $s(t)$ and the elements of the matrix $\mathbf{Q}$ are computed ($Q(t_i, t_j|\boldsymbol{\theta}) = \sigma^2 \exp\left(-\frac{(t_i - t_j)^2}{l^2}\right)$, where $(t_i - t_j)$ represents the separation distance in time units and the $\boldsymbol{\theta}$ are the variance (σ^2) and the characteristic time scale (l)); the error covariance matrix is $\mathbf{R} = \sigma_R^2 \cdot \mathbf{I}$, with $\sigma_R^2 = 1 \cdot 10^{-12}$.

Next, the numerical case is analyzed also through a 1-D numerical model, using the computer codes MODFLOW and MT3D that work with a finite difference scheme [*Harbaugh*, 2000; *Zheng and Wang*, 1999]. A column of porous material is numerically modeled; the length of the column is equal to 300 space units and the cross section size is unitary. The head boundary conditions are fixed in order to induce a unit effective velocity and unit dispersion coefficient. In the upstream cell a constant rate of solute is injected during the whole simulation, performed until the breakthrough curve is non zero at the downstream end of the column. At each cell of the domain the concentration value increases in time, following the S shape line shown in Figure 5.2a. At each location, marked in Figure 5.4, the breakthrough curve is processed in order to obtain the time derivative function that, as explained before, represents the numerically obtained TF. Figure 5.9 shows the comparison between the analytical and numerical TF at $x = 140$; the numerical TF is obtained through 3 different schemes (in order to find the optimal one) adopted to compute the first derivative:

$$\begin{aligned} \frac{dC_i}{dt} &= \frac{C_{i+1} - C_i}{\Delta t} \text{ Forward,} \\ \frac{dC_i}{dt} &= \frac{C_i - C_{i-1}}{\Delta t} \text{ Backward,} \\ \frac{dC_i}{dt} &= \frac{C_{i+1} - C_{i-1}}{2\Delta t} \text{ Central.} \end{aligned} \tag{5.13}$$

In Figure 5.5, the excellent agreement among the lines can be noticed; the time step used is $\Delta t = 1$. To better clarify the TF behavior, in Figure 5.6 the TFs, obtained with the backward scheme, in different locations are depicted.

The geostatistical methodology, in the constrained case, is then applied to recover the release history, given the concentration measurements, i.e. the $\mathbf{z}$ vector, and the TF numerically computed (by a backward scheme). In Figure 5.8 the results of the computation are compared: the true release function and those computed through the analytical and numerical TFs, respectively, are shown versus time. The results of Figure 5.8 are quite satisfactory: just a few differences are present among the lines. By inspection of the estimate error variance (Figure 5.9), a greater width of the interquantile strip can be noticed for the case of numerical TF. This fact is due both to the unavoidable errors coming from the numerical modeling of the dispersion process and to the numerical computation of the time derivative of the breakthrough curve. It is opinion of the author, however, that the chance for an extension of the geostatistical methodology to cases where the TF is not analytically known, overshadows the imperfections outlined in Figure 5.8 and 5.9.

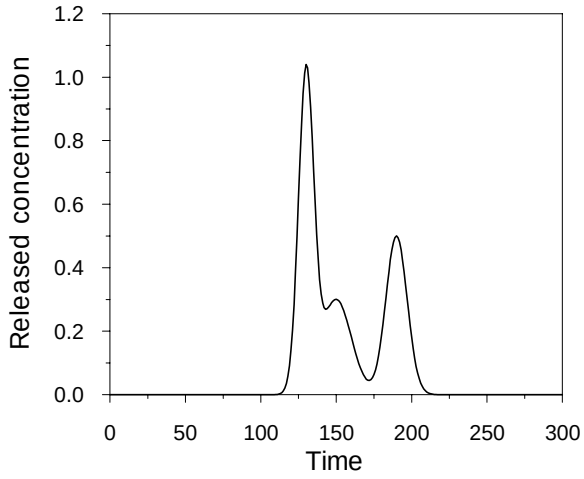

Figure 5.3: Released concentration proposed by *Skaggs and Kabala* [1994] based on the equation 5.12.

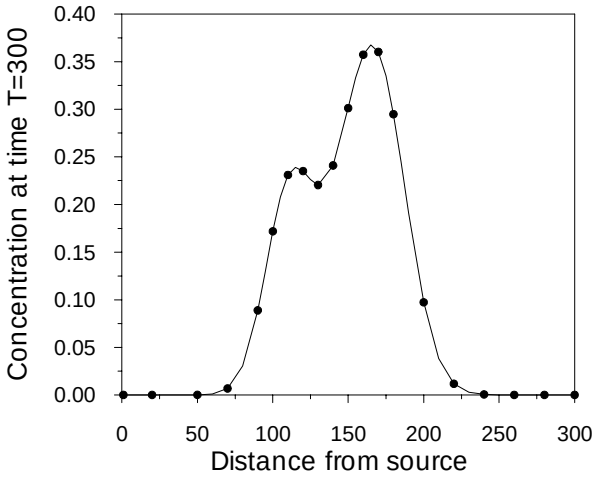

Figure 5.4: Concentration at time $T = 300$ and measurements locations (marks) in the 1-D homogeneous case.

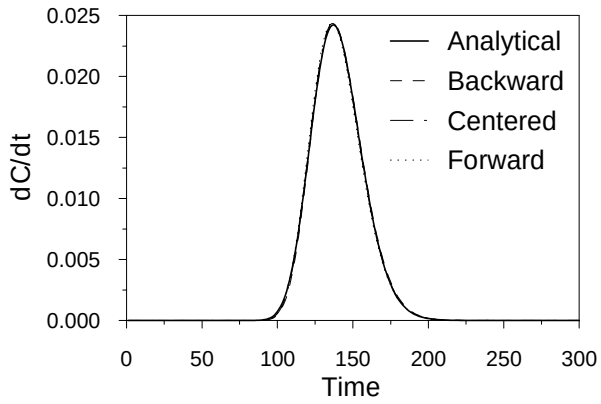

Figure 5.5: Analytical and Numerical (computed with three different derivative schemes) TFs in $x = 140$, 1-D.

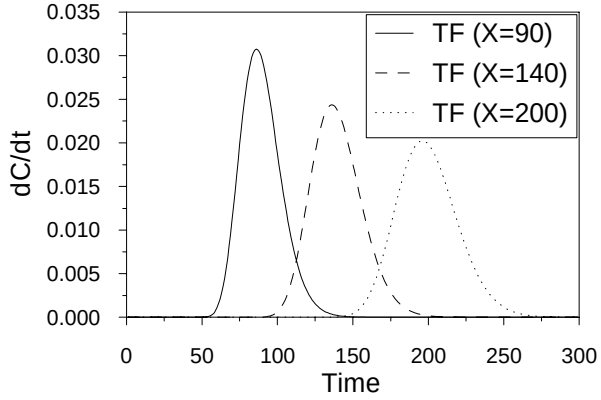

Figure 5.6: Numerical TF evaluated in three different locations using the backward scheme, 1-D.

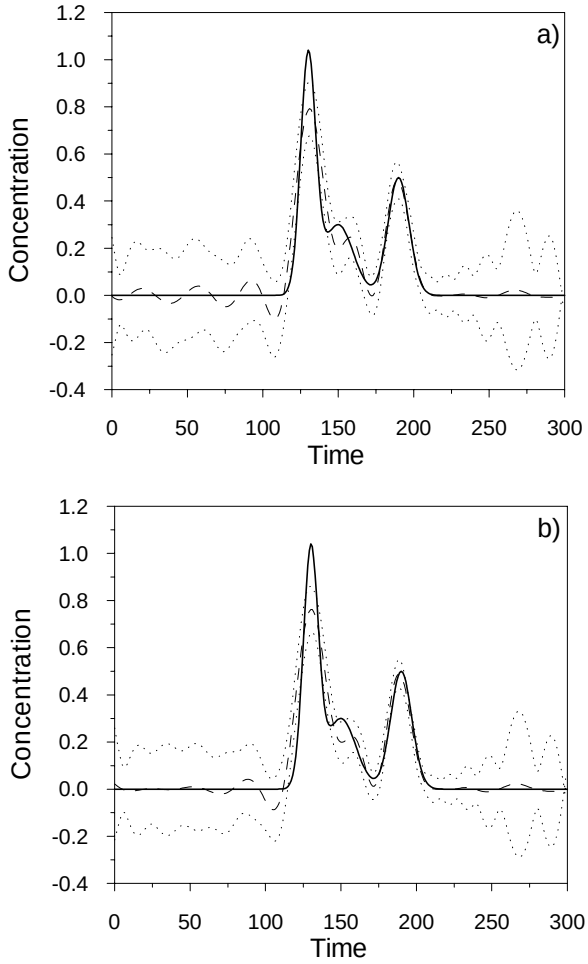

Figure 5.7: Unconstrained 1-D homogeneous case with Gaussian covariance: the true solution (solid line), best estimate (dashed line), and approximate 95% confidence interval (dotted line); a) Analytical, b) Numerical.

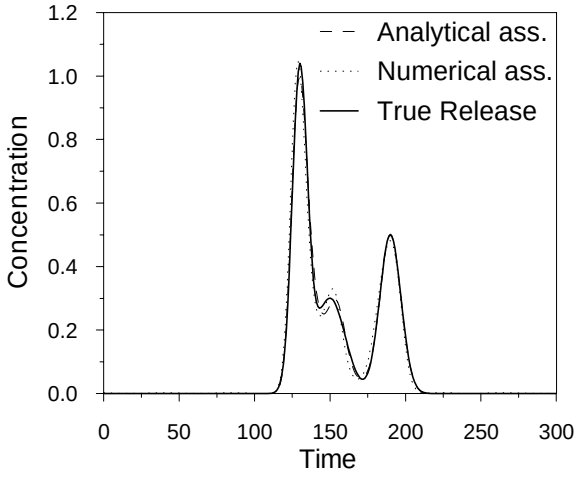

Figure 5.8: Analytical-Numerical assessment of the source release; 1-D constrained case.

5.2.3 Heterogeneous 1-D Flow and Transport - Numerical TF Only

The new example deals with a heterogeneous column made by two materials with the same thickness and different hydraulic properties, see Figure 5.10, but with a constant hydraulic conductivity. Two different sources of heterogeneity are considered: different porosity (that implies non uniform effective velocity) and different dispersion coefficients.

Three examples are considered using these sources of heterogeneity. The first case (a) deals with different porosities for the two materials. A 25% of variation of effective velocity is reached between the two zones. The test conditions are similar to the previously mentioned ones (section 5.2.1); a numerical flow and transport model is set up, subjected to a stepwise input function in the upstream cell. The time behavior of the concentration at the observation points (see Figure 5.4 for the locations) is used to calculate the numerical TFs. Next, a pollutant is injected in the upstream section of the model with the temporal distribution described by the function (5.12); at time $T = 300$ concentration data are sampled like in the previous example. These measured values are stored in the $\mathbf{z}$ vector and used in the geostatistical procedure. The check of the proposed methodology, for this case, is performed by comparison of the release function (5.12) to the recovered one; in fact, for this heterogeneous case no analytical TFs for comparison are available. Figure 5.11 shows the results using the constrained method; the recovered function is very close to the release one and the confidence interval seems reasonably narrow.

Figure 5.12 shows the results obtained in the second example (b). In this case the column is divided in two zones of material with different dispersion coefficients. The dispersion coefficient of the upper zone of the column is set to 1, in the remaining part it is equal to 0.5. Also in this case the recovered release function is satisfactory but the confidence interval is wider than in the previous heterogeneous analyzed case (a). This means a greater uncertainty of the results due to the bigger total dispersion obtained by the different dispersion coefficients.

The (c) example considers both the types of heterogeneity: different effective porosity and dispersion coefficient values. The values adopted are those of the two previous cases. Figure 5.13 shows the comparison between the true release history and the recovered one. The results obtained,

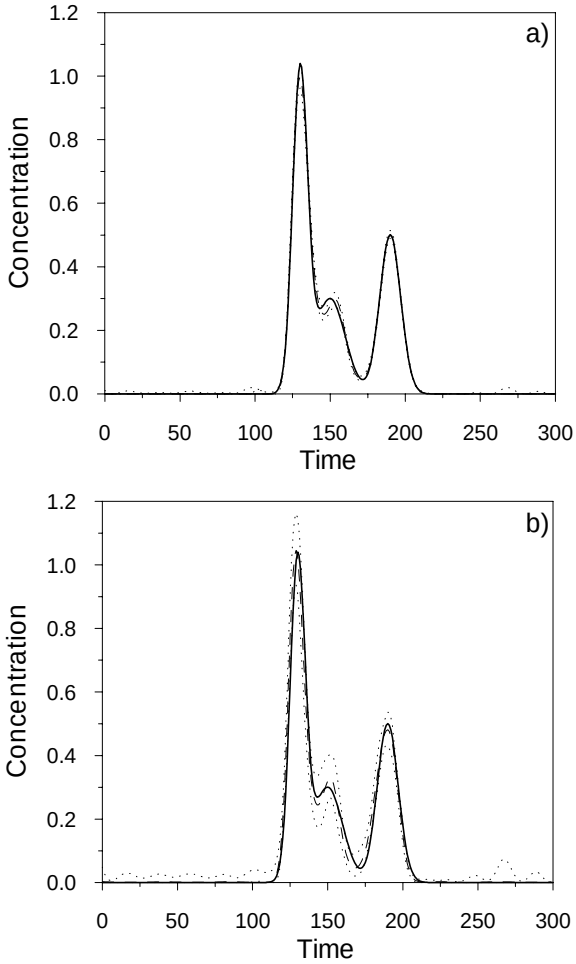

Figure 5.9: Constrained 1-D homogeneous case with Gaussian covariance: the true solution (solid line), best estimate (dashed line), and approximate 95% confidence interval (dotted line); a) Analytical, b) Numerical.

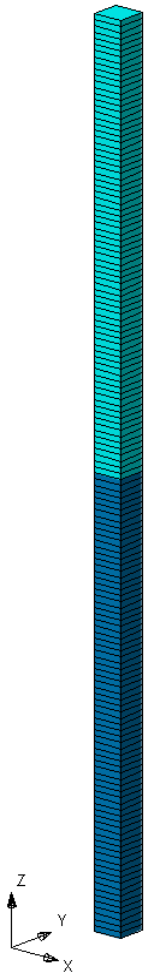

Figure 5.10: Grid of the 1-D numerical model, heterogeneous case: light blue: material 1; dark blue: material 2.

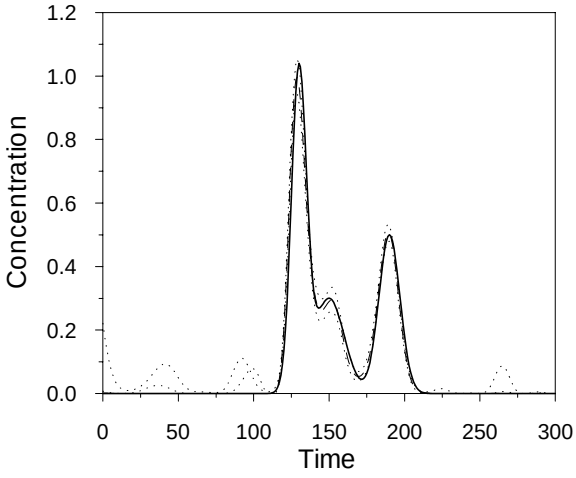

Figure 5.11: Constrained 1-D heterogeneous velocity case, with Gaussian covariance: the true solution (solid line), best estimate (dashed line), and approximate 95% confidence interval (dotted line).

as in the previous two cases, are very satisfactory.

5.2.4 Mathematical Statement of the 2-D Problem

Concerning the transport process of a conservative solute in a 2-D steady flow in the plane XY, the equation (5.1) can be specified in the following advection dispersion equation (ADE) describing the transport process, for a non reactive solute, due to a mass injection located in the position x_0, y_0 [*Bear*, 1972]:

$$\frac{\partial (nC)}{\partial t} = -\nabla' (nC \cdot \mathbf{v}) + \nabla' \left(\mathbf{D} \cdot n\nabla' C\right) + C_{in}(t)\, \delta (x - x_0)\, \delta (y - y_0) \tag{5.14}$$

In equation (5.14) $\mathbf{v}$ is the effective velocity, $\mathbf{D}$ is the dispersion tensor, $C_{in}(t)$ the rate of pollutant mass injected in time, $C(x, y, t)$ the concentration in the point with coordinates x, y at time t, n the porosity and the ∇' means the Nabla operator limited to x and y coordinates.

Equation (5.14) is a linear differential equation; for the boundary and initial condition: $C(x, y, 0) = 0$; $C(x, y, t) = 0$ $x \to \pm\infty$, $C(x, y, t) = 0$ $y \to \pm\infty$, it has the solution given by the following integral:

$$C(x, y, t) = \int_o^t C_{in}(\tau) \cdot f(x, y, t - \tau)\, d\tau \tag{5.15}$$

where $f(x, y, t - \tau)$ represents the TF. $f(x, y, t - \tau)$ describes the effect in the point x, y due to an impulsive release of the pollutant at time $t - \tau$. The solution of the ADE formulated in equation (5.14), that joins the input function (contaminant release) to the output function (distribution of solute in groundwater), holds also for complex cases, as well for those not analytically solvable, provided that the process preserve the linearity [*Jury and Roth*, 1990].

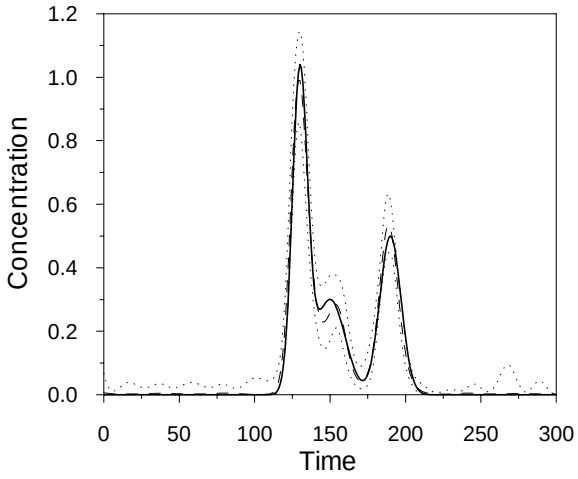

Figure 5.12: Constrained 1-D heterogeneous dispersion case, with Gaussian covariance: the true solution (solid line), best estimate (dashed line), and approximate 95% confidence interval (dotted line).

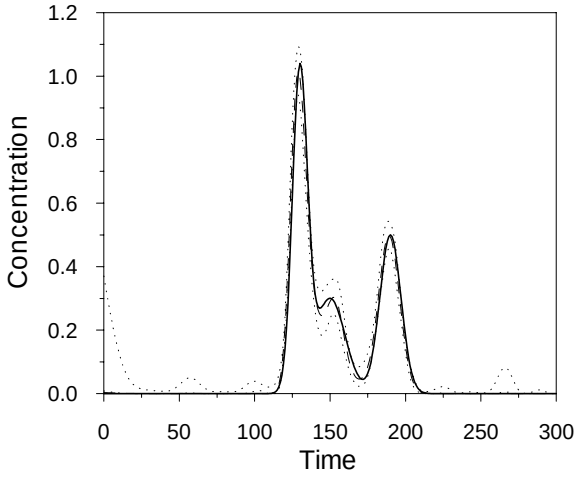

Figure 5.13: Constrained 1-D heterogeneous velocity + dispersion case, with Gaussian covariance: the true solution (solid line), best estimate (dashed line), and approximate 95% confidence interval (dotted line).

The TF, in the 2-D case with uniform flow ($v_y = 0$) and constant dispersion coefficient, can be analytically represented by the following [*Bear*, 1972]:

$$f(x,y,t) = \frac{1}{4\pi\sqrt{D_X D_Y(t)}} \cdot \exp\left[-\frac{(x-v(t))^2}{4D_X(t)} - \frac{y^2}{4D_Y(t)}\right] \tag{5.16}$$

being x and y the distance from the injection point along the longitudinal and transversal direction.

In 2-D problems $C_{in}(t)$ is given by the immission of a discharge $q(t)$ times the pollutant concentration $s(t)$ of the injected water: $C_{in}(t) = s(t) \cdot q(t)$. In general $s(t)$, $q(t)$ and $C_{in}(t)$ are functions of time. In the following 2-D examples $q(t)$ is taken as constant, known (unitary) and negligible compared with the groundwater flow [*Butera and Tanda*, 2003] in order to manage the unknown function to be the concentration time series $s(t)$.

5.2.5 Homogeneous 2-D Flow and Transport: Analytical Versus Numerical TF

The methodology outlined in section 5.2 has been tested also for the 2-D flow case through a numerical example proposed in the literature [*Butera and Tanda*, 2001]. It deals with the 2-D transport by advection and dispersion of a conservative solute through a homogeneous porous media. The release function adopted $s(t)$, shown in Figure 5.3, is the one proposed by *Skaggs and Kabala* [1994]. The source is located in the x_0, y_0 location. The 2-D problem described by the equation (5.14) plus the boundary conditions given in the previous section ($C_{in}(t) = s(t) \cdot q(t)$ with $q(t) = 1$) is solved through the integral (5.2) using the equation (5.16) as TF. All the variables are dimensionless and the hydrodispersive parameter values are $v_x = 1$, $v_y = 0$, $D_x = 1$, $D_y = 0.1$. At time $T = 300$ some concentration values are sampled at given locations of the flow field (Figure 5.16) and the collected data are stored in the $\mathbf{z}$ vector of measurements. Then, the release function is recovered by the geostatistical methodology adopting the expression (5.16) for the TF. As for the 1-D case the elements of the matrix $\mathbf{Q}$, the covariance matrix of $\mathbf{s}$, are computed by a Gaussian covariance function and the error covariance matrix is $\mathbf{R}$; assuming that the errors are taken to be independent and identically distributed with variance σ_R^2, the covariance matrix becomes $\mathbf{R} = \sigma_R^2 \cdot \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. The measurement error is known and it is negligible: $\sigma_R^2 = 1 \cdot 10^{-12}$.

The example in hand is then analyzed through the methodology outlined in section 4.2, ignoring the analytical expression of the TFs. A numerical model of the groundwater field is built by means of the MODFLOW and MT3D computer codes [*Harbaugh*, 2000; *Zheng and Wang*, 1999]. It has rectangular shape, 300 long and 20 units wide with unit thickness. The head boundary conditions are fixed and such to induce a unit seepage velocity in the positive x direction. The other hydrodispersive parameters are the same as considered in the analytical process. In the upstream cell a constant rate of solute is injected during the whole simulation, performed until the breakthrough curve reaches the downstream end of the model. At each cell of the domain involved in the pollution transport the concentration value increases in time, following a S shape line similar to that of Figure 5.2a. In some points located downstream the source (on the line $y = y_0$) and marked in Figure 5.16 (in agreement with the example proposed by *Butera and Tanda* [2001], the breakthrough curve is stored and processed in order to obtain the time derivative function that, as explained before, represents the numerically computed TF.

As for 1-D case, for comparison purposes the numerical TF is obtained through 3 different schemes of the first derivative: $\frac{dC_i}{dt} = \frac{C_{i+1} - C_i}{\Delta t}$ Forward, $\frac{dC_i}{dt} = \frac{C_i - C_{i-1}}{\Delta t}$ Backward and

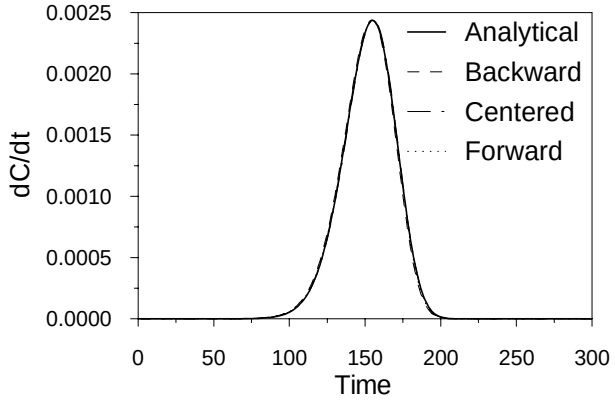

Figure 5.14: Analytical and Numerical (computed with three different derivative schemes) TFs in $x = 155$, $y = y_0$, 2-D.

$\frac{dC_i}{dt} = \frac{C_{i+1} - C_{i-1}}{2\Delta t}$ Central, in order to find the optimal scheme.

In Figure 5.14, the TF values, in $x = 155$ and $y = y_0$ node, computed through those different numerical schemes are shown and compared with those obtained from the analytical formulation equation (5.16); it can be seen that the agreement among the lines is fair good; the time step used is $\Delta t = 1$. Then, the geostatistical methodology (**constrained case**) is applied to recover the release history given the concentration measurements, i.e. the $\mathbf{z}$ vector, and the numerically obtained TF (backward scheme derivative). The results of such comparison are shown in the Figures 5.18, 5.19 and 5.20. In Figures 5.18 and 5.19 the true release function and those computed through the analytical and numerical TFs, respectively, are shown versus time. The results of Figure 5.19 are quite satisfactory: just a few differences are present among the lines. By inspection of the estimate error variance (Figures 5.18 and 5.19), a greater width of the interquantile strip can be observed for the case of numerical TF. This fact is due both to the unavoidable errors coming from the numerical modeling of the dispersion process and to the numerical computation of the time derivative of the breakthrough curve.

5.2.6 Heterogeneous 2-D Flow and Transport - Numerical TF only

Three tests are examined to check the performance of the outlined methodology in strongly heterogeneous fields. In cases 1 and 2 the source of the non-uniformity of the flow is the conductivity heterogeneity of the porous media; in case 3 a pumping well is active in the heterogeneous conductivity field. Due to the non uniformity of the flow the TFs do not have analytical expressions and, for the recovering of the release history, it is necessary to follow the new numerical procedure described in the previous sections.

A numerical model with rectangular shape (250 m long and 50 m wide, thickness equal to 10 m) has been built (Figure 5.21). The domain is discretized using a grid of cells of 2 m $\times$ 2 m dimensions, (125×25 cells) in the XY plane and one layer in the vertical direction.

About the boundary conditions, along two sides the head is fixed: upstream (East side) $h = 10$ m and downstream (West side) $h = 7.5$ m so that the flow is driven in the negative x direction; the other sides are impervious boundaries, see Figure 5.22. The hydraulic conductivity field is

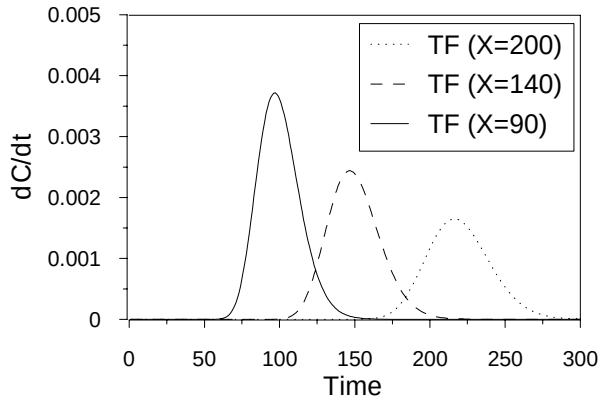

Figure 5.15: Numerical TF evaluated in three different locations using the backward methodology, 2-D.

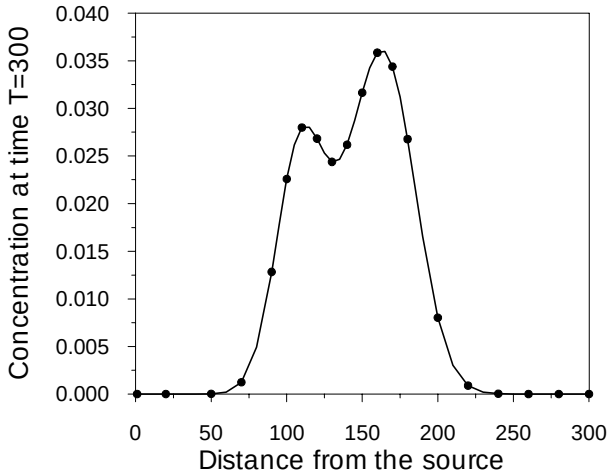

Figure 5.16: Concentration at time $T = 300$ and measurements locations (marks), along $y = y_0$, in the 2-D homogeneous case.

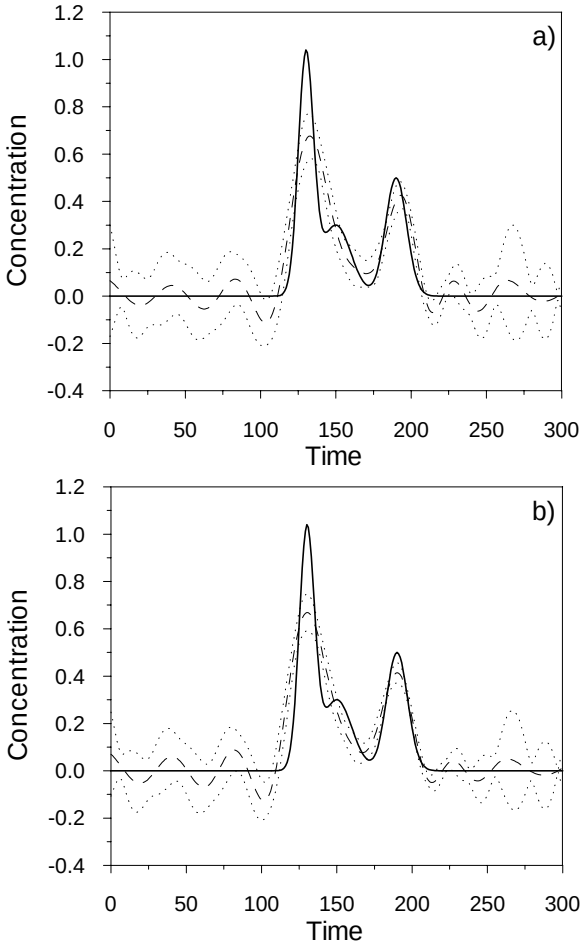

Figure 5.17: Unconstrained 2-D homogeneous case with Gaussian covariance: the true solution (solid line), best estimate (dashed line), and approximate 95% confidence interval (dotted line); a) Analytical, b) Numerical.

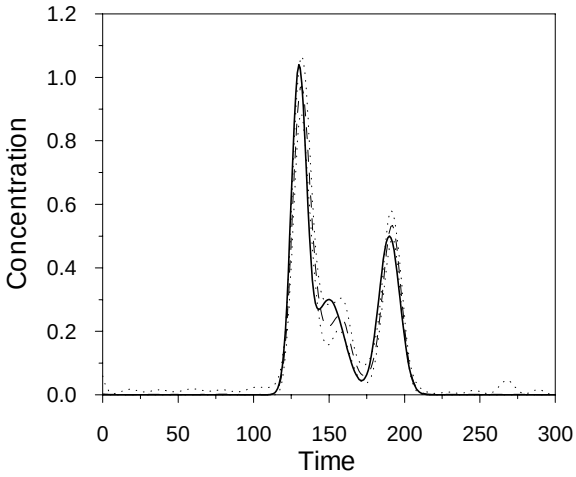

Figure 5.18: Constrained 2-D homogeneous case with Gaussian covariance: the true solution (solid line), best estimate (dashed line), and approximate 95% confidence interval (dotted line); Analytical TF.

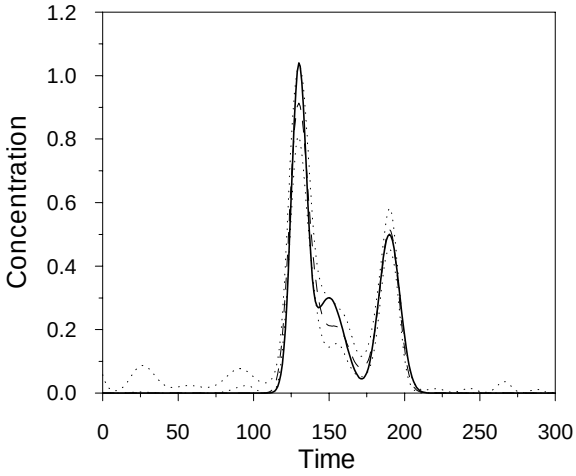

Figure 5.19: Constrained 2-D homogeneous case with Gaussian covariance: the true solution (solid line), best estimate (dashed line), and approximate 95% confidence interval (dotted line); Numerical TF.

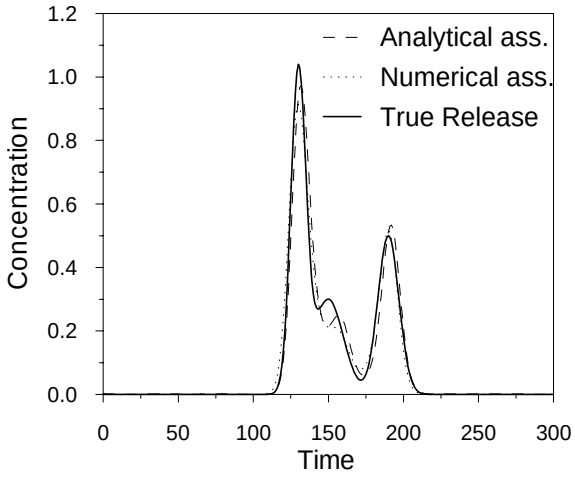

Figure 5.20: Analytical-Numerical assessment of the source release history; 2-D constrained case.

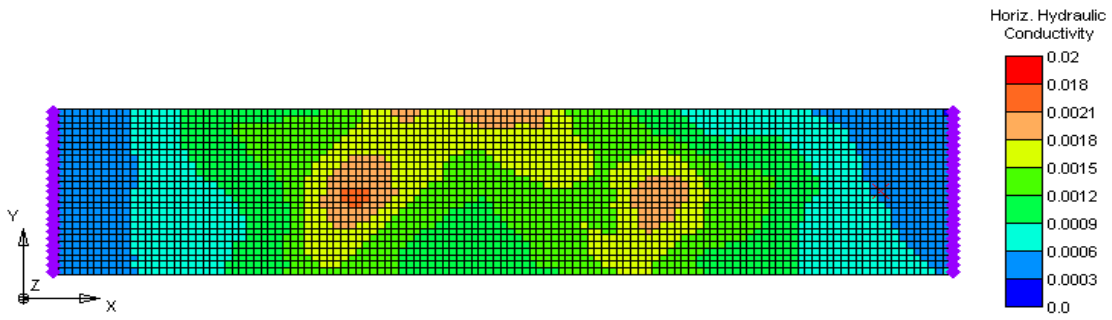

Figure 5.21: Grid of the numerical model with heterogeneous conductivity [m/s] ($\sigma_Y^2 = 0.22$).

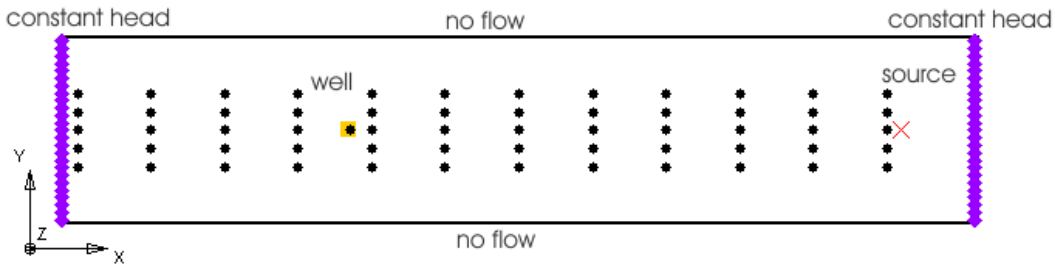

Figure 5.22: Scheme of the problem and monitoring grid; the well regards the case 3.

generated by a stochastic process with mean equal to 0.001 m/s. The obtained mean Darcy velocity is $1{\cdot}10^{-5}$ m/s. The longitudinal and transversal dispersivity are assumed constant ($\alpha_L = 1$ m and $\alpha_T = 0.1$ m). The release source is located in $x = 229$ m and $y = 25$ m The simulation of the transport process is carried out for 300 time intervals (ΔT) of 18000 s each. All the results are made dimensionless in the ratio with the characteristic time ΔT, length $\Delta l = \Delta x = \Delta y = 1$ m and concentration $C_0 = 1$ mg/L.

Case 1

In the first case of this section the conductivity field is weakly heterogeneous: the variance of the log-conductivity ($Y = lnK$) is equal to $\sigma_Y^2 = 0.22$. Following the outlined procedure a preliminary run was performed in order to calculate the breakthrough curves to be used for the computation of the numerical TFs. Then a new simulation with the injection, in the source location, of a pollutant release with time behavior given by equation (5.12) was carried out. The concentration data are monitored at the dimensionless time $T = 300$ in the nodes of a regular grid of 5 rows along the x direction and 12 columns along the y direction (Figure 5.22); the monitoring grid sides are $\Delta Lx = 20$ and $\Delta Ly = 5$. On the basis of the experience gained in previous studies, only some (24) data are used in the recovery procedure: 12 on $y = 25$ m (downstream the release source) and 12 on $y = 30$ m. Points where it is likely to detect higher concentration value are chosen as measurement points. The results of the recovery procedure are shown in Figure 5.23; the release is satisfactory reproduced and also the inter-quantile range is acceptable.

Case 2

In the Case 2 the procedure is applied to a conductivity heterogeneous field with $\sigma_Y^2 = 1.00$, a greater degree of heterogeneity; the monitoring network is the same of the Case 1. As required by the numerical procedure outlined in previous sections, a preliminary run was performed in order to calculate the breakthrough curves useful for the computation of the numerical TFs. Then a new simulation with the injection, in the source location, of a pollutant release with the trend described by the equation (5.12) has been carried out. The results of the procedure are shown in Figure 5.25; again the new procedure outlined by the writer seems to give a fair good recovered history, even if the higher degree of heterogeneity affects remarkably the non uniformity of the flow. The trends shown in Figures 5.23 and 5.25, respectively for $\sigma_Y^2 = 0.22$ and $\sigma_Y^2 = 1.00$, look similar. This fact can be explained considering that Case 1 and Case 2 present the same degree of difficulty with

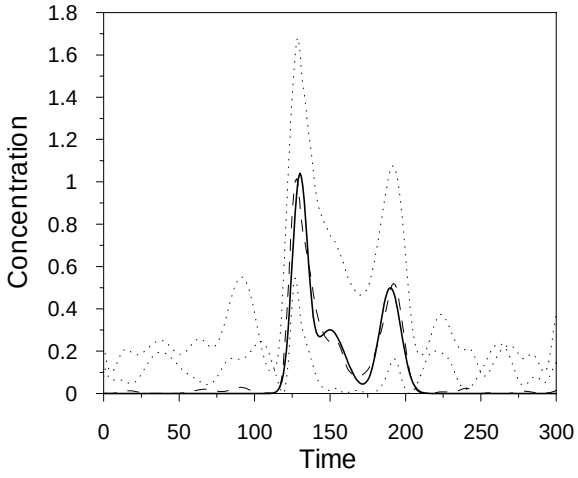

Figure 5.23: Constrained 2-D heterogeneous Case 1 ($\sigma^2_Y = 0.22$): the true solution (solid line), best estimate (dashed line), and approximate 95% confidence interval (dotted line).

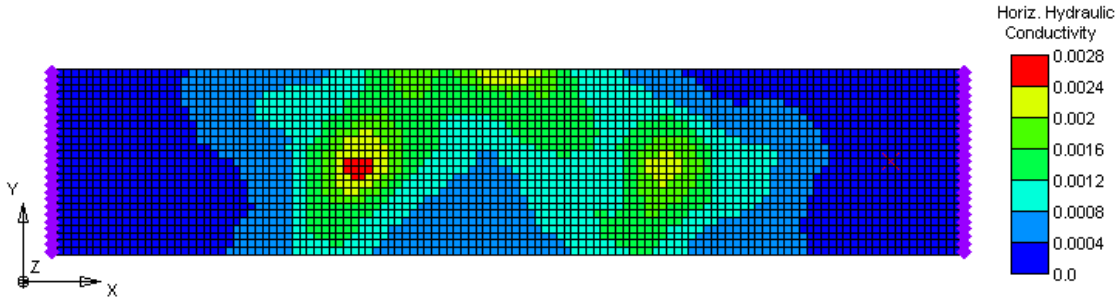

Figure 5.24: Grid of the numerical model with heterogeneous conductivity [m/s], ($\sigma^2_Y = 1.00$).

respect of the inverse problem in hand, once that the TFs are obtained, since the role of pore scale dispersion is the same. Conductivity heterogeneity, in fact, distorts the solute trajectories from the paths of the homogeneous case but does not increase dispersion phenomena.

Case 3

The Case 3 is a strongly non uniform flow: the heterogeneous conductivity field of Case 2 ($\sigma^2_Y = 1$) is considered but a pumping well is located downstream in $x = 79$ m, $y = 25$ m. The pumping discharge of the well is 1/5 of the total groundwater discharge of Case 2. Again the preliminary computations useful for the setting up of the numerical TFs have been carried out. Then the concentration values in 24 nodes have been considered at the time of $T = 300$ after the injection of the release (5.12) in the pollutant source. The points with highest maximum value of the computed TF have been selected. As expected, the most significant points are those located between the

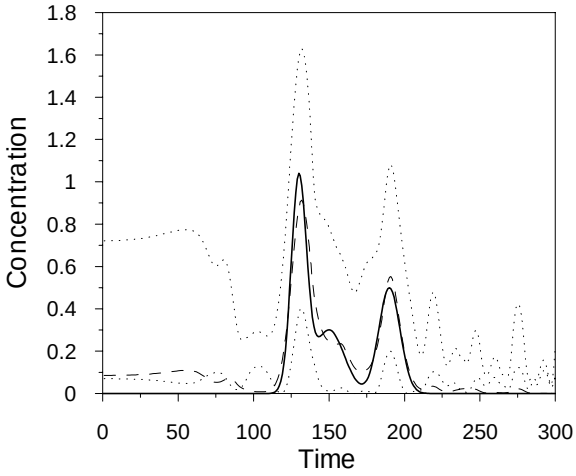

Figure 5.25: Constrained 2-D heterogeneous Case 2 ($\sigma^2_{Y'} = 1.00$): the true solution (solid line), best estimate (dashed line), and approximate 95% confidence interval (dotted line).

source and the pumping well at $y = 20$ m, 25 m, 30 m. The results of the recovery process for the Case 3 are shown in Figure 5.27.

Some spurious numerical oscillations are present but, in the whole, considering also the trend of the interquantile lines, the release is quite well recovered even if not as well as the Case 1 and 2 (5.23 and 5.25 respectively). The deterioration (in comparison with Cases 1 and 2) of the recovery results can be ascribed, rather than to the greater non uniformity of the flow, to the action of the well that increases the mean velocity increasing the dispersion coefficients as well. The Figure 5.26 shows the plume after 300 time units due to input step function, the heterogeneities and the pumping well; it is possible to see clearly the attracting effect of the well.

5.3 Conclusions

A new methodology for the numerical computation of the transfer functions (TFs) for problems where it is impossible to develop analytical solutions is here proposed and tested with good results in different conditions.

In the first tests the procedure has been applied to simple cases (from literature) with analytical solution in order to compare the differences using the analytical or numerical transfer function (TF). The scatters seem to be very moderate even if, in the opinion of the writer, the adopted numerical technique could be improved in order to reduce the highlighted discrepancies. In the other tests, concerning heterogeneous groundwater field with strongly non uniform flow and no analytical solution for TFs, the results were analyzed comparing the true release history with the restored one. Also in these cases the results can be considered widely satisfactory encouraging the extension of the new procedure to future application of practical importance.

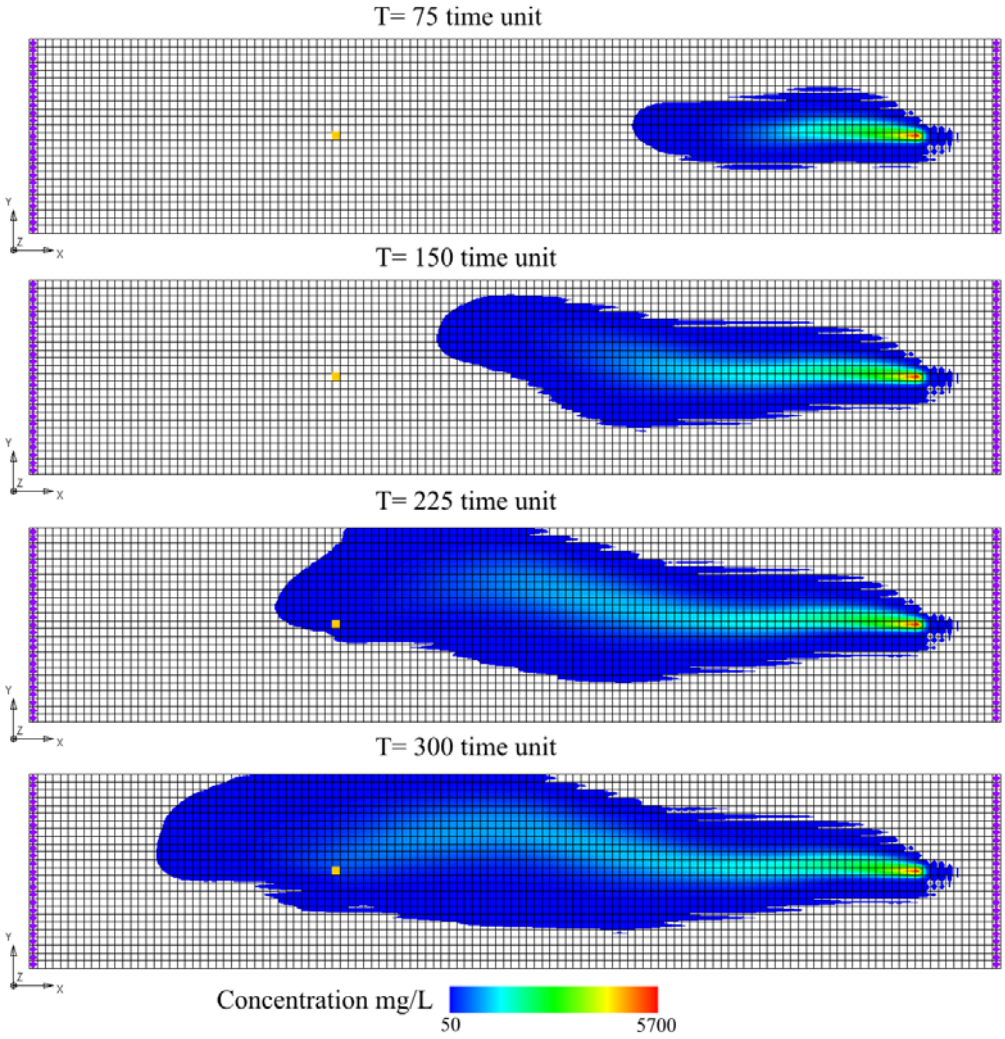

Figure 5.26: Evolution of the plume, case 3.

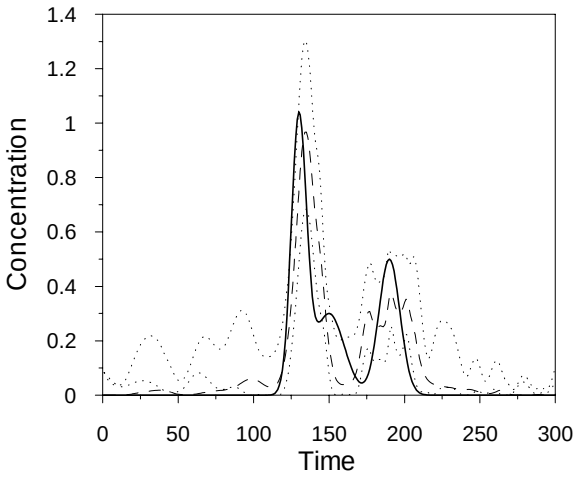

Figure 5.27: Constrained 2-D heterogeneous Case 3 ($\sigma_Y^2 = 1.00$, with pumping well): the true solution (solid line), best estimate (dashed line), and approximate 95% confidence interval (dotted line).

Chapter 6

Applications and new Developments in Aquifer Parameter Estimation

6.1 Introduction

One of the main problems in modeling groundwater flow processes is the estimation of hydraulic parameters such as the transmissivity. These parameters are normally estimated from scarce data because data is often difficult and costly to obtain. The accuracy of the estimate depends on the number of measurements, their locations in the study area, the observation error and the sensitivity of the observed quantity to, for instance, the characteristics of the transmissivity field.

This work applies the geostatistical approach (see section 4.2 for a summary of the methodology) to recovering the transmissivity field from hydraulic head measurements in the case of strong heterogeneity. The geostatistical technique to inverse modeling is a stochastic Bayesian approach, and over the last few decades, it has been applied in many studies in subsurface hydrology. Some examples are, the estimation of hydraulic conductivity. The geostatistical approach with the aim at recover the log-transmissivity field has been studied by several authors [*Kitanidis and Vomvoris*, 1983; *Hoeksema and Kitanidis*, 1984, 1989; *Dagan*, 1985; *Gutjahr and Wilson*, 1989; *Gutjahr et al*, 1994; *Harvey and Gorelick*, 1995; *Kitanidis*, 1996; *Rubin and Dagan*, 1987a, 1987b, 1988, 1989; *Yeh et al.*, 1995; *Yeh et al.* 1996; *Yeh and Zhang* 1996; *Yeh* et al. 2002; *Li and Yeh*, 1999; *Zhang and Yeh*, 1997; *Zimmerman et al.,* 1998; *Nowak et al.*, 2003; *Nowak and Cirpka*, 2004].

Stochastic approaches are often used in hydrology, and some literature regarding MCMC methods (which are used here), can be found in the works here briefly described. *Harter and Yeh* [1996a, 1996b] applied the Monte Carlo simulation to predict the solute transport; *Gòmez Hernández et al.* [1997] used the sequential simulations to generate conditional transmissivity fields; *Hanna and Yeh* [1998] developed a new co-conditional Monte Carlo simulation technique to derive mean and variance of transmissivity, head and Darcy's velocity, based on measurement of transmissivity and head in heterogeneous confined aquifer in steady state conditions; *Tamminen and Kyrölä* [2001] used Markov Chain Monte Carlo (MCMC) method for computing posterior distributions for nonlinear inverse problems; *Feyen et al.* [2003] used Monte Carlo sampling to generate transmissivity fields to delineate well capture zones; *Bates and Campbell* [2001] applied MCMC using the MH

algorithm to evaluate the parameters in conceptual rainfall-runoff models; *Marshall et al* [2004] compared four MCMC sampling algorithm in the context of rainfall-runoff modeling; *Vrugt et al.* [2003] proposed an MCMC sampler which was suited to infer the posterior distribution of hydrologic model parameters; *Chen et al.* [2004] used MCMC methods to estimate concentrations using ground-penetrating radar tomographic data.

This work follows the one about the quasi-linear (QL) methodology of *Kitanidis* [1995]; the aim is to test the QL methodology in a strongly nonlinear problem (e.g. that of a transmissivity field of high heterogeneity) and to propose a new approach to solve such problems. The QL methodology underestimates the variance of the recovered log-transmissivity field, and also may not be efficient in the case of strongly nonlinear problems. Therefore, in this work the QL approach is amended with a procedure that generates equally likely conditional realizations. Then an acceptance-rejection scheme known as the Metropolis-Hastings algorithm was implemented. This methodology was applied by *Michalak and Kitanidis* [2003, 2004a] to recover the release history of a pollutant source with excellent results.

The following section explains how the proposed methodology works and how it is implemented. Note that the distribution of transmissivity is usually found to be lognormal, and here this hypothesis has been assumed.

6.2 Optimization Routine

The first and most important improvement applied to the quasi-linear approach is the optimization routine; it consists of optimizing the step from an iteration to the next one. In fact, for strongly nonlinear field, the new estimate $\hat{\mathbf{s}}$ could be too different from the previous one and instead of converging to a value, the estimate could oscillate. A line search, see equation (6.1), along the line defined by the two values (the starting one and the new one that is obtained through the Gauss-Newton iteration) can be adopted to ensure monotonic improvement of the objective function. The parameter δ defines a point along this line:

$$\hat{\mathbf{s}} = \mathbf{s}_{old} \cdot \delta + \mathbf{s}_{new} \cdot (1 - \delta) \tag{6.1}$$

where $\mathbf{s}_{old}$ is the estimate at the previous iteration and $\mathbf{s}_{new}$ is evaluated by the equation (3.16) or (3.18).

A few tests regarding the values of the parameter δ were carried out. At first, δ was allowed to vary from -3 to $+3$; it was noticed that the optimization routine never uses negative values of δ, so the range has been reduced to $-0.1 \leq \delta \leq 1.1$. The reduction of the range allows to save computational resources and to converge to a value quickly. The value of δ is chosen by minimizing iteratively the following objective function:

$$objf = \left(\mathbf{z} - h\left(\hat{\mathbf{s}}\right)\right)^T \mathbf{R}^{-1} \left(\mathbf{z} - h\left(\hat{\mathbf{s}}\right)\right) + \hat{\mathbf{s}}^T \mathbf{G} \hat{\mathbf{s}} \tag{6.2}$$

This procedure ensures that the value of the objective function (6.2) decreases monotonically from an iteration to the next one. This minimization algorithm is based on the `fminbnd` function of Matlab. This routine minimizes a function (for example the one in equation (6.2)) of one variable, δ, on a fixed interval, $-0.1 \leq \delta \leq 1.1$. The Matlab routine has been chosen because it is highly efficient; it returns a value δ that is a local minimizer of the function that is described in (6.2) in the fixed interval ($-0.1 \leq \delta \leq 1.1$). The algorithm of the routine is based on Golden Section search and parabolic interpolation. On the downside, the Matlab routine tends to converge to local

minima, it often exhibits slow convergence when the solution is on a boundary of the interval, it only handles real variables, and the function to be minimized must be continuous. None of these disadvantages are important in this application.

6.3 Conditional Realization Applied to QL Approach

The procedure to generate conditional realizations is discussed in the works of *Gutjahr et al.* [1994], *Kitanidis* [1995, 2004], *Michalak* [2003] and *Michalak and Kitanidis* [2003, 2004a]. In this section the procedure to create and to accept the conditional realizations is summarized.

The first step is to generate an unconditional realization $\mathbf{s}_{u,i}$ with zero mean and covariance matrix $\mathbf{Q}$ and a realization of the measurement error $\mathbf{v}_i$ with zero mean and covariance matrix $\mathbf{R}$. The Choleski decomposition could be used to generate $\mathbf{s}_{u,i}$; this is not the most efficient method but it is the most straightforward. The first step is to generate $\mathbf{C}$ from $\mathbf{Q} = \mathbf{C}\mathbf{C}^T$ and then compute

$$\mathbf{s}_{u,i} = \mathbf{X}\boldsymbol{\lambda} + \mathbf{C}\mathbf{u}_i \tag{6.3}$$

where $\mathbf{u}_i$ is a vector of independent, identically distributed normal variables, $\boldsymbol{\lambda}$ could be set to 0 because the conditional realization finally produced is unaffected by the values of $\boldsymbol{\lambda}$. The realizations of $\mathbf{s}_{u,i}$ and $\mathbf{v}_i$ are generated independently. Then starting from this unconditional realization the conditional one can be found minimizing

$$\left(\mathbf{y} + \mathbf{v}_i - \mathbf{h}(\mathbf{s})\right)^T \mathbf{R}^{-1} \left(\mathbf{y} + \mathbf{v}_i - \mathbf{h}(\mathbf{s})\right) + \left(\mathbf{s}_{c,i} - \mathbf{s}_{u,i}\right)^T \mathbf{G} \left(\mathbf{s}_{c,i} - \mathbf{s}_{u,i}\right) \tag{6.4}$$

with respect to $\mathbf{s}_{c,i}$. Where

$$\mathbf{G} = \mathbf{Q}^{-1} - \mathbf{Q}^{-1}\mathbf{X}\left(\mathbf{X}^T\mathbf{Q}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Q}^{-1}$$

The optimum $\mathbf{s}_{c,i}$ gives the conditional realization. This results in an iterative method similar to iterative cokriging:

$$\mathbf{s}_{c,i} = \mathbf{s}_{u,i} + \mathbf{X}\overset{\wedge}{\boldsymbol{\beta}} + \mathbf{Q}\mathbf{H}^T\boldsymbol{\xi} \tag{6.5}$$

$\boldsymbol{\xi}$ and $\boldsymbol{\beta}$ may be obtained from the solution of the system of $m + p$ linear equations with $m + p$ unknowns:

$$\begin{bmatrix} \widetilde{\boldsymbol{\Sigma}} & \mathbf{H}\mathbf{X} \\ (\mathbf{H}\mathbf{X})^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi} \\ \overset{\wedge}{\boldsymbol{\beta}} \end{bmatrix} = \begin{bmatrix} \mathbf{y} + \mathbf{v}_i - \mathbf{h}\left(\widetilde{\mathbf{s}}_{c,i}\right) + \mathbf{H}\left(\widetilde{\mathbf{s}}_{c,i} - \mathbf{s}_{u,i}\right) \\ \mathbf{0} \end{bmatrix} \tag{6.6}$$

that procedure has to be repeated until convergence.

After the creation of the conditional realization an acceptance/rejection method has been applied. Instead of weighting the samples, it could be useful a procedure to screen them; some of the samples are rejected, but the ones accepted are equiprobable samples from the distribution p (pdf of a random variable $\mathbf{s}$). The steps of the used procedure ($0 \leq h(\mathbf{s}_{c,i}) \leq 1$) are the following:

1. Generate $\mathbf{s}_{u,i} \sim P$ (P is a distribution close to the target distribution) and then $\mathbf{s}_{c,i}$,
2. Generate $\mathbf{u}_i \sim Uniform[0, 1]$,
3. if $\mathbf{u}_i < h\left(\mathbf{s}_{c,i}\right)$ accept $\mathbf{s}_{c,i}$; else, throw it out,
4. Go back to step 1.

where

$$h\left(\mathbf{s}_c\right) = \frac{p\left(\mathbf{s}_c\right)}{MP\left(\mathbf{s}_c\right)} \tag{6.7}$$

M is a positive constant that satisfies:

$$M \geq \sup_{\mathbf{s}} \frac{p\left(\mathbf{s}_c\right)}{P\left(\mathbf{s}_c\right)}$$

in order to make sure that $h\left(\mathbf{s}_{c,i}\right)$ does not exceed one. Each candidate sample is generated using the quasi-linearization technique with an unconditional realization that depends on the unconditional realization of the last accepted sample. Consider that an unconditional realization $\mathbf{s}_{u,i}$ has been accepted. The (i+1)-th unconditional realization $\mathbf{s}_{u,i+1}$ is correlated to $\mathbf{s}_{u,i}$ by the following equation:

$$\mathbf{s}_{u,i+1} = \rho \mathbf{s}_{u,i} + \alpha \mathbf{u}_{i+1} \tag{6.8}$$

where ρ is a positive coefficient (for stability $\rho < 1$); α a real coefficient and $\mathbf{u}_{i+1}$ is an unconditional realization generated independently from previous realizations.

$$\alpha = \sqrt{1 - \rho^2}$$

For strongly nonlinear problems only a very low percentage of candidate will be accepted.

The Metropolis Hastings algorithm (MH) as acceptance/rejection method (see *Chib and Greenberg* [1995] and *Kitanidis* [2004] for a detailed description) is chosen in this work. The idea in MH algorithms is to generate a candidate sample that is dependent to the last accepted one and to accept or reject it based on a criterion that is the combination of two factors: (1) The ratio of the posterior pdf of the candidate to the last. (2) The ratio of the chance of a move from the candidate to the last. Methods that use MH algorithm involve the use of candidate pdf to obtain candidate realizations of the unknown function. The objective function is then used to accept or reject these realizations. The accepted realization can be shown to be equally likely samples from the pdf of interest. The algorithm follows this steps:

1. Initialization accepting some reasonable $\mathbf{s}_{u,o}$ (equation 6.3),
2. Generate candidate sample $\mathbf{s}_{u,i+1}$ based on a procedure that uses the last accepted one, $\mathbf{s}_{u,i}$ (see equation (6.8)),
3. Generate the conditional realization $\mathbf{s}_{c,i+1}$ using the equation (6.5),
4. Generate $\mathbf{u}_{i+1} \sim Uniform[0,1]$,
5. If $\mathbf{u}_{i+1} < \varsigma\left(\mathbf{s}_{c,i+1}|\mathbf{s}_{c,i}\right)$ accept $\mathbf{s}_{c,i+1}$; else accept the previous value $\mathbf{s}_{c,i}$,
6. Go back to step 2.

p'' is the target posterior distribution within a normalizing constant. The transition probability of $\mathbf{s}_{c,i+1}$ given $\mathbf{s}_{c,i}$, $q\left(\mathbf{s}_{c,i}|\mathbf{s}_{c,i+1}\right)$, is determined by the scheme used to move from the last to the candidate. The probability of acceptance is defined as :

$$\varsigma\left(\mathbf{s}_{c,i+1}|\mathbf{s}_{c,i}\right) = \left\{ \frac{p''\left(\mathbf{s}_{c,i+1}\right) q\left(\mathbf{s}_{c,i}|\mathbf{s}_{c,i+1}\right)}{p''\left(\mathbf{s}_{c,i}\right) q\left(\mathbf{s}_{c,i+1}|\mathbf{s}_{c,i}\right)}, 1 \right\}$$

The flow chart in Figure 6.1 summarizes the adopted procedure.

6.3.1 Metropolis Hastings Applied to the Recovery of the Release History

Michalak [2003] and *Michalak and Kitanidis* [2004a] applied the MH algorithm to unconstrained conditional realizations generated using linear geostatistical inverse modeling constrained to be non negative. The methodology is summarized in the following.

A list of the used symbols is given below: $\mathbf{s}_{cc,c}$ candidate conditional constrained realization, $\mathbf{s}_{cc,l}$ l-th accepted conditional constrained realization, $q(\mathbf{s}_{cc,.}|\mathbf{s}_{cc,.})$ the transition probability from one constrained conditional realization to another, $U(0,1)$ a uniform distribution in the range $[0,1]$ and $\varsigma(\mathbf{s}_{cc,c}|\mathbf{s}_{cc,l})$ as the acceptance probability of $\mathbf{s}_{cc,c}$ given the previous accepted realization $\mathbf{s}_{cc,l}$. In this context a constrained realization is one that is everywhere nonnegative and a conditional realization is one that has been conditioned on the available data. Initializing the chain with an arbitrary constrained realization $\mathbf{s}_{cc,0}$ the MH algorithm proceeds as follows:

1. for $l = 1, N$,
2. Generate $\mathbf{s}_{cc,c}$ from $q(\mathbf{s}_{cc,c}|\mathbf{s}_{cc,l})$ and u from $U(0,1)$,
3. if $u \leq \varsigma(\mathbf{s}_{cc,c}|\mathbf{s}_{cc,l})$, $\mathbf{s}_{cc,l+1} = \mathbf{s}_{cc,c}$,
4. Otherwise $\mathbf{s}_{cc,l+1} = \mathbf{s}_{cc,l}$,
5. Return the values $\{\mathbf{s}_{cc,1}, \mathbf{s}_{cc,2}, ..., \mathbf{s}_{cc,N}\}$.

The probability of acceptance is defined as :

$$\varsigma(\mathbf{s}_{cc,c}|\mathbf{s}_{cc,l}) = \min\left\{\frac{p''(\mathbf{s}_{cc,c})q(\mathbf{s}_{cc,l}|\mathbf{s}_{cc,c})}{p''(\mathbf{s}_{cc,l})q(\mathbf{s}_{cc,c}|\mathbf{s}_{cc,l})}, 1\right\}$$

where $p''(\mathbf{s}_{cc,c})$ is the posterior probability distribution. A new candidate conditional constrained realization is obtained by the first obtaining an unconditional realization $\mathbf{u}_c$ with zero mean and covariance $\mathbf{Q}$ in the standard geostatistical manner. The unconditional realizations used to obtain the candidate conditional realizations are sequentially correlated:

$$\mathbf{s}_{uu,c} = \rho\mathbf{s}_{uu,l} + \alpha\mathbf{u}_c$$

where $\mathbf{s}_{uu,l}$ is the unconditional realization used in the generation of the last accepted realization and $\alpha = \sqrt{1-\rho^2}$, $(0 < \rho < 1)$. A conditional unconstrained realization $\mathbf{s}_{cu,c}$ is generated from $\mathbf{s}_{uu,c}$ using the geostatistical procedure; the candidate conditional constrained realization $\mathbf{s}_{cc,c}$ is obtained by applying the method of Lagrange multipliers to $\mathbf{s}_{cu,c}$. The candidates, generated in this way, are accepted or rejected based on their posterior probability relative to that of the last accepted realization $\mathbf{s}_{cc,l}$, according to $\varsigma(\mathbf{s}_{cc,c}|\mathbf{s}_{cc,l})$. The posterior probability distribution is

$$p''(\mathbf{s}_{cc,.}) \propto \exp\left[-\frac{1}{2}\left(\mathbf{y} - \mathbf{H}\mathbf{s}_{cc,.}\right)^T \mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{H}\mathbf{s}_{cc,.}\right) - \frac{1}{2}\left(\mathbf{s}_{cc,.} - \mathbf{X}\boldsymbol{\beta}\right)^T \mathbf{Q}^{-1}\left(\mathbf{s}_{cc,.} - \mathbf{X}\boldsymbol{\beta}\right)\right]$$

The transition probability $q(\mathbf{s}_{cc,.}|\mathbf{s}_{cc,.})$ is approximated by $q(\mathbf{s}_{uu,.}|\mathbf{s}_{uu,.})$.

$$q(\mathbf{s}_{uu,c}|\mathbf{s}_{uu,l}) \propto \exp\left[-\frac{1}{2}\left(\mathbf{s}_{uu,c} - \rho\mathbf{s}_{uu,l}\right)^T \frac{\mathbf{Q}^{-1}}{\alpha^2}\left(\mathbf{s}_{uu,c} - \rho\mathbf{s}_{uu,l}\right)\right]$$

The chain is run until the probability space has been appropriately entirely sampled.

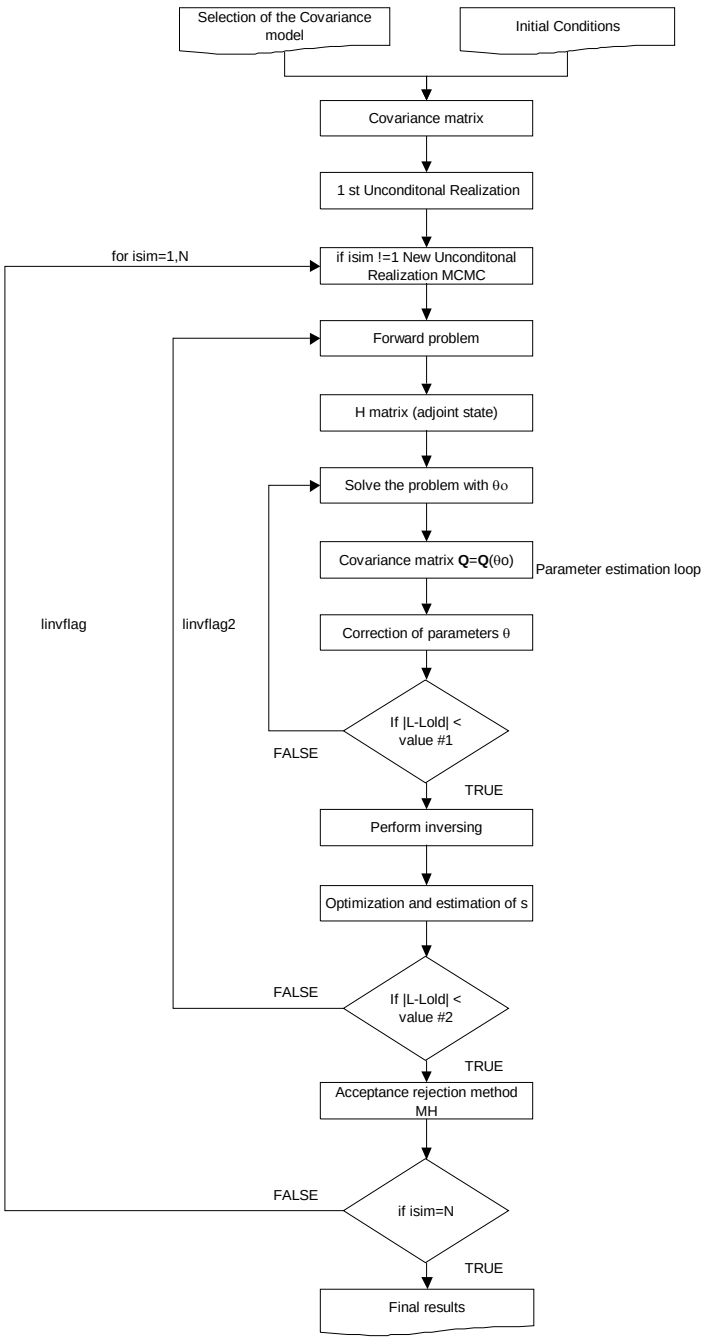

Figure 6.1: Flow chart of the methodology that applies the conditional realizations and the Metropolis-Hastings algorithm.

6.4 Applications of the Quasi-Linear Methodology

The applications proposed involve the finding of the transmissivity from hydraulic head data in a 2-D aquifer with steady flow. The governing equation (from Darcy's law and continuity) is

$$\frac{\partial}{\partial x}\left(T\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(T\frac{\partial h}{\partial y}\right)=-N$$

where T [L^2T^{-1}] is the transmissivity, h [L] is the head and N [LT^{-1}] is the specific recharge rate. The domain is rectangular ($L_1 \times L_2$, $L_1 = 1000$ m, $L_2 = 750$ m) and described by a grid of 24×32 cells; the head values at the western and eastern boundaries are constant ($h(0,y) = h_0 = 120$ m, and $h(L_1,y) = h_1 = 110$ m), respectively, and the northern and southern boundaries are considered impermeable ($\frac{\partial h}{\partial y} = 0$ at $y = 0$ and $y = L_2$). See Figure 6.2. A recharge rate ($N = 0.001$ md^{-1}) is added to the whole surface of the model.

In all problems that will be discussed, all quantities are made dimensionless in order to use an existing forward model (`mkls.m`, see *Kitanidis* [2004]). The dimensions of the model are divided by L_1 so they become $L_1^* = 1$ and $L_2^* = 0.75$. The dimensionless heads are evaluated using the following expression: $h^* = \frac{h - h_1}{h_0 - h_1}$, so the new western and eastern boundaries become: $h^*(0,y) = h_0^* = 1$, and $h^*(L_1,y) = h_1^* = 0$. The recharge contribution is transformed using the following relationship: $N^* = \frac{N \cdot t_0}{h_0 - h_1} = 0.0001$, where $t_0 = 1$ day. The transmissivity is made dimensionless as: $T^* = \frac{T}{L_1^2} t_0$. The results are shown in dimensional form in order to make comparison to real cases.

The objective of the test is to estimate the transmissivity field starting from few head measurements. At this aim a synthetic case was implemented: at first a transmissivity field was created then the forward problem was solved and few head measurements were sampled. Finally the inverse problem, starting from the head measurements, was analyzed to recover the transmissivity field.

The original ("true") log-transmissivity field, $\mathbf{s}$, (see Figures 6.3, 6.4 and 6.5) was created as an unconditional realization with zero mean and covariance matrix $\mathbf{Q}$, using the Cholesky decomposition (equation (6.3)) and forcing the mean to the value of 2.5: $\mathbf{s} = 2.5 + \mathbf{X}\boldsymbol{\lambda} + \mathbf{Cu}$; where $\boldsymbol{\lambda}$ =0, $\mathbf{Q} = \mathbf{CC}^T$ and $\mathbf{u}$ is a vector of independent, identically distributed normal variables. The cubic generalized covariance matrix: Q_{ij}=θh_{ij}^3 ($\theta = 4.386$) was assumed. Additional log-transmissivity fields, with an increasing contrast, were created from the original one ($\mathbf{s}$): **2s** ($\mathbf{2s} = 2.5 + 2 \cdot \mathbf{Cu}$), **4s** ($\mathbf{4s} = 2.5 + 4 \cdot \mathbf{Cu}$), **6s**, **8s** and **10s** (see table 6.1 and Figures 6.6, 6.7, 6.8 and 6.9); all the studied log-transmissivity fields present a mean value of 2.5. The variance increases from 0.1020 ($\mathbf{s}$) to 10.200 (**10s**), see table 6.1. The last analyzed case can be considered as a strongly nonlinear case (the variance of the transmissivity field is $1.08 \cdot 10^9$). **10s** was considered an extreme log-transmissivity field because it represents, considering the transmissivity field, a variance of 8 order of magnitude that is a wide range. Finding in nature this variability in the transmissivity field can be possible studying a very wide area. Moreover this large variability is highly difficult to model using commercial codes as MODFLOW because they requires very refined grid; this implies also a lot of computation time.

The boundary conditions and the recharge flow rate were considered the same for all the studied cases. Considering the first log-transmissivity field ($\mathbf{s}$) it is possible to notice Figures 6.3, 6.5 and table 6.1 that the mean of the transmissivity is 12.843 m^{2}d^{-1}; assuming 10 m as thickness of the

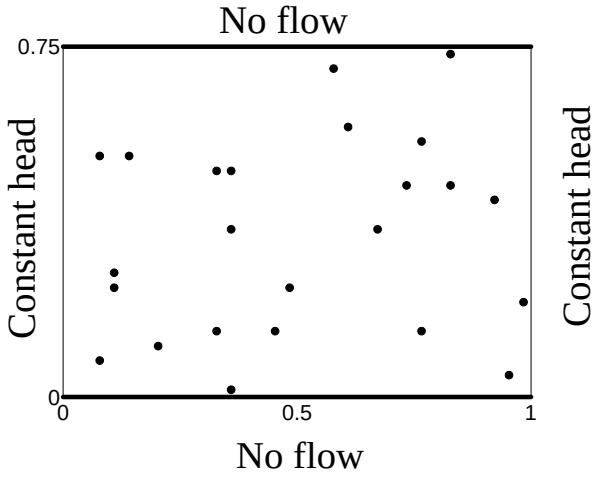

Figure 6.2: Domain of the problem and measurement locations.

Table 6.1: Statistics of the true fields

	Log-Transmissivity		Transmissivity	
	mean [**s**]	*var* [**s**]	*mean* [$\mathbf{T} = \exp(\mathbf{s})$] [$\mathrm{m^2 d^{-1}}$]	*var* [$\mathbf{T} = \exp(\mathbf{s})$]
s	2.5	0.1020	12.843	20.266
2s	2.5	0.4080	15.201	162.498
4s	2.5	1.6319	32.288	$6.79 \cdot 10^3$
6s	2.5	3.6718	119.86	$3.35 \cdot 10^5$
8s	2.5	6.5277	642.59	$1.85 \cdot 10^7$
10s	2.5	10.200	4113.2	$1.08 \cdot 10^9$

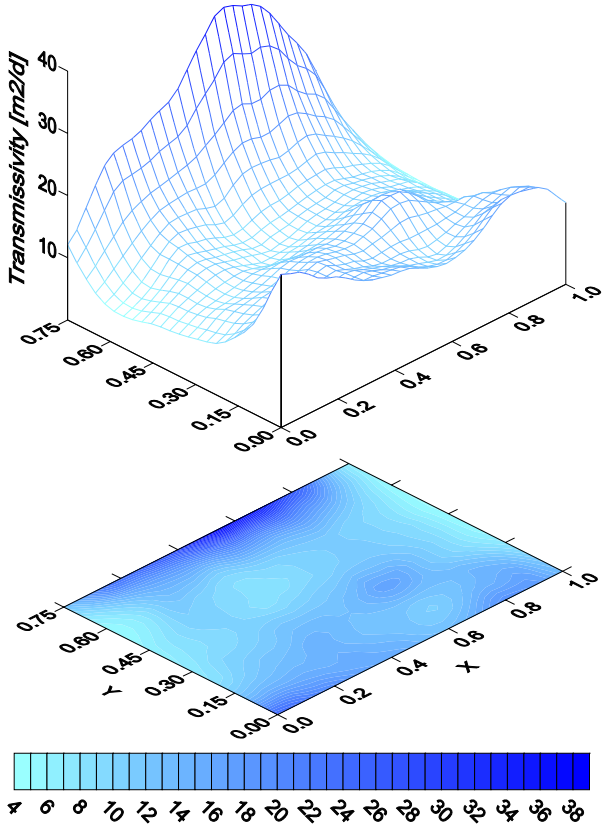

Figure 6.3: True transmissivity [m^2d^{-1}], **s** field.

aquifer, the resulting hydraulic conductivity can fit to a sand (1.284 m·d^{-1}). The head field that results from the forward problem, using as boundary and initial condition the ones described above, is shown in Figure 6.4. Considering the **10s**, the mean of the transmissivity field is 4113.2 m^2d^{-1}; if represents a high hydraulic conductivity value (411.3 m·d^{-1}) suitable for gravel. Analyzing the minimum and the maximum values it is possible to notice that the transmissivity goes from $2.8 \cdot 10^{-3}$ m^2d^{-1} to $4.7 \cdot 10^5$ m^2d^{-1}, due to the high variance ($1.08 \cdot 10^9$ m^4d^{-2}). This variability generates unrealistic heads, see Figure 6.9, because the recharge rate results a huge number respect to very low values of transmissivity. Analogous unrealistic heads were obtained also in the **6s** and **8s** fields. These unrealistic cases were analyzed in order to test the methodology in strongly non linear transmissivity fields.

The forward problem was solved for each transmissivity field and 24 hydraulic heads were measured in specific positions of the aquifer (vector of measurements **z**, see Figure 6.2) and noised by a measurement error. The measurement errors, dimensionless, were taken to be independent and identically distributed with four different variances of the error $\sigma_R^2 = 10^{-2}$, 10^{-4}, 10^{-6} and 10^{-8}; considering the dimensional problem, the measurement errors are 1, 0.1, and 0.01 and 0.001

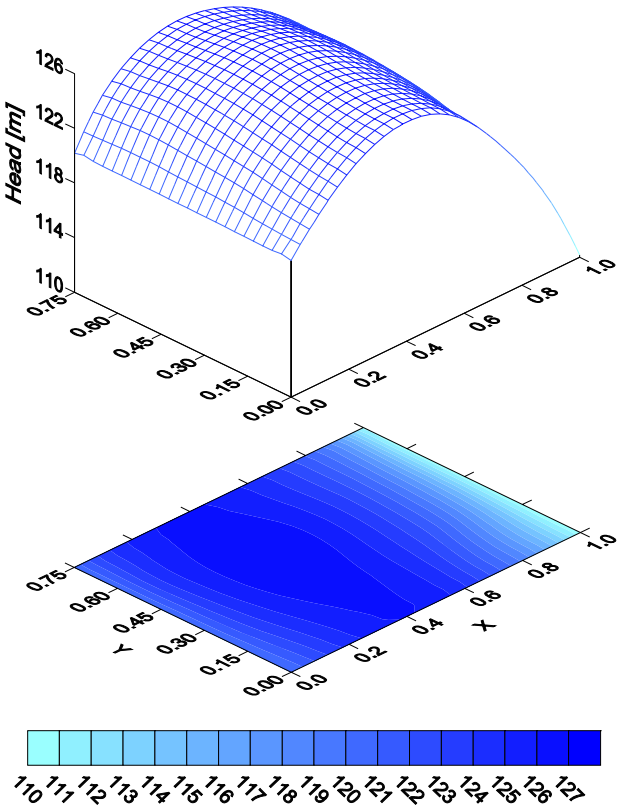

Figure 6.4: Head [m] related to the **s** transmissivity field.

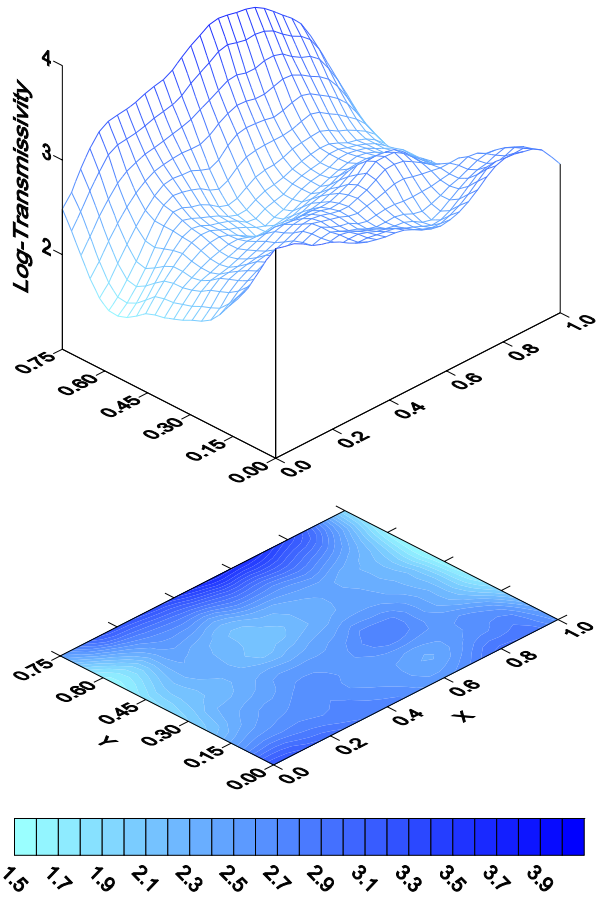

Figure 6.5: True log-transmissivity, **s** field.

m. Four different values of σ_R^2 were considered to highlight the influences of the measurement errors to the process and to make a comparison of the results. Assuming that the errors are taken to be independent and identically distributed with variance σ_R^2, the covariance matrix becomes $\mathbf{R} = \sigma_R^2 \cdot \mathbf{I}$, where $\mathbf{I}$ $(n \times n)$ is the identity matrix. The strongly nonlinear problems (**8s** and **10s**) were analyzed considering only a measurement error of $\sigma_R^2 = 10^{-2}$ and 10^{-4}, because a measurement error of 0.01 and 0.001 m added to head measurements related to these transmissivity fields resulted not significant. Analyzing the objective function (6.2) and comparing the results obtained using $\sigma_R^2 = 10^{-2}$ and 10^{-8}, it is possible to notice that the part of the objective function that represents the misform $(\left(\mathbf{z} - h\left(\hat{\mathbf{s}}\right)\right)^T \mathbf{R}^{-1} \left(\mathbf{z} - h\left(\hat{\mathbf{s}}\right)\right))$ is greater using $\sigma_R^2 = 10^{-2}$ than $\sigma_R^2 = 10^{-8}$, so the recovered field using the great error results flatter than using the small one.

The matrix $\mathbf{X}$ $(E[\mathbf{s}] = \mathbf{X}\boldsymbol{\beta})$ is a $n \times p$ matrix and $\boldsymbol{\beta}$ are are $p \times 1$ unknown coefficient. Using a linear generalized covariance function

$$Q_{ij} = -\theta h_{ij} \text{ and } p = 1$$

where h_{ij} means the distance between the point i and j. The matrix $\mathbf{X}$ is a known $n \times 1$ matrix and has coefficients equal to 1. Using the cubic generalized covariance function

$$Q_{ij} = \theta h_{ij}^3 \text{ and } p = 3$$

$\mathbf{X}$ has $n \times 3$ known coefficients; the first column consists of coefficients equal to 1, while the second and third column contains the spatial coordinates (x, y) of the grid nodes. In this work, only the results for the cubic covariance model are presented.

Assuming that the boundary condition of the problems in hand and the measurement errors are known it is possible to solve the inverse problems and to recover the log-transmissivity fields from the head measurements.

The log-transmissivity field was recovered using the QL method and also by sampling from the posterior to produce conditional realizations (CR) in order to make a comparison between the two methodologies and to find out the one that represents the true field in the better way.

A new improvement [*Nowak et al.*, 2003] was applied to the geostatistical process to speed up the multiplications between the covariance matrix $\mathbf{Q}$ and the other matrices/vectors. This methodology (described in chapter 7) uses the spectral method and the properties of the covariance matrix to reduce the computation time and to allow a very well defined computation grid (up to 2^{21} nodes, as reported in chapter 7).

A flat solution was assumed as starting values of the log-transmissivity fields for all the considered cases.

The first step of the work was the application of the quasi-linear procedure for each proposed cases (**s**, **2s**, ..., **10s**) to look for the limits of this procedure. The results of this process showed that the basic quasi-linear procedure was not so efficient; in fact it worked quite well for the **s** and **2s** cases (with weakly heterogeneous transmissivity field) but for the other cases (highly nonlinear field) it didn't converge to a solution. The process oscillated around the true solution and never converged to a value; this effect is due to great modifications from an iteration to the next one.

Then a parameter δ was introduced in the Gauss-Newton algorithm (see equation (3.14)) with the aim at avoiding oscillations in the convergence procedure. The new estimate is computed as follow:

$$\hat{\mathbf{s}}_i = \delta \hat{\mathbf{s}}_{i-1} + (1 - \delta)\hat{\mathbf{s}} \tag{6.9}$$

where $\hat{\mathbf{s}}$ is evaluated using the geostatistical procedure (equation (3.18)) and $\hat{\mathbf{s}}_{i-1}$ is the estimate at the previous iteration. Several simulations were carried out with a fixed value of δ (several

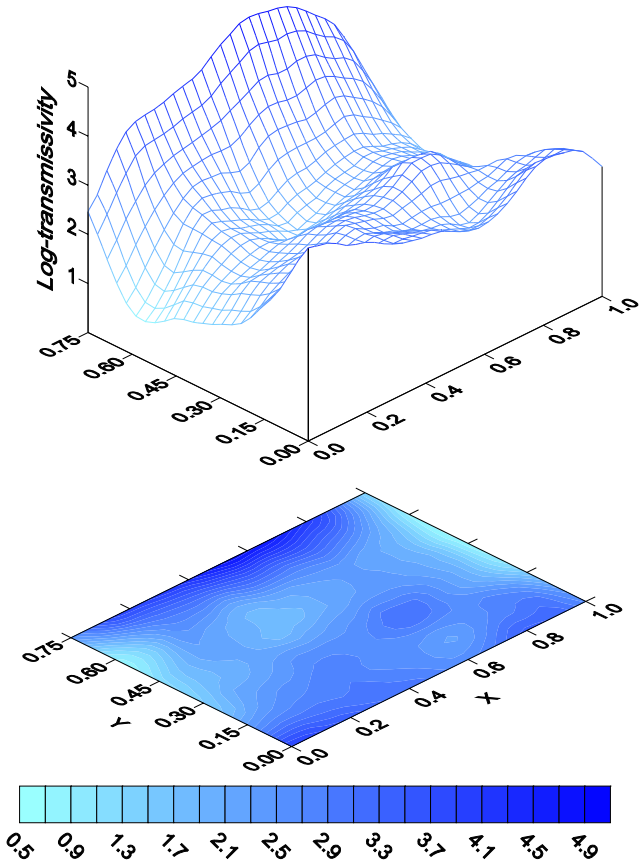

Figure 6.6: True log-transmissivity, **2s** field.

values in the interval $0.5 \leq \delta \leq 0.95$ were analyzed). The δ parameter modifies the velocity of the process: a small value gives big correction and fast results but not always the process converges; while a great value gives a slow solution (in this case the new estimate results very close to the previous one) but a more probable convergence. The use of a fixed value of δ resulted not optimal for strongly nonlinear problems; in fact, in this cases the use of a small value ($\delta = 0.5$) does not guarantee the convergence because of the great modification from an iteration to the next one. In the meanwhile the use of a great value ($\delta = 0.95$) means that the new estimate is very close to the new one, so the procedure requires a high number of iterations to reach the convergence. This fact is not optimal because it means a lot of computational time. Therefore, in order to avoid these problems related to the choice of δ, an optimization routine (see section 6.2) has been added to evaluate δ. This optimization procedure minimizes the objective function (6.4) modifying the value of δ (looking for the best one) at each iteration. The convergence of the algorithm results faster than the one with a fixed value. Moreover the objective function is monotonically decreasing from an iteration to the next one.

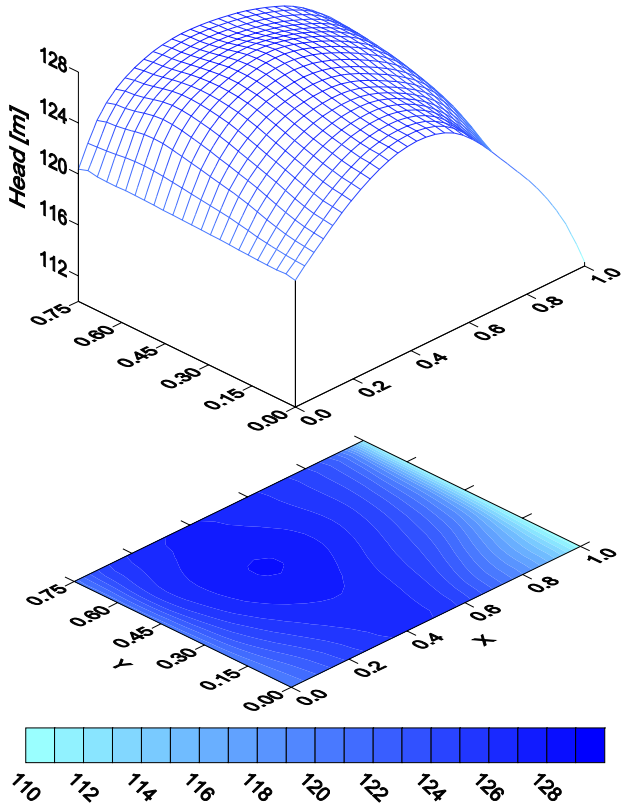

Figure 6.7: Head [m] related to the **2s** tranmissivity field.

The convergence criterion of the procedure is based on the differences between the actual value of the objective function and the old one ($L-L_{old}$); a specific value has been chosen in order to allow a comparison between all the studied cases. Initially a small convergence criterion was considered ($L - L_{old} \leq 0.01$). Performing several tests it was noticed that the weakly nonlinear cases were able to converge with small errors, but these values, for strongly nonlinear fields, implied a great number of iterations so it was decided to use, as convergence criterion $L-L_{old} \leq 1$ in order to avoid too many iterations and too long computation time. It is important to notice that this convergence criterion doesn't affect heavily the final results. The use of the optimization routine resulted very important in the convergence procedure; in fact it allowed to converge strongly nonlinear fields.

At first the QL approach was performed without the estimation of the parameter θ of the cubic generalized covariance function ($Q_{ij} = \theta h_{ij}^3$). Several fixed values of θ were considered but this configuration (QL without parameter estimation) was not satisfactory especially in the most nonlinear case. The choice of the right parameter means also the choice of the right model and that is crucial to reach the right solution. Then the procedure that allows to estimate the parameters of the covariance function $\mathbf{Q}$, described in *Kitanidis* [1995] and in section 6.2, was added with good results. Considering the statistics $Q_1 = \frac{1}{n-1}\sum_{i=2}^{n}\varepsilon_i$ and $Q_2 = \frac{1}{n-p}\sum_{i=2}^{n}\varepsilon_i^2$, where ε_i are the orthonormal residuals (see appendix D) it is possible to see if the chosen model is correct. The values reported in table 6.2 and in appendix A show that the parameter estimation routine works very well. The parameter estimation routine consists in the minimization of the function (3.12) with respect to θ (maximum likelihood). In the Gauss-Newton process the value of θ is updated at each iterations ($\theta_{l+1} = \theta_l - \mathbf{F}^{-1}\mathbf{g}$) till the convergence to a value is reached. The convergence criterion, in the parameter estimation routine, is also based on the differences between the actual and the old objective function. A convergence criterion of $L - L_{old} \leq 1$ was considered also in this case. The parameter estimation gives the best fitting of the covariance model to the data at each iteration.

For nonlinear cases it is necessary to use the Marquardt modification (see *Kitanidis* [1995] and section 3.2 for more details) in the Gauss-Newton iterative process (equation (3.14)) to allow an easy convergence of the procedure. The Marquardt modification consists in replacing $\mathbf{F}^{-1}$ with $(\mathbf{F} + \lambda\mathbf{I})^{-1}$, where $\mathbf{I}$ represents the identity matrix and λ is a positive parameter. The Marquardt parameter (λ) was fixed in each cases (see the appendix A for the used values). This modification helps [see *Kitanidis*, 1995] the quasi-linear approach to converge in strongly heterogeneous cases.

Increasing the nonlinearity of the transmissivity field the parameters of the algorithm have to be adjusted; see appendix A for the used values.

Then the inverse problems on the log-transmissivity fields were performed using all these improvements; this procedure will be briefly referred in the following as QL2.

The table 6.2 and the appendix A show all the statistics evaluated after the application of the QL2 to the log-transmissivity fields (**s**, **2s**, **4s**, **6s**, **8s** and **10s**). In the table 6.2 the following symbols are used:

- σ_R^2 is the variance of the measurement error added for noising purposes;
- θ_{est} is the parameter of the covariance matrix $\mathbf{Q}$, that allows the best fitting to the data, evaluated by the parameter estimation routine;
- Q_2 is the statistics that represents if the chosen model is good [see *Kitanidis*, 2004]. It should be close to 1 and in particular between $Q_{2crit1} < Q_2 < Q_{2crit2}$ (see appendix A); in the cases analyzed $Q_{2crit1} = 0.389$ and $Q_{2crit2} = 1.611$;

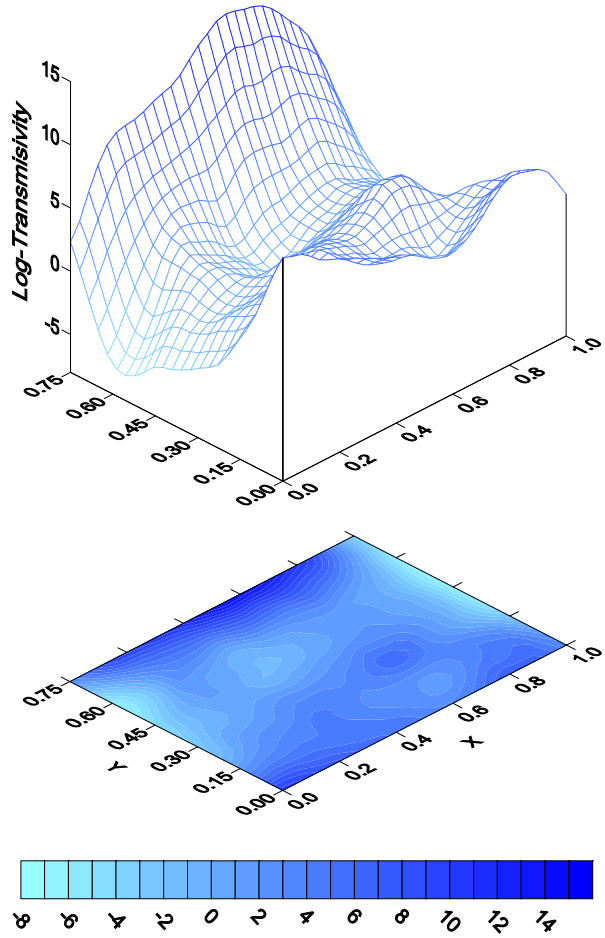

Figure 6.8: True log-transmissivity, **10s** field.

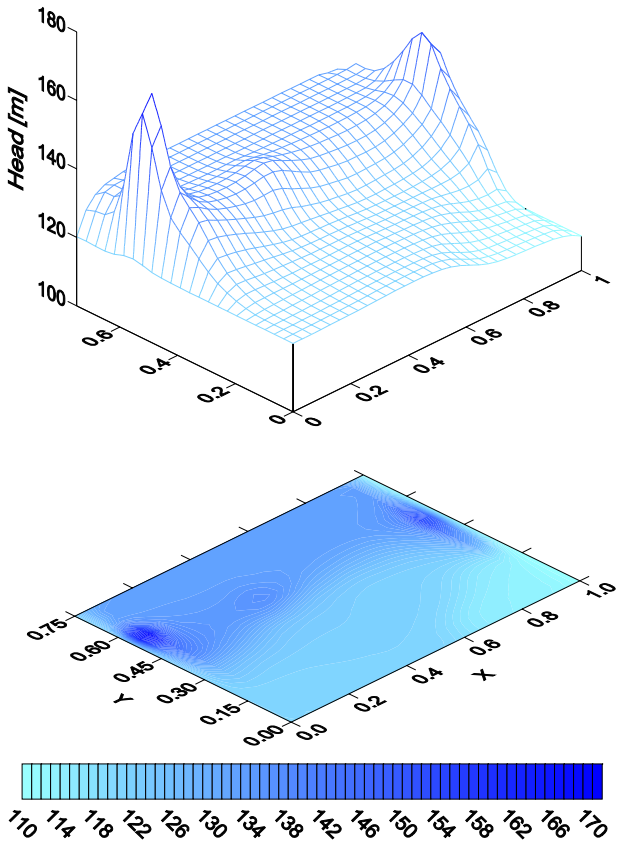

Figure 6.9: Head [m] related to the **10s** transmissivity field.

- mean and variance of the recovered log-transmissivity field, $mean\left[\overset{\wedge}{\mathbf{s}}\right]$, $var\left[\overset{\wedge}{\mathbf{s}}\right]$;
- mean of the differences between the true field and the recovered one, $mean\left[\mathbf{s}_{true}-\overset{\wedge}{\mathbf{s}}\right]$, and of the square differences, $mean\left[\left(\mathbf{s}_{true}-\overset{\wedge}{\mathbf{s}}\right)^2\right]$;
- $D_2 = mean\left[\dfrac{\left(\mathbf{s}_{true}-\overset{\wedge}{\mathbf{s}}\right)^2}{diag(\mathbf{V})}\right]$ is the mean of the square differences between the true field and the recovered one over the variance for each node; D_2 should be close to 1.

Analyzing the table 6.2 and comparing the obtained results it is possible to notice that using a measurement error of $\sigma_R^2 = 10^{-4}$ can be enough to get good results; in fact in all cases it is possible to see that the recovered field is very close to the true one. The possibility to use $\sigma_R^2 = 10^{-4}$ allows to speed up the computation and to reach the solution quickly.

The good agreement between the true log-transmissivity field and the recovered one is shown comparing the Figures 6.5 - 6.10 (**s** case, $\sigma_R^2 = 10^{-6}$), 6.6 - 6.11 (**2s** case, $\sigma_R^2 = 10^{-6}$) and 6.8 - 6.12 (**10s** case, $\sigma_R^2 = 10^{-4}$).

Considering the **s** log-transmissivity field, the efficiency of the parameter estimation routine is shown by the Q_2 parameter, see table 6.2. In fact considering the statistic Q_2 ($0.8741 \div 1.0007$) it is possible to see that the chosen model is correct. The mean of the recovered field is well reproduced (see tables 6.1 and 6.2); in the meanwhile the variance is underestimated for the cases with a small measurement error (this is due to the QL2 approach). Considering the mean of the difference between the true field and the recovered one, $mean\left[\mathbf{s}_{true}-\overset{\wedge}{\mathbf{s}}\right]$, it is easy to notice that is close to 0. Initially a parameter $D_1 = \dfrac{1}{m}\left(\mathbf{s}_{true}-\overset{\wedge}{\mathbf{s}}\right)^T \mathbf{V}^{-1}\left(\mathbf{s}_{true}-\overset{\wedge}{\mathbf{s}}\right)$ was considered (see appendix A) but that resulted too dependent on the measurement error (σ_R^2) and on nonlinearity, so was decided to use D_2 instead of D_1, that seemed more reliable. The parameter D_2, that means how the estimated error approximates the real one, is about 1 for all the analyzed cases even for the high nonlinear field (**10s**).

Considering the **2s** log-transmissivity field, see Figures 6.6 and 6.11, Q_2 is close to 1 for all the measurements errors analyzed that means that the chosen model is correct, also for this case this is due to the parameter estimation routine. The mean of the recovered field is close to 2.5 but the variance is underestimated. Analyzing the differences it is possible to notice that the true field is well recovered in fact the $mean\left[\mathbf{s}_{true}-\overset{\wedge}{\mathbf{s}}\right]$ and the $mean\left[\left(\mathbf{s}_{true}-\overset{\wedge}{\mathbf{s}}\right)^2\right]$ are small. The parameter D_2 results close to 1 in the case with a $\sigma_R^2 = 10^{-6}$.

The **10s** log-transmissivity field, see Figures 6.8 and 6.12, was considered only with the measurements error of $\sigma_R^2 = 10^{-2}$ and $\sigma_R^2 = 10^{-4}$ with the aim at avoiding numerical problem related to small values of σ_R^2, for instance too much ill-conditioned matrix and problems related to the convergence process. The Q_2 statistic is not so close to 1 but it is between the range of the two critical values, $Q_{2crit1} < Q_2 < Q_{2crit2}$. That means that the optimal parameter of the covariance function is not reached but it is acceptable. A greater number of iterations, in the parameter estimation routine, could help to find the optimal structural parameter. The mean of the recovered field is close to 2.5 and the $var\left[\overset{\wedge}{\mathbf{s}}\right] = 7.2431$ ($\sigma_R^2 = 10^{-4}$ case) is underestimated (true $var\,[s] = 10.200$).

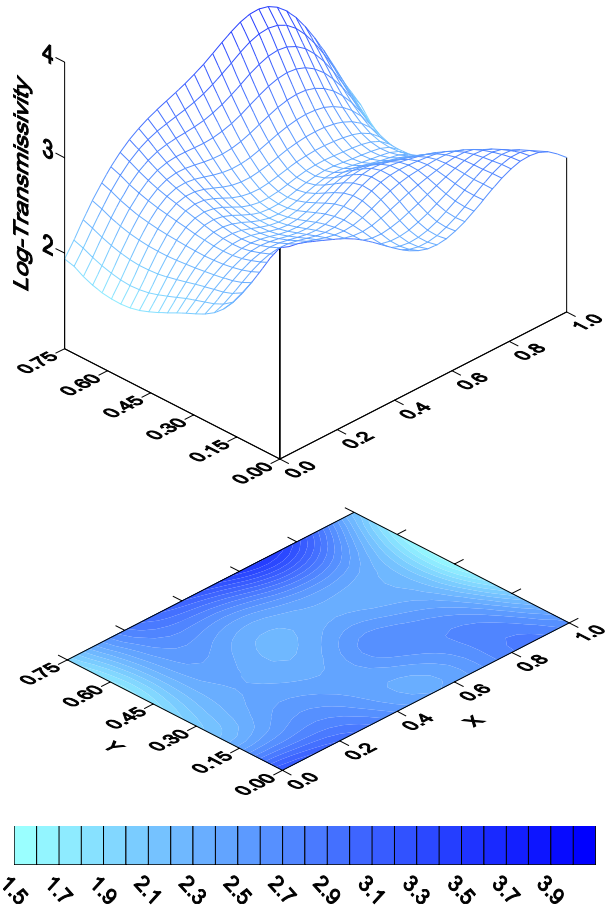

Figure 6.10: Recovery of the **s** log-transmissivity field obtained with the quasi linear approach with the addition of the optimization, parameter estimation and the Marquardt modification.

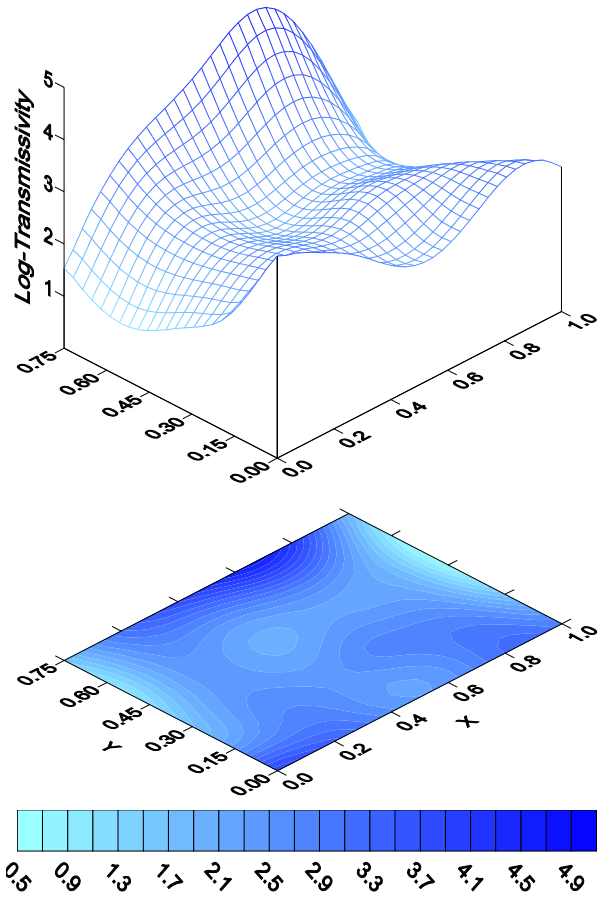

Figure 6.11: Recovery of the **2s** log-transmissivity field obtained with the quasi linear approach with the addition of the optimization, parameter estimation and the Marquardt modification.

Table 6.2: Summary of the statistics of the heterogeneous fields recovered with the quasi linear geostatistical approach with the addition of the optimization routine, the parameter estimation and the Marquardt modification

	σ_R^2	θ_{est}	Q_2	$mean\left[\hat{\mathbf{s}}\right]$	$var\left[\hat{\mathbf{s}}\right]$	$mean\left[\mathbf{s}_{true}-\hat{\mathbf{s}}\right]$	$mean\left[\left(\mathbf{s}_{true}-\hat{\mathbf{s}}\right)^2\right]$	D_2
	10^{-2}	6.1886	0.8741	2.6013	0.1145	−0.1013	0.1054	0.7294
s	10^{-4}	3.7084	1.0995	2.5003	0.1080	−0.0003	0.0246	1.4293
	10^{-6}	6.6892	1.0448	2.4905	0.0815	0.0095	0.0114	0.7447
	10^{-8}	6.1065	1.0007	2.4907	0.0775	0.0093	0.0129	0.9496
	10^{-2}	17.878	1.0171	2.3894	0.3318	0.1106	0.4020	2.0966
	10^{-4}	17.393	1.1595	2.5294	0.2956	−0.0294	0.1307	4.0596
2s	10^{-6}	23.376	1.0417	2.4889	0.3200	0.0111	0.0443	1.3863
	10^{-8}	23.215	1.0412	2.4913	0.3168	0.0087	0.0498	1.7213
	10^{-2}	70.101	1.2960	2.3215	1.3740	0.1785	1.3422	2.1644
	10^{-4}	74.426	1.1183	2.5629	1.3857	−0.0629	0.2438	1.0958
4s	10^{-6}	66.844	1.1498	2.3433	1.4385	0.1567	0.2190	1.7085
	10^{-8}	67.233	1.1544	2.3367	1.4444	0.1633	0.2243	1.7394
	10^{-2}	199.98	0.6875	2.4166	2.5144	0.0834	1.0436	0.9674
6s	10^{-4}	222.43	0.7815	2.3484	2.8770	0.1516	0.4213	0.7559
	10^{-6}	228.57	0.6311	1.9581	2.3641	0.5419	1.2930	1.5293
	10^{-8}	231.48	0.6255	1.9613	2.3700	0.5387	1.2737	1.4929
	10^{-2}	75.769	1.7221	2.3020	3.9898	0.1980	3.0753	5.1077
8s	10^{-4}	441.15	0.7497	2.0722	5.1689	0.4278	2.7678	0.9438
	10^{-2}	600.01	0.8978	1.6616	6.6663	0.8384	5.4171	1.5247
10s	10^{-4}	1137.6	0.5056	1.8665	7.2431	0.6335	4.2144	0.7245

Moreover the differences between the true and the estimated log-transmissivity field are close to 0. The parameter D_2 is close to 1; this means that the measurement error is quite well represented.

These brief analyses on the results can be generalized to all the studied cases; in fact, also the most heterogeneous transmissivity field (**10s**) is well recovered.

In conclusion the quasi-linear approach resulted not satisfactory in the application to strongly nonlinear fields while, with the addition of few improvements, it results efficient and versatile. In any case the QL methodology underestimates the variance.

6.5 Applications with Conditional Realizations

The previous cases (reported in 6.4) were also analyzed with the addition of the conditional realization using the procedure described in the section 6.3. In the following only three cases (**s** with $\sigma_R^2 = 10^{-6}$, **2s** with $\sigma_R^2 = 10^{-6}$ and **10s** with $\sigma_R^2 = 10^{-4}$) are reported. For the whole results see the appendix B. The application of MCMC to the quasi-linear approach is in order to avoid the underestimation of the variance of the recovered field.

Since the optimization routine and Marquardt modification (see section 6.2) resulted very efficient in the QL2 approach, they were applied also to the process that uses the conditional

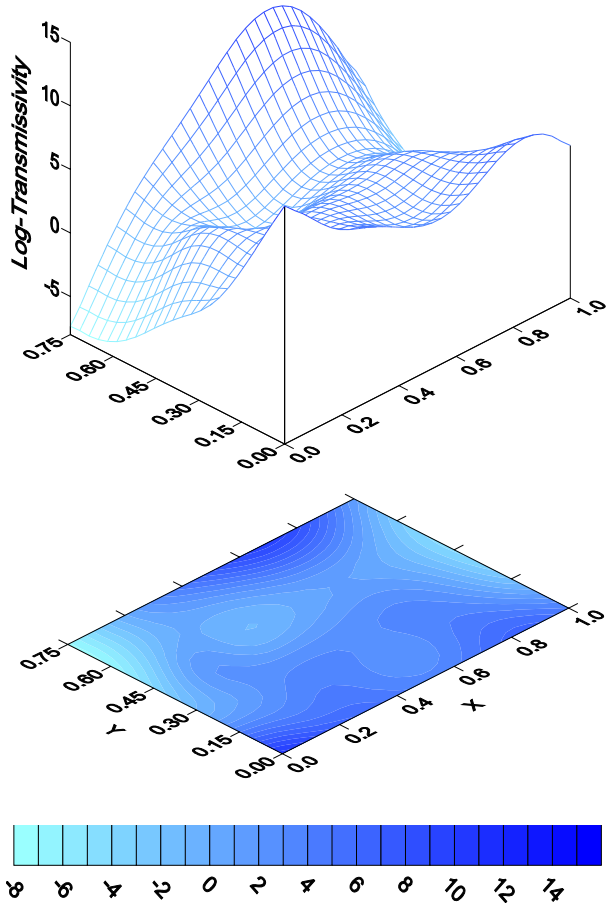

Figure 6.12: Recovery of the **10s** log-transmissivity field obtained with the quasi linear approach with the addition of the optimization, parameter estimation and the Marquardt modification.

realizations; the parameter estimation routine was not added to this methodology (it requires too much computation time) and as parameter of the covariance function (cubic) the one estimated by the QL2 method (see table 6.2) has been chosen. Moreover the same convergence criterion and Marquardt parameter, used in the QL2, were considered (see appendix A).

The first step of the methodology is to create an unconditional realization; at this aim a methodology, that is based on the eigenvalues decomposition (of the covariance matrix $\mathbf{Q}$ and of the drift matrix $\mathbf{X}$) and works with (conditionally) semidefinite covariance matrix, has been adopted. Then the MCMC chain (equation (6.8)) was set up; a value of ρ equal to 0.99 was assumed, so the new unconditional realization is very close to the last accepted one. 1000 conditional realizations (see equation (6.5)) for each transmissivity field were performed to test the method. The computation time varies from 2.5 hours of the **s** field to 10 hours of the **10s** field (simulations performed on an Intel Centrino M 760, 512 Mb RAM).

The results of the 1000 conditional realizations are summarized in the tables 6.3, 6.4, 6.5 and in appendix B. The MCMC chain needs a training period so the first simulations of the chain has to be discarded. For this reason in the tables 6.3, 6.4, 6.5 and in appendix B the results of the whole simulations, from 100 to 1000, 200 to 1000, 300 to 1000 and 500 to 1000 are shown; moreover in these tables the percentage of the accepted conditional realizations is shown.

As described in section 6.3 an acceptance/rejection algorithm was added to the process, in particular the Metropolis-Hastings. This algorithm allows to investigate a result of the geostatistical process with the conditional realizations and to accept it or to reject it in according to the method described in section 6.3. It is possible to see (tables 6.3, 6.4, 6.5 and appendix B) that the percentage of the accepted solution is low (excluding a couple of cases is less than 15%) for all the studied cases.

Considering the first field (**s**) with $\sigma_R^2 = 10^{-6}$ (see table 6.3), it is possible to notice that the statistics of the recovered field with 1000 CR are close to the one obtained with the QL2 approach. The variance of 1000 CR is 0.1098 instead of 0.1020 (true value), and avoiding the first 300 realizations the variance results 0.1028, that is very close to the true value. The mean of the differences between the recovered field and the true one $mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$ is close to zero and the variance of the differences $var\left[\hat{\mathbf{s}}\right]$ is small. The square difference over the variance (D_2) is close to one for all the ranges considered. See Figures 6.5 (true field), 6.10 (QL2) and 6.13 (conditional realizations) to compare the results obtained with the two methodologies. Both methodologies represent very well the shape of the true field but the solution obtained using the QL2 approach is smoother than the other one. Analyzing the differences between the true field and the estimated one, the mean differences using QL2 is 0.0095 (considering $\sigma_R^2 = 10^{-6}$), while using the CR the mean of the differences results -0.0026. Considering $mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$ instead of $mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$, it is possible to notice that the value obtained using the QL2 approach is 0.0114 while for the CR case is 0.0090 (considering 1000 realizations). Both values $mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$ and $mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$ result smaller using the CR approach. The value of D_2 for the QL2 approach is 0.7447, while using the CR approach the best value is 0.9315 and it is reached avoiding 500 realizations. This analysis means that for weakly nonlinear fields both the quasi-linear and the conditional realization approach are very efficient. The CR approach reproduces better the true log-transmissivity field, but the QL approach requires a few computation time (few minutes instead of 3 hours).

Considering the **2s** field (see table 6.4 and Figure 6.14) the mean of the recovered log-transmissivity results close to 2.5 and the variance is well estimated, considering the whole 1000 CR. $mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$

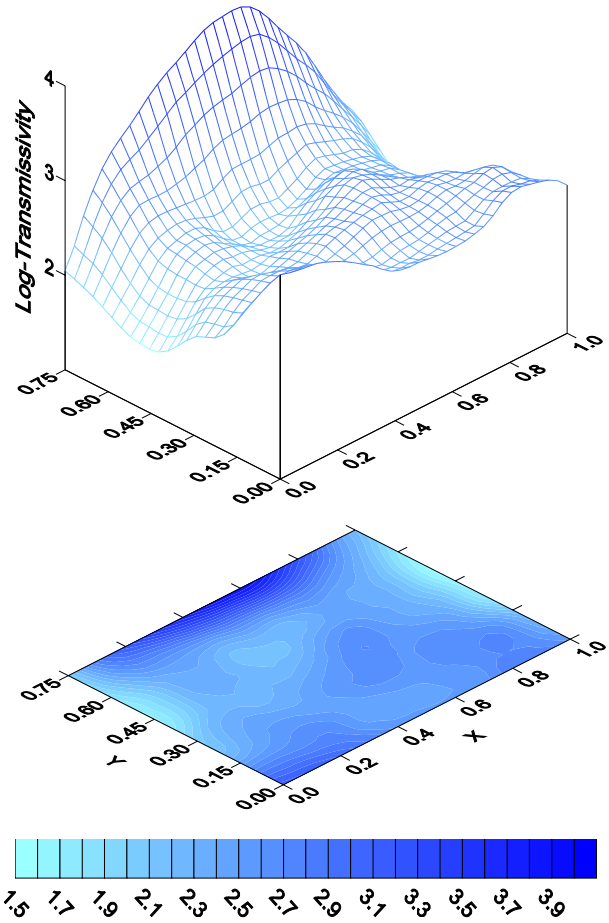

Figure 6.13: Recovery of the **s** log-transmissivity field obtained with the quasi linear approach with the addition of the 1000 conditional realizations.

Table 6.3: Summary of 1000 conditional realizations applied to the s transmissivity field.

$\sigma_R^2 = 10^{-6}$					
From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	10.700	9.8779	10.112	9.8431	12.575
$mean\left[\hat{\mathbf{s}}\right]$	2.4974	2.4947	2.4929	2.4907	2.4899
$var\left[\hat{\mathbf{s}}\right]$	0.1098	0.1079	0.1068	0.1028	0.1009
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.0026	0.0053	0.0071	0.0093	0.0101
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.0090	0.0093	0.0093	0.0095	0.0097
D_2	0.8031	0.8529	0.8662	0.9070	0.9315

Table 6.4: Summary of 1000 conditional realizations applied to the 4s transmissivity field.

$\sigma_R^2 = 10^{-6}$					
From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	11.600	10.877	11.610	11.840	12.176
$mean\left[\hat{\mathbf{s}}\right]$	2.4961	2.4915	2.4891	2.4861	2.4834
$var\left[\hat{\mathbf{s}}\right]$	0.4487	0.4386	0.4353	0.4262	0.4101
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.0039	0.0085	0.0109	0.0139	0.0166
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.0326	0.0325	0.0328	0.0333	0.0351
D_2	0.8544	0.8943	0.9164	0.9550	1.0351

results smaller than the one evaluated using the QL2 (0.0039 instead of -0.0111). Moreover the mean of the square difference is smaller using the CR approach 0.0326 instead of 0.0443. This means that the CR approach represents quite well the true log-transmissivity field. The D_2 parameter is close to 1 and similar to the one estimated with the QL2 approach. The best D_2 is reached avoiding 500 CR. The shape of the recovered field is well represented, see Figures 6.8 (true) and 6.14 (CR), and comparing the graphical results to the one obtained with the QL2 approach, see Figure 6.12, it is possible to notice, as the **s** log-transmissivity field, that the QL2 approach is smoother than the other one. Also in this case the solution is reached after several hours of computation time instead of few minutes needed by the QL2 approach.

Considering the **10s** field (see table 6.5 and Figure 6.15) the mean and the variance of the recovered log-transmissivity field are underestimated compared to the true one ($mean\left[\hat{\mathbf{s}}\right] = 2.1582$ instead of 2.5, $var\left[\hat{\mathbf{s}}\right] = 8.2021$ instead of 10.200). Both the value of $mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$ and $mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$ using the CR approach are smaller than using the QL2 methodology. The D_2 parameter is closer to 1 using the QL2 approach than the CR; that indicates that the QL2 represents

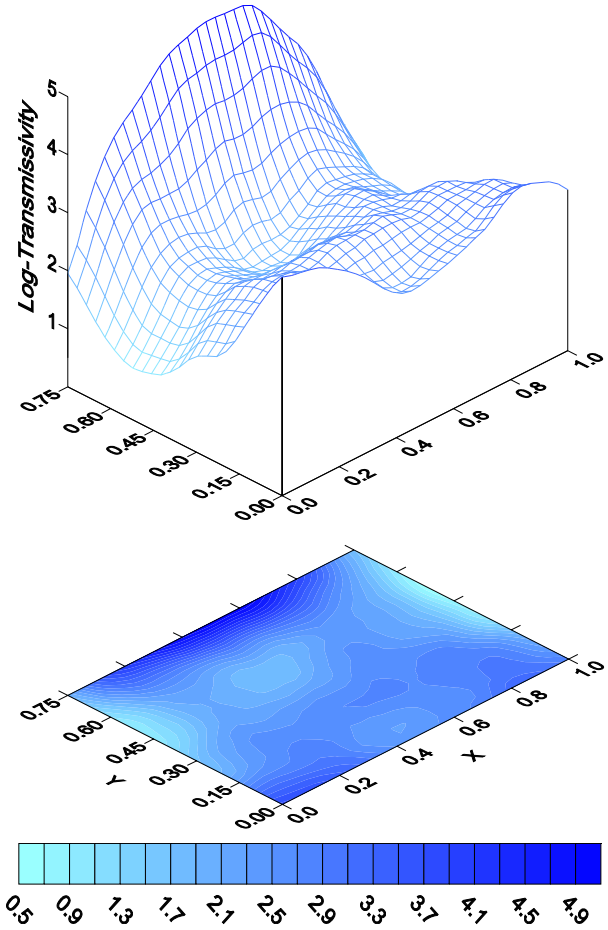

Figure 6.14: Recovery of the **2s** log-transmissivity field obtained with the quasi linear approach with the addition of the 1000 conditional realizations.

Table 6.5: Summary of 1000 conditional realizations applied to the 10s transmissivity field.

$\sigma_R^2 = 10^{-4}$					
From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	11.000	10.655	11.610	12.553	9.5808
$mean\left[\hat{\mathbf{s}}\right]$	2.1582	2.1197	2.1132	2.0994	2.0456
$var\left[\hat{\mathbf{s}}\right]$	8.2021	8.0912	8.0706	8.0510	8.1264
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.3418	0.3803	0.3868	0.4006	0.4544
$var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	1.3525	1.4987	1.5210	1.5699	1.7975
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	1.4676	1.6413	1.6687	1.7283	2.0016
D_2	1.5797	1.7853	1.8074	1.8490	2.1425

better the measurement error than the CR. The number of the accepted conditional realizations in the **10s** case is less than 13%, this is due to the strongly heterogeneous log-transmissivity field. The computation time is about 10 hours instead of few minutes for the QL2 approach. The CR needs so much time because for strongly nonlinear fields it requires a very great number of iterations to reach the convergence.

Increasing the nonlinearity the percentage of the accepted solution decreases, moreover using a too big measurement error the number of the accepted solution is very low. In fact considering the tables in appendix B it is possible to see that for the cases **4s**, **6s**, **8s** and **10s** with $\sigma_R^2 = 10^{-2}$ the percentage of the accepted solution is less than 3%. This is due to the great error considered associated to the high nonlinearity of the log-transmissivity field.

6.6 Conclusions

The first and most important result is that the quasi-linear procedure, with the additions of three simple routine, is efficient; it is applicable to strongly nonlinear field and it allows to reach the convergence quickly. The only disadvantage of the QL2 approach is that it underestimate the variance. The conditional realizations is a very expensive procedure, it requires a lot of computation efforts but, in all the studied cases, it recovered the log-transmissivity field better than the QL2 approach. The application of the conditional realizations is a very complex process; in particular the choice of the parameters. In this work the chosen parameters were the one evaluated by the parameter estimation of the QL2 approach. In fact, few analysis were performed to choose the value of the Marquardt parameter, and the starting value of the parameter estimation routine. Probably the best method to solve these kind of problems is to join the two methodologies; the first step regards the QL2 that is used to obtain a first estimate and allows to evaluate the correct parameters (it is important to remark that the QL2 allows to reach the solution in few minutes), then the second step applies the CR, with the evaluated parameter, in order to find the best solution. Moreover it is important to find a criterion that allows to specify the training period of the MCMC chain and the number of the realizations that are required to solve the process.

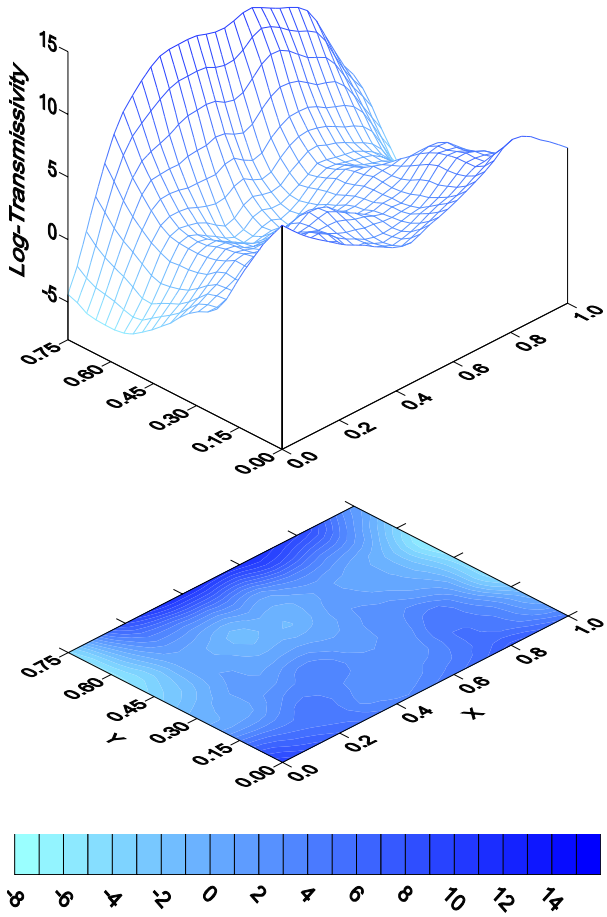

Figure 6.15: Recovery of the **10s** log-transmissivity field obtained with the quasi linear approach with the addition of the 1000 conditional realizations.

In this work several measurement errors (from 0.001 m to 1 m) were considered in order to compare the results; using a too small measurement error, the process is unrealistic as using a too wide error. Using a great error the solution results flatter than using a small error; this is due to the $\mathbf{R}$ matrix in the objective function: the misform part ($\left(\mathbf{z} - h\left(\hat{\mathbf{s}}\right)\right)^T \mathbf{R}^{-1} \left(\mathbf{z} - h\left(\hat{\mathbf{s}}\right)\right)$) becomes more important than the misfit ($\hat{\mathbf{s}}^T \mathbf{G}\hat{\mathbf{s}}$). A value of $\sigma_R^2 = 10^{-4}$ (0.1 m) or $\sigma_R^2 = 10^{-6}$ (0.01 m) are the most realistic cases for the measurement errors.

Both the QL2 and CR underestimated the variance; considering the CR approach this problem can be solved increasing the number of the iterations, but this means also the increasing of the computation time. Probably in the next years, thanks to the development of the personal computer, it will be possible to solve this kind of problems in a short time using a great number of realizations.

Few developments could be added to the CR approach; for instance a routine that allows to evaluate the Marquardt parameter could be helpful to the convergence routine. In fact it was noticed that the results are highly affected by the Marquardt parameter chosen. Moreover another improvement could be the use of the Fourier transforms to create unconditional realizations, that allows to extend the methodology to very refined grid.

Concluding, the CR approach, in all the studied cases, performs better than the QL2; but the QL2 methodology requires only few minutes to reach the solution instead of few hours. For weakly nonlinear case, as the tested case $\mathbf{s}$, the solution reached by the CR and QL2 are very similar, so the QL2 is a very good substitute of the CR. Using the CR approach, a relevant difference between the QL2 and the CR for the $\mathbf{s}$ case can be found after 1000 conditional realizations. While for all the other nonlinear studied cases, the CR approach is always better and not comparable to the QL2.

Chapter 7

New Developments: Matrix Multiplication

7.1 Introduction

The computation grid is one of the main aspects in the mathematical modeling, denser it is better the results are. But the main problems of dense grid are the costs in computation time and in storage memory. For this reason during the last years several Researchers studied new methodologies to speed up the matrix operations. One of the best methods is the application of the Fast Fourier Transform [*Nowak et al.*, 2003] to the matrix multiplication. The computational costs are reduced from $O(n^2)$ to $O(nlog_2 n)$ and the storage is reduced from $O(n^2)$ to $O(n)$. Geostatistics utilizes the cross-covariance matrix between the unknowns and the observations and the auto-covariance matrix among the observations to infer information from the observations onto the unknowns. This kind of condition requires a lot of matrix-matrix multiplication.

In geostatistics spatially distributed unknown are described as realizations of random processes and characterized by their mean values and covariance functions. A list of all the matrices and vectors used in common geostatistical processes (see section 3.2 or see *Kitanidis* [1995]) follows:

$\mathbf{s}$ $(n \times 1)$ vector of n unknowns

$\mathbf{t}$ $(m \times 1)$ vector of m observations

$\mathbf{Q}_{ss}$ $(n \times n)$ covariance matrix of $\mathbf{s}$

$\mathbf{Q}_{st}$ $(n \times m)$ cross-covariance matrix between observations and unknowns

$\mathbf{Q}_{tt}$ $(m \times m)$ auto covariance matrix of observation

$\mathbf{R}$ $(m \times m)$ auto covariance matrix of the measurement error

$\mathbf{Q}_{st} = \mathbf{Q}_{ss}\mathbf{H}^T$

$\mathbf{Q}_{tt} = \mathbf{H}\mathbf{Q}_{ss}\mathbf{H}^T + \mathbf{R} = \mathbf{H}\mathbf{Q}_{st} + \mathbf{R}$

Given the relationship $t = f(s)$ between observations and unknowns, $\mathbf{H}(m \times n)$ is the linearized sensitivity matrix ($\mathbf{H} = \partial t/\partial s$) computed using the adjoint-state method.

The construction of $\mathbf{Q}_{ss}$ is a computation $O(n^2)$, $\mathbf{Q}_{st}$ is $O(mn^2)$, and Q_{tt} via $\mathbf{Q}_{tt} = \mathbf{Q}_{ts}\mathbf{H}^T$ or $\mathbf{Q}_{tt} = \mathbf{H}\mathbf{Q}_{ss}\mathbf{H}^T$ is $O(nm^2)$ or $O(nm^2 + mn^2)$. These circumstances have limited the diffusion of methods like cokriging that requires explicit formulation of $\mathbf{Q}_{st}$ and $\mathbf{Q}_{tt}$. In most cases the unknown is a stationary random variable discretized on a regular and equispaced grid; this imposes

symmetric Toeplitz matrix (ST)[1] or symmetric block Toeplitz matrix with Toeplitz block (STT) structure onto $\mathbf{Q}_{ss}$ and reduces the storage requirements for $\mathbf{Q}_{ss}$ to $O(n)$. Toeplitz matrices can be embedded in larger circulant matrices[2].

1-D problems lead to symmetric circulant matrices (SC); 2-D problems lead to symmetric block circulant matrices with circulant blocks (SCC).

The overall approach is shown in the following and an example regarding the inverse problem is performed. The aim of the example is to determine the transmissivity of a flow field given measured heads in specific positions of the flow field.

7.2 Methodology [*Nowak et al.*, 2003]

Consider a finite regular grid $n'_y \times n'_x$ with constant spacing d_x and d_y, sized $L'_x \times L'_y$ and $\mathbf{x}'$ and $\mathbf{y}'$ the $n'_y n'_x \times 1$ vectors with the x and y coordinates of its nodes. $\mathbf{s}'$ denotes a stationary Gaussian random space variable on the regular grid, with zero mean and covariance function $R(h)$. The covariance matrix $\mathbf{Q'_{ss}}$ has STT Structure with $n'_x \times n'_x$ blocks sized $n'_y \times n'_y$.

The embedding of the matrix could be summarized as follows: to embed ST in SC matrices, extend the series $t_0, ..., t'_{nx-1}$ by appending the elements $t_1, ..., t'_{nx-2}$ in reverse order to obtain a series $c_o ... c_{n_x} ... c_1$, $n_x = n'_x - 1$, corresponding to mirroring the covariance function to render it periodic, so the dimensions of the symmetric circulant matrix are $2n_x \times 2n_x$. To embed STT in SCC matrices, embed the ST blocks $\mathbf{T}_i$ in SC blocks $\mathbf{C}_i$, and then extend the series of the blocks to obtain a periodic series of blocks; the result is a $4n_y n_x \times 4n_y n_x$ matrix, ($n_x = n'_x - 1$, $n_y = n'_y - 1$).

The Jacobian $\mathbf{H}$ has to undergo the same embedding by zero-padding all entries corresponding to the new entries in the SCC covariance matrix $\mathbf{Q}_{ss}$. Reshape the m rows sized $1 \times n'_y n'_x$ in to $n'_y \times n'_x$ matrices, pad them with zeros to obtain $n_y \times n_x$ matrices, and reshape them back to $1 \times n_y n_x$ rows to obtain the embedded version of $\mathbf{H}$ sized $m \times n_y n_x$.

7.2.1 Computation of the Eigenvalues

The diagonalization theorem gives eigenvalues of the $n_y n_x \times n_y n_x$ ($n_x = 2(n'_x - 1)$, $n_y = 2(n'_y - 1)$) SCC matrix $\mathbf{Q}_{ss} = \mathbf{F^H \Lambda F}$

$\mathbf{\Lambda}$: diagonal matrix of eigenvalues

$\mathbf{F}$: 2-D Fourier matrix

[1] The Toeplitz Matrix (ST) is structured as follow:

$$\mathbf{T} = \begin{bmatrix} t_0 & t_1 & ... & t_{n'_x-1} \\ t_1 & t_0 & . & t_{n'_x-2} \\ . & . & . & . \\ t_{n'_x-1} & t_{n'_x-2} & ... & t_0 \end{bmatrix}$$

STT matrices have the same structure of ST, but replacing t_i with $\mathbf{T}_i$.

[2] The circulant matrix (SC) is structured as follow

$$\mathbf{C} = \begin{bmatrix} c_0 & c_1 & ... & c_{n_x} & ... & c_1 \\ c_1 & c_0 & . & c_{n_x-1} & . & c_2 \\ . & . & . & . & . & . \\ c_{n_x} & c_{n_x-1} & ... & c_0 & ... & c_{n_x-1} \\ . & . & . & . & . & . \\ c_1 & c_2 & ... & c_{n_x-1} & ... & c_0 \end{bmatrix}$$

In SCC matrices c_i are replaced by $\mathbf{C}_i$, (which are SC submatrices).

$\mathbf{F^H} = \mathbf{F^{-1}}$: its Hermitian transpose[3], $\mathbf{F}^H\mathbf{F} = \mathbf{I}$

$\mathbf{FQ_{ss}} = \mathbf{\Lambda F}$, considering only the first column of $\mathbf{Q_{ss}}$ that contains all the informations $\mathbf{FQ_{ss,1}} = \mathbf{\Lambda F_1}$. All entries of $\mathbf{F_1}$ equal $(n_x n_y)^{-1/2}$:

$$\lambda = \sqrt{n_x n_y}\mathbf{FQ}_{ss,1} \tag{7.1}$$

where $\boldsymbol{\lambda}$ $(n_y n_x \times 1)$ is the vector of eigenvalues and $\mathbf{F}$ contains the eigenvectors. $\mathbf{FQ_{ss,1}}$ is computed using 2-D FFT.

The eigenvalues are computed once and then stored for all the subsequent step.

2-D Fourier matrix: The Discrete Fourier Transform (DFT) can be formulated in matrix notation. Consider $\mathbf{F}$ an $n_y n_x \times n_y n_x$ matrix with the entries:

$$F_{kl} = \frac{1}{\sqrt{n_x n_y}} \exp\left(-\frac{2\pi i}{n_y}\left[\frac{l-(l \bmod n_x)}{n_x}\right]\cdot\left[\frac{k-(k \bmod n_x)}{n_x}\right]\right) \tag{7.2}$$
$$\exp\left(-\frac{2\pi i}{nx}(l \bmod n_x)(k \bmod n_x)\right) \quad l, m = 0, 1..., n_x n_y - 1.$$

Then the 2-D Fourier transform $\mathbf{V}$ of any $n_y \times n_x$ matrix $\mathbf{U}$ is: $\mathbf{v} = \mathbf{Fu}$, in which $\mathbf{u}$ and $\mathbf{v}$ are $n_y n_y \times 1$ obtained from rearranging the matrices $\mathbf{U}$ and $\mathbf{V}$ column-wise.

The inverse Fourier transformation corresponds to the inverse of Fourier matrix. As $\mathbf{F}$ is unitary, $\mathbf{F}^{-1} = \mathbf{F}^H$ and $\mathbf{F}^H\mathbf{F} = \mathbf{I}$, where H denotes the Hermitian Transpose.

$\mathbf{Fu}$ is evaluated by reshaping $\mathbf{u}$ into an $n_y \times n_x$ matrix $\mathbf{U}$, computing the 2-D FFT $\mathbf{V} = F(\mathbf{U})$, and reshaping $\mathbf{V}$ back to an $n_y n_x \times 1$ vector $\mathbf{v} = \mathbf{Fu}$, which is $O(nlog_2 n)$ instead of $O(n^2)$, $n = n_y n_x$.

7.2.2 Matrix-Vector Multiplication

The product between $\mathbf{Q_{ss}}$ and a vector $\mathbf{u}$ $(n_y n_x \times 1)$ could be simplified using the previous equations:

$\mathbf{Q_{ss}u} = (\mathbf{F^H\Lambda F})\mathbf{u} = \mathbf{F^H\Lambda(Fu)}$, in which $\mathbf{Fu} = \mathbf{v}$ is computed via 2-D FFT.

As $\mathbf{\Lambda}$ is diagonal, $\mathbf{Q_{ss}u} = \mathbf{F^H}(\mathbf{\Lambda v}) = \mathbf{F^H}[\boldsymbol{\lambda}_1\mathbf{v}_1, \boldsymbol{\lambda}_1\mathbf{v}_1, ..., \boldsymbol{\lambda}_{n_x n_y}\mathbf{v}_{n_x n_y}]^T$

For $\mathbf{u}^T\mathbf{Q_{ss}}$, compute $(\mathbf{Q_{ss}u})^T$ as $\mathbf{Q_{ss}} = \mathbf{Q_{ss}^T}$.

This procedure is called spectral convolution.

7.2.3 Vector-Matrix-Vector Multiplication

To evaluate $\mathbf{u_1^T Q_{ss} u_2}$ with $\mathbf{u}_1$ and $\mathbf{u}_2$ sized $n_y n_x \times 1$ the procedure is:

$$\mathbf{u_1^T Q_{ss} u_2} = \mathbf{u_1^T}(\mathbf{F^H\Lambda F})\mathbf{u_2} = \mathbf{v_1^H \Lambda v_2} = \sum_{k=1}^{n_x n_y}(v_1^*)_k \lambda_k (v_2)_k \tag{7.3}$$

$\mathbf{v_1} = \mathbf{Fu}_1$ and $\mathbf{v_2} = \mathbf{Fu}_2$ are computed by 2-D FFT and $\mathbf{v_i^*}$ is the complex conjugate of $\mathbf{v_1}$.

For $\mathbf{u_1} = \mathbf{u_2}$:

$$\mathbf{u}^T\mathbf{Q_{ss}u} = \sum_{k=1}^{n_x n_y}\lambda_k |v_k|^2 \tag{7.4}$$

[3] Hermitian matrix [Weisstein]: A square matrix is called Hermitian if it is self-adjoint. Therefore, a Hermitian matrix $\mathbf{A} = (a_{ij})$ is defined as one for which $\mathbf{A} = \mathbf{A}^H$, where $\mathbf{A}^H$ denotes the conjugate transpose. This is equivalent to the condition $a_{ij} = \bar{a}_{ji}$, where $\bar{a}$ denotes the complexe conjugate. As a result of this definition, the diagonal elements a_{ii} of a Hermitian matrix are real numbers (since $a_{ii} = \bar{a}_{ii}$), while other elements may be complex.

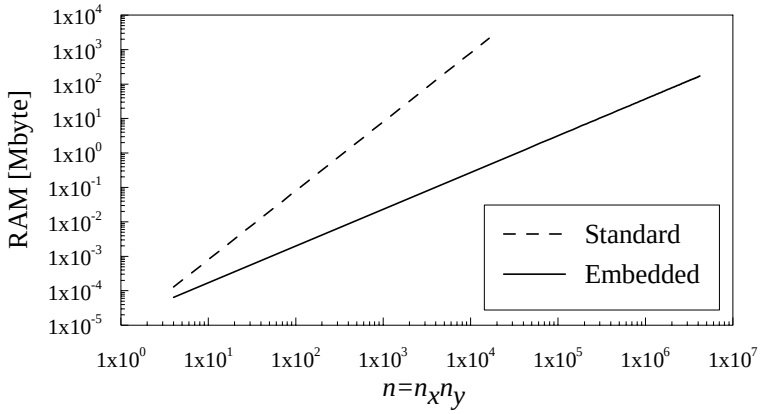

Figure 7.1: Log-Log plot for the comparison of memory consumption in MByte for the storage of $\mathbf{Q}_{ss}$ as a function of the number of unknowns $n = n_y n_x$, using the different methods.

7.2.4 Matrix-Matrix Multiplications

Considering $\mathbf{H}$ $(m \times n_y n_x)$, the computation of $\mathbf{Q_{st}} = \mathbf{Q_{ss}}\mathbf{H^T}$ can be split up into single vector-matrix multiplications:

$\mathbf{Q_{st}} = \mathbf{Q_{ss}}\mathbf{u}_k$, $k = 1...m$, where $\mathbf{u}_k$ is the transpose of the kth row of $\mathbf{H}$.

$\mathbf{Q_{tt}} = \mathbf{H}\mathbf{Q_{ss}}\mathbf{H^T}$ can be split up into m^2 subproblems: $\mathbf{Q_{tt},}_{kl} = \mathbf{u}_k^{\mathbf{T}}\mathbf{Q_{ss}}\mathbf{u}_l$

As $\mathbf{H}\mathbf{Q_{ss}}\mathbf{H^T}$ is symmetric, only the upper triangle (see equation (7.3)) and the diagonal (see equation (7.4)) has to be computed.

The spectral results of $\mathbf{Q_{st}} = \mathbf{Q}_{ss}\mathbf{H^T}$ $(m \times n_y n_x)$ correspond to the embedded version of $\mathbf{H}$. To extract the original size $\mathbf{Q_{st}}$, the process has to be reversed. $\mathbf{Q_{tt}}$ requires no extraction as it is sized $(m \times m)$.

The graph in Figure 7.1 shows the memory required for the storage of $\mathbf{Q}_{ss}$ using the standard method and the spectral method. The standard method runs out of memory at 2^{12} and the FFT method at 2^{21} (the test is performed on an Intel pentium III 1GHz with 384 Mb RAM).

As mentioned above the covariance matrix $\mathbf{Q}_{ss}$ has a STT structure, so this property can be used to embed the matrix in a fast way. The graphs 7.2 show the time (Clock and CPU) between the standard embedding method and the one that uses the STT properties. The method that uses the STT properties is always faster then the standard one.

7.3 Examples

The graphs in Figure 7.3 show the comparison of the computation time using standard and FFT methods including the generation of the matrix $\mathbf{Q}_{ss}$ for the matrix-matrix multiplication $\mathbf{Q}_{st} = \mathbf{Q}_{ss}\mathbf{H}^T$. The covariance matrix $\mathbf{Q}_{ss}$ is assumed to have the exponential form: $\mathbf{Q}_{ss} = \sigma^2 \cdot \exp\left(-\frac{h_{ij}}{l}\right)$. The covariance matrix has to be non negative. The graphs show that the FFT method it is faster than the standard one for grid with $n = n'_y n'_x > 2^8 = 256$. The graphs in Figure 7.4 show the computation time for the multiplication of the matrices used in inverse problems

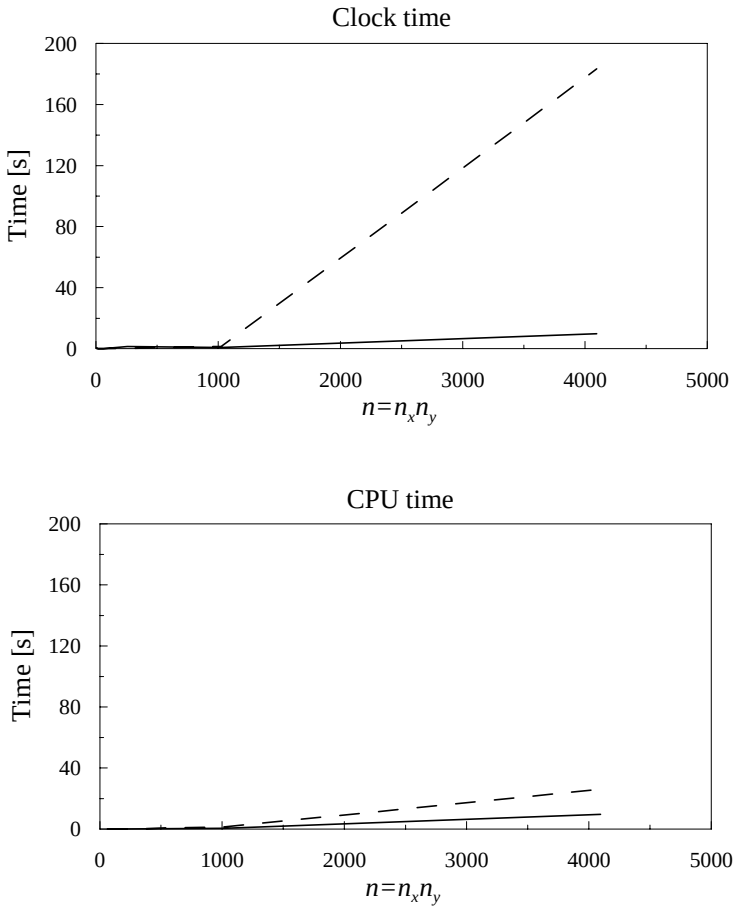

Figure 7.2: Time differences in embedding the covariance matrix. Continuos line is the embedding time using STT properties, Dashed line is the embedding time using standard method.

($\mathbf{Q}_{tt} = \mathbf{H}\mathbf{Q}_{ss}\mathbf{H}^T$ and $\mathbf{Q}_{st} = \mathbf{Q}_{ss}\mathbf{H}^T$) including the embedding of the matrix $\mathbf{Q}_{ss}$. In this case the FFT methods run faster for $n > 2^{10} = 1024$.

The great advantage of the spectral method in the inverse problem is the possibility to use very dense grid.

Two examples regarding the inverse problems were performed to test the spectral method. The aim of the first example (Application 4, CEE 267, Interpolation and Inverse Problems [*Kitanidis,* 2004]) is to determine the transmissivity of an aquifer starting from head measurements.

In this kind of inverse problems several iterations are performed to find the correct solution, but the covariance matrix $\mathbf{Q}_{ss}$ is built only once. So the spectral convolution speeds up problems with very dense grid and solved with a small number of iterations. For instance, to solve a problem in 9 iteration with a grid of $2^5 \times 2^5$ the standard method spends 130 s while the spectral convolution results slower (150 s). This is due to the fact that the covariance matrix is built only once affecting the total computation time in a moderate amount.

The second example deals with the inverse problem in recover the release history of a pollutant [*Snodgrass and Kitanidis*, 1997]. The example studied considers the 1-D unconstrained case [*Snodgrass and Kitanidis*, 1997] and compares the standard multiplication to the spectral method. In this example the covariance matrix $\mathbf{Q}_{ss}$ is assumed to have Gaussian form. The covariance matrix for this kind of problem is a ST matrix, so using the spectral methods, it is embedded in a larger SC matrix. The problem in hand has 300 computation nodes; the solution takes 10 s using the standard method of assembling and multiplying the matrices; with the use of the spectral method the computation time grows to 19 s because the dimension of the grid is out of the good working zone of the new method. The best performance has been obtained (3 s) assembling the covariance matrix $\mathbf{Q}_{ss}$ by means of the ST properties and computing the matrix multiplication with the standard method.

7.4 Conclusions

In the application of the first example the covariance matrix is built only once, so the multiplication time of the matrices has been considered without the embedding of the covariance matrix (see Figure 7.5). The standard multiplication method for grid up to $2^6 \times 2^6$ is always faster than the spectral one. So the only advantage of the procedure proposed by *Nowak et al.* [2003], for grid up to $2^6 \times 2^6$, is in the embedding of the covariance matrix $\mathbf{Q}ss$ (see Figure 7.2).

The results obtained applying these methods to the inverse problems are that the spectral method, for grid up to $2^6 \times 2^6$, is faster than the other one only for few iterations and for a large number of iterations the standard method results faster than the spectral one.

The best choice, for grid up to $2^6 \times 2^6$, is to use the STT properties to build the covariance matrix and to use the standard method for the matrix-matrix multiplication.

The main improvement of the spectral method is the possibility to use a computation grid up to 2^{21} nodes; note that the limit of the standard multiplication method is 2^{12}.

The first inverse problem analyzed needs to build the sensitivity matrix in each iteration, so this computation requires a lot of time, especially for large grid (the dimensions of the matrix $\mathbf{H}$ are $m \times n$ where $n = n_c \times n_r$ and m is the number of observations). A new improvement could be to find a faster method to embed the sensitivity matrix.

Regarding the second example, the set up of the covariance matrix using the properties of the Toeplitz matrix is the main advantage of the methodology proposed by *Nowak et al.* [2003]. Finally the solution of these inverse problems, using the ST properties with the standard multiplication method for grid up to 2^{12} nodes, is faster than the spectral method.

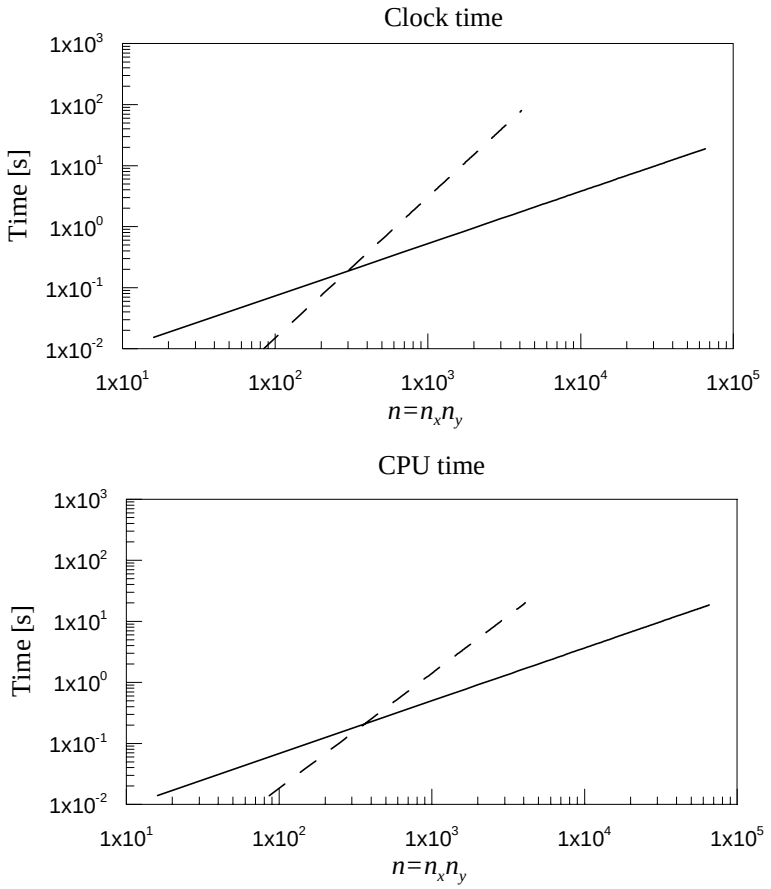

Figure 7.3: Log-Log plot of the computation time (CPU and clock) for the matrix multiplication: $\mathbf{Q_{st}} = \mathbf{Q_{ss}}\mathbf{H^T}$, including the embedding of the STT matrix $\mathbf{Q_{ss}}$. Spectral method (continuos line) and standard method (dashed line).

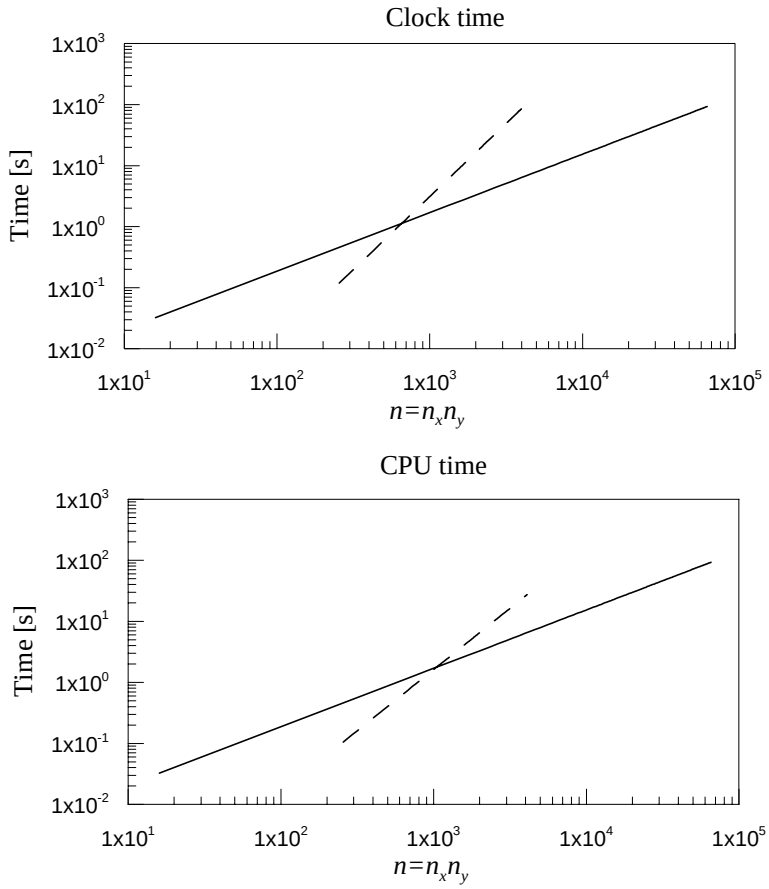

Figure 7.4: Log-Log plot of the computation time (CPU and clock) for the matrix multiplication: $\mathbf{HQ_{ss}H^T}$ and $\mathbf{Q_{ss}H^T}$, including the embedding of the STT matrix $\mathbf{Q_{ss}}$. Spectral method (continuos line) and standard method (dashed line).

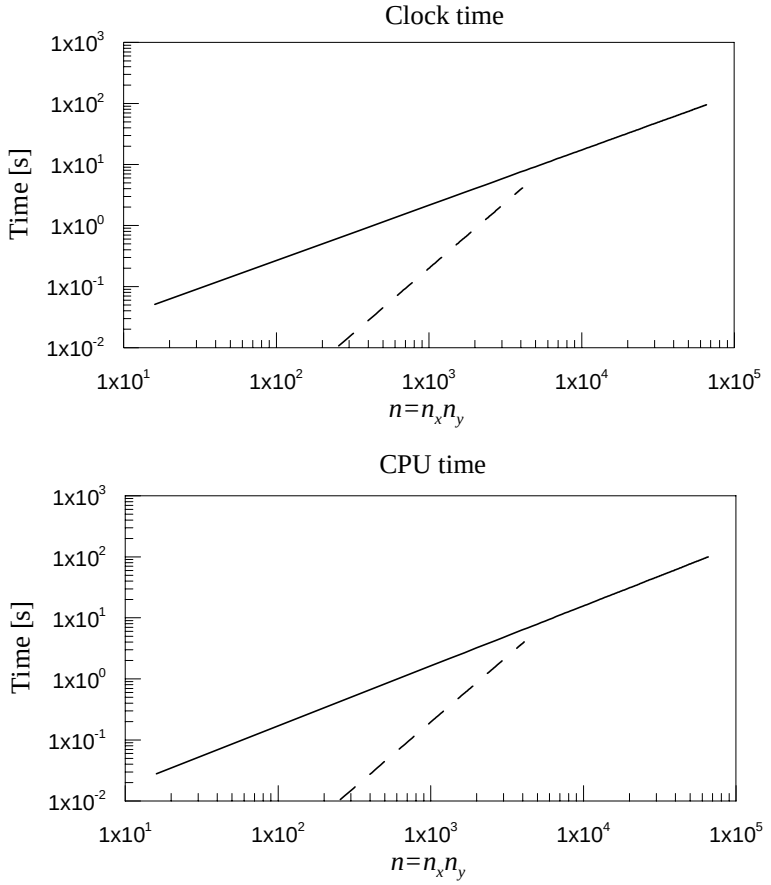

Figure 7.5: Log-Log plot of the computation time (CPU and clock) for the matrix multiplication: $\mathbf{HQ_{ss}H^T}$ and $\mathbf{Q_{ss}H^T}$. Spectral method (continuos line) and standard method (dashed line).

In conclusion the improvement of the procedure can be investigated in the following topics:

- As mentioned above, one improvement could be a new method to evaluate the sensitivity matrix $\mathbf{H}$. In the first example the covariance matrix is built only once and the sensitivity matrix is built in each iteration. For instance the building of the $\mathbf{H}$ matrix for a grid of 2^{12} nodes takes about 280 s. Considering that the $\mathbf{H}$ matrix has to be evaluated at each iteration, the speed up of this process can shorten the solution of the inverse problem.

- Another improvement could be the use of different (than the Gauss-Newton) iterative methods. Promising attempts are made by *Nowak and Cirpka* [2004] with the use of a modified Levenberg-Marquardt.

7.5 Note: Matlab FFT

In the following there is a brief summary of the Matlab function regarding the FFT.

The 2-D FFT in Matlab are not described by the equation (7.2). The regularization coefficient in Matlab is $n_x n_y$ and not $\sqrt{n_x n_y}$, so during the application of the spectral convolution the factor $\sqrt{n_x n_y}$ has to be considered. In Matlab the new equation for the eigenvalues results:

$$\lambda = \mathbf{F}\mathbf{Q}_{ss,1} \tag{7.5}$$

Conclusions

This thesis regards the application and new improvements of the quasi-linear geostatistical approach to two main inverse problems in groundwater: the recovery of the pollutant release history from scarce concentration measurements and the recovery of transmissivity fields from a few head measurements. The knowledge of the pollution release function allows to study the future pollution spread and to plan the remediation actions. Moreover, from a legal and regulatory point of view, it is also important to determine the release time and duration and the highest values of concentration of the injected solution. The second problem allows to save money; in fact, normally, the transmissivity measurements are more expensive than to obtain the head measurements. Starting from the boundary condition of the studied area and from a few head measurements it is possible to know the whole transmissivity field on a grid. Both of the problems are related to the mathematical modeling.

Regarding the recovery of the pollutant release history, the main improvement proposed in this work is the extension of the methodology proposed by *Snodgrass and Kitanidis* [1997] from a 1-D homogeneous problem to a multidimensional heterogeneous one. The transport of a non reactive pollutant through the groundwater is described by a convolution integral, where the transfer function represents the transport phenomena. For easy cases, such as 1-D flow and transport, it is possible to obtain an analytical transfer function, but for more complex cases, such as the presence of heterogeneity, wells, recharge, etc., it is necessary to use highly accurate numerical models to evaluate the transfer function. Knowing the source locations it is possible, using the numerical model, to release a constant concentration of a pollutant in the aquifer and at the same time to measure the response (breakthrough curve) in several measurement points. The transfer function is obtained making the time derivative of the breakthrough curve. Thanks to this improvement it is possible to apply the methodology proposed by *Snodgrass and Kitanidis* [1997] to any cases without limitations. In this work few synthetic cases were analyzed, with very good results, to test the methodology.

The setting up of a numerical model is a complex and costly procedure, but it is possible to save money and time evaluating the transmissivity of the aquifer using the procedure proposed by *Kitanidis* [1995]. Starting from few head measurements, knowing the boundary conditions of the problem in study, it is possible, using the quasi-linear approach, to evaluate the whole transmissivity fields. This work analyzes strongly heterogeneous transmissivity fields (6 cases with an increasing contrast were studied) and proposes few improvements to the methodology; the first regards the optimization of the Gauss-Newton convergence procedure and the second considers the conditional realizations. The results of 1000 conditional realizations and the quasi-linear approach, on each transmissivity field, were compared in order to find the best methodology. The quasi-linear with the addition of the optimization procedure results very efficient, but the methodology that uses the conditional realizations represents better the real transmissivity fields. Two questions raise

from this study; the first regards the number of the realizations that are required to obtain good results and the second is related to the number of iterations that represents the training period of the chain. The answers to these questions represent the future analyses of this methodology.

Considering the pollutant release history, the possibility to extend the proposed approach to evaluate the transfer functions are several. Initially the application of the methodology to reactive solutes will be consider, then to unsaturated cases. The ultimate objective of the author is to apply the methodology to a real case.

During the study of the quasi-linear geostatistical approach the improvements of *Nowak et al.* [2003] are reported and tested. They concern a method based on the fast Fourier transform, that allows to speed up the matrix multiplications. This issue results very important because, in the opinion of the author, the next step of the geostatistical approach is to develop tools that speed up the process, in order to save computation time.

Part III

Appendices

Appendix A

Results of the Quasi-Linear Geostatistical Approach Applied to the Estimation of the Aquifer Parameters

The following tables regard the chapter 6.3 and show all the statistics evaluated after the application of the quasi-linear geostatistical approach (see section 3.2) with the addition of the optimization, the parameter estimation routines and the Marquardt modification (see section 6.2) to all the log-transmissivity fields studied.

In the following tables this symbols are shown:

- σ_R^2 is the variance of the error added to the measurements;
- θ_{est} is the parameter of the covariance matrix $\mathbf{Q}$ (cubic $Q_{ij} = \theta h_{ij}^3$), that allows the best fitting to the data, evaluated by the routine of the parameter estimation;
- $Q_1 = \frac{1}{n-1}\sum_{i=2}^{n}\varepsilon_i$ and $Q_2 = \frac{1}{n-p}\sum_{i=2}^{n}\varepsilon_i^2$, where ε_i are the orthonormal residuals and $n-p$ represents the degrees of freedom [see *Kitanidis,* 2004]. This two statistics regard the goodness of the chosen model. The model is rejected if $|Q_1| > \frac{2}{\sqrt{n-p}} = Q_{1crit}$; for the studied cases $Q_{1crit} = 0.44$. Regarding the Q_2 statistic, the model is rejected if $|Q_2 - 1| > \frac{2.8}{\sqrt{n-p}}$, simplifying $Q_{2crit1} < Q_2 < Q_{2crit2}$; in the studied problems $Q_{2crit1} = 0.389$ and $Q_{2crit2} = 1.611$;
- mean and variance of the recovered log-transmissivity field ($mean\left[\hat{\mathbf{s}}\right]$, $var\left[\hat{\mathbf{s}}\right]$) and of the differences between the true field and the recovered one ($mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$, $var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$);
- mean and variance of the square differences, $mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$, $var\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$;

- $D_1 = \frac{1}{m}\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^T \mathbf{V}^{-1}\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)$, this value should be close to 1; this value at last resulted not reliable because it is too dependent to the measurement error σ_R^2;
- $D_2 = E\left[\frac{\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2}{diag(\mathbf{V})}\right]$ is the mean of the square differences between the true field and the recovered one over the variance for each node; D_2 should be close to 1;
- mean and variance of the recovered transmissivity field ($mean\left[\hat{\mathbf{T}}\right]$, $var\left[\hat{\mathbf{T}}\right]$) and of the differences between the true field and the recovered one ($mean\left[\mathbf{T}_{true} - \hat{\mathbf{T}}\right]$, $var\left[\mathbf{T}_{true} - \hat{\mathbf{T}}\right]$);
- θ_0 is the initial value of the covariance parameter;
- Parameter used in the Marquardt modification.

Table A.1: Summary of the statistics of the heterogeneous field s.

σ_R^2	10^{-2}	10^{-4}	10^{-6}	10^{-8}
θ_{est}	6.1886	3.7084	6.6892	6.1065
Q_1	−0.1733	−0.0749	−0.1659	0.1058
Q_2	0.8741	1.0995	1.0448	1.0007
$mean\left[\hat{\mathbf{s}}\right]$	2.6013	2.5003	2.4905	2.4907
$var\left[\hat{\mathbf{s}}\right]$	0.1145	0.1080	0.0815	0.0775
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	−0.1013	−0.0003	0.0095	0.0093
$var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	−0.1013	−0.0003	0.0095	0.0093
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.1054	0.0246	0.0114	0.0129
$var\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.8099	2.6111	0.8502	1.2663
D_1	0.6895	1.1489	0.7793	15.585
D_2	0.7294	1.4293	0.7447	0.9496
$mean\left[\hat{\mathbf{T}}\right]$	14.218	12.861	12.581	12.554
$var\left[\hat{\mathbf{T}}\right]$	19.267	18.623	14.845	13.784
$mean\left[\mathbf{T}_{true} - \hat{\mathbf{T}}\right]$	−1.3754	−0.0180	0.2621	0.2887
$var\left[\mathbf{T}_{true} - \hat{\mathbf{T}}\right]$	21.5257	4.5794	2.4041	2.8185
θ_0	1	1	1	1
Marquardt	1	1	1	1

Table A.2: Summary of the statistics of the heterogeneous field 2s.

σ_R^2	10^{-2}	10^{-4}	10^{-6}	10^{-8}
θ_{est}	17.878	17.393	23.376	23.215
Q_1	0.0372	0.1635	0.1845	−0.2050
Q_2	1.0171	1.1595	1.0417	1.0412
$mean\left[\hat{\mathbf{s}}\right]$	2.3894	2.5294	2.4889	2.4913
$var\left[\hat{\mathbf{s}}\right]$	0.3318	0.2956	0.3200	0.3168
$mean\left[\mathbf{s}_{true}-\hat{\mathbf{s}}\right]$	0.1106	−0.0294	0.0111	0.0087
$var\left[\mathbf{s}_{true}-\hat{\mathbf{s}}\right]$	0.3903	0.1300	0.0443	0.0498
$mean\left[\left(\mathbf{s}_{true}-\hat{\mathbf{s}}\right)^2\right]$	0.4020	0.1307	0.0443	0.0498
$var\left[\left(\mathbf{s}_{true}-\hat{\mathbf{s}}\right)^2\right]$	0.4848	0.0664	0.0077	0.0110
D_1	0.9540	1.1275	3.7836	1061.7
D_2	2.0966	4.0596	1.3863	1.7213
$mean\left[\hat{\mathbf{T}}\right]$	12.711	14.489	14.311	14.318
$var\left[\hat{\mathbf{T}}\right]$	45.413	70.814	109.23	108.23
$mean\left[\mathbf{T}_{true}-\hat{\mathbf{T}}\right]$	2.4898	0.7115	0.8894	0.8826
$var\left[\mathbf{T}_{true}-\hat{\mathbf{T}}\right]$	155.29	60.735	22.563	24.705
θ_0	1	1	1	1
Marquardt	1	1	1	1

Table A.3: Summary of the statistics of the heterogeneous field 4s.

σ_R^2	10^{-2}	10^{-4}	10^{-6}	10^{-8}
θ_{est}	70.101	74.426	66.844	67.233
Q_1	0.4656	−0.5078	−0.0935	−0.0843
Q_2	1.2960	1.1183	1.1498	1.1544
$mean\left[\hat{\mathbf{s}}\right]$	2.3215	2.5629	2.3433	2.3367
$var\left[\hat{\mathbf{s}}\right]$	1.3740	1.3857	1.4385	1.4444
$mean\left[\mathbf{s}_{true}-\hat{\mathbf{s}}\right]$	0.1785	−0.0629	0.1567	0.1633
$var\left[\mathbf{s}_{true}-\hat{\mathbf{s}}\right]$	1.3121	0.2401	0.1947	0.1979
$mean\left[\left(\mathbf{s}_{true}-\hat{\mathbf{s}}\right)^2\right]$	1.3422	0.2438	0.2190	0.2243
$var\left[\left(\mathbf{s}_{true}-\hat{\mathbf{s}}\right)^2\right]$	5.5996	0.1630	0.1210	0.1309
D_1	1.4285	2.5596	3.7698	237.18
D_2	2.1644	1.0958	1.7085	1.7394
$mean\left[\hat{\mathbf{T}}\right]$	17.398	27.353	20.181	20.128
$var\left[\hat{\mathbf{T}}\right]$	294.59	2813.3	851.99	857.28
$mean\left[\mathbf{T}_{true}-\hat{\mathbf{T}}\right]$	14.889	4.9343	12.106	12.159
$var\left[\mathbf{T}_{true}-\hat{\mathbf{T}}\right]$	6608.6	2048.3	3370.9	3394.6
θ_0	60	60	60	60
Marquardt	1	1	1	1

Table A.4: Summary of the statistics of the heterogeneous field 6s.

σ_R^2	10^{-2}	10^{-4}	10^{-6}	10^{-8}
θ_{est}	199.98	222.43	228.57	231.48
Q_1	0.1016	−0.1261	−0.1059	−0.1398
Q_2	0.6875	0.7815	0.6311	0.6255
$mean\left[\hat{\mathbf{s}}\right]$	2.4166	2.3484	1.9581	1.9613
$var\left[\hat{\mathbf{s}}\right]$	2.5144	2.8770	2.3641	2.3700
$mean\left[\mathbf{s}_{true}-\hat{\mathbf{s}}\right]$	0.0834	0.1516	0.5419	0.5387
$var\left[\mathbf{s}_{true}-\hat{\mathbf{s}}\right]$	1.0380	0.3988	1.0007	0.9848
$mean\left[\left(\mathbf{s}_{true}-\hat{\mathbf{s}}\right)^2\right]$	1.0436	0.4213	1.2930	1.2737
$var\left[\left(\mathbf{s}_{true}-\hat{\mathbf{s}}\right)^2\right]$	3.3809	0.7286	9.0656	8.7325
D_1	0.8428	25.222	1921.2	67101.0
D_2	0.9674	0.7559	1.5293	1.4929
$mean\left[\hat{\mathbf{T}}\right]$	33.762	46.738	18.731	18.759
$var\left[\hat{\mathbf{T}}\right]$	3365.4	22557.1	1316.9	1273.1
$mean\left[\mathbf{T}_{true}-\hat{\mathbf{T}}\right]$	86.103	73.127	101.13	101.11
$var\left[\mathbf{T}_{true}-\hat{\mathbf{T}}\right]$	309126.8	205355.1	326630.3	326265.7
θ_0	20	20	20	20
Marquardt	4	4	4	4

Table A.5: Summary of the statistics of the heterogeneous fields 8s and 10s.

	8s		10s	
σ_R^2	10^{-2}	10^{-4}	10^{-2}	10^{-4}
θ_{est}	75.769	441.15	600.01	1137.6
Q_1	-0.1845	-0.1747	0.4919	0.2282
Q_2	1.7221	0.7497	0.8978	0.5056
$mean[\hat{\mathbf{s}}]$	2.3020	2.0722	1.6616	1.8665
$var[\hat{\mathbf{s}}]$	3.9898	5.1689	6.6663	7.2431
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.1980	0.4278	0.8384	0.6335
$var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	3.0401	2.5881	4.7204	3.8181
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	3.0753	2.7678	5.4171	4.2144
$var\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	42.551	72.665	167.277	149.43
D_1	4.0028	146.69	16.122	1071.1
D_2	5.1077	0.9438	1.5247	0.7245
$mean[\hat{\mathbf{T}}]$	47.895	74.546	67.467	179.21
$var[\hat{\mathbf{T}}]$	9059.4	84754.1	42178.2	1080478.7
$mean[\mathbf{T}_{true} - \hat{\mathbf{T}}]$	594.6958	568.05	4045.7	3933.9
$var[\mathbf{T}_{true} - \hat{\mathbf{T}}]$	$1.8491 \cdot 10^7$	$1.7034 \cdot 10^7$	$1.0802 \cdot 10^9$	$1.0608 \cdot 10^9$
θ_0	600	600	600	600
Marquardt	15	15	15	15

Appendix B

Results of the Quasi-Linear Geostatistical Approach with Conditional Realizations Applied to the Estimation of the Aquifer Parameters

This appendix contains the results obtained using the quasi-linear approach with the addition of the MCMC method and the Metropolis Hastings, as acceptance rejection method (see chapter 6.3 for the theory and appendix E for the numerical codes) of all the transmissivity field studied.

Three different measurement error were considered: $\sigma_R^2 = 10^{-2}$, $\sigma_R^2 = 10^{-4}$ and $\sigma_R^2 = 10^{-6}$.

Summary of 1000 CR applied to the s field transmissivity field, $\sigma_R^2 = 10^{-2}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	23.200	21.976	21.473	19.258	16.966
$mean\left[\hat{\mathbf{s}}\right]$	2.6146	2.6131	2.6097	2.6087	2.5961
$var\left[\hat{\mathbf{s}}\right]$	0.1813	0.1927	0.2047	0.2139	0.2068
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	−0.1146	−0.1131	−0.1097	−0.1087	−0.0961
$var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.0834	0.0834	0.0840	0.0825	0.0745
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.0964	0.0960	0.0959	0.0942	0.0836
$var\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.0157	0.0150	0.0144	0.0130	0.0115
D_2	0.6709	0.6745	0.6826	0.6745	0.6035

Summary of 1000 CR applied to the s field transmissivity field, $\sigma_R^2 = 10^{-4}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	14.700	13.541	13.733	12.981	13.373
$mean\left[\hat{\mathbf{s}}\right]$	2.5008	2.5004	2.5007	2.4999	2.4962
$var\left[\hat{\mathbf{s}}\right]$	0.1434	0.1467	0.1478	0.1525	0.1522
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	−0.0008	−0.0004	−0.0007	0.0001	0.0038
$var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.0432	0.0457	0.0466	0.0483	0.0484
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.0431	0.0457	0.0465	0.0483	0.0483
$var\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.0038	0.0040	0.0041	0.0045	0.0044
D_2	5.2670	5.9173	6.2049	6.9348	7.2684

Summary of 1000 CR applied to the s field transmissivity field, $\sigma_R^2 = 10^{-6}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	10.700	9.8779	10.112	9.8431	12.575
$mean\left[\hat{\mathbf{s}}\right]$	2.4974	2.4947	2.4929	2.4907	2.4899
$var\left[\hat{\mathbf{s}}\right]$	0.1098	0.1079	0.1068	0.1028	0.1009
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.0026	0.0053	0.0071	0.0093	0.0101
$var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.1098	0.1079	0.1068	0.1028	0.1009
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.0090	0.0093	0.0093	0.0095	0.0097
$var\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.5908	0.5549	0.5243	0.5042	0.6221
D_2	0.8031	0.8529	0.8662	0.9070	0.9315

Summary of 1000 CR applied to the 2s field transmissivity field, $\sigma_R^2 = 10^{-2}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	27.100	26.193	24.844	25.678	25.349
$mean\left[\overset{\wedge}{\mathbf{s}}\right]$	2.3654	2.3688	2.3781	2.3843	2.3767
$var\left[\overset{\wedge}{\mathbf{s}}\right]$	0.4103	0.4109	0.4197	0.4279	0.4379
$mean\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	0.1346	0.1312	0.1219	0.1157	0.1233
$var\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	0.4622	0.4369	0.4113	0.3996	0.4137
$mean\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	0.4797	0.4535	0.4256	0.4124	0.4284
$var\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	0.5908	0.5549	0.5243	0.5042	0.6221
D_2	1.7190	1.6177	1.4980	1.4405	1.4082

Summary of 1000 CR applied to the 2s field transmissivity field, $\sigma_R^2 = 10^{-4}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	8.4000	7.7691	6.4919	6.2767	6.7864
$mean\left[\overset{\wedge}{\mathbf{s}}\right]$	2.5383	2.5386	2.5368	2.5350	2.5295
$var\left[\overset{\wedge}{\mathbf{s}}\right]$	0.3397	0.3398	0.3385	0.3351	0.3286
$mean\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	−0.0383	−0.0386	−0.0368	−0.0350	−0.0295
$var\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	0.1966	0.2098	0.2331	0.2425	0.2574
$mean\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	0.1978	0.2110	0.2341	0.2434	0.2579
$var\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	0.1782	0.2032	0.2394	0.2476	0.2728
D_2	2.1953	2.3043	2.5561	2.6893	2.8578

Summary of 1000 CR applied to the 2s field transmissivity field, $\sigma_R^2 = 10^{-6}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	11.600	10.877	11.610	11.840	12.176
$mean\left[\overset{\wedge}{\mathbf{s}}\right]$	2.4961	2.4915	2.4891	2.4861	2.4834
$var\left[\overset{\wedge}{\mathbf{s}}\right]$	0.4487	0.4386	0.4353	0.4262	0.4101
$mean\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	0.0039	0.0085	0.0109	0.0139	0.0166
$var\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	0.0326	0.0325	0.0328	0.0333	0.0351
$mean\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	0.0326	0.0325	0.0328	0.0333	0.0351
$var\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	0.0033	0.0032	0.0033	0.0035	0.0044
D_2	0.8544	0.8943	0.9164	0.9550	1.0351

Summary of 1000 CR applied to the 4s field transmissivity field, $\sigma_R^2 = 10^{-2}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	1.0000	0.3330	0.1248	0.1427	0.1996
$mean\left[\hat{\mathbf{s}}\right]$	2.5805	2.5979	2.4163	2.4163	2.4163
$var\left[\hat{\mathbf{s}}\right]$	1.9970	2.4831	2.6333	2.6333	2.6333
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	-0.0805	-0.0979	0.0837	0.0837	0.0837
$var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.5622	0.7369	1.1662	1.1662	1.1662
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.5680	0.7456	1.1717	1.1717	1.1717
$var\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.6402	0.6511	1.9476	1.9476	1.9476
D_2	0.8913	1.1931	2.7043	2.7043	2.7043

Summary of 1000 CR applied to the 4s field transmissivity field, $\sigma_R^2 = 10^{-4}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	12.700	12.098	12.484	12.411	10.579
$mean\left[\hat{\mathbf{s}}\right]$	2.5114	2.5111	2.5087	2.5045	2.4803
$var\left[\hat{\mathbf{s}}\right]$	1.7543	1.7615	1.7587	1.7429	1.7278
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	−0.0114	−0.0111	−0.0087	−0.0045	0.0197
$var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.2205	0.2298	0.2351	0.2394	0.2552
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.2204	0.2296	0.2349	0.2391	0.2553
$var\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.0964	0.1003	0.1046	0.1073	0.1228
D_2	1.3036	1.3458	1.3628	1.3823	1.4884

Summary of 1000 CR applied to the 4s field transmissivity field, $\sigma_R^2 = 10^{-6}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	8.7174	8.8988	8.3855	8.8698	11.200
$mean\left[\hat{\mathbf{s}}\right]$	2.3771	2.3691	2.3564	2.3529	2.3410
$var\left[\hat{\mathbf{s}}\right]$	1.4509	1.4474	1.4615	1.4894	1.5227
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.1229	0.1309	0.1436	0.1471	0.1590
$var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.2126	0.2175	0.2183	0.2234	0.2402
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.2275	0.2344	0.2387	0.2448	0.2651
$var\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	0.1309	0.1307	0.1021	0.1063	0.1338
D_2	1.7719	1.8357	2.0040	2.1188	2.3016

Summary of 1000 CR applied to the 6s field transmissivity field, $\sigma_R^2 = 10^{-2}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	1.2000	0.4440	0.3745	0.4280	0.3992
$mean\left[\overset{\wedge}{\mathbf{s}}\right]$	2.7488	3.0319	3.0483	3.0483	3.0176
$var\left[\overset{\wedge}{\mathbf{s}}\right]$	4.6245	5.8568	6.0155	6.0155	6.0928
$mean\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	−0.2488	−0.5319	−0.5483	−0.5483	−0.5176
$var\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	1.7014	2.1644	2.1770	2.1770	2.2343
$mean\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	1.7611	2.4446	2.4748	2.4748	2.4993
$var\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	8.7193	17.607	17.899	17.899	16.128
D_2	1.4179	1.4199	1.4257	1.4257	1.5982

Summary of 1000 CR applied to the 6s field transmissivity field, $\sigma_R^2 = 10^{-4}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	5.5000	4.8835	3.8702	3.8516	4.7904
$mean\left[\overset{\wedge}{\mathbf{s}}\right]$	2.0286	2.0054	2.0231	2.0336	2.0309
$var\left[\overset{\wedge}{\mathbf{s}}\right]$	3.9993	4.0939	3.8171	3.6978	3.6738
$mean\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	0.4714	0.4946	0.4769	0.4664	0.4691
$var\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	3.1499	3.3223	2.9676	2.7905	2.7368
$mean\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	3.3680	3.5626	3.1912	3.0044	2.9532
$var\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	109.84	121.14	88.199	73.840	68.933
D_2	2.1858	2.2904	2.2851	2.2917	2.3094

Summary of 1000 CR applied to the 6s field transmissivity field, $\sigma_R^2 = 10^{-6}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	10.721	9.5662	9.1364	10.300	11.623
$mean\left[\overset{\wedge}{\mathbf{s}}\right]$	2.2850	2.3503	2.3979	2.4022	2.4608
$var\left[\overset{\wedge}{\mathbf{s}}\right]$	5.2675	4.9383	4.8570	4.8541	4.8224
$mean\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	0.2150	0.1497	0.1021	0.0978	0.0392
$var\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	3.1548	2.5804	2.2625	2.2364	1.9016
$mean\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	3.1969	2.5995	2.2700	2.2431	1.9007
$var\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	133.00	81.992	61.300	59.743	40.767
D_2	2.7767	2.6761	2.6273	2.6212	2.5555

Summary of 1000 CR applied to the 8s field transmissivity field, $\sigma_R^2 = 10^{-2}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	2.4000	0.8879	0.2497	0.2853	0
$mean\left[\hat{\mathbf{s}}\right]$	2.1721	2.5652	2.5458	2.5458	
$var\left[\hat{\mathbf{s}}\right]$	4.7947	5.9960	5.9627	5.9627	
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.3279	−0.0652	−0.0458	−0.0458	
$var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	2.6049	1.5989	1.6946	1.6946	
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	2.7090	1.6010	1.6945	1.6945	
$var\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	41.951	8.5707	9.7299	9.7299	
D_2	1.0045	0.8738	0.9868	0.9868	

Summary of 1000 CR applied to the 8s field transmissivity field, $\sigma_R^2 = 10^{-4}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	12.500	11.765	11.860	10.842	8.5828
$mean\left[\hat{\mathbf{s}}\right]$	2.0439	1.9924	1.9708	1.9117	2.0104
$var\left[\hat{\mathbf{s}}\right]$	6.5280	6.8529	7.1166	7.4469	7.2114
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.4561	0.5076	0.5292	0.5883	0.4896
$var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	2.9071	3.3578	3.5900	4.0173	2.9229
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	3.1113	3.6111	3.8654	4.3582	3.1589
$var\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	127.87	181.07	213.56	260.71	124.78
D_2	1.3564	1.4490	1.4884	1.5861	1.4249

Summary of 1000 CR applied to the 10s field transmissivity field, $\sigma_R^2 = 10^{-2}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	1.2000	0.2220	0.1248	0.1427	0.1996
$mean\left[\hat{\mathbf{s}}\right]$	2.2084	2.0900	1.9407	1.9407	1.9407
$var\left[\hat{\mathbf{s}}\right]$	9.6817	8.9491	8.4978	8.4978	8.4978
$mean\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	0.2916	0.4100	0.5593	0.5593	0.5593
$var\left[\mathbf{s}_{true} - \hat{\mathbf{s}}\right]$	7.9642	7.3006	6.9196	6.9196	6.9196
$mean\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	8.0389	7.4593	7.2235	7.2235	7.2235
$var\left[\left(\mathbf{s}_{true} - \hat{\mathbf{s}}\right)^2\right]$	165.12	136.37	105.19	105.19	105.19
D_2	6.5316	6.4771	6.8691	6.8691	6.8691

Summary of 1000 CR applied to the 10s field transmissivity field, $\sigma_R^2 = 10^{-4}$

From CR	1	100	200	300	500
To CR	1000	1000	1000	1000	1000
% Accepted CR	11.000	10.655	11.610	12.553	9.5808
$mean\left[\overset{\wedge}{\mathbf{s}}\right]$	2.1582	2.1197	2.1132	2.0994	2.0456
$var\left[\overset{\wedge}{\mathbf{s}}\right]$	8.2021	8.0912	8.0706	8.0510	8.1264
$mean\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	0.3418	0.3803	0.3868	0.4006	0.4544
$var\left[\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right]$	1.3525	1.4987	1.5210	1.5699	1.7975
$mean\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	1.4676	1.6413	1.6687	1.7283	2.0016
$var\left[\left(\mathbf{s}_{true} - \overset{\wedge}{\mathbf{s}}\right)^2\right]$	9.1360	11.341	11.580	12.254	14.633
D_2	1.5797	1.7853	1.8074	1.8490	2.1425

Appendix C

Recovery of the Pollutants Release History

This appendix includes the source codes (in Matlab, .m files) used for the recovery of the pollutant release history, see *Snodgrass and Kitanidis* [1997] and sections 4.2, 5.1 for the details of the procedure. The unconstrained case and the constrained case are shown. Both main programs use few functions, in the following are reported only the most important.

C.1 Evaluation of the Numerical Transfer Function: `TF.m`

```
% This program evaluate the transfer function and extract the concentration data
% at a specified time from the output file of GMS.
% AZ last version: June 2005

clear,clc

flag = 0;
while flag == 0
    choice = input('Do you want to evaluate the TF, (a), or to measure the ...
            ...concentration at a specified time, (b),?\n','s');
    if isempty(choice)
        flag = 0;
    else
        if (choice == 'a' )|( choice == 'b')
            flag = 1;
        else
            flag = 0;
        end
    end
end

flag = 0;
```

```
while flag == 0
    file_in = input('Insert the name of the input file: ','s');
    if isempty(file_in)
        flag = 0;
    else
        flag = 1;
    end
end

flag = 0;
while flag == 0
    m = input('Insert the number of the measurements: ');
    if isempty(m)
        flag = 0;
    else
        flag = 1;
    end
end

% source location
source_X = 229;
source_Y = 25;

% input file containing the distance from the source
% 1# column point name P1, P2, ..., Pm
% 2# column X
% 3# column Y
point_m = load ('punti.txt');
dist = zeros(m,1);
for i = 1:m
    dist(i,1) = sqrt((point_m(i,2)-source_X)^2+(point_m(i,3)-source_Y)^2);
end

% input file exported from GMS
% 1# column point name
% 2# column measurement number
% 3# column Time
% 4# column Concentration
[Punto,T,C] = textread(file_in,'%f %*f %f %f');

p = size (T,1);
n = p/m;  % n: time intervals

deltat = T(3,1)-T(2,1);
disp(['delta T = ',num2str(deltat)]);

% evaluation of the transfer function
```

```
if choice == 'a'

% output MATRIX
% 1# column time
% 2# column fdt P1
% 3# column fdt P2
% 4# column fdt P3
% ................
% m+1# column fdt Pm

out = zeros(n,m+1);
out(:,1) = T (1:n,1);

% building of the TF, backward formulation
for i = 1:m
    iuh = zeros(n,1);
    iuh(1,1) = (C(1+n*(i-1),1)-0)/deltat;
    for j = 2:n
        iuh(j,1) = (C(j+n*(i-1),1)-C(j+n*(i-1)-1,1))/deltat;
    end
    col = Punto (1+n*(i-1),1) + 1;
    out(:,col) = iuh(:,1);
    figure (i)
    plot (out(:,1),out(:,col))
end

% measured concentration at time T
else

    flag = 0;
    while flag == 0
        TIME = input('Insert the time of the measurements as a multiple of ...
                    ...delta T ');
        if isempty(TIME)
            flag = 0;
        else
            tt = TIME/deltat;
            if tt > n
                disp '(Time exceed the measurement')
                flag = 0;
            else
                flag = 1;
            end
        end
    end

    % output file
```

```
    % concentration P1 at Time TIME
    % concentration P2 at Time TIME
    % concentration P3 at Time TIME
    % ..........................
    % concentration Pm at Time TIME
    out = zeros (m,1);
    for i = 1:m
        col = Punto (tt+n*(i-1));
        out(col,1) = C(tt+n*(i-1),1);
    end
    plot (dist,out,'or');
end

% results printed in an output file
flag = 0;
while flag == 0
    fileout = input('Insert the name of the output file: ','s');
    if isempty(fileout)
        flag = 0;
    else
        flag = 1;
    end
end
save (fileout,'out','-ascii')
```

C.2 Input File: `in.m`

```
% input file
% This file contains the parameter that are necessary for the application
% of the GA
% AZ last version: June 2005

function [TF_filename,point_filename,out_filename,startsol,thetamax,theta0,n,...
              ...m,lambda,sigma2r,TT,ST,maxerror,alfa] = in(a)

% name of the file that contains the transfer function
% transfer functions
% #1 column: Time
% #2 column: TF P1
% #3 column: TF P2
% #4 column: TF P3
% ................
% #m column: TF Pm
TF_filename = 'TF_1.txt';

% name of the file that contains the measurements
% #1 row: concentration P1
```

```
% #2 row: concentration P2
% #3 row: concentration P3
% ......................
% #m row: concentration Pm
point_filename = 'misure_1.txt';

% name of the file that contains the results
% #1 column: Time
% #2 column: confidence interval
% #3 column: estimate
% #4 column: confidence interval
out_filename = 'resultV_1.txt';

% convergence criterion on sigma2
thetamax(1,1) = 0.00001;
% convergence criterion on l
thetamax(2,1) = 0.000001;

% initial parameters
theta0(1,1) = 0.04;
theta0(2,1) = 13.0;

% number of time intervals
n = 300;

% number of measurements
m = 24;

% Marquardt modification
lambda = 0.5;

% error
sigma2r = 1E-12;

% Total time
TT = 300*10000;
% Sampling time
ST = 300*10000;

% CONSTRAINED CASE
% starting solution
startsol = 'result_NV.txt';

alfa = 2;
maxerror = 0.8;
```

C.3 Unconstrained Case: `unconstrained.m`

```
% program that applies the Quasi Linear Geostatistical approach,
% UNCONSTRAINED CASE, (Snodgrass & Kitanidis, 1997).
% AZ last version: June 2005

clear, clc
tic
% input function and data
[TF_filename,point_filename,out_filename,startsol,thetamax,theta,n,m,lambda,...
               ...sigma2r,TT,ST,maxerror] = in (1);

TF = load (TF_filename);  % TF (n,m+1)
TF = TF';
Z = load (point_filename);  % Z (m,1)
Zt = Z';
Time = TF (1,1:n);

hd = zeros(1,n);
for i = 1:n
    hd(1,i) = Time(1,i)-Time(1,1);
end

DT = TT/n;
R = sigma2r*eye(m);       % R (m,m)
X = ones(n,1);            % X (n,1)
Xt = X';

H = DT*TF(2:m+1,1:n);      % H (m,n)
Ht = H';

HX = H*X;
XtHt = HX';

% evaluation of parameters sigma2 and l
Q = zeros (n,n);

L = 1E100;
iteration = 1;
flag = 0;
while flag == 0 %parameter estimation loop

% covariance matrix
gen = exp(-(hd.^2)/theta(2,1)^2);
gen2 = gen.*hd.^2;
Qdevsig = toepls(gen);
Q = theta(1,1)*Qdevsig;
```

```
Qdevl = theta(1,1)*toepls(gen2)*2/(theta(2,1)^3);

% sigma matrix (m,m)
HQ = H*Q;
sigma = HQ*Ht+R;
sigmainv = inv(sigma);

% sigma derivate sigma2, theta(1,1) (m,m)
sigmadevsig = H*Qdevsig*Ht;

% sigma derivate l, theta(2,1)   (m,m)
sigmadevl = H*Qdevl*Ht;

% ksi matrix  (m,m)
core = XtHt*sigmainv*HX;
coreinv = inv(core);
ksi = sigmainv-sigmainv*HX*coreinv*XtHt*sigmainv;

a = ksi*sigmadevsig;
b = ksi*sigmadevl;

% vector g
g(1,1) = 0.5*trace(a)-0.5*Zt*(a*ksi)*Z;
g(2,1) = 0.5*trace(b)-0.5*Zt*(b*ksi)*Z;

% Fisher information matrix
Fisher(1,1) = 0.5*trace(a*a);
Fisher(1,2) = 0.5*trace(a*b);
Fisher(2,1) = 0.5*trace(b*a);
Fisher(2,2) = 0.5*trace(b*b);


% Marquardt modification
Marquardt = Fisher+lambda*eye(2);
Marquinv = inv(Marquardt);

correction = Marquinv*g;
thetaold = theta;
theta = thetaold-correction;

disp (['iteration #',num2str(iteration),'   sigma2 ',num2str(theta(1,1)),...
                '...'     l ',num2str(theta(2,1))])

if (abs(correction(1,1))<=thetamax(1,1))&(abs(correction(2,1))<=thetamax(2,1))
    flag =1;
else
    iteration = iteration+1;
```

```
end

end % flag loop

% Quasi linear inversion method

% The system (11) of the paper Snodgrass & Kitanidis, 1997, is simplyfied
% as: A*B=C, the matrix B contains the unknowns.

A = zeros (m+1,m+1);
B = zeros (m+1,n);
C = zeros (m+1,n);

A(1:m,1:m) = sigma;
A(m+1,1:m) = XtHt;
A(1:m,m+1) = HX;

C(1:m,1:n) = HQ;
C(m+1,1:n) = Xt;

B = A\C;
Lamt = zeros(m,n);
MUL = zeros(1,n);
Lamt = B(1:m,1:n);
Lam = Lamt';
% matrix of multipliers
MUL = B(m+1,1:n);

% best estimate
s = Lam*Z;

% Covariance
V = -X*MUL+Q-Q*Ht*Lamt;

% evaluation of the confidence interval

CIsup = s+sqrt(2*diag(V));
CIinf = s-sqrt(2*diag(V));

plot (Time,s,'r',Time,CIsup,'y',Time,CIinf,'y')

OUT = zeros (n,4);
OUT(1:n,1) = Time;
OUT(1:n,2) = CIinf;
OUT(1:n,3) = s;
OUT(1:n,4) = CIsup;
```

```
save (out_filename,'OUT','-ascii')
clear OUT
toc
```

C.4 Constrained Case: `constrained.m`

```
% program that applies the Quasi Linear Geostatistical approach,
% CONSTRAINED CASE, (Snodgrass & Kitanidis, 1997).
% AZ last version: June 2005

clear, clc
tic
% input function and data
[TF_filename,point_filename,out_filename,startsol,thetamax,theta,n,m,lambda,...
                 ...sigma2r,TT,ST,maxerror,alfa] = in (1);

TF = load (TF_filename);  % TF (n,m+1)
TF = TF';
Z_in = load (point_filename);  % Z (m,1)

Time = TF (1,1:n);

hd = zeros(1,n);
for i = 1:n
    hd(1,i) = Time(1,i)-Time(1,1);
end

DT = TT/n;
R = sigma2r*eye(m);       % R (m,m)
X = ones(n,1);            % X (n,1)
Xt = X';

% evaluation of parameters sigma2 and l
Q = zeros (n,n);

s0 = zeros(n,1);
% starting solution
stsol = load (startsol);
s0 = stsol(:,3);
clear stsol;
for i = 1:n;
    if s0(i,1) <= 0
    s0(i,1) = 0;
    end
end

outer_iteration = 1;
```

```
outer_flag = 0;
%_____________________________________________________________________________
while outer_flag == 0;

% building of hso
hso = zeros(m,1);
for i = 1:m
    sum = 0;
    for j = 1:n
        sum = sum+(((s0(j,1)+alfa)/alfa)^alfa)*TF(i+1,j);
    end
    hso(i,1) = sum;
end

% building of the H matrix
H = zeros(m,n);
in = 0;
for j = 1:n
    in = ((s0(j,1)+alfa)/alfa)^(alfa-1);
    for i = 1:m
        H(i,j) = in*TF(i+1,j)*DT;
    end
end

% H (m,n)
Ht = H';

HX = H*X;
XtHt = HX';

% building of Zol
Zol = zeros(m,1);
Zol = Z_in-hso+H*s0;

Z = Zol;
Zt = Z';

inner_iteration = 1;
inner_flag = 0;
%_____________________________________________________________________________
while inner_flag == 0 %parameter estimation loop

% covariance matrix
gen = exp(-(hd.^2)/theta(2,1)^2);
gen2 = gen.*hd.^2;
Qdevsig = toepls(gen);
Q = theta(1,1)*Qdevsig;
```

```
Qdevl = theta(1,1)*toepls(gen2)*2/(theta(2,1)^3);

% sigma matrix (m,m)
HQ = H*Q;
sigma = HQ*Ht+R;
sigmainv = inv(sigma);

% sigma derivate sigma2, theta(1,1) (m,m)
sigmadevsig = H*Qdevsig*Ht;

% sigma derivate l, theta(2,1)   (m,m)
sigmadevl = H*Qdevl*Ht;

% ksi matrix  (m,m)
core = XtHt*sigmainv*HX;
coreinv = inv(core);
ksi = sigmainv-sigmainv*HX*coreinv*XtHt*sigmainv;

a = ksi*sigmadevsig;
b = ksi*sigmadevl;

% vector g
g(1,1) = 0.5*trace(a)-0.5*Zt*(a*ksi)*Z;
g(2,1) = 0.5*trace(b)-0.5*Zt*(b*ksi)*Z;

% Fisher information matrix
Fisher(1,1) = 0.5*trace(a*a);
Fisher(1,2) = 0.5*trace(a*b);
Fisher(2,1) = 0.5*trace(b*a);
Fisher(2,2) = 0.5*trace(b*b);


% Marquardt modification
Marquardt = Fisher+lambda*eye(2);
Marquinv = inv(Marquardt);

correction = Marquinv*g;
thetaold = theta;
theta = thetaold-correction;

disp (['iteration #',num2str(inner_iteration),'   sigma2 ' ,num2str(theta(1,1)),...
                '...'     l ',num2str(theta(2,1))])
if (abs(correction(1,1))<=thetamax(1,1))&(abs(correction(2,1))<=thetamax(2,1))
    inner_flag =1;
else
    inner_iteration = inner_iteration+1;
end
```

```
end % inner flag loop
%_____________________________________________________________________________

% Quasi linear inversion method

% The system (11) of the paper Snodgrass & Kitanidis, 1997, is simplyfied
% as: A*B=C, the matrix B contains the unknowns.

A = zeros (m+1,m+1);
B = zeros (m+1,n);
C = zeros (m+1,n);

A(1:m,1:m) = sigma;
A(m+1,1:m) = XtHt;
A(1:m,m+1) = HX;

C(1:m,1:n) = HQ;
C(m+1,1:n) = Xt;

B = A\C;
Lamt = zeros(m,n);
MUL = zeros(1,n);
Lamt = B(1:m,1:n);
Lam = Lamt';
% matrix of multipliers
MUL = B(m+1,1:n);

% best estimate
s = Lam*Z;

% convergence criterion
diff = abs(s-s0);
maxdiff = max(diff);
disp(['max difference: ',num2str(maxdiff)])

if maxdiff<=maxerror
    outer_flag = 1;
else
    outer_iteration = outer_iteration+1;
    s0 = s;
end


end % external_flag
%_____________________________________________________________________________
```

```
% Covariance
V = -X*MUL+Q-Q*Ht*Lamt;

% evaluation of the confidence interval

CIsup = s+sqrt(2*diag(V));
CIinf = s-sqrt(2*diag(V));

CIsup_fin = ((CIsup+alfa)./alfa).^alfa;
CIinf_fin = ((CIinf+alfa)./alfa).^alfa;

s_fin = ((s+alfa)./alfa).^alfa;

figure(2)
plot (Time,s_fin,'r',Time,CIsup_fin,'y',Time,CIinf_fin,'y')

OUT = zeros (n,4);
OUT(1:n,1) = Time;
OUT(1:n,2) = CIinf_fin;
OUT(1:n,3) = s_fin;
OUT(1:n,4) = CIsup_fin;

save (out_filename,'OUT','-ascii')
clear OUT
toc
```

Appendix D

Aquifer Parameter Estimation: Quasi-Linear

This appendix includes the source codes (in Matlab, .m files) used for the transmissivity estimation from head measurements using the quasi-linear geostatistical approach, see *Kitanidis* [1995] and section 3.2 for the details of the procedure. In the following are reported only the most important functions.

D.1 Main Program: `QL.m`

```
% program that applies the Quasi Linear Geostatistical Approach,
% (Kitanidis, 1995).
% Original code: illust4.m from P. K. Kitanidis
% AZ last version: December 2005
clear,clc
%-----------------------------------------------------------
% Load Geometry
a=1;
[L2,Re,s_loc,nr,nc,xgrid,ygrid] = geometry(a);
% Load input data
[y_data,k_sample,initial,sig2,theta,thetamax,convergence,cov,lambda,Tm_true,...
...itermax,ro,N] = in(a);
n = size(y_data,1);
m=nr*nc;

tic

% evaluation of the distance between the grid point
hd = zeros(1,n);
for ii=1:m
      hd(1,ii) = sqrt((s_loc(1,1)-s_loc(ii,1))^2+(s_loc(1,2)-s_loc(ii,2))^2);
end
```

```
%INVERSE PROBLEM
%-------------------------------------------------------------
%PARAMETERS  - assumed known
%measurement error matrix
sig = sqrt(sig2); R = eye(n)*sig2; invR=inv(R);
%structure of unknown function:

% Covariance function built using the Toeplitz properties
if cov == 1      %linear
    p = 1;
    X = ones(m,p);
    gen = -hd;
    gen = reshape(gen,nc,nr);
    gen = gen';
    Qdev1 = toeplsbtstb(gen);
    Q = theta(1,1)*Qdev1;
elseif cov == 2   %cubic
    p = 3;
    X = ones(m,p);
    X(:,2:3) = s_loc(:,1:2);
    gen2 = hd.^3;
    gen2 = reshape(gen2,nc,nr);
    gen2 = gen2';
    Qdev2 = toeplsbtstb(gen2);
    Q = theta(1,1)*Qdev2;
end

%-------------------------------------------------------------
%Begin inverse problem
%Initial s estimate
sm = initial*ones(nr,nc); Tm = exp(sm); s = stackt(sm);

linvflag = 0; %linvflag = 0 indicates quasi-linear inversing has not converged
iter = 0; SSR = 1E100; Lm = 1E100; %initialize
%-------------------------------------------------------------


while linvflag == 0             %linvflag loop->
    iter = iter+1
    %solve forward problem for estimate
    figure(3)
    mesh(xgrid, ygrid, sm), xlabel('x'); ylabel('y'); zlabel('est. logtrans.');
    title('Estimated Logtransmissivity')
    pause(1)

    SSR_old = SSR;  sm_old = sm; L_old = Lm;
```

```
[A, b, CR, CC] = mkls(Tm,Re,L2);
h = full(A\b); head = reshapet(h,nr,nc);

figure(4)
mesh(xgrid, ygrid, head), xlabel('x'); ylabel('y'); zlabel('est. head');

%sample from predicted head, find SS of residuals
y_pred = h(k_sample);                 %y_pred is predicted observation
res = (y_data-y_pred); SSR = res'*res
if iter>1
    Lm = 0.5*(ksi'*HQHT*ksi+SSR/sig2);
    disp(['L is ', num2str(Lm)]);
    if Lm>L_old
        disp('L increased!')
    end
end
dh = adjstatm2(k_sample,Tm,L2,A,h,CR,CC);

for i = 1:n
    dh(:,:,i) = dh(:,:,i).*exp(sm);
end

H = zeros(n,m);
for i = 1:n
    H(i,:) = stackt(dh(:,:,i))';
end
y = y_data - y_pred + H*stackt(sm);

%parameter estimation
if iter>1
    [Q,thetanew] = PE(theta,H,X,R,ksi,SSR,sig2,thetamax,hd,n,y,cov,...
                  ...lambda,nc,nr);
    theta = thetanew
end

PHI = H*X; QHT = Q*H'; HQHT = H*QHT;
PSI = HQHT+R;

invQ = inv(Q);
G = invQ-invQ*X*inv(X'*invQ*X)*X'*invQ;

sol = [PSI,PHI;PHI',zeros(p,p)]\[y;zeros(p,1)];
ksi = sol(1:n,1); beta = sol(n+1:n+p,1);

sold = s;
snew = X*beta + QHT*ksi;
```

```
    %optimization along the line
    objft = @(delta)opti2(delta,invR,G,sold,snew,y_data,Re,L2,k_sample,nr,nc);
    [delta,val,fl,ou] = fminbnd(objft,-0.1,1.1);
    delta
    s = sold*delta+snew*(1-delta);

    sm = reshapet(s,nr,nc); Tm = exp(sm);
    if iter>1 & abs(Lm-L_old)<convergence
        linvflag = 1;
        INVMAT = inv([PSI,PHI;PHI',zeros(p,p)]);
        Pyy = INVMAT(1:n,1:n);
        Pyb = INVMAT(1:n,n+1:n+p);
        Pbb = INVMAT(n+1:n+p,n+1:n+p);
        V = Q - QHT*Pyy*QHT' - X*Pbb*X' - X*Pyb'*QHT'-QHT*Pyb*X';
    end
end              %linvflag loop<-

Res = zeros(m,3);
Res(:,2:3) = s_loc;
Res(:,1) = s;

diag_V=diag(V);
save('QL_s_err4.txt','A','-ascii')
save('V_s_err4.txt','diag_V','-ascii')
save('head_s_err4.txt','head','-ascii')

time_el = toc

% evaluation of the results
oo = evaluation(s,V,xgrid,ygrid,m,nr,nc,PSI,Pyy,y,n,p,Tm_true,theta,time_el);
```

D.2 Input File: geometry.m

```
% Input function, this function provides information regarding the
% geometry of the problem studied.
% AZ last version: March 2005

function [L2,Re,s_loc,nr,nc,xgrid,ygrid]=geometry(a);


%-------------------------------------------------------
%Dimensional data
L1D = 1000;     %length in x direction, meters
L2d = 750;         %length in y direction, meters

nr = 24; nc = 32;  %discretization of domain
```

```
m = nr*nc;          %m is the number of cells
Dxd = L1d/nc; Dyd=L2d/nr;

Red = 0.001*ones(nr,nc)*Dxd*Dyd; %recharge, m^3/day

h0d = 120; h1d = 110; % head in meters, west and east

t0 = 1; %unit time, day
%-------------------------------------------------------

% Make nondimensional;
L1 = 1; L2 = L2d/L1d; h0 = 1; h1 = 0;
Re = Red*t0/(h0d-h1d); Dx = Dxd/L1d; Dy = Dyd/L1d;

%-------------------------------------------------------

% Grid characteristics
xgrid = -Dx/2+Dx*(1:nc); ygrid = -Dy/2+Dy*(1:nr);
xgrid = xgrid'; ygrid = ygrid'; ygrid = flipud(ygrid);
[XG,YG] = meshgrid(xgrid,ygrid);

s_loc = [stackt(XG),stackt(YG)];

%-------------------------------------------------------
```

D.3 Input File: `in.m`

```
% input function, this function provides information regarding the
% input file, covariance model, convergence criterion and so on.
% AZ last version: March 2005

function [y_data,k_sample,initial,sig2,theta,thetamax,convergence,cov,lambda,...
              ...Tm_true,itermax,ro,N]=in(a)


y_data = load ('y_dataT_err4.txt');
load k_sample.txt
Tm_true = load ('Tm.txt');

sig2 = 1E-4; %R matrix
initial = 2.5; %initial estimate
% Starting parameters of the covariance model
theta(1,1) = 1;
%theta(2,1) = 0.1;
lambda = 1; %Marquardt
% convergence criterion of the parameter
thetamax(1,1) = 1;
```

```
% thetamax(2,1) = 0.1;
% Convergence criterion
convergence = 1;
% covariance model 1 linear 2 cubic 3 linear+cubic
cov = 2;
%-------------------------------------------------------------------
%CR
itermax = 60; % CR max number of iteration allowed
ro = 0.99; % MCMC
N = 1000; % number of conditional realizations
```

D.4 Parameter Estimation Routine: `PE.m`

```
%parameter estimation function
% AZ last version: March 2005

function [Q,thetanew] = PE(theta,H,X,R,ksi,SSR,sig2,convergence,hd,n,y,param,...
              ...lambda,nc,nr)

Lm = 1e100;

iteration = 0; linvflag2 = 0;
while linvflag2 == 0
iteration = iteration+1;

    if param == 1 %linear covariance
        gen = -hd; %linear
        gen = reshape(gen,nc,nr);
        gen = gen';
        Qdev1 = toeplsbtstb(gen);
        Q = theta(1,1)*Qdev1;
    elseif param == 2 %cubic covariance
        gen = hd.^3; %cubic
        gen = reshape(gen,nc,nr);
        gen = gen';
        Qdev1 = toeplsbtstb(gen);
        Q = theta(1,1)*Qdev1;
    elseif param == 3
        disp('Under construction...')
        return
    else
        disp('something is wrong.... please introduce the correct ...
          ...value for the covariance model')
        return
    end
    %R=theta(2,1)*eye(n);
```

```
    %perform now one iteration of linear inversing (xi form)
    PHI = H*X; QHT = Q*H'; HQHT = H*QHT;
    PSI = HQHT+R;

    PSIdev1 = H*Qdev1*H';
    %PSIdev2=eye(n);

    invPSI = inv(PSI);
    PP = invPSI-invPSI*H*X*inv(X'*H'*invPSI*H*X)*X'*H'*invPSI;
    g(1,1) = 0.5*trace(PP*PSIdev1)-0.5*y'*(PP*PSIdev1*PP)*y;
    %g(2,1)=0.5*trace(PP*PSIdev2)-0.5*y'*(PP*PSIdev2*PP)*y;

    F(1,1) = 0.5*trace(PP*PSIdev1*PP*PSIdev1);
    %F(1,2)=0.5*trace(PP*PSIdev1*PP*PSIdev2);
    %F(2,1)=0.5*trace(PP*PSIdev2*PP*PSIdev1);
    %F(2,2)=0.5*trace(PP*PSIdev2*PP*PSIdev2);

    Marq = F+lambda;%*eye(2);

    thetaold = theta;
    corr = -inv(Marq)*g;
    theta = thetaold+corr;

    Lm_old = Lm;
    Lm = 0.5*(ksi'*HQHT*ksi+SSR/sig2); % ksi form

    if abs(Lm-Lm_old)<convergence
         linvflag2 = 1;
    end

    %if (abs(corr)<convergence)
     %   linvflag2=1;
    %end

    if iteration>50
        disp('do not converge')
        linvflag2 = 1;
    end

end %linvflag2

if param == 1 %linear covariance
Q = theta(1,1)*Qdev1; %linear
elseif param == 2 %cubic covariance
Q = theta(1,1)*Qdev1; %cubic
end
```

```
thetanew = theta;
%R=theta(2,1)*eye(n);
```

D.5 Optimization Routine: `opti2.m`

```
% This function optimizies the delta value to minimize the objective function
% Quasi Linear case
% AZ last version: March 2005

function objft=opti2(delta,invR,G,sold,snew,y_data,Re,L2,k_sample,nr,nc)

% new estimate
s=sold*delta+snew*(1-delta);
sm = reshapet(s,nr,nc); Tm = exp(sm);

% Forward problem
[A, b, CR, CC] = mkls(Tm,Re,L2);
h = full(A\b);
%sample from predicted head
y_pred = h(k_sample); %y_pred is predicted observation

y = y_data - y_pred;

   % a=y'*inv(R)*y
   % b=s'*G*s
    %objft=a+b;
objft = y'*invR*y+s'*G*s;
```

D.6 Evaluation of the Results: `evaluation.m`

```
% Evaluation of the results
% AZ last version: December 2005

function [oo]=evaluation(s,V,xgrid,ygrid,m,nr,nc,PSI,Pyy,y,n,p,Tm_true,theta,...
...time_el)
% Evaluation of the results
% s estimate
% V variance
%-----------------------------------------------------------------
% residual testing

iPSI = inv(PSI);
T = (orth(Pyy))'; %should produce an (n-p) by n matrix (verify that: T*PHI=0)
delta_r = T*y;    %delta residuals
varia = diag(T*PSI*T');
se = sqrt(varia); %variance and standard error of delta
```

```
epsilon_r = delta_r./se;  %epsilon residuals

Q1 = mean(epsilon_r)
Q22 = mean(epsilon_r.^2)
%T*PHI

Q1crit=2/sqrt(n-p)
Q2crit1=1+2.8/sqrt(n-p)
Q2crit2=1-2.8/sqrt(n-p)

%-----------------------------------------------------------------

sm=reshapet(s,nr,nc); %estimate
Tm=exp(sm);TT=exp(s);

conint = sqrt(diag(V))*2;
conintm = reshapet(conint,nr,nc);

T_true=stackt(Tm_true);
sm_true=log(Tm_true);
s_true=stackt(sm_true);

invV=inv(V);
DDD=(1/m)*(s_true-s)'*invV*(s_true-s)
%DDD2=(1/m)*(s_true-s)'*pinv(V)*(s_true-s)

mean_s=mean(s)
var_s=var(s)

diff_s=s_true-s;
diff_sm=reshapet(diff_s,nr,nc);

mean_diff_s=mean(diff_s)
var_diff_s=var(diff_s)

diff_s2=(diff_s).^2;

mean_diff_s2=mean(diff_s2)
var_diff_s2=var(diff_s2)

D2=((diff_s).^2)./diag(V);

meanD2=mean(D2)
varD2=var(D2)

mean_T=mean(Tm(:))
var_T=var(Tm(:))
```

```
diff_T=T_true-TT;
meandiffT=mean(diff_T)
vardiffT=var(diff_T)


output(1,1) = theta;
output(2,1) = mean_s;
output(3,1) = var_s;
output(4,1) = mean_diff_s;
output(5,1) = var_diff_s;
output(6,1) = mean_diff_s2;
output(7,1) = var_diff_s2;
output(8,1) = meanD2;
output(9,1) = varD2;
output(10,1) = mean_T;
output(11,1) = var_T;
output(12,1) = meandiffT;
output(13,1) = vardiffT;
output(14,1) = Q1;
output(15,1) = Q22;
output(16,1) = DDD;
output(17,1) = time_el;

xlswrite('out_QL_s_err4' ,output)

figure(35)
mesh(xgrid, ygrid, sm), xlabel('x'); ylabel('y'); zlabel('Log T');
figure(36)
title('Log Transmissivity')
mesh(xgrid, ygrid, diff_sm), xlabel('x'); ylabel('y'); zlabel('diff Log T');
title('Difference Log Transmissivity')
figure(6)
mesh(xgrid, ygrid, log(Tm_true)), xlabel('x'); ylabel('y');
zlabel(' log True_Transmissivity');
title('Log True Transmissivity')
%figure(8)
%mesh(xgrid, ygrid, exp(sm+conintm)), xlabel('x'); ylabel('y');
zlabel('Transmissivity');
%title('Upper 95% Confidence Interval')
%figure(9)
%mesh(xgrid, ygrid, exp(sm-conintm)), xlabel('x'); ylabel('y');
zlabel('Transmissivity');
%title('Lower 95% Confidence Interval')

oo=1;
```

Appendix E

Aquifer Parameter Estimation: Conditional Realizations

This appendix includes the source codes (in Matlab, .m files) used for the transmissivity estimation from head measurements using the quasi-linear geostatistical approach with the addition of the Conditional Realizations, see *Kitanidis* [1995], *Michalak and Kitanidis* [2003] and sections 3.2, 6.3 for the details of the procedure. In the following are reported only the most important functions.

E.1 Main Program: `CR.m`

```
% program that applies the Quasi Linear Geostatistical approach with the adding of
% the Conditional realizations (that it is explained in Kitanidis, 1995).
% AZ last version: December 2005

% Main program Conditional realizations
clear, clc
tic
% Load Geometry
a = 1;
[L2,Re,s_loc,nr,nc,xgrid,ygrid] = geometry(a);

% Load input data
[y_data,k_sample,initial,sig2,theta,thetamax,convergence,cov,lambda,Tm_true,...
...itermax,ro,N] = in(a);
n = size(y_data,1);
m = nr*nc;

%INVERSE PROBLEM
%-------------------------------------------------------------
%PARAMETERS  - assumed known
%measurement error matrix
sig = sqrt(sig2); R = eye(n)*sig2; invR=inv(R);
```

```
%Covariance matrix & structure of unknown function:
hd = zeros(1,n);
for ii=1:m
      hd(1,ii) = sqrt((s_loc(1,1)-s_loc(ii,1))^2+(s_loc(1,2)-s_loc(ii,2))^2);
end

if cov == 1
    p = 1;
    X = ones(m,p);
gen = -theta(1,1)*hd;     %linear covariance
elseif cov == 2
    p = 3;
    X = ones(m,p);
    X(:,2:3) = s_loc(:,1:2);
gen = theta(1,1)*(hd.^3); %cubic covariance
end
clear s_loc;

gen = reshape(gen,nc,nr);
gen = gen';
Q = toeplsbtstb(gen);

invQ = inv(Q);
G = invQ-invQ*X*inv(X'*invQ*X)*X'*invQ;
%---------------------------------------------------------------
%Begin inverse problem

%Initial s estimate
sm=initial*ones(nr,nc);

linvflag = 0;
%linvflag = 0 indicates quasi-linear inversing has not converged
iter = 0; SSR = 1E100; Lm = 1E100; %initialize

%---------------------------------------------------------------------
% Conditional realizations
s_c = zeros(m,N); s_cm = zeros(nr,nc,N); L = 1.E100*ones(1,N);

alfa = sqrt(1-ro^2);

% Unconditional realizations
s_u2 = simvm2(N,Q,X);


%initialization
AR = zeros(N,1);
```

```
flag = ones(N,1);
s_u = zeros(m,N);
s_c = zeros(m,N);
s_uA = zeros(m,N);
s_cA = zeros(m,N);
Va = zeros(m,m);
Va_diag = zeros(m,N);
pl = -1000;

for isim = 1:N

   disp(['ISIM = ',num2str(isim)])

   %unconditional realization


   if isim==1
       s_u(:,1) = s_u2(:,1);
       s_uA(:,1) = s_u2(:,1);
       v = randn(n,1)*sig;
   else

   s_u(:,isim) = ro*s_uA(:,isim-1)+alfa*s_u2(:,isim); %MCMC
   v = randn(n,1)*sig;

   %mesu = mean(s_u);
   %s_u(:,isim) = s_u-mesu;
   end


   %Initial cond simulation estimate
   s_cm(:,:,isim) = sm;
   Tc = exp(s_cm(:,:,isim)); s_c(:,isim) = stackt(s_cm(:,:,isim));

   linvflag = 0; %linvflag = 0 means not converged
   iter = 0; SSR = 1E100; %initialize
   %---------------------------------------------------------------
   while linvflag == 0             %linvflag loop->
      iter = iter+1;
      disp (['ITERATION = ',num2str(iter)])
      %solve forward problem for estimate
      SSR_old = SSR;   L_old = L(isim); %s_cm_old = s_cm;
      [A, b, CR, CC] = mkls(Tc,Re,L2);
      h = full(A\b); head = reshapet(h,nr,nc);

      y_pred = h(k_sample);          %y_pred is predicted observation
      res = (y_data+v-y_pred); SSR = res'*res;
```

```
disp (['SSR = ',num2str(SSR)])
if iter>1
   L(isim) = 0.5*(ksi'*HQHT*ksi+SSR/sig2);
   disp(['L is ', num2str(L(isim))]);
   if L(isim)>L_old
      disp('L increased!')
   end
end
dh = adjstatm2(k_sample,Tc,L2,A,h,CR,CC);
for i = 1:n
   dh(:,:,i) = dh(:,:,i).*Tc;
end
H = zeros(n,m);
for i = 1:n
   H(i,:) = stackt(dh(:,:,i))';
end
y = y_data - y_pred + H*s_c(:,isim);
%perform now one iteration of linear inversing (xi form)
PHI = H*X; QHT = Q*H'; HQHT = H*QHT;
PSI = HQHT+R;
sol = [PSI,PHI;PHI',zeros(p,p)]\[y+v-H*s_u(:,isim);zeros(p,1)];
ksi = sol(1:n,1); beta = sol(n+1:n+p,1);

sold = s_c(:,isim);
snew = s_u(:,isim)+ X*beta + QHT*ksi;

%optimization along the line
objft = @(delta)opti3(delta,invR,Re,L2,G,sold,snew,y_data,k_sample...
        ...,v,s_u(:,isim),nr,nc);
[delta,val,fl,ou] = fminbnd(objft,-0.5,1.5);
delta

s_c(:,isim) = delta*sold+(1-delta)*snew;
s_cm(:,:,isim) = reshapet(s_c(:,isim),nr,nc);
Tc = exp(s_cm(:,:,isim));
if iter>1 & abs(L(isim)-L_old)<convergence
   linvflag = 1;
   figure(8)
   mesh(xgrid, ygrid, s_cm(:,:,isim)), xlabel('x'); ylabel('y');
   title('Estimated Log Transmissivity' ),zlabel(['cond real ltrans,...
   ... isim = ', num2str(isim)]);
   pause(0.5)
end

if iter == itermax
   flag(isim,1) = 0; % do not converge
   disp('does not converge')
```

```
         linvflag = 1;
         flag(isim,1) = 0;
      end

      if s_c(:,isim)>100|s_c(:,isim)<0
         flag(isim,1) = 0; % do not converge
         disp('Tc ->00 or Tc->0')
         linvflag=1;
      end

   end              %linvflag loop<-

if isim == 1 %do not perform the acceptance rejection on the first realization
s_cA(:,isim) = s_c(:,isim);
s_uA(:,isim) = s_u(:,isim);
AR(isim,1) = 1;
        % Evaluation of the posterior probability p = exp (pl)
        pl = -0.5*(s_c(:,isim)-X*beta)'*invQ*(s_c(:,isim)-X*beta)-...
        ...0.5*(y_data-H*s_c(:,isim))'*invR*(y_data-H*s_c(:,isim));
else
if flag(isim,1) == 0 %did not converge
s_cA(:,isim) = s_cA(:,isim-1); %last conditional realization accepted
s_uA(:,isim) = s_uA(:,isim-1); %last unconditional realization accepted
AR(isim,1) = 0;
else %MH acceptance/rejection algorithm
   [AR(isim,1),s_cA(:,isim),s_uA(:,isim),pl_new] = MH2(s_u(:,isim),s_uA(:,isim-1),...
   ...s_c(:,isim),s_cA(:,isim-1),invQ,ro,X,beta,y,H,invR,pl);
            pl = pl_new; % pl log of the posterior probability of the
                         % last accepted realization
        end
end

if AR(isim,1) == 1
INVMAT = inv([PSI,PHI;PHI',zeros(p,p)]);
Pyy = INVMAT(1:n,1:n);
Pyb = INVMAT(1:n,n+1:n+p);
Pbb = INVMAT(n+1:n+p,n+1:n+p);
Va = Q - QHT*Pyy*QHT' - X*Pbb*X' - X*Pyb'*QHT'-QHT*Pyb*X';
Va_diag(:,isim) = diag(Va);
else
Va_diag(:,isim) = Va_diag(:,isim-1);
    end

end % loop N Realizations
toc
time_cpu=toc
oo = evaluation_CR1(s_cA,flag,Va_diag,AR,xgrid,ygrid,m,nr,nc,n,Tm_true,...
```

```
     ...N,time_cpu);

save('AR_CR_s_err4.txt','AR','-ascii')
save('flag_CR_s_err4.txt','flag','-ascii')
save('Va_diag_CR_s_err4.txt','Va_diag','-ASCII')
save('s_c_CR_s_err4.txt','s_c','-ascii')
save('s_cA_CR_s_err4.txt','s_cA','-ascii')
save('s_u_CR_s_err4.txt','s_u','-ascii')
save('s_uA_CR_s_err4.txt','s_uA','-ascii')

%hot start
save('v_last_CR_s_err4.txt','v','-ascii')
```

E.2 Optimization Routine: `opti3.m`

```
% This function optimizies the delta value to minimize the objective function
% Quasi Linear + Conditional realizations case
% AZ last version: March 2005

function objft=opti3(delta,invR,Re,L2,G,sold,snew,y_data,k_sample,v,su,nr,nc)

n = size(y_data,1);
m = nr*nc;

% new estimate
s = sold*delta+snew*(1-delta);
sm = reshapet(s,nr,nc); Tm = exp(sm);

% Forward problem
[A, b, CR, CC] = mkls(Tm,Re,L2);
h = full(A\b); head = reshapet(h,nr,nc);

%sample from predicted head
y_pred = h(k_sample);  %y_pred is predicted observation

y=y_data+v-y_pred;

objft=y'*invR*y+(s-su)'*G*(s-su);
```

E.3 Acceptance/Rejection Algorithm: `MH2.m`

```
% This function perform the Metropolis-Hastings acceptance/rejection algorithm
% AZ last version: March 2005

function [AR,s_cAccepted,s_uAccepted,pl_new] = ...
                    ...MH2(s_u,s_uA,s_c,s_cA,invQ,ro,X,beta,y_data,H,invR,pl)
```

```
% s_u unconditional to test
% s_c conditional to test
% s_uA last uncoditional accepted
% s_cA last conditional accepted
% AR flag that indicates if a realizations is accepted (AR=1) or rejected (AR=0)
% pl  log of the posterior probability of the last accepted

alfa = sqrt(1-ro^2);

% computation of the acceptance probability
% transition probability
% c candidate, l last accepted
qcl = -0.5*(s_u-ro*s_uA)'*invQ*(s_u-ro*s_uA)/(alfa^2);
qlc = -0.5*(s_uA-ro*s_u)'*invQ*(s_uA-ro*s_u)/(alfa^2);

% posterior probability function
pc = -0.5*(s_c-X*beta)'*invQ*(s_c-X*beta)-0.5*(y_data-H*s_c)'*invR*(y_data-H*s_c);

praccept = exp(pc+qlc-pl-qcl);
praccept = min(praccept,1);

Unif=rand;
if Unif<praccept          % realization is accepted
s_cAccepted = s_c;    % conditional realization accepted
s_uAccepted = s_u;    % unconditional realization accepted
AR = 1;
    pl_new = pc;
else                      % realization is rejected
s_cAccepted = s_cA;   % last conditional realization accepted
s_uAccepted = s_uA;   % last unconditional realization accepted
AR = 0;
    pl_new = pl;
end
```

E.4 Evaluation of the Results: `evaluation_CR.m`

```
% Evaluation of the results
% AZ last version: December 2005

function [oo]=evaluation_CR(s_cA,flag,Va_diag,AR,xgrid,ygrid,m,nr,nc,n,Tm_true,...
...N,time_cpu)
% Evaluation of the results
% Va_diag=diag(V)

output=zeros(18,6);
```

```
simul(1)=1;
simul(2)=100;
simul(3)=200;
simul(4)=300;
simul(5)=500;
simul(6)=700;


T_true=stackt(Tm_true);
sm_true=log(Tm_true);
s_true=log(T_true);


for cont=1:6

AREAL=0;NNN=0;
Va_mean=zeros(m,1); %diag V
s_mean=zeros(m,1);


ST = simul(cont)

EN= N;
for i=ST:EN
    s_mean=s_mean+s_cA(:,i)*AR(i,1); %*flag(i,1);
    Va_mean=Va_mean+Va_diag(:,i)*AR(i,1); %*flag(i,1);
    AREAL=AREAL+AR(i,1);
end
NNN=(ST-EN)-((ST-EN)-sum(flag(ST:EN,1))); %number of converged realizations
AREAL; %number of accepted realizations


percentAR = AREAL / NNN*100;

s_mean=s_mean/AREAL; % NNN;
sm_mean=reshapet(s_mean,nr,nc);
Va_mean=Va_mean/AREAL; % NNN;

conint_mean = sqrt(Va_mean)*2;
conintm_mean=reshapet(conint_mean,nr,nc);


mean_s=mean(s_mean);
var_s=var(s_mean);

diff_s=s_true-s_mean;
```

```
diff_sm=reshapet(diff_s,nr,nc);

diff_s2=diff_s.^2;


mean_diff_s=mean(diff_s);
var_diff_s=var(diff_s);


mean_diff_s2=mean(diff_s2);
var_diff_s2=var(diff_s2);


D2=(diff_s.^2)./Va_mean;

meanD2=mean(D2);
varD2=var(D2);

T_mean=exp(s_mean);
Tm_mean=reshapet(T_mean,nr,nc);
mean_T=mean(T_mean);
var_T=var(T_mean);

diff_T=T_true-T_mean;
meandiffT=mean(diff_T);
vardiffT=var(diff_T);


output(1,cont) = ST;
output(2,cont) = EN;
output(3,cont) = NNN;
output(4,cont) = AREAL;
output(5,cont) = percentAR;
output(6,cont) = mean_s;
output(7,cont) = var_s;
output(8,cont) = mean_diff_s;
output(9,cont) = var_diff_s;
output(10,cont) = mean_diff_s2;
output(11,cont) = var_diff_s2;
output(12,cont) = meanD2;
output(13,cont) = varD2;
output(14,cont) = mean_T;
output(15,cont) = var_T;
output(16,cont) = meandiffT;
output(17,cont) = vardiffT;
output(18,cont) = time_cpu;
```

```
end

xlswrite('out_CR_s_err4',output)


figure(35)
mesh(xgrid, ygrid, sm_mean), xlabel('x'); ylabel('y'); zlabel('mean cond real');
figure(36)
mesh(xgrid, ygrid, diff_sm), xlabel('x'); ylabel('y'); zlabel('diff Log');
figure(6)
mesh(xgrid, ygrid, Tm_mean), xlabel('x'); ylabel('y'); zlabel('Transmissivity');

figure(25)
mesh(xgrid, ygrid, exp(sm_mean+conintm_mean)), xlabel('x'); ylabel('y');
zlabel('Transmissivity');
title('Upper 95% Confidence Interval')
figure(27)
mesh(xgrid, ygrid, exp(sm_mean-conintm_mean)), xlabel('x'); ylabel('y');
zlabel('Transmissivity');
title('Lower 95% Confidence Interval')

oo=1;
```

Appendix F

Matrix Multiplication

This appendix includes the source codes (in Matlab, .m files) used for the spectral matrix multiplications, see *Nowak et al.* [2003] and section 7 for the details of the procedure. In the following are reported only the most important functions.

F.1 Main Function, Evaluation of $\mathbf{Q}_{tt} = \mathbf{HQH}^T$ and $\mathbf{Q}_{st} = \mathbf{QH}^T$: `multiplication.m`

```
function [Qtt,Qst]=multiplication(gen,H);
% This function performs the matrix multiplication using spectral
% convolution
% gen: (nr,nc) generator matrix of the Qss STT and of the SCC matrix
% H: (m,n) sensitivity matrix
% m number of observation
% n=nr*nc number of unknown
% nc= number of columns of the grid
% nr= number of rows of the grid
% Qtt=HQssHt matrix
% Qst=QssHt matrix
% AZ last version: December 2004

[nr,nc]=size(gen);
m=size(H,1);
n=nr*nc;
n1=size(H,2);

if n1~=n, disp('ERROR, arguments must have same number of columns'), return, end

QssSCC=circulantsb_a(gen); % this function stores only the first column of the ...
                ...SCC matrix O(n)
% dimension of the QssSCC (nynx,1)
```

```
nx=2*nr-2; %numbers of SC blocks
ny=2*nc-2; %dimensions of the SC blocks
coefficient=1/(sqrt(nx*ny));

% computation of eigenvalues eigQ=sqrt(nxny)FQss,1
% Qss,1 (nynx,1)
u=reshape(QssSCC,ny,nx);
v=fft2(u);
eigQ=reshape(v,ny*nx,1); % v=F*QssSCC,1
%LAM=diag(eigQ); %eigenvalues matrix
clear u; clear v; clear QssSCC;

% computation of Qtt=HQssHt
%Qtt is symmetric, only the upper triangle and diagonal have to be computed
Qtt=zeros(m,m);

% computation of Qst=Qss*Ht
Qst=zeros(nr*nc,m);

% upper / lower triangle
for k=1:m %row1 index
    H1=reshape(H(k,:),nr,nc);  %(nr,nc)
    % FFT of the embedded matrix of H
    v1=coefficient*fft2(H1,nx,ny);
    v1=reshape(v1,ny*nx,1); clear H1;
    v1star=real(v1)-i*imag(v1); %complex conjugate of v1

    % computation of Qst=Qss*Ht
    w=eigQ.*v1; %(nynx,1)
    w_re=reshape(w,nx,ny);
    Qst1=sqrt(nx*ny)*ifft2(w_re); %(nx,ny)
    %estraction from the embedded matrix
    Qst2=Qst1(1:nr,1:nc);
    Qst(:,k)=real(reshape(Qst2,nr*nc,1));
    clear Qst1; clear Qst2; clear w; clear w_re;

    % computation of Qtt
    for l=k:m %row2 index
        if k==l %diagonal
            p=eigQ.*(abs(v1).^2); %(nynx,1)
            Qtt(k,k)=sum(p);
            clear p;
        else
            H2=reshape(H(l,:),nr,nc);
            % FFT of the embedded matrix of H
            v2=coefficient*fft2(H2,nx,ny);
            v2=reshape(v2,ny*nx,1);
```

```
            p=v1star.*eigQ.*v2; %(nynx,1)
            Qtt(k,l)=real(sum(p));
            Qtt(l,k)=Qtt(k,l);
            clear H2; clear v2; clear p;
        end
    end
    clear v1; clear v1star;
end
```

F.2 Main Function, Evaluation of $\mathbf{Q}_{st} = \mathbf{Q}\mathbf{H}^T$: `mulQst.m`

```
function Qst=mulQst(gen,H);

% gen: (nr,nc) generator matrix of the Qss STT and of the SCC matrix
% H: (m,n) sensitivity matrix
% m number of observation
% n=nr*nc number of unknown
% nc= number of columns of the grid
% nr= number of rows of the grid
% AZ last version: December 2004

[nr,nc]=size(gen);
m=size(H,1);
n=nr*nc;
n1=size(H,2);

if n1~=n, disp('ERROR, arguments must have same number of columns'), return, end

QssSCC=circulantsb_a(gen); % this function stores only the first column of ...
          ...the SCC matrix O(n)
% dimension of the QssSCC (nynx,1)

nx=2*nr-2; %numbers of SC blocks
ny=2*nc-2; %dimensions of the SC blocks

% computation of eigenvalues eigQ=sqrt(nxny)FQss,1
% Qss,1 (nynx,1)
u=reshape(QssSCC,ny,nx);
v=fft2(u);
eigQ=reshape(v,ny*nx,1); % v=F*QssSCC,1
%LAM=diag(eigQ); %eigenvalues matrix
clear u; clear v; clear QssSCC;

% computation of Qst=Qss*Ht
Qst=zeros(nr*nc,m);

% upper / lower triangle
```

```
for k=1:m %row1 index
    H1=reshape(H(k,:),nr,nc);  %(nr,nc)
    % FFT of the embedded matrix of H
    v1=1/(sqrt(nx*ny))*fft2(H1,nx,ny);
    v1=reshape(v1,ny*nx,1); clear H1;

    % computation of Qst=Qss*Ht
    w=eigQ.*v1; %(nynx,1)
    w_re=reshape(w,nx,ny);
    Qst1=sqrt(nx*ny)*ifft2(w_re); %(nx,ny)
    %estraction from the embedded matrix
    Qst2=Qst1(1:nr,1:nc);
    Qst(:,k)=reshape(Qst2,nr*nc,1);
    clear Qst1; clear Qst2; clear w; clear w_re;
    clear v1;
end
```

F.3 Symmetric Circulant Matrix: `circulant_a.m`

```
function SC = circulant_a(G1);
% Symmetric Circulant Matrix
% G1 input row vector 1xn
% C output matrix 2n-2x2n-2
% Warning !!! In Nowak's paper the dimensions of the circulant matrix is
% defined as 2nx2n because the generator vector G1= [C0 C1 .... Cn]
% starts from 0 and not from 1, the input vector should be 1xm, where m=n+1.
% nx=nx1-1
% AZ last version: December 2004

n =size (G1,2);
G2a= fliplr(G1(2:n-1));
G2b=[G1 G2a]; %dimension of G2b 1x2n-2, (n+n-2)
G2c=fliplr(G2b(2:end)); %dimension of G2c 1x2n-3, (2n-2-1)
G2=[G2c G2b]; %dimension of G2 1x4n-5, (2n-3 + 2n-2)

SC=zeros(2*n-2,2*n-2);

%embedding of the circulant matrix
for i=1:2*n-2
  SC(i,:) = G2(2*n-1-i:4*n-4-i);%G2(2*n-3+2-i:2*n-3+2*n-2+1-i)
end
```

F.4 Symmetric Circulant Matrix with Circulant Blocks: `circulantsb_a.m`

```
function SCC = circulantsb_a(Gb11);
```

```
% Symmetric Circulant Matrix with circulant blocks
% This program evaluate only the first row of the SCC matrix
% Gb11 input matrix p,m
% SCC output vector SCC(1,(2m-2)*(2p-2))

% Gb11 : generator of circulant blocks for first generator
% m    : size of square blocks
% p    : number of blocks

% SCC    : output first row of SCC matrix

% AZ last version: December 2004

m     = size(Gb11,2);
p     = size(Gb11,1);
n     = m*p;

%SCC((2*m-2)*(2*p-2),(2*m-2)*(2*p-2))

% In the FFT multiplication only the first row/column is useful, so only
% this row/column it is stored. In this way the memory required is O(n)

% generator of the first row

G1=zeros(1,(2*m-2)*(2*p-2));
for i=1:p
    C=circulant_a(Gb11(i,:));
    G1(1,(i-1)*(2*m-2)+1:i*(2*m-2))=C(1,:);
end

for i=p+1:2*p-2
    C=circulant_a(Gb11(2*p-2-i+2,:));
    G1(1,(i-1)*(2*m-2)+1:i*(2*m-2))=C(1,:);
end

SCC=G1';% storage of the only first column
```

F.5 Symmetric Block Circulant Matrix with Circulant Blocks: `circulantsbt_a.m`

```
function SCC = circulantsbt_a(Gb11);
% Symmetric Circulant Matrix with circulant blocks
% This program embed the STT matrix in a larger SCC matrix using the
% generator of the STT matrix
% Gb11 input matrix p,m
% SCC output vector SCC(1,(2m-2)*(2p-2))
```

```
% Gb11 : generator of circulant blocks for first generator
% m    : size of square blocks
% p    : number of blocks

% AZ last version: December 2004

m     = size(Gb11,2);
p     = size(Gb11,1);

SCC=zeros((2*m-2)*(2*p-2),(2*m-2)*(2*p-2));

%building the SC blocks
G1=zeros((2*m-2),(2*m-2)*p);
G2=zeros((2*m-2),(2*m-2)*p);
for i=1:p
    C=circulant_a(Gb11(i,:));
    G1(:,(i-1)*(2*m-2)+1:i*(2*m-2))=C(:,:); %(2m-2),(2m-2)p
    G2(:,p*(2*m-2)-(2*m-2)*i+1:p*(2*m-2)-(2*m-2)*(i-1))=C(:,:); %(2m-2),(2m-2)p
end

G2b=[G1 G2(:,(2*m-2)+1:p*(2*m-2)-(2*m-2))]; %(2m-2),(2p-2)(2m-2)
G3=[G2(:,1:p*(2*m-2)-(2*m-2)) G2b];   %(2m-2),(2m-2)(3p-3)
G4=[G1(:,2*m-2+1:(p-1)*(2*m-2)) G3];   %(2m-2,4p-5)

%embedding of the circulant matrix
for i=1:2*p-2
    SCC((i-1)*(2*m-2)+1:i*(2*m-2),:)=G4(:,(2*m-2)*(2*p-3)+1+(1-i)*...
                ...(2*m-2):(2*m-2)*(4*p-5)+(1-i)*(2*m-2));
end
```

Bibliography

[1] Alapati, S., and Z. J. Kabala (2000), Recovering the release history of a groundwater contaminant via the nonlinear least-squares estimation, *Hydrological process*, *14*(6), 1003-1016.

[2] Amerio, L. (1974), *Analisi Matematica*, Tamburini Ed., (in Italian).

[3] Aral, M., and J. Guan (1996), Genetic algorithms in search of groundwater pollution sources, *Advances in Groundwater Pollution Control and Remediation, NATO ASI Series 2 Environment, 9*, 347-369, Kluwer Academic Publishers, The Netherlands.

[4] Atmadja, J., and A. C. Bagtzoglou (2000), Groundwater pollution source identification using the backward beam equation method, in: *Computational methods for Subsurface Flow and Transport*, 397-404, A.A. Balkema, Rotterdam, The Netherlands.

[5] Atmadja, J., and A. C. Bagtzoglou (2001), Pollution source identification in heterogeneous porous media, *Water Resour. Res., 37*(8), 2113-2125.

[6] Atmadja, J., and A. C. Bagtzoglou (2001), State of the art report on mathematical methods for groundwater pollution source identification, *Environ. Forensics, 2*(3), 205-214.

[7] Bagtzoglou, A. C., and J. Atmadja (2003), Marching-Jury backward beam equation and quasi-reversibility methods for hydrologic inversion: Application to contaminant plume spatial distribution recovery, *Water Resources Res., 39*(2), 1038.

[8] Ball, W. P., C. Liu, G. Xia, and D. F. Young (1997), A diffusion-based interpretation of tetrachloroethene and trichloroethene concentration profiles in a grounwater aquitard, *Water Resources Research, 33*, 2741-2757.

[9] Bates, B. C., and E. P. Campbell (2001), A Markov Chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling, *Water Resources Research, 37*(4), 937-947.

[10] Bear J. (1972), *Dynamics of fluids in porous media*, p. 762, American Elsevier Publishing Company Inc., New York.

[11] Bear, J., and Y. Bachmat (1990), *Introduction to modeling of transport phenomena in porous media*, p. 553, Kluwer Academic Publishers, The Netherlands.

[12] Bear, J., and A. Verruijt (1987), *modeling groundwater flow pollution*, D. Reidel Publishing Company.

[13] Birchwood, R. A. (1999), Identifying the location and the release characteristics of groundwater pollution source using spectral analysis, paper presented at the 19th Annual Hydrology Days Conference, AGU, Colorado State University, Fort Collins Colorado.

[14] Boano, F., R. Revelli, and L. Ridolfi (2005), Source identification in river pollution problems: A geostatistical approach, *Water Resources Research, 41*, W07023, doi:10.1029/2004WR003754.

[15] Butera, I., E. Grella, U. Maione, and M. G. Tanda (1998), Applicazione comparativa di tre metodi di risoluzione del problema inverso in un acquifero eterogeno bidimensionale, proceedings of "XXVI Convegno di Idraulica e Costruzioni Idrauliche", Catania September 9-12, 1998, (in Italian).

[16] Butera, I., G. Monti, C. Ruggeri and M. G. Tanda (2001), Restoring the release history of a pollutant in a 2D aquifer, Proceedings of 3^{rd} FGR01, June 25-27, 2001, Lisbon, Portugal.

[17] Butera, I., and M. G.Tanda (2001), L'approccio geostatistico per la ricostruzione della storia di rilascio di inquinanti in falda, *L'Acqua, 4*, (in Italian).

[18] Butera, I., and M. G. Tanda (2002), Ricostruzione della storia temporale del rilascio di un inquinante in falda mediante l'approccio geostatistico, Proceedings of "XXVIII Convegno di Idraulica e Costruzioni Idrauliche", Potenza (I) September 16-19, 2002, (in Italian).

[19] Butera, I., and M.G. Tanda (2003), A geostatistical approach to recover the release history of groundwater pollutants, *Water Resources Research, 39*(12), 1372, doi:10.1029/2003WR002314.

[20] Butera, I., and M. G. Tanda (2004), *L*a ricostruzione della storia del rilascio di inquinanti negli acquiferi. Impiego di rilevazioni a tempi diversi e impatto dell'eterogeneità dell'acquifero, *L'Acqua, 3*, 27-36, (in Italian).

[21] Butera, I., M. G. Tanda, and A. Zanini (2004), La ricostruzione della storia del rilascio di inquinanti in acquiferi sede di moto non uniforme mediante approccio geostatistico, Proceedings of "XXIX Convegno di Idraulica e Costruzioni Idrauliche", Trento (I) September 7-10 2004, (in Italian).

[22] Butera, I., M. G. Tanda, and A. Zanini (2005), The recovering of the pollutant release history in groundwater: use of the numerical modeling for the identification of the transfer function in non-uniform flow field, submitted at *Water Resources Research.*

[23] Butera, I., M. G. Tanda, and A. Zanini (2005), Using numerical modelling to identify the transfer function and application to the geostatistical procedure in the solution of inverse problems in groundwater, submitted at *Journal of Ill Posed Problems.*

[24] Capilla J. E., J. J. Gòmez-Hernández, and A. Sahuquillo (1997), Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data 2. Demonstration on a synthetic aquifer*, Journal of Hydrology, 203*, 175-188.

[25] Capilla J. E., J. J. Gòmez-Hernández, and A. Sahuquillo (1998), Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data 3. Application to the Culebra formation at the Waste Isolation Pilot Plant (WIPP), New Mexico, USA*, Journal of Hydrology, 207*, 254-269.

[26] Chen, J., S. Hubbard, Y. Rubin, C. Murray, E. Roden, and E. Majer (2004), Geochemical characterization using geophysical data and Markov Chain Monte Carlo methods: A case study at the South Oyster bacterial transport site in Virginia, *Water Resour. Res., 40*, W12412, doi:10.1029/2003WR002883.

[27] Chib, S., and E. Greenberg (1995), Understanding the Metropolis-Hastings Algorithm, *The American Statistician, 49*(4), 327-335.

[28] Chow, V. T. (1964), *Handbook of applied hydrology: a compendium of water-resources technology*, McGraw-Hill Book Company, New York.

[29] Citrini, C. (1992), *Analisi matematica 2*, p. 516, Bollati Boringhieri, Torino.

[30] Dagan, G. (1985), Stochastic modeling of groundwater-flow by unconditional and conditional probabilities: The inverse problem, *Water Resource Research, 21*(1), 65-72.

[31] Dagan, G. (1989), *Flow and Transport in Porous Formation*, Springer-Verlag.

[32] De Marsily, G. (1986), *Quantitative Hydrogeology*, Academic Press, Inc..

[33] De Marsily, G., J. P. Delhomme, A. M. Coudrain-Ribstein, and A. M. Lavenue (1999), 40 years of inverse problems in hydrogeology, *Comptes Rendus de l'Académie des Sciences - Series IIA - Earth and Planetary Science, Volume 329, Issue 2*, 30, 73-87.

[34] Domenico, P. A., and F. W. Schwartz (1998), *Physical and chemical hydrogeology*, John Wiley & Sons.

[35] Feyen, L., P. J. Ribeiro Jr, J. J. Gòmez-Hernández, K. J. Beven, and F. De Smedt (2003), Bayesian methodology for stochastic capture zone delineation incorporating transmissivity measurements and hydraulic head observations, *Journal of Hydrology, 271*, 156-170.

[36] Fienen, M. N., P. K. Kitanidis, D. Watson, and P. Jardine (2004), An application of Bayesian Inverse Methods to vertical deconvolution of hydraulic conductivity in a heterogeneous aquifer at Oak Ridge National Laboratory, *Mathematical Geology, 36*(1).

[37] Frigo, M. (1998), A Fast Fourier Transform Compiler, Proceedings of the "1999 ACM SIGPLAN" Conference on Programming Language Design and Implementation (PLDI), Atlanta, Georgia, May 1999. http://www.fftw.org.

[38] Giudici, M., G. Morossi, G. Parravicini, and G. Ponzini (1988), A new method for the identification of distributed transmissivities, *Water Resources Research, 31*(8).

[39] Gòmez-Hernández, J. J., A. Sahuquillo, J. E. Capilla (1997), Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data- 1 Theory, *Journal of Hydrology, 203*, 162-174.

[40] Gorelick, S. M., B. Evans, and I. Remson (1983), Identifying sources of groundwater pollution: An optimization approach, *Water Resources Research, 19*(3), 779-790.

[41] Grella, E. (1996), Metodi di risoluzione del problema inverso con riferimento a problemi di trasporto in acquiferi eterogenei, Phd Thesis, Politecnico di Milano, Milano, (in Italian).

[42] Gutjahr, A. L., and J. L. Wilson (1989), Co-kriging for stochastic models, *Transp. Porous Media, 4*(6), 585-598.

[43] Gutjahr, A., B. Bulard, S. Hatch, and L. Hughson (1994), Joint conditional simulations and the spectral approach for flow modeling, *Stoch. Hydrol. Hydraul., 8*(1), 79-108.

[44] Hadamard, J. (1902), Sur les problèmes aux dérivées partielles et leur signification physique. Princeton University Bulletin, pp. 49-52 (in French).

[45] Hanna, S., and T.-C. J. Yeh (1998), Estimation of co-conditional moments of transmissivity, hydraulic head, and velocity fields, *Advances in Water Resources*, *22*(1), 87-95.

[46] Harbaugh, A. W., E. R. Banta., M. C. Hill, and M. G. McDonald (2000), MODFLOW-2000 The U.S. Geological Survey modular Ground-Water model - User Guide to modularization concepts and the ground-water flow process, U.S. Geological Survey Open-File Report 00-92, p. 121.

[47] Harter, T., and T. C. J. Yeh (1996a), Stochastic analysis of solute transport in heterogeneous, variability saturated soils, *Water Resources Research, 32*(6), 1585-1595.

[48] Harter, T., T. C. J. Yeh (1996b), *Conditional stochastic analysis of solute transport in heterogeneous, variably saturated soils*, Water Resources Research, Vol 32, no 6 1597-1609.

[49] Harvey, C. F., and S. Gorelick (1995), Mapping hydraulic conductivity: Sequential conditioning with measurements of solute arrival time, hydraulic head, and local conductivity, *Water Resour. Res., 31*(7), 1615-1626.

[50] Hastings, W. K. (1970), Monte Carlo sampling methods using Markov chains and their applications, *Biometrika, 87*, 97-109.

[51] Hoeksema, R. J., and P. K. Kitanidis (1984), An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling, *Water Resources Research, 20*(7), 1003-1020.

[52] Hoeksema, R. J., and P. K. Kitanidis (1985), Comparison of Gaussian Conditional Mean and Kriging Estimation in the Geostatistical Approach to the Inverse Problem, *Water Resources Research, 21*(6), 825-836.

[53] Hoeksema, R. J., and P. K. Kitanidis (1989), Prediction of transmissivities, heads, and seepage velocities using mathematical modeling and geostatistics, *Adv. Water Resour., 12*, 90-101.

[54] Hughson, D. L., and T. C. Yeh (2000), An inverse model for three-dimensional flow in variability saturated porous media, *Water Resources Research, 36*(4), 829-840.

[55] IMSL: User's Manual, Stat/Library, Version 1.1, (1989), IMSL, Inc, Houston Texas, USA.

[56] Jury, W. A. (1982), Simulation of solute Transport using a transfer function model, *Water Resour. Res., 18*, 363-368.

[57] Jury, W. A., and K. Roth (1990), *Transfer functions and solute movement through soil: Theory and applications*, Birkhauser, Boston, Mass..

[58] Jury, W.A., G. Sposito, and R.E. White (1986), A transfer function model of solute transport through soil. 1.Fundamental Concepts, *Water Resour. Res., 22*, 243-247.

[59] Jury, W. A., L. H. Stolzy, and P. Shouse (1982), A field test of the transfer function model for predicting solute transport. *Water Resour. Res., 18*, 369-375.

[60] Kabala, Z. J., and T. H. Skaggs (1998), Comment on "Minimum relative entropy inversion: theory and application to recovering the release history of a groundwater contaminant" by A. D. Woodbury and T. J. Ulrych, *Water Resources Research, 34*, 2077-2079.

[61] Kitanidis, P. K. (1986), Parameter uncertainty in estimation of spatial functions: Bayesian analysis, *Water Resources Research, 22*(4), 499-507.

[62] Kitanidis, P. K. (1991), Orthonormal residuals in geostatistics: Model criticism and parameter estimation, *Math. Geol., 23*(5), 741-758.

[63] Kitanidis, P. K. (1993), Generalized covariance functions in estimation*, Math. Geol, 25*(5), 525-540.

[64] Kitanidis, P. K. (1995), Quasi-linear geostatistical theory for inversing*, Water Resources Research, 31*(10), 2411-2419.

[65] Kitanidis, P. K. (1996a), On the geostatistical approach to the inverse problem, *Advances in water resources, 19*(6), 333-342.

[66] Kitanidis, P. K. (1996b), Analytical expressions of conditional mean, covariance, and sample functions in geostatistics, *J. Stoch. Hydrol. Hydraul., 10*(4), 279-294.

[67] Kitanidis, P. K. (1997a), *Introduction to Geostatistics, Applications in Hydrogeology*, Cambridge University Press, New York.

[68] Kitanidis, P. K. (1997b), The minimum Structure solution to the inverse problem, *Water Resources Research, 33*(10), 2263-2272.

[69] Kitanidis, P. K. (1997c), Comment on "A reassessment of the groundwater inverse problem" by D. McLaughlin and L. R. Townley, *Water Resources Research, 33*(9), pp. 2199-2202.

[70] Kitanidis, P. K. (1998), How observations and structure affect the geostatistical solution to the steady state inverse problem, *Ground Water, 36*(5), 754-763.

[71] Kitanidis, P. K. (1999), Generalized covariance functions associated with the Laplace equation and their use in interpolation and inverse problems, *Water Resources Research, 35*(5), 1361-1367.

[72] Kitanidis, P. K. (2004), Applied Stochastic Inverse Methods*,* notes and practices for CEE 267, Interpolation and Inverse Methods.

[73] Kitanidis, P. K., and R. W. Lane (1985), Maximum Likelihood parameter estimation of hydrologic spatial process by the Gauss-Newton method, *Journal of Hydrology, 79* 53-71.

[74] Kitanidis, P. K., and K. F. Shen (1996), Geostatistical estimation of chemical concentration, *Advance Water Resources, 19*(6) 369-378.

[75] Kitanidis, P. K., and E. G. Vomvoris (1983), A Geostatistical Approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations, *Water Resources Research, 19*(3), 677-690.

[76] LaVenue, A. M., B. S. RamaRao, G. de Marsily, and M. G. Marietta (1995), Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields. 2. Application, *Water Resour. Res., 31*(3), 495-516.

[77] Li, B., and T.-C. J Yeh (1999), Cokriging estimation of the conductivity field under variably saturated flow conditions, *Water Resour. Res., 35*(12) 3663-3674.

[78] Li, W., W. Nowak, and O. A. Cirpka (2005), Geostatistical inverse modeling of transient pumping tests using temporal moments of drawdown, *Water Resour. Res., 41*, W08403, doi:10.1029/2004WR003874.

[79] Liu, C., and W. P. Ball (1998), Analytical modeling of diffusion-limited contamination and decontamination in a two layer porous medium, *Adv. Water Resours., 21*, 297-313.

[80] Liu, C., and W. P. Ball (1999), Application of inverse methods to contaminant source identification from aquitard diffusion profiles at Dover AFB, Delaware, *Water Resources Research, 35*(7), 1975-1985.

[81] Liu, C. C. K., C. Ji, and D. Neupane (2000), Linear systems approach to subsurface pollutant analysis, Proceedings of XIV Engineering Mechanics Conference, ASCE, Austin, 21-24 May 2000.

[82] Luo, J., O. A. Cirpka, M. N. Fienen, W. Wu, T. L. Mehlhorn, J. Carley, P. M. Jardine, P.M., C. S. Criddle, and P. K. Kitanidis (2006), A parametric transfer function methodology for analyzing reactive transport in nonuniform flow, *Journal of Contaminant Hydrology*, *83*(1-2), 27-41.

[83] Mac Donald, J. R. (1995), Solution of an "impossible" diffusion-inversion problem, *Computers in Physics, 9*(5), 546-553.

[84] McDonald, M. G., and A. W. Harbaugh (1988), A modular three-dimensional finite difference groundwater flow model, *U.S. Geol. Surv. Tech. Water Resour. Invest.*, book 6, chap. A1, 586 pp.

[85] Mahar, P. S., and B. Datta (1997), Optimal monitoring network and groundwater pollution source identification, *Journal of Water Resource Plng. and Mgmt., ASCE 123*(4), 199-207.

[86] Mahar, P. S., and B. Datta (2000), Identification of pollution sources in transient groundwater systems, *Water Resources Management, 14*, 209-227.

[87] Mahar, P. S., and B. Datta (2001), Optimal identification of groundwater pollution sources and parameter estimation, *Journal of Water Resource Plng. and Mgmt., ASCE 127,* 20-29.

[88] Marquardt, D.W. (1963), An algorithm for least squares estimation nonlinear parameters, *SIAM, J., II* 431-441.

[89] Marshall, L., D. Nott, and A. Sharma (2004), A comparative study of Markov Chain Monte Carlo methods for conceptual rainfall-runoff modeling, *Water Resources Research, 40*, W02501.

[90] Marsily, de G. (1986), *Quantitative Hydrogeology*, Academic Press, San Diego, CA.

[91] Matheron, G. (1963), Principles of geostatistics, *Econ. Geol., 58*, 1246-1266.

[92] Matheron, G. (1971), *The theory of regionalized variables and its applications*, 212 pp, Ecole de mines, Fointainebleau, France.

[93] Matheron, G. (1973), The intrinsic random functions and their applications, *Appl. Probab., 5*, 439-468.

[94] Learning MATLAB 7, (2004), The MathWorks Inc.

[95] Michalak, A. M. (2003), Application of Bayesian Inference Methods to Inverse Modeling for Contaminant Source Identification, Phd Thesis, Stanford University, Stanford CA.

[96] Michalak, A. M., and P. K. Kitanidis (2002). Application of Bayesian Inference Methods to Inverse Modeling for Contaminant Source Identification at Gloucester Landfill, Canada, in Computational Methods in Water Resources XIV, Volume 2, pp. 1259-1266, edited by S.M. Hassanizadeh, R.J. Schotting, W.G. Gray and G.F. Pinder, Elsevier, Amsterdam, The Netherlands.

[97] Michalak, A. M., and P. K. Kitanidis (2003), A method for enforcing parameter nonnegativity in Bayesian inverse problems with an application to contaminant source identification, *Water Resources Research, 39* (2), SBH 7-1, 7-14.

[98] Michalak, A. M., and P. K. Kitanidis (2004a), Application of geostatistical inverse modeling to contaminant source identification at Dover AFB, Delaware, *Journal of Hydraulic research, 42*, EXTRA ISSUE , 9-18.

[99] Michalak, A. M., and P. K. Kitanidis (2004b), Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling, *Water Resources Research, 40*.

[100] Michalak, A. M., and P. K. Kitanidis (2005), A Method for the Interpolation of Nonnegative Functions with an Application to Contaminant Load Estimation, *Stochastic Environmental Research and Risk Assessment, 19*, doi:10.1007/s00477-004-0189-1, 8-23.

[101] Monti, G. and C. Ruggeri (2000), Ricostruzione della storia di rilascio di un inquinante in falda mediante l'analisi dell'evoluzione spaziale del pennacchio, M.S., Politecnico di Milano, Milano A.A. 1999-2000, (in Italian).

[102] Morrison, R. D. (2000), Application of forensic techniques for age dating and source identification in environmental litigation, *Environ. Forensics, 1*(3), 131-153.

[103] Neuaper, R. M., B. Borchers, and J. Wilson (2000), Comparison of inverse methods for reconstructing the release history of a groundwater contamination source, *Water Resources Research 36*(9), 2469-2475.

[104] Neuaper, R. M., and J. Wilson (1999), Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant, *Water Resources Research, 35*(11), 3389-3398.

[105] Neuaper, R. M., and J. Wilson (2005), Backward probability model using multiple observations of contamination to identify groundwater contamination sources at the Massachussetts Military Reservation, *Water Resources Research, 41*, W02015.

[106] Nowak, W., and O. A. Cirpka (2004), A modified Levenberg-Marquardt algorithm for quasi-linear geostatistical inversing, *Advances in Water Resources, 27*, 737-750.

[107] Nowak, W., S. Tenklenve, and O. A. Cirpka (2003), Efficient Computation of Linearized Cross-Covariance Matrices of interdependent Quantities, *Mathematical Geology, 35*(1), 53-66.

[108] Obenchain, R. L. (1977), Classical F-tests and confidence regions for ridge regression, *Technometics, 19*(4), 429-439.

[109] Pandolfi, L. (2004), Class Notes from "Inverse and Ill-posed Problems".

[110] Philip, R. D., and P. K. Kitanidis (1989), Geostatistical estimation of hydraulic head gradients, *Ground Water, 27*(6).

[111] Press, W. H. et al (1989), *Numerical recipes*, Cambridge University Press.

[112] Provencher, S. W. (1982), A constrained regularization method for inverting data represented by linear algebraic or integral equations, *Comput. Phys. Commmun., 27*, 213-227.

[113] RamaRao, B. S., A. M. LaVenue, G. de Marsily, and M. G. Marietta (1995), Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields. 1. Theory and computational experiments, *Water Resour. Res., 31*(3), 475-494.

[114] Rinaldo, A., A. Marani, and A. Bellin (1997), On mass response functions, *Water Resources Research, 25*(7), 1603-1617.

[115] Robin, M. J. L., A. L. Gutjahr, E. A. Sudicky, and J. L. Wilson (1993), Cross-correlated random field generation with the direct Fourier transform method, *Water Resour. Res., 29*(7), 2385-2398.

[116] Rubin, Y. (1991a), Prediction of Tracer Plume Migration in Disordered Porous Media by the Method of Conditional Probabilities, *Water Resour. Res., 27*(6), 1291-1308.

[117] Rubin, Y. (1991b), Transport in heterogeneous porous media: Prediction and uncertainty, *Water Resour. Res., 27*(7), 1723-1738.

[118] Rubin, Y., M. A Cushey, and A. Wilson (1997), The moments of the breakthrough curves of instantaneously and kinetically sorbing solutes in heterogeneous geologic media: Prediction and parameter inference from field measurements, *Water Resources Research, 33*(11), 2465-2481.

[119] Rubin, Y., and G. Dagan (1987a), Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 1, Theory, *Water Resour. Res., 23*(7), 1185-1192.

[120] Rubin, Y., and G. Dagan (1987b), Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 2, Case-study, *Water Resour. Res., 23*(7), 1193-1200.

[121] Rubin, Y., and G. Dagan (1988), Stochastic analysis of boundaries effects on head spatial variability in heterogeneous aquifers, 1, Constant head boundary, *Water Resour. Res., 24*(10), 1689-1697.

[122] Rubin, Y., and G. Dagan (1989), Stochastic analysis of boundaries effects on head spatial variability in heterogeneous aquifers, 2, Impervious boundary, *Water Resour. Res., 25*(4), 707-712.

[123] Rubin, Y., and G. Dagan (1992), Conditional Estimation of Solute Travel Time in Heterogeneous Formations: Impact of Transmissivity Measurements, *Water Resour. Res., 28*(4), 1033-1040.

[124] Scheiddeger A. E. (1961), General theory of dispersion in porous media, *J. Gephys. Res., 66*, 3273-3278.

[125] Skaggs, T. H., and Z. J. Kabala (1994), Recovering the release history of a groundwater contaminant, *Water Resources Research, 30*, 71-79.

[126] Skaggs, T. H., and Z. J. Kabala (1995), Recovering the release history of a groundwater contaminant plume: method of quasi-reversibility, *Water Resources Research, 31*, 2669-2673.

[127] Skaggs, T. H., and Z. J. Kabala (1998), Limitations in recovering the history of a groundwater contaminant plume, *Journal of contaminant Hydrology, 33*, 347-359.

[128] Skaggs, T. H., Z. J. Kabala, and W. A. Jury (1998), Deconvolution of a nonparametric transfer function for solute transport in soils, *Journal of Hydrology, 207*(3-4), 170-178.

[129] Snodgrass, M. (2005), Advances in parameters estimation for subsurface process, Phd Thesis, Stanford University, Stanford CA.

[130] Snodgrass, M. F., and P. K. Kitanidis (1997), A geostatistical approach to contaminant source identification, *Water Resources Research, 33*, 537-546.

[131] Sposito, G., R.E. White, P.R. Darrah, and W.A. Jury (1986), A transfer function model of solute transport through soil. 3. The convection-Dispersion Equation, *Water Resour. Res. 22*, 255-262.

[132] Sudicky, E.A. (1986), A natural-gradient experiment on solute transport in a sand aquifer: spatial variability of hydraulic conductivity and its role on dispersion process, *Water Resources Research, 22*, 2069-1082.

[133] Sun, N. -Z. (1994), *Inverse Problems in Groundwater modeling*, Kluwer Academic Publishers, The Netherlands.

[134] Sun, N. -Z., and W. W-G. Yeh (1990a), Coupled inverse problem in groundwater modeling, 1, Sensitivity analysis and parameter identification, *Water Resour. Res. 26*(10), 2507-2525.

[135] Sun, N. -Z., and W. W-G. Yeh (1990b), Coupled inverse problem in groundwater modeling, 1, Sensitivity analysis and parameter identification, *Water Resour. Res. 26*(10), 2527-2540.

[136] Sun, N. -Z., and W. W-G. Yeh (1992), A stochastic inverse solution for transient groundwater flow: Parameter identification and reliability analysis, *Water Resour. Res. 28*(12), 3269-3280.

[137] Tamminen, J., and E. Kyrölä (2001), Bayesian solution for nonlinear and non-Gaussian inverse problems by Markov chain Monte Carlo method, *Journal of Geophysica l research, 106*(D13), 14,377-14,390.

[138] Tanda, M. G., I. Butera, and A. Zanini (2005), The recovering of the pollutant release history in aquifer with non uniform transport process, Proceedings of "Aquifer Vulnerability and Risk 04" 2^{nd} International workshop, Parma (I), September 21-23 2005.

[139] Tarantola, A. (2005), *Inverse Problems Theory and Methods for Model Parameter Estimation*, Siam, Philadelphia.

[140] Tikhonov, A. N. and V. Y. Arsenin (1977), *Solutions of Ill-Posed Problems*, Winston and Sons, New York.

[141] Trefethen, L. N. (2000), *Spectral Methods in Matlab*, Siam, Philadelphia.

[142] Vrugt, J.A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A Suhffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resources Research, 39*(8).

[143] Yeh, T.-C. J., A. L. Gutjahr, and M. Jin (1995), An iterative cokriging-like technique for groundwater flow modeling, *Ground Water, 33*, 33-41.

[144] Yeh, T.-C. J., and S. Y. Liu (2000), Hydraulic tomography: Development of a new aquifer test method, *Water Resour. Res., 36*(8), 2095-2105.

[145] Yeh, S. Liu, R. J. Glass, K. Baker, J. R. Brainard, D. Alumbaugh, and D. LaBrecque (2002), A geostatistically based inverse model for electrical resistivity surveys and its applications to vadose zone hydrology, *Water Resour. Res., 38*(12).

[146] Yeh, T.-C., Minghui J., and S. Hanna (1996), An iterative stochastic inverse method: Conditional effective transmissivity and hydraulic head fields, *Water Resour. Res., 32*(1), 85-92.

[147] Yeh, T.-C., and J. Zhang (1996), A geostatistical inverse method for variably saturated flow in the vadose zone, *Water Resour. Res., 32*(9), 2757-2766.

[148] Wagner, B. J. (1992), Simultaneous parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modeling, *Journal of Hydrology, 135*, 275-303.

[149] Wagner, B. J., and S.M. Gorelick (1989), Reliable aquifer remediation in the presence of spatially variable hydraulic conductivity: From data to design, *Water Resour. Res. 25*(10), 2211-2225.

[150] Walton, W. C. (1970), *Groundwater Resource evaluation*, Mc Graw Hill.

[151] Weese, J. (1992), A reliable and fast method for the solution of Fredholm integral equations of the first kind based on Tikhonov regularization, *Computer Physics Communications, 69*, 99-111.

[152] Weisstein, E. W., Hermitian Matrix, From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/HermitianMatrix.html.

[153] White, R.E., J.S. Dyson, R.A. Haigh, W.A. Jury, and G. Sposito (1986), A transfer function model of solute transport through soil. 2. Illustrative applications, *Water Resour. Res., 22,* 248-254.

[154] Wilson, E. M. (1990), *Engineering Hydrology*, IV Ed. McMillan Press Ltd.

[155] Woodbury, A. D., E. Sudicky, T. J. Ulrych, and R. Ludwig (1998), Three dimensional plume source reconstruction using minimum relative entropy inversion, *Journal of Contaminant Hydrology 32*, 131-158.

[156] Woodbury, A. D., and T. J. Ulrych (1996), Minimum relative entropy inversion: theory and application to recovering the release history of a groundwater contaminant, *Water Resources Research, 32*, 2671-2681.

[157] Woodbury, A. D., and T. J. Ulrych (1998), Reply to Comment on: Minimum relative entropy inversion: theory and application to recovering the release history of a groundwater contaminant, *Water Resources Research, 34*(8), 2081-2084.

[158] Woodbury, A. D., and T. J. Ulrych (2000), A full-Bayesian approach to the groundwater inverse problem for steady state flow, *Water Resources Research, 36*(8), 2081-2093.

[159] Zanini, A., I. Butera, M. G. Tanda (2005), Recovering the pollutant release history in aquifers with non uniform flow field, Proceedings of "AGU Fall meeting 2005", San Francisco, 5-9 December, 2005.

[160] Zhang, R. (2000), Generalized transfer function model for solute transport in heterogeneous soil, *Soil Sci. Soc. Am. J. 64*, 1595-1602.

[161] Zhang, J., and T.-C. J. Yeh (1997), An iterative geostatistical inverse method for steady flow in the vadose zone, *Water Resour. Res., 33*, 63-71.

[162] Zheng, C., and P.P. Wang (1999), MT3DMS: A Modular Three-Dimensional Multispecies Transport Models; Documentation and User's Guide, Contract Report SERDP-99-1, U.S. Army Engineer Research and Development Center, Vicksburg, MS.

[163] Zimmerman, D. A., G. DeMarsily, C. A. Gotway, M. G. Marietta, C. L. Axness, R. L. Beauheim, R. L. Bras, J. Carrera, G. Dagan, P. B. Davies, D. P. Gallegos, A. Galli, J. Gòmez-Hernández, P. Grindrod, A. L. Gutjahr, P. K. Kitanidis, A. M. Lavenue, D. McLaughlin, S. P. Neuman, B. S. RamaRao, C. Ravenne, and Y. Rubin (1998), A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow, *Water Resour. Res., 34*(6), 1373-1413.