UNIVERSITÀ DEGLI STUDI DI MILANO

FACOLTÀ DI SCIENZE E TECNOLOGIE

Scuola di Dottorato Terra, Ambiente e Biodiversità

Dipartimento di Scienze della Terra

Corso di Dottorato in Scienze della Terra

Ciclo XXVI

Characterization of the subsurface through joint hydrogeological and geophysical inversion

Tutor: Professor Mauro GIUDICI

Cotutor (esterno): Professor Roberto DE FRANCO

Coordinatore: Professoressa Elisabetta ERBA

Tesi di Dottorato di:

Laura CATTANEO

Anno Accademico 2013 - 2014

Commission of the final examination:

Professoressa Annalisa Zaja, Università degli Studi di Padova

Professor Sergio Negri, Università del Salento

Professor Riccardo Tribuzio, Università degli Studi di Pavia

Final examination:

9 Settembre 2014

Università degli Studi di Milano, Dipartimento di Scienze della Terra, Milano, Italy

Contents

List of Figures **v**

List of Tables **ix**

Introduction **ix**
Motivations xi
Goals xii
Methods and approaches xiii

1 A hydrogeophysical integrated approach **1**
1.1 Hydraulic field theory 1
1.2 Geoelectrical field theory 3
1.3 Phenomenological relationship between hydraulic conductivity, electrical resistivity and lithological facies 8
1.4 An original procedure of joint interpretation of hydraulic and geoelectrical data 10

2 Methods **13**
2.1 The Ensemble Kalman Filter Procedure 13
2.2 The stochastic simulation of HGFs 16
2.3 Solving the hydraulic forward problem: the `YAGMod` code 17
2.4 Solving the electrical forward problem: the `YAELMod` code 31
2.5 The implementation of the simplified EnKF for parallel computing: the `JoMod` code 39

3 Application **41**
3.1 The case study 41
3.2 Sensitivity Analysis on boundary conditions 45
3.3 Sensitivity Analysis on parameters 46
3.4 Application of the EnKF procedure 49

4 Discussion and conclusions **63**
4.1 Discussion 63
4.2 Conclusions 65

Bibliography **67**

List of Figures

1.1 Wenner Array Scheme 6
1.2 Schlumberger Array Scheme 6
1.3 Semi-Schlumberger Array Scheme 7
1.4 Dipole-dipole Array Scheme 7
1.5 Pseudosection of apparent resistivity for a dipole-dipole array with equally spaced electrodes. 7
1.6 Conceptual scheme of the methodology adopted in the original procedure. HFP and EFP correspond to, respectively, hydraulic and electrical forward problems. The grey block shows the steps that are used in this work to generate synthetic data in order to perform several tests of the proposed method (see section 3.4.1). 10

2.1 Scheme of the simplified EnKF procedure. 16
2.2 `YAGMod`: a) Plan view of a domain's layer, b) Vertical section view. Red arrows are examples of groundwater fluxes considered in the discrete model. 18
2.3 Domain code example in 2D: `I`, `E` and `D` codes correspond, respectively, to internal, external and prescribed-head (Dirichlet boundary conditions) cells; the blue line denotes the border of the domain. 19
2.4 Scheme of the domain used for the comparison test. 24
2.5 Map of the balance error given by equation (2.29), for the Doherty (2001) approach, as a function of the hydraulic heads h_2 (x axis) and h_5 (y axis), for different values of m (left, $m = 1$; centre, $m = 10$; right, $m = 100$). At the bottom right corner of each graph, the zone containing the minimum error value is enlarged. 27
2.6 Map of the total quadratic balance error given by equation (2.29), as a function of the piezometric heads h_2 (x axis) and h_5 (y axis). From left to right, respectively the results obtained with `YAGMod` and the approaches by Doherty (2001), Keating and Zyvoloski (2009) and Niswonger et al. (2011). From top down, results obtained with extraction rates of $0.1\ \mathrm{m^3\,s^{-1}}$, $0.2\ \mathrm{m^3\,s^{-1}}$ and $0.3\ \mathrm{m^3\,s^{-1}}$. The blue circles point out the zone where the least value of the total quadratic balance error is located. 28
2.7 The GD basin model: Location of some source terms in the GD basin (left) and plot of the zones characterized by different values of hydraulic conductivity. See Table 2.1 for the values of K in each zone. Red diamonds are active wells. Black arrows represents flow lines 29

2.8 Comparison between (a) the piezometric head estimated for the heterogeneous medium obtained for the reference test, (b) that obtained by decreasing the inflow rate due to rain infiltration by 10% and (c) that obtained by increasing the pumping extraction rate due to agricultural use by 20%. 32

2.9 `YAELMod`: Domain discretization. 33

2.10 The block of sediments of the Ticino Valley (Northern Italy) with the electrodic line and a superposition of the apparent resistivity section obtained with `YAELMod` 39

3.1 Aerial map and localization of the well field in Pozzuolo Martesana (Northern Italy). 42

3.2 Sketch of the buried infrastructures of the well field, reproduced from "tavola 13" of Minotta (2002). Wells positions are denoted by brown dots. Green and blue lines correspond to water conduits for, respectively, shallow and deep wells; thin cyan line corresponds to the drinkable water pipe; thick cyan lines correspond to irrigation ditches; the red double line, crossing the area from west/north-west to east/south-east, corresponds to a gas pipeline; red dotted lines correspond to electrical-power lines. 43

3.3 The experimental variogram along z direction. 44

3.4 The geological discretized domain: the eight digitized stratigraphic logs (left) and the corresponding statistical simulation (right). Coarse sediments in cyan, fine sediments in red. 45

3.5 Geometry of the volume domain for sensitivity analysis on BCs: the distribution of HGFs (cyan: coarse material; red: fine material) is shown on two vertical sections, crossing at the pumping well position (yellow vertical bar). 46

3.6 Vertical section of the saturation field (saturated cells in blue, dry cells in brown) along x direction crossing the well: 1) two borders with Dirichlet BCs and two borders with Neumann null flux BCs; 2) two borders with Robin BCs and two borders with Neumann null flux BCs; 3) four borders with Dirichlet BCs; 4) four borders with Robin BCs. 47

3.7 Elevation (in meters) of the water table (contour interval $0.5\ \mathrm{m}$): 1) two borders with Dirichlet BCs and two borders with Neumann null flux BCs; 2) two borders with Robin BCs and two borders with Neumann null flux BCs; 3) four borders with Dirichlet BCs; 4) four borders with Robin BCs. The yellow dotted lines indicates the position of the electrodic lines used for electrical surveys. 47

3.8 Pseudosections of apparent resistivity (the electrodic line is shown by the yellow dotted line in Fig 3.7): 1) the apparent resistivity values obtained with simulation #1 (two borders with Dirichlet BCs and two borders with Neumann null flux BCs) are shown , legend on the top; then, the difference between the results for scenario #1 and other scenarios are mapped: 2) two borders with Robin BCs and two borders with Neumann null flux BCs; 3) four borders with Dirichlet BCs; 4) four borders with Robin BCs, legend on the bottom. 48

3.9 Root Mean Square Variability calculated on model parameters: scale of RMSV for electrical resistivity (C-ρ_w, F-ρ_w, C-ρ_d, F-ρ_d) is given on the left axis, for hydraulic conductivity (C-K, F-K) on the right axis. 50

3.10 Root Mean Variability calculated on model parameters: scale of RMV for electrical resistivity (C-ρ_w, F-ρ_w, C-ρ_d, F-ρ_d) is given on the left axis, for hydraulic conductivity (C-K, F-K) on the right axis. 50

3.11 Sketch of the position of the three activated wells (red dots) and of the electrodic line (blue dotted lines). Black dots represents the other five inactive wells. 51

3.12 Elevation (in meters) of the water table (contour interval $0.5\ \mathrm{m}$): the numbers of the simulations correspond to those listed in table 3.3. The yellow dotted line represents the electrodic line. 51

3.13 Pseudosections of apparent resistivity (the electrodic line is shown by the yellow dotted line in Fig 3.12) obtained for the three different simulations of the synthetic example: 1) the apparent resistivity values obtained with simulation #1 are shown, legend on the top; then, the difference between the results for simulation #1 and other simulations are mapped: 2) simulation #2 3)simulation #3, see Table 3.3 for stress condition details. 52

3.14 Box plot of model parameters as a function of k for the first test. The ensemble mean is represented by an horizontal dark blue bar; the second and third quartiles are represented respectively by light and dark blue boxes; minimum and maximum values are indicated by the thin dark blue error bars. The red lines correspond to the reference values. 54

3.15 Ensemble mean of the total error (left axis) and coefficient of variation of the ensemble mean of the total error with respect to the same simulation at the previous iteration of the "external" loop of the simplified EnKF procedure for the first test. 55

3.16 Box plot of model parameters as a function of k for the second test. The ensemble mean is represented by an horizontal dark blue bar; the second and third quartiles are represented respectively by light and dark blue boxes; minimum and maximum values are indicated by the thin dark blue error bars. The red lines correspond to the reference values. 56

3.17 Ensemble mean of the total error (left axis) and coefficient of variation of the ensemble mean of the total error with respect to the same simulation at the previous iteration of the "external" loop of the simplified EnKF procedure for the second test. 57

3.18 Box plot of model parameters as a function of k for the third test. The ensemble mean is represented by an horizontal dark blue bar; the second and third quartiles are represented respectively by light and dark blue boxes; minimum and maximum values are indicated by the thin dark blue error bars. The red lines correspond to the reference values. 58

3.19 Ensemble mean of the total error (left axis) and coefficient of variation of the ensemble mean of the total error with respect to the same simulation at the previous iteration of the "external" loop of the simplified EnKF procedure for the third test. 59

3.20 Box plot of model parameters as a function of k for the third test. The ensemble mean is represented by an horizontal dark blue bar; the second and third quartiles are represented respectively by light and dark blue boxes; minimum and maximum values are indicated by the thin dark blue error bars. The red lines correspond to the reference values 60

3.21 Ensemble mean of the total error (left axis) and coefficient of variation of the ensemble mean of the total error with respect to the same simulation at the previous iteration of the "external" loop of the simplified EnKF procedure for the fourth test. 61

List of Tables

1.1 Typical scale dimension for studying porous media 1

2.1 Hydraulic conductivity values used in the GD model. 29

3.1 Parameters for the computation of the experimental variogram and for the theoretical variogram. 44
3.2 Hydrogeological parameters values assigned to HGFs. 45
3.3 Extraction rates of the three wells for the three different simulated scenarios. 51
3.4 Numerical results to test the invertibility of the covariance matrix $\mathrm{Cov}^{(k)}[YY]$. 53
3.5 Values used for the generation of the initial parameters for the fourth test. 59

Introduction

Motivations

Characterization of the subsurface heterogeneity, monitoring groundwater dynamics, modelling flow and transport in the subsoil are of paramount importance for protection of groundwater quality, design of remediation plans, control of restoration activities.

One of the key physical parameters that control groundwater flow and solute transport is hydraulic conductivity, K. The possibility of predicting K from geophysical measurements has resulted in a large volume of literature (Archie, 1942; Biella et al., 1983; Frohlich and Summerly, 1996; de Limaa and Sri Niwas, 2000; Hubbard and Rubin, 2000; Purvance and Andricevic, 2000; Niwas and de Lima, 2002; Purvance, 2003; Slater and Lesmes, 2002; Slater, 2007; Soupios et al., 2007; Sènèchal and Sènèchal, 2009; Hinnell et al., 2010; Camporese et al., 2011; Fowler and Moysey, 2011; Böhm et al., 2013) focused on this objective over the last 60 years. These research efforts have been largely motivated by two basic factors.

First, several geophysical parameters and, in particular, those related to electrical and electromagnetic methods, depend on porosity, water content and textural properties (including specific surface and grain-size distribution), which are of paramount importance also to determine the K values of soils, unconsolidated sediments and rocks.

Second, but not of minor relevance, it is difficult to directly and effectively measure K. In fact, invasive direct tools, as well as geotechnical and direct probe surveys, are mostly based on the drilling of piezometers and boreholes, are sparsely distributed or permit to acquire data at scales inappropriate to study flow and transport processes. Moreover, borehole data often yield information with accurate vertical resolution, but require horizontal interpolation or correlation. Hydrogeophysics provides useful complementary techniques, both for hydrostratigraphic and hydrogeological characterization and for monitoring. It provides a minimally invasive approach to obtaining spatially-continuous data-sets, at a relatively high temporal and spatial sampling density.

Therefore, a geophysical approach to the estimation of K and other hydrogeological parameters is attractive because it is non-invasive or minimally invasive, it permits to investigate a large volume, depending on the inherent limitations of any particular geophysical technique, and geophysical measurements are relatively quick and inexpensive when compared to other field tests, including pumping and slug tests. The same motivations support the application of geophysical prospecting, in particular geoelectrical and electromagnetic methods, as a monitoring tool for hydrogeological applications (De Franco et al., 2009; Hayley et al., 2009; Wilkinson et al., 2010; Nimmer et al., 2007).

A large number of empirical equations has been proposed to convert geophysical measurements into hydraulic parameters (Kelly, 1977; Kosinky and Kelly, 1981; Hinnell et al., 2010; Pollock and Cirpka, 2010; Fowler and Moysey, 2011; Straface et al., 2011; Jardani et al., 2013; Lochbuehler et al., 2013). One of the key problems is how to differentiate the effects that the solid matrix and the pore fluids have on the bulk geophysical and hydraulic properties.

Many different multidisciplinary approaches have been adopted in recent years to asses hydraulic properties of the soil from geophysical investigation. Among the others, Böhm et al. (2013) combine high-resolution hydraulic and geophysical tomographic measurements to define site-specific relationships between geophysical and hydraulic parameters; Frohlich and Summerly (1996) combine fracture trace analysis and geoelectrical depth soundings for locating bedrock fractures and they obtain a relationship between resistivity and hydraulic conductivity from comparison of observation (electrical and hydraulic) data.

Moreover, many coupled inversion methods have been tested with many different approaches. Pollock and Cirpka (2010) perform a numerical experiment in order to simulate salt tracer tests and then determine hydraulic conductivity from monitoring the process by electrical resistivity tomography (ERT). Straface et al. (2011) interpret hydraulic head data from direct measurements in boreholes and self-potential measurements, through the assumption of a linear relationship between the saturated thickness of the aquifer and the measured self potential signals. All the collected hydraulic head data are then embedded jointly within a three-dimensional (pseudo-steady state) inverse model of groundwater. Fowler and Moysey (2011) evaluate whether electrical resistivity monitoring data can constrain the hydraulic conductivity, porosity, and dispersivity of a homogeneous aquifer within the framework of coupled inversion, where the advection-dispersion equation of the transport process is solved analytically and bulk electrical resistivity is assigned following Archie's law (Archie, 1942). The hydrologic and electrical conductivity models developed by Hinnell et al. (2010) are linked by the dependence of the soil electrical conductivity on the soil water content. Through a geophysical inversion they obtain the spatial distribution of a geophysical property (e.g., electrical conductivity). The geophysical property is then converted to a hydraulic property (e.g., water content) through Archie's law, and the inferred hydraulic property is then used either independently or together with direct hydrologic observations to constrain a hydrologic inversion.

Furthermore, many studies involved the ensemble Kalman filter (EnKF) technique (Evensen, 2009; Gillijns et al., 2006) as a method for parameter estimation. In Camporese et al. (2011) an approach based on the Lagrangian formulation of transport and the EnKF technique is applied to assess the spatial distribution of hydraulic conductivity by incorporating time-lapse cross-hole ERT data; Ferraresi et al. (1996) characterize the hydrodynamic behaviour of aquifers given their geometrical description, the piezometric field, the net infiltration and the boundary conditions, considering only hydrogeological models.

Goals

The research conducted during this PhD work can be introduced in this panorama and is mainly aimed to applications for the study of alluvial aquifers, even if a lot of the concepts that are introduced and a lot of the methods that are developed and tested could be applied to fractured or karst aquifers in sedimentary rocks (e.g., limestone or sandstone).

The general-purpose objective of the work is the development of a modelling tool for the subsurface characterization, in order to improve studies on groundwater flow and contaminant transport, with the specific goal of obtaining a spatial 3D parameter distribution of hydraulic conductivity K and electrical resistivity ρ. Such a tool profits from DC geoelectrical and hydraulic collected data, which are used in a joint geophysical and hydrological data inversion, with an approach similar to the EnKF.

The basic idea is to extend the concept of hydrofacies (Anderson, 2009; Klingbei et al., 1999). Namely, the subsoil is described as a collection of sub-domains, each of which is occupied by a hydrogeoelectrical facies (HGF), i.e., sediments or rocks which share similar characteristics from the lithological, hydraulic and electrical point of view. The two physical parameters that are investigated, hydraulic conductivity K and electrical resistivity ρ, depend on the properties of HGFs through semi-empirical or phenomenological formulas, like, e.g., Archie's law. In principle, such semi-empirical relations should take into account not only textural, petrographical and mineralogical properties of the solid matrix but also the content and properties of pore fluids; this requires the introduction of fitting (empirical or phenomenological) parameters, which can be different for each HGF. Finally, the spatial distributions of HGFs and of pore water characteristics determine the spatial variability of K and ρ.

Methods and approaches

The application of the proposed modelling tool requires an accurate monitoring activity. In fact, the EnKF method, or other similar approaches, need different sets of experimental data at different times: both hydraulic observed data and geoelectrical field data must be collected simultaneously. However, the technology, the standards for field data acquisition, the software for data processing are already well developed and progressively updated, not only from academic research centres, but also from private companies. On the other hand, there is also a great variety of modelling tools that can be used in the geophysical interpretation, but this is the specific aspect on which this PhD work is concentrated.

In fact, two models, YAGMod (Yet Another Groundwater flow Model) and YAELMod (Yet Another ELectrical resistivity Model), are proposed to solve, respectively, the 3D hydraulic and electrical forward problems and are implemented with original computer codes (`YAGMod` and `YAELMod`) developed in Fortran programming language. The predicted results are compared with field measurements in an iterative process, in order to reach a more reliable estimate of the empirical parameters for each HGF and, ultimately, of the hydraulic conductivity and the electrical resistivity. For this goal, an original code (`JoMod`) developed with Octave, a programming language compatible with Matlab®, has been implemented to run on a multi-processor parallel computer.

In practical applications, a variety of computer codes, based on the classical numerical methods of solution of partial differential equations (finite differences, finite elements, finite volumes, etc.), are used to solve the forward problems of groundwater hydrology and geoelectrical prospecting.

Modelling groundwater dynamics

In the field of groundwater hydrology MODFLOW (MODFLOW, 2005) is among the most powerful and widely applied models. It is based on a conservative finite difference scheme, for which the conservation equation and a discrete version of Darcy's law are applied to a 3D network. A similar scheme, applied to blocks with arbitrary shape,

is given by integrated finite differences (Narashiman and Witherspoon, 1976), that are implemented to simulate density-dependent groundwater flow with variable saturation with the Tough2 set of codes (TOUGH2, 1991). Other popular codes, e.g., Feflow (Trefry and Muffels, 2007), Femwater (Lin et al., 1997), SUTRA (SUTRA, 2010), simulate density-dependent flow of groundwater for variable saturation with the finite element method. In some cases they also model transport of conservative and reactive solutes and energy. Some packages, e.g., Mike (MIKE, 2003), are also available to integrate models that simulate water flow in surface bodies (rivers and lakes) and in buried aquifers and the water exchange between surface and subsurface water bodies.

Therefore, in principle, there is such a wide set of available models that there is no need of developing new codes. However, it should be realised that two kinds of complexities should be dealt with, while working on flow and transport in porous media: (1) the complexity of the processes which might interact to determine the groundwater flow field, the distribution of a solute or the thermal field in the subsurface, and (2) the complexity of the hydrostratigraphic structure. In order to handle complex processes, e.g., coupling water, solute and heat transport, it is necessary to estimate a great number of phenomenological parameters that are seldom measured in the field. Therefore several assumptions or guesses have to be introduced, the optimal values can be found with inverse methods and the effects of their uncertainties on the outcome can be estimated with a sensitivity analysis (Hill and Tiedeman, 2007; Saltelli et al., 2008). On the other hand, if heterogeneous media are considered, then the spatial distribution of the physical parameters has to be known; if the heterogeneity causes the physical parameters to vary by some orders of magnitudes at small distances, further difficulties arise, both from the practical and the numerical points of view. Complexities related to both physical processes and heterogeneity might cause the model to work with input data that correspond to extreme, critical conditions for which the model run could crash or fail to find a physically acceptable solution.

Therefore, there is still place to develop numerical codes which can handle particular situations more easily than general purpose software packages. In particular this work was prompted by practical experiences with two different modelling frameworks. First, the use of a computer code for inverse modelling and sensitivity analysis might require the simulation of conductivity fields, which exhibit great heterogeneity and great contrast of parameters between adjacent blocks and which could make it difficult to handle non-linear source terms. Second, regional aquifers could require the simulation of complex situations, if the source terms depend on the aquifer water head, e.g., for drainage terms or river-aquifer interaction, and if the well extraction is so great as to dry some portions of the subsurface.

Within this background, the `YAGMod` code was developed in Fortran90, for the simulation of constant density, groundwater flow under stationary conditions, which is the extension of the codes developed by Università degli Studi di Milano over the years (Ponzini et al., 1989b; Bersezio et al., 1999; Giudici et al., 2000, 2007, 2012; Lunati et al., 2001; Valota et al., 2002; Zappa et al., 2006; Felletti et al., 2006; Vassena et al., 2008, 2010, 2011). `YAGMod` is based on a conservative finite difference scheme for stationary conditions and is oriented to the simulation of flow in saturated media, but it takes into account the possible drying of shallow blocks of the domain with an original approach. With respect to most of the existing codes, `YAGMod` handles different types of source terms and boundary conditions. In particular it considers both prescribed distributed sources, which could simulate aquifer recharge, and variable point sources, which can be used to simulate draining systems, but also the effects of the water head drawdown on a water well discharge.

Modelling geoelectrical fields

Also for the solution of the electrical forward problem, different published models are available and they were carefully considered. In particular, the attention was addressed to: RES3DMOD (Geotomo Software, 2001; Locke, 2001), developed by GEOTOMO Software Sdn. Bhd.; R3t (Binley, 2013a,b), developed by Lancaster University; RESINVM3D (Pidlisecky et al., 2013, 2007), developed by Stanford University.

The impossibility of obtaining source codes, the complexity of developing tools for pre- and post- processing of I/O files in order to merge the codes with `YAGMod` or other codes for the forward problem of groundwater hydrology and the uncertainties on the accuracy of the model outputs supported the decision to develop the original model YAELMod. The latter is based on assumptions and methods for the discretization of the domain and of the equations which are very similar to those used for YAGMod, so that the two codes are easily coupled to perform simulations in cascade.

Modelling groundwater dynamics

Finally, the approach proposed to join hydrogeological and geoelectrical modelling can be summarized with the steps, described in the following list.

1. Collection of hydrostratigraphic data (typically, litho-stratigraphic borehole logs), definition of HGFs, choice of the semi-empirical formulas relating different physical quantities for each HGF and definition of the empirical parameters to be identified.

2. Reconstruction of the 3D facies distribution, possibly through conditional geostatistical simulations.

3. Field monitoring and experimental activity, i.e., collection of hydrogeological (hydraulic heads) and geoelectrical (apparent resistivities) data at different times, under different conditions of aquifer exploitation.

4. Application of an iterative procedure, analogous to the EnKF and based on two nested loops, in order to obtain the optimal sets of the empirical parameters for each HGF and possibly an estimate of the uncertainty on such values. The "external" iterative loop consists of the following steps.

   (a) Selection of an ensemble of sets of empirical parameters.

   (b) Improving the estimate of the empirical parameters, by using all the available field data with an "internal" loop which consists of the following steps:

      i. solution of the hydraulic and the electrical forward problems, in cascade, to model the hydraulic and saturation fields and successively to obtain the modelled values of apparent resistivity for each time at which field data are available;

      ii. ranking the sets of parameters through a suitably defined objective function, based on the comparison of field data and model predictions;

      iii. computing the modified ensemble of sets of phenomenological parameters, in order to improve their estimate and reduce the uncertainty;

      iv. iterative repetition from step 4(b)i for each time at which field data are available.

(c) Checking if the estimates of the empirical parameters and the fitting error (difference between modelled and measured quantities) are significantly changed with respect to the previous iteration of the external loop; if not, then the external loop continues from step 4b.

CHAPTER 1

A hydrogeophysical integrated approach

1.1 Hydraulic field theory

1.1.1 The porous media

Porous media and, in particular, alluvial sediments, could be studied at different scales (Neuman and Di Federico, 2003; Giudici, 2010): in table 1.1 some typical scale dimensions are listed.

Scale	Characteristic length	Main Feature
Molecular	from $3 \cdot 10^{-10}$ m to $5 \cdot 10^{-10}$ m	Typical dimension of water molecule
Microscopic	from 10^{-6} m to 10^{-3} m	Typical dimension of pores
Macroscopic	from 10^{-4} m to 1 m	Typical dimension of REV
Megascopic	from 1 m to 10^{2} m	Typical dimension for a field problem at local scale (flow test)
Gigascopic	from 10^{2} m to $> 10^{3}$ m	Typical dimension for a field problem at regional scale (aquifer system)

Table 1.1: Typical scale dimension for studying porous media

At the macroscopic scale, physical quantities could be considered as mean values of the corresponding physical quantities at the microscopic scale, calculated over a *representative elementary volume* (REV[1]). Therefore, physical properties of porous media and fluids are described through functions that enter in the equations of geophysical fluid dynamics.

1.1.2 Geophysical fluids dynamics

At the microscopic scale, groundwater flow is governed by the *continuity equation* that expresses the principle of mass conservation in mathematical terms (de Marsily, 1986)

$$\mathrm{div}(\zeta \mathbf{u}) + \frac{\partial \zeta}{\partial t} = 0, \tag{1.1}$$

where ζ $(\mathrm{kg/m^3})$ is fluid density and $\mathbf{u}$ $(\mathrm{m/s})$ the microscopic velocity, and by *Navier-Stokes' equation*, which expresses the second principle of dynamics, that under steady

[1]For every volume of porous medium which is smaller than the REV, the mean value of any physical quantity depends on the volume size. The REV is the smallest unit volume containing granules and pores, over which it is possible to calculate a statistically significant mean.

conditions and for an incompressible fluid assumes the form

$$\mathbf{grad}p - \mu\nabla^2\mathbf{u} = \zeta\mathbf{F}, \tag{1.2}$$

where p (Pa) is pressure, μ ($\mathrm{kg/(m\,s)}$) is dynamic viscosity of the fluid, ∇^2 is the Laplace operator and $\mathbf{F}$ ($\mathrm{N/m^3}$) is the resultant force per unit volume.
At the macroscopic scale, it is necessary to rewrite the basic equations in terms of mean quantities. In fact, Navier-Stokes' equation could not be solved inside each REV, and it is substituted by *Darcy's law* which relates pressure, velocity and external forces for groundwater flux through a porous medium. Darcy's law is a simple proportional relationship between the instantaneous discharge rate through a porous medium and the hydraulic gradient. Since water flow occurs through the volume of connected pores, the discharge rate is related to the mean pore water velocity by the dynamic porosity ϕ [2]:

$$\mathbf{q} = \phi\mathbf{u}. \tag{1.3}$$

The phenomenological Darcy's law takes the form

$$Q = -K\mathcal{A}\frac{\Delta h}{\ell}, \tag{1.4}$$

where $\mathcal{A}$ ($\mathrm{m^2}$) is the section of the porous sample for which this physical law is tested, ℓ (m) the length of the sample, Δh (m) the difference in hydraulic head between the edges of the sample, K ($\mathrm{m/s}$) the hydraulic conductivity and $\mathbf{Q}$ ($\mathrm{m^3/s}$) the total flux. The laboratory experiment was conducted with a totally saturated sample and under stationary conditions. From equation (1.4) it is possible to write the expression for the specific discharge, i.e., the flow rate per unit surface,

$$q = \frac{Q}{\mathcal{A}} = -K\frac{\Delta h}{\ell} = -KJ, \tag{1.5}$$

where J represents the hydraulic gradient. In a continuum domain, equation (1.5) could be written in differential form

$$\mathbf{q} = -K\mathbf{grad}h. \tag{1.6}$$

Repeating the Darcy's experiment for different fluids in the same porous medium, it is easy to asses that the hydraulic conductivity, i.e. the proportionality coefficient between specific discharge and hydraulic gradient, depends on both the porous medium and the nature of the fluids, so that it could be expressed through the relation

$$K = \frac{\kappa\zeta g}{\mu}, \tag{1.7}$$

where g is the gravity acceleration ($\mathrm{m/s^2}$) and κ is the intrinsic permeability ($\mathrm{m^2}$).

Recalling that the hydraulic (or piezometric) head could be expressed by

$$h = z + \frac{p}{\zeta g}, \tag{1.8}$$

[2] If V is the total volume of porous medium and V_P the pore's volume in that elementary volume the *total porosity* is defined as $\phi = V_P/V$. Dynamic porosity is the ratio between the volume of the pores, through which water can effectively flow, and V. Hereinafter, for the simplicity's sake, the same symbol will be used for total and dynamic porosity.

Darcy's law could be written as

$$\mathbf{q} = -\frac{\kappa}{\mu}\left(\mathbf{grad}p + \zeta g\mathbf{grad}z\right). \tag{1.9}$$

The balance equation for a steady flow in a porous medium is now

$$-\mathrm{div}\mathbf{q} + f = 0, \tag{1.10}$$

where f represents the *source term*, i.e. the water discharge per unit volume.
With relation (1.5), equation (1.10) becomes

$$\mathrm{div}(K\mathbf{grad}h) + f = 0. \tag{1.11}$$

Notice that for partially saturated porous media, the matric potential, ψ (m), should substitute the pressure head $p/(\zeta g)$ in (1.8). Moreover, both ψ and κ (and therefore K) should depend on the volumetric water content in a non-linear fashion; these characteristic curves could exhibit an hysteretic behaviour and could be fitted with formulas that include several phenomenological parameters.

1.2 Geoelectrical field theory

1.2.1 Elementary theory of stationary electrical currents

Electric current density $\mathbf{j}$ ($\mathrm{A/m^2}$) is related to electric field $\mathbf{E}$ ($\mathrm{V/m}$) through Ohm's law

$$\mathbf{j} = \frac{1}{\rho}\mathbf{E}. \tag{1.12}$$

Under stationary conditions, the electric field is conservative, so that it could be expressed as the gradient of the scalar potential, i.e. $\mathbf{E} = -\mathbf{grad}V$, and equation (1.12) becomes

$$\mathbf{j} = -\frac{1}{\rho}\mathbf{grad}V. \tag{1.13}$$

To describe the electric current in heterogeneous, but continuous and isotropic media, in absence of sources or sinks of electric charge, an arbitrary closed volume Ω, delimited by the surface Σ, is considered: for the principle of charge conservation, the surface integral representing the total flux of charge entering the volume vanishes. If $\mathbf{n}$ is the unit vector perpendicular to the surface Σ, pointing inside the volume Ω,

$$\int_{\Sigma} \mathbf{j}\cdot\mathbf{n}\,\mathrm{d}\Sigma = 0, \tag{1.14}$$

which, by applying the divergence theorem, becomes

$$-\int_{\Omega} \mathrm{div}\mathbf{j}\,\mathrm{d}\Omega = 0. \tag{1.15}$$

Since (1.15) holds for any arbitrary, but fixed, volume Ω, to satisfy equation (1.15) it is necessary that the integrand vanishes everywhere, so that

$$\mathrm{div}\mathbf{j} = 0. \tag{1.16}$$

As done for equation (1.13), the current density could be expressed in terms of the electric potential, so

$$\mathrm{div}\left(\frac{1}{\rho}\mathbf{grad}V\right) = 0. \tag{1.17}$$

In a homogeneous and isotropic medium, equation (1.17) assumes a simpler form, due to the fact that the electrical conductivity is constant all over the domain:

$$\mathrm{div}\left(\frac{1}{\rho}\mathbf{grad}V\right) = \frac{1}{\rho}\nabla^2 V \tag{1.18}$$

and Laplace's equation is obtained:

$$\nabla^2 V = 0. \tag{1.19}$$

When a point source of charge is introduced in the domain, supposing to be centred at a point $\mathbf{x}_s \in \Omega$, the surface integral over the volume Ω becomes

$$\int_\Sigma \mathbf{j}\cdot\mathbf{n}\,\mathrm{d}\Sigma = -I \tag{1.20}$$

where I is the injected electrical current.

The flux through the surface Σ equals the electric current injected by the point source of charge, under stationary conditions, so that the differential equation to be solved becomes

$$\mathrm{div}\left(\frac{1}{\rho}\mathbf{grad}V\right) = I\delta(\mathbf{x}-\mathbf{x}_s) \tag{1.21}$$

in heterogeneous media, while for homogeneous media the equation reduces to Poisson's equation

$$\nabla^2 V = \rho I\delta(\mathbf{x}-\mathbf{x}_s). \tag{1.22}$$

Analytical solutions of Poisson's equation (1.22) can be obtained by considering a single point charge source, inserted in a homogeneous and infinitely extended domain; in this case the solution depends on the distance r from the point source. Equipotential surfaces have spherical shape and the electric field $\mathbf{E}$ has radial direction, normal to equipotential surface. If Σ is a spherical surface with radius r, centred on the point source, the electric charge flux is given by

$$\mathbf{j}\cdot\frac{\mathbf{r}}{r}4\pi r^2 = -\frac{1}{\rho}\frac{\partial V}{\partial r}4\pi r^2, \tag{1.23}$$

where $\mathbf{r}$ is the radial vector starting from the charge source. In absence of other charge sources, the charge flux given by (1.23) must be equal to the electric current intensity injected by the electrode:

$$-\frac{1}{\rho}\frac{\partial V}{\partial r}4\pi r^2 = I. \tag{1.24}$$

Then, integrating equation (1.23) with respect to r, the electric potential attains the form

$$V(r) = \frac{I\rho}{4\pi r} + c, \tag{1.25}$$

where the constant c could be calculated by considering that $V \to 0$ for $r \to \infty$, so that $c = 0$.

The formula

$$V(r) = \frac{I\rho}{4\pi r} \tag{1.26}$$

represents the electric potential produced by a point electrode, injecting a stationary current I in a homogeneous and isotropic, infinite medium, characterized by resistivity ρ. The analytical solution of equation (1.26) could be simply calculated for a point electrode, posed on a plane delimiting a semi-space, such as a plain ground, for example. In this case, since air can be considered as an insulator, the electrical current flows through a solid angle which is exactly half of that considered before, the surface Σ is a semi-sphere. So, the potential is given by

$$V(r) = \frac{I\rho}{2\pi r}. \tag{1.27}$$

1.2.2 Geoelectrical prospecting techniques

These results are the basic properties of geoelectrical prospecting in order to investigate subsurface characteristics and properties. Most of the geoelectrical prospecting techniques use four electrodes: two power, or current electrodes, named A and B, and two potential electrodes, M and N. The current electrodes (A and B), between which an electrical potential difference is kept constant in time, generate a stationary electric field. Under the hypothesis of independence of the electrical resistivity from electric field or electric potential (process linearity), the superposition principle could be used to calculate the total electric potential as the sum of the electric potentials due to each current electrode separately. The potential electrodes (M and N) are used to measure the *epd* (electrical potential difference) existing between the two points where they are located.

Every quadrupole, formed by the current and potential dipoles, allows to calculate a value of *apparent resistivity*, from the field measurements of electric current and epd. ΔV, the epd measured between M and N, can be expressed by applying the superposition principle,

$$\Delta V = V(M) - V(N) = [V_A(M) + V_B(M)] - [V_A(N) + V_B(N)], \tag{1.28}$$

where V_A and V_B represent the potential fields generated by each individual electrode of the current dipole. Then, using the solution to Poisson's equation (1.22) calculated for a single point electrode posed at the surface of a semi-infinite domain as in equation (1.27), ΔV given by (1.28) becomes:

$$\Delta V = \frac{I\rho}{2\pi\overline{AM}} - \frac{I\rho}{2\pi\overline{AN}} - \frac{I\rho}{2\pi\overline{BM}} - \frac{I\rho}{2\pi\overline{BN}} \tag{1.29}$$

and so

$$\Delta V = \frac{I\rho}{2\pi}\left(\frac{1}{\overline{AM}} - \frac{1}{\overline{AN}} - \frac{1}{\overline{BM}} - \frac{1}{\overline{BN}}\right). \tag{1.30}$$

Distances between electrodes, electric current I and epd ΔV are all measurable quantities and the only parameter of equation (1.30) to be determined is ρ. Under the above mentioned assumptions, ρ is the hypothetical resistivity for a homogeneous domain and corresponds with the real resistivity of the medium; in practice, the value obtained from (1.30) is named *apparent resistivity* ρ_{app}, because in general it does not coincide with a "real value" of resistivity of the ground. The apparent resistivity can be easily calculated from equation (1.30) as:

$$\rho_{app} = \frac{\Delta V}{I} 2\pi P, \tag{1.31}$$

where

$$P = \left(\frac{1}{\overline{AM}} - \frac{1}{\overline{AN}} - \frac{1}{\overline{BM}} - \frac{1}{\overline{BN}} \right)^{-1} \tag{1.32}$$

is the geometric parameter, which depends only on the geometrical configuration of the electrodes.

In the following, the most common quadrupole setups (or geoelectrical arrays) are recalled. All the analysed configurations assume that the four electrodes are aligned.

- Wenner array, shown in figure (1.1), is symmetrical with respect to the centre of the array and the distance between adjacent electrodes a is kept constant and the geometric parameter is given by

$$P = \left(\frac{1}{a} - \frac{1}{2a} - \frac{1}{2a} + \frac{1}{a} \right)^{-1} = a. \tag{1.33}$$

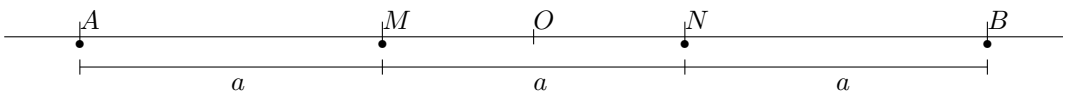


Figure 1.1: Wenner Array Scheme

- Schlumberger array, shown in figure (1.2), is symmetrical with respect to O, but the distance between potential electrodes $2b$ is very small respect to the distance between current electrodes $2a$, i.e. $\overline{MN} \ll \overline{AB}$, so that $b \ll a$: in this case,

$$P = \left(\frac{1}{a-b} - \frac{1}{a+b} - \frac{1}{a+b} + \frac{1}{a-b} \right)^{-1} = \frac{a^2 - b^2}{4b}. \tag{1.34}$$

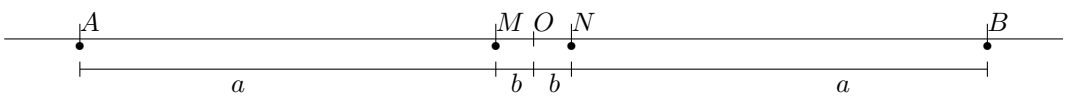


Figure 1.2: Schlumberger Array Scheme

- Semi-Schlumberger array, shown in figure (1.3) has the same disposition of the Schlumberger array, but the electrode B is posed very far from other electrodes, in order to make its effect negligible. In this case

$$P = \left(\frac{1}{a-b} - \frac{1}{a+b} \right)^{-1} = \frac{a^2 - b^2}{2b}. \tag{1.35}$$

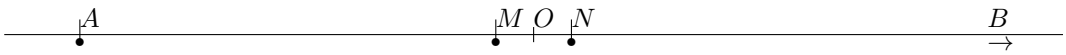


Figure 1.3: Semi-Schlumberger Array Scheme

- Dipole-dipole arrays consider generic dispositions of the two dipoles; here the polar dipole-dipole array is considered and shown in figure (1.4). It is not symmetrical and it is usually used with automatic systems of data acquisition, working with multiple equally spaced electrodes. The typical length of both AB and MN dipole is a, the electrode spacing, and the distance between the two dipoles is na, so that the geometric parameter assume the form

$$P = n(n+1)(n+2)\frac{a}{2}. \tag{1.36}$$

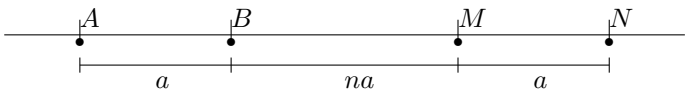


Figure 1.4: Dipole-dipole Array Scheme

Results obtained from measures made with multi-electrode acquisition systems are usually represented with *pseudosections* of apparent resistivity. Every quadrupole yields a value of apparent resistivity, which is associated at a point in the vertical section under the electrodes array. For dipole-dipole arrays this value is located at the intersection of the straight lines which start from the midpoints of the two dipoles and form a fixed angle with the ground surface (45 degrees for the sketch in figure 1.5).

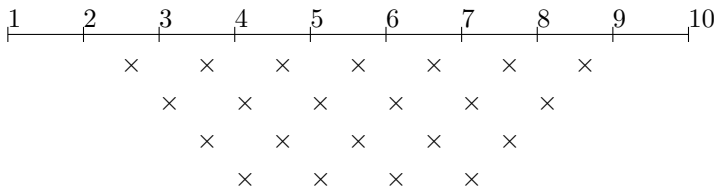


Figure 1.5: Pseudosection of apparent resistivity for a dipole-dipole array with equally spaced electrodes.

1.3 Phenomenological relationship between hydraulic conductivity, electrical resistivity and lithological facies

This work is based on a model of the subsurface as a collection of regions occupied by geological materials, called hydro-geoelectrical-facies, HGF. An HGF is a block of sediments that share the same geoelectrical and hydrodynamic characteristics. Electrical resistivity and hydraulic conductivity are related to texture, soil saturation, pore connectivity, pore water electrical conductivity, etc., by phenomenological laws that include specific parameters (Archie, 1942; Slater, 2007). In particular, the electrical resistivity or its inverse, the electrical conductivity, is a function both of the aquifer's petrophysical properties and of the electrical properties of pore fluids (Keller and Frischknecht, 1966; Telford et al., 1990; Slater, 2007; Reynolds, 2011).

Electrical resistivity measurements have often been used to map soil water content (Zhou et al., 2001; Michot et al., 2003; Di Maio and Piegari, 2011). On the other hand, resistivity measurements could be useful to address also the two major, challenging tasks mentioned in the Introduction: (a) the delineation of the external geometries and of the internal heterogeneities of hydrostratigraphic systems (i.e., the succession of aquifers and aquitards) and (b) the estimation of the subsurface physical properties (Bernabe and Revil, 1995; Huggenberger and T., 1999; Hubbard and Rubin, 2000; Bridge and Hyndman, 2004; Binley and Deiana, 2010; Hubbard and Rubin, 2005). Borehole electrical logs and laboratory experiments on core samples allow to analyse the electrical properties of the hydrostratigraphic sequences at the microscopic (pores and grains) and mesoscopic (representative elementary volume, REV) scales (Giudici, 2010), see table 1.1. At these scales, quantitative information about aquifer physical properties can be provided by linking the electrical and sedimentary properties through predictive petrophysical models based on the integration of the traditional facies and architectural analysis of depositional systems (Galloway and Sharp, 1998; Bridge and Hyndman, 2004) with the knowledge of the physical processes of electrical charge transport occurring within the porous media. Porosity, water saturation and pore connectivity control hydraulic properties (intrinsic permeability, hydraulic conductivity), but they also control electric properties (electrical resistivity) of porous media. In fact, the solid phase of porous media is generally characterized by low electrical conductivity, with the exception of clay minerals. Porosity allows fluids with a certain brine concentration to penetrate the porous media and high pore connectivity enhances the continuity of the electrolytic solution. In presence of an electric field, electric current is generated, mainly due to electrolytic effects, i.e., to the movement of electrical charges dispersed in the pore fluid.

For porous systems (rocks and unconsolidated sediments) the relation between electrical resistivity of the fluids and of the sediments and rocks is often described with the empirical *Archie's law* (Archie, 1942), that could be applied only to rocks characterized by salt water in the pores and a non conductive solid matrix (Klein and Sill, 1982; de Lima and Sharma, 1990; Worthington, 1993; Purvance and Andricevic, 2000). Under these conditions, electrical current is due primarily to electrolysis and Archie's law is given by:

$$\rho_s = a\rho_f \phi^{-m} S^{-n}, \tag{1.37}$$

where ρ_s is the electrical resistivity of solid matrix, ρ_f is the electrical resistivity of the fluid, S is the saturation index, defined as the ratio between fluid volume and pore volume[3], and ϕ is the porosity; a is a numerical coefficient depending on the type of forma-

[3]More precisely, S is given by $S = (\theta - \theta_r)(\theta_s - \theta_r)^{-1}$, where θ, θ_r and θ_s are the actual, residual and saturated volumetric water content.

tion, m is the cementation factor and n a numerical value, that typically equals 2. The ratio

$$F = \frac{\rho_s}{\rho_f} = a\phi^{-m}S^{-n} \tag{1.38}$$

is called *formation factor* and does not depend on the nature of the fluid that occupies pores.

Archie's law (1.37) is the generalization of the result obtained by a simple model which considers a fully-saturated porous medium as composed by two distinct phases (the solid granular phase and the liquid phase that fully occupies the pores); when an epd is applied at the edge of a cylindrical sample of porous medium, the sample behaves as an equivalent resistor consisting of two parallel resistances:

$$\frac{1}{R_{eq}} = \frac{1}{R_f} + \frac{1}{R_s}, \tag{1.39}$$

where R_{eq}, R_f and R_s are respectively the resistances of the sample, the fluid and the solid matrix. Ohm's laws

$$\Delta V = RI \text{ and } R = \rho\frac{\ell}{\mathcal{A}} \tag{1.40}$$

can be applied to estimate the total resistance of the sample, whose length and section are respectively ℓ and $\mathcal{A}$. By considering that the effective fluid section is $\phi\mathcal{A}$ and the effective grain section is $(\phi - 1)\mathcal{A}$, each term of (1.39) can be substituted with the corresponding form of equations (1.40), so that (1.39) can be written as

$$\frac{1}{\rho\ell\mathcal{A}^{-1}} = \frac{1}{\rho_f\ell(\phi\mathcal{A})^{-1}} + \frac{1}{\rho_s\ell[(\phi-1)\mathcal{A}]^{-1}}. \tag{1.41}$$

Then,

$$\frac{1}{\rho} = \frac{\phi}{\rho_f} + \frac{(\phi-1)}{\rho_s}, \tag{1.42}$$

and, under the hypothesis that $\rho_f \ll \rho_s$, the simplest formulation of Archie's law is finally obtained:

$$\rho = \rho_f\phi^{-1}. \tag{1.43}$$

Unfortunately, freshwater alluvial aquifers are often formed by a mixture of coarse-grained granules (sand and gravel) and fine-grained particles (silt and clay) at fine scale (Keller and Frischknecht, 1966; Schön, 2004; Ellis and Singer, 2007) and by a mixture of textures at the larger scales. Moreover, the salinity of pore-water contained in freshwater aquifers is rarely so high as to make electrolytic conduction the fully dominant process. These remarks limit the validity of Archie's law in its simplest formulations.

Several early works for predicting intrinsic permeability of the medium were based on capillary tubes models, whose fundamental basis is Kozeny-Carman's equation (Carman, 1939; Kozeny, 1927) or its modified form obtained by Börner and Schön (1991), Lesmes and Friedman (2005) and Slater (2007).

Recently, a relationship between electrical and hydraulic conductivity has been proposed by assuming that equations from percolation theory (Katz and Thompson, 1986; Thompson et al., 1987) could be applied also to unsaturated media (Doussan and Ruy, 2009). Moreover, Mele et al. (2014) proposed a phenomenological model able to describe the different electrical conduction mechanisms (electrolytic and shale), in order to investigate the relationship between the bulk electrical properties, the pore fluids and the textural properties of samples of alluvial sediments characterized by a homogeneous

texture. Among all factors that contribute to the global electrical response of a porous medium, fluid properties and textural properties are those which mostly affect the bulk electrical response of the samples. Mele et al. (2014) introduced an additive term to a linear expression of Archie's law and proposed a three-component model that establishes a phenomenological relationship among bulk electrical conductivity, pore-water electrical conductivity and textural properties, expressed by porosity and grain-size distribution, and permits to quantify the prevalence of electrolytic or shale conduction for fully saturated samples.

1.4 An original procedure of joint interpretation of hydraulic and geoelectrical data

The main deliverable of this research work is an original procedure for the integrated interpretation of hydraulic and geoelectrical field data. The procedure is sketched in Fig. 1.6. The principal steps are briefly outlined and summarized in this section, together with some notation that will be used later, whereas the details are fully described in the following chapters.

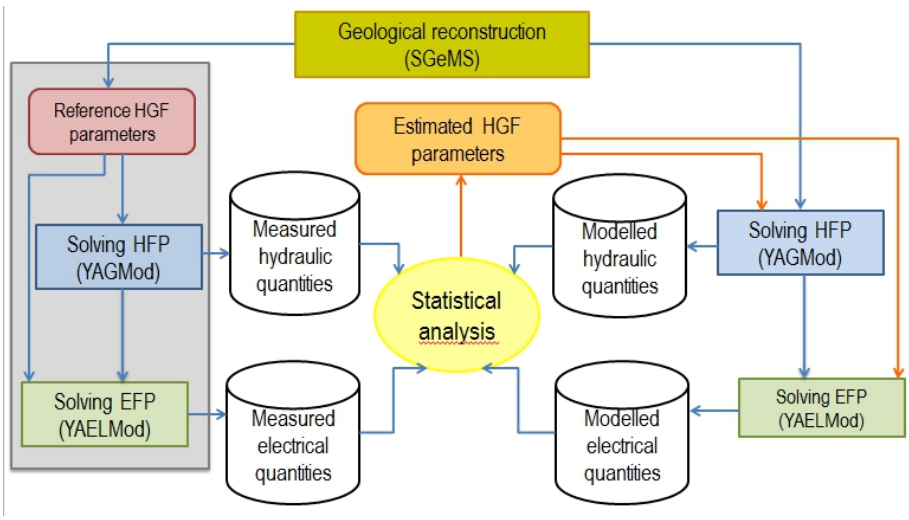


Figure 1.6: Conceptual scheme of the methodology adopted in the original procedure. HFP and EFP correspond to, respectively, hydraulic and electrical forward problems. The grey block shows the steps that are used in this work to generate synthetic data in order to perform several tests of the proposed method (see section 3.4.1).

1. **Geological Reconstruction**
   The 3D spatial distribution of HGFs, $\varphi(\mathbf{x})$, can be obtained from a collection of lithological data (stratigraphic logs and pumping tests) with the use of geostatistical simulation methods, that yield an ensemble of equiprobable distributions of HGFs. In particular, the study volume is subdivided in cells, each of which is associated with an HGF. This step is described in section 2.2.

2. **Definition of physical parameters for HGFs**
   From the literature, some formulas to compute hydraulic conductivity and electrical resistivity of the HGFs are selected and the corresponding "hydrogeological"

and "geoelectrical" phenomenological parameters are defined. They are cast in two arrays, $\mathbf{m}^{(K)}$ and $\mathbf{m}^{(\rho)}$. In principle, these parameters depend both on the HGF and on some physical properties, e.g., pore water saturation. The 3D distribution of physical parameters, namely hydraulic conductivity K and electrical resistivity ρ, is obtained from $\varphi(\mathbf{x})$ with the application of the empirical formulas mentioned above, so that formally one can write

$$\begin{aligned} K(\mathbf{x}, t_{\mathsf{k}}) &= \varkappa\left(\mathbf{m}^{(K)}, \varphi(\mathbf{x}), S(\mathbf{x}, t_{\mathsf{k}})\right) \\ \rho(\mathbf{x}, t_{\mathsf{k}}) &= \varrho\left(\mathbf{m}^{(\rho)}, \varphi(\mathbf{x}), S(\mathbf{x}, t_{\mathsf{k}})\right), \end{aligned} \tag{1.44}$$

where $S(\mathbf{x}, t_{\mathsf{k}})$ is the saturation field at different times t_{k}.

3. **Field data**
   In this work it is assumed that field data correspond to measurements of hydraulic head at some piezometers and apparent resistivity from the surface; in particular, it is assumed that a set of data is collected for each groundwater flow situation corresponding to the stationary conditions established by water withdrawal with some pumping wells. In principle, more complex experimental setups could also be considered. The observed hydrological and electrical data are stored in two arrays,

   $$\left\{h^{(\text{meas})}(\mathbf{x}_m, t_{\mathsf{k}}), m = 1, \ldots, N_M^{(h)}\right\}$$

   and

   $$\left\{\rho_{app}^{(\text{meas})}(A_m, B_m, M_m, N_m, t_{\mathsf{k}}), m = 1, \ldots, N_M^{(\rho)}\right\},$$

   where $N_M^{(h)}$ and $N_M^{(\rho)}$ denote the number of measurements of, respectively, hydraulic head and apparent resistivity. These arrays are merged in a single array $\hat{\mathbf{Y}}^{(\mathsf{k})}$. The index k is used to identify each experiment, which corresponds to different stress conditions at time t_{k}.

   For the PhD research, no field data were available to test the procedure, but synthetic data have been generated, in order to have a full control of the results (see section 3.4.1), with the steps which are sketched in the grey box of figure 1.6. In particular, the site geometry and the experimental setup correspond to those of a real site in the Po plain (Northern Italy), whereas the f field has been obtained by a single stochastic simulation of facies distribution, starting from field lithostratigraphic logs from the study site. By choosing a given functional dependence of K and ρ upon φ and S (i.e., the functions $\varkappa$ and ϱ), the values of the reference "true" parameters $\mathbf{m}^{(K)}$ and $\mathbf{m}^{(\rho)}$ have been selected from literature data. The synthetic data are finally generated by the solution of hydraulic and electrical forward problems (`YAGMod` + `YAELMod` steps).

4. **External loop of the EnKF procedure**

   (a) **Initial distributions of parameters**
   N_R initial equiprobable sets of parameters $\mathbf{m}^{(K)}$ and $\mathbf{m}^{(\rho)}$ are created with a given probability distribution: $\mathbf{m}_q^{(K)}$ and $\mathbf{m}_q^{(\rho)}$ are collected in the array of model parameters $\mathbf{X}_q^{(0)}$, where $q = 1, \ldots, N_R$ is the code used to identify each parameter realization.

(b) **Internal loop** If N_S sets of data, corresponding to t_{k} are available, then the following steps are executed iteratively for $\mathsf{k} = 1, \ldots, N_S$.

i. **Predicting step**
The solution to N_R hydraulic and electrical forward problems yields the set of predicted hydraulic head and apparent resistivity values

$$\left\{ h^{(\mathrm{mod})}(\mathbf{x}_m, t_{\mathsf{k}}, q), m = 1, \ldots, N_M^{(h)} \right\}$$

and

$$\left\{ \rho_{app}^{(\mathrm{mod})}(A_m, B_m, M_m, N_m, t_{\mathsf{k}}, q), m = 1, \ldots, N_M^{(\rho)} \right\},$$

where $q = 1, \ldots, N_R$ is the code used to identify the individual realization. The two sets of predicted data are merged in the array of the estimated data $\mathbf{Y}_q^{(\mathsf{k})}$, which depends on the realization of model parameters (index q) and on time (index k).

ii. **Statistical Analysis**
With a procedure similar to the ensemble Kalman Filter, at every iterative step k, the observed values $\tilde{\mathbf{Y}}^{(\mathsf{k})}$ are used to obtain an estimation of $\mathbf{X}_q^{(\mathsf{k}+1)}$, a parameter set used to predict $\mathbf{Y}_q^{(\mathsf{k}+1)}$.

The external loop is repeated till the difference between the distributions $\mathbf{X}_q^0$ and $\mathbf{X}_q^{N_S}$, namely the parameter distributions before and after the inner loop, is satisfactory according to criteria which are described later (see sections 2.1.1 and 2.1.2).

CHAPTER 2

Methods

2.1 The Ensemble Kalman Filter Procedure

2.1.1 The Ensemble Kalman Filter for parameter identification

In this work the approach to non-linear state estimation known as the *ensemble Kalman filter* (EnKF) (Evensen, 2009; Gillijns et al., 2006) has been considered. The standard Kalman filter (Kalman, 1960) was developed for linear models. The EnKF, instead, is an extension and modification of the original version and is widely used if the models are non-linear and the initial states are uncertain.

The application of the EnKF in the novel procedure proposed in this work assumes that a sequence of steps is performed to use data that correspond to different times and possibly different physical situations. Each step is denoted with the index k (see point 3 of the list in section 1.4). The application of EnKF to inverse modelling requires that the "state variable" of the standard Kalman filter approach is substituted with the set of model parameters, i.e., an array of N_P values. At a given step k, an ensemble of N_R forecasted state estimates (prior estimate) is assumed to be available and each element of this ensemble is denoted as $\mathbf{X}_q^{(\mathsf{k})} \in \mathbb{R}^{N_P}$,

$$\mathbf{X}_q^{(\mathsf{k})} = \{X_{1q}^{(\mathsf{k})}, X_{2q}^{(\mathsf{k})}, \ldots, X_{N_P q}^{(\mathsf{k})}\}, \tag{2.1}$$

where the index $q = 1, \ldots, N_R$ is used to identify the realizations of model parameters.

Then, the ensemble mean of a single model parameter is defined as

$$< X_l^{(\mathsf{k})} >= \frac{1}{N_R} \sum_{q=1}^{N_R} X_{lq}^{(\mathsf{k})}, \tag{2.2}$$

with $l = 1, \ldots, N_P$.

For every iterative step k, the measurements of hydraulic head and apparent resistivity are collected in the array $\tilde{\mathbf{Y}}^{(\mathsf{k})} \in \mathbb{R}^{N_M}$, where $N_M = N_M^{(h)} + N_M^{(\rho)}$ is the total number of measured quantities.

At every iterative step, for each realization of model parameters, $\mathbf{X}_q^{(\mathsf{k})}$, the combined hydraulic and electrical forward problems are solved to obtain, under determined stress and boundary conditions, the array containing the predicted variables $\mathbf{Y}_q^{(\mathsf{k})} \in \mathbb{R}^{N_M}$

$$\mathbf{Y}_q^{(\mathsf{k})} = \{Y_{1q}^{(\mathsf{k})}, Y_{2q}^{(\mathsf{k})}, \ldots, Y_{N_M q}^{(\mathsf{k})}\}, \tag{2.3}$$

where the index $q = 1, \ldots, N_R$ identifies the realization of model parameters which is used to compute the modelled values.

The mean of the predicted quantities is trivially defined as

$$< Y_m^{(\mathsf{k})} >= \frac{1}{N_R} \sum_{q=1}^{N_R} Y_{mq}^{(\mathsf{k})}, \tag{2.4}$$

with $m = 1, \ldots, N_M$.

Posterior estimates $\mathbf{X}_q^{(\mathsf{k}+1)}$ of the model parameters are obtained using the available measurements $\tilde{\mathbf{Y}}^{(\mathsf{k})}$: at each iterative step, the parameter distribution $\mathbf{X}_q^{(\mathsf{k})}$ is updated using a *corrective term*, as follows

$$\mathbf{X}_q^{(\mathsf{k}+1)} = \mathbf{X}_q^{(\mathsf{k})} + G^{(\mathsf{k})} \left(\tilde{\mathbf{Y}}^{(\mathsf{k})} - \mathbf{Y}_q^{(\mathsf{k})} \right). \tag{2.5}$$

For the standard Kalman filter, the matrix G, appearing in (2.5) is called the gain matrix. Its expression for the EnKF is

$$G^{(\mathsf{k})} = \mathrm{Cov}^{(\mathsf{k})}[XY] \cdot \mathrm{Cov}^{(\mathsf{k})}[YY]^{-1}, \tag{2.6}$$

where $\mathrm{Cov}^{(\mathsf{k})}[XY]$ and $\mathrm{Cov}^{(\mathsf{k})}[YY]$ are covariance matrices, whose elements are computed as

$$\begin{aligned} \mathrm{Cov}^{(\mathsf{k})}[XY]_{lj} &= \frac{1}{N_R - 1} \sum_{q=1}^{N_R} \left(X_{lq}^{(\mathsf{k})} - < X_l^{(\mathsf{k})} >) \right) \left(Y_{jq}^{(\mathsf{k})} - < Y_j^{(\mathsf{k})} >) \right) \\ \mathrm{Cov}^{(\mathsf{k})}[YY]_{ij} &= \frac{1}{N_R - 1} \sum_{q=1}^{N_R} \left(Y_{iq}^{(\mathsf{k})} - < X_i^{(\mathsf{k})} >) \right) \left(Y_{jq}^{(\mathsf{k})} - < Y_j^{(\mathsf{k})} >) \right), \\ l &= 1, \ldots, N_P \text{ and } i, j = 1, \ldots, N_M \end{aligned} \tag{2.7}$$

The procedure is repeated iteratively, so that a succession of parameter distributions $\mathbf{X}_q^{(0)}, \mathbf{X}_q^{(1)}, \ldots, \mathbf{X}_q^{(N_S)}$ is obtained. This is the inner loop described in section 1.4. Such a loop is repeated again and the procedure ends when the following condition is verified:

$$\max_{l=1,\ldots,N_P} \left| \frac{< X_l^{(N_S)} > - < X_l^{(0)} >}{\sqrt{\frac{\left(\sigma_l^{N_S}\right)^2}{N_R} + \frac{\left(\sigma_l^0\right)^2}{N_R}}} \right| < 0.5, \tag{2.8}$$

where

$$\left(\sigma_l^{\mathsf{k}}\right)^2 = \frac{1}{N_R - 1} \sum_{q=1}^{N_R} \left(X_{lq}^{(\mathsf{k})} - < X_l^{(\mathsf{k})} >) \right)^2,$$

for $\mathsf{k} = 1, \ldots, N_S$.

2.1.2 The simplified EnKF Procedure

Many tests performed for this PhD research have shown that the matrix $\mathrm{Cov}^{\mathsf{k}}[YY]$ of equation (2.7) is singular or badly conditioned (Rencher and Christensen, 2012)[1]. This

[1] an example is shown in section 3.4.2; Table 3.4 lists some numerical values that explicitly illustrate this issue

means that the inversion of that matrix and therefore the computation of the matrix G with equation (2.6) are not stable and therefore do not provide reliable results. For this reason a *simplified EnKF procedure* has been developed in this work and is described below.

The definition of G in the standard Kalman filter minimizes the sum of the variances of the updated errors on the system state. In other words the objective is to reduce the differences between observed values and predicted values. With the EnKF, this is done by the computation of the G matrix with equation (2.6), which cannot be applied in a stable way in the specific application considered in this research work. Then a different, a simpler approach is proposed to modify the input parameters in order to reduce the differences between modelled and measured quantities.

At iterative step k, a *total error* generated by every distribution is calculated as

$$\begin{aligned}
\epsilon_q^{(h)} &= \sum_{i=m}^{N_M^{(h)}} \left[\frac{\left(\tilde{Y}_{(h)}^{(\mathsf{k})}\right)_m - \left(Y_{q,(h)}^{(\mathsf{k})}\right)_m}{\max_m \left(\tilde{Y}_{(h)}^{(\mathsf{k})}\right)_m - \min_m \left(\tilde{Y}_{(h)}^{(\mathsf{k})}\right)_m} \right]^2 \\
\epsilon_q^{(\rho)} &= \sum_{m=1}^{N_M^{(\rho)}} \left[\frac{\left(\tilde{Y}_{(\rho)}^{(\mathsf{k})}\right)_m - \left(Y_{q,(\rho)}^{(\mathsf{k})}\right)_m}{\tilde{Y}_{(\rho)}^{(\mathsf{k})}{}_m} \right]^2 \\
\epsilon_q^{(\mathrm{tot})} &= \frac{\epsilon_q^{(h)}}{N_M^{(h)}} + \frac{\epsilon_q^{(\rho)}}{N_M^{(\rho)}},
\end{aligned} \tag{2.9}$$

where the measurements have been subdivided between the $N_M^{(h)}$ measurements of hydraulic head, $\tilde{Y}_{(h)}^{(\mathsf{k})}$, and the $N_M^{(\rho)}$ measurements of apparent resistivity, $\tilde{Y}_{(\rho)}^{(\mathsf{k})}$, in order to normalize the errors calculated on different physical quantities, remembering that

$$N_M^{(h)} + N_M^{(\rho)} = N_M. \tag{2.10}$$

Notice that the normalization introduced in the definitions of $\epsilon_q^{(h)}$ and $\epsilon_q^{(\rho)}$ differ from each other, because it is necessary to take into account the physical meanings of the measured quantities. In fact, resistivity is a positive physical parameter, which might vary over several orders of magnitude; therefore relative variations are significant. On the other hand, hydraulic head is a potential for which only variations from a point to another are physically significant, whereas the value could be changed arbitrarily, simply by choosing a different reference state. Therefore for hydraulic head, the differences between modelled and measured values are normalized with respect to the differences between maximal and minimal values measured in the study area.

In this way every distribution $\mathbf{X}_q^{(\mathsf{k})}$ is characterized by a corresponding total error $\epsilon_q^{(\mathrm{tot})}$. The elements $\mathbf{X}_q^{(\mathsf{k})}$, $q = 1, \ldots, N_R$, can be sorted according to the increasing value of the total error. The realizations belonging to the fourth quartile, i.e., those that yield the worst values of $\epsilon_q^{(\mathrm{tot})}$, are erased from the ensemble and substituted with new realizations calculated with an appropriate mean (arithmetic or geometric) between the realization taken from the first quartile and a realization randomly selected from those of the second and third quartiles.

An example of this corrective method for 8 realizations is sketched in Fig. 2.1. To

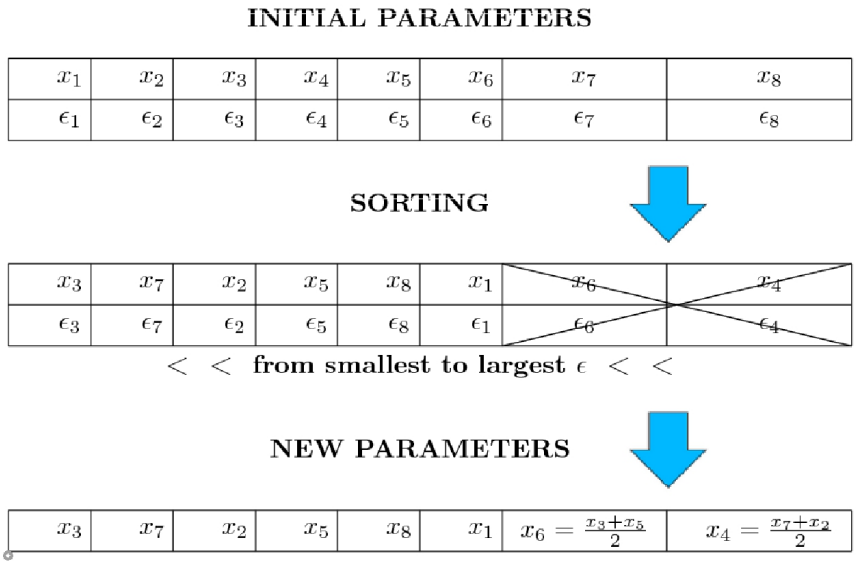


Figure 2.1: Scheme of the simplified EnKF procedure.

put into practice this procedure, it is necessary to have an estimation of the spatial distribution fo HGFs, to create an initial distribution of parameters and then to solve the hydraulic and the electric forward problem. In following sections all methods used to carry out these goals are examined separately.

2.2 The stochastic simulation of HGFs

The geostatistical simulation method used to estimate the HGFs distribution is the Sequential Indicator Simulation (SISIM) (Deutsch and Journel, 1992), a conditioned non gaussian simulation which is widely used. In SISIM the starting data are the codes of the "facies" that are found at different positions, as it is possible to obtain, for instance, from the lithological log of a well. In the case which is studied in this thesis, a code is assigned to each HGF and a variogram model is estimated for the indicator function of each of them[2]. SISIM permits to honour both the proportions among different HGFs inferred from field data and the position in space of the observed HGF. The analysis of the spatial correlation of every HGF along the three cartesian directions is performed through the computation of experimental indicator variograms and the successive fitting with theoretical models, i.e., with simple functions that can describe the variation of the variogram as a function of the space lag, and the calculus of relative proportions between the HGFs identified by collected data.

The experimental variogram is a geostatistic tool used to evaluate the spatial autocorrelation of observed data, through the semivariogram function $\gamma(\mathbf{d})$ (Matheron, 1965; de Marsily, 1986):

$$\gamma(\mathbf{d}) = \frac{1}{2n(\mathbf{d})} \cdot \sum_{i=1}^{2n(\mathbf{d})} \left[I(\mathbf{x}_i + \mathbf{d}) - I(\mathbf{x}_i)\right]^2 \tag{2.11}$$

[2]The indicator function of a given facies is equal to 1 if the point is occupied by that facies and to 0 otherwise.

where I are the observed values of the indicator function, $\mathbf{d}$ is the distance between two measurement points (lag) and $n(\mathbf{d})$ is the number of couples of observed points which are spaced by a distance $\mathbf{d}$, within a prescribed tolerance. If the semivariogram depends only on the absolute value of the lag and is independent from the direction, the function is isotropic. The indicator variogram function quantifies the spatial correlation of a variable and allow to define a distance over which there is no more correlation. In an experimental variogram the observed values near the origin typically show a strong correlation with distance, till $|\mathbf{d}|$ exceeds the *range* value, over which the data are organized around a mean value of $\gamma(|\mathbf{d}|)$ called *sill*, which correspond to experimental data variance. The range is the most important parameter of the experimental variogram, because it represents the correlation zone of a measured point: at distance greater than range value, the correlation between two measured points is null.

Simulating the HGFs distribution with SISIM is conducted with the software SGeMS (Stanford Geostatistical Modeling Software) which provides simple interfaces and tools to compute the experimental variograms, fit the model variograms and perform geostatistical conditional simulations (Remy et al., 2011).

2.3 Solving the hydraulic forward problem: the `YAGMod` code

One of the most challenging problems encountered while modelling groundwater flow processes is the occurrence of a dry cell or element, under the influence of an extraction source term. A thorough and fully rigorous solution would require the modelling of variably saturated groundwater flow, but this needs the knowledge of the non-linear relationships between conductivity and matric potential with soil water content for all the lithologies found in the subsoil. Therefore, approximated approaches, which introduce relatively simple modifications of the classical equations for saturated groundwater flow, are often applied.

When a cell becomes dry, i.e., its calculated water level falls below the bottom of the cell, two main problems arise (Doherty, 2001; Keating and Zyvoloski, 2009; Bedekar et al., 2012; Niswonger et al., 2011). First, if the dry cell is declared inactive, it can not receive external water, neither can contain any extraction source term, unless the water level rises above a prescribed value, so that the cell could be considered "re-wetted". Second, drying and re-wetting functionality often yields difficulties for the convergence of iterative algorithms used for the solution of the algebraic equations of the discrete model. Doherty (2001) proposes an asymptotically small transmissivity to avoid dry cells to be deactivated, even if they actually become dry: this approach uses a function that prevents cell transmissivity from becoming negative. Keating and Zyvoloski (2009)'s innovative idea is a weak scaling for vertical connectivity, from partially saturated to dry conditions. On the other hand, Niswonger et al. (2011) use a quadratic approximation of the function that relates horizontal conductance to hydraulic head, over small intervals close to the fully-dry and fully-saturated limits.

Within this background, the model YAGMod is developed and implemented in the `YAGMod` code, in Fortran90 programming language, for the simulation of constant density, groundwater flow under stationary conditions.

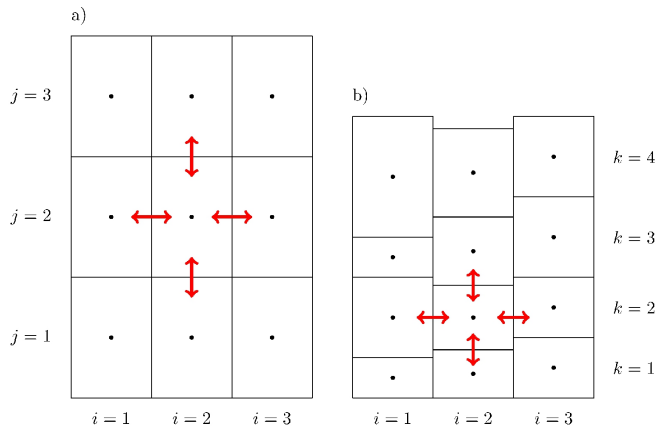


Figure 2.2: YAGMod: a) Plan view of a domain's layer, b) Vertical section view. Red arrows are examples of groundwater fluxes considered in the discrete model.

2.3.1 Mathematical Model and Discretization

Equation (1.11) can be written as

$$\frac{\partial}{\partial x}\left(K\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(K\frac{\partial h}{\partial y}\right) + \frac{\partial}{\partial z}\left(K\frac{\partial h}{\partial z}\right) + f = 0. \tag{2.12}$$

The numerical solution of (2.12) is found with the finite difference method. The continuous physical system is replaced by a finite set of cells or blocks, which are identified by three integer indices (i, j, k) ($1 \leq i \leq Nx$, $1 \leq j \leq Ny$, $1 \leq k \leq Nz$) and could be rectangular in the horizontal plane (the side lengths of the cells along x and y directions, Δx and Δy, are assumed to be constant for all the grid), but vertically distorted, with a cell thickness denoted by $\Delta z_{(i,j,k)}$. The centre of a cell is called a node and is denoted with the same indices as the corresponding cell. Values of hydraulic head are assigned at each node and the spatially varying hydraulic conductivities are considered to be effective parameters of a cell.

The saturated thickness of a cell is given by

$$\vartheta_{(i,j,k)} = \begin{cases} \Delta z_{(i,j,k)} & \text{if} \quad h_{(i,j,k)} > top_{(i,j,k)}, \\ h_{(i,j,k)} - bot_{(i,j,k)} & \text{if} \quad bot_{(i,j,k)} < h_{(i,j,k)} \leq top_{(i,j,k)}, \\ 0 & \text{if} \quad h_{(i,j,k)} \leq bot_{(i,j,k)}, \end{cases} \tag{2.13}$$

where $top_{(i,j,k)}$ and $bot_{(i,j,k)}$ represent, respectively the height of the top and bottom surfaces of a cell. In other words, $\vartheta_{(i,j,k)}$ corresponds to the thickness of the cell $\Delta z_{(i,j,k)} = top_{(i,j,k)} - bot_{(i,j,k)}$ for a fully saturated cell; it is calculated as the difference between the hydraulic head and the bottom of the cell, in case of partially saturated cell; it vanishes when the cell is dry.

For each cell an integral balance equation can be written. The water discharge into or from a cell is calculated considering only the six adjacent (first-neighbourhood) cells (see figure 2.2).

In order to handle complex aquifers' geometries, each cell is identified by a domain code (see figure 2.3):

E	E	D	D	D	D	D	D	E	E
E	I	I	I	I	I	I	I	I	E
I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I
I	I	I	I	I	I	I	I	I	I
E	E	E	I	I	I	I	I	I	E
E	E	E	D	D	D	D	D	D	E

Figure 2.3: Domain code example in 2D: `I`, `E` and `D` codes correspond, respectively, to internal, external and prescribed-head (Dirichlet boundary conditions) cells; the blue line denotes the border of the domain.

- `I` identifies the internal cells, for which the hydraulic head can vary freely;
- `D` identifies the cells with Dirichlet conditions where the hydraulic head is prescribed;
- `E` identifies the cells external to the domain or where no flow takes place.

The integral version of equation (2.12) could be discretized for each cell as

$$
\begin{aligned}
& K_{(i+1/2,j,k)} \frac{h_{(i+1,j,k)} - h_{(i,j,k)}}{\Delta x} \Delta y \, \vartheta_{(i+1/2,j,k)} \quad + \\
& + K_{(i-1/2,j,k)} \frac{h_{(i-1,j,k)} - h_{(i,j,k)}}{\Delta x} \Delta y \, \vartheta_{(i-1/2,j,k)} \quad + \\
& + K_{(i,j+1/2,k)} \frac{h_{(i,j+1,k)} - h_{(i,j,k)}}{\Delta y} \Delta x \, \vartheta_{(i,j+1/2,k)} \quad + \\
& + K_{(i,j-1/2,k)} \frac{h_{(i,j-1,k)} - h_{(i,j,k)}}{\Delta y} \Delta x \, \vartheta_{(i,j-1/2,k)} \quad + \\
& + K_{(i,j,k+1/2)} \frac{h_{(i,j,k+1)} - h_{(i,j,k)}}{\Delta z_{(i,j,k+1/2)}} \Delta x \, \Delta y \quad + \\
& + K_{(i,j,k-1/2)} \frac{h_{(i,j,k-1)} - h_{(i,j,k)}}{\Delta z_{(i,j,k-1/2)}} \Delta x \, \Delta y \quad + \quad F_{(i,j,k)} = 0
\end{aligned}
\tag{2.14}
$$

where:

- $h_{(i,j,k)}$ [L] is the hydraulic head at the node (i, j, k);
- $K_{(i-1/2,j,k)}$ and $K_{(i+1/2,j,k)}$ $\left[\mathrm{LT}^{-1}\right]$ are called internode (or interblock) hydraulic conductivities along the x direction (analogous definitions are used for the similar terms along the y and z directions);

- $\vartheta_{(i+1/2,j,k)} = \left(\vartheta_{(i,j,k)} + \vartheta_{(i+1,j,k)}\right)/2$ [L] is the arithmetic mean of the saturated thickness of the cells (i,j,k) and $(i+1,j,k)$ (an analogous definition is used for cells along the y direction) ;
- $\Delta z_{(i,j,k+1/2)} = \left(\Delta z_{(i,j,k)} + \Delta z_{(i+1,j,k)}\right)/2$ [L] is the distance between two adjacent nodes along the vertical;
- $F_{(i,j,k)}$ $\left[\mathrm{L}^3\mathrm{T}^{-1}\right]$ is the cell source term, i.e., the volume of water injected (negative if extracted) per unit time from the cell.

Each of the nine terms appearing in the left-hand-side of (2.14) represents the water flux through the interface separating two cells.

A single value of hydraulic conductivity is assigned to every cell of the domain and the internode hydraulic conductivity is calculated as the harmonic mean of the hydraulic conductivities of adjacent cells.

Equation (2.14) can be synthetically written for the most general case, by denoting the coefficients of the differences of hydraulic heads appearing in (2.14) as internode transmittances:

$$\begin{aligned}
\mathcal{T}_{(i+1/2,j,k)} &= K_{(i+1/2,j,k)}\vartheta_{(i+1/2,j,k)}\,\frac{\Delta y}{\Delta x},\\
\mathcal{T}_{(i,j+1/2,k)} &= K_{(i,j+1/2,k)}\vartheta_{(i,j+1/2,k)}\,\frac{\Delta x}{\Delta y},\\
\mathcal{T}_{(i,j,k+1/2)} &= K_{(i,j,k+1/2)}\,\frac{\Delta x\,\Delta y}{\Delta z_{(i,j,k+1/2)}}.
\end{aligned} \tag{2.15}$$

Then, equation (2.14) becomes

$$\begin{aligned}
&\mathcal{T}_{(i+1/2,j,k)}\left(h_{(i+1,j,k)} - h_{(i,j,k)}\right) + \mathcal{T}_{(i-1/2,j,k)}\left(h_{(i-1,j,k)} - h_{(i,j,k)}\right) +\\
+\;&\mathcal{T}_{(i,j+1/2,k)}\left(h_{(i,j+1,k)} - h_{(i,j,k)}\right) + \mathcal{T}_{(i,j-1/2,k)}\left(h_{(i,j-1,k)} - h_{(i,j,k)}\right) +\\
+\;&\mathcal{T}_{(i,j,k+1/2)}\left(h_{(i,j,k+1)} - h_{(i,j,k)}\right) + \mathcal{T}_{(i,j,k-1/2)}\left(h_{(i,j,k-1)} - h_{(i,j,k)}\right) =\\
=\;&-F_{(i,j,k)}.
\end{aligned} \tag{2.16}$$

2.3.2 Boundary conditions and source terms

Dirichlet, Neumann or Robin boundary conditions can be assigned. The cells, where Dirichlet boundary conditions (prescribed head) are assigned, are simply identified by using a `D` label for the domain code: in that case the hydraulic head does not change during the computation of the solution. Neumann and Robin boundary conditions are implemented as specific types of source terms.

Distributed Source/Sink Terms

This type of source or sink terms simulates spatially-distributed fixed source-terms, such as rainfall recharge. For each contribution an array of $Nx \times Ny \times Nz$ elements $F^{(\mathrm{d})}_{(i,j,k)}$ represents the flow rate into each cell of the domain; it is independent from $h_{(i,j,k)}$ and its dimensions are $[M^3T^{-1}]$. The user must consider that this type of source remains constant even if the hydraulic head of a single cell becomes lower than the bottom of the cell during the iterative search of a solution.

Local Source/Sink Terms

In the `YAGMod` code, local sources or sinks, i.e., those which are concentrated in a single cell, are modelled with the paradigmatic equation:

$$F^{(\mathrm{loc})} = \begin{cases} F_1 + \mathcal{K}_1 \cdot (h - \mathcal{H}^{(\mathrm{cal})}) & \text{if} \quad h \geq \mathcal{H}^{(\mathrm{act})}, \\ F_2 + \mathcal{K}_2 \cdot (h - \mathcal{H}^{(\mathrm{cal})}) & \text{if} \quad h < \mathcal{H}^{(\mathrm{act})}, \end{cases} \tag{2.17}$$

where $F^{(\mathrm{loc})}$ is the contribution that the individual source or sink gives to the source/sink term in the cell where it is located. $F^{(\mathrm{loc})}$ depends on the hydraulic head in that cell, h, according to the difference $(h - \mathcal{H}^{(\mathrm{act})})$. F_1 and F_2 are fixed fluxes $[M^3T^{-1}]$, $\mathcal{K}_1$ and $\mathcal{K}_2$ are conductances $[M^2T^{-1}]$: these parameters could be different for each source or sink. $\mathcal{H}^{(\mathrm{cal})}$ and $\mathcal{H}^{(\mathrm{act})}$ are two reference head values, which can vary for each source/sink: $\mathcal{H}^{(\mathrm{act})}$ is a threshold which establishes if a source or sink is active or which couple of fluxes and conductances, $(F_1, \mathcal{K}_1)$ or $(F_2, \mathcal{K}_2)$, should be used to compute $F^{(\mathrm{loc})}$; the $\mathcal{H}^{(\mathrm{cal})}$ threshold, which in many cases could be equal to $\mathcal{H}^{(\mathrm{act})}$, is used to compute the contribution to the source term, which linearly depends on h.

Different combinations of F_1, F_2, $\mathcal{K}_1$, $\mathcal{K}_2$, $\mathcal{H}^{(\mathrm{act})}$ and $\mathcal{H}^{(\mathrm{cal})}$ allow the user to generate a great variety of source terms, some of which are listed and shortly described below.

- Drain
  $F_1 = F_2 = 0$, $\mathcal{K}_2 = 0$; $\mathcal{K}_1$ represents the drain conductance and must be negative to obtain the outgoing flux of the cell; $\mathcal{H}^{(\mathrm{act})} = \mathcal{H}^{(\mathrm{cal})}$ represents the drain elevation.

- Robin boundary conditions
  These conditions can be used if the aquifer interacts with another water body and water exchange is controlled by the difference of hydraulic head in the aquifer and in the external water body. They are introduced through (2.17), by assigning the following parameters: $F_1 = F_2 = 0$; $\mathcal{H}^{(\mathrm{act})} = \mathcal{H}^{(\mathrm{cal})}$ are the reference hydraulic heads; $\mathcal{K}_1$ and $\mathcal{K}_2$ represent the conductances for flux out or into the cell, must be negative and depend upon the conductivity of the materials that separate the aquifer from the water body at the reference hydraulic head and upon the distance from this water body. Notice that for the simulation of limited domains of aquifers with a large extension, it is usually impossible to prescribe physically based boundary conditions. In those cases, Robin boundary conditions are very useful to introduce fictitious boundary conditions, which are more flexible than prescribed head (Dirichlet) or flux (Neumann) boundary conditions. In these situations, $\mathcal{H}^{(\mathrm{act})} = \mathcal{H}^{(\mathrm{cal})}$ should be close to the estimated hydraulic head far from the aquifer system and the conductances, $\mathcal{K}_1$ and $\mathcal{K}_2$, could assume different values, in order to keep into account the geometry or the geological and hydrological characteristics of the aquifer.

- River/aquifer interaction
  $\mathcal{H}^{(\mathrm{act})}$ is the height of the bottom of the river: therefore, if $h \geq \mathcal{H}^{(\mathrm{act})}$ the river and groundwater are in contact, whereas, if $h < \mathcal{H}^{(\mathrm{act})}$ then they are separated by a vadose zone, i.e., partly saturated sediments or rocks. In the first situation, $\mathcal{H}^{(\mathrm{cal})}$ is the water head in the river, so that the river is draining the aquifer if $h > \mathcal{H}^{(\mathrm{cal})}$ and it is recharging the aquifer if $h < \mathcal{H}^{(\mathrm{cal})}$; it is quite common to assume $F_1 = 0$ and $\mathcal{K}_1$ depends on the conductivity of the river bed sediments, their thickness and the area of the contact surface between the river bed and the

aquifer in the considered cell. In the second situation, the river bed is assumed to be composed of fine-grained materials, which could be almost saturated but poorly permeable, whereas the vadose zone between the river bed and the water table could be more permeable than the river bed sediments and approximated as dry. Therefore, the water flows through the river bed under a gravity controlled, unit hydraulic gradient, and freely flows through the relatively permeable vadose zone: then $\mathcal{K}_2 = 0$, whereas F_2 depends upon the conductivity, thickness and extension of the river bed sediments in the considered cell and on the river water head.

Screened Wells

YAGMod considers a new source type that permits to simulate wells, whose extraction rate depends on the aquifer water head, and, in particular, to take into account that no water can be pumped out if a cell becomes dry. Sources in this category are denoted as "screened wells", as the user has to give as input data not only the (x, y) coordinates of the well, i.e., the node indices i_W and j_W, but also the top and bottom elevation of the screened interval (top_W and bot_W) and the maximum well extraction rate, q_w.

The maximum extraction rate is subdivided among the cells occupied by screened intervals, as

$$q^{(\mathrm{scr})}_{(i_W,j_W,k)} = q_W \cdot \frac{K_{(i_W,j_W,k)}\,\Delta\mathcal{L}^{(\mathrm{scr})}_{(i_W,j_W,k)}}{\sum_{k'=n_{min}}^{n_{max}} K_{(i_W,j_W,k')}\,\Delta\mathcal{L}^{(\mathrm{scr})}_{(i_W,j_W,k')}}, \tag{2.18}$$

with $k = n_{min}, \ldots, n_{max}$, where: n_{min} and n_{max} are the indices along the vertical direction of the cells where are located the top and bottom of the screened interval of the well; $\Delta\mathcal{L}^{(\mathrm{scr})}_{(i_W,j_W,k)}$ is the screened thickness of the well corresponding to a fully saturated porous medium within the (i_W, j_W, k) cell and is computed as

$$\Delta\mathcal{L}^{(\mathrm{scr})}_{(i_W,j_W,k)} = \min\left(h_{(i_W,j_W,k)}, s^{top}_{(i_W,j_W,k)}\right) - \min\left(h_{(i_W,j_W,k)}, s^{bot}_{(i_W,j_W,k)}\right), \tag{2.19}$$

where

$$s^{top}_{(i_W,j_W,k)} = \begin{cases} top_{(i_W,j_W,k)} & k = n_{min}, \ldots, n_{max}-1, \\ top_W & k = n_{max}, \end{cases} \tag{2.20}$$

and

$$s^{bot}_{(i_W,j_W,k)} = \begin{cases} bot_{(i_W,j_W,k)} & k = n_{min}+1, \ldots, n_{max}, \\ bot_W & k = n_{min}. \end{cases} \tag{2.21}$$

Notice that if $h_{(i_W,j_W,k)} < s^{bot}_{(i_W,j_W,k)}$, then $\Delta\mathcal{L}^{(\mathrm{scr})}_{(i_W,j_W,k)} = 0$.

If the cell (i_W, j_W, k) desaturates, the value of $q^{(\mathrm{scr})}_{(i_W,j_W,k)}$ is corrected, so that the contribution of the well to the source term of the cell (i_W, j_W, k) is given by:

$$F^{(\mathrm{s})}_{(i_W,j_W,k)} = q^{(\mathrm{scr})}_{(i_W,j_W,k)} \cdot \frac{\sqrt{\Delta\mathcal{L}^{(\mathrm{scr})}_{(i_W,j_W,k)}}}{s^{top}_{(i_W,j_W,k)} - s^{bot}_{(i_W,j_W,k)}}. \tag{2.22}$$

The latter equation implies that the water extracted from a cell reduces as the square root of the thickness of the screened interval of the well that intersects a fully saturated portion of the aquifer in that cell.

2.3.3 Solution of the balance equations

Equation (2.16) can be written for each internal cell and therefore yields a system of possibly non-linear equations, that can be written in matrix formulation as:

$$\mathbf{A}\left(\mathbf{x}\right)\mathbf{x} = \mathbf{b}^{(\mathrm{fix})} + \mathbf{b}^{(\mathrm{var})}(\mathbf{x}), \tag{2.23}$$

where the array $\mathbf{x}$ includes the values of piezometric head in the internal nodes, the array $\mathbf{b}^{(\mathrm{fix})}$ includes the source/sink terms that are fixed (section 2.3.2), the array $\mathbf{b}^{(\mathrm{var})}$ includes the source/sink terms that depend on the hydraulic head of the aquifer (sections 2.3.2 and 2.3.2) and the terms appearing in the left hand side of (2.16) that involve the piezometric head of `D` nodes. $\mathbf{A}$ is a sparse, symmetric, diagonally dominant matrix, which is strictly diagonally dominant if at least one `D` node is present in the domain its elements are built with transmittances and therefore depends on $\mathbf{x}$, as shown by (2.15) and (2.13).

The solution to (2.23) could be obtained with any of the methods of solution for non-linear equations that can be found in textbooks of numerical analysis. Here a simple approach, based on a generalization of the relaxation methods (Young, 1971) for the solution of systems of algebraic linear equations, is proposed.

Starting from an initial guess $\mathbf{x}^{(0)}$, the approximated solution to (2.23) at a given iteration ℓ is obtained by:

- factorizing the coefficient matrix of 2.23 as $\mathbf{A}\left(\mathbf{x}^{(\ell-1)}\right) = \mathbf{D}\left(\mathbf{x}^{(\ell-1)}\right) - \mathbf{L}\left(\mathbf{x}^{(\ell-1)}\right) - \mathbf{U}\left(\mathbf{x}^{(\ell-1)}\right)$, where $\mathbf{D}$, $\mathbf{L}$ and $\mathbf{U} = \mathbf{L}^t$ are, respectively, diagonal, lower triangular and upper triangular matrices;

- computing $\mathbf{x}^\star$ with the Gauss-Seidel approach:

$$\begin{aligned}\mathbf{x}^\star \quad = \quad & \left[\mathbf{D}\left(\mathbf{x}^{(\ell-1)}\right) - \mathbf{L}\left(\mathbf{x}^{(\ell-1)}\right)\right]^{-1} \cdot \\ & \left[\mathbf{U}\left(\mathbf{x}^{(\ell-1)}\right)\left(\mathbf{x}^{(\ell-1)}\right) + \mathbf{b}^{(\mathrm{fix})} + \mathbf{b}^{(\mathrm{var})}\left(\mathbf{x}^{(\ell-1)}\right)\right];\end{aligned} \tag{2.24}$$

- relaxing the correction $\left(\mathbf{x}^\star - \mathbf{x}^{(\ell-1)}\right)$ with a parameter ω, $0 < \omega < 2$, and computing $\mathbf{x}^{(\ell)}$ as

$$\mathbf{x}^{(\ell)} = \mathbf{x}^{(\ell-1)} + \omega\left(\mathbf{x}^\star - \mathbf{x}^{(\ell-1)}\right) = \left(1-\omega\right)\mathbf{x}^{(\ell-1)} + \omega\mathbf{x}^\star;$$

- checking whether

$$\max_i \left|\mathbf{x}_i^{(\ell)} - \mathbf{x}_i^{(\ell-1)}\right| < \epsilon,$$

  where ϵ is a prescribed tolerance, to stop the iterative procedure.

Recall that over-relaxation is obtained for $\omega > 1$ and it accelerates convergence toward the correct solution for a linear problem, whereas under-relaxation, i.e., $\omega < 1$, is often more robust for non-linear problems, even if it reduces the speed of convergence.

The choice of this method of solution is not optimal from the computational point of view, because other, more sophisticated approaches could require less CPU time, but it proved to be more robust, in particular for complex situations, e.g., related to very heterogeneous media.

2.3.4 Check of physical consistency of the solutions

The proposed model does not solve equations for variably saturated conditions, but aims at finding a solution for fully saturated groundwater flow: the cells which become dry during the iterative algorithm of solution are not eliminated from the domain, but are used as "auxiliary" cells in the sense to be specified below.

If $h_{(i+1,j,k)} < bot_{(i+1,j,k)}$, then $\vartheta_{(i+1,j,k)} = 0$. If also the adjacent cells along the horizontal directions are dry, then the terms corresponding to horizontal fluxes in (2.16) vanish and therefore the cell under examination is involved only for a balance along the vertical direction. This choice permits to transfer the fixed source terms to deeper cells: this is necessary, e.g., to permit to the aquifer recharge, which is assigned at the top layer of cells, to reach the water table. Instead, if the adjacent cells have a non vanishing, possibly small, thickness, then the physical situation implies that there is an horizontal transfer of water.

When the solution procedure has reached convergence criterion, a recursive function checks for every continuous path connecting partially or totally desaturated cells with the top layer. At the end of the checking subroutine run, every totally or partially desaturated cell need to be connected with the surface, in order to allow air to infiltrate into the porous media. A warning message is given by YAGMod if this conditions is not satisfied for every cell, so that the solution is not physically acceptable.

2.3.5 A simple test of YAGMod

Different approaches proposed to manage dry cells use different ways to calculate internode conductivities and, in some cases, also effective extraction rates, depending on saturated thickness (Doherty, 2001; Keating and Zyvoloski, 2009; Niswonger et al., 2011). The algorithm implemented in YAGMod is now compared with those models by means of a very simple test, which nevertheless permits to emphasize some properties of the different methods. In particular the basic characteristics of the analysed algorithms are briefly recalled using a simplified notation based on this example.

A simple two-dimensional domain has been constructed with a grid of $3 \times 1 \times 2$ cells whose size is $100\ \mathrm{m} \times 100\ \mathrm{m} \times 20\ \mathrm{m}$. This two-dimensional domain is illustrated in Fig. 2.4, together with the cell numbering which is used in the following for the sake of simplicity. At cells (1), (3), (4) and (6) hydraulic head is prescribed in such a way as to generate an hydraulic gradient along the x direction: $h_1 = h_4 = 40\ \mathrm{m}$, $h_3 = h_6 = 39\ \mathrm{m}$. At cell (2) an extraction source term is assigned.

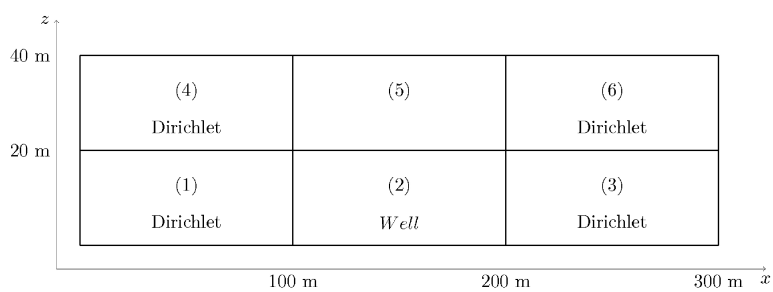


Figure 2.4: Scheme of the domain used for the comparison test.

The balance equation for the "free" cells (2) and (5) could be written as

$$\begin{aligned} K_{(1,2)}\vartheta_{(1,2)}\frac{h_{(1)}-h_{(2)}}{\Delta x}\Delta y &+ \\ K_{(3,2)}\vartheta_{(3,2)}\frac{h_{(3)}-h_{(2)}}{\Delta x}\Delta y &+ \\ K_{(5,2)}\frac{h_{(5)}-h_{(2)}}{\Delta z}\Delta x\Delta y &= -F_{(2)}, \end{aligned} \tag{2.25}$$

and

$$\begin{aligned} K_{(4,5)}\vartheta_{(4,5)}\frac{h_{(4)}-h_{(5)}}{\Delta x}\Delta y &+ \\ K_{(6,5)}\vartheta_{(6,5)}\frac{h_{(6)}-h_{(5)}}{\Delta x}\Delta y &+ \\ K_{(2,5)}\frac{h_{(2)}-h_{(5)}}{\Delta z}\Delta x\Delta y &= 0. \end{aligned} \tag{2.26}$$

In the numerical experiment conducted in this work, varying the values of h_2 and h_5 of equations (2.25) and (2.26) from minimum to maximum values, i.e. in the interval $[0, 40]\,m$, the balance errors ϵ_2 and ϵ_5 are calculated

$$\begin{aligned} \epsilon_2 = \; & K_{(1,2)}\vartheta_{(1,2)}\frac{h_{(1)}-h_{(2)}}{\Delta x}\Delta y + \\ & K_{(3,2)}\vartheta_{(3,2)}\frac{h_{(3)}-h_{(2)}}{\Delta x}\Delta y + \\ & K_{(5,2)}\frac{h_{(5)}-h_{(2)}}{\Delta z}\Delta x\Delta y + F_{(2)}, \end{aligned} \tag{2.27}$$

$$\begin{aligned} \epsilon_5 = \; & K_{(4,5)}\vartheta_{(4,5)}\frac{h_{(4)}-h_{(5)}}{\Delta x}\Delta y + \\ & K_{(6,5)}\vartheta_{(6,5)}\frac{h_{(6)}-h_{(5)}}{\Delta x}\Delta y + \\ & K_{(2,5)}\frac{h_{(2)}-h_{(5)}}{\Delta z}\Delta x\Delta y. \end{aligned} \tag{2.28}$$

The study of existence and uniqueness of the solution of the problem is based on the analysis of the total quadratic balance error

$$\epsilon_{\text{tot}}^2 = \epsilon_2^2 + \epsilon_5^2. \tag{2.29}$$

The method proposed by Doherty (2001) uses for the horizontal interblock transmissivity an asymptotically small transmissivity function, in order to keep every cell active even if it actually becomes dry. This approach uses a decay function that prevents the transmissivity of a dry cell from becoming non-positive:

$$T = \begin{cases} K\Theta_r e^{-g\Theta} + K\Theta & \text{if } \Theta > 0, \quad \text{(a)} \\ \Theta_r e^{f\Theta} & \text{if } \Theta < 0, \quad \text{(b)} \end{cases} \tag{2.30}$$

where: T is the transmissivity; K is the hydraulic conductivity of a cell; Θ is the saturated thickness, as for `YAGMod`; g and f, which are numerical parameters, and Θ_r, the residual saturated thickness, are parameters supplied by the user. To ensure that the function of equation (2.30) is continuous and continuously differentiable, the following relationship must be satisfied

$$g = \frac{1}{\Theta_r} - f, \tag{2.31}$$

so that the user must specify only two parameters, f and Θ_r. The transmissivity, calculated with equation (2.30) for every cell (i, j, k) of the domain, is used to calculate interblock transmissivity with harmonic average.

For the vertical water balance, Doherty (2001) considers that if any cell in the domain becomes dry, then water inputs from the upper layer remain active; to improve vertical water exchange with the lower layers, i.e. to permit that the water introduced in the model reaches deeper cells, a linear reduction of vertical interblock resistance (reciprocal of conductance) is introduced using the following equations

$$R_{(2,5)} = \begin{cases} R^{(\mathrm{u})}_{(2,5)} & \text{for } h_5 > h_u, \\ R^{(\mathrm{b})}_{(2,5)} + \dfrac{h_5 - bot_5}{h_u - bot_5}\left(R^{(\mathrm{u})}_{(2,5)} - R^{(\mathrm{b})}_{(2,5)}\right) & \text{for } h_u > h_5 > bot_5, \\ R^{(\mathrm{b})}_{(2,5)} & \text{for } h_5 \leq bot_5, \end{cases} \tag{2.32}$$

where

$R_{(2,5)}$ is the interblock resistance, reciprocal of conductance;

h_u is the water level below which the linear reduction of resistance is activated (supplied by the user);

$R^{(\mathrm{u})}_{(2,5)}$ is the "standard" interblock resistance given by:

$$R^{(\mathrm{u})}_{2,5} = \frac{1}{2}\left(\frac{\vartheta_{(2)}}{K_2} + \frac{\vartheta_5}{K_5}\right)\frac{1}{\Delta x \Delta y};$$

$R^{(\mathrm{b})}_{(2,5)}$ is the modified interblock resistance, calculated as the reciprocal of the "enhanced interblock conductance" $\mathcal{C}^{(\mathrm{b})}_{(2,5)}$, given by:

$$\mathcal{C}^{(\mathrm{b})}_{(2,5)} = m\mathcal{C}_{(2,5)} = \frac{m}{R^{(\mathrm{u})}_{(2,5)}},$$

where m is a user-supplied multiplier depending on simulation.

In Fig. 2.5, the results obtained for different values of m from 1 to 100 and $q = 0.1\ \mathrm{m^3\ s^{-1}}$, are plotted: no significant difference was noticed among the simulations, so that in further tests $m = 1$ was assigned.

The second investigated approach was proposed by Keating and Zyvoloski (2009). Horizontal interblock transmittance is calculated as follows

$$\begin{aligned} \mathcal{T}_{(1,2)} \quad = \quad & \frac{1}{2}\left(\frac{\vartheta_1}{top_1 - bot_1} + \frac{\vartheta_2}{top_2 - bot_2}\right) \cdot 2\frac{K_1 K_2}{K_1 + K_2} \\ & \cdot \frac{1}{2}\left(top_1 - bot_1 + top_2 - bot_2\right)\frac{\Delta y}{\Delta y}. \end{aligned} \tag{2.33}$$

For vertical transmittance, changing from intrinsic vertical internode conductivity in a partially saturated cell to 0 in a dry cell would lead to a discontinuity. To improve numerical stability, Keating and Zyvoloski (2009) allow a weak scaling controlled by a user-specified parameter, ξ

$$\begin{aligned} \mathcal{T}_{(2,5)} \quad = \quad & \left[\min\left(1, \xi\frac{\vartheta_2}{top_2 - bot_2}\right) + \min\left(1, \xi\frac{\vartheta_5}{top_5 - bot_5}\right)\right] \\ & \cdot 2\frac{K_2 K_5}{K_2 + K_5} \cdot 2\frac{\Delta x \Delta y}{top_5 - bot_2}. \end{aligned} \tag{2.34}$$

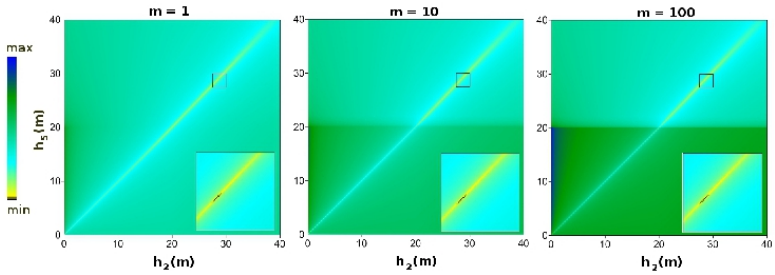


Figure 2.5: Map of the balance error given by equation (2.29), for the Doherty (2001) approach, as a function of the hydraulic heads h_2 (x axis) and h_5 (y axis), for different values of m (left, $m = 1$; centre, $m = 10$; right, $m = 100$). At the bottom right corner of each graph, the zone containing the minimum error value is enlarged.

Keating and Zyvoloski (2009) suggest to use a value of 10 for ξ, which provides both accuracy and stability for all problems.

The last approach considered for this comparison exercise is implemented in the UPW Package of the `MODFLOW NTW` model (Niswonger et al., 2011), a standalone version of MODFLOW 2005. The UPW Package smooths the horizontal conductance function during wetting and drying of a cell. Using this method, horizontal interblock conductance for this test is calculated as follows:

$$C_{(1,2)}^{NTW} = \begin{cases} \eta & \text{if } \chi \leq 0, \\ \alpha K_{(1,2)} \left[\dfrac{0.5\beta\chi^2}{\Omega} \right] & \text{if } 0 < \chi \leq \Omega, \\ \alpha K_{(1,2)} \left[AX + 0.5(1-\beta) \right] & \text{if } \Omega < \chi \leq (1-\Omega), \\ \alpha K_{(1,2)} \left[\dfrac{1 - 0.5\beta(1-\chi)^2}{\Omega} \right] & \text{if } (1-\Omega) < \chi \leq 1, \\ \alpha K_{(1,2)} & \text{if } \chi \geq 1, \end{cases} \tag{2.35}$$

where: $\chi = (h_{up} - bot_{up}) (top_{up} - bot_{up})^{-1}$; h_{up} is the maximum between h_1 and h_2; bot_{up} and top_{up} are respectively the bottom and top level of the cell corresponding to h_{up}; η is a small value, usually taken as $\eta = 1 \times 10^{-9}\ \mathrm{m^2\,s^{-1}}$; $\alpha = (top_{up} - bot_{up})\, \Delta y \Delta x^{-1}$; $K_{(1,2)}$ the internode conductivity calculated as $K_{(1,2)} = 2K_1K_2 (K_1 + K_2)^{-1}$; Ω is the smoothing interval, that is suggested to be very small (10^{-5}) and $\beta = \Omega(1-\Omega)^{-1}$.

In this approach vertical conductance is calculated as in standard MODFLOW 2005:

$$C_{(2,5)} = 2\Delta x \Delta y \frac{K_2 K_5}{(h_2 - bot_2)K_5 + (h_5 - bot_5)K_2}. \tag{2.36}$$

The pumping rate is reduced as the head in the cell drops below a user-specified percentage of the cell thickness, as

$$q^{NWT} = \begin{cases} 0 & \text{if } \delta \leq 0, \\ q\left(-2\delta^3 + 3\delta^2\right) & \text{if } 0 < \delta < 1/\Phi, \\ q & \text{if } \delta \geq 1/\Phi, \end{cases} \tag{2.37}$$

where

$$\delta = \frac{1}{\Phi} \frac{h - bot}{top - bot}$$

and Φ is a user specified fraction of the cell thickness, typically calculated as $\Phi = 0.25$.

The results obtained with the application of the four described approaches for three cases corresponding to an extraction rate varying from $0.1\ \mathrm{m^3\,s^{-1}}$ to $0.2\ \mathrm{m^3\,s^{-1}}$ and $0.3\ \mathrm{m^3\,s^{-1}}$ are represented in Fig. 2.6. Notice that with the approach by Keating and Zyvoloski (2009) h_5 cannot drop below the cell bottom: therefore, the results obtained with this method for extraction rates of $0.2\ \mathrm{m^3\,s^{-1}}$ and $0.3\ \mathrm{m^3\,s^{-1}}$ could not be significantly compared with those from other algorithms. Besides this remark, all the methods give realistic results, even if the values that yield the least total error balance for alternative algorithms differ from each other.

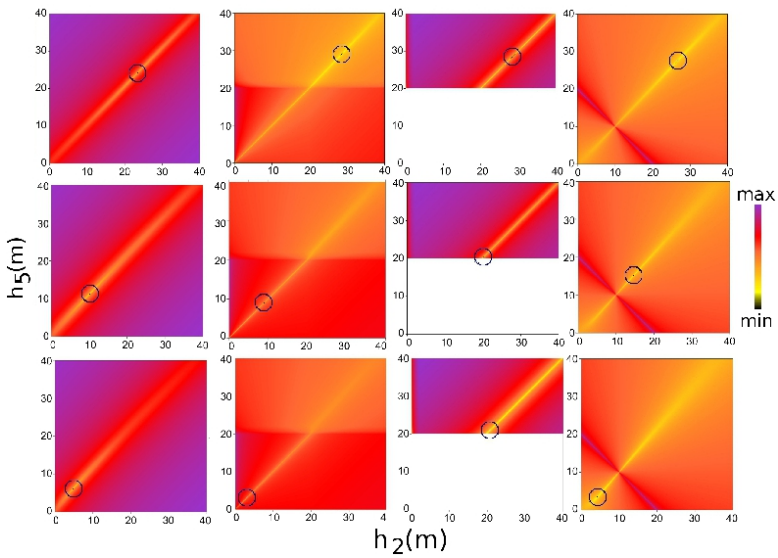


Figure 2.6: Map of the total quadratic balance error given by equation (2.29), as a function of the piezometric heads h_2 (x axis) and h_5 (y axis). From left to right, respectively the results obtained with YAGMod and the approaches by Doherty (2001), Keating and Zyvoloski (2009) and Niswonger et al. (2011). From top down, results obtained with extraction rates of $0.1\ \mathrm{m^3\,s^{-1}}$, $0.2\ \mathrm{m^3\,s^{-1}}$ and $0.3\ \mathrm{m^3\,s^{-1}}$. The blue circles point out the zone where the least value of the total quadratic balance error is located.

The hydraulic heads computed with YAGMod for an extraction rate of $0.1\ \mathrm{m^3\,s^{-1}}$ are smaller than those obtained with the other approaches: on the other hand, for higher extraction rates, the behaviour is more complex and no systematic difference is shown.

The colour scales of the plots of Fig. 2.6 are normalized with respect to the minimum and maximum total quadratic errors, separately for each method. Therefore, the images show that all the methods yield a single minimum and for YAGMod $\epsilon^2_{\mathrm{tot}}$ increases from the least value more rapidly than for other methods.

2.3.6 Application of YAGMod to real aquifers

In this section two applications of the YAGMod code to real aquifers are considered. The first one refers to a groundwater alluvial basin in East Africa, whose extension is about $20\ \mathrm{km^2}$ (regional scale), with the principal goal of understanding groundwater recharge mechanism (Cattaneo et al., 2013); the second one refers to a fractured and

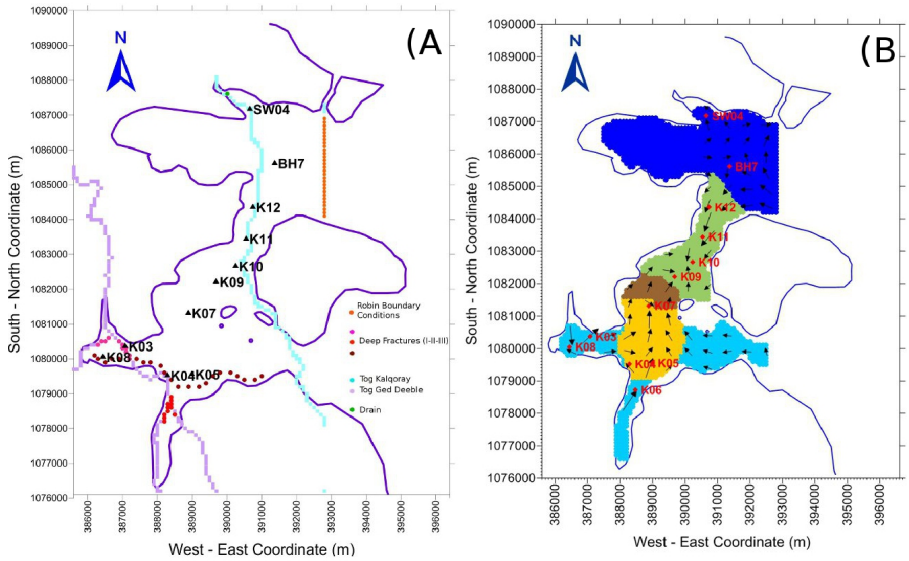


Figure 2.7: The GD basin model: Location of some source terms in the GD basin (left) and plot of the zones characterized by different values of hydraulic conductivity. See Table 2.1 for the values of K in each zone. Red diamonds are active wells. Black arrows represents flow lines

ZONE	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5
K (m/s)	$5.0 \cdot 10^{-4}$	$4.0 \cdot 10^{-4}$	$3.5 \cdot 10^{-4}$	$5.0 \cdot 10^{-5}$	$1.0 \cdot 10^{-5}$

Table 2.1: Hydraulic conductivity values used in the GD model.

karst, carbonate-rock aquifer in Southern Italy, whose extension is more than $200\ \mathrm{km}^2$ (large scale), in order to evaluate main feeding of fresh water and to determine the impact of pumping water on salt water intrusion De Filippis et al. (2013).

- The alluvial aquifer of the Ged Deeble (GD) basin supports the water demand of the city of Hargeisa (Somalia, East Africa). The `YAGMod` code has been used for the mathematical model, by considering a 2D hydraulic flow approximation, pseudo-steady conditions corresponding to the average annual flow, no-flow boundary conditions in correspondence of the crystalline bedrock and fixed head at the edge with the widespread and thick Laas Dhuurre-Damal (LDD) basin. The latter boundary condition is modelled as a Robin boundary condition, with two different conductance values for inflow and outflow water exchange: this choice permits to take into account the different geometries of the two connected basins, the smallest GD basin and the largest LDD basin. The interpretation of the data and the model calibration performed with YAGMod have given some important suggestions to describe the state of the GD basin. From the geological point of view, the basin has been divided in two sections, separated by an area of low permeability (see Fig 2.7 for a map of the hydraulic conductivity zones and Table 2.1 for corresponding values). From the hydrological point of view, recharge could not be limited to rain and wadis infiltration only: increasing extraction generates additional recharge sources, that in the upstream (southern) section probably comes

from an underground fracture-fault network, modelled as local-source terms depending on water level, whereas in the downstream (northern) section comes from the LDD basin, through Robin boundary conditions. Quantification of flux through the fracture networks has come from model calibration.

- The Salento peninsula (southern Italy), like other coastal areas, is subjected to the risk of desertification and a proper management of groundwater resources requires tools to analyse and predict the water balance and the evolution of the physical system in response to human activities (e.g., ground water withdrawals) and climatic factors. The Salento peninsula is a typical Mediterranean basin, where the main water resource is the aquifer hosted in Cretaceous carbonate-rocks (Calcare di Altamura, Altamura limestone): this is a fractured and karst aquifer, with a poor recharge and complex relationships with the sea. In this study, YAGMod was adapted to be used to solve the balance equation for a fluid moving in a saturated porous medium under pseudo-steady 2D flow conditions and when the bottom of the aquifer is characterized by the interface between fresh and salt water. In addition, for this application, the module that is included in YAGMod to perform model calibration with the Comparison Model Method (CMM) for 2D flow conditions has been applied. This method was originally proposed by Scarascia and Ponzini (1972), successively developed by Ponzini and Lozej (1982), cast in a more formal mathematical framework by Ponzini and Crosta (1988) and applied so far for 2D hydraulic flow regime in regional aquifers (Associazione Irrigazione Est Sesia, 1979; Beatrizotti et al., 1983; Benoit et al., 2005; Giudici et al., 2012; Vassena et al., 2008, 2011). The module that implements it in YAGMod is not described here, because it is not used in the mainstream of the PhD work.

  The model has been built up with 286×220 squared cells with $500\ \mathrm{m}$ side length. Dirichlet boundary conditions have been assigned along the boundary between the Salento peninsula and the Murge hills, by using the reference piezometric level. Along the coast, two types of boundary conditions have been assigned:

  1. where the aquifer is under phreatic conditions ($t > 0\ \mathrm{m}$, with t as the top of Altamura limestone with respect to the mean sea level), the outgoing flow from cells is modelled as a drain

     $$Q_{drain} = Ch, \tag{2.38}$$

     where C is an appropriate conductance;
  2. where the aquifer is confined near the coast ($t \leq 0\mathrm{m}$), the simple strategy of assign Dirichlet boundary conditions by using the reference piezometric level is chosen.

  Aquifer recharge is mostly due to rain infiltration, which has been estimated from meteorological and hydrostratigraphic data, and to infiltration through sinkholes and dolines, for which a constant flow rate ($10^{-3}\mathrm{m}^3\mathrm{s}^{-1}$) has been assigned. Groundwater extractions for agricultural purposes, categorized according to the land use over the whole area, have been estimated. The model has shown that the flow through the boundary between the Salento peninsula and the Murge hills is mainly incoming, while that through the coasts is generally outgoing. The incoming flow through the Adriatic coast is due to the presence of a south-eastern zone with high h values, while that through the Ionian coast is due to its complex geometry. Furthermore, the main recharge term of the aquifer is related to rain infiltration, while

the recharge due to sinkholes and dolines is of minor relevance. Finally, withdrawals are mainly due to agricultural uses: on the other hand, the extractions for industrial purposes mainly affect the shallow aquifer, which has not been modelled by De Filippis et al. (2013). Furthermore, the numerical model has correctly identified the areas where the aquifer is saturated with salt water and provided information about the global balance of the aquifer itself. Moreover, the model has been used to quantify the consequences that a decrease of the mean annual rainfall (natural change) or an increase of extractions for irrigation purposes (man induced change) are expected to have in terms of depletion and deterioration of groundwater quality. In both cases, the model has predicted a lowering of hydraulic head in the central part of the peninsula and an increase of the extension of the area where the aquifer is saturated with salt water (see Fig 2.8).

2.4 Solving the electrical forward problem: the `YAELMod` code

For the solution of the electrical forward problem, different published models have been evaluated, before taking the decision to develop a new original code. In particular the attention was addressed on the following models:

- RES3DMOD, developed by GEOTOMO Software Sdn. Bhd. (Geotomo Software, 2001; Locke, 2001);
- R3t, developed by Andrew Binley, Lancaster University (Binley, 2013a,b);
- RESINVM3D, developed by Adam Pidliseckj, Eldad Haber and Rosemary Knight, Stanford University (Pidlisecky et al., 2013, 2007).

RES3DMOD is a Windows based 3-D resistivity and IP modelling program that calculates the apparent resistivity values for a survey carried out with a rectangular grid of electrodes using the finite-difference (Dey and Morrison, 1979) or the finite-element (Silvester and Ferrari, 1990) methods. It is a robust and fast code, which is based on the finite-difference method and is widely used. Unfortunately, the source codes for this model can not be downloaded and it was very complex to extract resistivity values from the output files containing the solutions of the forward problem.

R3t is a forward/inverse solution for 3D current flow in a tetrahedral or triangular prism mesh. R3t stores calculated parameters (resistivity) for the entire mesh and the user must extract results for the region under study. However, this model works with finite element mesh, only with triangular based shapes. This complicated the joint modelling with the hydraulic model solver YAGMOD.

RESINVM3D is instead a 3D resistivity-inversion MATLAB-based package and the source codes are available to download. It uses a finite-difference method to solve differential equations, and the forward problem solver package is relatively easy to manage. Some simple tests have been performed to evaluate the efficiency of the forward problem solver. The numerical solution has been compared with the analytical solution obtained for three different types of resistivity domain: homogeneous, vertical discontinuity, horizontal discontinuity. The analytical solution was computed with the method of image charges (also known as the method of images and method of mirror charges; see, e.g., Feynman et al. (1989)) for a polar-dipole array, with a current dipole and a single potentiometric electrode (the other electrode is assumed to be at great distance). The comparison between the numerical and analytical solutions showed that the errors were quite great in the neighbourhood of the two current electrodes.

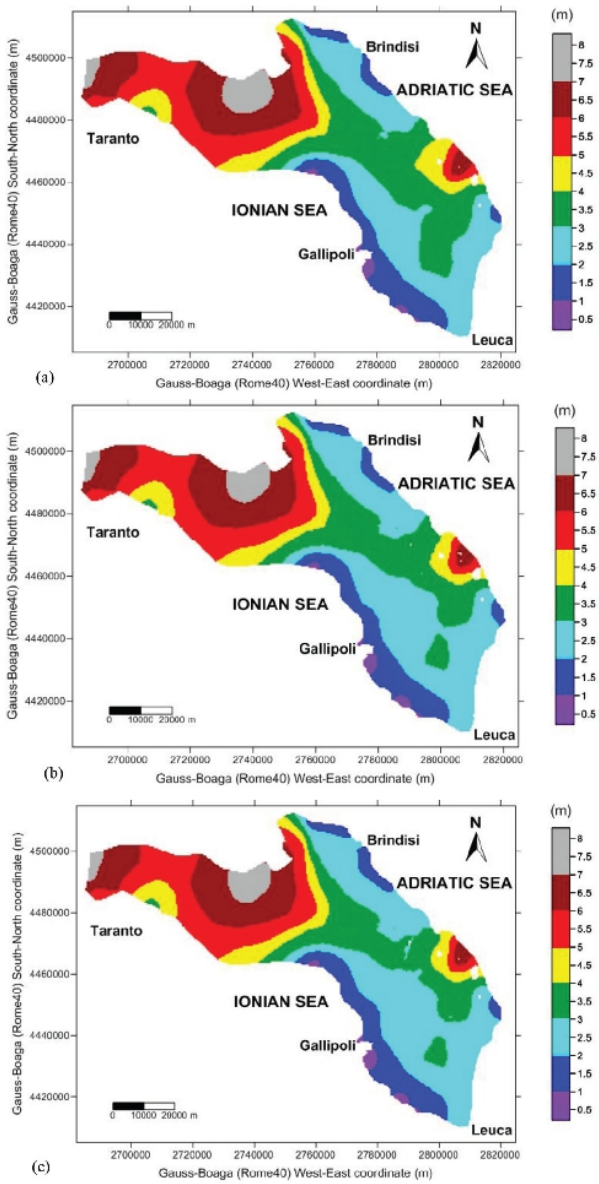


Figure 2.8: Comparison between (a) the piezometric head estimated for the heterogeneous medium obtained for the reference test, (b) that obtained by decreasing the inflow rate due to rain infiltration by 10% and (c) that obtained by increasing the pumping extraction rate due to agricultural use by 20%.

For the reasons just mentioned, developing a new code to simulate the geoelectrical field seemed to be unavoidable, in order to manage easily and efficiently the joint modelling of hydraulic and electric forward problems.

The starting point for this objective was a preliminary model developed for the simulation and the interpretation of geoelectrical data, acquired on heterogeneous domains with complex geometries of the ground surface (Dell'Oro, 2011). In that work, several validation tests were performed, comparing the analytical solution with the numerical solutions obtained for some simple examples. Moreover, the model was applied to a real case study, i.e., a block of glacio-fluvial sediments dug in a quarry site in Varallo Pombia (Northern Italy; see section 2.4.5).

2.4.1 Mathematical Model and Discretization

The basic equation to be solved is (1.21), in case of a point source charge, centred in $\mathbf{x}_s = (x_s, y_s, z_s)$, in a isotropic but heterogeneous domain. By recalling the relationship between electrical resistivity ρ and conductivity σ, $\sigma = 1/\rho$, (1.21) becomes:

$$\mathrm{div}(\sigma \mathbf{grad} V) = I\delta(\mathbf{x} - \mathbf{x}_s) \tag{2.39}$$

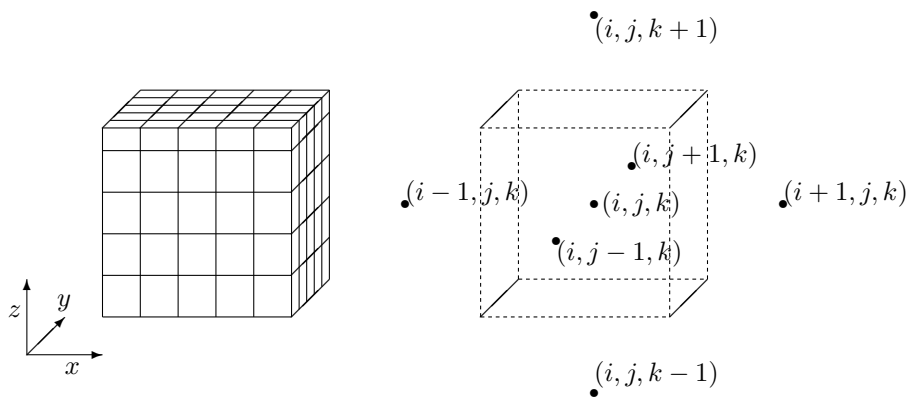


Figure 2.9: `YAELMod`: Domain discretization.

As done for the hydrological model, also in this case the continuum domain is subdivided in cells or blocks, which are identified by three indices (i, j, k), $1 \leq i \leq Nx$, $1 \leq j \leq Ny$, $1 \leq k \leq Nz$. The ground surface is assumed to be plain. Moreover, the side lengths of the cells along x, y and z directions, Δx, Δy and Δz, are assumed to be constant for the whole grid, but for the shallowest layer: in fact, the thickness of the last layer, i.e., the surface layer, is half that of other layers, i.e. $\Delta z(k = Nz) = \frac{1}{2}\Delta z(k = i), i = 1, \ldots, Nz-1$, so that the nodes of this layer are positioned exactly on the ground surface. Resistivity values are assigned to every cell, whereas the electric potential is calculated at the nodes.

The integral version of equation (2.39) could be discretized for each cell as

$$\begin{aligned} &C_{(i,j,k+1/2)}(V_{(i,j,k+1} - V_{(i,j,k)}) + C_{(i,j,k-1/2)}(V_{(i,j,k-1} - V_{(i,j,k)}) + \\ &+C_{(i+1/2,j,k)}(V_{(i+1,j,k)} - V_{(i,j,k)}) + C_{(i-1/2,j,k)}(V_{(i-1,j,k)} - V_{(i,j,k)}) + \\ &+C_{(i,j+1/2,k)}(V_{i,j+1,k} - V_{(i,j,k)}) + C_{(i,j-1/2,k)}(V_{i,j-1,k} - V_{(i,j,k)}) = 0. \end{aligned} \tag{2.40}$$

where the internode conductances $C_{(i,j,k+1/2)}$ and similar are defined as

$$C = \sigma' \frac{\mathcal{A}}{\Delta\ell} \tag{2.41}$$

with σ' the internode conductivity, $\mathcal{A}$ the lateral surface of a cell and $\Delta\ell$ the internode distance.

For the sake of simplicity, equation (2.40) could be written in a symbolic way, where V represents the potential in the cell (i, j, k) and V' the potential in adjacent cells:

$$\sum C \cdot (V' - V) = 0, \tag{2.42}$$

where the summation is extended to the six interblock surfaces separating the central cell with its neighbouring cells. When the node is occupied by a point source charge, which injects a current I,

$$\sum C \cdot (V' - V) = I. \tag{2.43}$$

Some remarks about the most meaningful average to be adopted for the computation of σ' are necessary. The current flowing through the surface that separates two adjacent cells, which are characterized by electrical resistivities ρ_1 and ρ_2, could be modelled as the electrical current in a circuit with two resistances in series: this mean that the internode resistivity, ρ', is the arithmetic mean of individual resistivities. By recalling that $\rho = 1/\sigma$, then:

$$\rho' = \frac{\rho_1 + \rho_2}{2}, \tag{2.44}$$

so that

$$\sigma' = \frac{2}{\frac{1}{\sigma_1} + \frac{1}{\sigma_2}} = \frac{2\sigma_1\sigma_2}{\sigma_1 + \sigma_2}. \tag{2.45}$$

2.4.2 Boundary Conditions

For the simulation of the electrical field generated by an electrode posed on the ground surface, the boundary conditions to be considered are:

- no flow on the ground surface Σ in contact with the air,

$$\mathbf{grad}V \cdot \mathbf{n} = 0, \tag{2.46}$$

  where $\mathbf{n}$ is the unit vector orthogonal to the surface Σ;

- null potential at infinity,

$$V = 0, r \to \infty, \tag{2.47}$$

  where $\mathbf{r} = \mathbf{x} - \mathbf{x}_s$ and $r = |\mathbf{r}| = ((x - x_s)^2 + (y - y_s)^2 + (z - z_s)^2)^{1/2}$.

To model BCs of no flow, the flux corresponding to the cell face in contact with air must be null and is eliminated from (2.43).

When working with discrete domains, the main problem is how to manage the null potential condition at great distance (infinity). Assuming that the potential is null on the borders of the domain would introduce great errors, unless the dimension of the domain is very large in order to keep great the distance between the sources and the borders. To solve this problem, Dey and Morrison (1979) proposed a mixed boundary condition, explained in the following.

The electric potential and field for an homogeneous medium asymptotically behave as

$$V(r) \sim \frac{1}{r} \text{ and } |\mathbf{E}(r)| \sim \frac{1}{r^2} \tag{2.48}$$

at great distances r from the charge source. At great distances from the source, it can be assumed that the domain can be considered as an equivalent medium, characterized by a mean value of resistivity, calculated over all the domain, and that the behaviour of V and $\mathbf{E}$ is given by (2.48). From a mathematical point of view, the field is supposed to be $V(x, y, z) = cr^{-1}$, where c is a constant, for $r \to \infty$. Then the derivative of the potential with respect to x could be calculated as

$$\begin{aligned} \frac{\partial}{\partial x} V(r) = \frac{\partial}{\partial x} \frac{c}{r} = c \frac{\partial}{\partial x} \frac{1}{((x - x_s)^2 + (y - y_s)^2 + (z - z_s)^2)^{1/2}} = \\ = -c \frac{(x - x_s)}{r^3} = -V(r) \frac{(x - x_s)}{r^2}. \end{aligned} \tag{2.49}$$

Then, under the hypotheses mentioned above, the potential gradient at the borders of the domain could be approximated as

$$\mathbf{grad} V = -V \frac{\mathbf{r}}{r^2}. \tag{2.50}$$

If $\mathbf{n}$ is the unit vector orthogonal to the border, the mixed BCs become

$$\mathbf{grad} V \cdot \mathbf{n} = -V \frac{\mathbf{r} \cdot \mathbf{n}}{r^2}. \tag{2.51}$$

Using (2.51) the internode conductance of a cell next to the borders becomes

$$C_{mix} = -\sigma' \frac{\mathcal{A}}{r^2} \mathbf{r} \cdot \mathbf{n} \tag{2.52}$$

and the electric current through these types of faces are calculated as $-\sigma' V \mathcal{A} \mathbf{r} \cdot \mathbf{n}\, r^{-2}$, and it would replace one of the terms of equations (2.42) or (2.43).

As done for the hydraulic problem, equation (2.40) can be written for each internal cell and therefore yields a system of possibly non-linear equations, that can be solved with the same relaxation method for the solution of systems of algebraic linear equations, explained in section 2.3.3. The application of the Gauss-Seidel method (2.24) practically requires to solve (2.40) for the calculation of the electric potential at the node of the cell where such equation is applied. The electric potential V at that node could be written in different ways, depending on the type and position of the cell:

- internal node centred on the source

$$V = \frac{\sum CV' - I}{\sum C}; \tag{2.53}$$

- internal node without source terms

$$V = \frac{\sum CV'}{\sum C}; \tag{2.54}$$

- node with no flow boundary condition

$$V = \frac{\sum_0 CV'}{\sum_0 C};, \tag{2.55}$$

where $\sum_0$ means that the surface in contact with the air is excluded from the calculus, because there is no flux through this surface;

- node with mixed boundary condition

$$V = \frac{\sum_c CV'}{\sum_c C + C_{mix}}, \tag{2.56}$$

where $\sum_c$ means that the face of the cell in contact with the border, where a mixed BCs is imposed, is excluded from this sum and such flux is calculated using mixed BCs, see equation (2.51).

2.4.3 Primary and Secondary Field

Solving equation (2.40) with the finite difference technique might yield a large approximation error, because at the source point the potential is singular (Lowry et al., 1989; Zhao and Yedlin, 1996): rapid variations of the electrical potential close to the source charge generates high approximation errors due to the discretization with finite differences. In this work, the modification suggested by Lowry et al. (1989) and Zhao and Yedlin (1996) to finite difference modelling is applied. It consists in splitting the potential in two contributions: the primary field (V_p), caused by the current source in a uniform half-space and calculated from equation (1.27), and the secondary field (V_s), caused by the subsurface inhomogeneities. In a similar way, also electrical conductivity is split as the sum of two contributions, a constant value σ_0 and a spatially varying term $\hat{\sigma}$:

$$V = V_p + V_s \tag{2.57}$$

and

$$\sigma = \sigma_0 + \hat{\sigma}, \tag{2.58}$$

where σ_0 could be the mean value of σ all over the domain or, as done in this work, the value of electrical conductivity at the source point (Zhao and Yedlin, 1996). BCs for V_p and V_s are the same as those explained in section 2.4.2.

Equation (2.39) becomes

$$\mathrm{div}\left((\sigma_0 + \hat{\sigma})\mathbf{grad}(V_p + V_s)\right) = I\delta\left(\mathbf{x} - \mathbf{x}_s\right), \tag{2.59}$$

that is

$$\mathrm{div}\left(\sigma_0\mathbf{grad}V_p + \sigma_0\mathbf{grad}V_s + \hat{\sigma}\mathbf{grad}V_p + \hat{\sigma}\mathbf{grad}V_p\right) = I\delta\left(\mathbf{x} - \mathbf{x}_s\right). \tag{2.60}$$

Remembering that V_p is a solution of Poisson's equation

$$\sigma_0\nabla^2 V_p = I\delta\left(\mathbf{x} - \mathbf{x}_s\right), \tag{2.61}$$

and σ_0 is a constant value, then equation (2.60) becomes

$$\mathrm{div}(\sigma V_s) + \sigma_0\nabla^2 V_p + \mathrm{div}(\hat{\sigma}\mathbf{grad}V_p) = I\delta\left(\mathbf{x} - \mathbf{x}_s\right). \tag{2.62}$$

Moreover, by considering equation (2.61), (2.62) can be written as

$$\mathrm{div}(\sigma V_s) = -\mathrm{div}(\hat{\sigma}\mathbf{grad}V_p). \tag{2.63}$$

On the right hand side of equation (2.63), a new effective source term, as a consequence of the splitting in two fields, now appears.

2.4.4 Integral approximation for flux calculation

A consequence of the decomposition of the electric potential is, therefore, the introduction of an additional effective source term. The analytic evaluation of the volume integral of this term requires a further modification of the finite difference scheme.

The current flux between a cell occupied by a point charge source and its adjacent cells, associated to the primary potential is given by

$$\Phi = -\hat{\sigma}' \iint \frac{\partial V_p}{\partial \ell} \, \mathrm{d}\mathcal{A}, \tag{2.64}$$

where the double integral extends on the face separating two adjacent cells, while ℓ denotes the direction perpendicular to this face. Such an integral is calculated analytically, following Lowry et al. (1989) and Zhao and Yedlin (1996), to limit the errors caused by finite-difference approximations in a region where the potential V is rapidly varying. As a paradigm, let the integral of equation (2.64) be considered for the surface area $d\mathcal{A} = \mathrm{d}y \cdot \mathrm{d}z$ as

$$\Phi = -\frac{\rho_s I}{2\pi} \hat{\sigma}' \left[\int_{z_s - \Delta z}^{z_s} \int_{y_s - \frac{\Delta y}{2}}^{y_s + \frac{\Delta y}{2}} \frac{\partial}{\partial x} \frac{1}{r} \, \mathrm{d}z \mathrm{d}y \right]_{x = x_s + \frac{\Delta x}{2}}. \tag{2.65}$$

Then, the computation reduces to solve the integral

$$\digamma(y, z; x) = \iint \frac{x}{r^3} \, \mathrm{d}z \mathrm{d}y, \tag{2.66}$$

that is considered as a function of y and z and that parametrically depends on x, because it is calculated for fixed x values. The function in equation (2.66) satisfies

$$\frac{\partial^2 \digamma}{\partial z \partial y} = \frac{x}{r^3}, \tag{2.67}$$

and

$$\digamma(y, z; x) = \arctan \frac{yz}{xr}. \tag{2.68}$$

So,

$$\begin{aligned} \Phi = -\frac{\rho_s I}{2\pi} \hat{\sigma}' \cdot \Bigg\{ & \arctan \frac{(y_s + \Delta y/2) z_s}{(x_s + \Delta x/2)[(x_s + \Delta x/2)^2 + (y_s + \Delta y/2)^2]^{1/2}} \\ & - \arctan \frac{(y_s + \Delta y/2)(z_s - \Delta z)}{(x_s + \Delta x/2)[(x_s + \Delta x/2)^2 + (y_s + \Delta y/2)^2 + (z_s - \Delta z)^2]^{1/2}} \\ & - \arctan \frac{(y_s - \Delta y/2) z_s}{(x_s + \Delta x/2)[(x_s + \Delta x/2)^2 + (y_s - \Delta y/2)^2 + z_s^2]^{1/2}} \\ & + \arctan \frac{(y_s - \Delta y/2)(z_s - \Delta z)}{(x_s + \Delta x/2)[(x_s + \Delta x/2)^2 + (y_s - \Delta y/2)^2 + (z_s - \Delta z)^2]^{1/2}} \Bigg\}. \end{aligned} \tag{2.69}$$

This integral method used for the calculation of the primary potential V_p is easily extended to all the cells of the domain, even if they are not occupied by a point charge-source, because the potential V_p is known everywhere. In this way, the net flux Ψ through

all the faces of those adjacent cells of the domain, which are not occupied by a source term, could be calculated with this integral method, in place of using finite difference scheme.

Let the decomposition $V = V_p + V_s$ and the calculation of Φ and Ψ with the integral method be introduced into equations (2.42) and (2.43). This yields the following cases for the computation of V_S at each cell and the application of (2.24):

- node adjacent to the source

$$V_s = \frac{\sum CV_s' + \sum_0 \Psi + \Phi}{\sum C}; \tag{2.70}$$

- Internal node non-adjacent to the source

$$V_s = \frac{\sum CV_s' + \Psi}{\sum C}; \tag{2.71}$$

- Node with no flow boundary condition

$$V_s = \frac{\sum_0 CV_s' + \sum_0 \Psi}{\sum_0 C}; \tag{2.72}$$

- Node with mixed boundary condition

$$V_s = \frac{\sum_c CV_s' + \sum_c \Psi}{\sum_c C + C_{mix}}. \tag{2.73}$$

Here the internode electrical conductivity which appears in the definition of the conductances (2.41) is calculated as

$$\hat{\sigma}' = \frac{2\sigma_1\sigma_2}{(\sigma_1 + \sigma_2)} - \sigma_0, \tag{2.74}$$

Finally, it is worth noticing that the same iterative method described in section 2.3.3 is applied to solve the system of equations that is obtained for the secondary potential, but in this case the system is linear and therefore the choice of the accelerating parameter ω is less critical to obtain a good approximate solution than for YAGMod.

2.4.5 Application of YAELMod to a real case

The YAELMod code has been used to interpret non conventional electrical survey, in a region where the hypotheses of infinitely wide domain and plane surface of the terrain are not valid. The work has been conducted on a block of sediment dug in a quarry site in Varallo Pombia (Ticino Valley, Northern Italy). The block had the form of a parallelepiped, whose height varied between 1.5 m and 1.9 m and which was 3.6 m $\times$ 7.4 m wide. On this block many multidisciplinary studies have been conducted, including geoelectrical data acquisition - ERGI (Baines et al., 2002), with a multi-electrode system, in order to characterize subsurface heterogeneity and its effects on water flow and transports. The dipole-dipole electrical survey was performed with 16 electrodes, along eight parallel lines on the top surface of the block, with a spacing of 0.40 m. The output of the numerical model have been used to correct field data, in order to take into account the effects of the specific geometry.

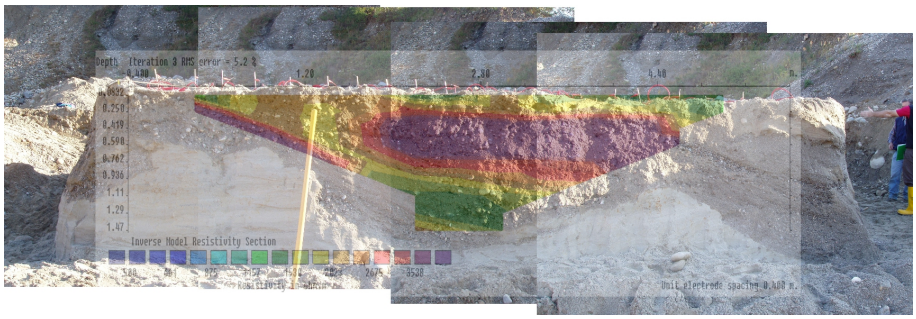


Figure 2.10: The block of sediments of the Ticino Valley (Northern Italy) with the electrodic line and a superposition of the apparent resistivity section obtained with `YAELMod`

2.5 The implementation of the simplified EnKF for parallel computing: the `JoMod` code

The procedure described in section 1.4 is implemented with a code developed with the Octave interpreted programming language and designed to run on a parallel computer[3], under Linux operating system. In fact inverse methods often require a huge number of runs of the codes that solve the forward problems, with different sets of data. The approach proposed in this work, which is somehow akin to a Monte-Carlo approach, can be easily implemented on a parallel computer, because the simulations corresponding to the realizations of a given ensemble can run independently from each other. This permits to keep the execution time within reasonable limits, which is not the case if the codes are executed on standard personal computers, even with high performance processors.

The implementation of the `JoMod` code can be summarized in the following work flow scheme.

A. Data pre-processing.
The measured data are read from input data and stored for further analysis. The initial distribution of the phenomenological parameters is built up starting from a set of initial parameter guess, inserted by the user.

B. Solving the hydraulic and electric forward problems.
This step of the implementation is organized in an "internal loop" in order to use all available field data.

1. Creation of N_R directories where the input and output files of `YAGMod` and `YAELMod` will be stored. An appropriate string command submits a job to a queue that manages the execution of N_R parallel jobs, one for each realization, to:
   - prepare input files for `YAGMod`;
   - run `YAGMod` and solve the hydraulic forward problem;
   - run a *read-and-write* Fortran-compiled code to create the spatial distribution of resistivity and the input file for `YAELMod` with the saturation field obtained from `YAGMod`;

[3]The results shown in chapter 3 are obtained by running `JoMod` on the 32-processors cluster of the Laboratorio di Calcolo parallelo at the Dipartimento di Scienze della Terra dell'Università degli Studi di Milano.

- run YAELMod and solve the electrical forward problem.

2. Then a very simple control cycle verifies whether all the N_R runs are completed (hydraulic and electric modelling) before moving to the next step. Once the cycle of all the N_R parallel runs is completed, output files are stored in a dedicated directory.
3. JoMod performs a *statistical analysis* (see Fig. 1.6) on the obtained results, by applying the EnKF technique (section 2.1.1) or the simplified EnKF technique (section 2.1.2). The parameter distribution is updated in order to improve the estimate and reduce the uncertainty.
4. JoMod iteratively repeats this loop from step 1 for every set of data; all the distribution of parameters are saved at every step in a file that could be used for *ex post* data analysis.

C. Data post-processing.

The JoMod code continues the "external loop" by restarting from point B. till the criterion expressed by equation (2.8) is fulfilled.

CHAPTER 3

Application

3.1 The case study

The site chosen to perform some tests on the proposed procedure is a well field located in Pozzuolo Martesana, about 20 km east of Milan, in the Po plain. The well field (rectangular yellow area, Fig. 3.1) has been built to provide drinkable water. In this area eight boreholes, each corresponding to a couple of water wells, have been drilled and water is pumped from both shallow and deep aquifers. For each borehole stratigraphic logs, which are used for the geostatistical reconstruction, and pumping tests of the two wells are available.

The geological evolution of the area is characterized by a regressive mega-cycle, with a progressive migration from North to South during the whole Pleistocene (Regione Lombardia and Eni Divisione Agip, 2001). The marine depositional environment (lower Pleistocene) is progressively substituted, in a fluctuating way, by transitional (lower Pleistocene) and then continental (medium-upper Pleistocene) depositional environments. The continental deposits were dominated by the alternation of glacial and interglacial periods, at the end of which the post-glacial and Holocene sediments sealed the relict incisions and covered the valleys of the present-day fluvial network.

This sedimentological evolution created an hydrostratigraphic structure that is characterized by alternations of fine- and coarse-grained sediments (Cavalli, 2011): the former are represented by silt and clay originated in marine platforms, coastal lagoons or flooding plains or due to pedogenetic episodes, are characterized by high electrical conductivity and low hydraulic conductivity; the latter are represented by sand and gravel deposits, which show high hydraulic conductivity and greater variability of electrical conductivity between drained and wet conditions, as shown, e.g., by Mele et al. (2012).

For each couple of wells, water is extracted separately from permeable deep ($> 80\ \mathrm{m}$ below ground surface) and shallow ($< 100\ \mathrm{m}$ below ground surface) levels. Pumped water is collected in two tanks, located in the central building of the pumping station. Automatic electronic systems control the pumping schedule: the water level in the tanks causes the activation of two deep wells and two shallow wells and determines their pumping rates. The selection of the pumping wells is changed every 15 days. However, it is possible to design and activate different pumping schemes, but pumping rates cannot be selected by the operators.

During the months of May and June 2011, some field campaigns were conducted to verify the possibility of performing some experiments in that area. The results of resistivity surveys (ERGI and SEV) showed several effects that could be attributed to artificial structures, e.g., buried pipes, electrical power lines (see Fig. 3.2).

The area (see Fig. 3.2) is obviously crossed by two networks of pipelines (with a diameter of $200\ \mathrm{mm}$), which take water from the deep and shallow wells to the pumping

Figure 3.1: Aerial map and localization of the well field in Pozzuolo Martesana (Northern Italy).

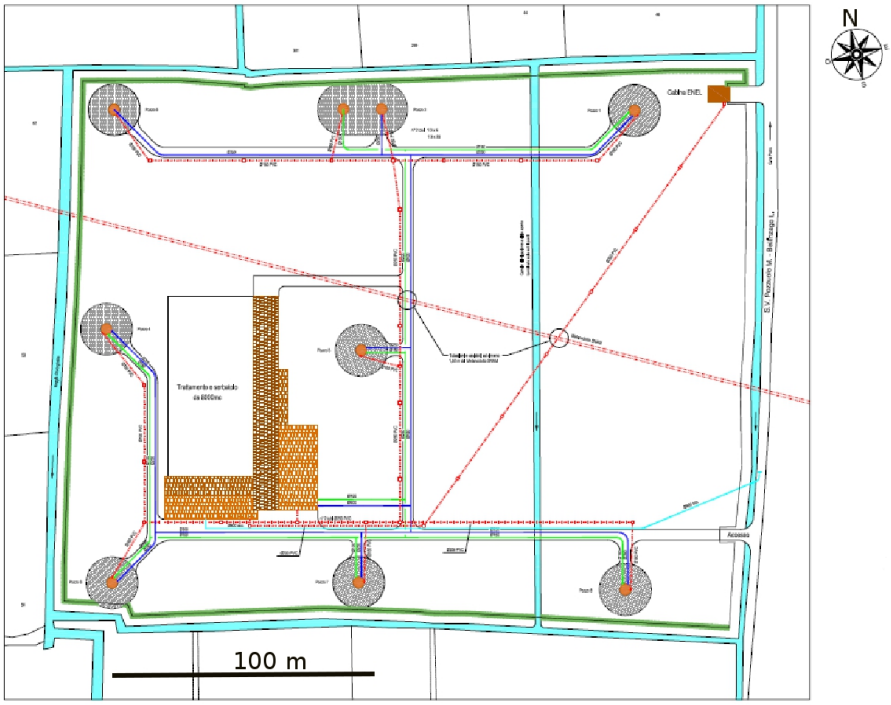
station. Unfortunately, before the field measurements it was not known the presence of a gas pipeline that crosses the area from west/north-west to east/south-east, and the presence of the drinkable water pipeline (with a diameter of $800\ \mathrm{mm}$).

Unfortunately, these buried structures caused artefacts on the geoelectrical data that masked the geophysical signature of the natural system and of the dynamics of the water table and thus prevented from the possibility of using these data or to organize experiments. Then, the hydrostratigraphic data collected at the site were used to make a realistic geological reconstruction of the domain, but a synthetic example is preferred to test the proposed methods in order to keep a better control on the proposed procedure.

Following the sedimentological characterization proposed by Cavalli (2011) and for the sake of simplicity, the preliminary reconstruction of the hydrogeological structure of the site has been obtained by grouping the sediments identified in the stratigraphic logs into two main categories, corresponding to the two HGFs used in the simulations: *coarse* and *fine*.

Then, the eight stratigraphic logs have been digitized with a half-meter resolution. The file including x, y, z coordinates[1] and HGF codes is then elaborated with the geostatistical software SGEMS (see section 2.2). Fig. 3.3 shows the experimental input variogram for the z direction. The experimental variograms along x and y cannot be computed in a reliable way with only eight logs; in particular, for this specific case, they are both flat and close to zero, due to the input HGFs data. Therefore the experimental variogam for the z direction was fitted with an exponential model, whereas the range of the model variogram for x and y directions has been estimated on the basis of literature data and results for similar examples, as, for instance Perulero Serrano et al. (2014). In

[1]The cartesian coordinates are chosen by taking into account the geometry of the study site, namely, with the x axis following the west/north-west to east/south-east direction, which corresponds to the longest side along which the boreholes are aligned.



ht

Figure 3.2: Sketch of the buried infrastructures of the well field, reproduced from "tavola 13" of Minotta (2002). Wells positions are denoted by brown dots. Green and blue lines correspond to water conduits for, respectively, shallow and deep wells; thin cyan line corresponds to the drinkable water pipe; thick cyan lines correspond to irrigation ditches; the red double line, crossing the area from west/north-west to east/south-east, corresponds to a gas pipeline; red dotted lines correspond to electrical-power lines.

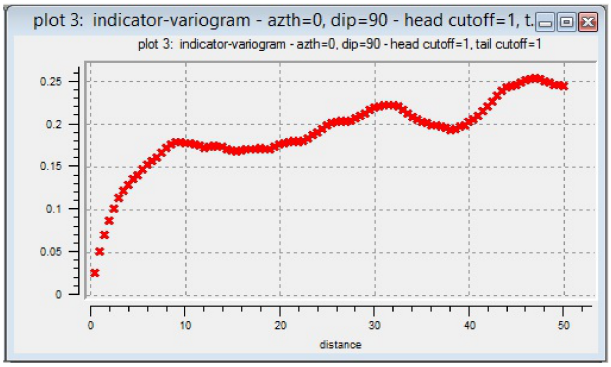


Figure 3.3: The experimental variogram along z direction.

Table 3.1 the input parameters for the experimental variogram and the model variogram are summarised. Recall that the parameters of the model variogram are among the input data for SISIM.

The domain has been discretized with $70 \times 50 \times 40$ cubic cells of 5 m side. In Fig. 3.4 the eight digitized stratigraphic logs and the corresponding statistical simulation are shown.

Experimental variogram input parameters	
# of lags	100
lag separation	0.5 m
lag tolerance	0.25 m
Model variogram parameters	
Sill	0.175
Type	*Exponential*
Ranges	8 m along z direction 500 m along x and y directions

Table 3.1: Parameters for the computation of the experimental variogram and for the theoretical variogram.

If the hydraulic head field, $h(\mathbf{x}, t)$, is known as a function of position and time, as a first analysis, in this work, the saturation function assumes the following simple form:

$$S(\mathbf{x}, t) = \begin{cases} 1, \ \forall \mathbf{x} = (x, y, z) : z \leq h(\mathbf{x}, t) \\ \\ 0, \ \forall \mathbf{x} = (x, y, z) : z > h(\mathbf{x}, t). \end{cases} \tag{3.1}$$

In other words, the porous medium is assumed to be either dry or fully saturated and the effect of partial saturation is neglected. Correspondingly, the semi-empirical functions $\varkappa$ and ϱ connecting K and ρ to the HGFs and to the saturation degree simply depend on three values: a single value of hydraulic conductivity and two values of electrical resistivity, which refer to saturated and dry conditions, are assigned for each HGF.

Therefore, the model parameters for every HGF are the hydraulic conductivity ($K = K_s$) only in the limit case of saturation $S(\mathbf{x}, t) = 1$ and the electrical resistivities (ρ_w and ρ_d) for the two limit cases $S(\mathbf{x}, t) = 1$ and $S(\mathbf{x}, t) = 0$. The assigned values for the hydraulic and electric input parameters are reported in table 3.2.

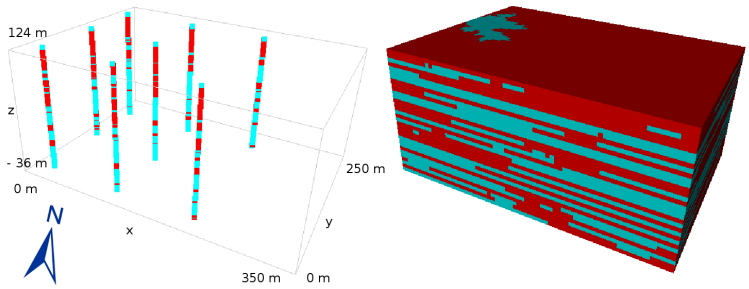


Figure 3.4: The geological discretized domain: the eight digitized stratigraphic logs (left) and the corresponding statistical simulation (right). Coarse sediments in cyan, fine sediments in red.

HGF	Hydraulic conductivity (m/s)	Wet electrical resistivity (ohm · m)	Dry electrical resistivity (ohm · m)
COARSE	C-K = 0.0015	C-ρ_w = 300	C-ρ_d = 100
FINE	F-K = 0.00005	F-ρ_w = 20	F-ρ_d = 10

Table 3.2: Hydrogeological parameters values assigned to HGFs.

3.2 Sensitivity Analysis on boundary conditions

A first test was conducted to assess the sensitivity of the method to different types of boundary conditions (BCs) for the hydraulic problem. In fact, it should be recalled that in a situation like that of a well field, the area is part of a very wide aquifer and no physical BC can be assigned to the hydraulic problem.

The "natural" flow field is modified by a single well located at $x = 122\ \mathrm{m}, y = 21\ \mathrm{m}$ with a filtered interval between $z_{min} = 80\ \mathrm{m}$ and $z_{max} = 110\ \mathrm{m}$ and with an extraction rate of $-0.5\ \mathrm{m}^3/\mathrm{s}$ (see Fig. 3.5). Such a value of extraction rate is very high, much greater than the typical discharge rates of the real wells, but it is chosen in order to enhance the effects of the BCs and facilitate the analysis of the results.

Three different types of BCs, Dirichlet (prescribed head), Neumann (imposed flux) and Robin BCs (a combination of prescribed head and flux), are tested in four different combinations to qualitatively assess the influence of BCs on the simulated saturation field, water level and pseudosection of apparent resistivity.

- Scenario #1: Dirichlet BCs along y borders and Neumann (no flow) BCs along x borders;
- Scenario #2: Robin BCs along y borders and Neumann (no flow) BCs along x borders;
- Scenario #3: Dirichlet BCs along both x and y borders;
- Scenario #4: Robin BCs along both x and y borders.

The results are illustrated in Figg. 3.6, 3.7 and 3.8. For their interpretation it is important to remember that Dirichlet BCs are physically equivalent to the presence of a

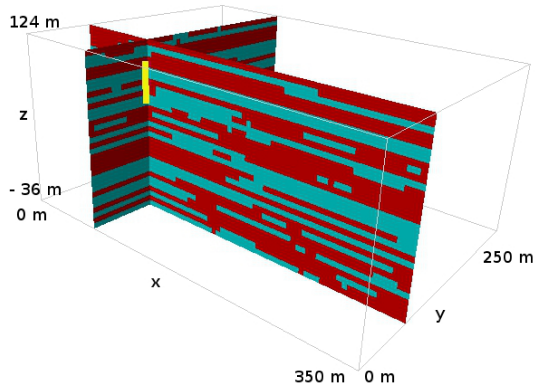


Figure 3.5: Geometry of the volume domain for sensitivity analysis on BCs: the distribution of HGFs (cyan: coarse material; red: fine material) is shown on two vertical sections, crossing at the pumping well position (yellow vertical bar).

reservoir that permits to exchange any flux without changing the water level: this is a very strong and restricting BC for a situation like that considered in this case study.

In Fig. 3.6 the saturation field is mapped for a vertical section, along x direction, crossing the well. Imposing Robin BCs on two (plot 2 of Fig. 3.6) or four borders (plot 4 of Fig. 3.6) of the domain does not create sensitive variation of the saturation field, while the assignment of Dirichlet BCs all around the domain (plot 3 of Fig. 3.6) prevents the formation of an extended desaturated area.

The water level (Fig. 3.7) is the most sensitive quantity to the BCs: the shape of the hydraulic head field and the values of piezometric head change sensitively from Dirichlet BCs (plots 1 and 3 of Fig. 3.7) to Robin BCs (plots 2 and 4 of Fig. 3.7); changes could be noticed also from cases when Dirichlet BCs are assigned along two borders (plot 1 of Fig. 3.7) or four borders (plot 3 of Fig. 3.7). This is more evident in hydraulic head values than in the shape of the contour lines. On the other hand no relevant change could be appreciated if Robin BCs are applied along two (plot 2 Fig. 3.7) or four borders (plot 4 of Fig. 3.7).

The variations obtained for the pseudosection of apparent resistivity (see Fig. 3.8), along the section indicated by the yellow dotted line in Fig. 3.7 are obviously related to the variations of the saturation field. The observed variations are in agreement with those observed for saturation fields and water levels. An increment in apparent resistivity values is noticed by passing from two Dirichlet BCs to four Dirichlet BCs (plots 1 and 3 of Fig. 3.8). A decrease in apparent resistivity is obtained from Dirichlet BCs (plots 1 and 3 of Fig. 3.8) to Robin BCs (plots 2 and 4 of Fig. 3.8). There are no sensible changes between simulations with two or four Robin BCs (plots 2 and 4 of Fig. 3.8).

3.3 Sensitivity Analysis on parameters

To asses the sensitivity that model outcomes have on input parameters, several simulations have been conducted by varying one-at-a-time the six input parameters by a given quantity with respect to their reference values (Table 3.2) and the corresponding variation of the apparent resistivity for every quadrupole of a dipole-dipole electrode configuration is considered.

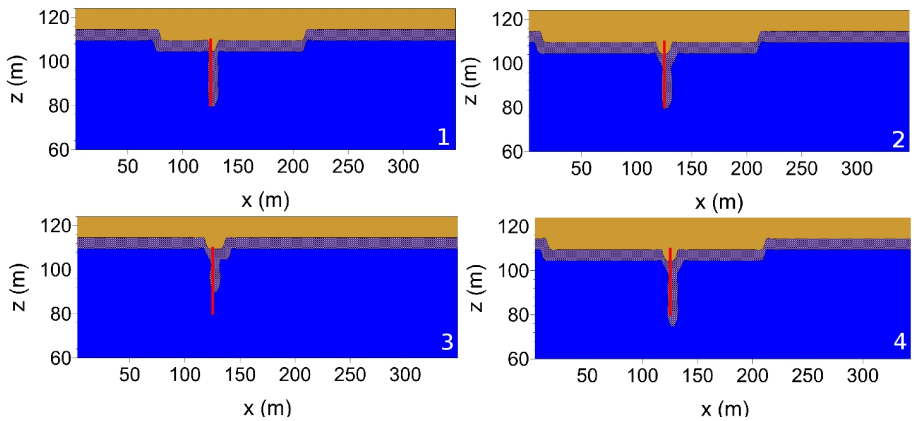


Figure 3.6: Vertical section of the saturation field (saturated cells in blue, dry cells in brown) along x direction crossing the well: 1) two borders with Dirichlet BCs and two borders with Neumann null flux BCs; 2) two borders with Robin BCs and two borders with Neumann null flux BCs; 3) four borders with Dirichlet BCs; 4) four borders with Robin BCs.

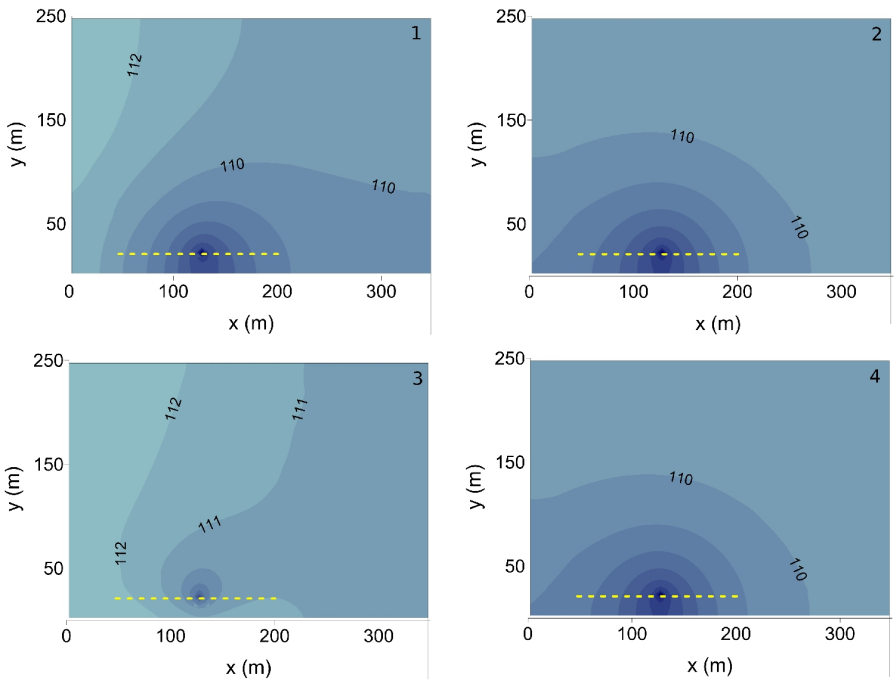


Figure 3.7: Elevation (in meters) of the water table (contour interval 0.5 m): 1) two borders with Dirichlet BCs and two borders with Neumann null flux BCs; 2) two borders with Robin BCs and two borders with Neumann null flux BCs; 3) four borders with Dirichlet BCs; 4) four borders with Robin BCs. The yellow dotted lines indicates the position of the electrodic lines used for electrical surveys.

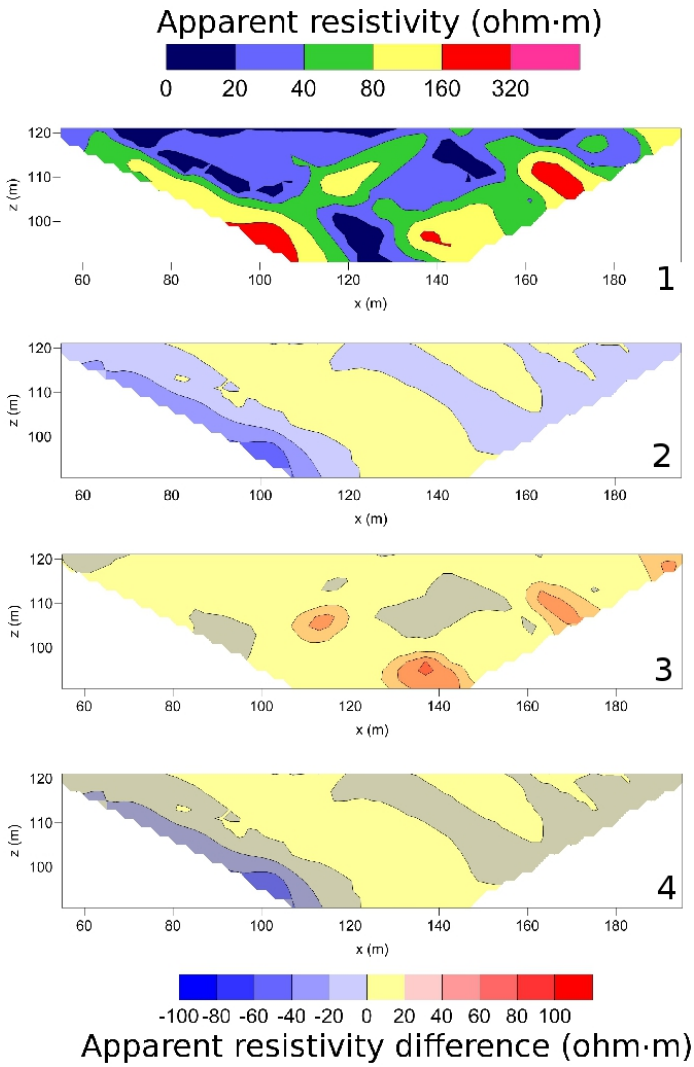


Figure 3.8: Pseudosections of apparent resistivity (the electrodic line is shown by the yellow dotted line in Fig 3.7): 1) the apparent resistivity values obtained with simulation #1 (two borders with Dirichlet BCs and two borders with Neumann null flux BCs) are shown , legend on the top; then, the difference between the results for scenario #1 and other scenarios are mapped: 2) two borders with Robin BCs and two borders with Neumann null flux BCs; 3) four borders with Dirichlet BCs; 4) four borders with Robin BCs, legend on the bottom.

In order to simplify the notation, the six input parameters are denoted with X_l, $l = 1, \ldots, 6$, where X_1 = C-ρ_w, X_2 = F-ρ_w, X_3 = C-ρ_d, X_4 = F-ρ_d, X_5 = C-K, X_6 = F-K. The model parameters are increased and decreased by 20% of $\mathbf{X}_{(\mathrm{ref})}$, the reference values listed in Table 3.2.

Two variability indices were calculated: the Relative Mean Variability (RMV)

$$\mathrm{RMV}_l = \frac{1}{N_M^{(\rho)}} \sum_{m=1}^{N_M^{(\rho)}} \frac{\rho_{app_m}(\mathbf{X}_l) - \rho_{app_m}(\mathbf{X}_{(\mathrm{ref})})}{\rho_{app_m}(\mathbf{X}_{(\mathrm{ref})})} \tag{3.2}$$

and the Root Mean Square Variability (RMSV)

$$\mathrm{RMSV}_l = \frac{1}{N_M^{(\rho)}} \sum_{m=1}^{N_M^{(\rho)}} \left(\frac{\rho_{app_m}(\mathbf{X}_l) - \rho_{app_m}(\mathbf{X}_{(\mathrm{ref})})}{\rho_{app_m}(\mathbf{X}_{(\mathrm{ref})})} \right)^2, \tag{3.3}$$

where $N_M^{(\rho)}$ is the number of quadrupoles for which an apparent resistivity value is computed, $\rho_{app_m}(\mathbf{X}_{(\mathrm{ref})})$, $m = 1, \ldots, N_M^{(\rho)}$ is the value of apparent resistivity of the m-th quadrupole computed with the reference values of the parameters and $\rho_{app_m}(X_l)$ the corresponding apparent resistivity computed when the l-th parameter is changed ($l = 1, \ldots, 6$).

The results are shown in Figg. 3.9 and 3.10. Both plots show that the model outcomes are more sensitive to variations of the resistivity of dry cell. This is due to the fact that the cells that become dry are shallow, and therefore close to the electrodes positions. Model outcomes are more sensitive to decreasing variation of input parameters, when the variation is induced by resistivity, while a very small variability could be observed for an increase of hydraulic conductivity of the coarse HGF. This is related to the fact that the cone of depression generated by the pumping well becomes greater if the hydraulic conductivity increases. Moreover, model outcomes seem not to be affected at all by hydraulic conductivity variability, but this remark is an obvious consequence of the relation between resistivity and saturation.

3.4 Application of the EnKF procedure

3.4.1 Synthetic data

As discussed in section 3.1, the impossibility of collecting reliable field data in the study site induced to generate synthetic data to test and check the applicability of the proposed procedure. To this aim, three different scenarios were simulated, and the hydraulic and electric forward problems were solved for the HGFs distribution described in section 3.1, with the reference input parameters that are listed in table 3.2. For the three scenarios, Robin BCs are prescribed all around the domain, by using the prototype equation (2.17): $\mathcal{H}^{(\mathrm{act})}$ is assigned as the height above mean sea level of the bottom of the cell, which this term is referred to; one of the two conductances $\mathcal{K}_1$ and $\mathcal{K}_2$ is set to zero, in order to deactivate the inflow or outflow source term when the cell is completely dry, whereas the other one is assigned the value $1.0 \cdot 10^{-5}\ \mathrm{m}^2/\mathrm{s}$; the prescribed value of $\mathcal{H}^{(\mathrm{calc})}$ in each cell aims at reproducing an hydraulic gradient along the x direction.

Three out of the 16 wells of the pumping station have been activated with different extraction rates for each simulated situation. The position of the three wells with respect to the electrodic line for the resistivity measurements is shown in Fig. 3.11. The hydraulic

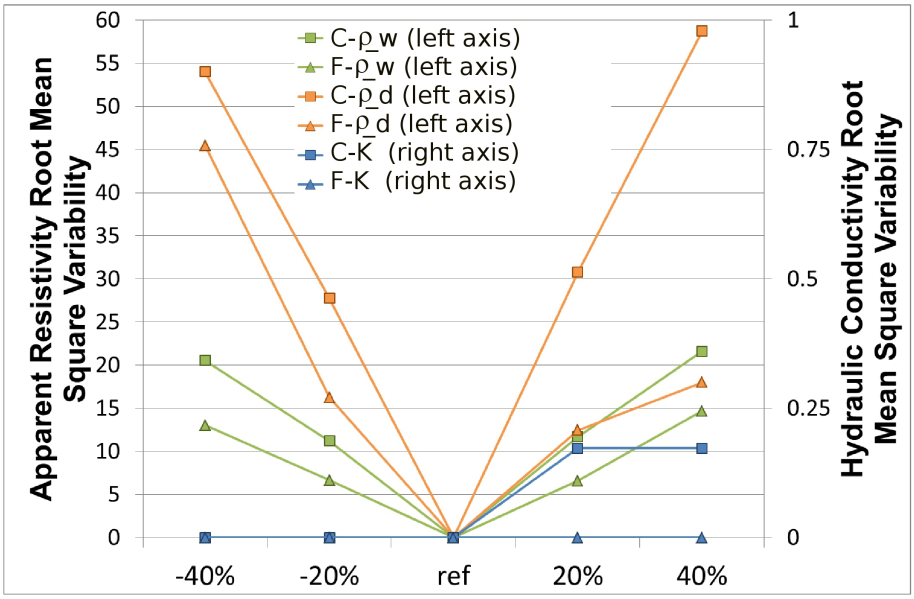


Figure 3.9: Root Mean Square Variability calculated on model parameters: scale of RMSV for electrical resistivity (C-ρ_w, F-ρ_w, C-ρ_d, F-ρ_d) is given on the left axis, for hydraulic conductivity (C-K, F-K) on the right axis.

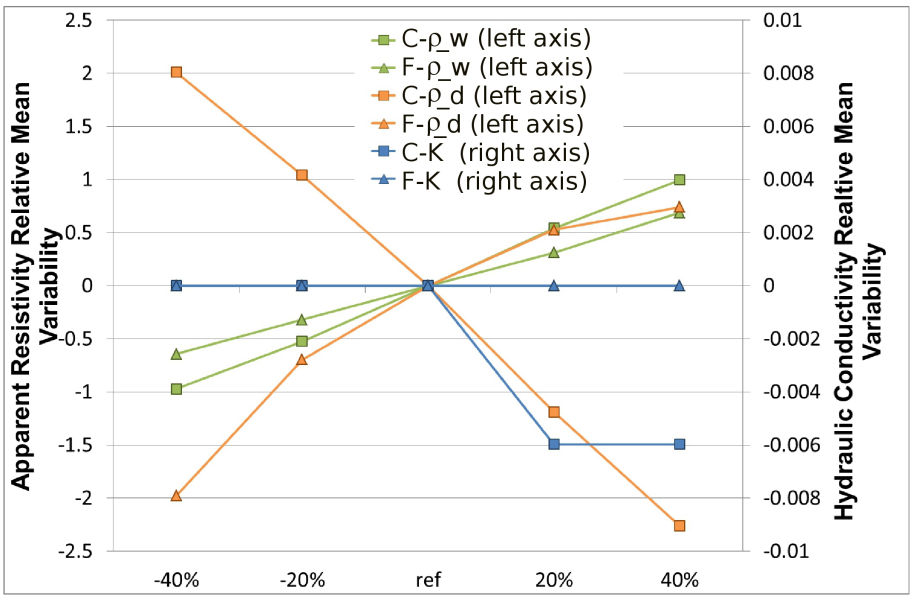


Figure 3.10: Root Mean Variability calculated on model parameters: scale of RMV for electrical resistivity (C-ρ_w, F-ρ_w, C-ρ_d, F-ρ_d) is given on the left axis, for hydraulic conductivity (C-K, F-K) on the right axis.

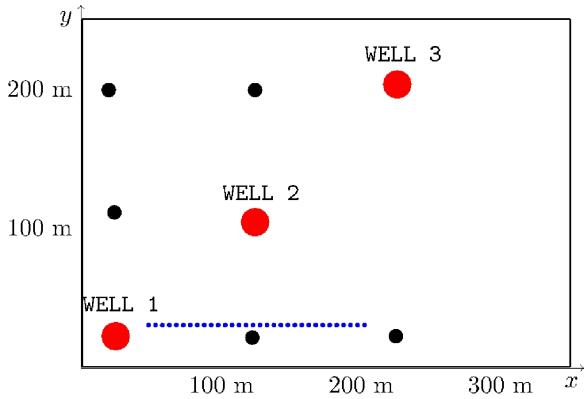


Figure 3.11: Sketch of the position of the three activated wells (red dots) and of the electrodic line (blue dotted lines). Black dots represents the other five inactive wells.

Well #	(x, y) coordinates	Simulation #1	Simulation #2	Simulation #3
1	$(24\ \mathrm{m}, 22\ \mathrm{m})$	$-0.2\ \mathrm{m}^3/\mathrm{s}$	$-0.002\ \mathrm{m}^3/\mathrm{s}$	$-0.002\ \mathrm{m}^3/\mathrm{s}$
2	$(124\ \mathrm{m}, 104\ \mathrm{m})$	$-0.002\ \mathrm{m}^3/\mathrm{s}$	$-0.2\ \mathrm{m}^3/\mathrm{s}$	$-0.002\ \mathrm{m}^3/\mathrm{s}$
3	$(226\ \mathrm{m}, 203\ \mathrm{m})$	$-0.002\ \mathrm{m}^3/\mathrm{s}$	$-0.002\ \mathrm{m}^3/\mathrm{s}$	$-0.2\ \mathrm{m}^3/\mathrm{s}$

Table 3.3: Extraction rates of the three wells for the three different simulated scenarios.

heads modelled at the three wells are taken as the hydraulic observed values for each situation. A scheme of the three different stress conditions is sketched in table 3.3.

The elevation of the water table obtained for the three different simulations is shown in Fig. 3.12.

An electrical survey is simulated along the electrodic line, which is shown in Fig. 3.11and also as a yellow dotted line in Fig. 3.12, with 32 electrodes equally spaced by $5\ \mathrm{m}$. The electric forward problem was then solved and the values of apparent resistivity for 368 electrodic quadrupole were obtained, by simulating a dipole-dipole survey.

Fig. 3.13 (plots 2 and 3) shows the variation of apparent resistivity for simulations # 2 and 3 with respect to that of simulation #1 (see Table 3.3 for stress condition details). The pseudosections of apparent resistivity for simulations # 2 and #3 are both quite different from that for simulation #1, but there are small differences between them. In fact, the

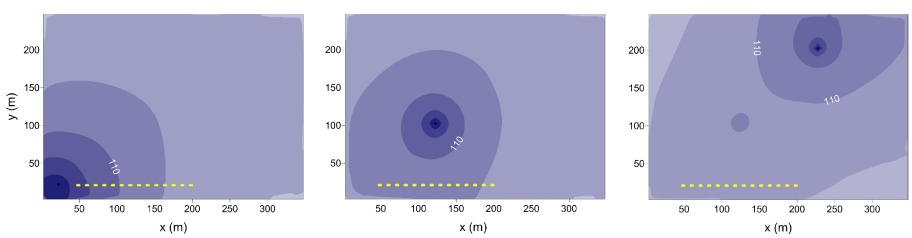


Figure 3.12: Elevation (in meters) of the water table (contour interval $0.5\ \mathrm{m}$): the numbers of the simulations correspond to those listed in table 3.3. The yellow dotted line represents the electrodic line.

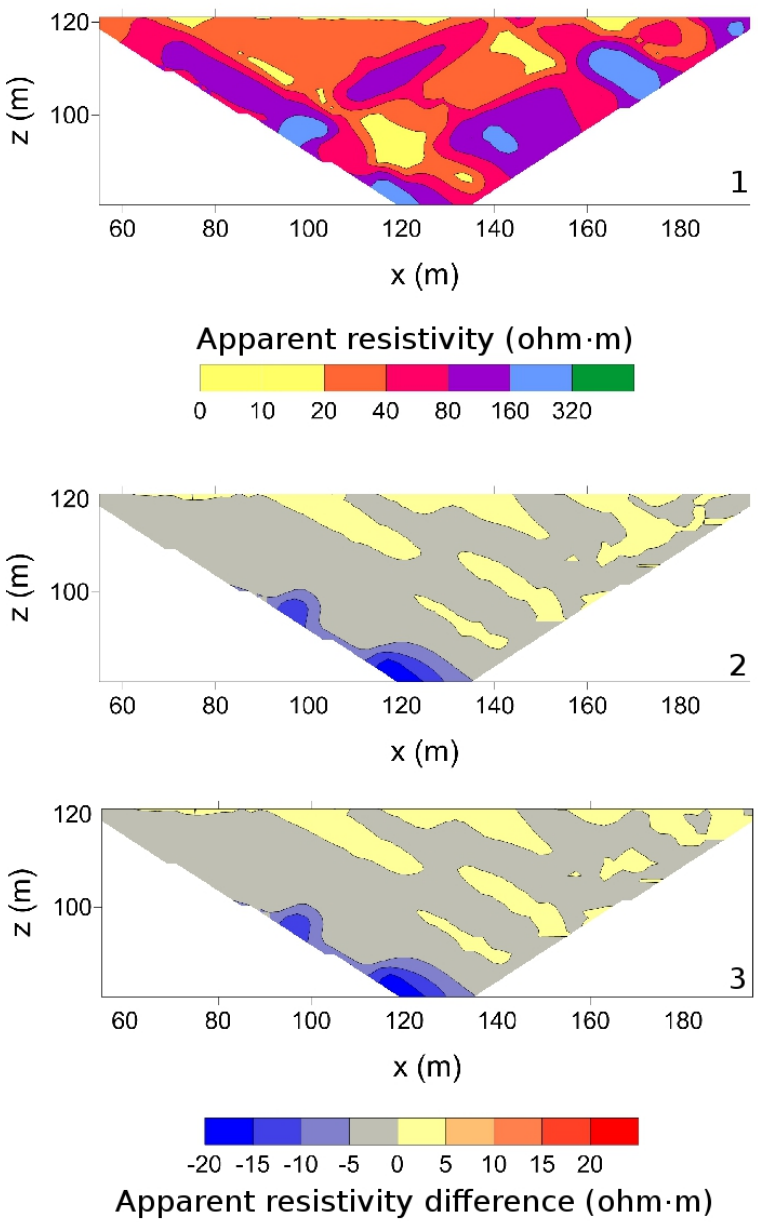


Figure 3.13: Pseudosections of apparent resistivity (the electrodic line is shown by the yellow dotted line in Fig 3.12) obtained for the three different simulations of the synthetic example: 1) the apparent resistivity values obtained with simulation #1 are shown, legend on the top; then, the difference between the results for simulation #1 and other simulations are mapped: 2) simulation #2 3)simulation #3, see Table 3.3 for stress condition details.

electrodic line (see Fig. 3.12) is quite far from the wells with the highest extraction rate for simulations #2 and #3; therefore, the measurements of apparent resistivity, that are strictly related to the saturation field, are not significantly affected by the depression cones of those two simulations.

The elevation of the water table obtained from the hydraulic forward problem in the three wells and the 368 apparent resistivity values obtained from the electrical forward problem for every different stress condition are then considered as the real observed field values.

3.4.2 The results

Several tests have been conducted with the `JoMod` code, in order to assess the properties of the proposed method. For each test, the cycle described at point [B.] of section 2.5 to update the estimate of the input parameters with the data corresponding to the three simulations described in section 3.4.1 was iteratively repeated, till the convergence criterion given by equation (2.8) is satisfied.

Reference test

The first test of the procedure started with the extraction of 20 equiprobable sets of the parameters X_l, $l = 1, \ldots, 6$, with a uniform distribution in the interval

$$\left[0.5 \cdot \mathbf{X}_{(\mathrm{ref})_l}, 1.5 \cdot \mathbf{X}_{(\mathrm{ref})_l}\right].$$

In other words, the initial values are extracted from uniform distributions centred on the reference values of the corresponding parameters and with ranges whose widths are equal to the reference values themselves.

As a first analysis, the invertibility of the gain matrix (see section 2.1.1) is investigated. Some numerical results, obtained with standard routines of Octave, for the conditioning number, the determinant and the eigenvalues λ_m, $m = 1, \ldots, N_M$, of the matrix $\mathrm{Cov}^{(\mathsf{k})}[YY]$, defined by equation (2.7), are reported in Table 3.4. They show that it is impossible to compute the inverse of that matrix. In fact, the matrix is singular and badly conditioned; even if the matrix should be positive definite by definition, the numerical approximations yield negative eigenvalues. Then, the simplified EnKF procedure (section 2.1.2) is adopted for all the successive numerical tests.

Quantity	Numerical value
conditioning number	$\approx +5.0 \cdot 10^{+14}$
$\det(\mathrm{Cov}^{(\mathsf{k})}[YY])$	0
$\min_m \lambda_m$	$\approx -2.0 \cdot 10^{-6}$
$\max_m \lambda_m$	$\approx +6.0 \cdot 10^{-4}$

Table 3.4: Numerical results to test the invertibility of the covariance matrix $\mathrm{Cov}^{(\mathsf{k})}[YY]$.

Following the Simplified EnKF procedure, the distribution of parameters are updated at every iterative step, by using arithmetic means of realizations extracted from the first and from the second and third quartiles with the lowest values of $\epsilon_q^{(\mathrm{tot})}$ (see section 2.1.2).

The results obtained for this first test are synthetically represented in Fig. 3.14, where the trend of the ensemble mean and its error bar (± 1 standard deviation of the ensemble mean) of each model parameter is shown as a function of the index k.

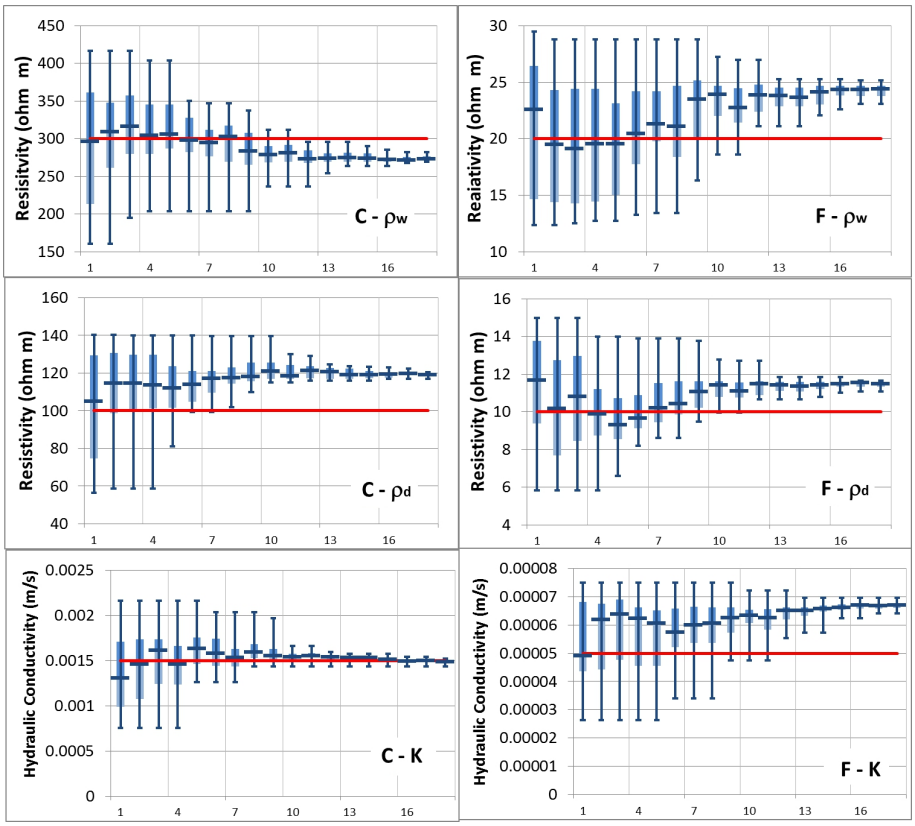


Figure 3.14: Box plot of model parameters as a function of k for the first test. The ensemble mean is represented by an horizontal dark blue bar; the second and third quartiles are represented respectively by light and dark blue boxes; minimum and maximum values are indicated by the thin dark blue error bars. The red lines correspond to the reference values.

After four iterations of the "external" loop of the proposed procedure, the average values of the parameters attain a value which remains approximately constant for the successive iterations. Instead the standard error on this average value continuously decreases. Notice that the asymptotic values for the phenomenological parameters differ from the reference values, but for C-K which is perfectly reproduced: the differences are less than 20% of the reference value. Fig. 3.15 shows the trend of the ensemble average of $\epsilon_q^{(\mathrm{tot})}$ together with the coefficient of variation of $\epsilon_q^{(\mathrm{tot})}$ for a given simulation between two consecutive iterations of the "external" loop. It is clear that simulation #1 is the one with the highest errors, whereas simulations #2 and #3 yield smaller errors. In particular the reduction of the total error after two "external" loops is limited, above all for simulation #1. If the plots of Fig. 3.14 are examined up to $\mathsf{k} = 6$, it is apparent that the first iterations permit to reduce $\epsilon_q^{(\mathrm{tot})}$, by keeping the ensemble of arrays of parameter values close to the reference values. Instead, the procedure successively reduces $\epsilon_q^{(\mathrm{tot})}$ at the expense of a departure of the average values from the reference ones. This is quite common in inverse modelling, when the reduction of the fitting error might cause an over-fitting effect (Schoups et al., 2008)

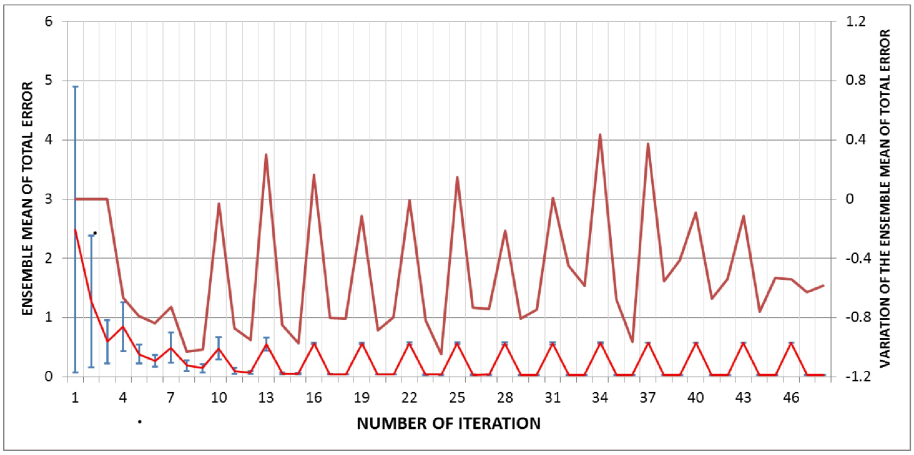


Figure 3.15: Ensemble mean of the total error (left axis) and coefficient of variation of the ensemble mean of the total error with respect to the same simulation at the previous iteration of the "external" loop of the simplified EnKF procedure for the first test.

Initial uniform distribution of the logarithm of empirical parameters

The second test concerned with the extraction of 12 equiprobable sets of the parameters X_l, $l = 1, \ldots, 6$, with a uniform distribution of their logarithm in the interval

$$\left[\log(0.5 \cdot \mathbf{X}_{(\mathrm{ref})_l}), \log(1.5 \cdot \mathbf{X}_{(\mathrm{ref})_l})\right] . \tag{3.4}$$

In other words, the initial values are uniformly distributed on a logarithmic scale. In this case, the average mean used to calculate updated realizations is the geometric mean, that corresponds to the arithmetic mean of the logarithmic values.

Fig. 3.17, as before, shows the trend of the ensemble average of $\epsilon_q^{(\mathrm{tot})}$ together with the coefficient of variation of the total error for a given simulation between two consecutive iterations of the "external" loop. In this second test the convergence criterion of equation (2.8) is satisfied after five iterations of the "external" loop. This strongly reduced the time of computation. Moreover, the average values of the parameters remain approximately constant after three iteration of the external loop and the standard errors on these average values continuously decrease. Nevertheless, notice that the asymptotic values for the phenomenological parameters differ from the reference values, and C-K is no more perfectly reproduced.

Other remarks similar to those given for the first test, are still valid. For instance, Fig. 3.17 shows that simulation #1 is the one with the highest errors, even if after three iterations of the "external loop" the error is negligible. The cause of this fact could be attributed to the position of the electrodic lines with respect to the position of the wells of the domains (see Fig. 3.11): for simulation #1 the electrodic lines is near to the well which have the greatest extraction rate (see Table 3.3); the depression cone due to this stress condition is located beneath the electrodic line (see Fig. 3.12) and highly influences the apparent resistivity measures.

Effects of the order of scenarios in the simplified EnKF

As a third test, the order of the stress condition scenarios has been changed, whereas the other characteristics of the tests are the same as for the second test. The procedure

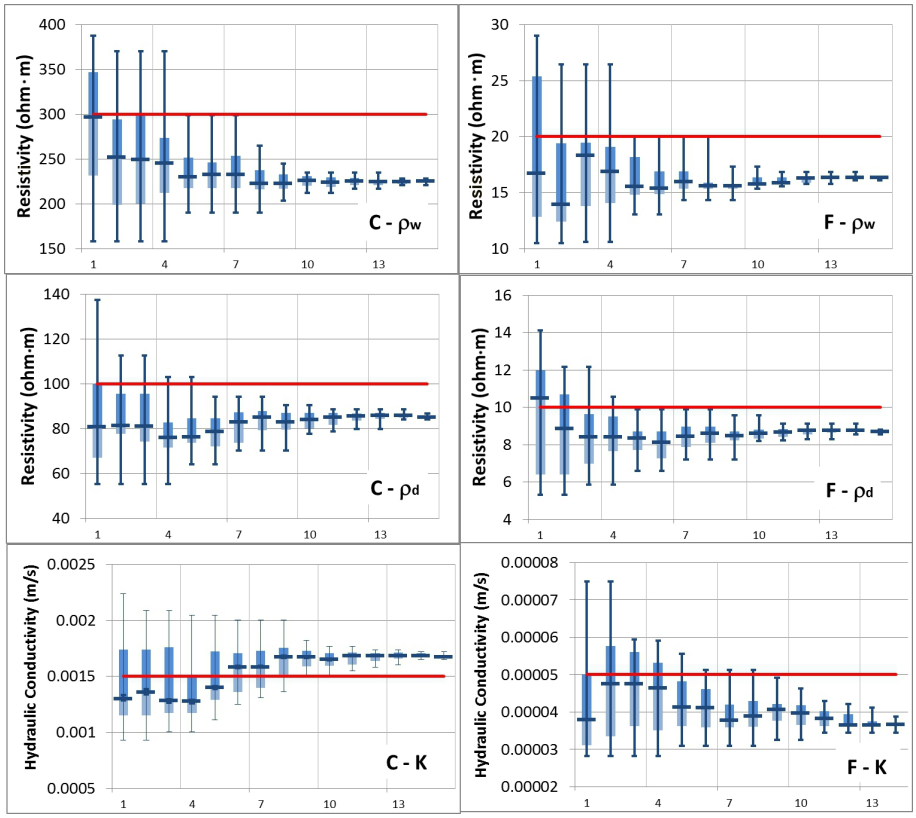


Figure 3.16: Box plot of model parameters as a function of k for the second test. The ensemble mean is represented by an horizontal dark blue bar; the second and third quartiles are represented respectively by light and dark blue boxes; minimum and maximum values are indicated by the thin dark blue error bars. The red lines correspond to the reference values.

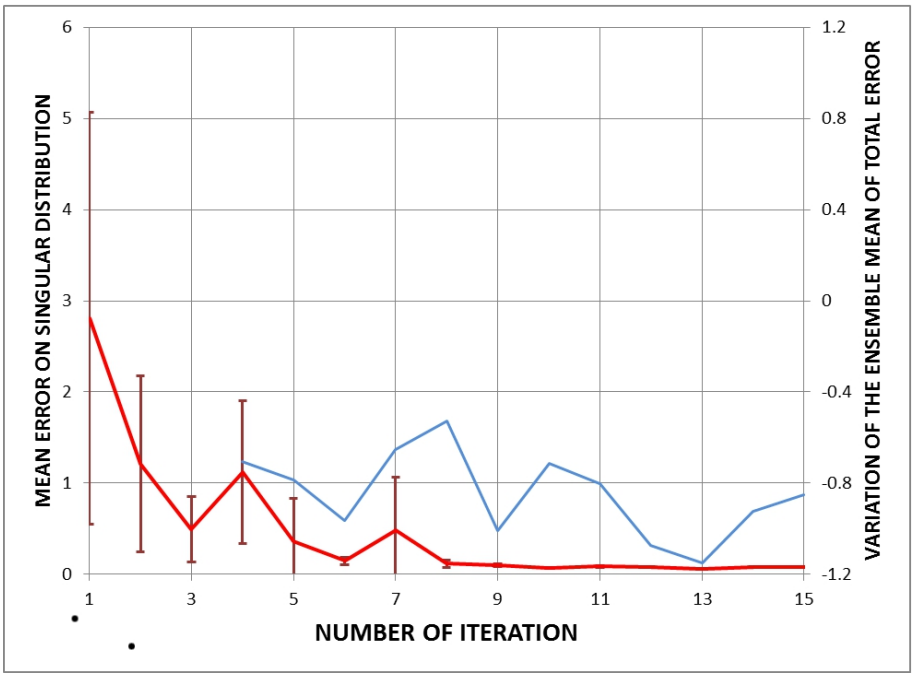


Figure 3.17: Ensemble mean of the total error (left axis) and coefficient of variation of the ensemble mean of the total error with respect to the same simulation at the previous iteration of the "external" loop of the simplified EnKF procedure for the second test.

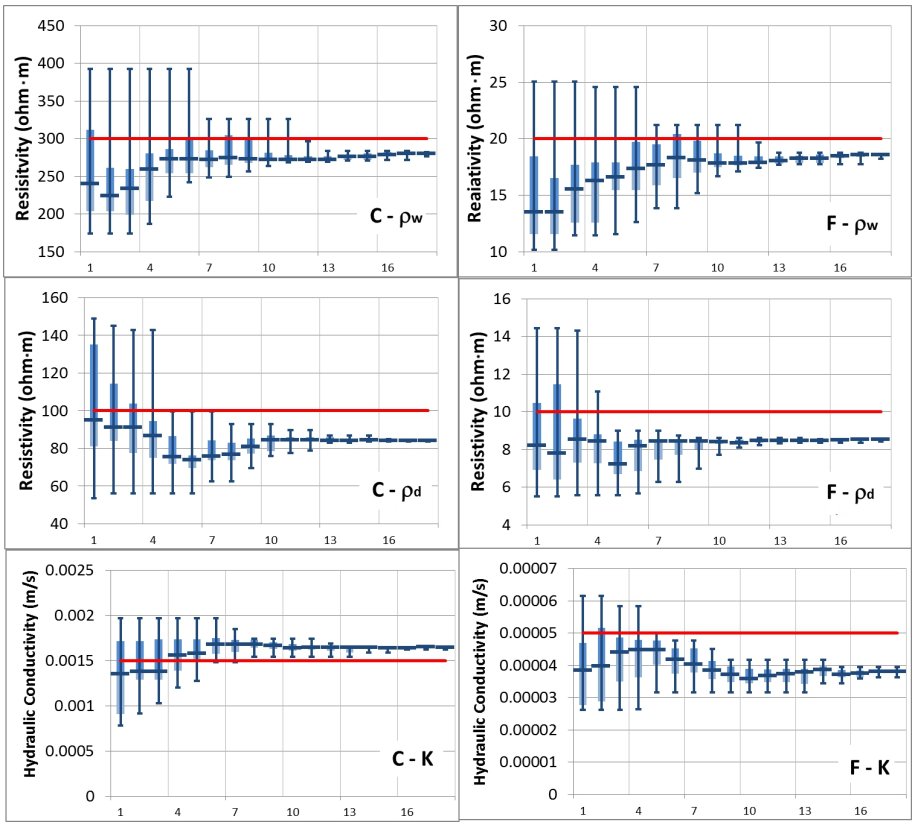


Figure 3.18: Box plot of model parameters as a function of k for the third test. The ensemble mean is represented by an horizontal dark blue bar; the second and third quartiles are represented respectively by light and dark blue boxes; minimum and maximum values are indicated by the thin dark blue error bars. The red lines correspond to the reference values.

considers the following order of simulations: simulation #2, simulation #1 and then simulation #3. The results for this third test are shown in Fig. 3.18. Looking at the mean values for the first two temporal step, comparing Figg. 3.18 and 3.16, it is easy to notice that the two values are closer in this third test than in the second one. The update of the parameter distribution keeps the average values closer to each others, even for the first steps. The position of simulation #1, which is affected by the highest errors, at the second step of the "inner loop" seems to improve the parameter estimation. This could be a joint effect of the small number of initial simulations and of the fact that the simulation considered at $k = 1$ has mostly small errors: this seems to modify the ensemble of realizations in such a way as to keep the median close to the initial values. Moreover, as noticed for the first test, from Fig. 3.19 it is clear that the reduction of the total error after two iterations of the "external" loop is limited, in this case for both simulation #2 and #1.

Effects of a bias in the initial parameter distribution

For the fourth test, 20 initial equiprobable sets of parameters X_l, $l = 1, \ldots, 6$, are extracted from uniform distributions on logarithmic scale, i.e., by extracting $\log X_l$ from

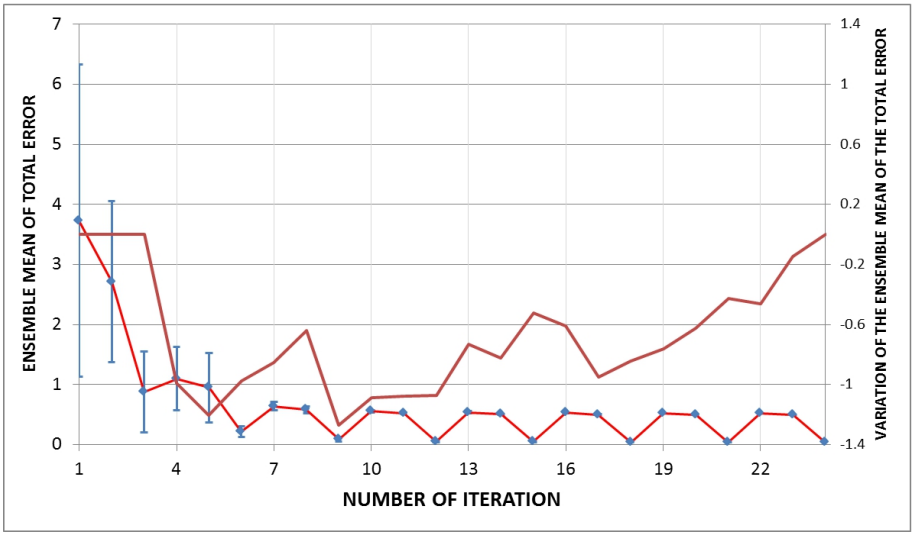


Figure 3.19: Ensemble mean of the total error (left axis) and coefficient of variation of the ensemble mean of the total error with respect to the same simulation at the previous iteration of the "external" loop of the simplified EnKF procedure for the third test.

Input Parameter	Numerical value
X_1 = C-ρ_w	350 ohm $\cdot$ m
X_2 = F-ρ_w	25 ohm $\cdot$ m
X_3 = C-ρ_d	150 ohm $\cdot$ m
X_4 = F-ρ_d	14 ohm $\cdot$ m
X_5 = C-K	0.0013 m/s
X_6 = F-K	0.000065 m/s

Table 3.5: Values used for the generation of the initial parameters for the fourth test.

the interval shown in equation (3.4), but $\mathbf{X}_{(\mathrm{ref})_l}$, $l = \ldots, 6$, have been substituted by the values listed in Table 3.5. In this way a bias in the initial values is introduced. The results, shown in Fig. 3.20, are very encouraging. The asymptotic values for the phenomenological parameters are very similar to those obtained with other initial distributions. The initial shifting does not seem to affect the parameter estimation in a very sensitive way.

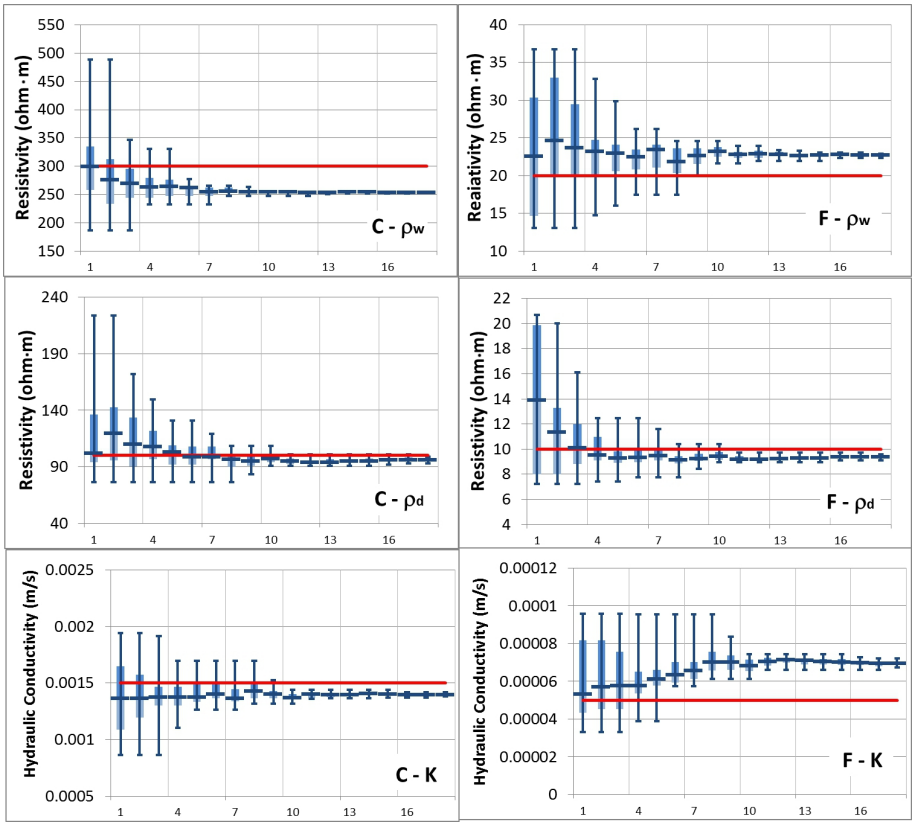


Figure 3.20: Box plot of model parameters as a function of k for the third test. The ensemble mean is represented by an horizontal dark blue bar; the second and third quartiles are represented respectively by light and dark blue boxes; minimum and maximum values are indicated by the thin dark blue error bars. The red lines correspond to the reference values

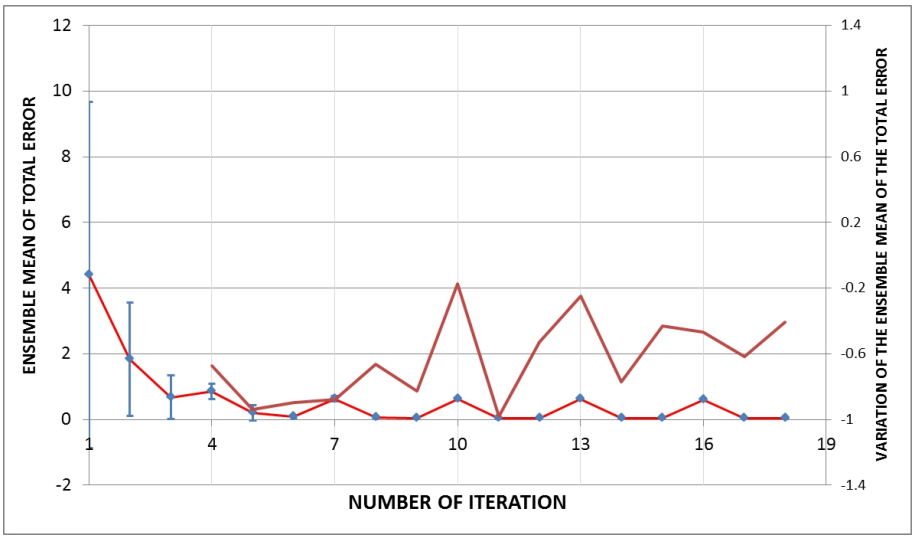


Figure 3.21: Ensemble mean of the total error (left axis) and coefficient of variation of the ensemble mean of the total error with respect to the same simulation at the previous iteration of the "external" loop of the simplified EnKF procedure for the fourth test.

CHAPTER 4

Discussion and conclusions

4.1 Discussion

This section is devoted to a critical analysis of the method developed and tested with this PhD research. In order to facilitate the reading of these final remarks, they are organized by applying a SWOT analysis, a tool which was developed for strategic planning and which is based on the identification of **s**trengths, **w**eaknesses, **o**pportunities and **t**hreats of a project or a business venture. Strengths and weaknesses are related to "internal factors" (the developed method), whereas opportunities and threats are related to "external factors" (the potential use and improvement of the method); on the other hand, strengths and opportunities are related to helpful aspects of the research, whereas weaknesses and threats are related to harmful problems. Roughly speaking, here the SWOT analysis is devoted to answer the following questions:

1. What is good in the methodology developed with this research work (strengths)?
2. Which are the difficulties and the limitations of the method (weaknesses)?
3. Which are the potentials of improvement and upgrading of the method (opportunities)?
4. Which are the possible causes of failure of the method in practical applications (threats)?

For the sake of simplicity, a numbered list of aspects is examined for each of the four items in the following subsections.

4.1.1 Strengths

1. The models YAGMod and YAELMod are based on very similar approaches.
2. Both models adopt finite-differences schemes, which are physically-based, conservative and found on the same basic approximations.
3. `JoMod` permits to run the integrated modelling on parallel computers in a very efficient way.
4. `JoMod` permits to apply both the classical EnKF and the simplified EnKF procedures.
5. `YAGMod` and `YAELMod` share the same discretization structure, so that it is quite easy to prepare input files for both codes.

6. In principle `JoMod` can run with an arbitrary number of input data and of empirical parameters, the limits being imposed by memory and CPU time, only.

7. The results seem to be quite robust with respect to the initialization.

4.1.2 Weaknesses

1. Initial distributions of the tentative parameters should be estimated *a priori*; however (see point 7 of the strengths' list), the method is quite robust with respect to this issue.

2. The initial ensemble must be large enough to span such a big sub-space of the possible values of the parameters, that several possible optimal configurations of parameters could be tested.

3. No physical condition has been identified to introduce a physically-based stopping criterion for the "external loop".

4. The solution to the hydraulic forward problem introduces non-linear effects, which can yield problems with the convergence of the iterative relaxation algorithm used for the solution of the system of non-linear equations. Under some physical situations (e.g., spatial distributions and values of physical parameters, groundwater extraction schedules, etc.) these problems could be related also to the physical inconsistency of the problem setting.

5. YAELMod requires quite a fine grid in order to obtain reliable results. Unfortunately it is difficult to provide prior estimates of the optimal spacing.

6. Despite the possibility of running `JoMod` with multi-processors computers, some physical situations could require a long CPU time, especially for `YAGMod`, and they could constitute a bottleneck for the execution time.

7. The models YAGMod and YAELMod are presently developed for stationary conditions only.

4.1.3 Opportunities

1. In principle it is very easy to modify `JoMod` in order to run different models as alternatives to YAGMod and YAELMod: this innovation would require tools for I/O formatting.

2. It is also very easy to adapt `JoMod` in order to take into account different functional dependencies of K and ρ on φ and S, i.e., more complex functional forms of the functions $\varkappa$ and ϱ.

3. In order to overcome the difficulties mentioned at point 5 of the weaknesses list, `YAELMod` could be modified in order to work on a grid which is refined with respect to that used for `YAGMod`.

4. Different kinds of measured and modelled quantities could be considered. Moreover, also solute transport in groundwater could be added.

5. It is quite easy to test different initial distributions of the empirical parameters.

6. Since several different configurations are tested during `JoMod` runs, such a code could be modified in order to profit from the big mass of available results from `YAGMod` and `YELMod` and perform a thorough sensitivity analysis of model output on the empirical parameters.

4.1.4 Threats

1. The use of complex functions $\varkappa$ and ϱ, which take into account in a better way the link between saturation and physical parameters, or the use of a large number of HGFs could make the procedure difficult to be applied, because field data of very high quality would be necessary, CPU times could be very long, and identifiability of the individual phenomenological parameters could become a very limiting issue.

2. The simplified EnKF yields an artificially low standard deviation around the average value if the procedure is not stopped after few external iterations. This could deceive not-well-educated professionals (see point 3 of the weaknesses' list).

4.2 Conclusions

The goal of this PhD research was to develop a tool that could be useful for the characterization of alluvial aquifers in order to improve the knowledge on the major features governing groundwater flow and contaminant transport in practical applications. The method for joint modelling and inversion of hydrological and electrical data proposed in this work has been tested on synthetic examples and the results support the strengths that have been summarised in the previous section.

It is worth adding that the method is based on field data whose acquisition requires techniques that are quite standard and almost routine for professionals. Therefore, the method could be useful in practical applications, if it is possible to organise the water pumping schedule in such a way as to stress the aquifer under study in such a way as to produce independent sets of data with a "tomographic-like" approach.

Finally, the method is very attractive for its potential future improvements (see the list of opportunities listed in the previous section).

Bibliography

Anderson, M., 2009. Hydrogeological facies models to delineate large-scale spatial trends in glacial and glaciofluvial sediments. Geological Society of America Bulletin 101, 501–511.

Archie, G. E., 1942. The electrical resistivity log as an aid in determining some reservoir characteristics. Trans. Vol. 146, 56–62.

Associazione Irrigazione Est Sesia, 1979. Le acque sotterranee della pianura irrigua Novarese-Lomellina. Studi e ricerche per la realizzazione di un modello gestionale (Subsurface water of irrigated Novara-Lomellina plain. Study and research for a managing model realization.). Report.

Baines, D., Smith, D., Froese, D., Bauman, P., Nimeck, G., 2002. Electrical resistivity ground imaging (ergi): a new tool for mapping the lithology and geometry of channelbelts and valleyfills. Sedimentology, 49, 441–449.

Beatrizotti, G., Hansen, W. J., Spocci, R., 1983. Optimierung der Benötigten daten für ein numerisches Modell der Grundwasserbewirtschaftung im Lockergestein (Optimization of the required data for a numerical model of groundwater management in soil). Gas Wasser Abwasser 63, 469–476.

Bedekar, V., Niswonger, R. G., Kipp, K., Panday, S., Tonkin, M., 2012. Approaches to the Simulation of Unconfined Flow and Perched Groundwater Flow in MODFLOW. Groundwater 50, 187–198.

Benoit, N., Pasquier, P., Marcotte, D., Nastev, M. (Eds.), 2005. Conditional stochastic inverse modelling of the Châteauguay river aquifers. MODELCARE 2005, prepublished proceedings, 515-521.

Bernabe, Y., Revil, A., 1995. Pore-scaleh eterogeneity, energy dissipation and the transport properties of rocks. Geophysical Research Letters, 22(12), 1529–1532.

Bersezio, R., Bini, A., Giudici, M., 1999. Effects of sedimentary heterogeneity on groundwater flow in a quaternary pro-glacial delta environment: joining facies analysis and numerical modelling. Sedimentary Geology 129, 327–344.

Böhm, G., Brauchler, R., Nieto, D. Y., Baradello, L., Affatato, A., Sauter, M., 2013. A field assessment of site-specific correlations between hydraulic and geophysical parameters. Near Surface Geophysics, doi:10.3997/1873-0604.2013034.

Biella, G., Lozej, A., Tabacco, I., 1983. Experimental study of some hydrogeophysical properties of unconsolidated porous media. GROUND WATER Vol. 21 - No.6, 741–751.

Binley, A., 2013a. R3t 1.8,.
URL `http://www.es.lancs.ac.uk/people/amb/Freeware/R3t/R3t.htm`

Binley, A., 2013b. R3t version 1.8. Lancaster University, User Manual.

Binley, A.and Cassiani, G., Deiana, R., 2010. Hydrogeophysics: opportunities and challenges. Bollettino Di Geofisica Teorica Ed Applicata Vol. 51, 267–284.

Bridge, J., Hyndman, D., 2004. Aquifer characterization. Vol. 80. In: Aquifer characterization (eds. J.S. Bridge and D.S. Hyndman).

Börner, F. D., Schön, J., 1991. A relation between the quadrature component of electrical conductivity and the specific surface area of sedimentary rocks. The Log Analyst Vol. 32, 612–613.

Camporese, M., Cassiani, G., Deiana, R., Salandin, P., 2011. Assessment of local hydraulic properties from electrical resistivity tomography monitoring of a three-dimensional synthetic tracer test experiment. WATER RESOURCES RESEARCH, doi:10.1029/2011WR010528.

Carman, P. C., 1939. Permeability of saturated sands, soils and clays. The Journal of Agricultural Science Vol. 29, 263–273.

Cattaneo, L., Vassena, C., Giudici, M., Petrucci, B., 2013. Modelling groundwater recharge in an alluvial aquifer of Somaliland with the groundwater flow model YAGMOD - Modellizzazione della ricarica in un acquifero alluvionale del Somaliland con il modello di flusso idrico sotterraneo YAGMOD. Acque Sotterranee - Italian Journal of Groundwater, doi: 10.7343/AS-018-13-0042.

Cavalli, E., 2011. Messa a punto di una metodologia per la modellazione tridimensionale e multiscala dell'idrostratigrafia, su base GIS. Università degli Studi di Milano, Ph.d. Thesis.

De Filippis, G., Giudici, M., Margiotta, S., Mazzone, F., Negri, S., Vassena, C., 2013. Modellazione numerica del flusso idrico nell'acquifero fratturato e carsico della penisola Salentina (Sud Italia) - Numerical modeling of the groundwater flow in the fractured and karst aquifer of the Salento peninsula (Southern Italy). Acque Sotterranee - Italian Journal of Groundwater, doi : 10.7343/AS-016-013-0040.

De Franco, R., Biella, G., Tosi, L., Teatini, P., Lozej, A., Chiozzotto, B., Giada, M., Rizzetto, F., Claude, C., Mayer, A., et al., 2009. Monitoring the saltwater intrusion by time lapse electrical resistivity tomography: The chioggia test site (venice lagoon, italy). Journal of Applied Geophysics Vol. 69 - n. 3, 117–130.

de Lima, O., Sharma, M., 1990. A grain conductivity approach to shaly sandstones. GEOPHYSICS Vol. 55, 1347–1356.

de Limaa, O. A. L., Sri Niwas, 2000. Estimation of hydraulic parameters of shaly sandstone aquifers from geoelectrical measurements. Journal of Hydrology, Vol. 235, 12–26.

de Marsily, G., 1986. Quantitative Hydrogeology - Groundwater hydrology for engineers. Academic Press.

Dell'Oro, M., 2011. Modellazione 3D per l'interpretazione di indagini geoelettriche non convenzionali. Università degli Studi di Milano, master Thesis.

Deutsch, C. V., Journel, A. G., 1992. Geostatistical Software Library and User's Guide. Oxford University press, New York.

Dey, A., Morrison, H. F., 1979. Resistivity modeling for arbitrary shaped three-dimensional structures. GEOPHYSICS Vol. 44, 753–780.

Di Maio, R., Piegari, E., 2011. Water storage mapping of pyroclastic covers through electrical resistivity measurements. Journal of Applied Geophysics 75, 196–202.

Doherty, J., 2001. Improved Calculations for Dewatered Cells in MODFLOW. Groundwater 39, 863–869.

Doussan, C., Ruy, S., 2009. Prediction of unsaturated soil hydraulic conductivity with

electrical conductivity. WATER RESOURCES RESEARCH, 45.

Ellis, D., Singer, J., 2007. Well Logging for Earth Scientists. Springer.

Evensen, G., 2009. The ensemble kalman filter for combined state and parameter estimation - monte carlo techniques for data assimilation in large system. IEEE Control System Magazine, doi: 10.1109/MCS.2009.932223.

Felletti, F., Bersezio, R., Giudici, M., 2006. Geostatistical simulation and numerical upscaling, to model ground water flow in a sandy-gravel, braided river, aquifer analogue. Journal of Sedimentary Research 76, 1215–1229, doi:10.2110/jsr.2006.091.

Ferraresi, M., Todini, E., Vignoli, R., 1996. A solution to the inverse problem in groundwater hydrology based on kalman filtering. Journal of Hydrology Vol. 175, 567–58.

Feynman, R., Leighton, R., Sands, M., 1989. Feynman Lectures on Physics, Mainly Electromagnetism and Matter. Addison-Wesley., ISBN 0-201-51003-0.

Fowler, E. D., Moysey, M. J. S., 2011. Estimation of aquifer transport parameters from resistivity monitoring data within a coupled inversion framework. Journal of Hydrology, doi:10.1016/j.jhydrol.2011.08.063.

Frohlich, R. K.and Fisher, J., Summerly, E., 1996. Electric-hydraulic conductivity correlation in fractured crystalline bedrock: Central Landfill, Rhode Island, USA. Journal of Applied Geophysics Vol. 35, 249–259.

Galloway, W., Sharp, J., 1998. Characterizing aquifer heterogeneity within terrigenous clastic depositional systems., 85–90In: Hydrogeologic models of sedimentary aquifers (eds. F. J.S. and J.M. David),.

Geotomo Software, 2001. RES3DMODx64 ver. 3.04 with 3D I.P. modeling using complex resistivity method.
URL `http://www.geotomosoft.com/`

Gillijns, S., Barrero Mendoza, O., Chandrasekar, J., De Moor, B. L. R., Bernstein, D. S., Ridley, A., 2006. What Is the Ensemble Kalman Filter and How Well Does it Work? Proceedings of the 2006 American Control Conference.

Giudici, M., 2010. Modeling water flow and solute transport in alluvial sediments: scaling and hydrostratigraphy from the hydrological point of view. Memorie descrittive della Carta Geologica d'Italia Vol. XC, 113–119, in: Proceedings of the Second National Workshop '"Multidisciplinary approach for porous aquifer characterization" (eds. R. Bersezio and M. Amanti).

Giudici, M., Foglia, L., Parravicini, G., Ponzini, G., Sincich, B., 2000. A quasi three dimensional model of water flow in the subsurface of Milano (Italy): the stationary flow. Hydrology and Earth System Sciences 4, 113–124.

Giudici, M., Margiotta, S., Mazzone, F., Negri, S., Vassena, C., 2012. Modeling hydrostratigraphy and groundwater flow of a fractured and karst aquifer in a Mediterranean basin (Salento peninsula, southeastern Italy). Environmental Earth Sciences, doi:10.1007/s12665-012-1631-1.

Giudici, M., Ponzini, G., Romano, E., Vassena, C., 2007. Some lessons from modelling ground water flow in the metropolitan area of Milano (Italy) at different scales. Memorie Descrittive della Carta Geologica d'Italia 76, 207–218.

Hayley, K., Bentley, L. R., Gharibi, M., 2009. Time-lapse electrical resistivity monitoring of salt-affected soil and groundwater. WATER RESOURCES RESEARCH.

Hill, M. C., Tiedeman, C. R., 2007. Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty. J. Wiley & Sons, New York.

Hinnell, A. C., Ferré, T. P. A., Vrugt, J. A., Huisman, J. A., Moysey, S., Rings, J., Kowalsky, M. B., 2010. Improved extraction of hydrologic information from geo-

physical data through coupled hydrogeophysical inversion. WATER RESOURCES RESEARCH, doi:10.1029/2008WR007060.

Hubbard, S., Rubin, Y., 2005. Introduction to Hydrogeophysics. Springer, in: Hydrogeophysics (eds. Y. Rubin and S.S. Hubbard).

Hubbard, S. S., Rubin, Y., 2000. Hydrogeological parameter estimation using geophysical data: a review of selected techniques. Journal of Contaminant Hydrology Vol. 45, 3–34.

Huggenberger, P., T., A., 1999. Introduction to the special issue on aquifer-sedimentology: problems, perspectives andmodern approaches. Sedimentary Geology, 179–186.

Jardani, A., Revil, A., Dupont, J., 2013. Stochastic joint inversion of hydrogeophysical data for salt tracer test monitoring and hydraulic conductivity imaging. Advanced in Water Resources, doi:10.1016/j.advwatres.2012.08.005.

Kalman, R. E., 1960. A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME–Journal of Basic Engineering, Vol. 82 (Series D), 35–45.

Katz, A. J., Thompson, A., 1986. Quantitative prediction of permeability in porous rock. Physical Review B 34(11), 8179–8181.

Keating, E., Zyvoloski, G., 2009. A stable and efficient numerical algorithm for unconfined aquifer analysis. Groundwater 47, 569–579.

Keller, G., Frischknecht, F., 1966. Electrical methods in geophysical prospecting. Pergamon Press.

Kelly, W. E., 1977. Geoelectric sounding for estimating aquifer hydraulic conductivity. GROUND WATER Vol. 15, 420–425.

Klein, J., Sill, W., 1982. Electrical properties of artificial claybearing sandstone. GEOPHYSICS Vol. 47, 1593–1601.

Klingbei, R., Kleineidam, S., Asprion, U., Aigner, T., Teutsch, G., 1999. Relating lithofacies to hydrofacies: outcrop-based hydrogeological characterisation of Quaternary gravel deposits. Sedimentary Geology 129 (3–4), 299 – 310, doi:10.1016/S0037-0738(99)00067-6.

Kosinky, W. K., Kelly, W. E., 1981. Geoelectric sounding for predicting aquifer properties. GROUND WATER Vol. 19, 163–171.

Kozeny, J., 1927. Ueber kapillare Leitung des Wassers in Boden. Akademie der Wissenschaften in Wien. Sitzungsberichte. Mathematisch-naturwissenschaftliche Klasse. Abteilung 1 Vol. 136, 271–306, Wien.

Lesmes, D. P., Friedman, S. P., 2005. Relationships between the electrical and hydrogeological properties of rocks and soils. Hydrogeophysics.

Lin, H. J., Richards, D. R., Talbot, C. A., Yeh, G., Cheng, J., Cheng, H., Jones, N. L., 1997. FEMWATER: A Three-Dimensional Finite Element Computer Model for Simulating Density-Dependent Flow and Transport in Variably Saturated Media. Report CHL-97-12.

Lochbuehler, T., Doetsch, J., Brauchler, R., Linde, N., 2013. Structure-coupled joint inversion of geophysical and hydrological data. GEOPHYSICS, doi:10.1190/geo2012-0460.1.

Locke, M. H., 2001. RES3DMOD ver. 2.1 for Windows 95/98/Me/2000/NT, 3-D resistivity and IP forward modeling using the finite-difference and finite-element methods. User Manual.

Lowry, T., Allen, M., Shive, P. N., 1989. Singularity removal: A refinement of resistivity modeling techniques. GEOPHYSICS Vol. 54, 766–774.

Lunati, I., Bernard, D., Giudici, M., Parravicini, G., Ponzini, G., 2001. A numerical com-

parison between two upscaling techniques: non-local inverse based scaling and simplified renormalization. Advances in Water Resources 24, 913–929.

Matheron, G., 1965. Les variables regionaliséés et leur estimation: une application de la théorie des fonctions aléatoires aux sciences de la nature. Masson et Cie, Paris.

Mele, M., Bersezio, R., Giudici, M., 2012. Hydrogeophysical imaging of alluvial aquifers: electrostratigraphic units in the quaternary po alluvial plain (italy). International Journal of Earth Sciences Vol. 101 - n. 7, 2005–2025.

Mele, M., Inzoli, S., Giudici, M., Bersezio, R., 2014. Relating electrical conduction of alluvial sediments to textural properties and pore-fluid conductivity. Geophysical Prospecting, doi: 10.1111/1365-2478.12102.

Michot, D., Benderitter, Y., Dorigny, A., Nicoullaud, B., King, D., Tabbagh, A., 2003. WATER RESOURCES RESEARCH, Vol. 39(5).

MIKE, 2003. Danish Hydraulic Institute Water & Environment, MIKE by - DHI software. http://www.mikebydhi.com/.

Minotta, N., 2002. Interventi per il miglioramento della qualità dell'acqua negli acquedotti con presenza di nitrati - primo lotto. relazione tecnica generale. comune di pozzuolo martesana. Progetto n.2056 del PT 1996.
URL `http://www.cartografia.regione.lombardia.it/silvia/jsp/schede/schedaSintesi.jsf?idProcedura=030100000891&titolo=VIA%20Regionale%20elenco%20studi%20per%20categoria%20di%20opera%20%20Derivazioni%20di%20acque%20sotterranee`

MODFLOW, 2005. MODFLOW-2005 Version 1.8, Three-dimensional finite-difference ground-water model, U.S. Geological Survey, Reston, Virginia. http://water.usgs.gov/nrp/gwsoftware/modflow2005/modflow2005.html.

Narashiman, T., Witherspoon, P. A., 1976. An integrated finite-difference method for analyzing fluid flow in porous media. Water Resources Research 12, 57–64.

Neuman, S. P., Di Federico, V., 2003. Multifaceted nature of hydrogeologic scaling and its interpretation. Reviews of Geophysics 41 (3), doi:10.1029/2003RG000130.
URL `http://dx.doi.org/10.1029/2003RG000130`

Nimmer, R. E., Osiensky, J. L., Binley, A. M., Sprenke, K. F., Williams, B. C., 2007. Electrical resistivity imaging of conductive plume dilution in fractured rock. Hydrogeology Journal.

Niswonger, R. G., Panday, S., Ibaraki, M., 2011. MODFLOW-NWT, A Newton Formulation for MODFLOW-2005. Groundwater Resources Program.

Niwas, S., de Lima, O. A. L., 2002. Aquifer parameter estimation from surface resistivity data. GROUND WATER Vol. 41 - No.1, 94–99, techical Note.

Perulero Serrano, R., Guadagnini, L., Riva, M., Giudici, M., Guadagnini, A., 2014. Impact of two geostatistical hydro-facies simulation strategies on head statistics under non-uniform groundwater flow. Journal of Hydrology 508, 343 – 355, doi:10.1016/j.jhydrol.2013.11.009.
URL `http://www.sciencedirect.com/science/article/pii/S0022169413008251`

Pidlisecky, A., Haber, E., Knight, R., 2007. Resinvm3d: A 3d resistivity inversion package. GEOPHYSICS Vol. 72, NO. 2, 1–10.

Pidlisecky, A., Haber, E., Knight, R., 2013. 2007-0001: Resinvm3d.
URL `http://software.seg.org/2007/0001`

Pollock, D., Cirpka, O. A., 2010. Fully coupled hydrogeophysical inversion of synthetic salt tracer experiments. WATER RESOURCES RESEARCH,

doi:10.1029/2009WR008575.

Ponzini, G., Crosta, G., 1988. The comparison model method: a new arithmetic approach to the discrete inverse problem of groundwater hydrology, 1, One-dimensional flow. Transport in Porous Media 3, 415–436.

Ponzini, G., Crosta, G., Giudici, M., 1989b. The hydrogeological role of an aquitard in preventing drinkable water well contamination: a case study. Environmental Health Perspectives 83, 77–95.

Ponzini, G., Lozej, A., 1982. Identification of aquifer transmissivities: the comparison model method. Water Resources Research 18, 597–622.

Purvance, D. T., 2003. Roles apparent resistivity amplitude and phase play in an aquifer's electrical-hydraulic conductivity correlation. GEOPHYSICAL RESEARCH LETTERS, doi:10.1029/2002GL016131.

Purvance, D. T., Andricevic, R., 2000. On the electrical-hydraulic conductivity correlation in aquifers. WATER RESOURCES RESEARCH, Vol. 36, NO. 10, 2905–2913.

Regione Lombardia, Eni Divisione Agip, 2001. Geologia degli acquiferi Padani della Regione Lombardia. S.EL.CA.

Remy, N., Boucher, A., Wu, J., 2011. Applied Geostatistics with SGeMS: A User's Guide. Cambridge University Press, ISBN-13: 978-0521514149, ISBN-10: 0521514142.

Rencher, A. C., Christensen, W. F., 2012. Methods of Multivariate Analysis, 3rd Edition. John Wiley and Sons, New York., ISBN: 978-0-470-17896-6.

Reynolds, J., 2011. An Introduction to Applied and Environmental Geophysics. Wiley-Blackwell, ISBN 978–0–471–48536–0.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global Sensitivity analysis – The Primer. John Wiley & Sons, Chichester, UK.

Scarascia, S., Ponzini, G., 1972. An approximate solution for the inverse problem in hydraulics. L'Energia Elettrica 49, 518–531.

Schön, J., 2004. Physical Properties of Rocks: Fundamentals and Principles of Petrophysics. Elsevier, ISBN 9780080443461.

Schoups, G., van de Giesen, N. C., Savenije, H. H. G., 2008. Model complexity control for hydrologic prediction. Water Resources Research 44 (12), n/a–n/a.
URL `http://dx.doi.org/10.1029/2008WR006836`

Silvester, P., Ferrari, R., 1990. Finite elements for electrical engineers (2nd.ed.).

Slater, L., 2007. Near surface electrical characterization of hydraulic conductivity: from petrophysical properties to aquifer geometries - A review. Survey in Geophysics Vol. 28, 169–197.

Slater, L., Lesmes, D. P., 2002. Electrical-hydraulic relationships observed for unconsolidated sediments. WATER RESOURCES RESEARCH, doi:10.1029/2001WR001075.

Sènèchal, P., Sènèchal, G., 2009. Relationships between water flow rate and geophysical measurements in an alluvial aquifer. Acta Geophysica doi: 10.2478/s11600-009-0024-7.

Soupios, P., Kouli, M., Vallianatos, F., Vafidis, A., Stavroulakis, G., 2007. Estimation of aquifer hydraulic parameters from surficial geophysical methods: A case study of keritis basin in chania (crete – greece). Journal of Hydrology, doi:10.1016/j.jhydrol.2007.02.028.

Straface, S., Chidichimo, F., Rizzo, E., Riva, M., Barrash, W., Revil, A., Cardiff, M., Guadagnini, A., 2011. Joint inversion of steady-state hydrologic and self-potential data for 3d hydraulic conductivity distribution at the boise hydrogeophysical research site. Journal of Hydrology, doi:10.1016/j.jhydrol.2011.07.013.

SUTRA, 2010. SUTRA: A Model for Saturated-Unsaturated Variable-Density Ground-Water Flow with Solute or Energy Transportl, U.S. Geological Survey, Reston, Virginia.

Telford, W., Geldart, L., Sheriff, R., 1990. Applied geophysics. Cambridge University Press, ISBN 9780521339384.

Thompson, A. H., Katz, A. J., Krohn, C. E., 1987. The microgeometry and transport properties of sedimentary rock. Advances in Physics 36, 625– 694.

TOUGH2, 1991. TOUGH2 Software Qualification, Lawrence Berkeley National Laboratory, Earth Sciences Division, berkeley, California. http://esd.lbl.gov/research/projects/tough/.

Trefry, M. G., Muffels, C., 2007. FEFLOW: A Finite-Element Ground Water Flow and Transport Modeling Tool. doi:10.1111/j.1745-6584.2007.00358.x.

Valota, G., Giudici, M., Parravicini, G., Ponzini, G., Romano, E., 2002. Is the forward problem of ground water hydrology always well posed? Ground water 40, 500–508.

Vassena, C., Cattaneo, L., Giudici, M., 2010. Assessment of the role of facies heterogeneity at the fine scale by numerical transport experiments and connectivity indicators. Hydrogeology Journal 18, 651–668, doi:10.1007/s10040-009-0523-2.

Vassena, C., Durante, C., Giudici, M., Ponzini, G., 2008. The importance of observations on fluxes to constrain ground water model calibration. Physics and Chemistry of the Earth, Parts A/B/C 33, 1105–1110, doi:10.1016/j.pce.2008.01.004.

Vassena, C., Rienzner, M., Ponzini, G., Giudici, M., Gandolfi, C., Durante, C., Agostani, D., 2011. Modeling water resources of an highly irrigated alluvial plain: coupling and calibrating soil and ground water models. Hydrogeology Journal, 20, 449–467, doi:10.1007/s10040-011-0822-2.

Wilkinson, P., Meldrum, P., Kuras, O., Chambers, J. E., Holyoake, S. J., Ogilvy, R. D., 2010. High-resolution electrical resistivity tomography monitoring of a tracer test in a confined aquifer. Journal of Applied Geophysics.

Worthington, P., 1993. The uses and abuses of the archie equations. 1. the formation factor porosity relationship. Journal of Applied Geophysics 30, 215–228.

Young, D. M., 1971. Iterative solution of large linear systems. Academic Press, New York, NY, USA.

Zappa, G., Bersezio, R., Felletti, F., Giudici, M., 2006. Modeling heterogeneity of gravel-sand, braided stream, alluvial aquifers at the facies scale. Journal of Hydrology 325, 134–153.

Zhao, S., Yedlin, M., 1996. Some refinements on the finite-difference method for 3-D dc resistivity modeling. GEOPHYSICS Vol. 61, 1301–1307.

Zhou, Q. I., Shimada, J., Sato, A., 2001. Three-dimensional spatial and temporal monitoring of soil water content using electrical resistivity tomography. WATER RESOURCES RESEARCH, Vol. 37, 273 – 285.