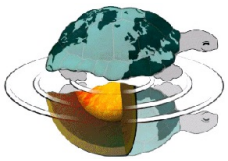
UNIVERSITÀ DEGLI STUDI DI MILANO



Dottorato di Ricerca in Scienze della Terra
Ciclo XXVIII

Calibration of the groundwater flow model and assessment of the saltwater intrusion in a multi-layered aquifer system of the Ionian coastal area (Taranto gulf, southern Italy)

PhD Thesis

Giovanna De Filippis

ID nr. R10219

Tutor	**Academic Year**	*Coordinator*
Prof. Mauro Giudici	**2014-2015**	**Prof. Elisabetta Erba**

Contents

Research Aim

In some Mediterranean karst areas, groundwater is often the only available supply for freshwater. Besides the contamination induced by human activities, coastal aquifers often suffer from the saltwater intrusion phenomenon, which can be enhanced by both extensive withdrawals and climatic changes. Establishing an effective set of regulatory and management measures to ensure the sustainability of coastal aquifers requires a deep knowledge about natural and anthropic stresses involved in groundwater dynamics. In this regard, a prior conceptualization of aquifer systems and a deeper characterization of balance terms through mathematical modelling are of paramount importance.

In the gulf of Taranto (southern Italy), these issues are particularly pressing, as the multi-layered, carbonatic aquifer is the only available resource of freshwater and satisfies most of the human water-related activities. Especially during the last decades, proper management plans and decisions seem to be compelling, as the national government included Taranto in the list of the contaminated sites of national importance, due to the presence of highly-polluting activities nearby the Mar Grande and Mar Piccolo seawater bodies, whose relationship with the underground resources is matter of concern, as they host important freshwater springs. Furthermore, the Taranto area is particularly sensitive to the phenomenon of seawater intrusion, both for the specific hydrostratigraphic configuration and for the presence of highly water-demanding industrial activities. These problems, strictly related to the protection and preservation of groundwater quality and quantity, have triggered several actions. Among them, the Flagship Project RITMARE (la Ricerca Italiana per il Mare - the Italian Research for the Sea) took into account criticalities involving several environmental components within the Mar Piccolo ecosystem, including groundwater.

In this thesis, a full charactrization of the multi-layered aquifer system of the whole Province of Taranto is presented, with the purpose of supporting monitoring activities, land-use plans and management decisions. The preliminary outcomes refer to the identification of the conceptual model, namely the reconstruction of the hydrostratigraphic structure of the underground and the qualitative assessment of the groundwater dynamics. The successive development of a numerical model permits to produce a tool for quantifying the hydrogeological balance and simulating the system response to climate or man-induced changes.

Generally speaking, thorough evaluation of model adequacy and/or accuracy is an important step in the study of environmental systems, due to the uncertainties on hydrodynamic properties and boundary conditions and to the scarcity of good-quality field data. This commonly results in groundwater models being calibrated and often leads to the development of many candidate models that can differ in the

analysed processes, representation of boundary conditions, distribution of system characteristics, and parameter values. In this framework, calibration of alternative models allowed to identify the main challenges which limit the reliability of model outcomes and test model adequacy while proposing a new calibration methodology, which represents tha major scientific contribution of this thesis.

Chapter 1

Introduction

1.1 Scientific background and motivations

Freshwater is one of the most universally required natural resource for human life, either at an individual level or at the community level (e.g., for irrigation, industrial and tourism activities, sanitation, electricity production), or for the preservation of ecosystems as a whole [7]. For this reason, its availability is a crucial component in socio-economic development, because freshwater from springs and other sources has always played a major role for human settlement and social development [31].
Since freshwater represents a little fraction only of the total water reserves on Earth and approximately three-quarters of this amount is frozen in the cryosphere, the only available freshwater to provide for human well-being is stored in surface water bodies or in the underground. In particular, groundwater is, by far, the world's largest source of fresh accessible water, representing over 95% of the world's available freshwater reserves and supplying, among the others, basic drinking and sanitation needs. In fact, where aquifers are available, groundwater is normally preferred over surface sources of water supply because, with respect to surface water bodies (e.g., lakes and rivers), it is relatively well protected from pollutant sources and is less influenced by droughts and climate changes [74].
Focusing on groundwater, there is evidence of important changes due to extensive withdrawals to supply human activities (e.g., industry, tourism, irrigation) [93]. In addition, society is currently faced with unparalleled challenges imposed by pollution, ecosystem degradation, water scarcity and insecurity due to climate variability, growing population pressure and/or conflict among water users: all these factors are expected to potentially affect the ecosystem functioning [7]. For this reason, due to its vital significance for the whole socio-economic sector, groundwater is seen in recent years as a vulnerable natural resource to be protected and managed in an environmentally sustainable way, satisfying the present-days needs without compromising the ability of future generations to meet their own needs [7, 30]. This is particularly true in heavily urbanized coastal Mediterranean areas, where the request for freshwater is very acute, due to a high concentration of water-demanding activities. Here coastal aquifers sometimes are the only available water resources, which largely cover the demand for drinking water in most countries and contribute to a more variable extent to the supply of irrigation water, so that extensive withdrawals

of freshwater from the subsurface can cause concern in terms of groundwater quality and quantity.
Freshwater in coastal aquifers is very susceptible to degradation, due to its proximity to the sea, in combination with the intensive water demands by the increasing population in coastal zones [7, 37]. In fact, freshwater and seawater are miscible fluids: under natural undisturbed conditions, a state of equilibrium is maintained with a stationary contact zone and the freshwater flowing towards the sea; however, permanent or temporary perturbations of the equilibrium due to groundwater withdrawals or changes in aquifer recharge require special attention and special management techniques, in view of the danger of seawater intrusion phenomenon, namely the motion of seawater inland.
Saltwater intrusion is enhanced by prolonged changes in coastal groundwater levels, due to anthropic (e.g., overexploitation of coastal aquifers, land-use modification) or natural (e.g., variations of the mean sea level, climate change) factors. These can cause the water table of the aquifer in the vicinity of the coast to drop so that the hydraulic head in the freshwater body becomes lower than that in the adjacent seawater wedge. In this case, the interface advances inland until a new equilibrium is reached and, when the advancing interface reaches pumping wells, the salt content of groundwater could exceed threshold levels prescribed to avoid damages to human health and activities (e.g., agriculture). So the primary effects of seawater intrusion are a reduction in the available freshwater storage volume and a degradation of freshwater quality [141]. Since original salinity values in aquifers can be restored over decades or even centuries, a correct policy of exploitation of groundwater resources and monitoring activities are of paramount importance.
These effects are severe in fractured and karst aquifers [37], as water rapidly infiltrates through the systems of discontinuities in the soluble rock mass, so creating the underground network of conduits and caves, which is the most typical feature of karst environments [98].
Karst aquifers constitute 12% of the global land surface and store water on which approximately 20-25% of the world's population depends [23, 135]. In most Mediterranean countries carbonate rocks, mainly karstified, offer the main groundwater resources, which have been exploited for several millenniums, supplying today at least 25% of domestic water [10], not counting their crucial importance for touristic and economically sustainable development [49]. Mediterranean karst is often considered to be the common reference for scientific community, as here karst landforms display their main hydrological characteristics [10].
In coastal carbonatic aquifers, closely linked to the sea, conduit flow is dominant and two essential mechanisms are involved in the functioning of these systems, i.e., aquifer discharge through submarine springs and saline intrusion through conduits open to the sea [50]. In this regard, Mediterranean coastal and submarine springs, which represent at least 90% of the known submarine karst springs in the world, are undoubtedly the most important features of Mediterranean karst [10]. Many large towns along the shoreline were built up in the vicinity of large karstic springs [49]. Moreover, many karst springs contribute to surface waters and play a major role in maintaining numerous aquatic ecosystems and wetlands [104].
From the standpoint of water quality, in principle, groundwater in carbonatic aquifers has high natural quality and does not require expensive treatment if not artificially polluted [121, 135]. However, carbon-

atic aquifers are highly vulnerable to contaminant spills and degradation, if wellhead protection areas are not properly detected [67]. So these aquifers are particularly sensitive to pollution from point and diffuse sources, which allow rapid infiltration and transport of dissolved substances along highly permeable conduits. Frequently, the ingress of pollutants is favoured by the absence of any clayey soil layer to act as a protective barrier against the passage of contaminants [135].

Due to the special intrinsic characteristics of karst (i.e., rapid hydraulic responses, focused channelized flow regimes, fast groundwater flow and transport in conduits), the effects of climate change on recharge, together with increasing pressures on karst groundwater quality, mean that the exploitation of aquifers in various Mediterranean regions has already caused deterioration of water resources. Despite increasing pressures on the quality and quantity of water sources, particularly in coastal areas, there is lack of coherent policy for proper monitoring of possible natural and human impacts [104]. This is particularly worrying in highly populated areas characterized by industrial and agricultural production activities, due to the release of chemical compounds entering soil and groundwater ecosystems from a variety of sources (e.g., industrial wastes, drain from urbanized areas and transportation systems, sewage treatment plants, irrigation) [25]. Groundwater and contaminants in carbonatic aquifers, indeed, can easily travel more than a kilometer per day, compared with velocities in porous media that are typically less than a meter per day. The rapid velocities and short travel times in carbonatic aquifers are a double-edged sword [67], so it is imperative in today's political environment to recognize that these characteristics and features render carbonatic aquifers particularly vulnerable to degradation resulting from different sources of pollution.

These problems are even more relevant than in non-karst regions, because the dissolution of carbonate matrix causes a limited surface runoff, so that surface water supplies in karst areas are generally more limited than in other natural settings. As a consequence, groundwater supplies are often the only available source of freshwater in karst areas for human health, food security and the economic sector [49, 50, 104].

Sustainable management of groundwater resources can be supported by numerical models to increase understanding of such complex natural systems and improve planning strategies [58]. However, the problem often lies in the questionable uncertainty and incompleteness of existing data, which limit the implementation of these models [81]. In fact, the difficulty in characterizing subsurface heterogeneity with commonly sparse datasets severely limits the accuracy and realism of groundwater flow and transport models [9]. Anyway, the main trends of the hydrological behaviour may be identified, as long as the underlying conceptual model is correct, and the parameters of the model have physically reasonable values. Once these conditions have been met, the numerical models may be used as decision-making tools in water-management planning [145].

Hydrogeological modelling is generally based on three steps: (1) the geometry of the different hydrostratigraphic units constituting the aquifer system under exam must be known, (2) their hydraulic parameters must be correctly estimated, and (3) these geometry and properties must be adequately implemented into a hydrogeological model [138]. Steps (1) and (2) allow setting a preliminary conceptual model and modelers are firstly faced with challenges related to them. In this regard, difficulties are mostly due to

scarcity of good-quality data, so making models predictive function often limited by a poor characterization of geometry and hydrogeological parameters at the model scale [44].
Correct estimation of aquifer parameters is important in modeling groundwater flow and transport and in groundwater resources assessment, protection and management. Aquifer transmissivity is commonly estimated from pumping tests, which are generally expensive and time intensive [143]. Consequently, the availability and spatial distribution of transmissivity measurements are often lacking. Thus, simple and inexpensive parameter estimation methods are desired. A common approach is to correlate transmissivity with the more easily measured specific capacity of a well, which is the ratio of pumping rate to drawdown in the well [13]. Furthermore, the combination of indirect geophysical data with direct hydrogeological parameters shows a high potential of providing hydraulic information at a scale and accuracy required for the developement of groundwater flow and transport models [18].
In addition, an effective management strategy requires a proper quantification of groundwater recharge to ensure the protection of groundwater resources from an unavoidable climate change impact and other stresses [94]. For this reason, many efforts are addressed to evaluate groundwater recharge, by identifying statistical correlations between groundwater levels and hydrometeorological parameters [42, 86], possibly using integrated approaches of remote sensing and GIS tools [96, 134].
The approach applied in this thesis to identify model parameter values consists in the solution of inverse problems by conditioning on observed, available data. A series of methods have been proposed over the decades to solve the inverse problem, ranging from trial-and-error calibration, based on subjective evaluations, to the complex automatic data assimilation algorithms [149]. A comprehensive list of different inverse methods proposed in groundwater hydrology and a full discussion of the problems related to their application can be found in several review papers (see, e.g., Carrera [26], Carrera et al. [27], de Marsily et al. [40], Giudici [62], McLaughlin et al. [92], Sun [124] and Yeh [144]). Practical examples and applications, instead, can be found, among the others, in Brouwer et al. [19], El-Rawy et al. [47], Foglia et al. [52, 53], Giudici et al. [61], Giudici & Vassena [63], Hendricks et al. [70], Parravicini et al. [99], Trykozko et al. [127, 128], Zijl et al. [150], Zimmerman et al. [151].
Several studies concerning numerical modelling of coastal carbonatic aquifers have been conducted on local and regional scales (see, e.g., Masciopinto [87] and Romanazzi et al. [107] and references therein). However, application of numerical models in carbonatic aquifers is problematic, because these natural systems are generally very heterogeneous, dominated by secondary (fracture) or tertiary (conduit) porosity and may exhibit hierarchical permeability structures or flow paths. Furthermore, they are likely to have a turbulent flow component, so that adapting traditional equations, which assume laminar flow in porous media, may be problematic. To solve these issues, the equivalent porous medium approach can be used, if the conduits are uniformly distributed and well interconnected [112]. It consists, indeed, in assuming that fractures and channels may be locally important, but their density is so great that the aquifer can be treated and modeled as an equivalent porous medium [142]. This approach performs better at large scales, where the effects of many conduits can be integrated [4]. As an example, it was adopted to the coastal, carbonatic aquifer of the Salento peninsula (Apulia region, southern Italy) to assess possible impacts of natural and anthropic stresses on quality and quantity of groundwater resources at the regional

scale [39, 65]. In those papers, a 2D sharp-interface approach based on the Ghyben-Herzberg approximation was used to detect the areas affected by the saltwater intrusion phenomenon, while a dispersive approach, based on the solution of the transport equation, was adopted in Romanazzi et al. [107] to the same region.

1.2 Specific objectives of the thesis

This thesis proposes a scientific contribution within the framework of modeling groundwater resources, in the view of assessing, understanding and managing in a sustainable way carbonatic coastal aquifers, crucial components of the Earth ecosystems. First and foremost, this research has a relevant technical significance, as it takes into account all the steps necessary for the developement of a hydrogeological model, with particular focus on the evaluation of model parameters in the challenging context of inverse modeling. Not less important is the practical application of mathematical tools to a real case study, which is a complex, highly non-linear natural system and subject to several criticalities, affecting either the environmental and socio-economic frameworks.
Such case study is represented by the Taranto area (Apulia region, southern Italy), whose groundwater resources, which satisfy mainly industrial and agriculture water demands, are stored in a multi-layered aquifer system. The latter is very complex from the hydrostratigraphic point of view, due to the presence of a deep, karst aquifer (which is the most exploited for human activities) and of overlying hydrostratigraphic units, whose sediments are poorly permeable or host shallow aquifers of local relevance. Previous studies [39, 65] highlighted that the Taranto area is particularly sensitive to the phenomenon of seawater intrusion, mostly due to extensive withdrawals for industrial and agricultural activities.
Another cause of concern is the relationship between the aquifer system and the highly vulnerable Mar Piccolo and Mar Grande surface seawater bodies, which host the most important freshwater springs in Apulia region [37] and whose quality is endangered by civil and military wastes (see, among the others, Alabiso et al. [1], Buccolieri et al. [21], Cardellicchio et al. [24], Storelli & Marcotrigiano [122]). The Mar Piccolo basin, in particular, plays a fundamental role also for fishing and mussel cultivation activities, so its quality, and therefore the wellness of the whole ecosystem, is a matter of major concern and depends not only on direct pollution, due to industrial, commercial and shipping activities practiced in the surroundings, but also on groundwater quality, through the submarine springs. Based on what said above and due to the socio-economic implications related to these issues, it is necessary to develop numerical tools to analyze the groundwater balance and the evolution of the aquifer system in response to anthropic activities and to climate changes, in view of supporting sustainable water management plans. The research activities presented in this thesis contributed to a wide multidisciplinary framework, represented by the National Flagship Project *RITMARE* (*la Ricerca Italiana per il MARE - the Italian Research for the Sea*). The main objective of this scientific initiative is combining maritime and marine policies, namely promoting a deep knowledge of marine environmental resources in order to develop innovation and technologies for their exploitation, while ensuring respect for their equilibrium. This major scope can be explicited as follows:

- supporting policies for marine protection;
- allowing a sustainable exploitation of marine resources;
- defining a strategy to prevent and mitigate natural and anthropic impacts.

Within the RITMARE project, a sub-program is specifically dedicated to coastal areas. Among coastal case studies, the gulf of Taranto (southern Italy) is taken into account within this sub-program.
In this thesis, a numerical model of groundwater variable-density flow for the coastal area of the Taranto gulf is presented. The specific aims are the following: to characterize groundwater resources of this area from a quantitative and qualitative point of view, to evaluate both the stresses due to climate changes and human water-demanding activities, and the occurrence of the saltwater intrusion phenomenon. In doing so, considering the carbonatic aquifer system at the scale of the whole catchment, the equivalent porous medium approach was used.
As a first step, a preliminary conceptual model was necessary in order to describe the groundwater dynamics from a qualitative point of view: this required to map the geometry of the aquifer system itself through the reconstruction of the hydrostratigraphic setup of the subsurface and to identify all kinds of source terms (recharge or withdrawals) influencing the groundwater balance.
Then a first groundwater flow model was developed by adapting the YAGMod code [28], developed at the Department of Earth Sciences of the Università degli Studi di Milano, to simulate the dynamics of a coastal aquifer under pseudo-stationary conditions, i.e., representing the average behavior during a hydrological year. This required the solution of an inverse problem to estimate the optimal hydraulic conductivity field of the deep, karst aquifer, and then the solution of the forward problem to simulate the hydraulic head and the water fluxes in the aquifer system under exam. In particular, to detect which zones are threatened by the saltwater intrusion phenomenon, the YAGMod code was modified and adapted to simulate a sharp interface between freshwater and saltwater through the Ghyben-Herzberg approximation.
The application of the YAGMod code allowed to identify also the main criticalities in terms of numerical model setup and conceptual model definition. These difficulties were analyzed and overcome through a further calibration of the groundwater flow model, in order to obtain the "optimal" values of the most uncertain model parameters. In this regard, a new methodology, so far unexplored, was proposed and applied: it consists in merging direct and indirect approaches for inverse modeling, and represents the major contribution and the most challenging aspect of this research activity from the scientific point of view. For this purpose, UCODE was used for indirect calibration and coupled with MODFLOW for the simulation process.
The occurrence of the saltwater intrusion phenomenon was then assessed through a dispersive solute-transport approach by using the SEAWAT code. A sensitivity analysis was further carried on to show the effects of climate changes and human overexploitation on the saline plume spreading inland.
It is worth stressing that this research involved remarkable computational and interpretation efforts, due to the use of different numerical tools to analyze a complex and highly non-linear aquifer system, for which scarce and often uncertain data are available.

1.3 Structure of the thesis

The first part of the thesis is devoted to the description of materials and methods used to face the described problems.
In particular, chapter 2 introduces some basic definitions of the systems and processes taken into account. Specifically, the focus is on porous media, because the equivalent porous medium approach is adopted. Groundwater flow in porous media is described introducing the continuum approach to define physical quantities, so that processes can be described at the mesoscopic scale. Darcy's law (which can be formulated also for fractured media) and the mass conservation law are then derived and Hubbert's potential is defined, to introduce balance equation for a fluid moving into a compressible and saturated porous medium, which can be adapted also for fractured materials. Such equation is then formulated in terms of variable-density flow, to be coupled with the equation for solute transport. For the aims of this thesis, transport is simulated to assess the occurrence of the saltwater intrusion phenomenon, which can be treated also with a sharp-interface approach, as that adopted in the Ghyben-Herzberg approximation.
The purpose of chapter 3 is to introduce some basic concepts related to the numerical modeling of groundwater flow. The discretization of the analyzed physical system and the balance equations are treated with a finite-difference formulation, which leads to simulate the processes by numerically solving a system of possibly non-linear equations. Properly setting a numerical model further requires introducing boundary conditions and initial conditions for the simulated processes. Moreover, calibration is an important step for evaluation of the most uncertain model parameters and, for this scope, the solution of direct and indirect inverse problems is proposed here. Finally, as numerical models are often used as decision support tools, evaluating their predictive utility through sensitivity analysis is of paramount importance.
Chapter 4 is devoted to describe the structure and mathematical statements of the codes used in this thesis: specifically, details about the discretization and solution of the equations adopted, as well as the definition of the model setup are provided. The codes described here are MODFLOW [69] for the simulation of groundwater flow, YAGMod [28] for model calibration with a direct method of inversion and for the simulation of groundwater flow adopting a sharp-interface approach to assess saltwater intrusion, SEAWAT [66] for the simulation of the saltwater intrusion phenomenon adopting a diffusive approach, and UCODE [72, 100] for sensitivity analysis and calibration through an indirect, inverse approach.
The second part of this thesis is devoted to the application of these methods to the Ionian aquifer system of the Taranto area.
In chapter 5 the study area is described from the geographical, environmental, geomorphological and hydrogeological points of view.
Chapter 6 aims at identifying the hydrostratigraphic setup and the conceptual model of the aquifer system under exam, so getting a qualitative overview of the groundwater flow dynamics and aquifer balance and setting the stage for a further numerical characterization.
In chapter 7, application of the numerical codes YAGMod, MODFLOW and UCODE is described. In particular, YAGMod allowed the solution of an inverse problem with a direct approach through the CMM

for the estimation of the hydraulic conductivity of the carbonatic, deep aquifer. The forward problem was then solved and allowed to simulate the groundwater flow dynamics, detecting also the areas affected by the saltwater intrusion, as the original code was modified to embed the Ghyben-Herzberg approximation in the estimation of the saturated thickness of the aquifer. Major challenges in the conceptual model and model setup identified with YAGMod were then taken into account in the following calibration phase. In this regard, a merging of direct and indirect inverse approaches is proposed, through the use of the CMM and UCODE, and using MODFLOW to perform simulations so comparing results of calibrated models with the available observations. The last paragraph of chapter 7 is devoted to assess the occurrence of the saltwater intrusion phenomenon by adopting a dispersive solute-transport approach with the SEAWAT code. Two scenarios are further presented to show the effects of climate changes and human impacts on the only freshwater resource of the study area.

Part I

Materials and methods

Chapter 2

Groundwater balance in coastal aquifers: continuum approach and basic equations

This chapter deals with the basic laws governing the motion of groundwater in aquifers, and with the aquifer properties appearing in these laws. In deriving these basic laws, the continuum approach is employed and all variables and parameters have their average meaning in the porous medium as a continuum.

2.1 Basic definitions

In this section some fundamental definitions are introduced, in order to clarify the basic terminology used in what follows, and the most important physical processes involved [14].

The flow of a fluid through a rock occurs across a complex network of interconnected pores or fractures which develop within a solid and heterogeneous matrix. Unconsolidated sediments (e.g., sands, gravels, clays) and sedimentary rocks (e.g., sandstones) are *porous media*, namely heterogeneous multiphase materials where at least one phase is a persistent, possibly deformable, solid phase. The latter is called the *solid matrix*, while the spaces among solid grains of the rock are called the *pores* and the *pore space* (or *void space*) is the portion of rock formation which is not occupied by solid matter. On the other hand, compact rocks (sedimentary, metamorphic or igneous) allow fluids' flow through *fractures*, whose origin is due to the alteration of the rock caused by atmospheric agents or tectonic movements.
Within this framework, an *aquifer* is a geologic formation or a group of formations made of several rock layers or unconsolidated sediments which contain water and permit significant amounts of fluid to move through pores or fractures under ordinary field conditions, namely under the effect of a hydraulic gradient lower than 10^{-2}.
On the other hand, an *aquiclude* is a formation which may contain water (sometimes in appreciable quantities), but is not able to transmit significant amounts of fluid under ordinary field conditions (e.g., a clay layer). For all practical purposes, an aquiclude is considered as an impervious formation.

An *aquitard* is a semipervious geologic formation: it transmits water at a very low rate compared to the aquifer, but, over a large (horizontal) area, it may permit the motion of large amounts of fluid between adjacent aquifers which it separates from each other.
Finally, an *aquifuge* is an impervious formation which neither contains nor transmits water.
Subsurface formations containing water may be divided vertically into two main zones, according to the relative proportion of the pore space which is filled with water. In the *saturated zone* all pores are completely filled with water, while in the overlying *unsaturated zone* (or *vadose zone*) the pores contain both gases (mainly air and water vapour) and water. The saturated zone is bounded from above by the *water table* (or *phreatic surface*), namely an imaginary surface at all points of which the pressure is atmospheric. The water table can be revealed by the level at which water stands in a well just penetrating the aquifer.
Aquifers may be classified as confined or unconfined, depending upon the presence or absence of overlying, poorly permeable, confining layers (aquitards, aquicludes or aquifuges).

- A *confined aquifer* (or *pressure aquifer*) is bounded both from above and from below by impervious formations. In a well just penetrating such an aquifer, the water level will rise above the base of the upper confining formation; it may reach the ground surface (*artesian aquifer*) and in that case water will flow freely from a well tapping the aquifer, without pumping (artesian or flowing well).

- A *phreatic aquifer* (or *unconfined aquifer*) is bounded from below by impervious formations, while the water table represents its upper boundary. A phreatic aquifer is directly recharged from the ground surface, unless impervious layers exist between the water table and the ground surface itself. Actually, above the water table, there is a capillary fringe, namely a zone within which a gradual decrease in moisture content with height occurs: just above the phreatic surface, the pores are completely saturated, while moving higher the water seeps up by capillary action to fill only the smallest connected pores. However, the capillary fringe is often neglected in groundwater studies.

An aquifer system consisting of a phreatic aquifer overlying one or more confined aquifers, separated from each other by aquitards, is defined as a *multilayered aquifer*.
For the aims of this thesis, it is worth defining also a *coastal aquifer* like an aquifer partially or fully bounded by the sea. The most important properties of coastal aquifers are listed below [131]:

- high permeability of rocks outcropping at the surface and/or at the sea bottom, along the coastline;

- hydraulic head values decrease from inland towards the coast, where they vanish as the water table reaches the height of the mean sea level;

- groundwater outflow locally occurs through submarine or subaerial springs, which can drain brackish water;

- owing to the presence of seawater into the aquifer formation under the sea bottom, a zone of contact is formed between the low-density freshwater flowing towards the sea and the high-density, underlying, seawater.

2.2 Groundwater flow in porous media

The continuum approach to groundwater flow is firstly described for porous media [64], whereas the flow processes occurring in fractured media will be discussed in subsections 2.2.5 and 2.3.3.

2.2.1 Continuum approach to porous media

When dealing with flow into a porous aquifer, through a complex network of interconnected pores, we overlook the microscopic flow patterns inside individual pores and consider an average flow which takes place in the porous medium as a whole. This requires introducing some average physical quantities characterizing the medium itself. By doing so, a *continuum approach* is employed. Instead, applying the concept of continuum at the microscopic (pores-grains) scale (relevant horizontal and vertical resolution lengths ranging between 10^{-5} m and 10^{-3} m) would require a detailed knowledge of the pores boundaries and their properties at the same scale.
Based on what said, a porous medium can be described as a continuum medium at the *mesoscopic scale* (relevant horizontal and vertical resolution lengths ranging between 10^{-3} m and 10^{-1} m) [64]. In passing from the microscopic level (at which we consider what happens at each point inside each pore) to the mesoscopic one (at which only averaged phenomena are considered), the definition of a *Representative Elementary Volume* (*REV*) is necessary. This means that, at the mesoscopic scale, the considered physical quantities are defined as averages, over a REV of the porous medium, of the physical quantities themselves defined at the microscopic scale. Therefore, the physical quantities so defined (e.g., porosity, volumetric water content, solute concentration,etc.) are representative of the physics of the porous medium, in the sense that they take into account the presence of several phases (gas, liquid, and solid phases) and possibly several constituents within each phase.
The definition of the REV and the physical quantities at the mesoscopic scale is introduced in the next subsection by using the definition of porosity as an example [14, 59].

2.2.2 Definition of REV

Let P be a point inside a porous medium and consider an elementary spherical volume centered at P and whose radius is r. Then, its volume is given by $V(r;P) = \frac{4}{3}\pi r^3$. Let $V_P(r;P)$ be the volume of the pores contained in this elementary volume. The following ratio,

$$n(r;P) = \frac{V_P(r;P)}{V(r;P)} \tag{2.1}$$

represents the fraction of the elementary volume taken by pores: $0 \leq n(r;P) \leq 1$.
If P lies into a pore, then $n(r;P) = 1$ if r is lower then the pore dimension. Instead, if P is within a grain, then $n(r;P) = 0$ for r lower than the grain dimension. As r increases, the elementary volume could include several pores and grains. So, even for a small variation of r, the value of $n(r;P)$ oscillates quite strongly at first and then it might show a smooth trend, due to large scale heterogeneity, because the elementary volume would include sediments corresponding to different lithological properties.

If $n(r;P)$ is almost constant for a given interval $r_{min} \leq r \leq r_{max}$, then the elementary volume, whose radius is within this interval, is assumed to be "representative" of the characteristics of the porous medium. Therefore, a REV is defined as an elementary volume, whose radius falls within the interval (r_{min}, r_{max}). Since the value of $n(r;P)$ does not depend on the radius (r) of the REV, it is the same for every REV and describes a property of the porous medium at the mesoscopic scale, the total porosity, which is thus indicated with $n(P)$.

In this way, a conceptual procedure to define a *continuous* and *punctual* property of the porous medium has been introduced. This property is continuous because it is defined everywhere in the porous medium, and, therefore, on a continuous domain. It is punctual because it is representative of the porous medium at the mesoscopic scale and varies from point to point, namely inside a volume which is so small as to be considered as infinitesimal, but so big as to include a great number of pores and grains [64].

The above procedure can be generalised in a straightforward way and the concept of a *Representative Elementary Area* (*REA*) can be introduced for some quantities (e.g., the filtration flux through the porous medium or specific discharge or Darcy's velocity).

The definition of the REV allows associating physical quantities to each point. These quantities are significant at the mesoscopic scale and can be formalized mathematically as functions of space and time [59]. This permits applying all the tools of mathematical analysis, like derivatives and integrals.

The concepts of REV and REA are useful, for instance, to define the link between the fluid average velocity at the mesoscopic scale, $\mathbf{v}^{(m)}$, defined as a volume-averaged quantity, and Darcy's velocity $\mathbf{q}$, namely the flow rate per unit surface across the porous medium, defined as a surface-averaged quantity. Such link is usually expressed through the drainage or effective porosity, n, by the following expression:

$$\mathbf{q} = n\mathbf{v}^{(m)}. \tag{2.2}$$

2.2.3 Darcy's law in saturated porous media

Darcy's law is a basic physical or phenomenological law describing the hydraulic flux in porous or fractured media.

To formulate Darcy's law, it is worth describing an experiment, whose schematic representation is shown in Figure 2.1.

The system shown in Figure 2.1 consists of a shaped tube filled with a sandy sample in the middle, at the horizontal side. Let L and A be the length and the base area of the sample cylinder, respectively. The left reservoir is constantly fed with water, but an opening allows the water level (h_1) staying constant during the experiment, both if filtration through the sample occurs or not. The right reservoir, instead, has multiple openings and the water flowing out of one of them is collected in a graduated vessel, which allows measuring the filtration flow rate Q (namely the volume of fluid which circulates through the sample per unit time, $[L^3T^{-1}]$). Let h_2 be the water level in the right reservoir, according to which opening is active. The water levels h_1 and h_2 are measured with respect to some arbitrary (horizontal) datum level. Under these conditions, the hydraulic load at the left side of the sample is constant, while that at its right side can vary.

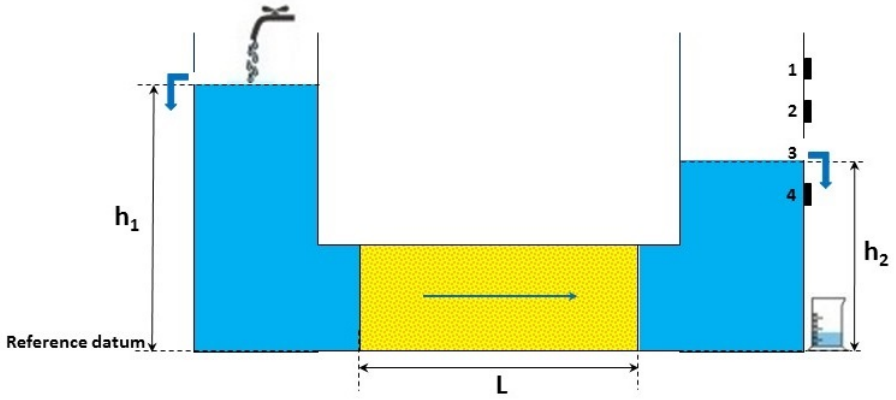


Figure 2.1: Representation of the equipment used for Darcy's experiment.

At first, suppose that outlet 1 is open, so that $h_1 = h_2$ and no difference of hydraulic load is detected between the extreme sides of the sample. In this case, no filtration of fluid occurs through it and $Q = 0$. If outlet 2 is open, a difference in the hydraulic load between the extreme sides of the sample is generated ($\Delta h = h_1 - h_2$) and a filtration process is activated. If so, the water flowing out from outlet 2 can be collected in the graduated vessel and Q can be measured.
If outlets 3 and 4 are open, the difference of hydraulic load (Δh) increases accordingly and so is for the filtration flow rate (Q). Experimental data show that Q linearly varies when Δh varies:

$$Q \propto \Delta h.$$

On the contrary, by repeating the experiment with a constant Δh but with different values for the length of the sample cylinder (L) one can observe that Q and L are inversely proportional:

$$Q \propto \frac{1}{L}.$$

So we can write:

$$Q \propto \frac{\Delta h}{L} = i,$$

where i is called the *hydraulic gradient*, it is dimensionless and it is the ratio between the difference of hydraulic load between two points and their distance. In physical terms, i defines the driving force which allows fluid movement inside the aquifer, against the resistence of the terrain.
Furthermore, if the base area of the sample cylinder (A) varies, while Δh stays constant, one can get a linear relation between Q and A:

$$Q \propto A.$$

Definitively, we can write:

$$Q = -KAi, \tag{2.3}$$

where the proportionality coefficient, K, is called the *hydraulic conductivity* $[LT^{-1}]$, and represents the ability of a porous medium to be crossed by water. K depends upon the properties of both the medium and the fluid and can be expressed as:

$$K = \kappa \frac{\rho g}{\mu}, \tag{2.4}$$

where κ is the *intrinsic permeability* of the porous medium $[L^2]$, and depends upon the pores dimension, while ρ and μ are the fluid density $[ML^{-3}]$ and dynamic viscosity $[ML^{-1}T^{-1})]$, respectively.
The negative sign into equation 2.3 takes into account that the hydraulic flow occurs from the left reservoir (where the water level is h_1) towards the right one (where the water level is $h_2 < h_1$).
The formulation 2.3 for Darcy's law refers to the experiment described above, with the same geometry shown in Figure 2.1. In order to describe the hydraulic flow which occurs in the subsurface, equation 2.3 can be generalized by using the specific discharge (or Darcy's velocity), $q = Q/A$, namely the flow rate per unit surface across the porous medium. By doing so, equation 2.3 can be written as follows:

$$q = -Ki,$$

which, by using vectors, becomes:

$$\mathbf{q} = -K\nabla h, \tag{2.5}$$

where ∇h is the vector formulation of the hydraulic gradient (i). Equation 2.5 is the differential *Darcy's law* for saturated porous media.
The flow rate across the porous medium is usually so small that, if the two reservoirs are large enough, it can be assumed that water in the reservoirs is static. Therefore, the pressure at the top of a reservoir is nothing but p_a, the atmospheric pressure, whereas at any height (z) above a given reference horizontal plane it is given by the law of hydrostatic pressure: $p = p_a + \rho g(h - z)$, where ρ is the water density, g is the gravity acceleration and h is the height of the water level in the reservoir with respect to the reference level. Notice that, if the water is considered as incompressible, than ρ does not vary with pressure. Consequently, h can be expressed as:

$$h = z + \frac{p - p_a}{\rho g}, \tag{2.6}$$

which is called the *water head* or *hydraulic head*. Notice that the hydraulic head is constant within each of the two reservoirs of Darcy's experiment (see Figure 2.1, where $h = h_1$ into the left reservoir and $h = h_2$ into the right one) and the water flow through the rock sample is driven by the difference $\Delta h = h_1 - h_2$ between the extreme sides of the sample itself. Moreover, it is possible to measure the water level into an aquifer by measuring the height (with respect to a reference horizontal plane, usually the mean sea level) of the water level into a borehole. In fact, if a borehole is drilled into the ground, until a depth where the rocks are saturated with water, and if the borehole is left open at the bottom, the water will rise through the borehole up to the height h, where the pressure is equal to p_a.
Darcy's law 2.5 is not universally valid for all conditions of fluid flow in porous media [73].

First of all, it has long been recognized that the linearity of the flux versus hydraulic gradient relatonship fails at high flow velocities, where inertial forces are no longer negligible compared to viscous forces. Darcy's law applies only as long as flow is laminar (i.e., non turbulent movement of adjacent layers of the fluid relative to one other occurs) and where soil-water interction does not result in a variation of permeability with a change of gradient. Laminar flow prevails in silts and finer materials for most commonly occurring hydraulic gradients found in nature. In coarse sands and gravels, however, hydraulic gradients much in excess of unity may result in non-laminar flow conditions, so in such cases Darcy's law may not be applicable. Non-laminar flow can also occur in macropores, such as wide cracks or wormholes.
The quantitative criterion for onset of turbulent flow is given by the *Reynolds number* (R_e), which is a dimensionless number expressing the ratio of inertial to viscous forces acting on the fluid:

$$R_e = dq\rho/\mu,$$

where d is the effective pore diameter, q is the mean flow velocity, ρ is the fluid density and μ its dynamic viscosity. For porous media, it is safe to assume that flux remains linear with hydraulic gradient only as long as R_e is smaller than unity.
As flow velocity increases, especially in systems with large pores, the occurrence of turbulent eddies results in "waste" of effective energy, namely some energy is dissipated by the internal turbolence of the fluid, so that the hydraulic potential gradient becomes less effective in inducing flow.
Deviations from Darcy's law may also occur at the opposite end of the flow-velocity range, namely at low gradients and in narrow pores: in clayey soils, for example, small hydraulic gradients would not cause any flow, or flow rates that are less than proportional to the gradient. A possible reason for this anomaly is that the water in proximity to the particle surfaces and subject to their adsorptive force fields may be more rigid than ordinary bulk water.
Still, another important limitation of Darcy's law relates to the scale of the flow system. The law may fail at the microscopic scale, where flow processes take place across short distances comparable to the sizes of individual pores. On the other extreme, it may also fail at the macroscopic scale, where the soil medium is very heterogeneous and may include domains of both laminar and turbulent flows in several directions and tortuosities. So, it is tipically relevant in the volume range from a few cubic millimeters to a few cubic meters.
In conclusion, Darcy's law 2.5 is applicable mainly to relatively homogeneous systems of intermediate scale and pore sizes [73].

2.2.4 Hubbert's potential

Since water flows through the soil very slowly, its kinetic energy is generally considered to be negligible. On the other hand, the potential energy is of primary importance in determining the state and movement of water in soils. Into the soil, indeed, the moving force causing flow is related to differences in potential energy per unit weight of water between two points and acts from high to low potential energy ("high" and "low" are referred to a reference level) [73].

If ϕ denotes the potential energy per unit weight, groundwater flow is caused by the negative potential gradient ($-\nabla\phi$), namely the variation of ϕ with distance. Similarly, according to Darcy's law 2.5, groundwater flow is caused by the negative hydraulic gradient ($-\nabla h$), namely the variation of water level (h) over distance. By comparison, we want to check that h expresses a potential energy per unit weight of water.

Let m be the mass of water inside the volume (V) of a porous medium. In order to get its potential energy, we should consider the mechanic energy which allows the water to move inside the medium. The meaningful quantity to be considered is not the energy itself, but the variation of energy with respect to a reference state, which will be defined below while considering different contributions to the total energy. For the sake of simplicity, we will consider a one-dimensional motion.

The first contribution is related to the kinetic energy, given by $\frac{mv^2}{2}$, where v is the absolute value of the average velocity of the water inside the medium. At the reference state, the water is considered at rest, so that $v_{ref} = 0$.

Moreover, the potential energy due to the gravity is expressed by mgz, where g is the gravity acceleration and z is the height of the water mass with respect to the reference level, for example the mean sea level.

In order to express the third contribution to the energy, we should refer to the thermodynamic state of the water, which can be described by the Gibbs free energy, if we assume that the processes occurring in the underground are isothermal. Variations of Gibbs free energy are related to variations of pressure (p), which allow getting variations in hydraulic head and so, according to Darcy's law, flux through the medium. Gibbs free energy is expressed by:

$$G = U - TS + pV,$$

where U is the internal energy of the thermodynamic system, while T and S are its temperature and entropy, respectively. Since we are interested in variations of Gibbs free energy, we can express dG by using the first principle of thermodynamics in the form $dU = TdS - pdV$:

$$dG = dU - (SdT + TdS) + (Vdp + pdV) =$$

$$= TdS - pdV - SdT - TdS + Vdp + pdV = Vdp - SdT.$$

Furthermore, assuming isothermal conditions in the subsurface (constant T, so that $dT = 0$), the formulation $dG = Vdp$ is particularly convenient. At the reference state, pressure and temperature are assumed to be the atmospheric pressure (p_a) and the (constant) temperature of the water in the subsurface, respectively.

The last contribution we will consider is related to eventual variations in the chemical composition of the water solution. These variations can be expressed as $\sum_k \varphi_k dn_k$, where the sum is extended to all k components of the solution, φ_k represents chemical potentials and n_k is the mole number for the k component. In this case, the reference state is described by pure water, without any component dissolved into the solution.

Now we can define Hubbert's potential (ϕ^*) by adding all these contributions and (eventually) other

contributions here neglected, per unit weight:

$$\phi^* = \frac{v^2}{2g} + z + \frac{G}{mg} + \frac{1}{mg}\sum_k \varphi_k n_k + ...$$

By expressing the variation of Hubbert's potential, we get:

$$d\phi^* = \frac{1}{2g}dv^2 + dz + \frac{V}{mg}dp + \frac{1}{mg}\sum_k \varphi_k dn_k + ... \tag{2.7}$$

If we assume that the average velocity in a porous medium is negligible and so is for the chemical potentials, the variation of Hubbert's potential can be approximated as $d\phi^* = dz + \frac{dp}{\rho g}$, where $\rho = m/V$ is the water density and can vary with pressure. By considering the reference state (pure water, at rest, at the mean sea level, at atmospheric pressure and at the same temperature of the underground), the final equation for *Hubbert's potential* is:

$$\phi^* = z + \int_{p_a}^{p} \frac{dp}{\rho(p)g},$$

which is the generalization of the hydraulic head concept (see equation 2.6) and represents the energy necessary to bring a mass of water from the reference state to that characterizing the underground.

2.2.5 Groundwater flow and Darcy's law in fractured media

For this subsection refer to Domenico & Schwartz [46].

Darcy's law 2.5 can be applied also when water flow occurs in a fractured medium. In principle, this case would require the solution of the flow equations in each fracture, but if the fracture density (namely the number of fractures per unit rock volume) is high and if they have small openings, then the continuum approach, the concept of REV and Darcy's law 2.5 can still be applied. In this way, it is assumed that the fractured medium is *hydraulically equivalent* to a porous one. However, in this case the orientation of the fractures could yield anisotropy, so that K and κ are second-order tensors.
For fractured media with flat and parallel fractures, it is possible to define hydraulic conductivity and intrinsic permeability thanks to Snow's laws, as follows:

$$K = \frac{\rho g N b^3}{12\mu}, \qquad \kappa = \frac{N b^3}{12},$$

where b is the opening of fractures and N is the number of fractures per unit length perpendicularly to the fractures plane.
Moreover, it can be demonstrated by experiments that, once fixed the hydraulic gradient ($i = \nabla h$), the flow through a fracture is proportional to the cube of its opening (b). In fact, for a laminar flux through flat and parallel plates, the volumetric flow rate is expressed by the following cubic law:

$$Q = \frac{\rho g b^2}{12\mu}(bw)i,$$

where w is the fractures width perpendicularly to the flow direction, while i is the hydraulic gradient into the flow direction. The previous equation is nothing but Darcy's law (see equation 2.5), where A is replaced by bw and K is given by:

$$K = \frac{\rho g b^2}{12\mu}.$$

By considering the roughness of the fracture, it can be obtained the following expression:

$$K = \frac{\rho g b^2}{12\mu[1 + CR_f^m]},$$

where C is a constant higher than one, R_f describes the roughness and $m > 1$.
As an example, in Table 2.1 K values for some porous and fractured materials are listed:

Lithology	K (m/s)
Gravel [14]	$10^{-4} - 10^{-1}$
Sand [14]	$10^{-6} - 10^{-3}$
Silt [14]	$10^{-8} - 10^{-5}$
Clay [14]	$10^{-11} - 10^{-7}$
Fractured basalt [46]	$4 \cdot 10^{-7} - 2 \cdot 10^{-2}$
Fractured granite [46]	$5 \cdot 10^{-8} - 3 \cdot 10^{-4}$
Limestone [46]	$10^{-9} - 6 \cdot 10^{-6}$
Karst limestone [46]	$10^{-6} - 2 \cdot 10^{-2}$

Table 2.1: Hydraulic conductivity values for some porous and fractured materials.

2.3 Mass conservation laws

2.3.1 Mass conservation equation for a porous medium

In order to get the governing equation for a fluid moving into a porous medium, the mass conservation law is necessary, in addition to Darcy's law 2.5.
For a mathematical formulation of the mass conservation principle, consider a saturated porous medium and let n be its total porosity. Let Ω be a fixed, arbitrary subset of the fluid domain.
The fluid mass inside Ω at a certain time instant (t) can be expressed as:

$$M(t) = \int_\Omega n\rho d^3\mathbf{x}. \tag{2.8}$$

Let $\Delta M = M(t + \Delta t) - M(t)$ be the mass variation inside Ω between the time instants t and $(t + \Delta t)$. By taking into account the mass conservation principle, if there are no source terms inside Ω, the mass

variation (ΔM) must be produced by a net mass flux through the boundary surface ($\partial\Omega$). This statement can be expressed in mathematical terms as follows:

$$M(t+\Delta t) - M(t) = \int_t^{t+\Delta t} \int_{\partial\Omega} n\rho(\mathbf{x}, t')\mathbf{v}(\mathbf{x}, t') \cdot \mathbf{n}^{in}(\mathbf{x}) d^2\mathbf{x} dt',$$

where $\mathbf{n}^{in}$ is the unit vector perpendicular to $\partial\Omega$ and oriented inside Ω and $\mathbf{v}$ is the average velocity of the fluid inside Ω.
By dividing both sides by Δt and assuming that Δt is infinitely small, the previuos equation can be written as:

$$\frac{dM}{dt} = \int_{\partial\Omega} \rho\mathbf{q} \cdot \mathbf{n}^{in}(\mathbf{x}) d^2\mathbf{x},$$

where we used the relationship between $\mathbf{v}$ and the fluid velocity inside the pores ($\mathbf{q}$, see equation 2.2). By using equation 2.8, we get:

$$\frac{d}{dt}\int_\Omega n\rho d^3\mathbf{x} = \int_{\partial\Omega} \rho\mathbf{q} \cdot \mathbf{n}^{in}(\mathbf{x}) d^2\mathbf{x}. \tag{2.9}$$

Since Ω is fixed and arbitrary, the left-hand side can be written as:

$$\frac{d}{dt}\int_\Omega n\rho d^3\mathbf{x} = \int_\Omega \frac{\partial(n\rho)}{\partial t} d^3\mathbf{x},$$

while the term at the right-hand side can be transformed by using the divergence theorem:

$$\int_{\partial\Omega} \rho\mathbf{q} \cdot \mathbf{n}^{in}(\mathbf{x}) d^2\mathbf{x} = -\int_\Omega \mathrm{div}(\rho\mathbf{q}) d^3\mathbf{x}.$$

With these modifications, equation 2.9 becomes:

$$\int_\Omega \frac{\partial(n\rho)}{\partial t} d^3\mathbf{x} = -\int_\Omega \mathrm{div}(\rho\mathbf{q}) d^3\mathbf{x}.$$

If Ω is arbitrary, we get:

$$\frac{\partial(n\rho)}{\partial t} = -\mathrm{div}(\rho\mathbf{q}), \tag{2.10}$$

which is the *mass conservation law* for a fluid inside a porous medium with an average velocity $\mathbf{v}$ and a total porosity n.

2.3.2 Balance equation for a fluid moving into a compressible and saturated porous medium

Consider a compressible and saturated porous medium. The balance equation for a fluid moving throughout the pores network expresses in mathematical terms the mass conservation principle.
In order to derive this continuity equation, consider equation 2.10. For the sake of simplicity, we will

suppose that the fluid density (ρ) is constant over time (e.g., the fluid is incompressible) and doesn't vary sharply over space, so that equation 2.10 becomes:

$$\frac{\partial n}{\partial t} = -\mathrm{div}\mathbf{q}.$$

Furthermore, while considering groundwater flow in porous aquifers, we should take into account eventual variations of fluid mass related to source terms. For this reason, we should introduce an additional term (f) to the right-hand side of the previous equation:

$$\frac{\partial n}{\partial t} = -\mathrm{div}\mathbf{q} + f.$$

Moreover, by using Darcy's law 2.5, we get:

$$\frac{\partial n}{\partial t} = \mathrm{div}(K\nabla h) + f.$$

We can express the left-hand side of this latter equation by introducing the *specific storage coefficient* (S_0), which is the amount of water that can be injected, per unit volume of the aquifer, per unit change in head:

$$S_0 = \frac{\partial n}{\partial t}\left(\frac{\partial h}{\partial t}\right)^{-1}.$$

So, finally, we get:

$$S_0\frac{\partial h}{\partial t} = \mathrm{div}(K\nabla h) + f, \tag{2.11}$$

which is the *balance equation for a fluid moving into a compressible and saturated porous medium*, obtained by joining the mass conservation law 2.10 and Darcy's law 2.5.

2.3.3 Balance equation in fractured media

In analogy with the discussion in section 2.2.5 for Darcy's law, equation 2.11, which has been deduced for a porous medium, is often applied also to represent flow in fractured and karst media. In large-scale models, indeed, fractured aquifers are treated as equivalent heterogeneous porous media and the fracture zones, lying along preferential planes, are distinguished from the porous rock matrix by the assignment of long-range correlation structures associated with high hydraulic conductivity. Anyway, idealization in a continuum model may be possible only when there are sufficient interactions (or fluxes) between fluids in the low and high conductivity zones; that is, the continuum model assumption may be idealized as the number of fractures is very high. In this framework, stochastic methods are often used to represent preferential flow pathways in fractured aquifers [129].
Fractured systems are characterized by extreme variability in porosity and permeability, with most of the fluid being located in low-permeability, disjointed matrix blocks. In contrast, most of the fluid mobility is in a small volume of high-permeability, interconnected fractures [77]. In this framework, carbonatic and

fractured aquifers are sometimes viewed as double-continuum or double-porosity systems: the porous matrix of carbonate rock, indeed, is characterized by low permeability and a high storage capacity and is cut by conduits ranging in size from millimeters or less (for fractures) to tens of meters (for karst caves and conduits), characterized in turn by a high permeability and a low storage capacity [4]. Earlier double-porosity models rely on parallel-fracture assumptions, while other models represent blocky fractured systems (i.e., systems with suborthogonal fractures) by using a spherical idealization of matrix blocks [77].

When developing a double-porosity model, the two overlapping continua (fractures and matrix blocks) are allowed to coexist and interact with each other. In particular, the differences in pressure between the porous blocks and the fractures lead to flow of fluid from porous blocks to adjacent fractures [116]. Two approaches exist to treat the matrix-fracture flow transfer: in the first approach, this transfer is related to a shape factor, which depends on matrix block geometry; the other approach is to treat the flow transfer term explicitly through boundary conditions, assuming that porous matrix blocks act as sources (or sinks) that feed (or drain) the fractures [32].

The differential equations describing the confined flow within the fractures and matrix block can be expressed as:

$$S_{\mathfrak{f}}\frac{\partial h_{\mathfrak{f}}}{\partial t} + S_{\mathfrak{m}}\frac{\partial h_{\mathfrak{m}}}{\partial t} = div(K_{\mathfrak{f}}\nabla h_{\mathfrak{f}}) + f, \tag{2.12}$$

$$S_{\mathfrak{m}}\frac{\partial h_{\mathfrak{m}}}{\partial t} = \frac{K_{\mathfrak{m}}}{\mathcal{L}}(h_{\mathfrak{f}} - h_{\mathfrak{m}}),$$

where subscripts $\mathfrak{f}$ and $\mathfrak{m}$ refer to fractures and matrix block, respectively, $\mathcal{L}$ is the characteristic length and the meaning of the other symbols is obvious from subsection 2.3.2. Furthermore, hydraulic condictivities $K_{\mathfrak{f}}$ and $K_{\mathfrak{m}}$ are related through the above mentioned shape factor. In the following, we will generally assume an equilibrium at large scale between matrix and fracture network, so that $h_{\mathfrak{f}} = h_{\mathfrak{m}}$ and equation 2.12 reduces to equation 2.11.

2.4 Variable-density flow

The content of this section is based on Bear [14] and Tulipano & Sappa [131], unless otherwise specified.

A number of computer codes exist for the simulation of seawater intrusion in coastal aquifers. Seawater intrusion models may be divided into interface and dispersive solute-transport models. In *dispersive solute-transport models*, fluid density can vary continuously or from cell to cell in a model domain, while in *interface models*, freshwater and seawater are separated by a sharp interface.

The following subsections deal with mathematical statements about variable-density groudwater flow, as treated in both dispersive solute-transport (subsections 2.4.1 and 2.4.3) and interface (subsection 2.4.4) models. Such discussion can be important in many types of studies of coastal aquifers, such as contaminated site remediation, freshwater discharge into saltwater bodies, aquifer storage and recovery, besides saltwater intrusion itself [66].

2.4.1 Balance equation for variable-density groundwater flow

For the content of this subsection refer to Guo & Langevin [66].

In order to derive the balance equation for variable-density flow, we should take into account mass conservation law 2.10 and add at the right-hand side a mass-exchange term due to sources or sinks:

$$\frac{\partial(n\rho)}{\partial t} = -\mathrm{div}(\rho\mathbf{q}) + \rho_s q_s, \tag{2.13}$$

where ρ_s is the density of water entering from a source or leaving through a sink and q_s is the volumetric flow rate per unit volume of aquifer representing sources and sinks $[L^3T^{-1}]$.
The right-hand side of equation 2.13 is the sum of the net flux of mass through the faces of the REV ($\mathrm{div}(\rho\mathbf{q})$), plus the rate at which mass enters from sources or leaves through sinks located within the REV ($\rho_s q_s$). The left-hand side of equation 2.13 is the time rate of change in the mass stored within the REV and can be expanded as:

$$\frac{\partial(n\rho)}{\partial t} = \rho\frac{\partial n}{\partial t} + n\frac{\partial \rho}{\partial t}. \tag{2.14}$$

The changes of porosity considered here are restricted to those associated with the change of fluid pressure (p), therefore:

$$\frac{\partial n}{\partial t} = \frac{\partial n}{\partial p}\frac{\partial p}{\partial t}. \tag{2.15}$$

Moreover, under isothermal conditions, fluid density is a function of fluid pore pressure and solute concentration ($\mathcal{C}$), therefore:

$$\frac{\partial \rho}{\partial t} = \frac{\partial \rho}{\partial p}\frac{\partial p}{\partial t} + \frac{\partial \rho}{\partial \mathcal{C}}\frac{\partial \mathcal{C}}{\partial t}. \tag{2.16}$$

Substituting equations 2.15 and 2.16 into equation 2.14 gives:

$$\frac{\partial(n\rho)}{\partial t} = \rho\frac{\partial n}{\partial p}\frac{\partial p}{\partial t} + n\frac{\partial \rho}{\partial p}\frac{\partial p}{\partial t} + n\frac{\partial \rho}{\partial \mathcal{C}}\frac{\partial \mathcal{C}}{\partial t}, \tag{2.17}$$

where the first two terms at the right-hand side represent the rate of fluid mass accumulation due to groundwater storage effects (i.e., due to the compressibility of the bulk porous material and fluid), while the third term represents the rate of fluid mass accumulation due to the change of solute concentration.
We can express the compressibility of a bulk porous material $[M^{-1}LT^2]$ as:

$$\xi = \frac{1}{1-n}\frac{\partial n}{\partial p} \tag{2.18}$$

and the coefficient of fluid compressibility $[M^{-1}LT^2]$ as:

$$\zeta = \frac{1}{\rho}\frac{\partial \rho}{\partial p}. \tag{2.19}$$

Using equations 2.18 and 2.19, equation 2.17 becomes:

$$\frac{\partial(n\rho)}{\partial t} = \rho[\xi(1-n) + \zeta n]\frac{\partial p}{\partial t} + n\frac{\partial \rho}{\partial \mathcal{C}}\frac{\partial \mathcal{C}}{\partial t}. \tag{2.20}$$

The term $\rho[\xi(1-n) + \zeta n]$ represents the volume of water released from storage in a unit volume of a confined and compressible porous aquifer per unit change in pressure. Let $S_p = [\xi(1-n) + \zeta n]$ be the specific storage in terms of pressure $[M^{-1}LT^2]$. With this definition, equation 2.20 gives:

$$\frac{\partial(n\rho)}{\partial t} = \rho S_p\frac{\partial p}{\partial t} + n\frac{\partial \rho}{\partial \mathcal{C}}\frac{\partial \mathcal{C}}{\partial t}. \tag{2.21}$$

The first term at the right-hand side of equation 2.21 is the rate of fluid mass accumulation due to fluid pore pressure change, while the second one is the rate of fluid mass accumulation due to the change of solute concentration. Substituting equation 2.21 into equation 2.13, the flow equation finally becomes:

$$\rho S_p\frac{\partial p}{\partial t} + n\frac{\partial \rho}{\partial \mathcal{C}}\frac{\partial \mathcal{C}}{\partial t} = -\mathrm{div}(\rho\mathbf{q}) + \rho_s q_s, \tag{2.22}$$

which is the *governing equation for variable-density groundwater flow in porous media*.

2.4.2 Darcy's law for variable-density groundwater flow

The governing equation for variable-density groundwater flow 2.22 includes a term for specific discharge (**q**), which is calculated with Darcy's law.

In order to get Darcy's law for variable-density flows into coastal aquifers affected by seawater intrusion, we should consider again the variation of Hubbert's potential, as expressed by equation 2.7. In this case, we will assume again that the average fluid velocity into the porous medium is negligible, while we will take into account the chemical potential due to the presence of salt ($NaCl$) dissolved in water. We can explicit the mole number of salt (n_{NaCl}) as follows:

$$n_{NaCl} = \frac{m_{NaCl}}{M_{NaCl}} = \frac{\mathcal{C}_{NaCl}V_{sol}}{M_{NaCl}}, \tag{2.23}$$

where m_{NaCl} and $M_{NaCl} = 58\ \mathrm{g/mol}$ are the mass of dissolved $NaCl$ and its molecular mass, respectively, while $\mathcal{C}_{NaCl}$ is the salt concentration within the solution $[ML^{-3}]$, expressed as the ratio between its mass (m_{NaCl}) and the volume of the solution (V_{sol}).

Equation 2.23 allows obtaining:

$$dn_{NaCl} = \frac{V_{sol}}{M_{NaCl}}d\mathcal{C}_{NaCl},$$

so that the variation of Hubbert's potential can be expressed as:

$$d\phi^* = dz + \frac{dp}{\rho_{sol}g} + \frac{\varphi}{\rho_{sol}gM_{NaCl}}d\mathcal{C}_{NaCl},$$

where ρ_{sol} is the density of the solution, i.e.:

$$\rho_{sol} = \frac{m_{sol}}{V_{sol}} = \frac{m_f + m_{NaCl}}{V_{sol}} = \frac{m_f}{V_{sol}} + \frac{m_{NaCl}}{V_{sol}} = \rho_f + \mathcal{C}_{NaCl}, \tag{2.24}$$

where the subscript f refers to freshwater. From equation 2.24, because ρ_f assumes a constant value, we can infer that $d\mathcal{C}_{NaCl} = d\rho_{sol}$, so that the variation of Hubbert's potential becomes:

$$d\phi^* = dz + \frac{dp}{\rho_{sol} g} + \frac{\varphi}{g M_{NaCl}} \frac{d\rho_{sol}}{\rho_{sol}}.$$

Again, by considering the reference state (pure water, at rest, at the mean sea level, at atmospheric pressure and at the same temperature of the underground), the final equation for Hubbert's potential for variable-density flow is:

$$\phi^* = \int_0^z dz + \int_{p_a}^{p} \frac{dp}{\rho(p) g} + \frac{\varphi}{g M_{NaCl}} \int_{\rho_f}^{\rho} \frac{d\rho}{\rho} = z + \int_{p_a}^{p} \frac{dp}{\rho(p) g} + \frac{\varphi}{g M_{NaCl}} \ln \frac{\rho}{\rho_f},$$

where we used ρ rather than ρ_{sol} for the sake of simplicity. The last term at the right-hand side allows coupling flow and solute-transport processes.
The density of the solution (ρ) generally depends on pressure, but if we assume that the fluid solution is incompressible, Darcy's law for variable-density flow becomes:

$$\mathbf{q} = -\kappa \frac{\rho g}{\mu} \mathbf{grad} \left(z + \frac{p - p_a}{\rho g} + \frac{\varphi}{g M_{NaCl}} \ln \frac{\rho}{\rho_f} \right),$$

where we used equation 2.4 for hydraulic conductivity.

2.4.3 Governing equation for solute transport

For the content of this subsection refer to Giudici [59] and Zheng & Wang [148].

For variable-density groundwater conditions, flow and solute transport are linked processes. So, in addition to the flow equation 2.22, a second equation is required to describe *solute transport* in the aquifer. In fact, groundwater flow causes the redistribution of solute concentration and this in turn alters the density field, thus affecting groundwater movement. Therefore, the movement of groundwater and the transport of solutes in the aquifer are coupled processes.
Solute mass is transported in porous media by average groundwater flow (advection), molecular diffusion, hydrodynamic dispersion and reactions of chemical (e.g., redox reactions), biological (e.g., biodegradation) and physical (e.g., radioactive decay) nature.
Advection describes solute transport due to the fluid flow as a whole, so occurring at the same average velocity as the groundwater, and it generally is the dominant mechanism. We can express advective flux of a solute through a unit surface area of the porous medium as:

$$\mathbf{q}_{adv} = \mathcal{C}\mathbf{q} \qquad [ML^{-2}T^{-1}], \tag{2.25}$$

where $\mathcal{C}$ is the mass concentration of solute within the solution $[ML^{-3}]$ and $\mathbf{q}$ is the specific discharge of the solution itself within the porous medium.

Molecular diffusion mechanism is due to molecular motion at the microscopic scale and becomes relevant when the fluid is at rest. The flux of a solute through a unit surface area of the porous medium caused by molecular diffusion can be expressed by *Fick's law*:

$$\mathbf{q}_{diff} = -nd\nabla\mathcal{C} \qquad [ML^{-2}T^{-1}], \tag{2.26}$$

where n is the medium porosity and d is the so-called *coefficient of molecular diffusion* $[L^2T^{-1}]$.
Hydrodynamic dispersion is caused by deviations of velocity at the microscopic scale from the average groundwater velocity which induces advection. This is caused by heterogeneities in the velocity field at the microscopic (i.e., inside the pores) and mesoscopic (i.e., among the pores) scales and heterogeneities of the conductivity field at fine scales. The flux of a solute through a unit surface area of the porous medium caused by hydrodynamic dispersion can be expressed as:

$$\mathbf{q}_{disp} = -n\mathbf{D}\nabla\mathcal{C} \qquad [ML^{-2}T^{-1}], \tag{2.27}$$

where D is the *hydrodynamic dispersion tensor* $[L^2T^{-1}]$. It is a symmetric second-order tensor admitting principal axes. The axis along which the *longitudinal dispersion* (i.e., its maximum value D_L) is obtained, corresponds to the direction of Darcy's velocity, whereas the tensor is isotropic in the plane perpendicular to $\mathbf{q}$, and the corresponding components are given by the *transversal dispersion* (D_T). As a first approximation, D_L and D_T are proportional to the average velocity of fluid ($\mathbf{v}$), so that $D_L = \alpha_L\mathbf{v}$ and $D_T = \alpha_T\mathbf{v}$, where α_L and α_T are the *longitudinal and transversal dispersivities*, respectively (generally $\alpha_T \ll \alpha_L$). The parameter α_L can be interpreted qualitatively as a measure of the scale length of the heterogeneity of the flow field, which is not captured by the convective term and whose effects on transport are described by the hydrodynamic dispersion.
Now consider a porous medium saturated with a certain solution. We can write the mass conservation law as:

$$\frac{d}{dt}\int_\Omega n\mathcal{C}d^3\mathbf{x} = \int_{\partial\Omega}(\mathbf{q}_{adv} + \mathbf{q}_{diff} + \mathbf{q}_{disp})\cdot\mathbf{n}^{in}d^2\mathbf{x} + \int_\Omega nQ_\mathcal{C}d^3\mathbf{x},$$

where Ω, $\partial\Omega$ and $\mathbf{n}^{in}$ are the same as defined in subsection 2.3.1 and $Q_\mathcal{C}$ represents the production rate of solute due to eventual reactions inside Ω $[ML^3T^{-1}]$.
By considering the porosity as independent of time, the divergence theorem and the arbitrary nature of Ω give:

$$n\frac{\partial\mathcal{C}}{\partial t} = -\nabla\cdot(\mathcal{C}\mathbf{q} - n\mathbf{D}^*\nabla\mathcal{C}) + nQ_\mathcal{C}, \tag{2.28}$$

where we used equations 2.25, 2.26 and 2.27 and we introduced $\mathbf{D}^* = d\mathbf{I} + \mathbf{D}$.
When considering solute transport in porous aquifers, we should take into account eventual variations of solute mass related to source terms. For this reason, we should introduce an additional term to the right-hand side of equation 2.28:

$$n\frac{\partial\mathcal{C}}{\partial t} = -\nabla\cdot(\mathcal{C}\mathbf{q} - n\mathbf{D}^*\nabla\mathcal{C}) + nQ_\mathcal{C} - q_s\mathcal{C}_s, \tag{2.29}$$

where $\mathcal{C}_s$ is the solute concentration of the source or sink flux $[ML^{-3}]$ and q_s is the volumetric flow rate per unit volume of aquifer representing fluid sources or sinks $[T^{-1}]$.

2.4.4 The Ghyben-Herzberg approximation

Since freshwater and seawater are miscible fluids, the contact zone between them takes the form of a *transition zone* caused by mixing and molecular diffusion. Across this zone, the salinity of the mixed water varies from that of freshwater (about $0.4\ \mathrm{g/L}$) to that of seawater (about $40\ \mathrm{g/L}$). However, under certain conditions, the width of this zone is relatively small (e.g., when compared with the thickness of the aquifer), so that a *sharp interface approximation* can be introduced. This assumption greatly simplifies the problem in many cases of practical interest.
Beginning with Ghyben (1888) and Herzberg (1901), investigations of the interface in a coastal aquifer have aimed at determining the relationship between its shape and position.
The calculation of the depth of such interface was carried out by the two researchers assuming static equilibrium between two water columns with different densities (see Figure 2.2). Furthermore, some approximations were introduced:

- freshwater and saltwater are immiscible fluids, so that a sharp interface separates them, as already mentioned;
- vertical flow components for freshwater can be neglected;
- the seawater is stationary.

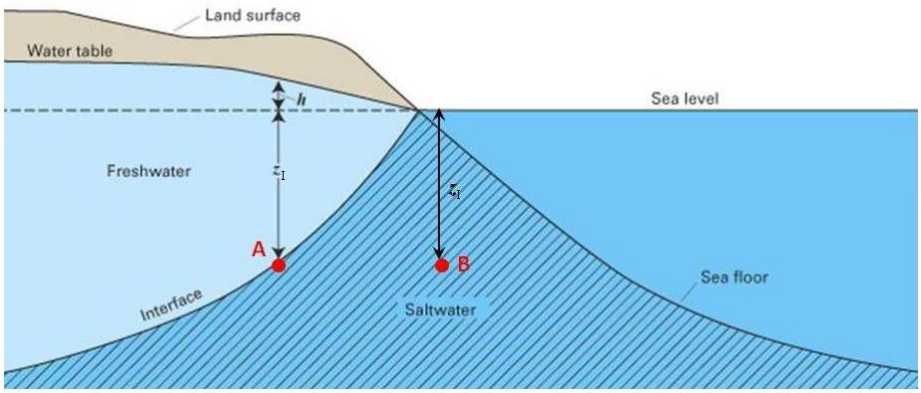


Figure 2.2: Schematic representation of the equilibrium relationship between freshwater and saltwater according to Ghyben and Herzberg [75].

Let A and B be two points, located at the depth z_I below the mean sea level, at the salt-fresh water inland interface and in the aquifer below the sea floor off-shore, respectively (Figure 2.2). According to hydrostatic pressure law, the pressure due to the freshwater column at point A is given by:

$$p_A = p_a + \rho_f g(z_I + h),$$

where p_a is the atmospheric pressure, ρ_f is the freshwater density and h is the hydraulic head of the aquifer at point A.

In the same way, the pressure due to the seawater column at point B is given by:

$$p_A = p_a + \rho_s g z_I,$$

where ρ_s is the seawater density.
If the seawater is stationary, p_A and p_B must be equal, so that:

$$\rho_f(z_I + h) = \rho_s z_I$$

and

$$z_I = \frac{\rho_f}{\rho_s - \rho_f} h = \delta h. \tag{2.30}$$

This is the *Ghyben-Herzberg approximation*, according to which the depth (below the mean sea level) of the interface between freshwater and seawater at each point of the aquifer is related to the hydraulic head at that point. The dimensionless proportionality coefficient (δ) depends on the densities of freshwater and seawater and varies between 33 and 50 [108]. For Italian seas, common estimations are $\rho_f = 1.000$ $\mathrm{g/cm^3}$ and $\rho_s = 1.027$ $\mathrm{g/cm^3}$ [130], so that $\delta = 37$: this means that, at each point of the aquifer, the depth of the stationary interface below the mean sea level is 37 times the height of the freshwater table above it.

Limitations of the Ghyben-Herzberg approximation

The Ghyben-Herzberg approximation has some limitations related to the hypoteses introduced to obtain it.
First of all, as already assessed, freshwater and seawater are miscible fluids, so that the interface between them is actually a transition zone within which the salinity increases with depth. The most important reasons why this zone exists and extends vertically and horizontally are the fluctuations of both the hydraulic head of the aquifer (due to withdrawals and variations of recharge terms) and the mean sea level (due to tides and climate changes). These fluctuations lead seawater to continuously go up and down into the aquifer, so causing a mixing with freshwater: this phenomenon, togheter with molecular diffusion, contributes to increase the thickness of the transition zone.
Moreover, as the sea is approached, the assumption of horizontal flow into the aquifer is no longer valid, because vertical flow components can no longer be neglected.
Finally, seawater cannot be considered as stationary, especially nearby the coast, where a cyclic flow of seawater can occur locally through coastal springs. These latter, in fact, normally drain brackish water from the transition zone: the amount of saltwater so conveyed towards the sea should be replaced by seawater intruding inland. In this way, a cyclic flow of seawater occurs through local springs, at least nearby the coast.
Despite these limitations, the Gyben-Herzberg approximation can be valid far from the coast, where the thickness of the transition zone can be considered as negligible with respect to the thickness of the aquifer.

Chapter 3

Discretization of natural systems and related processes

For this chapter refer to Giudici [60] and Harbaugh [69], unless otherwise specified.

Natural systems (e.g., biological, geological, ecological) and the related relevant processes are often very complex to reproduce accurately. For this reason, they can be described only by using simplifications and approximations through the development of *simulation models*, which in turn reflect the complexities of the simulated processes. As a first step towards the description of a complex process, a clear idea about the aim of the model is of paramount importance, in order to correctly define the structure of the model itself. In this regard, a model can be used to deepen the knowledge about a physical process for which a theory has been developed. The outcomes of a model can also help in decision support, by simulating different scenarios which could affect the evolution of the physical system under exam. Furthermore, a model can provide useful indications in order to design monitoring and testing activities.
A complete mathematical and physical setup of a simulation model consists of several phases, listed below:

a. specifying the *geometry* of the model domain and the *time interval* within which the physical process will be simulated;

b. defining the *characteristic space* and *time scales* of the model, which should be consistent with spatial and temporal distribution of the available information;

c. defining a *conceptual model*, namely a schematic representation of the physical system under exam and the processes related to it;

d. specifying the *initial conditions* at some initial time;

e. specifying the *boundary conditions* at the border of the model domain;

f. defining properties and parameters involved in the description of the system itself, which (eventually) need a *sensitivity analysis* and a *calibration* phase.

Obviously, while defining the general structure of the model through these steps, it is very important that the approximations so introduced are consistent with the purpose of the model itself and the spatial and temporal distribution of the available information.
Once defined the aim and the structure of the simulation model, the proper equations describing the physics of the process under exam need to be identified. Generally speaking, these are partial differential equations based on a set of assumptions which simplify the complex real system. In principle, two classes of methods exist for solving such equations.

- ***Analytical methods***. In this case, the equations are solved analytically, but this is possible just for problems when homogeneous media with simple geometry and boundary conditions are involved. So, these methods are often used for the implementation of simple experiments or to provide useful indications about the order of magnitude of the effects related to phenomena which, eventually, should be better analysed.

- ***Numerical methods***. In this case, the equations are replaced with systems of discrete, possibly non linear, algebraic or trascendent equations, solved by numerical algorithms.

Furthermore, simulation models are classified according to the approach to the problem as follows.

- ***Deterministic models***. In this case, once given boundary and initial conditions, an input stress causes a (generally unique) response of the physical system, which can be simulated through the resolution of proper equations.

- ***Stochastic models***. In this case, the response of the physical system to an input stress is simulated togheter with the relative uncertainty, due to the approximations introduced in the description of the system itself.

In the following sections, the basic information about numerical and deterministic models of groundwater flow developed and/or applied in this work are described.

3.1 Discretization of the model domain

The first approximation necessary for the analysis of natural systems and processes is the discretization of the physical system with a *computational grid*, namely a discrete set of *cells* identified by three labels (i, j, k) in the space domain and a label (l) in the time domain. The center of a cell is called a *node* and is identified by means of the same labels. Such a discretization results in a simplification of the physical system because, in this way, its properties are no longer continuous variables, but they are known just at each node of the grid and at certain time instants.
More specifically, in view of the successive discretization of the equations with the finite-difference method (see section 3.2.1), consider a 3D computational grid composed by $Nx \times Ny \times Nz$ cells, each of them identified by labels (i, j, k), where $i = 1, ..., Ny$ identifies rows in the horizontal plane, $j = 1, ..., Nx$ identifies columns in the horizontal plane and $k = 1, ..., Nz$ identifies vertical layers. Let

Δx, Δy and Δz be the sizes of each cell along the x, y and z directions, respectively. Notice that it is not necessary to have constant dimensions of all cells over the whole grid. This means that, in the horizontal plane, the grid can be locally made of smaller cells in zones where a greater detail is required, while in the vertical plane, it can consists of non flat layers. Furthermore, for the sake of simplicity, rows, columns and layers are assumed to be aligned along the three Cartesian directions, even if it is not necessary in general.

Especially when defining a grid with a complex geometry, it is necessary to distinguish active cells, non-active cells and boundary cells (see Figure 3.1). Active cells are those internal to the model domain and within them the related physical processes are simulated by solving proper equations. For example, if groundwater flow dynamics is considered, in active cells the hydraulic head can vary with time and according to the water exchange with neighboring active cells, quantifyied by the balance equation 2.11. Non active cells are those external to the model domain and for them the simulation equations are not solved. So, if groundwater flow dynamics is considered, no water exchange occurs between an active cell and a non-active one. Finally, boundary cells are internal to the model domain, but particular boundary conditions are imposed as explained in detail in subsection 3.2.2.

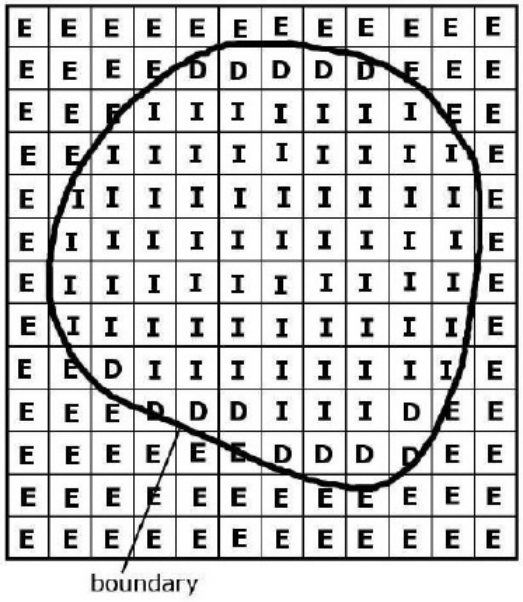


Figure 3.1: Example of a 2D model grid. Labels I, E and D denote active, non-active and boundary cells, respectively.

3.2 Discretization and numerical resolution of groundwater flow equations

Once an approximation of the geometry of the natural system is introduced through a computational grid, all relevant properties of the system are now discrete spatial and temporal functions, defined at each node

of the grid and at certain time instants. Therefore, a discrete approach is necessary for the analysis of the physical processes, which require proper equations referred at each node of the grid and at certain time instants, resulting in a set of discrete equations. The dependent variables obtained from the solution of such system of equations represent average values within respective cells.
Several numerical techniques can be implemented for the discretization of continuous equations (e.g., finite-difference method, finite-elements method, boundary-elements method, finite-volumes method). For the aims of this thesis, the finite-difference method will be introduced for the discretization of the groundwater flow equations discussed in chapter 2.

3.2.1 The finite-difference method

Natural processes are often described by differential equations, whose discretization requires the discretization of differential operators. The main idea of the *finite-difference method* consists in replacing the derivatives of continuous functions (describing physical properties on the whole space-time domain), with incremental ratios of discrete functions (describing physical properties at the nodes of a grid and at certain time instants).
If we want to characterize the groundwater flow dynamics in a porous aquifer, we must adapt equation 2.11, which formalizes the physics of this process, to a discrete form within each cell of the domain and for each time-step. For this goal, we should remember that equation 2.11 was derived by using the mass conservation law, equation 2.9, where the flow density (ρ) can be simplified if the fluid is considered as incompressible. Recalling the definition of specific discharge (S_0) and that the term at the right-hand side of equation 2.9 is nothing but the net mass flux through the boundary surface of the porous domain, we can write the discrete form of equation 2.11 within each cell as follows:

$$SS\frac{\Delta h}{\Delta t}\Delta V = \sum_n Q_n + F, \tag{3.1}$$

where SS is the discrete formulation of the specific discharge coefficient (S_0), Δh is the head change into the cell over the time-step of length Δt, $\Delta V = \Delta x \Delta y \Delta z$ is the volume of the cell, $\sum_n$ is referred to all the faces of the cell ($n = 1, ..., 6$), Q_n is the flow rate into or out of the cell through its n-th face and F is the sum of the source terms inside the cell itself.
Each of the terms Q_n ($n = 1, ..., 6$) can be computed by using the discrete formulation of Darcy's law 2.3 and considering the involved cell and one of the six neighbouring cells. For example, the inflow rate across the face separating the cell C and the next cell toward north, denoted as N, can be expresses as:

$$Q_n = K_n \cdot A \cdot \frac{(h_N - h_C)}{\Delta x},$$

where K_n is called the *interblock hydraulic conductivity*, namely the hydraulic conductivity along the direction perpendicular to the considered face, A is the area of the cell face and $(h_N - h_C)$ is the head change between the nodes C and N, placed at a distance Δx from each other.
In order to approximate the time derivative of head, let $t^{(l)}$ denote the time instant at which the flow terms

are evaluated and let $t^{(l-1)}$ be the time instant preceding $t^{(l)}$, so that $\Delta t = t^{(l)} - t^{(l-1)}$. So, the head values within a cell associated with these time instants are defined by superscripts as $h^{(l)}$ and $h^{(l-1)}$. In this way, the time derivative of the head in equation 3.1 can be replaced by the ratio $(h^{(l)} - h^{(l-1)})/\Delta t$. Equation 3.1 cannot be solved independently at each node, because it is a single equation in seven unknowns, i.e., the hydraulic head at time instant $t^{(l)}$ at the involved cell and at its six neighbouring cells. However, equation 3.1 can be written similarly for each active cell of the grid and, since only one unknown (hydraulic head) exists for each cell, if we assume that SS, interblock hydraulic conductivities and F are known, we will have a system of m equations in m unknowns, where m is the number of active grid cells.
The entire system of equations may be written in matrix form as:

$$\mathbf{Ah} = \mathbf{b}, \tag{3.2}$$

where **A** is the symmetric and diagonally dominant matrix of the coefficients of head at each node of the grid, **h** is the vector of head values at the end of the time-step l for all nodes of the grid and **b** is the vector of the constant terms at each node of the grid.
Numerical codes often use iterative methods to obtain the solution to the system of finite-difference equations for each time-step. In these methods, the calculation of head values at the end of a given time-step is started by assigning a trial value, or estimate, for the head at each node of the grid and at the beginning of that time-step. These estimated values are successively modified, producing a new set of head values that are in closer agreement with the system of equations. These new head values then take the place of the initially assumed heads and the procedure of calculation is repeated, producing a third set of head values. This procedure is repeated successively at each stage, producing a new set of heads that more nearly satisfies the system of equations. Each repetition of the calculation is termed an *iteration*. Ultimately, the changes produced by succeeding stages of calculation become very small. This behavior is used to determine when iterations can be stopped.
Ideally, one would like to specify that iterations stop when the calculated heads are suitably close to the exact solution. Since the actual solution is unknown, however, an indirect stopping criterion must be used. The method most commonly employed is to specify that the changes in computed head occurring from one iteration to the next one must be less than a certain prescribed quantity, termed the *convergence* (or *closure*) *criterion* (or *threshold*), which is specified by the user. After each iteration, absolute values of computed head change in that iteration are examined for all nodes of the grid and the largest of these absolute head change values is compared with the convergence criterion. If the largest value exceeds the convergence criterion, iterations continue, otherwise the iterative process has *converged* for that time-step. As a rule of thumb, it is wise to use a value of convergence criterion that is an order of magnitude smaller than the level of accuracy desired in the head results. The initial estimate of head distribution may affect the number of iterations required to obtain an acceptable approximation of the solution.
Besides the convergence criterion, it is possible to define also a maximum permissible number of iterations, which is specified by the user. If convergence is not achieved within this maximum number of iterations, then the iterative process is terminated.
The discrete flow equation 3.1 was developed assuming transient conditions. A steady-state problem is

defined when the storage term is zero, so that the sum of all inflows (outflows are nothing but negative inflows) from adjacent cells and external stresses must be zero for each cell of the model. A steady-state problem requires only a single solution of simultaneous equations, rather than multiple solutions for multiple time-steps. Since iterative methods work by successively improving the estimated head values, also in this case an initial estimate is required to start the iterative process. Then, the same discussion about the convergence of the iterative method can be repeated as in transient problems.

3.2.2 Initial and boundary conditions

The discrete balance equation 3.1 presented in the previous section derives from a partial differential equation which contains no information related to any specific case of flow through a porous medium, e.g., the shape of the flow domain. Therefore, this equation has an infinite number of possible solutions, each of which corresponds to a particular case of flow through a porous medium domain.
A mathematical problem that corresponds to a physical reality should satisfy the following basic requirements:

- the solution must exist (*existence*);
- the solution must be uniquely determined (*uniqueness*);
- the solution should depend continuously on the data (*stability*). This means that a variation of the given data in a sufficiently small range leads to a small change in the solution.

Any problem that satisfies these three requirements is called a *well-posed problem*.
Regarding the uniqueness, among all possible solutions, obtaining one particular solution to a specific problem is possible by providing supplementary information not contained into the equation itself. Such information should include specifications of:

- the geometry of the domain in which the flow takes place;
- values of all relevant physical coefficients (e.g., K, S_0);
- the initial state of the fluid in the flow domain (*initial conditions*);
- statements on how the fluid in the domain interacts with its surroundings, namely conditions on the boundaries of the flow domain (*boundary conditions*).

The first two items have already been discussed in the previous section, while boundary and initial conditions for flow through porous media should take into account the physical reality of the flow. So, they are first determined from field data or assumed on the basis of available information and past experience, and then expressed in mathematical forms. Notice that different initial and boundary conditions lead to different solutions, hence the importance of correctly determining the proper conditions along the real boundaries.
Initial conditions include the specification of h at all points within the domain at some initial time (t^0).

With this specification, the iteration process can start in order to evaluate head distributions at successive time instants. As already mentioned, an initial estimate for h is required also for steady-state problems, but in that case this initial guess is necessary only to start the iterative method of solution of the system of equations and it does not have a strict physical meaning.
Boundary conditions assigned along the boundaries of the domain (boundary cells) are classified as follows. Of course, different boundary conditions can be assigned to different parts of the domain boundary.

- ***Dirichlet boundary conditions*** (***prescribed head***). In this case, the head values (h) are fixed along the boundary and cannot vary during the computation of the solution to the system of equations. This kind of boundary condition can be assigned whenever a perfect hydraulic contact occurs between the flow domain (the aquifer) and a superficial water body (e.g., sea, river, lake). Namely, if no impermeable layer separates the aquifer and the superficial body, these can exchange water, so that an equilibrium exists between their water levels. Under this configuration, the water level along the aquifer boundary is forced to be the same as that of the superficial water body.

- ***Neumann boundary conditions*** (***prescribed flux***). In this case, the component of ∇h perpendicular to the boundary is prescribed for all points and cannot vary during the computation of the solution to the system of equations. According to Darcy's law 2.5, if the hydraulic conductivity is known, this means forcing the flux normal to the boundary surface. For example, one can impose no flow wherever there are barriers to the hydraulic flow (e.g., impermeable layers).

- ***Robin boundary conditions***. It is possible to merge the previous conditions by assigning the coefficient of a linear combination of h and ∇h along the boundary. For example, one can impose Robin boundary condition if less permeable sediments prevent from a perfect hydraulic contact between the aquifer and a superficial water basin: in this case, the exchange flow rates linearly depend on the difference between the water levels in the aquifer and in the superficial body.

Besides groundwater flow, if the solute-transport process is considered, initial and boundary conditions must be specified also for this particular process. In this case, mathematical boundaries are commonly defined in three categories:

- a Dirichlet boundary is one in which the value of solute concentration is specified at all points along the boundary. This implies that the dispersive flux toward or away from the boundary occurs in response to the difference between the specified boundary concentration and the calculated concentration at points directly adjacent to the boundary itself;

- a Neumann boundary represents the condition in which the concentration gradient is specified normal to the boundary: this results in a specified dispersive flux of solute across the boundary itself;

- a Cauchy boundary represents a boundary on which both concentration and concentration gradient are specified. This results in a specified total flux of solute mass across the boundary itself and

this condition in the transport model coincides with a Neumann boundary condition in the flow regime.

3.3 Model calibration

A common goal of modeling is to provide accurate predictions and clear guidance for future planning. Thus, models can quantitatively test ideas about properties and processes in a manner which is not possible for complex systems by any other method. While this contribution to the toolbox of understanding natural systems is important, work over the last 30 years has shown that models typically do not achieve the level of prediction accuracy (i.e., the closeness of model predictions to real occurring processes) that was idealized. For this reason, understanding which system features are likely to be important and unimportant (including identifying dominant parameters, important observations, and controlling processes) is of paramount importance in order to provide insights into model performance [100].
During calibration, model inputs such as the values of parameters used to quantify physical properties, initial and boundary conditions, and stresses are modified so that the model output matches related measured values. The model inputs that need to be estimated are often spatially and/or temporally distributed, so that the number of parameter values could be infinite, in principle. The observations, however, are generally limited in number and support the estimation of relatively few parameters, often over scales very different from the model scale. For example, hydraulic conductivity values used for a groundwater flow model at the regional scale must allow the simulation of flow over length scales which are much greater than the characteristic scales of pumping tests.
Historically, observed and simulated quantities, such as hydraulic heads, flows, and solute concentrations for groundwater systems, were often compared subjectively, so that it was difficult to determine how well one model was calibrated relative to another. In addition, adjustments of parameter values were accomplished mostly by trial-and-error method, by comparing observations with several outcomes simulated by modifying input data, until the achievement of a satisfactory agreement [72].
Formal methods have been developed that attempt to estimate parameter values, given a mathematical model of system processes and a set of relevant observations. These are called *inverse methods*, and generally they are limited to the estimation of parameters as defined above. Thus, the terms “inverse modeling” and “parameter estimation” are commonly synonymous [72].
For some models, the inverse problem is linear, i.e., the predicted quantities are linear functions of the parameters. In many circumstances of practical interest, however, the inverse problem is non-linear, and its solution is not as straightforward as for linear problems. The simplest form of the groundwater flow equation, Darcy’s law, can be used to demonstrate the linearity of hydraulic head with respect to the spatial dimension and flow rate and the non-linearity with respect to hydraulic conductivity. Consider Darcy’s law in the form 2.3, where $i = \frac{\partial h}{\partial x}$. By integrating both sides between $x = 0$ and any distance $x = X$, under the conditions of stationary flow and of absence of external source/sink terms, we get:

$$h = h_0 - \frac{Q}{KA}X, \tag{3.3}$$

where h and h_0 are the hydraulic heads at $x = 0$ and $x = X$, respectively. By using partial derivative notation, the derivatives of equation 3.3 with respect to X, Q and K are:

$$\frac{\partial h}{\partial X} = -\frac{Q}{KA},$$
$$\frac{\partial h}{\partial Q} = -\frac{1}{KA}X,$$
$$\frac{\partial h}{\partial K} = -\frac{Q}{K^2A}X.$$

The hydraulic head is a linear function of X and Q because $\frac{\partial h}{\partial X}$ is independent of X and $\frac{\partial h}{\partial Q}$ is independent of Q, respectively. However, hydraulic head is a non-linear function of K because $\frac{\partial h}{\partial K}$ is a function of K [72].

The complexity of many real systems and the scarcity of available data result in inversions that are often plagued by problems of insensitivity, non-uniqueness, and instability, regardless of how model calibration is achieved. Insensitivity occurs when the observations do not contain enough information to support estimation of the parameters. Non-uniqueness occurs when different combinations of parameter values match the observations equally well. Instability occurs when even very small changes in observations radically change the inversion results. All these problems are exacerbated when the system is non-linear [72].

As already said, a discrete model can be described by a set of equations linking model parameters and the state of the system. These discrete equations can be generally formulated as:

$$\mathbf{F}(\mathbf{p}^{(fix)}, \mathbf{p}^{(cal)}, \mathbf{s}) = \mathbf{0},$$

or

$$\mathbf{s} = \mathbf{G}(\mathbf{p}^{(fix)}, \mathbf{p}^{(cal)}), \tag{3.4}$$

where $\mathbf{p}^{(fix)}$ and $\mathbf{p}^{(cal)}$ consist of fixed parameters and parameters which need calibration, respectively, while $\mathbf{s}$ consists of quantities which describe the state of the system and operators $\mathbf{F}$ and $\mathbf{G}$ describe the relationship between them. Calibration requires the solution of equation 3.4 for $\mathbf{p}^{(cal)}$, which is often non-linear, on the basis of measurements and observations about the state of the system, and this is called an inverse problem.

We can easily infer that calibration depends on observations, but also on hypotheses and approximations introduced during the definition of the model, so that if calibration provides inconsistent outcomes, previous steps should be revised.

For the aims of this thesis, the classification made by Neuman [95], Yeh [144] and Sun [123] between indirect and direct methods is taken into account.

Indirect inverse modeling allows to condition parameter values and model construction on measurements of model output quantities [52], compared with the available observation data, and this is done by minimizing an objective function, i.e., a measure of misfit between observed and simulated values (residuals).

Generally, residuals are weighted to take into account the different units and precision of different contributions to the objective function itself: in this way, residuals on both head and flow observations can be combined in a single objective function [53]. Indirect methods include also the computation of statistics for model evaluation [72], thanks to which the modeler can explore the relations between different types of data and the processes represented in the model, including the testing of different hypotheses about system structure [52].
In *direct inverse modeling*, instead, no objective function is postulated and measured heads and flow rates are used to directly compute conductivities as the solution of the standard balance equation 2.11 [19]. Although the theory is straightforward, it is often impossible to obtain a realistic solution of the inverse problem, due to serious ill-posedness and singularity of the matrices involved in the numerical formulation [149]. The main shortcoming of direct methods is that they require that hydraulic heads are available for all nodes of the discretized system. Furthermore, despite a direct method is much cheaper in terms of computation times, direct inversion may become mathematically unstable or ill-conditioned, which means that small errors in head data can result in large errors in hydraulic conductivity [124].
In subsections 4.2.5 and 4.4.3, a direct and an indirect calibration methods will be described and they will be used for application to a real case study in part II.

3.4 Sensitivity analysis

As already assessed at the beginning of this chapter, the outcomes of a simulation model can be useful in decision support, both by simulating different scenarios or providing information in order to design monitoring and testing activities. Therefore, these outcomes should be as accurate as possible, namely as close as possible to the actual simulated processes. To achieve this objective, taking into account the uncertainties of input parameters is of paramount importance, because they reflect on the model outcomes. However, the non linearity of simulated processes makes difficult the identification of relevant parameters. In this framework, *sensitivity analysis* aims at quantifying the reliability of model predictions and identifying which parameters mostly affect them and require a better estimate through calibration.
In order to determine how well parameters are estimated, some statistics can be calculated after calibration is successfully completed. For models with lengthy execution times, methods that do not require completion of calibration can be very helpful, as exclusion of insensitive parameters from the calibration phase improves estimation performance, rarely affects results and reduces execution times. These sensitivity analysis methods are based on observations, which provide information about parameter definition and model construction [72].
Examples of some statistics used will be provided in subsection 4.4.2.

Chapter 4

Numerical codes applied for this study

Since early 1990s, MODFLOW [69] has been among the most powerful and widely applied groundwater flow models. It is based on a conservative finite-difference scheme, for which the balance equation 2.11 is applied to a 3D network within porous media. MODFLOW's capabilities have been further expanded over the years including, among the others, transport simulation and parameter estimation.
Here, the basics of the MODFLOW code will be discussed, in order to show how the equations introduced in the previuos chapter can be arranged by using the formalism of finite-difference method. The same formalism will be used also to introduce the basics of the YAGMod and SEAWAT codes applied to a real case study, as described in detail in Part II. The last section of this chapter will be dedicated to the UCODE software for sensitivity analysis and calibration of the developed models.

4.1 MODFLOW

As already mentioned at the beginning of this chapter, *MODFLOW* can handle several physical processes and features that control groundwater flow models, all integrated into a single program. For this reason, the algorithm is organized into "processes", within which proper equations are solved according to the simulated process (e.g., flow or transport).
In this section we will refer to the *GroundWater Flow* (*GWF*) *Process*, within which the groundwater flow equation 2.11 is solved. What follows is based on Harbaugh [69], unless otherwise specified.
In the GWF Process, MODFLOW simulates the movement of groundwater of constant density through porous media by solving equation 3.1 within each cell of the model domain and at each time instant.
To introduce a more practical formulation of equation 3.1, consider the 3D computational grid defined in section 3.1 and, as a convention, assume that increments in the row label (i) would correspond to decreases in y, while increments in the column label (j) would correspond to increases in x and layer number (k label) from the top down, so that increments in the k label would correspond to decreases in elevation (Figure 4.1). Moreover, let l be the label for the time discretization of the process under exam: $l = 1, ..., Nt$, where Nt is the maximum number of time-steps whose time spacing is Δt. Therefore, $Nt \times \Delta t$ denotes the time length of the simulation.

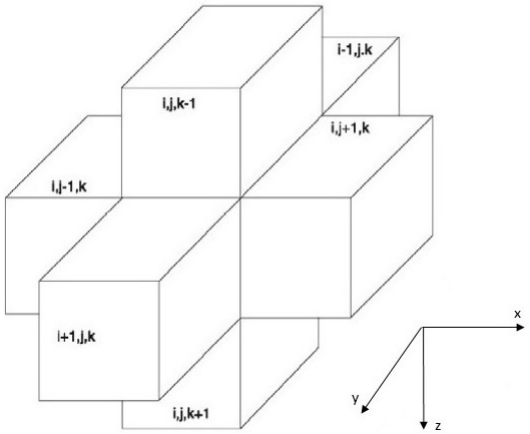


Figure 4.1: Labels for the six neighboring cells surrounding the cell (i, j, k) (hidden). Modified after Harbaugh [69].

The terms Q_n $(n = 1, ..., 6)$ are expressed by using Darcy's law (equation 2.3) and the discretization described in section 3.1, and they represent the incoming flux in the cell (i, j, k) from its six neighboring cells $(i, j - 1, k)$, $(i, j + 1, k)$, $(i - 1, j, k)$, $(i + 1, j, k)$, $(i, j, k - 1)$ and $(i, j, k + 1)$ (Figure 4.1). Then, the flow into cell (i, j, k) along the x direction from cell $(i, j - 1, k)$ is given by:

$$Q^{(l)}_{i,j-1/2,k} = K_{i,j-1/2,k} \frac{h^{(l)}_{i,j-1,k} - h^{(l)}_{i,j,k}}{\Delta x} \Delta y \Delta z,$$

where $Q^{(l)}_{i,j-1/2,k}$ is the volumetric flow rate through the face separating cells (i, j, k) and $(i, j - 1, k)$, $K_{i,j-1/2,k}$ is called the *interblock hydraulic conductivity* along the x direction between nodes (i, j, k) and $(i, j - 1, k)$, $h_{i,j-1,k} - h_{i,j,k}$ is the head change along the x direction between nodes (i, j, k) and $(i, j - 1, k)$, placed at the distance Δx from each other, and $\Delta y \Delta z$ is the area of the cell face normal to the x direction. In the previous expression, the flow term has been specified at time $t^{(l)}$.
Similar expressions can be written to approximate the flow into the cell (i, j, k) through the remaining five faces:

$$Q^{(l)}_{i,j+1/2,k} = K_{i,j+1/2,k} \frac{h^{(l)}_{i,j+1,k} - h^{(l)}_{i,j,k}}{\Delta x} \Delta y \Delta z$$

$$Q^{(l)}_{i-1/2,j,k} = K_{i-1/2,j,k} \frac{h^{(l)}_{i-1,j,k} - h^{(l)}_{i,j,k}}{\Delta y} \Delta x \Delta z$$

$$Q^{(l)}_{i+1/2,j,k} = K_{i+1/2,j,k} \frac{h^{(l)}_{i+1,j,k} - h^{(l)}_{i,j,k}}{\Delta y} \Delta x \Delta z$$

$$Q^{(l)}_{i,j,k-1/2} = K_{i,j,k-1/2} \frac{h^{(l)}_{i,j,k-1} - h^{(l)}_{i,j,k}}{\Delta z} \Delta x \Delta y$$

$$Q^{(l)}_{i,j,k+1/2} = K_{i,j,k+1/2} \frac{h^{(l)}_{i,j,k+1} - h^{(l)}_{i,j,k}}{\Delta z} \Delta x \Delta y$$

Finally, regarding the source term (F) within the cell (i, j, k), we should account for processes and systems external to the aquifer, which may be areally distributed (e.g., recharge and evapotranspiration) or localized (e.g., wells, drains and rivers). The related flows within the cell itself may be dependent on its head (e.g., when evaluating exchange flow with a river) or not (e.g., pumping from a well). Then, the source term at the right-hand side of equation 3.1 may be represented as: $F^{(l)}_{i,j,k} = P_{i,j,k}h^{(l)}_{i,j,k} + Q_{i,j,k}$, where $P_{(i,j,k)}$ and $Q_{(i,j,k)}$ are constant quantities, whose dimensions are $[L^2/T]$ and $[L^3/T]$, respectively. With previous modifications, equation 3.1 can be written for cell (i, j, k) as follows:

$$
\begin{aligned}
&SS_{i,j,k}\frac{h^{(l)}_{i,j,k} - h^{(l-1)}_{i,j,k}}{\Delta t}\Delta x\Delta y\Delta z = \\
&K_{i,j-1/2,k}\frac{h^{(l)}_{i,j-1,k} - h^{(l)}_{i,j,k}}{\Delta x}\Delta y\Delta z + K_{i,j+1/2,k}\frac{h^{(l)}_{i,j+1,k} - h^{(l)}_{i,j,k}}{\Delta x}\Delta y\Delta z + \\
&K_{i-1/2,j,k}\frac{h^{(l)}_{i-1,j,k} - h^{(l)}_{i,j,k}}{\Delta y}\Delta x\Delta z + K_{i+1/2,j,k}\frac{h^{(l)}_{i+1,j,k} - h^{(l)}_{i,j,k}}{\Delta y}\Delta x\Delta z + \\
&K_{i,j,k-1/2}\frac{h^{(l)}_{i,j,k-1} - h^{(l)}_{i,j,k}}{\Delta z}\Delta x\Delta y + K_{i,j,k+1/2}\frac{h^{(l)}_{i,j,k+1} - h^{(l)}_{i,j,k}}{\Delta z}\Delta x\Delta y + \\
&P_{i,j,k}h^{(l)}_{i,j,k} + Q_{i,j,k}.
\end{aligned}
\tag{4.1}
$$

In MODFLOW some simplifications of the discrete equation 4.1 are introduced by defining the *interblock hydraulic conductances* $[L^2T^{-1}]$ as a combination of grid dimensions and interblock hydraulic conductivities. For example,

$$C_{i,j-1/2,k} = \frac{K_{i,j-1/2,k}}{\Delta x}\Delta y\Delta z$$

is the hydraulic conductance along the x direction between nodes (i, j, k) and $(i, j - 1, k)$, expressed as the product of interblock hydraulic conductivity along the x direction ($K_{i,j-1/2,k}$) and cross-sectional area of flow ($\Delta y\Delta z$) divided by the length of the flow path (in this case Δx, the distance between the nodes). Similar expressions can be obtained for $C_{i,j+1/2,k}$, $C_{i-1/2,j,k}$, $C_{i+1/2,j,k}$, $C_{i,j,k-1/2}$ and $C_{i,j,k+1/2}$, so that equation 4.1 can be transformed as:

$$
\begin{aligned}
&SS_{i,j,k}\frac{h^{l}_{i,j,k} - h^{l-1}_{i,j,k}}{\Delta t}(\Delta x\Delta y\Delta z) = \\
&C_{i,j-1/2,k}(h^{l}_{i,j-1,k} - h^{l}_{i,j,k}) + C_{i,j+1/2,k}(h^{l}_{i,j+1,k} - h^{l}_{i,j,k}) + \\
&C_{i-1/2,j,k}(h^{l}_{i-1,j,k} - h^{l}_{i,j,k}) + C_{i+1/2,j,k}(h^{l}_{i+1,j,k} - h^{l}_{i,j,k}) + \\
&C_{i,j,k-1/2}(h^{l}_{i,j,k-1} - h^{l}_{i,j,k}) + C_{i,j,k+1/2}(h^{l}_{i,j,k+1} - h^{l}_{i,j,k}) + \\
&P_{i,j,k}h^{l}_{i,j,k} + Q_{i,j,k}.
\end{aligned}
\tag{4.2}
$$

It is convenient to re-arrange equation 4.2 so that all terms containing heads at the end of the current time-step are grouped on the left-hand side of the equation, and all terms that are independent of head at the end of the current time-step are on the right-hand side. All coefficients of $h^{(l)}_{i,j,k}$ that do not include

conductance between nodes are combined into a single term, $HCOF$, and all right-hand-side terms are combined into the term RHS. Furthermore, the complexity can be reduced by assuming that the time superscript is l unless otherwise specified. The resulting equation is:

$$C_{i,j-1/2,k}h_{i,j-1,k} + C_{i-1/2,j,k}h_{i-1,j,k} + C_{i,j,k-1/2}h_{i,j,k-1}$$
$$-(C_{i,j-1/2,k} + C_{i,j+1/2,k} + C_{i-1/2,j,k} + C_{i+1/2,j,k} + C_{i,j,k-1/2} + C_{i,j,k+1/2} - HCOF_{i,j,k})h_{i,j,k}$$
$$+C_{i,j+1/2,k}h_{i,j+1,k} + C_{i+1/2,j,k}h_{i+1,j,k} + C_{i,j,k+1/2}h_{i,j,k+1} = RHS_{i,j,k}, \tag{4.3}$$

where

$$HCOF_{i,j,k} = P_{i,j,k} - \frac{SS_{i,j,k}\Delta x\Delta y\Delta z}{\Delta t} \qquad [L^2T^{-1}],$$

$$RHS_{i,j,k} = -Q_{i,j,k} - SS_{i,j,k}\Delta x\Delta y\Delta z\frac{h_{i,j,k}^{(l-1)}}{\Delta t} \qquad [L^3T^{-1}].$$

Following the discussion in subsection 3.2.1, this equation can be written for each active cell of the grid and the resulting system of equations can be written in matrix form (equation 3.2) and solved iteratively to get values of head at specific points and times.

The whole period of simulation is divided into a series of *stress-periods*, within which specified stress data are constant (i.e., flow rates and boundary heads). Each stress-period, in turn, is divided into a series of *time-steps*. Thus, MODFLOW is organized in order to include three nested loops: a stress-period loop, within which there is a time-step loop, which in turn contains an iteration loop to yield head values at each node at the end of each time-step. In steady-state simulations a single stress-period having a single time-step within it is defined.

Regarding the (active, non-active or boundary) status of a certain cell, MODFLOW takes as input user-defined integer codes. Active cells are called *variable-head* cells, while cells used to simulate boundary conditions are grouped into two categories: *constant-head* cells and *no-flow* cells. Constant-head cells are those for which the head is specified at each time and the head value does not change as a result of solving the flow equation (Dirichlet boundary condition). No-flow cells are those for which no flow into or out of the cell is permitted (Neumann boundary conditions).

Since each capability of the MODFLOW code is widely documented, in the following subsections only details useful for the aims of this thesis will be provided. Further discussion can be found in Appendix 9.3, while for a full description the reader should refer to Harbaugh [69] and references therein.

4.1.1 Program structure

MODFLOW has many capabilities for solving hydrologic problems and to facilitate developing an easily manageable program, different processes are subdivided into *packages*, each of which can perform particular functions. The MAIN Program controls the order in which these functions are executed through subroutines.

First of all, a *Basic Package* addresses a variety of tasks in support of the entire simulated process (e.g.,

declaring and allocating memory for variables and for calculating an overall water budget). Moreover, some *Solver Packages* implement different algorithms for solution of the systems of proper discrete equations. Additional packages are specific for the process to be simulated. For the aims of this discussion, we will deal with packages required by the GWF Process, *Hydrologic Packages*, which calculate the coefficients of the discrete equation 4.3 within each variable-head cell. There are two types of Hydrologic Packages: *Internal Flow Packages* simulate flow between adjacent cells, while *Stress Packages* simulate several kinds of stress terms (e.g., rivers, wells, recharge).

Internal Flow Packages

Internal Flow Packages calculate conductance coefficients and the groundwater storage terms for the solution of the discrete equation 4.3.
Horizontal interblock conductances ($C_{i,j\pm1/2,k}$ and $C_{i\pm1/2,j,k}$) in equation 4.3 are calculated as the harmonic mean of hydraulic conductances of adjacent cells (see equation 9.6 in Appendix 9.3). In turn, hydraulic conductance of a single cell depends upon its hydraulic conductivity and geometry (see equation 9.3 in Appendix 9.3). Interblock conductance between cells (i, j, k) and $(i, j + 1, k)$ is so expressed as:

$$C_{i,j+1/2,k} = \frac{\frac{T_{i,j,k}\Delta y}{1/2\Delta x}\frac{T_{i,j+1,k}\Delta y}{1/2\Delta x}}{\frac{T_{i,j,k}\Delta y}{1/2\Delta x} + \frac{T_{i,j+1,k}\Delta y}{1/2\Delta x}},$$

where $T_{i,j,k}$ is the *hydraulic transmissivity* at cell (i, j, k) along the x direction, while Δy and Δx are the grid widths of rows and columns, respectively. The hydraulic transmissivity of an aquifer is defined as the integral of hydraulic conductivity over its saturated thickness; calculation of $T_{i,j,k}$ in MODFLOW is explicited hereinafter. Moreover, notice that, for the sake of simplicity, here we are assuming that rows and columns have uniform widths all over the domain and that rows, columns and layers are aligned along the three Cartesian directions (see section 3.1), but of course rows and columns can assume different widths in different parts of the grid (e.g., if a finer detail is needed) and rows, columns and layers do not need to be necessarily aligned along the three Cartesian directions. For this reason, in Harbaugh [69] a general notation is found, with Δy replaced by Δr (the row width), Δx replaced by Δc (the column width) and Δz replaced by Δv (the cell width along the vertical direction) and subscripts (i, j, k) are used to identify widths of specified rows, columns and layers (e.g., Δc_i denotes the grid width of column i). Similar expressions can be obtained for $C_{i,j-1/2,k}$ and $C_{i\pm1/2,j,k}$.
In order to apply the harmonic-mean computation, transmissivity is computed as:

$$T_{i,j,k} = K_{i,j,k}\vartheta_{i,j,k},$$

where $K_{i,j,k}$ is the hydraulic conductivity of cell (i, j, k), which must be provided as input data, and $\vartheta_{i,j,k}$ is the thickness of cell (i, j, k) itself saturated with freshwater (*saturated thickness*).
Notice that in all the preceeding equations, hydraulic conductivity is a tensor when dealing with anisotropic media, so that different components must be used along row and column directions.

For confined cells, $\vartheta_{i,j,k}$ is simply the cell thickness (vertical cell width): $\vartheta_{i,j,k} = Top_{i,j,k} - Bot_{i,j,k}$, where $Top_{i,j,k}$ and $Bot_{i,j,k}$ are the elevations of top and bottom surfaces of cell (i, j, k), respectively.
For water-table conditions, $\vartheta_{i,j,k}$ depends on head. Generally speaking, since a cell can convert between confined and water-table within a simulation, the following approach for computing saturated thickness can be applied:

$$\vartheta_{i,j,k} = \begin{cases} Top_{i,j,k} - Bot_{i,j,k} & \text{if } h_{i,j,k} \geq Top_{i,j,k} \\ h_{i,j,k} - Bot_{i,j,k} & \text{if } Bot_{i,j,k} < h_{i,j,k} < Top_{i,j,k}, \\ 0 & \text{if } h_{i,j,k} \leq Bot_{i,j,k} \end{cases} \tag{4.4}$$

When water-table conditions are simulated, cell saturated thickness is recalculated at the start of each iteration. Hydraulic transmissivity is then calculated as the product of hydraulic conductivity and saturated thickness itself and, finally, hydraulic conductance is re-calculated. In MODFLOW there are three more methods to calculate horizontal interblock conductances if hydraulic transmissivity and conductivity are not uniform within a cell, as assumed, instead, in the harmonic-mean method.
Regarding vertical interblock conductances, the following expression derives from equations 9.1 and 9.5 in Appendix 9.3:

$$\frac{1}{C_{i,j,k\pm1/2}} = \frac{1}{\frac{\Delta x \Delta y K_{i,j,k}}{1/2\vartheta_{i,j,k}}} + \frac{1}{\frac{\Delta x \Delta y K_{i,j,k\pm1}}{1/2\vartheta_{i,j,k\pm1}}},$$

where $K_{i,j,k}$ is supposed to be the vertical hydraulic conductivity of cell (i, j, k) (of course in isotropic conditions it will assume the same value as horizontal hydraulic conductivity), while $\vartheta_{i,j,k}$ is the saturated thickness of cell (i, j, k).
Internal Flow Packages handle also situations where the head in a cell is below its top elevation while the cell directly above is fully or partially saturated. In this case, the term calculating the vertical flow between the involved cells in equation 4.3 is modified so that the flow is downward and the unconfined cell below is fully saturated. A correction term is then introduced to guarantee that the matrix of coefficients is symmetric. Details can be found in Appendix 9.3.
The last capability of the Internal Flow Packages deals with re-wetting of a dry cell (i.e., a cell where head falls below its bottom elevation, so that its saturated thickness is zero, according to the last case of equation 4.4). In this case, the dry cell can be re-wetted by any of the four cells adjacent in the horizontal plane, or by the cell below. Anyway, the head of the dry cell is set in the current iteration to an elevation which equals its bottom plus a quantity which is specified by the user or depends on the hydraulic head in the neighbouring cell with respect the bottom itself (see Appendix 9.3 for details).

Stress Packages

In the GWF Process, hydrologic stresses to a groundwater system representing inflows or outflows are simulated by specific Stress Packages. Mathematically, these are boundary conditions.
The six packages documented [69] are: Well Package (WEL), Recharge Package (RCH), General-Head Boundary Package (GHB), River Package (RIV), Drain Package (DRN) and Evapotranspiration Package (EVT). WEL and RCH Packages define constant stress terms, which are subtracted from RHS in

equation 4.3, while the other Packages define stress terms that are coefficients of head ($h_{i,j,k}$), so that they are added to $HCOF$ in equation 4.3. For the purposes of this thesis, we will provide mathematical statements only about WEL, RCH and DRN Packages. Details about the other Stress Packages can be found in Harbaugh [69].

The WEL Package is designed to simulate features, such as wells, that withdraw water from or add water to the aquifer at a constant rate (i.e., independent of both the cell area and the head in the cell) during a stress-period. The flow rate (Q) for a well is specified by the user within an input file (`filename.wel`), as a fluid volume per unit time at which water is extracted from or added to the aquifer. Negative values of Q are used to indicate well discharge (pumping), whereas positive values of Q indicate a recharging well. Additional data required for each well are row, column and layer labels of the cell in which the well is located. The wells are re-defined each stress-period and, within a stress-period, at each iteration the Q value for each well is subtracted from the RHS value (see equation 4.3) for the cell containing that well. Where more than one well falls within a single cell, the calculation is repeated for each well.

For wells that are open to more than one layer of the model, the total flow rate must be divided in some way among the individual layers, externally to the model program. A common method of doing this is to divide Q in proportion to the layer transmissivities.

The RCH Package is designed to simulate areally distributed recharge to the groundwater system. Most commonly, areal recharge occurs as a result of precipitation that percolates to the groundwater system, but the RCH Package can be used also to simulate recharge from sources other than precipitation (e.g., artificial recharge or discharge). Recharge flow rate applied to the model at horizontal cell location (i, j) is defined as:

$$Q_{i,j}^{RCH} = P_{i,j}\Delta x\Delta y,$$

where $P_{i,j}$ is the recharge flux $[LT^{-1}]$ applicable to the map area ($\Delta x\Delta y$) of the cell (usually $P_{i,j}$ is expressed by a fraction of precipitation). $N_x \times N_y$ values of recharge flux ($P_{i,j}$) are specified by the user within an input file (`filename.rch`) at each stress-period. Than these values are multiplied by horizontal cell areas inside the code and the flow rate ($Q_{i,j}^{RCH}$) is subtracted from the RHS value (see equation 4.3) at each iteration for all cells that receive recharge.

The recharge ($Q_{i,j}^{RCH}$) is applied to a single cell within the vertical column of cells located at (i, j) location. The RCH Package does not allow for recharge to occur simultaneously at multiple depths in the same vertical column, because natural recharge enters the groundwater system at the top. In the simplest situation, the top of the groundwater system will occur in model layer 1. However, two further options are implemented in MODFLOW: application of the recharge to any cell in the vertical column as specified by the user in the input file, or application of the recharge to the uppermost variable-head cell in the vertical column, provided no constant-head cell is above the variable-head cell itself in that column. Anyway, if a cell designated to receive recharge is no-flow or constant-head, then no recharge is added.

The DRN Package is designed to simulate the effects of features, such as drains, which remove water from the aquifer at a rate proportional to the difference between the head in the aquifer and some fixed

stage, as long as the head in the aquifer is above that stage. If, however, the aquifer head falls below the drain stage, then the drain has no effect on the aquifer. The constant of proportionality is called the *drain conductance*. A mathematical statement of this situation is:

$$Q^{DRN} = \begin{cases} C^{DRN}(H^{DRN} - h_{i,j,k}) & \text{if } h_{i,j,k} > H^{DRN}, \\ 0 & \text{if } h_{i,j,k} \leq H^{DRN} \end{cases} \tag{4.5}$$

where Q^{DRN} is the flow rate from the drain into the aquifer $[L^3T^{-1}]$, C^{DRN} is the drain conductance $[L^2T^{-1}]$, H^{DRN} is the stage of the drain and $h_{i,j,k}$ is the hydraulic head within the cell (i, j, k) containing the drain. Thus, drain flow rate is either negative or zero.
Drain conductance and stage, togheter with row, column and layer labels of the cell where the drain is located are specified by the user within an input file (`filename.drn`). Then flow rates are calculated as in equation 4.5 and the term $(-C^{DRN})$ is added to $HCOF$, while $(C^{DRN}HDRN)$ is subtracted from RHS (see equation 4.3). Figure 4.2 shows a graph of flow rate from a drain depending on hydraulic head in the cell containing the drain itself, as defined by equation 4.5:

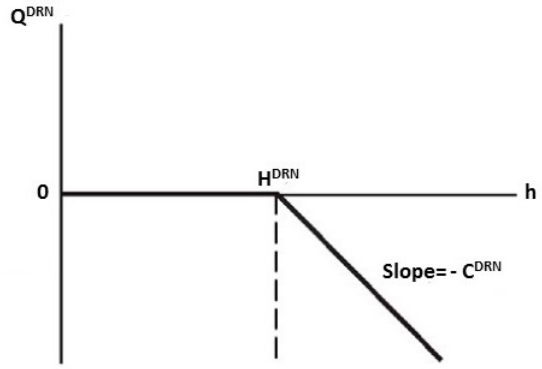


Figure 4.2: Plot of flow rate into a drain as a function of hydraulic head in the cell containing the drain. H^{DRN} and C^{DRN} are the elevation and the conductance of the drain, respectively. Modified after Harbaugh [69].

Solver Packages

In MODFLOW three Solver Packages are implemented: Strongly Implicit Procedure (SIP) Package [91], Direct Solver (DE4) Package [68] and Preconditioned Conjugate-Gradient Solver (PCG) Package [71]. The best solver to be applied depends on the characteristics of the specific problem. Actually, all the solvers can solve the equations of typical simulations, but some highly non-linear problems cannot be solved by all of them. In fact, the SIP Package is more suitable to solve large systems of simultaneous linear equations, while DE4 and PCG Packages can support also non-linear models.
Assuming that any algorithm can solve the equations, the next most important objective for solver selection is minimizing the required computer time. Solution time of course varies with the power of the computer, but for a given computer, different solvers can take substantially different amounts of time.

Another consideration regarding solver selection is the amount of computer memory required. Each solver uses a different amount of memory and this can be of concern for very large simulations: insufficient memory, in fact, will prevent the solution of the equations. Moreover, minimizing the amount of memory used for the model when other programs are running concurrently can be advantageous.
All of the three solvers incorporate iterations, which can occur at two levels. The primary level includes *outer iterations*: here any head-dependent (non-linear) terms will change as a result of the head computed in the prior outer iteration; then one partial solution of the flow equations is performed. The second iteration level includes *inner iterations*: they are used to improve accuracy of the solution without changing the non-linear terms. The SIP Package supports just outer iterations, while DE4 Package uses inner or outer iteration, but not at the same time, based on the lineariry or non-linearity of the system. On the other hand, the PCG Package supports both inner and outer iterations.
The outer iterations are repeated until a user-specified maximum number of iterations is reached or a satisfactory solution is obtained (closure criterion, see subsection 3.2.1). In addition, the PCG Package requires also meeting the *residual criterion*, i.e., the maximum absolute value of the difference between the inflows and outflows (residual) for each cell and iteration must be lower than or equal to a user-specified value.
For the aims of this thesis, some details about the PCG solver are provided in Appendix 9.3.

4.2 YAGMod

Working on flow and transport in porous media, the modeler is faced with the high complexity of the hydrostratigraphic structure, which induces heterogeneity of the conductivity field, and processes which might interact to determine the groundwater flow field or the distribution of a solute in the subsurface. In order to handle these complex processes (e.g., coupling groundwater flow and solute transport), it is necessary to estimate a great number of phenomenological parameters, which are seldom measured in the field. Therefore, several assumptions have to be introduced, the optimal values of parameters can be found with inverse methods and the effects produced by their uncertainty on the model outcome can be estimated with a sensitivity analysis. Anyway, if heterogeneous media are considered, then the spatial distribution of the physical parameters has to be known and this introduces further difficulties, both from practical and numerical points of views.
Based on these remarks, despite the wide set of available models, there is still place to develop numerical codes which can be oriented to the solution of specific problems, so that handling some particular situations is easier than with commercial software packages. Within this framework, the code *YAGMod* (*Yet Another Groundwater flow Model*) was developed at the Department of Earth Sciences of the Università degli Studi di Milano, for the simulation of constant-density groundwater flow in saturated media under stationary conditions [28]. This code was developed in Fortran90 and is based on a conservative finite-difference scheme, very similar to the one applied in MODFLOW. It also includes a module for the model calibration with the Comparison Model Method (CMM) for 2D flow conditions (see subsection 4.2.5). Hereinafter, the most important differences with respect to exiting codes will be highlighted.

Although the method could be easily extended to 3D flows, practical applications have been successfully performed so far for 2D hydraulic flows in regional aquifers [39, 65].
What follows is based on Cattaneo et al. [28], unless otherwise specified.

4.2.1 Mathematical statement

In YAGMod a different convention for discretization is used with respect to MODFLOW, so a different notation will be used to avoid confusion: here m is the label for columns, n is the label for rows and p is the label for layers. Moreover, in contrast to what stated in subsection 3.2.1, increments in the column label (m) would correspond to increases in x (like in MODFLOW), but increments in the row label (n) would correspond to increases in y and increments in the p label would correspond to increases in elevation. With reference to this convention, simplifying the discrete equation 4.1 gives:

$$\begin{aligned}
&\mathcal{T}_{m+1/2,n,p}(h_{m+1,n,p} - h_{m,n,p}) + \mathcal{T}_{m-1/2,n,p}(h_{m-1,n,p} - h_{m,n,p}) + \\
&\mathcal{T}_{m,n+1/2,p}(h_{m,n+1,p} - h_{m,n,p}) + \mathcal{T}_{m,n-1/2,p}(h_{m,n-1,p} - h_{m,n,p}) + \\
&\mathcal{T}_{m,n,p+1/2}(h_{m,n,p+1} - h_{m,n,p}) + \mathcal{T}_{m,n,p-1/2}(h_{m,n,p-1} - h_{m,n,p}) = -F_{m,n,p},
\end{aligned} \tag{4.6}$$

where the subscripts $(m \pm 1/2, n, p)$ identify the groundwater flow terms between cells (m, n, p) and $(m \pm 1, n, p)$ along the x direction (similar definitions can be given along y and z directions), and $F_{m,n,p}$ is the source term at cell (m, n, p). Equation 4.6 has the same form of equation 4.3, as solved in MODFLOW, with the storage term set to zero (steady-state process), but here *interblock hydraulic transmittances* are used instead of interblock hydraulic conductances. Interblock hydraulic transmittances along x, y and z directions are defined as $[L^2T^{-1}]$:

$$\mathcal{T}_{m\pm1/2,n,p} = K_{m\pm1/2,n,p}\vartheta_{m\pm1/2,n,p}\frac{\Delta y}{\Delta x}$$

$$\mathcal{T}_{m,n\pm1/2,p} = K_{m,n\pm1/2,p}\vartheta_{m,n\pm1/2,p}\frac{\Delta x}{\Delta y}$$

$$\mathcal{T}_{m,n,p\pm1/2} = K_{m,n,p\pm1/2}\frac{\Delta x \Delta y}{\Delta z_{m,n,p\pm1/2}},$$

where $K_{m\pm1/2,n,p}$, $K_{m,n\pm1/2,p}$ and $K_{m,n,p\pm1/2}$ are called the *interblock hydraulic conductivities* along x, y and z directions, respectively, $\vartheta_{m\pm1/2,n,p}$ and $\vartheta_{m,n\pm1/2,p}$ are nothing but the arithmetic mean of saturated thicknesses of involved cells, as defined by equation 4.4, and $\Delta z_{m,n,p\pm1/2} = (\Delta z_{m,n,p} - \Delta z_{m,n,p\pm1})/2$ denotes the vertical distance between adjacent nodes. Interblock hydraulic conductivities are calculated as the harmonic mean of hydraulic conductivities of involved cells along the proper direction.
After writing equation 4.6 for each active cell of the domain, a system of simultaneous finite-difference equations (see matrix equation 3.2) is formulated and solved to yield head values at each node of the model grid, by using iterative solution methods. Since YAGMod performs stedy-state simulations only, a single set of simultaneous equations will be solved iteratively. Like in MODFLOW, the convergence is achieved when the convergence criterion (see subsection 3.2.1) is met, but also a maximum permissible

number of iterations is specified by the user to attempt convergence.
Regarding the (active, non-active or boundary) status of a certain cell, YAGMod takes as input user-defined capital characters: active cells are called as *internal cells* (I), non-active cells are called as *external cells* (E; no flow into or out of the cell is permitted and no solution of the flow equation is attempted), while boundary cells used to simulate Dirichlet boundary conditions are called as *Dirichlet cells* (D; the head values are specified and do not change as a result of solving the flow equation). Of course, the terms "internal" and "external" refer to the model domain within which groundwater flow has to be simulated.

4.2.2 Handling dry cells

The proposed model does not solve equations for variably-saturated conditions, but aims at finding a solution for fully saturated media: 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 specified here.
If $h_{m,n,p} < Bot_{m,n,p}$, then $\vartheta_{m,n,p} = 0$ (see equation 4.4). If also the adjacent cells along the horizontal directions are dry, then the terms corresponding to horizontal fluxes in equation 4.6 vanish and the cell (m, n, p) 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 uppermost active cell of the column, to reach the water table. Instead, if the adjacent cells have non vanishing thicknesses, then the physical situation implies that there is a horizontal transfer of water.
As stated in Appendix 9.1, when the convergence criteria are met, the code checks for the physical reliability of the solution reached by updating an array of alphanumeric codes, named $Sat(m, n, p)$, thanks to which every internal cell (I) of the domain is associated with a character label which describes its saturation status. First of all, $Sat(m, n, p)$ is initialized with the following codes (a for totally dry cells not connected with surface, w for partially saturated cells not connected with surface and S for fully saturated cells):

$$Sat_{(m,n,p)} = \begin{cases} a & \text{if } h_{(m,n,p)} < Bot_{(m,n,p)} \\ w & \text{if } Bot_{(m,n,p)} \leq h_{(m,n,p)} < Top_{(m,n,p)}, \\ S & \text{if } h_{(m,n,p)} \geq Top_{(m,n,p)} \end{cases}$$

Then a recursive function, starting from the top layer, checks for every continuous path connecting partially or totally desaturated cells with the surface. Firstly, all a or w cells of the top layer are transformed respectively into A (totally dry cells connected with surface) or W (partially saturated cells connected with surface). Then, a check is performed for every cell if a continuous path connects partially or totally desaturated cells with the top layer. At the end of the checking subroutine run, every totally or partially desaturated cell needs to be connected with the surface, in order to allow air to infiltrate into the porous media. Isolated cells or subsets of cells where $Sat = a$ or $Sat = w$ do not have a continuous path joining air in the cell with the surface. This is considered not acceptable from a physical point of view and a warning message is displayed.

4.2.3 Source terms

In YAGMod, Neumann and Robin boundary conditions are implemented as specific types of source terms. These can be classified according to location and dependence on the water level. Hereinafter, a list of the most common source terms which can be simulated is provided.
Areally distributed and fixed (i.e, not depend on water level within the cell) source terms (e.g., rainfall recharge) are represented by $F^{(d)}_{m,n,p}$ hydraulic flow $[L^3T^{-1}]$ within each cell.
Point source terms (i.e., located within a single cell) are defined by the following equation:

$$F^{(loc)} = \begin{cases} F_1 + K_1(h - H^{(cal)}) & \text{if } h \geq H^{(act)} \\ F_2 + K_2(h - H^{(cal)}) & \text{if } h < H^{(act)}, \end{cases} \tag{4.7}$$

where $F^{(loc)}$ is the contribution of the point source term within the cell where it is located, F_1 and F_2 are fixed fluxes inside the cell $[L^3T^{-1}]$, K_1 and K_2 are hydraulic conductances $[L^2T^{-1}]$, h is the water level within the cell, $H^{(cal)}$ (which usually equals $H^{(act)}$) is used to calculate the terms which linearly depend on h and $H^{(act)}$ is a threshold defining if the source term is active and which couple between $(F_1; K_1)$ and $(F_2; K_2)$ must be used.
Since equation 4.7 defines point source terms, besides F_1, F_2, K_1, K_2, $H^{(cal)}$ and $H^{(act)}$, the code requires the labels (m, n, p) of the cell where $F^{(loc)}$ is located.
Equation 4.7 allows defining several kinds of point source terms listed below.

- Well with fixed extracted flow rate. In this case, $F_1 = F_2 < 0$ quantify the extracted flow rate, while $K_1 = K_2 = 0$, so that $F^{(loc)}$ does not depend on h and $H^{(cal)}$ and $H^{(act)}$ values are not relevant.

- Drain. In this case, $F_1 = F_2 = 0$, $K_2 = 0$, K_1 represents the drain conductance and must be negative to obtain outflow from the cell and $H^{(cal)} = H^{(act)}$ represent the drain elevation (they could be different from each other to account for situations in which the threshold on the aquifer water head to start the draining process is different from the threshold which controls the water flux).

- Robin boundary conditions. These conditions can be used if the aquifer interacts with another water body and water exchange is controlled by the difference between water head in the aquifer and water level in the external water body. In this case, $F_1 = F_2 = 0$, K_1 and K_2 represent the conductances for flow out or into the cell (they must be negative and depend upon the conductivity of the materials that separate the aquifer from the water body at the reference water head and upon the distance from this water body), and $H^{(cal)} = H^{(act)}$ are the reference hydraulic heads of the involved water bodies.

- River/aquifer interaction. In this case, $H^{(act)}$ is the height of the river bottom, so that the aquifer is in contact with the river if $h \geq H^{(act)}$, otherwise (i.e., if $h < H^{(act)}$) they are separated by a vadose zone. In the first case, $H^{(cal)}$ is the water level into the river, so that the river is draining

the aquifer if $h > H^{(cal)}$ and it is recharging the aquifer if $h < H^{(cal)}$; for this situation it is quite common to assume $F_1 = 0$ and K_1 dependent on the conductivity of the river bed sediments, their thickness and the area of the contact surface between the river bed and the aquifer into the considered cell. In the second case, the river bed sediments are assumed to be fine-grained, almost saturated and impermeable, while the vadose zone is more permeable and dry; for this situation, $K_2 = 0$ and F_2 depends upon the conductivity, thickness and extension of the river bed sediments in the considered cell.

YAGMod further considers screened wells, namely wells whose extracted flow rate depends on the hydraulic head on the aquifer, so that no outflow occurs if the cell becomes dry. The input data required to define this kind of wells are: coodinates (x, y) of the well, labels (m_W, n_W) of the cell where the screened well is located, top and bottom elevations of the screened interval (top_W and bot_W, respectively) and the maximum well extraction rate (q_W).
The maximum extraction rate is subdivided among cells occupied by screened intervals, as:

$$q^{(scr)}_{m_W,n_W,p} = q_W \frac{K_{m_W,n_W,p}\Delta\mathcal{L}^{(scr)}_{m_W,n_W,p}}{\sum_{p'=s_{min}}^{s_{max}} K_{m_W,n_W,p'}\Delta\mathcal{L}^{(scr)}_{m_W,n_W,p'}},$$

where s_{min} and s_{max} are the indices (along the vertical direction) of the cells where top and bottom of the screened interval of the well are located, respectively, while $\mathcal{L}^{(scr)}_{m_W,n_W,p}$ is the screened thickness of the well corresponding to a fully saturated porous medium within the cell (m_W, n_W, p) and is computed as:

$$\mathcal{L}^{(scr)}_{m_W,n_W,p} = \min(h_{m_W,n_W,p}, s^{(top)}_{m_W,n_W,p}) - \min(h_{m_W,n_W,p}, s^{(bot)}_{m_W,n_W,p}),$$

where

$$s^{(top)}_{m_W,n_W,p} = \begin{cases} top_{m_W,n_W,p} & p = s_{min}, ..., s_{max} - 1 \\ top_W & p = s_{max}, \end{cases}$$

and

$$s^{(bot)}_{m_W,n_W,p} = \begin{cases} bot_{m_W,n_W,p} & p = s_{min} + 1, ..., s_{max} \\ bot_W & p = s_{min} \end{cases}$$

Notice that if $h(m_W, n_W, p) < s^{(bot)}_{m_W,n_W,p}$, then $\mathcal{L}^{(scr)}_{m_W,n_W,p} = 0$.
If the cell (m_W, n_W, p) desaturates, the contribution of the screened well to the source term within the cell (m_W, n_W, p) is given by:

$$F^{(scr)}_{m_W,n_W,p} = q^{(scr)}_{m_W,n_W,p} \frac{\sqrt{\mathcal{L}^{(scr)}_{m_W,n_W,p}}}{s^{(top)}_{m_W,n_W,p} - s^{(bot)}_{m_W,n_W,p}}. \tag{4.8}$$

Equation 4.8 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.

4.2.4 Solution with the relaxation method

Equation 4.6 can be written in matrix form (see equation 3.2) as:

$$\mathbf{A}(\mathbf{h})\mathbf{h} = \mathbf{b}^{(fix)} + \mathbf{b}^{(var)}(\mathbf{h}), \tag{4.9}$$

where the matrix $\mathbf{A}$ is built with transmittances and therefore depends on $\mathbf{h}$, $\mathbf{b}^{(fix)}$ includes distributed and point source terms, which are constant with respect to the water level, and $\mathbf{b}^{(var)}$ includes distributed and point source terms, which depend on the hydraulic head within the cell where they are located.
Equation 4.9 is solved by using a generalization of the *relaxation method* for linear algebraic equations: the *classic Gauss-Seidel method*. This method might be not the optimal choice from the computational point of view, but it proved to be quite robust, in particular for complex situations.
The idea consists in decomposing the coefficient matrix as $\mathbf{A} = \mathbf{D} - \mathbf{L} - \mathbf{U}$, where $\mathbf{D}$ is a diagonal matrix, $\mathbf{L}$ is a lower-triangular matrix and $\mathbf{U}$ an upper-triangular matrix. So, equation 4.9 can be written as:

$$\mathbf{D}(\mathbf{h})\mathbf{h} - \mathbf{L}(\mathbf{h})\mathbf{h} = \mathbf{U}(\mathbf{h})\mathbf{h} + \mathbf{b}^{(fix)} + \mathbf{b}^{(var)}(\mathbf{h}).$$

Starting from an initial guess ($\mathbf{h}^0$), then successive approximations of the solution to the equation 4.9 ($\mathbf{h}^r$ $(r = 1, 2, ...)$) are computed. The generalization of the classic Gauss-Seidel method to non-linear problems allows calculating:

$$\mathbf{h}^* = [\mathbf{D}(\mathbf{h}^{r-1}) - \mathbf{L}(\mathbf{h}^{r-1})]^{-1} \cdot [\mathbf{U}(\mathbf{h}^{r-1})\mathbf{h}^{r-1} + \mathbf{b}^{(fix)} + \mathbf{b}^{(var)}(\mathbf{h}^{r-1})]$$

and

$$\mathbf{h}^r = \mathbf{h}^{r-1} + \omega(\mathbf{h}^* - \mathbf{h}^{r-1}) = (1 - \omega)\mathbf{h}^{r-1} + \omega\mathbf{h}^*,$$

where $\omega \in (0, 2)$ is a weight called the *relaxation parameter*.
This approach is usually referred to as *over-relaxation method* when $1 < \omega < 2$, namely if the correction given by the Gauss-Seidel method approach is increased in the attempt to accelerate convergence to the solution, and *under-relaxation method* when $0 < \omega < 1$. This latter case is more robust for systems of algebraic, non-linear equations.
The iterative procedure is stopped when the convergence criterion is met, namely, $\| \mathbf{h}^r - \mathbf{h}^{r-1} \|_\infty < \varepsilon$, where ε is a small, user-defined number.
Bad choices of the relaxation parameter (ω) and the initial guess ($\mathbf{h}^0$) might prevent convergence to a solution or, for highly non-linear problems, lead to different solutions.

4.2.5 Inverse model with the CMM

YAGMod includes also a subroutine for calibration through the *Comparison Model Method* (*CMM*). It is a direct inversion method, which allows getting the hydraulic conductivity field for 2D flows under steady-state conditions. At first, this method was proposed by Scarascia and Ponzini in 1972 [113], then further developed by Ponzini and Lozej in 1982 [102], finally formalized by Ponzini and Crosta in 1988

[103] and successfully applied [39, 65, 136, 137].
The basic idea consists in solving the forward problem for a "Comparison" Model (CM), namely with a starting conductivity field, usually assumed as uniform. Then, the "real" conductivity field can be get by supposing that flows calculated for the CM are a good approximation of the "real" ones, calculated from the unknown conductivity field and the reference hydraulic gradient.
First of all, the CMM method requires knowing the reference hydraulic head ($h_{m,n}^{(ref)}$), obtained by interpolation of field data and further filtering, in order to limit some instability or ill-conditioning effects of the inverse problem, related to low wavelength irregularities of the $h^{(ref)}$ field [136]. Also the reference source term within each cell of the domain ($F_{m,n}^{(ref)}$) need to be estimated after proper filtering process, in order to have consistent data. Notice that, since the CMM is defined for 2D flows, here the p label equals 1 and is simply omitted. The generalization and application of the CMM method to 3D flows would require knowing the reference hydraulic head inside each cell of the model domain.
The Comparison Model has the same geometry, source terms and boundary conditions as the model to be calibrated. So, with reference to the cell (m, n), the balance equation for the CM has the same form as equation 4.6, where $\mathcal{T}_{m,n} = \mathcal{T}_{m,n}^{(CM)}$ is determined through an initial estimate of the conductivity field ($K_{m,n}^{(CM)}$ obtained, e.g., by field data), $h_{m,n} = h_{m,n}^{(CM)}$ is the unknown variable to be calculated, and $F_{m,n} = F_{m,n}^{(CM)} = F_{m,n}^{(ref)}$ is the reference source term. Of course, if (m, n) is a Dirichlet cell we have $h_{m,n} = h_{m,n}^{(CM)} = h_{m,n}^{(ref)}$.
Once calculated $h_{m,n}^{(CM)}$ from the resolution of this equation, the volumetric flow rate within the cell per unit horizontal length is given by:

$$\mathbf{Q}_{m,n}^{(CM)} = -K_{m,n}^{(CM)} \vartheta_{m,n}^{(CM)} \nabla h_{m,n}^{(CM)} = -T_{m,n}^{(CM)} \nabla h_{m,n}^{(CM)}, \tag{4.10}$$

where $\vartheta_{m,n}^{(CM)}$ is calculated by using equation 4.4 with $h_{m,n,p} = h_{m,n}^{(CM)}$ and the concept of hydraulic transmissivity was used. In the discretization of equation 4.10, $\nabla h_{m,n}^{(CM)}$ is defined by using the block-centered finite-difference method, which is unuseful within Dirichlet cells, where the arithmetic mean of the K values in neighboring internal cells is so assigned.
Now suppose that reference hydraulic head is a good approximation of the real hydraulic head within the cell, so that the real volumetric flow rate within the cell per unit horizontal length is given by:

$$\mathbf{Q}_{m,n} = -K_{m,n} \vartheta_{m,n}^{(ref)} \nabla h_{m,n}^{(ref)} = -T_{m,n} \nabla h_{m,n}^{(ref)}, \tag{4.11}$$

where $\vartheta_{m,n}^{(ref)}$ is calculated by using equation 4.4 with $h_{m,n,p} = h_{m,n}^{(ref)}$ and, again, the concept of hydraulic transmissivity was used. Here $T_{m,n}$ (or $K_{m,n}$) is the hydraulic transmissivity (or conductivity) field to be estimated.
If $K_{m,n}^{(CM)}$ is a good approximation of the real conductivity field ($K_{m,n}$), then we can assume that $\mathbf{Q}_{m,n}^{(CM)} \simeq \mathbf{Q}_{m,n}$. From this comparison, considering absolute values of equations 4.10 and 4.11, we can obtain the following estimate of the real conductivity field within the cell (m, n):

$$K_{m,n} = K_{m,n}^{(est)} = K_{m,n}^{(CM)} \frac{\vartheta_{m,n}^{(CM)}}{\vartheta_{m,n}^{(ref)}} \frac{|\nabla h_{m,n}^{(CM)}|}{\nabla h_{m,n}^{(ref)}}, \tag{4.12}$$

or, in terms of the hydraulic transmissivity:

$$T_{m,n} = T_{m,n}^{(est)} = T_{m,n}^{(CM)} \frac{|\nabla h_{m,n}^{(CM)}|}{\nabla h_{m,n}^{(ref)}}.$$

The procedure just introduced defines the *integral approach* of the CMM.
With the alternative *differential approach*, instead, the volumetric flow rate $Q_{m,n} = K_{m,n}\vartheta_{m,n}|\nabla h_{m,n}|$ is a function of $h_{m,n}$, both explicitly (through $|\nabla h_{m,n}|$) and implicitly (through $\vartheta_{m,n}$). So small variations of $K_{m,n}$ and $h_{m,n}$ produce small variations of $Q_{m,n}$:

$$dQ_{m,n} = dK_{m,n}\vartheta_{m,n}|\nabla h_{m,n}| + K_{m,n}d\vartheta_{m,n}|\nabla h_{m,n}| + K_{m,n}\vartheta_{m,n}d|\nabla h_{m,n}|. \tag{4.13}$$

If

$$dK_{m,n} = K_{m,n}^{(est)} - K_{m,n}^{(CM)}, \qquad d\vartheta_{m,n} = \vartheta_{m,n}^{(ref)} - \vartheta_{m,n}^{(CM)}, \qquad d|\nabla h_{m,n}| = |\nabla h_{m,n}^{(ref)}| - |\nabla h_{m,n}^{(CM)}|$$

and the relation $\mathbf{Q}_{m,n}^{(CM)} \simeq \mathbf{Q}_{m,n}$ is still valid, then $dQ_{m,n} = 0$ and

$$K_{m,n}^{(est)} = K_{m,n}^{(CM)} \left(1 - \frac{\vartheta_{m,n}^{(ref)} - \vartheta_{m,n}^{(CM)}}{\vartheta_{m,n}^{(CM)}} - \frac{|\nabla h_{m,n}^{(ref)}| - |\nabla h_{m,n}^{(CM)}|}{|\nabla h_{m,n}^{(CM)}|}\right),$$

or, in terms of the hydraulic transmissivity:

$$T_{m,n}^{(est)} = T_{m,n}^{(CM)}(1 - \frac{|\nabla h_{m,n}^{(ref)}| - |\nabla h_{m,n}^{(CM)}|}{|\nabla h_{m,n}^{(CM)}|}).$$

In a confined aquifer, cells are always saturated, so that $d\vartheta_{m,n} = 0$. In this case, if $dQ_{m,n} = 0$, from equation 4.13 we get:

$$\frac{dK_{m,n}}{K_{m,n}} = -\frac{d|\nabla h_{m,n}|}{|\nabla h_{m,n}|}.$$

By integrating the left-hand side between $K_{m,n}^{(CM)}$ and $K_{m,n}^{(est)}$ and the right-hand side between $|\nabla h_{m,n}^{(CM)}|$ and $|\nabla h_{m,n}^{(ref)}|$, we finally get:

$$\log \frac{K_{m,n}^{(est)}}{K_{m,n}^{(CM)}} = -\log \frac{|\nabla h_{m,n}^{(ref)}|}{|\nabla h_{m,n}^{(CM)}|},$$

and so equation 4.12. In this way, the relationship between the integral and differential approaches is demostrated.
The inverse problem aims at finding the conductivity field by an iterative procedure, which yields the following expressions for the estimated K field:

$$K_{m,n}^{q+1} = K_{m,n}^{q}\left[w\left(\frac{\vartheta_{m,n}^{q}}{\vartheta_{m,n}^{(ref)}}\frac{|\nabla h_{m,n}^{q}|}{|\nabla h_{m,n}^{(ref)}|} - 1\right) + 1\right],$$

for the integral approach, and

$$K_{m,n}^{q+1} = K_{m,n}^{q}\left[1 - w\left(\frac{\vartheta_{m,n}^{(ref)} - \vartheta_{m,n}^{q}}{\vartheta_{m,n}^{q}} + \frac{|\nabla h_{m,n}^{(ref)}| - |\nabla h_{m,n}^{q}|}{|\nabla h_{m,n}^{q}|} - 1\right) + 1\right],$$

for the differential approach. In these latter equations, superscripts refer to iteration. The weight $w = \min\left(c|\nabla h_{m,n}^{(ref)}|, 1\right)$, where c is a positive constant, is useful to avoid ill-conditioning, due to excessive growth of some values where the reference hydraulic gradient is small. Similar equations can be written for hydraulic transmissivity.
The initial guess $K^{(CM)}$ (i.e, K^0 in the iterative process) can be estimated by using alternative methods:

- a constant value all over the whole domain;
- a heterogeneous field read from an external file;
- a constant value so that ∇h is as similar as possible to $\nabla h^{(ref)}$ [136].

4.3 SEAWAT

Among dispersive solute-transport models (see subsection 2.4.1), some programs (e.g., SUTRA [139] and FEFLOW [126]) solve the flow and transport equations simultaneously, while others (e.g., SEAWAT) solve the transport equation separately using particle-based or finite-difference methods and compute a new flow field in order to represent a changing density field. In this section, some basic information on the SEAWAT program will be provided. For further details refer to Guo & Langevin [66].

4.3.1 Equivalent freshwater head

For the content of this subsection refer to Bakker et al. [11] and Guo & Langevin [66].

The *SEAWAT* program was developed to simulate 3D, variable-density groundwater flow in porous media. The source code for SEAWAT was developed by combining MODFLOW and MT3DMS [148] into a single program that solves the coupled flow and solute-transport equations (see equations 2.22 and 2.29, respectively).
SEAWAT is based on the concept of the (*equivalent*) *freshwater head* in a saline groundwater environment, as specified hereinafter.
Consider two piezometers open to a given point (N) located at the height Z_N with respect to a reference horizontal plane, as shown in Figure 4.3. Piezometer A contains freshwater with density ρ_f, while piezometer B contains saline water with density ρ_s. The height of the water level in piezometer A above point N is $P_N/\rho_f g$, where P_N is the hydrostatic pressure at point N and g is the gravity acceleration. So, the (equivalent) freshwater head in piezometer A above the datum is given by:

$$h_f = \frac{P_N}{\rho_f g} + Z_N. \tag{4.14}$$

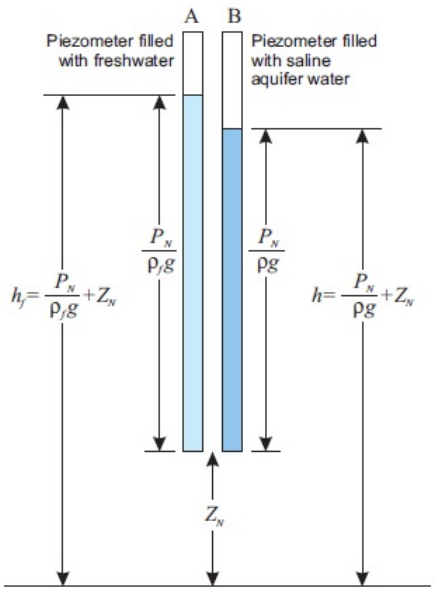


Figure 4.3: Two piezometers, filled with freshwater (A) and saline aquifer water (B), open to the same point above the reference datum. Modified after Guo & Langevin [66].

Similarly, the saline water head in piezometer B above the datum is given by:

$$h_s = \frac{P_N}{\rho_s g} + Z_N. \tag{4.15}$$

Generally speaking, ρ_s is the water density within the aquifer, which is not necessarily the saltwater density. For this reason, hereinafter ρ_s is simply replaced with ρ and, similarly, h_s is simply replaced with h.

The balance equation for groundwater flow can be formulated in terms of h_s, but the result includes expressions involving density and its derivatives and no computational advantage is gained. On the other hand, formulation of the flow equation in terms of freshwater head causes no increase in complexity and allows the use of software, such as MODFLOW, with relatively little modification. However, the calculated value of freshwater head at a given point in the aquifer does not represent the level to which native groundwater would rise in a piezometer open at that point. Conversion between head as measured for the native aquifer water and equivalent freshwater head is, therefore, necessary. In this regard, by eliminating pressure between equations 4.14 and 4.15 and solving for the respective head value, we get:

$$h_f = \frac{\rho}{\rho_f} h - \frac{\rho - \rho_f}{\rho_f} Z = \frac{\rho}{\rho_f} h - \nu Z,$$

$$h = \frac{\rho_f}{\rho} h_f + \frac{\rho - \rho_f}{\rho_f} Z = \frac{\rho_f}{\rho} h_f + \nu Z,$$

where ν is the inverse of parameter δ appearing in equation 2.30 and Z is the elevation of the bottom of the piezometer above the reference datum.

Darcy's law for variable-density flow can be written in terms of the freshwater head as:

$$q_x = -K_f \frac{\partial h_f}{\partial x}, \tag{4.16}$$

$$q_y = -K_f \frac{\partial h_f}{\partial y}, \tag{4.17}$$

$$q_z = -K_f \left(\frac{\partial h_f}{\partial z} + \nu \right), \tag{4.18}$$

where K_f is the freshwater hydraulic conductivity along the three Cartesian directions, obtained by using density and viscosity of freshwater in equation 2.4.
Furthermore, the second term on the left-hand side of equation 2.22 represents the change of fluid mass in the REV due to the change in solute concentration. To evaluate this term, the relation between solute concentration and fluid density is required. For isothermal conditions, fluid density is predominantly affected by the solute concentration and fluid pore pressure. The effects of pore pressure on fluid density are included in the storage term (S_p), while the effect of solute concentration can be expressed through the following empirical relation: $\rho = \rho_f + E\mathcal{C}$, where $E \approx 0.7143$ is a dimensionless constant. By using this relation, equation 2.22 can finally be written in terms of the equivalent freshwater head as:

$$\rho S_f \frac{\partial h_f}{\partial t} + nE \frac{\partial \mathcal{C}}{\partial t} = -\mathrm{div}(\rho \mathbf{q}) + \rho_s q_s, \tag{4.19}$$

where equation 4.14 was used to express $\partial p / \partial t$, with $S_f = S_p \rho_f g$, and the components of $\mathbf{q}$ are given by equations 4.16, 4.17 and 4.18. Equation 4.19 is the governing equation for variable-density flow in terms of freshwater head as used in SEAWAT but, with respect to equation 42 in Guo & Langevin [66], here we assumed that the principal directions of permeability coincide with the Cartesian coordinate system. The first term on the left-hand side represents the time rate of change of liquid mass (which includes both water and solute) per unit volume of the aquifer, due to pressure changes in the system. The second term on the left-hand side represents the time rate of change of fluid mass per unit volume of aquifer due to the change in solute concentration. This term is calculated from the concentrations obtained from the solution of the solute-transport equation 2.29 and, as the concentrations reach dynamic equilibrium, this term becomes negligible. Thus, the flow field does not reach steady state until the solute concentrations remain constant in time. The first term on the right-hand side of equation 4.19 accounts for the difference between inflow and outflow of mass per unit volume of aquifer across the faces of a differential element of aquifer (for example, a model cell). Finally, the second term on the right-hand side represents the mass flux due to sources and sinks.
Equation 4.19 is approximated and solved in SEAWAT with a central finite-difference scheme in space and a backward finite-difference scheme in time, so obtaining a relation of the same form as equation 4.3 solved in MODFLOW: the difference is that each term of the RHS accumulator is multiplied by fluid density. Thus, the RHS accumulator in SEAWAT has dimensions $[MT^{-1}]$ rather than $[L^3T^{-1}]$: the flow equation 4.3 is so reformulated to conserve mass rather than fluid volume.
After writing the approximated form of equation 4.19 for each cell of the domain, the matrix equation 3.2 can be obtained, with the equivalent freshwater head instead of the "actual" hydraulic head within the cell.

4.3.2 Approximation of the solute-transport equation

MT3DMS code [148] includes different techniques to approximate the solute-transport equation 2.29. First of all, the standard finite-difference method offers the advantage and convenience of a fixed grid and handles dispersion/reaction-dominated problems effectively. Some mixed Eulerian-Lagrangian approaches are implemented to combine the advantages of both the Eulerian and the Lagrangian approaches by solving the advection term with a Lagrangian method (particle tracking) and the dispersion and reaction terms with an Eulerian method (e.g., finite-difference or finite-elements). Finally, an Eulerian TVD (Total-Variation-Diminishing) method is also available, with the condition that the sum of concentration differences between adjacent nodes diminishes over successive transport-steps.

4.3.3 Program structure

The SEAWAT code combines MODFLOW and MT3DMS [148] into a single program that solves the coupled flow and solute-transport equations (4.19 and 2.29). The coupling between flow and transport is performed through a synchronous time-stepping approach that cycles between MODFLOW solution of the flow equation and MT3DMS solution of the transport equation. SEAWAT includes both explicit and implicit methods for coupling the flow and solute-transport equations. With the *explicit method*, fluid densities are calculated with solute concentrations from the previous time-step; advective fluxes from the flow solution for the current time-step are then used in the current solution to the transport equation. With the *implicit method*, solutions to the flow and transport equations are repeated and concentrations and densities are updated within each time-step, until the maximum difference in fluid density at a single cell for consecutive iterations is less than a user-specified value.
The temporal discretization of the simulation is implemented with a combination of the schemes used by MODFLOW and MT3DMS. As in MODFLOW, in SEAWAT stress-periods are subdivided into time-steps and, as in MT3DMS, MODFLOW time-steps are further subdivided into transport-steps, whose maximum length is calculated according to specific stability and accuracy requirements based on flow velocities.
With the explicit coupling approach, some drawbacks could arise: instability problems during solution of the flow equation, as densities are calculated using the previous time-step concentrations; the lengths of transport-steps are based on flow velocities calculated for the end of the preceeding time-step and, thus, there is a lag of one time-step in the application of the stability constraints and accuracy requirements. On the other hand, with the implicit coupling approach, the user may specify the lengths of the transport-steps.

4.4 UCODE

For the content of this section refer to Hill and Tiedeman [72] and Poeter et al. [100].

UCODE can be combined with existing models to perform sensitivity analysis, data needs assessment,

calibration, prediction and uncertainty analysis. Any process model or set of models can be used; the only requirements are that models have numerical (ASCII or text only) input and output files, that the values in these files are stored with sufficient significant digits, that all required models can be run from a single batch file or script and that simulated values are continuous functions of the parameter values. Process models can include pre-processors and post-processors as well as multiple models related to the processes of interest (e.g., physical, chemical, etc.), making UCODE extremely powerful.
An other feature that makes UCODE a universal inverse code with broad applicability is the inclusion of informative statistics to evaluate the importance of observations to parameters and the importance of parameters to predictions.
UCODE can be used fruitfully in indirect model calibration through its sensitivity analysis capabilities and its ability to estimate parameter values that result in the best possible fit to the observations. UCODE can also be used to calculate non-linear confidence and predictions intervals, which quantify the uncertainty of model simulated values when the model is non-linear.
Specifically, UCODE was developed to perform the following operations.

- Manipulate process-model input files and read values from process-model output files.
- Compare user-provided observations with equivalent simulated values derived from the process-model output files using a number of summary statistics, including a weighted least-squares objective function.
- Use optimization methods to adjust the values of user-selected input parameters in an iterative procedure to minimize the value of the weighted least-squares objective function, by using non-linear regression.
- Report the estimated parameter values.
- Calculate and print statistics calculated with perturbation methods in order to: diagnose inadequate data or identify parameters that probably cannot be estimated, evaluate estimated parameter values, evaluate model fit to observations, evaluate how accurately the model represents the processes.

For this research activity, UCODE input files were prepared with the user-interface ModelMate, as it easily interacts with ModelMuse, a free user-interface for MODFLOW. Details on input files are not provided here, but can be found in Poeter et al. [100].
Once the required input files are created, UCODE can be executed in several modes:

- the *forward mode* is generally run first to check for errors in model construction, path specification and calculation of simulated values;
- the *sensitivity-analysis mode* produces statistics calculated for the parameter values specified in the main input file;
- the *parameter-estimation mode* is used to calculate parameter values that provide a better fit to the objective function;

- the *test-model-linearity mode* uses parameter values generated in the parameter-estimation mode to test the accuracy of linear confidence intervals on parameters and predictions;
- the *prediction mode* is used to calculate predictions and associated sensitivities;
- the *advanced-test-model-linearity mode* conducts additional tests of model linearity in order to account for non-linearity with respect to predictions;
- the *non-linear-uncertainty mode* calculates non-linear confidence or prediction intervals;
- the *investigate-objective-function mode* is used to investigate difficulties with the regression, such as insensitivity, local minima and extreme parameter correlation.

4.4.1 Weighted least-squares objective function

The match of simulated to observed values, measured through objective functions, is one of the most important indicators of how well a model represents an actual system. Model calibration efforts largely involve attempting to construct a model that produces a good fit, which means that the parameters estimated by model calibration produce an objective function as small as possible.
In UCODE, the *weighted least-squares objective function* is used. In the context of a groundwater model, using hydraulic heads and flows as observations, the weighted least-squares objective function, $S(\mathbf{b})$, can be expressed as:

$$S(\mathbf{b}) = \sum_{i=1}^{NH} \omega_{h_i}[y_{h_i} - y'_{h_i}(\mathbf{b})]^2 + \sum_{j=1}^{NQ} \omega_{q_j}[y_{q_j} - y'_{q_j}(\mathbf{b})]^2 + \sum_{k=1}^{NPR} \omega_{p_k}[y_{p_k} - y'_{p_k}(\mathbf{b})]^2,$$

where the vector $\mathbf{b}$ contains NP parameters to be estimated, NH, NQ and NPR are the number of hydraulic head observations, flow observations and prior information values, respectively, y_{h_i} is the i-th observed hydraulic head, y_{q_j} is the j-th observed flow, $y'_{h_i}(\mathbf{b})$ ($y'_{q_j}(\mathbf{b})$) is the simulated hydraulic head (flow) that corresponds to the i-th (j-th) observed hydraulic head (flow) and depends on $\mathbf{b}$, y_{p_k} is the k-th prior estimate, $y'_{p_k}(\mathbf{b})$ is the k-th simulated value, ω_{h_i}, ω_{q_j} and ω_{p_k} are the weights for the i-th head observation, j-th flow observation and k-th prior estimate, respectively.
Using y to indicate a generic contribution of any kind (observation or prior information) and ω to indicate its weight, the objective function is more commonly expressed as:

$$S(\mathbf{b}) = \sum_{i=1}^{ND+NPR} \omega_i[y_i - y'_i(\mathbf{b})]^2 = \sum_{i=1}^{ND+NPR} \omega_i e_i^2, \tag{4.20}$$

where ND is the number of observations and $e_i = [y_i - y'_i(\mathbf{b})]$ is called the i-th *residual*.
For prior information, UCODE supports equations of the form:

$$P'_p(\mathbf{b}) = \sum_{j=1}^{NP}(a_{p,j}b_j),$$

where p indicates the p-th prior information equation, $a_{p,j}$ are coefficients and b_j is the j-th parameter value.

More generally, the weighting requires a full weight-matrix, in which one or more of the off-diagonal matrix entries are non-zero. These off-diagonal entries are needed to represent correlated observation errors. For a full weight-matrix, the least-squares objective function in equation 4.20 is written as:

$$S(\mathbf{b}) = [\mathbf{y} - \mathbf{y}'(\mathbf{b})]^T \omega [\mathbf{y} - \mathbf{y}'(\mathbf{b})] = \mathbf{e}^T \omega \mathbf{e}, \tag{4.21}$$

where ω is a square matrix dimensioned $(ND + NPR) \times (ND + NPR)$ containing weights, $\mathbf{y}$ is a vector of $(ND + NPR)$ elements containing observations and prior information, $\mathbf{y}'(\mathbf{b})$ is a vector of $(ND + NPR)$ elements containing simulated values, and $\mathbf{e}$ is a vector of $(ND + NPR)$ elements containing residuals.

Theoretically, the least-squares objective function can be used to produce a model that accurately represents a system and provides accurate measures of model uncertainty only if three conditions are met: (1) relevant processes, system geometry and so on are adequately represented and simulated; (2) true errors (i.e., the unknown amounts by which an observation or prior information differs from the value in the actual system) of the observations and prior information are random, normally distributed and have a zero mean; (3) weighting is proportional to the inverse of the variance-covariance matrix on the true observation errors. The true errors cannot be analyzed, so weighted residuals are investigated and the characteristics of the true errors are inferred. To estimate parameter values with the least-squares objective function, the true errors must follow a normal distribution, which allows calculation of observation error variances and covariances from field data and construction of linear confidence intervals.

In equations 4.20 and 4.21, weighting is introduced for two reasons: weighted residuals must be squared and summed, so that they must have the same units; weighting will reduce the influence of less accurate observations and prior information. The latter reason allows relating requirements (2) and (3), which can be mathematically expressed as:

$$E(\varepsilon) = 0,$$

$$\omega_i \propto 1/\sigma_i^2 \qquad for\ a\ diagonal\ weight - matrix,$$

$$\omega^{1/2} \propto \mathbf{V}(\varepsilon)^{-1} \qquad for\ a\ full\ weight - matrix,$$

where ε is a vector of true errors, σ_i^2 is the variance of the true error of the i-th observation and $\mathbf{V}(\varepsilon)$ is the variance-covariance matrix of the true errors, with variances along the diagonal and covariances off the diagonal.

In UCODE, other statistics can be used instead of variance, e.g., standard deviation or coefficient of variation, to compute the weights.

4.4.2 Fit-independent statistics for sensitivity analysis

The sensitivity analysis performed in UCODE uses *fit-independent statistics*, meaning that residuals (see equations 4.20 and 4.21) are not used to calculate them. Furthermore, using these statistics allows

measuring the information provided by observations for parameter values or, viceversa, the sensitivity of the model (i.e., of the observed values) when parameters are perturbed. Moreover, sensitivities calculated by UCODE are *local sensitivities*, because they are calculated with perturbation methods for one set of parameter values. Furthermore, sensitivities are calculated by perturbing parameters one by one, i.e., they do not account for simultaneous changes in multiple parameters.
In model calibration, most of the effort is generally spent trying to use the information provided by the observations to adjust parameter values. For this reason, some statistics are calculated after calibration is successfully completed, in order to determine how well parameters are estimated.
Fit-independent statistics defined in UCODE are calculated using sensitivities, namely the derivatives of simulated quantities with respect to the model parameters:

$$\left(\frac{\partial y_i'}{\partial b_j}\right)\Bigg|_{\mathbf{b}},$$

where y_i' is the i-th simulated value that corresponds to an observation or prior information, b_j is the j-th parameter and the notation indicates that sensitivities are calculated for the parameter values listed in vector $\mathbf{b}$ (local sensitivities).
UCODE approximates sensitivities by using forward, backward or central differences. For example, the forward-difference approximation is formulated as:

$$\left(\frac{\partial y_i'}{\partial b_j}\right)\Bigg|_{\mathbf{b}} \simeq \frac{y_i'(\mathbf{b}+\mathbf{\Delta b}) - y_i'(\mathbf{b})}{\Delta b_j},$$

where $\mathbf{\Delta b}$ is a vector of zeros except for the j-th element, which is equal to Δb_j. This latter equation is calculated by running the model once using the parameter values in $\mathbf{b}$ to obtain $y_i'(\mathbf{b})$, then again after changing the j-th parameter value to obtain $y_i'(\mathbf{b}+\mathbf{\Delta b})$, and finally taking the difference and dividing by the change in the j-th parameter value (Δb_j).
Sensitivities can be used to indicate the importance of observations to the estimation of parameter values: observations are likely to be very valuable in estimating a parameter value if their simulated equivalents change substantially, given a small change in the parameter value; on the other hand, observations contribute very little to estimating a parameter if their simulated equivalents change very little, even with a large change in the parameter value.
It is often useful to compare the relative importance of different observations. A problem with this comparison through sensitivities is that the values of sensitivities are expressed in units, which depend on the units of the simulated quantities and of the parameters. The problem is overcome by scaling sensitivities. Hereinafter, the dimensionless sensitivities, used in this thesis to measure the information provided by observations for parameter values, will be defined. For further details refer to Poeter et al. [100].

Composite Scaled Sensitivities

Composite Scaled Sensitivities reflect the total amount of information provided by observations for the estimation of one parameter.

The composite scaled sensitivity for the j-th parameter calculated for ND observations is:

$$css_j = \left[\frac{1}{ND} \sum_{i=1}^{ND} \left(\frac{\partial y_i'}{\partial b_j} \right) \Big|_{\mathbf{b}} b_j (\omega^{1/2})_{ij} \right]^{1/2}.$$

Composite scaled sensitivities are often used in a comparative manner, whereby larger values indicate parameters for which the observations provide more information. If there are composite scaled sensitivities that are less than one percent of the largest value, calibration procedure often will have trouble converging. In this situation, parameter values with small composite scaled sensitivities may need to be assigned prior information or have the parameter specified.
Composite scaled sensitivities are also meaningful individually: in fact they can be interpreted as the amount by which the simulated values change, expressed as a percent of the standard deviation of the observation error, given a one-percent change in the parameter value. This interpretation shows clearly that a parameter can be estimated only if the information provided by observations, as expressed through their sensitivities, dominates the effects of observation error (noise in the data): if a css_j value is too small, the observation data may be too noisy relative to the sensitivity information provided and the calibration may not be able to estimate a value of b_j.
The interaction between the information content of the observations, as reflected in their sensitivities, and the noise in the observations suggests that there is some css_j value below which observations provide insufficient information to estimate parameter b_j. As suggested, $css_j = 1.0$ means that a one-percent change in the parameter value produces, on average, a change in simulated values that is equivalent to one percent of the measurement error standard deviation. So, parameter values with composite scaled sensitivities less than 1.0 are more likely to be poorly estimated.
Scaled sensitivities have three major limitations.

- They do not account for the possibility that while the observations may provide substantial information about individual parameters, simultaneous changes in values of more than one parameter may produce the same model fit: thus, it cannot be determined if observations can be used to estimate each parameter uniquely. This occurs when parameters are highly correlated and can be detected by calculating the parameter correlation coefficients and leverage statistics.

- The scaling by the parameter value is useful when the effect of changing parameter values by a multiplicative factor is of interest. For example, it is common to quantify errors in flow parameters, such as recharge, as a percentage of the reference value, rather than a fixed error value. Instead, for some types of parameters, scaling by the parameter value can produce misleading results. For example, parameters that represent hydraulic head at constant-head boundaries pose a special problem.

- The effects of both non-linearity and scaling by the parameter value cause scaled sensitivities to be different for different sets of parameter values: if the differences that occur for a reasonable range of parameter values are too extreme (so that different parameters are rated as important

when calculated at one set of parameter values and not important when calculated at another set), the scaled sensitivities do not perform well.

Parameter Correlation Coefficients

Problems with uniqueness occur when coordinated changes in parameter values produce the same fit to the observations. *Parameter Correlation Coefficients* indicate whether observations provide information for estimating parameters uniquely. They are calculated as:

$$pcc_{ij} = \frac{Cov(\mathbf{b})_{ij}}{Cov(\mathbf{b})_{ii} Cov(\mathbf{b})_{jj}},$$

where $Cov(\mathbf{b})_{ij}$ is the covariance between parameters b_i and b_j and $Cov(\mathbf{b})_{ii}$ and $Cov(\mathbf{b})_{jj}$ are simply the variances of b_i and b_j, respectively.
Parameter correlation coefficients are calculated for each possible pair of model parameters. They indicate whether parameter values can be estimated uniquely by regression, given the constructed model and the observations and prior information provided.
The pcc_{ij} values can vary between -1 and $+1$, while pcc_{ii} values for a parameter with itself are always $+1$. If the pcc_{ij} for a pair of parameters is equal to or very close to -1 or $+1$, the two parameters generally cannot be estimated uniquely. Accordingly, extreme correlation between more than two parameters is indicated by values of pcc_{ij} near -1 or $+1$ for all pairs of parameters. Under these conditions, the parameters involved generally cannot be estimated uniquely. Furthermore, if the absolute values of all pcc_{ij} are less than about 0.95, then it is likely that all parameter values can be estimated uniquely. However, this is a rule of thumb: experience has shown that unique estimates sometimes can be obtained even with absolute values of pcc_{ij} that are very close to 1.
In UCODE, correlation coefficients are displayed in the main output file as a matrix, which is always symmetric, with diagonal elements equal to 1. Moderate correlations between 0.85 and 0.95 are also included because parameter correlation coefficients can change substantially during calibration and it is useful to know whether a previously high correlation has become just moderately high or low.
Parameter correlation coefficients suffer from three main disadvantages:

- the non-linearity of inverse problems can cause parameter correlation coefficients to be quite different for different sets of parameter values;
- they can be inaccurate when calculated using sensitivities with an inadequate number of correct significant digits;
- as parameter sensitivity decreases (in general this occurs as more parameters are defined), greater sensitivity precision is required for the pcc_{ij} to be accurate.

4.4.3 Parameter estimation

Parameter estimation with UCODE aims at determining parameter values that produce the smallest possible value of the objective function. Most of the methods used to perform such an estimation are *gra-*

dient methods (or *regression methods*), because they use the gradient of the objective function in order to determine how to proceed toward its minimum. In UCODE, the modified Gauss-Newton non-linear regression method is used, as specified below.

Consider the weighted least-squares objective function as defined in equation 4.21. Regression methods allow finding its minimum by using its gradient, namely:

$$\frac{\partial S(\mathbf{b})}{\partial \mathbf{b}} = \frac{\partial}{\partial \mathbf{b}}\Big[[\mathbf{y} - \mathbf{y}'(\mathbf{b})]^T \omega [\mathbf{y} - \mathbf{y}'(\mathbf{b})]\Big] = \mathbf{0}, \tag{4.22}$$

where $\mathbf{0}$ is a vector of NP elements with zero value.

When $\mathbf{y}'(\mathbf{b})$ is non-linear, equation 4.22 is solved by approximating it as a linear function, using two terms of a Taylor series expansion:

$$\mathbf{y}'(\mathbf{b}) \simeq \mathbf{y}^{\ell}(\mathbf{b}) = \mathbf{y}'(\mathbf{b}_0) + \left.\frac{\partial \mathbf{y}'(\mathbf{b})}{\partial \mathbf{b}}\right|_{\mathbf{b}=\mathbf{b}_0} (\mathbf{b} - \mathbf{b}_0) = \mathbf{y}'(\mathbf{b}_0) + \mathbf{X}|_{\mathbf{b}=\mathbf{b}_0}(\mathbf{b} - \mathbf{b}_0),$$

where $\mathbf{y}^{\ell}(\mathbf{b})$ is the linearized form of $\mathbf{y}'(\mathbf{b})$, $\mathbf{b}_0$ is the vector of parameter values about which $\mathbf{y}'(\mathbf{b})$ is linearized and $\mathbf{X} = \left.\frac{\partial \mathbf{y}'(\mathbf{b})}{\partial \mathbf{b}}\right|_{\mathbf{b}=\mathbf{b}_0}$ is the sensitivity matrix calculated using the parameter values listed in vector $\mathbf{b}_0$. Notice that $\mathbf{y}'(\mathbf{b})$ has $(ND + NPR)$ elements, the vector $\mathbf{b}$ has NP elements, so that $\mathbf{X}$ has $(ND + NPR)$ rows and NP columns.

With this approximation, the least-squares objective function can be espressed as:

$$S^{\ell}(\mathbf{b}) = [\mathbf{y} - \mathbf{y}^{\ell}(\mathbf{b})]^T \omega [\mathbf{y} - \mathbf{y}^{\ell}(\mathbf{b})]$$

and equation 4.22 is solved, so obtaining what is called a *normal equation* for $\mathbf{b}$, whose matrix and iterative form is:

$$(\mathbf{X}_r^T \omega \mathbf{X}_r)\mathbf{d}_r = \mathbf{X}_r^T \omega (\mathbf{y} - \mathbf{y}'(\mathbf{b}_r)), \tag{4.23}$$

where r is the iteration index, $\mathbf{X}_r$ is the matrix of sensitivities calculated for parameter values included in $\mathbf{b}_r$, $\mathbf{d}_r$ is the NP-dimensional parameter change vector used to update the parameter values and $\mathbf{b}_r$ is the vector including parameter values at the beginning of r-th iteration.

For the first iteration, the model is linearized around starting parameter values defined by the modeler. In each subsequent iteration, the model is linearized around parameter values estimated at the previous iteration. For each iteration, equation 4.23 is solved for $\mathbf{d}_r$, and then $\mathbf{d}_r$ is used to update the parameter values for the beginning of iteration $r + 1$, using the equation $\mathbf{b}_{r+1} = \mathbf{b}_r + \mathbf{d}_r$.

The right-hand side of equation 4.23 is proportional to the gradient of the linearized objective function, so without the $(\mathbf{X}_r^T \omega \mathbf{X}_r)$ term on the left-hand side of equation 4.23, the parameter change vector ($\mathbf{d}_r$) would point directly down the gradient of the linearized objective-function surface, as shown by arrow A in Figure 4.4. This is called the *steepest descent direction*. The $(\mathbf{X}_r^T \omega \mathbf{X}_r)$ term modifies the direction of $\mathbf{d}_r$ to point toward the minimum of the linearized objective-function surface, as shown by arrow B in Figure 4.4.

The basic Gauss-Newton method presented in equation 4.23 is prone to difficulties, such as oscillations due to overshooting the optimal parameter values. It only works well when modified by using scaling, the Marquardt parameter and damping.

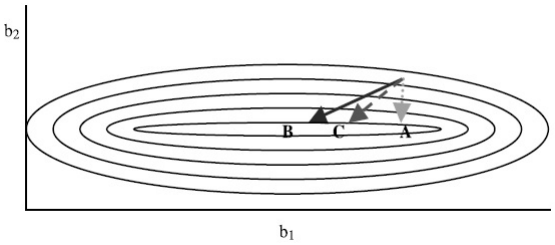


Figure 4.4: A linearized objective-function surface for a hypothetical two-parameters problem, illustrating the geometry of the normal equations. The arrows represent the direction relevant to the parameter change vector ($\mathbf{d}_r$). Arrow A points down gradient in a direction defined by the right-hand side of equation 4.23. Arrow B points in the direction of $\mathbf{d}_r$, as described by equation 4.23 or 4.24. Arrow C shows that the direction of $\mathbf{d}_r$, obtained by using a non-zero Marquardt parameter in equation 4.25, is between arrows A and B [72].

Scaling Often parameter values, and thus sensitivities, differ by many orders of magnitude. This can cause great difficulties to obtain an accurate solution of equation 4.23.
The accuracy of the parameter change vector ($\mathbf{d}_r$) can be improved by scaling equation 4.23 as follows:

$$(\mathbf{C}^T\mathbf{X}_r^T\omega\mathbf{X}_r\mathbf{C})\mathbf{C}^{-1}\mathbf{d}_r = \mathbf{C}^T\mathbf{X}_r^T\omega(\mathbf{y} - \mathbf{y}'(\mathbf{b}_r)), \tag{4.24}$$

where $\mathbf{C}$ is a diagonal scaling-matrix with elements $c_{ii} = [(\mathbf{X}^T\omega\mathbf{X})_{ii}]^{1/2}$.
Scaling with $\mathbf{C}$ changes the magnitude but not the direction of $\mathbf{d}_r$. Therefore, in Figure 4.4 the parameter change vector ($\mathbf{d}_r$) still points in the direction of arrow B after scaling has been implemented.

Marquardt Parameter In some circumstances, the direction of $\mathbf{d}_r$ is nearly tangent to the iso-surfaces of the objective function and changing the parameter values using $\mathbf{d}_r$ yields little progress toward estimating optimal parameter values. In this case, changing the direction of $\mathbf{d}_r$ can be advantageous. For this reason, the second modification involves the introduction of a term that causes the direction of vector $\mathbf{d}_r$ to move toward the steepest descent direction. Such a term is called the Marquardt parameter (m_r) and is included in equation 4.24 as follows:

$$(\mathbf{C}^T\mathbf{X}_r^T\omega\mathbf{X}_r\mathbf{C} + \mathbf{I}m_r)\mathbf{C}^{-1}\mathbf{d}_r = \mathbf{C}^T\mathbf{X}_r^T\omega(\mathbf{y} - \mathbf{y}'(\mathbf{b}_r)),$$

where $\mathbf{I}$ is the $(NP \times NP)$ identity matrix.
In Figure 4.4 a non-zero Marquardt parameter moves the direction of $\mathbf{d}_r$ from the direction of arrow B to the direction of arrow C.
The Marquardt parameter is initially zero; then, if $\mathbf{d}_r$ is nearly orthogonal to the steepest descent direction, it is changed to a non-zero value. In particular, if the cosine of the angle between the vector $\mathbf{d}_r$ and the vector orthogonal to the steepest descent direction is less than a threshold value (commonly 0.08), m_r is increased using a linear relation.

Damping Overshoot is a common problem with the Gauss-Newton method, when $\mathbf{d}_r$ points toward locations on the objective-function surface that are closer to its minimum, but then extends beyond these locations to larger objective-function values. Under these circumstances, damping prevents overshoot by allowing the parameters to change less than the full amount calculated by $\mathbf{d}_r$ and is applied by multiplication of $\mathbf{d}_r$ times a damping factor (ρ_r).
Finally, the matrix and iterative form of the *modified Gauss-Newton non-linear regression normal equations* is:

$$(\mathbf{C}^T\mathbf{X}_r^T\omega\mathbf{X}_r\mathbf{C} + \mathbf{I}m_r)\mathbf{C}^{-1}\mathbf{d}_r = \mathbf{C}^T\mathbf{X}_r^T\omega(\mathbf{y} - \mathbf{y}'(\mathbf{b}_r)), \tag{4.25}$$

$$\mathbf{b}_{r+1} = \rho_r\mathbf{d}_r + \mathbf{b}_r.$$

The damping parameter (ρ_r) shortens $\mathbf{d}_r$, modifying all values inside it by the same factor, and can vary in value from 0.0 to 1.0. Thus, the direction of $\mathbf{d}_r$ (represented by arrow C in Figure 4.4) is preserved.
For each parameter-estimation iteration, the damping parameter is initially 1, then it is changed to a smaller value to ensure that all quantities $d_j^r/|b_j^r|$ ($j = 1, ..., NP$) are less than a prescribed value specified by the user ($maxchange$):

$$\rho_r = \frac{maxchange}{\max(d_j^r/|b_j^r|)} \qquad j = 1, ..., NP.$$

After that, $\mathbf{b}_{r+1}$ is calculated and contains the parameter values for starting the next parameter-estimation iteration.

4.4.4 Convergence criteria

Convergence criteria are needed to determine when to stop the modified Gauss-Newton iterative process. In UCODE, parameter estimation converges if either one of two convergence criteria are satisfied.
By the first criterion, convergence is achieved when the parameter values change only a small amount from one parameter-estimation iteration to the next. This indicates that, at the current regression iteration, the parameter values lie in a relatively flat area that is a minimum in the objective-function space. This condition is satisfied if:

$$\max\left(\frac{d_j^r}{|b_j^r|}\right) < maxchange \qquad j = 1, ..., NP.$$

Tipically, $maxchange$ is 0.01 or 0.001 for final regressions, indicating that convergence is reached when parameter values are changing between parameter-estimation iterations no more than 1% or 0.1%. Generally, $maxchange$ is for all parameters, but UCODE allows assigning different values of $maxchange$ to different parameters.
By the second convergence criterion, the non-linear regression converges if the model fit changes little over the course of two parameter-estimation iterations: if three consecutive values of the least-squares objective function (see equations 4.20 or 4.21) change less than a user-defined amount, non linear regression converges.

The second criterion often is useful early in the calibration process to avoid lengthy simulations that fail to improve model fit. However, satisfying this criterion does not provide as strong an indication that a minimum has been reached as the parameter-value criterion. Therefore, for final regression runs, it is preferable that the parameter-value criterion be satisfied.
Modified Gauss-Newton optimization typically converges within a number of iterations equal twice the number of parameters, even if well-conditioned problems (commonly those with large css_j values and little correlation) tend to converge in fewer iterations.
UCODE allows parameter values to be log-transformed. Log-transformed parameters are often useful because the uncertainty of many parameters is best represented by a log-normal probability distribution: when the parameter is log-transformed, the uncertainty is then best represented by a normal distribution, which is convenient to use. Log-transforming parameters can produce an inverse problem that converges more easily and prevents the native parameter values from becoming negative. Tipically, log-tranformation is used for parameters whose values can vary over several order of magnitudes, such as hydraulic conductivity.
UCODE supports also upper and lower limits on parameters that constrain possible estimated values. Such limiting constraints may be necessary given the unrealistic parameter values that can be estimated through inverse modeling, if data do not contain enough information to estimate them or there is a fundamental model error. In the first circumstance, the best response is to use prior information on the parameter value, which tends to produce an estimate that is close to the most likely value. In the second circumstance, the best response is to find and resolve the error.
Often parameters are estimated only after using starting parameter values to evaluate model fit and perform a sensitivity analysis to identify insensitive and correlated parameters. Execution of UCODE for these purposes proceeds through several steps.

- First of all, after instruction files are created with starting parameter values, these values are used to run the model in order to obtain equivalent (unperturbed) simulated values to be compared to the observations. Then, the unperturbed residuals are squared, weighted and summed to produce the least-squared objective function, which measures how well the model fits the observations (forward mode).

- Sensitivities are then calculated using forward-difference perturbation: the process model is executed once for each parameter and, for each execution, one parameter value is slightly increased (perturbed) from its unperturbed value, while the other parameter values are not perturbed. Thus, the differences between the resulting perturbed simulated values and the unperturbed simulated values are used to calculate forward-difference sensitivities (sensitivity-analysis mode).

- The goal of regression is to find parameter values that produce the smallest value of the least-squared objective function. The non-linear regression method performed in UCODE uses sensitivities to find those parameter values. In parameter-estimation mode, sensitivities are usually calculated with central-difference method for accuracy reasons. Furthermore, UCODE allows omitting insensitive parameters so that they do not disrupt regression performance.

- If parameter estimation does not converge and the maximum number of iterations has not been reached, the updated parameter values are substituted into the template files and the next parameter-estimation iteration is performed. When parameter estimation converges or the maximum number of iterations has been reached, these final parameter values are considered to be optimized and final sensitivities are calculated.

Part II

Application to the Taranto area

Chapter 5

Description of the study area

Among coastal case studies taken into account within the RITMARE project, the gulf of Taranto (southern Italy) is very representative, because here commercial, shipping, industrial and urban activities coexist at a short distance, thus producing an ideal case to study the interaction among these different potential sources of contamination and the maritime space. According to the purposes of RITMARE and the issues about groundwater quality in the Taranto area (see section 5.1), the present research project aims at providing a characterization of the multi-layered aquifer and an insight about the relationship between groundwater and seawater.
In this chapter, the study area is described from the environmental, geomorphological, geological and hydrogeological point of views. This will include also an overview of the major issues that provide the reasons for such a study.

5.1 Geographical and environmental setting

The study area (Figure 5.1) is located in the northern Ionian side of the Apulia region (south-eastern Italy) and extends for about $2\,800\ \mathrm{km}^2$, including almost the whole Province of Taranto. This area is bounded by the Ionian coast to the south side, by the Murge plateau to the north side, by the line which roughly connects Manduria and Ostuni to the east side and includes part of the Laterza municipality to the west side. Notice that, although the RITMARE project deals with issues related to a more restricted area (as specified hereinafter), for the aims of this study the target area extends over the whole catchment. Specifically, the extreme boundaries of the described area have the following metric coordinates with reference to the UTM system: $X_{min} = 651\,050\ \mathrm{m}$ and $X_{max} = 724\,550\ \mathrm{m}$ along the WE direction, $Y_{min} = 4\,463\,250\ \mathrm{m}$ and $Y_{max} = 4\,513\,250\ \mathrm{m}$ along the SN direction.
Within the area shown in Figure 5.1, Taranto ($4028'N$; $1714'E$) is the main town of the Province and the third more densely-populated town of southern Italy (about $198\,000$ inhabitants) [2]. During the last decades, Taranto has been undergoing quick and rather uncontrolled urban and industrial development, which has deeply modified the original environmental asset. In fact, a military and trade harbour, an important industrial centre with petrochemical and cement plants, naval ship-building industry and one

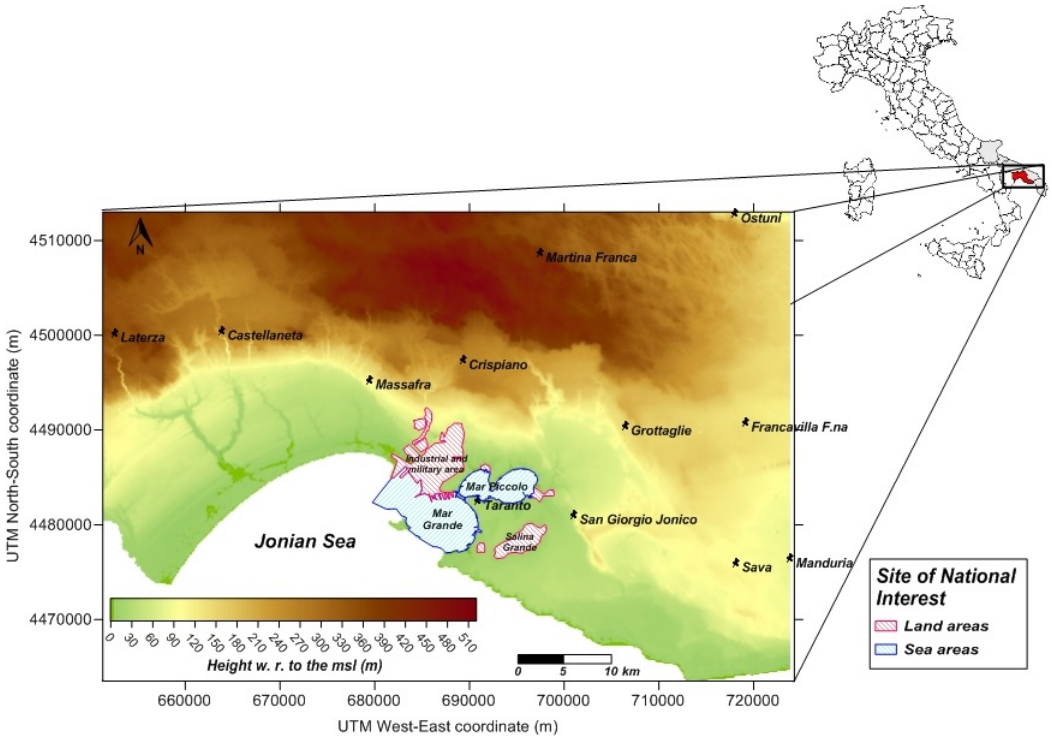


Figure 5.1: Location of the study area. The color map refers to the Digital Terrain Model (spatial resolution of $8\ \mathrm{m}$ in the horizontal plane) [117]. Bounded areas belong to the Site of National Interest (Ministerial Decree dated September 20$^{\text{th}}$ 2001).

of the biggest steel-manufacturer plants in Europe (Ilva) are located in the same area, close to the Taranto urban centre.

The presence of a wide industrial area close to the town and the numerous maritime and military activities in the harbour area represent factors of great interest which are worth attention. In fact, these activities heavily affect both the landscape and the socio-economic and environmental frameworks, thus representing a threat for ecosystems, due also to the ongoing anthropization. In this regard, several studies were carried out in order to highlight significant effects on air quality in this area, due to productive activities and vehicular traffic [2, 3, 12, 22, 56]. Besides this, several environmental components are involved, namely sediments, surface and subsurface hydrography [21]. In fact, in addition to air emissions, landfills and sites of illegal disposal of waste are present. Moreover, not-purified wastewater is ofted delivered in superficial, possibly intermittent, torrential watercourses, such as Paternisco and Canale d'Aiedda [132], which in turn discharge into the Mar Piccolo basin.

The *Mar Piccolo* (see Figure 5.1) is a semi-enclosed basin with lagoon features, whose surface area is $20.72\ \mathrm{km}^2$ and it is divided by two promontories into two smaller basins, named the first and second inlets (Figure 5.2). The maximum depth is $15\ \mathrm{m}$ in the centre of the first inlet and $10\ \mathrm{m}$ in the central part of the second inlet [132]. The basin is the most important area of mussel farming in Italy and is characterized by the presence of some submarine freshwater springs, locally called "*citri*", whose flow rates are consistent (overall more than $1\ \mathrm{m}^3/\mathrm{s}$, see section 6.3). Furthermore, it is characterized by scarce

Figure 5.2: Panoramic picture of the second inlet.

water fluxes, moderate sea-tides (tidal range does not exceed $30 - 40$ cm [24]) and low water exchange with the nearby Mar Grande; these features determine a high water stratification especially in summer. The scarce water circulation inside the Mar Piccolo makes it particularly vulnerable to human polluting activities. The main problems of environmental impact are: urban pollution due to many sewage outlets irregularly processed, shipyard activities of the Italian Navy located into the first inlet, aquaculture and commercial fishing activities in both inlets, small rivers and freshwater springs which drain urban waste and brakish groundwater, respectively. So, human activities seriously affect the well-being of this ecosystem, due also to the particular morphology of the basin itself [41], with obvious implications on sediments and water quality [1, 24, 29, 122] and, in turn, on human health.
The coastal ecosystem of the Taranto Sea includes also the *Mar Grande* (see Figure 5.1), directly connected with the open sea and with the neighbouring Mar Piccolo through two channels, called as "Navigabile" channel and "Porta Napoli" channel. The Mar Grande covers an area of $35\ \mathrm{km}^2$, with a maximum depth of about 35 m and a mean depth of about 15 m [132] and hosts a freshwater spring, the so called "Anello di San Cataldo". Quality of sediments and seawater is a remarkable issue also for this basin, mostly due to commercial and industrial harbour activities, but also because it receives both purified wastewater from the biggest industrial settlements of the area and non purified wastewater from the town sewer network. For these reasons, studies have been carried out in order to understand factors which mostly affect water exchanges with the Mar Piccolo (e.g., tides and wind) [41, 132].
The Ministerial Decree dated September 20$^{\text{th}}$ 2001 officially declared the area nearby Taranto as "at high risk of environmental crisis". This area covers approximately $115\ \mathrm{km}^2$, along a coastline 17 km long, and specifically contains (see Figure 5.1):

- the industrial pole including the Ilva steel industry, the Agip refinery, the Cementir cement industry and several manufacturing small to medium enterprises;
- the Mar Piccolo and Mar Grande basins;
- several military and civil yards;
- a saline located south-east of Taranto (Salina Grande);
- some disused quarries, which show signs of local instability;
- some landfills not properly waterproofed, whose percolate threatens the quality of groundwater resources.

To sum up, the effects due to commercial, shipping, industrial and urban activities occurring at a short distance nearby Taranto are noteworthy and involve several environmental components. This makes the study area "at high risk of environmental crisis" and an ideal target for scientific research, in order to assess and restore critical situations.

5.2 Geological and geomorphological setting

The Apulia region is located on the Apulian platform, which is situated on the south-western margin of the Adriatic tectonic plate, in the central Mediterranean, and forms a peninsula which separates the Adriatic from the Ionian Sea [90]. The study area (Figure 5.1), in turn, is located in the south-western part of the Apulian foreland, which represents the emerged part of the foreland domain of both the Apennine and Dinar orogens and is close to the Bradanic foredeep (Figure 5.3), namely the southern sector of the Apennine foredeep, whose formation is related to that of the southern Apennines, which began to rise during the late Oligocene, when the Adriatic plate started to subduct under the European plate [111]. The region is characterised by a Variscan basement covered by Mesozoic limestones, whose thickness reaches $5\ \mathrm{km}$, and thin, discontinuous deposits of Pliocene and Pleistocene ages, up to $70\ \mathrm{m}$ thick, which are represented by the Gravina calcarenites and the Subappennine clays formations. The Mesozoic and Plio-Pleistocene units are then covered by middle-to-upper Pleistocene beach deposits [8, 15, 16, 17, 43, 88, 89, 90, 111].

The Apulian foreland is slightly deformed and is affected by Apenninic (NW-SE) and anti-Apenninic (NE-SW) trending faults, which cut it into three different geodynamic units characterised by different uplift rates: the Gargano promontory, the Murge plateau and the "Serre" of the Salento Peninsula (Figure 5.3). These blocks are further separated by intermediate low plains which are subjected to relatively lower uplift rates: the Tavoliere plain, the Ofanto graben and the Taranto-Brindisi plains [8, 15, 80, 88, 89, 90, 114]. The subsidence of the Apulian foreland, which started during the early Miocene, was followed by the general uplift of the area which began in the middle Pleistocene, when the thick continental lithosphere - the Apulian swell - reached the Apennine subduction hinge, thereby offering greater resistance to flexure. The subsequent slowing down of the eastward rollback of the subduction hinge and the penetrating slab induced the buckling of the foreland, which was ultimately responsible for the uplift of this region. However, the uplift rates of the Gargano promontory, the Murge plateau and the "Serre" of the Salento Peninsula seem to decrease from north to south on the Ionian side (from about $0.4\ \mathrm{mm/y}$ in the Basilicata coastal area to about $0.2\ \mathrm{mm/y}$ in the Taranto coastal area) and approach subsidence close to the southernmost end of Apulia and, generally, from Murge to Salento [8, 15]. From the middle Pleistocene onwards, the superimposition of regional uplift and glacioeustatic sea level changes produced along the Apulian coastaline a number of staircase arranged marine terraces, stretching from an elevation of $400\ \mathrm{m}$ to a few meters above the mean sea level [8, 16, 17, 88, 90, 111, 147].

Besides tectonic causes and sea level changes, further modifications were due to karstic processes during periodic uplifts of the foreland, when the water flowing on the surface infiltrated through fractures, cracks and bedding planes within the limestone stretching them more and more. This is due to the corro-

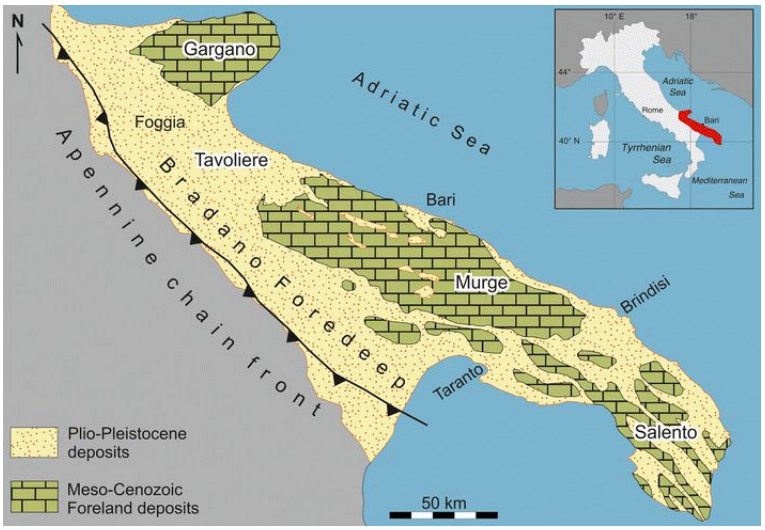


Figure 5.3: Geological sketch map of Apulia region [115].

sive action of atmospheric carbon dioxide dissolved in water, which increases with pressure, thus being responsible for the formation of karst structures, such as sinkholes and dolines [20, 82].
Focusing on the study area (Figure 5.1), two zones can be identified, separated by the Crispiano-Grottaglie alignment. The northern zone corresponds to the southernmost area of the Murge plateau and is characterised by the outcrop of the Mesozoic limestone unit, while the southern one is characterised by the outcrop of the Gravina calcarenites and Subappennine clays, cutted by a staircase of marine terraces. Hereinafter a detailed description of the geological units recognized within the study area, some of which include several lithostratigraphic units, is provided (see also Table 5.1) [39, 65, 83].

- The oldest rock formations are represented by Mesozoic carbonatic rocks made up predominantly of white, dark brown or grey compact limestones and dolomitic limestones. These formations, which constitute the geological foundation of the whole Apulia region, are characterized by evident stratification and fracturing and they have also been subject to widespread karst phenomena, which make them particularly permeable. This geological unit is represented in the study area by the ***Altamura limestone*** (***AltLim***), dated back between the upper Cretaceous and Oligocene. It widely outcrops inland, where its top surface reaches a height of about $500\ \mathrm{m}$ above the mean sea level (msl) in the upper central part, and a depth of about $300\ \mathrm{m}$ below msl near the coast, south of the Mar Piccolo (see section 6.1).

- The ***Galatone formation*** (***Gal***) dated back to the Oligocene is characterized by thin layers, made of grey-white, fine-grained and compact limestones, marls and sandy clays. Due to this wide lithologic heterogeneity, this unit shows a strongly variable porosity. It is locally present within the study area, along the coast north-west and south-east of Taranto, with a thickness of about $80\ \mathrm{m}$ near the coast, south of the Mar Piccolo (see section 6.1).

- Proceeding in the stratigraphic succession from bottom up, we find lower-Pleistocenic ***Gravina calcarenites*** (***GCalc***), transgressive on limestones. This unit is made of yellow and coarse-grained

calcarenites, whose thickness widely varies in the study area until a maximum of about $60\ \mathrm{m}$ (see section 6.1).

- Proceeding further upward, the ***Subappennine clays*** (***SubCl***), dated back to the middle-lower Pleistocene, are made of poorly permeable clayey and sandy-clayey sediments. This unit outcrops at discontinuous locations along an arc from the western to the south-eastern part of the study area and its thickness increases from inland towards the coast, where it reaches values of about $240\ \mathrm{m}$ south of the Mar Piccolo (see section 6.1).
- On top of the Subappennine clays we can find two geological units dated back to the upper Pleistocene: coarse-grained calcarenites on the eastern side and conglomerates and marine sands on the western side. Overall, they are grouped and known as the ***Terrace deposits*** (***Terr***). They are widely outcropping along the coastline, with a maximum thickness of about $25\ \mathrm{m}$ on the western side of the study area (see section 6.1).
- Finally, lagoon and marshy silts are found, predominantly in the western sector of the study area, along with current and recent alluvial deposits. These sediments are known as the ***Continental deposits*** (***Cont***), with generally low permeability and thicknesses of few meters (see section 6.1).

HU code	**Lithostratigraphic units**	**Lithological features**	**Geological age**	**Hydrogeological features**
AltLim	Altamura limestone Torre Specchia La Guardia, Torre Tiggiano and Castro limestones	Limestones and dolomitic limestones	Cretaceous Eocene Oligocene	Fractured and karst main aquifer
Gal	Galatone formation	Limestones, marls and sandy-clays	Oligocene	Impermeable deposits hosting sulphur waters
GCalc	Gravina calcarenites	Biodetritic calcarenites	Lower Pleistocene	Porous aquifer
SubCl	Subappennine clays	Clays	Lower to Middle Pleistocene	Aquitard
Terr	Brindisi sands Terrace deposits	Sands and biodetritic calcarenites	Upper Pleistocene	Porous shallow aquifers
Cont	Continental deposits	Clays and sands	Holocene to Recent	Generally impermeable

Table 5.1: List of the recognized hydrostratigraphic units (HUs) and their characteristics.

5.3 Hydrogeological setting

Due to its karstic nature, the study area (Figure 5.1) shows a poor surface hydrography characterized by torrential channels which, during the rainy periods, collect the meteoric water that does not infiltrate

the subsoil and rapidly carry it toward the sea. On the contrary, the underground is particularly rich in freshwater resources, thanks to the presence of a multi-layered aquifer system, which satisfies most of the water demand of the region, especially for irrigation and industrial purposes (see section 6.3) [80].
The overall underground hydrography is quite simple and characterized by two separate systems of circulation, based on porosity of different hydrostratigraphic units (HUs) [80]. The first system is constituted of water flowing in the Pleistocenic soils (whose permeability is due to pore porosity) overlying the impermeable SubCl HU and is generally known as the *shallow groundwater*. The second system is constituted of water flowing in the AltLim HU (whose permeability is due to fractures and karst processes) and is known as the *deep groundwater* [33, 39, 65, 83, 101]. The deep aquifer may be in contact with the Gravina calcarenites, which show a various permeability, both due to porosity and karst processes, and can host local porous aquifers.
The deep aquifer is fed by rain infiltration inland, where the AltLim HU outcrops, and by groundwater inflow from the adjacent Murge plateau and by vertical exchange with shallow aquifers [39, 65]. Moreover, the shallow aquifer is under phreatic conditions, while the deep one is confined or unconfined based on the presence of the SubCl HU which overlies the AltLim HU.
Near the coast, submarine and subaerial freshwater springs are present (e.g., "citri" springs within the Mar Piccolo basin): these outflows occur where the deep aquifer is confined but the impermeable clayey coverage leaks or is missing [101]. In addition, subaerial freshwater springs fed by the surface aquifer have much lower flow rates and often are dry during arid periods (see section 6.3). Among the activities related to the RITMARE project, qualitative and quantitative monitoring of some freshwater springs was carried out by the CNR-IRPI Bari Operative Unit [152].
The transition zone between freshwater and saltwater in this multi-layered, carbonatic, coastal system can be found within the AltLim HU [39, 65]. During the last decades, a gradual and continuous decrease of the deep aquifer volume has been occurring because of the overexploitation of deep freshwater resources, so causing an increase of salinity values of groundwater [80, 84]. In this framework, studying the seawater intrusion phenomenon is of paramount importance in order to provide guidelines for management of the most important freshwater resource of the region [39, 65]. For this purpose, developing detailed hydrostratigraphic and groundwater flow models in the study area is necessary: this would point out the main directions of the groundwater outflows, thus helping to prevent eventual circulation of contaminants, as a result of accidental or intentional spills.

5.4 Data used for the conceptual and numerical model

For the aims of this discussion, different datasets and information were considered and properly integrated, including technical reports, well data and information deduced from official maps.
First of all, the development of the hydrostratigraphic model (see section 6.1) started by collecting stratigraphic information from wells scattered over the study area. For the current research activity, 40 wells mostly located along the coastline (red dots in Figure 5.5, courtesy of Ente Irrigazione) were selected from a dataset already used in previous works to characterize the groundwater dynamics in the deep,

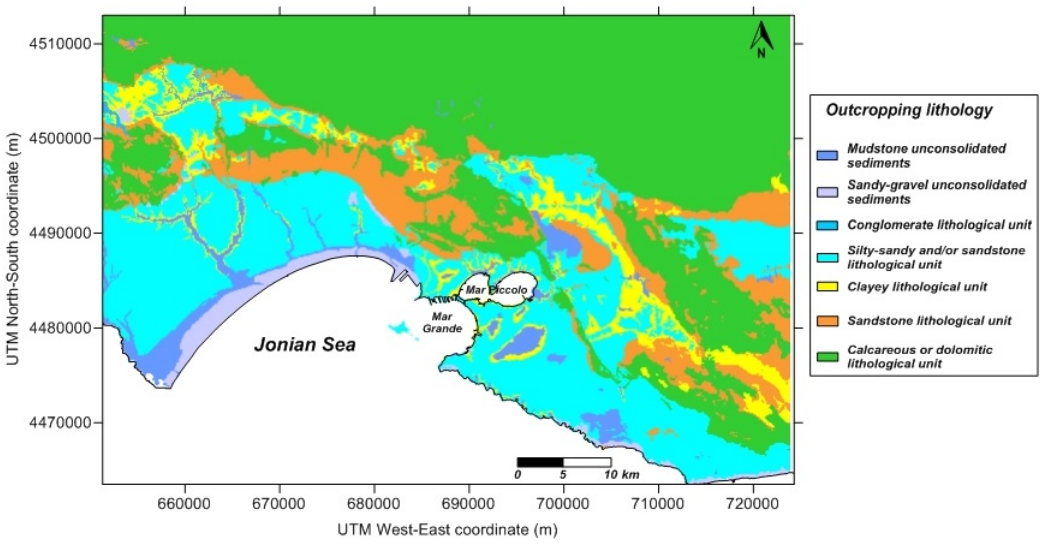


Figure 5.4: Map of the outcropping lithological units (spatial resolution of 8 m in the horizontal plane) [117].

karst aquifer of the whole Salento peninsula at the regional scale [39, 65]. The lack of stratigraphic data inland required the consultation of official maps. For this scope, the metadata catalog provided by regional authorities through the GIS "*SIT-Puglia*" [117] was consulted: among several cartographic (e.g., ortophoto maps) and thematic (e.g., land use maps, geomorphological maps) data, the Digital Terrain Model (DTM) map (see Figure 5.1) and the map of the outcropping lithological units (Figure 5.4) were selected. Since these maps provided data with a spatial resolution of 8 m in the horizontal plane, they were adapted to the computational grid (spacing of 500 m in the horizontal plane, see section 6.1) by considering the arithmetic mean for DTM values (Figure 5.5) and the mode for lithology codes. Furthermore, for the current analysis, the lithological units listed in Figure 5.4 were interpreted in terms of the recognized HUs (see Table 5.1), after a comparison with the geological map of the Murge plateau and Salento peninsula [106]. The map of the outcropping HUs obtained after such a comparison is shown in Figure 5.6.

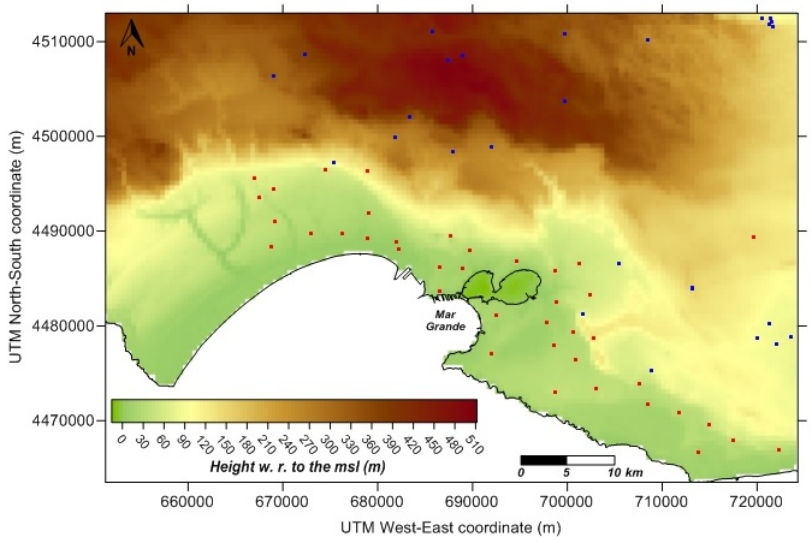


Figure 5.5: DTM values (spatial reslution of $500\ \mathrm{m}$ in the horizontal plane). Red dots indicate locations of wells (courtesy of the Ente Irrigazione), whose data were used to get the hydrostratigraphic model and the reference hydraulic head map for the deep aquifer. Blue dots indicate locations of wells [120], whose data were used to get the reference hydraulic head map for the deep aquifer (see section 6.2).

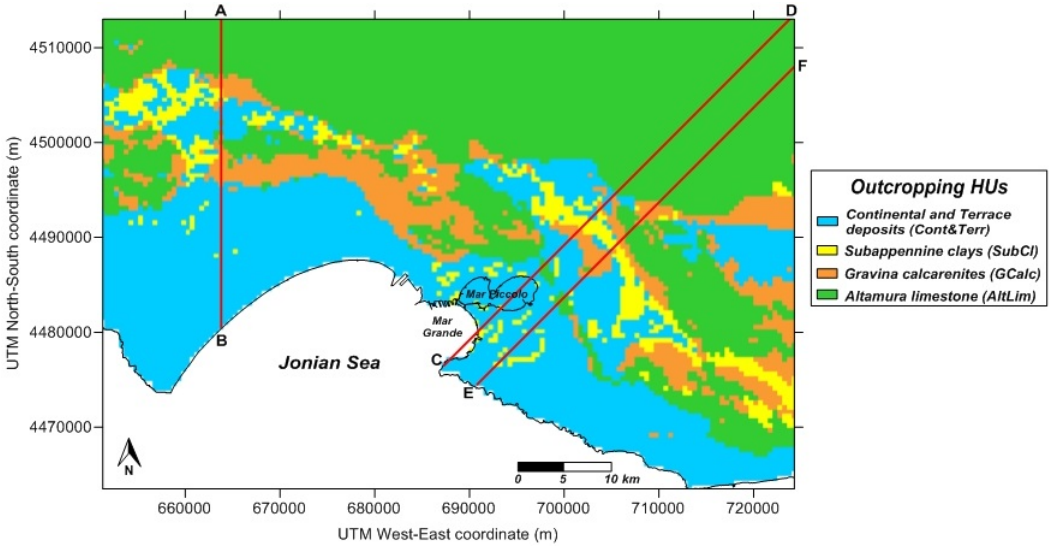


Figure 5.6: Map of the outcropping hydrostratigraphic units (spatial reslution of $500\ \mathrm{m}$ in the horizontal plane). The red profiles refer to the cross-sections AB, CD and EF drawn in Figure 6.6.

Notice that, in Figure 5.6, the most recent HUs (Cont and Terr) are joined into a single layer (Cont&Terr), due to the lack of available well data to infer the thickness of the Cont HU. However, despite this choice is not completely representative of the hydraulic features of these HUs, it is justified from a more practical point of view both by the aims of this study and the low values of thickness for the Cont HU (just few meters), as reported in previous papers [39, 65] and observed from field surveys.
Furthermore, for the aims of this discussion, the previous maps and all the following ones include also the

Mar Piccolo basin. In this regard, the DTM values inside the Mar Piccolo (Figure 5.5) were calculated as an arithmetic mean of the bathymetry values of the sea floor, provided by the CNR-ISMAR Venezia Operative Unit within the RITMARE project (personal communication). On the other hand, several lithological surveys inside the Mar Piccolo were found in literature [125] and the described units were coupled to the Cont&Terr HU (see Figure 5.6).
For 38 out of 40 wells from the dataset granted by Ente Irrigazione (red dots in Figure 5.5), data about the hydraulic head (h) related to the deep, karst aquifer under static conditions (i.e., without pumping) are known. Further data were found in technical reports [120] and selected according to the physical consistency of their values with information obtained by official maps [5, 6] used as a comparison (see blue dots in Figure 5.5). The published hydraulic head maps [5, 6] were used also to constrain the h values in zones where well data are missing.
The evaluation of the recharge term due to rain infiltration required integrating different datasets as well: the mean annual rainfall [118], the lithology map (Figure 5.6) and the land use map (Figure 5.7). First of all, the estimate of the mean annual rainfall was possible thanks to several data recorded by 20 pluviometric stations within the study area and around it (black dots in Figure 6.8). These stations are managed by Servizio Idrografico e Mareografico Nazionale (Bari district) and the mean of the total annual rainfall recorded over more than 20 years until 1996 [118] was computed. Furthermore, as already done for the lithology map, the land use map from the SIT-Puglia [117], with a spatial resolution of 8 m in the horizontal plane, was adapted to the grid with a spacing of 500 m in the horizontal plane by considering the mode of the land use codes for each pixel of the coarse scale map. In addition, the following categories for land use were introduced in order to merge similar typologies together (Figure 5.7):

- urban areas;
- green urban areas (e.g., parks and public gardens);
- uncultivated areas;
- wet areas (including the Mar Piccolo);
- vineyards;
- olive trees;
- other crops;
- wooded areas.

Regarding withdrawals from the deep aquifer for irrigation, drinking and industrial purposes, the technical report provided by SOGESID S.p.A. & Regione Puglia [119] was consulted to get the total extracted flow rates for the Province of Taranto.
Finally, outflows from freshwater springs were esimated from measures found in the literature [34, 45,

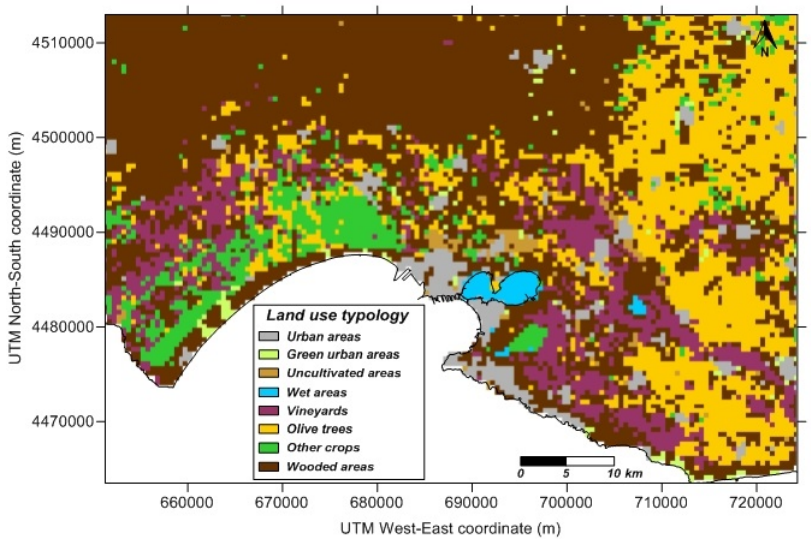


Figure 5.7: Land use map.

105], while for some of them measurements were performed by the CNR-IRPI Bari Operative Unit within the RITMARE project [152].

Chapter 6

Conceptual model for the multi-layered aquifer system

The accuracy of a numerical flow model depends on the definition of the *conceptual model* of the aquifer and on a proper estimation of the water balance [78]. In fact, conceptual models act as the basis for modelling groundwater systems, as they allow understanding their physical behavior [146]. Furthermore, hydrogeological conceptualization makes a major contribution to the sustainability and management of water resources and has become an invaluable tool in understanding water-related mechanisms and communicating with practitioners, researchers, stake-holders, decision makers and society in order to implement water resources programs within an environmental, sustainable and geoethical framework [31].

6.1 Hydrostratigraphic model of the underground

Aquifer characterization is a matter of importance in groundwater modeling applications because the representation of spatial variability has a substantial influence on the hydraulic behavior of the system. However, the difficulty in characterizing subsurface heterogeneity with commonly sparse data sets severely limits the accuracy and realism of numerical models. Moreover, lithologic log data may provide excellent information about the vertical variability of the sediments but only limited lateral variability information. For this reason, several methods used for modeling heterogeneity and spatial connectivity are necessary, such as interpolation techniques [9].
The data processing described in this chapter was performed over a discretization grid of the study area consisting of 147×100 square cells with $500\ \mathrm{m}$ side length in the horizontal plane. The choice of such a spatial detail aims at matching both the spatial distribution of the available data and the primary object of this study. First of all, the amount of data used for this analysis is not such as to allow finer spatial resolution (see dots in Figure 5.5). On the other hand, an overall characterization over the scale of the whole catchment does not require such a great spatial detail as for studies at the local scale, when focusing on areas of particular interest (e.g., the industrial area).

To get a reliable hydrostratigraphic model of the underground, the top surface of the AltLim HU and the thicknesses maps of the other identified HUs were first obtained by interpolation of the available well data granted by Ente Irrigazione (red dots in Figure 5.5), by using the kriging method. Then, the lithology map shown in Figure 5.6 was consulted to verify that the interpolation did not provide inconsistent values, namely negative or null thicknesses for each HU in its outcropping areas, or to constrain the thicknesses values for the missing HUs. After this check, the top surfaces of every HU were obtained from the DTM values (Figure 5.5) in the outcropping areas and by subtraction of the interpolated thicknesses from the DTM surface (Figure 5.5) elsewhere. Figures 6.1 through 6.5 show the thicknesses of the Cont&Terr, SubCl, GCalc and Gal HUs and the top surface of the AltLim HU, obtained after these processings.

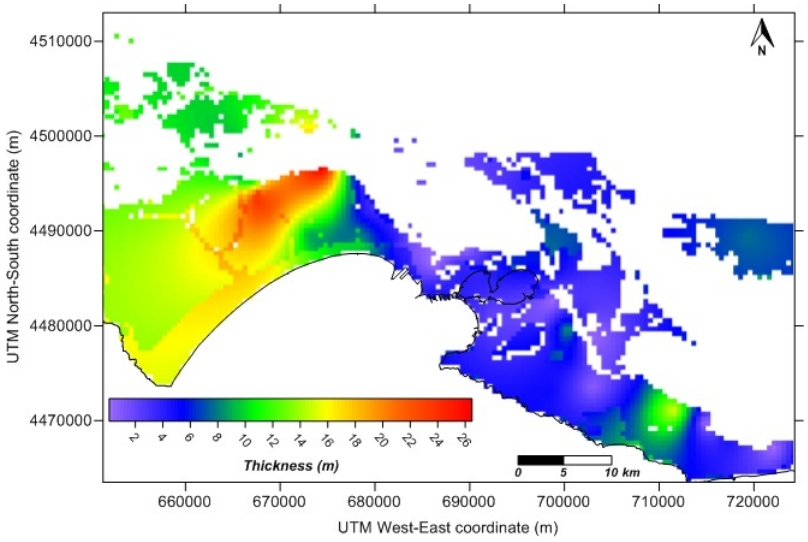


Figure 6.1: Tickness of the Cont&Terr HU.

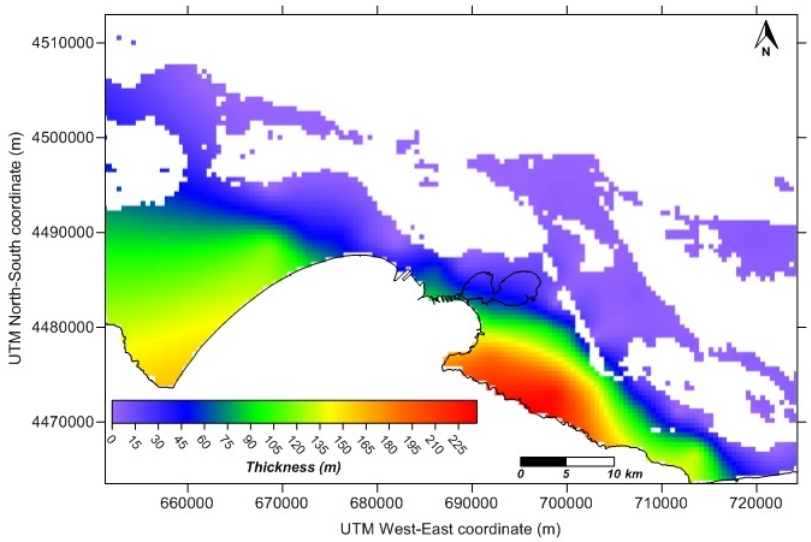


Figure 6.2: Tickness of the SubCl HU.

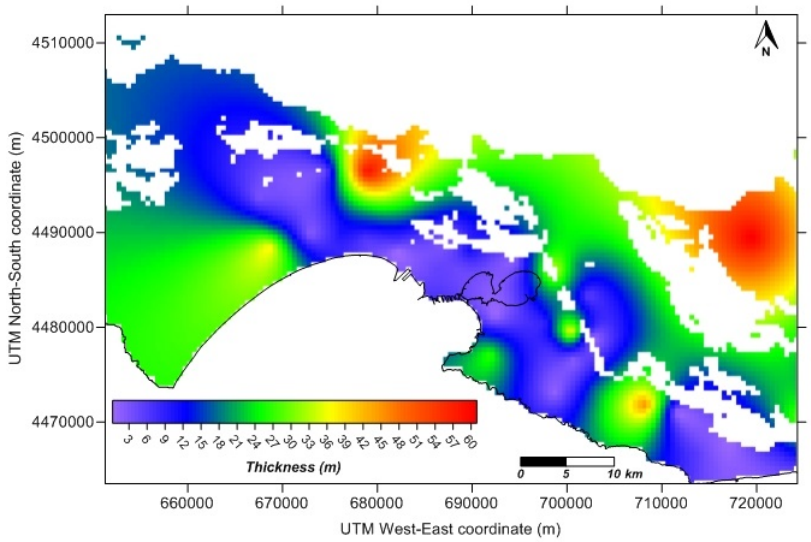


Figure 6.3: Tickness of the GCalc HU.

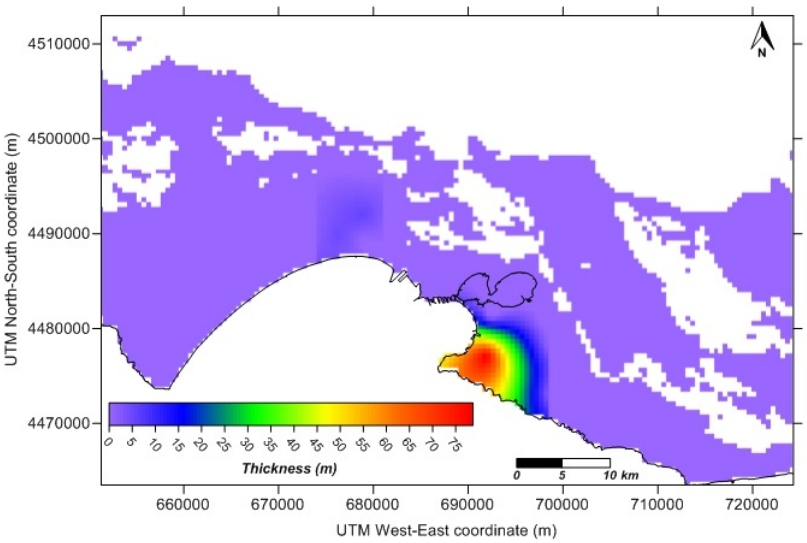


Figure 6.4: Tickness of the Gal HU.

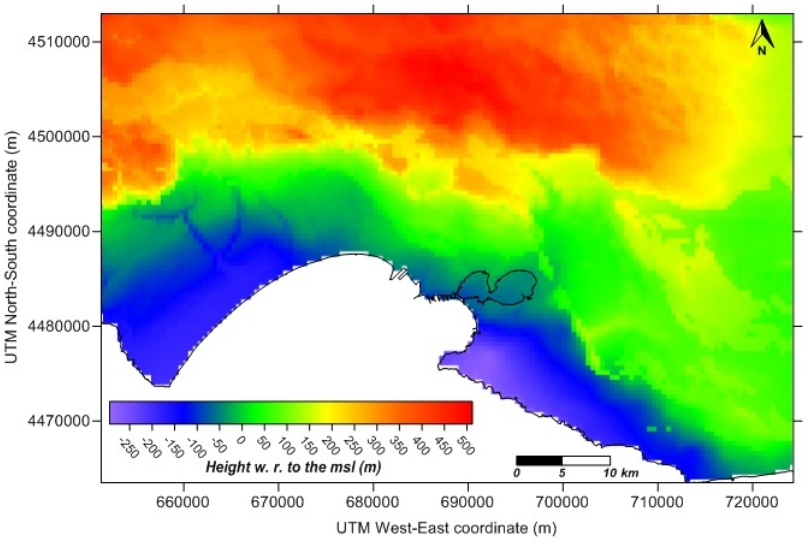


Figure 6.5: Top surface of the AltLim HU.

In order to better understand the hydrostratigraphic setup of the study area, three vertical cross-sections are sketched in Figure 6.6 (see Figure 5.6 for their location): section AB (Figure 6.6 (top)) is along the NS direction, while sections CD (Figure 6.6 (center)) and EF (Figure 6.6 (bottom)) are along the SE-NW direction, crossing the second inlet of the Mar Piccolo and the Salina Grande, respectively.
Along section AB (Figure 6.6 (top)), the GCal HU reaches a maximum thickness of about $20\ \mathrm{m}$ near the coast and is present until about $10\ \mathrm{km}$ inland. The SubCl HU, instead, is present until about $15\ \mathrm{km}$ inland, with a significant thickness (of the order of $100\ \mathrm{m}$) within the first $9\ \mathrm{km}$ from the coast. Furthermore, the Cont&Terr HU is variably present along the profile with thicknesses of the order of $10\ \mathrm{m}$.
Along the CD profile (Figure 6.6 (center)), the Gal HU is present until about $7\ \mathrm{km}$ inland, with a maximum thickness of about $50\ \mathrm{m}$ near the coast. The GCal HU is mostly present from the upper boundary of the second inlet of the Mar Piccolo until about $26\ \mathrm{km}$ far from the coast, with a thickness of about $20\ \mathrm{m}$. Again, the SubCl HU has significant thicknesses (more than $150\ \mathrm{m}$) near the coast, which decrease more and more towards inland, with values of about $10 - 50\ \mathrm{m}$ within the second inlet of the Mar Piccolo. The importance of the GCal and SubCl HUs near the coast south of the Mar Piccolo can be inferred also from section EF (Figure 6.6 (bottom)).
Generally speaking, it can be noticed quite easily that the height of the ground surface largerly varies within the study area from the mean sea level to more than $300\ \mathrm{m}$ msl inland, where the AltLim HU widely outcrops. Moreover, the presence of the impermeable SubCl HU near the coast is of paramount importance from the hydrogeological and hydrogeochemical point of views, because it lays between the deep aquifer hosted in the AltLim HU and the shallow aquifers, which are in contact with urban and industrial sources of contamination. Therefore, the SubCl HU naturally protects the most important groundwater resource of the area and the whole Salento peninsula.

6.2 Reference hydraulic head

As a further step towards the numerical characterization of the groundwater flow in the aquifer system of the study area, groundwater dynamics of the deep aquifer was qualitatively described through the map of the spatial distribution of water head (contour lines in Figure 6.7), obtained as a result of the interpolation of the available well data (dots in Figure 5.5) and official information [5, 6] by using the kriging method. Furthermore, a gaussian low-pass filter was then applied in order to avoid effects due to some irregularities during the model calibration (see subsection 4.2.5).
In Figure 6.7, blue arrows are perpendicular to contour lines and allow inferring the general dynamics of the groundwater within the AltLim HU: the deep aquifer would be mostly fed by inflow from the Murge plateau and discharges towards the sea through the Ionian coast. Generally speaking, the flow lines trend shows a good agreement with that of the topography surface (see Figure 5.5). Hereinafter, Figure 6.7 will be referred to as the *reference hydraulic head map for the deep aquifer*, which is a further outcome of the conceptual model introduced here.

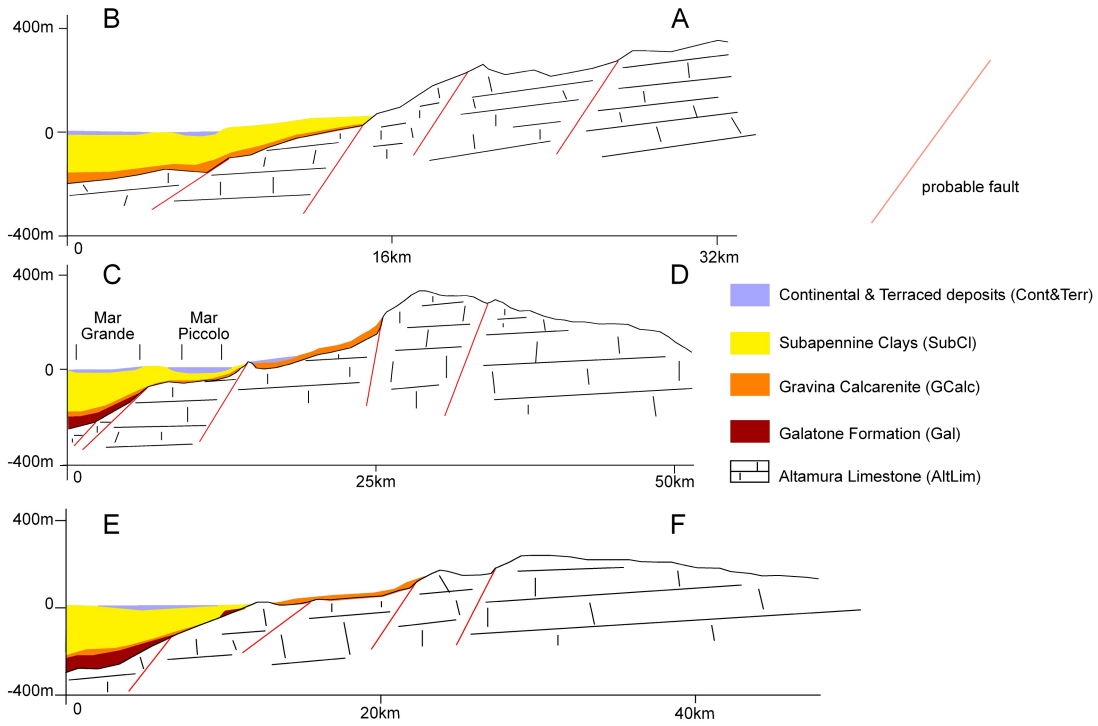


Figure 6.6: Hydrostratigraphic section AB (top); hydrostratigraphic section CD (center); hydrostratigraphic section EF (bottom); 10-fold vertical exaggeration.

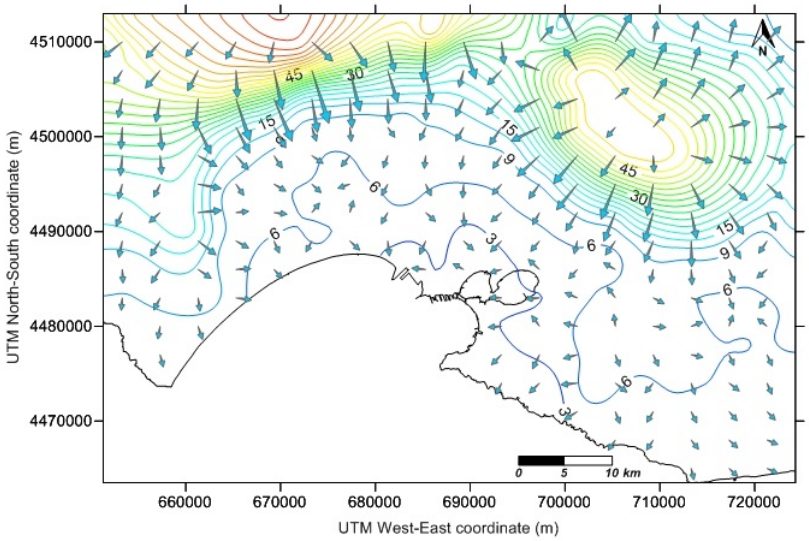


Figure 6.7: Reference hydraulic head map for the deep aquifer (values in meters). Blue arrows indicate the general trend of the flow lines.

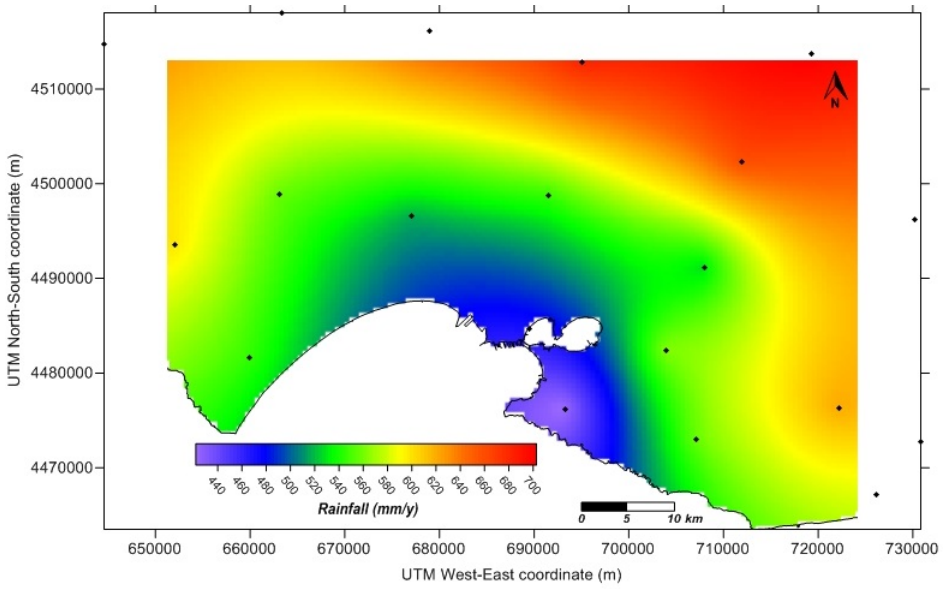


Figure 6.8: Mean annual rainfall interpolated over the studya area. Black dots indicate the locations of pluviometric stations, whose data were used for interpolation.

6.3 Terms of the water budget

The aquifer recharge is also due to rain infiltration through the stack of different HUs. In order to evaluate this source term over the study area, data of the mean annual rainfall, processed as described in section 5.4, were interpolated with the kriging method (see Figure 6.8) and multiplied by proper infiltration coefficients in order to take into account the outcropping lithological units (Figure 5.6) and the land use map (Figure 5.7): 0.4 in the areas where the AltLim, GCalc and Cont&Terr HUs outcrop, 0.1 in urban areas and where the SubCl HU outcrops. A rain infiltration of 37% was estimated with respect to the total rainfall obtained by interpolation ($50.534\ \mathrm{m}^3/\mathrm{s}$). The recharge distribution so achieved is represented in Figure 6.9. Notice that a null recharge was assigned within the Mar Piccolo.
Within the study area, withdrawals from the deep aquifer hosted in the AltLim HU were classified according to the needs for agriculture, drinking and industrial purposes. In order to estimate them, technical reports providing total extracted flow rates were consulted [119], as explained hereinafter.
Regarding withdrawals for irrigation purposes, a total extracted flow rate of $42\ \mathrm{Mm}^3/\mathrm{y}$ [119] was distributed over the study area according to the land use map (Figure 5.7) and taking into account the water needs for different crops, as already assessed in previous papers [39, 65]. The extracted flow rates assigned to satisfy the water needs of each typology of land use in each cell of the grid are listed below:

- $7.93 \cdot 10^{-7}\ \mathrm{m}^3/\mathrm{s}$ for urban areas;
- $7.93 \cdot 10^{-6}\ \mathrm{m}^3/\mathrm{s}$ for green urban areas;
- $0\ \mathrm{m}^3/\mathrm{s}$ for uncultivated, wet and wooded areas;
- $3.17 \cdot 10^{-3}\ \mathrm{m}^3/\mathrm{s}$ for vineyards;

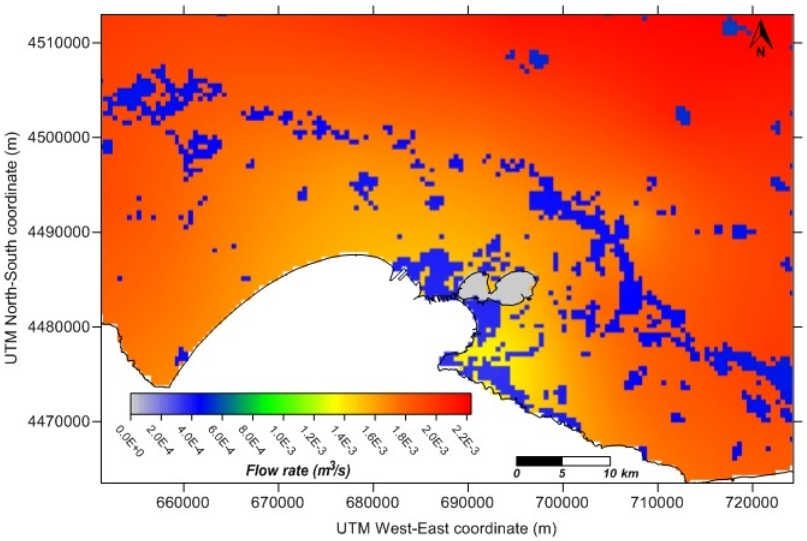


Figure 6.9: Recharge distribution due to rain infiltration and estimated for the aquifer system under exam.

- $3.96 \cdot 10^{-5}\ \mathrm{m}^3/\mathrm{s}$ for olive trees;
- $3.17 \cdot 10^{-5}\ \mathrm{m}^3/\mathrm{s}$ for other crops.

The distribution of withdrawals for agriculture purposes in the study area is shown in Figure 6.10. Notice that the western municipalities of Castellaneta, Palagiano (south-east of Castellaneta) and Palagianello (south of Castellaneta) were excluded from this estimate (null extracted flow rate) because they are fed by extra-regional sources for irrigation purposes (Consorzio di Bonifica Stornara e Tara, personal communication).
On the other hand, withdrawals from the deep aquifer for drinking purposes occur through 25 wells listed in technical reports [120]. In this case, the total extracted flow rate of $8\ \mathrm{Mm}^3/\mathrm{y}$ [119] was equally distributed among these wells (Figure 6.11), which are mostly located in the middle-western part of the domain, between Castellaneta and Massafra, while most of the drinking needs of the Province is satisfied by extra-regional sources [119].
Finally, a total extraction flow rate of $27\ \mathrm{Mm}^3/\mathrm{y}$ was indicated for industrial purposes and this estimate refers to industrial activities of the whole Province, except for the Ilva steel plant, which is fed by the Tara freshwater spring [119]. So, this total extracted flow rate was distributed among the other industrial activities according to their extension (Figure 6.12), as indicated in the official land use map [117]. The map shown in Figure 6.12 clearly shows the distinction between large and medium-small plants.

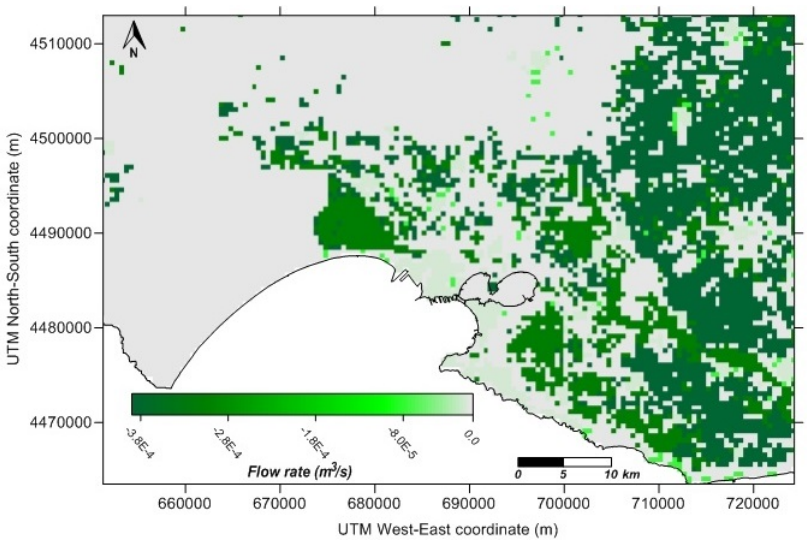


Figure 6.10: Withdrawals due to irrigation purposes and estimated for the deep aquifer of the study area.

As already mentioned in sections 5.1 and 5.3, another important factor for the groundwater balance of the aquifer system under exam is represented by subaerial and submarine freshwater springs, some of which contribute to the water balance with significant flow rates (up to some hundreds liters per second). One can distinguish between freshwater springs fed by the deep aquifer hosted in the AltLim HU and those fed by shallow aquifers hosted in the overlying porous HUs. The first group includes submarine springs located in the Mar Piccolo and Mar Grande basins and some subaerial springs mostly located near Taranto, whereas the second group includes subaerial springs located in the coastal plains NW and SE of Taranto. The flow rates of these latter springs usually decrease during summer and they sometimes are completely dry [34].

The location of freshwater springs considered for this research activity is shown in Figure 6.13 and gathered according to their mean flow rates. Most of them were calculated as an arithmetic average of the estimates found in the literature [34, 45, 105], while for the Battentieri, Galeso, Riso, Chianca and Tre Fontane springs, the measurements performed by the CNR-IRPI Bari Operative Unit within the RITMARE project [152] were considered (see Table 6.1).

Based on previous assessments about the source terms affecting the aquifer system under exam, a first estimate of the groundwater balance was achieved (see Figure 6.14). This estimate indicates a positive net balance: the recharge due to rain infiltration, in fact, would balance the extraction terms with a surplus of about $10\ \mathrm{m}^3/\mathrm{s}$. Therefore, it is expected that the groundwater flow towards the sea is quite substantial, with a flow rate per unit length of about $0.1\ \mathrm{m}^3/(\mathrm{s}\cdot\mathrm{km})$ along the coastline. Notice that leaks from sewage systems and waterworks were considered negligible and so was the infiltration during irrigation. Moreover, as for withdrawals, since the drinking needs are mainly satisfied by extra-regional sources and the biggest industrial plant (Ilva) does not affect the aquifer system [119], the most important extraction term is related to agriculture. Finally, one can infer that the freshwater springs have not negligible outgoing flow rates.

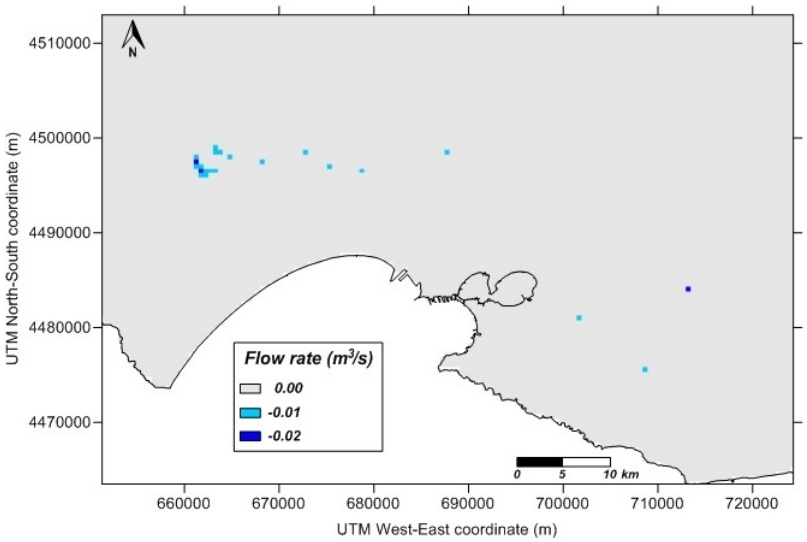


Figure 6.11: Withdrawals due to drinking purposes and estimated for the deep aquifer of the study area.

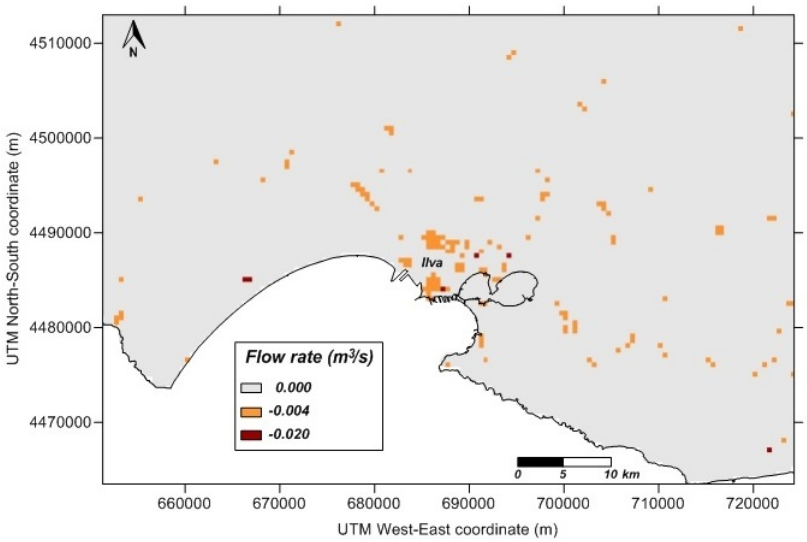


Figure 6.12: Withdrawals due to industrial purposes and estimated for the deep aquifer of the study area.

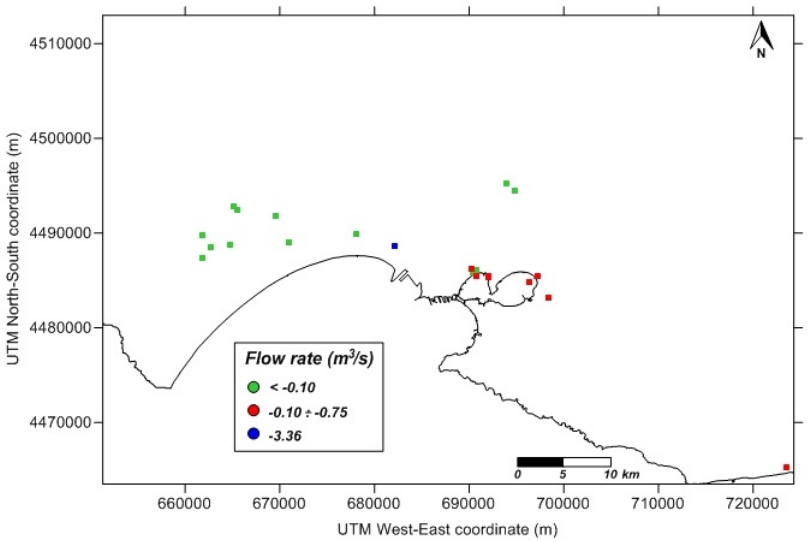


Figure 6.13: Location of freshwater springs in the study area.

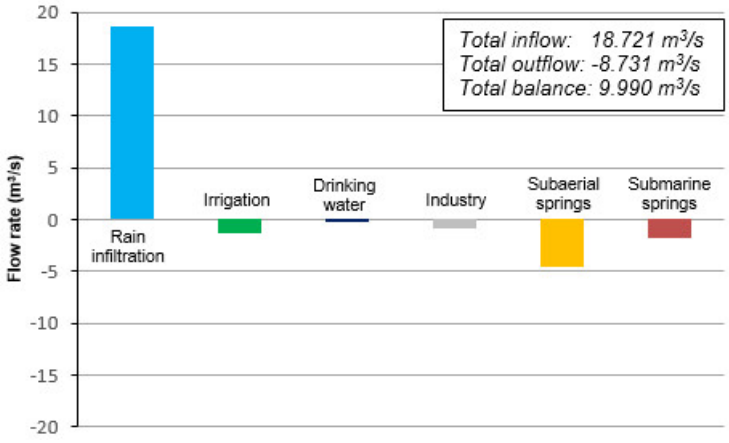


Figure 6.14: A first estimate of the groundwater balance for the aquifer system under exam, as a result of the conceptual model.

Name	x coordinate (m, UTM)	y coordinate (m, UTM)	Typology	Supplying aquifer	Mean flow rate (m^3/s)
Battentieri	697 201	4 485 423	Subaerial	Deep	-0.200
Burago	723 504	4 465 243	Subaerial	Surface	-0.214
Casamassima	661 863	4 489 765	Subaerial	Surface	-0.002
Cavallara	664 733	4 488 731	Subaerial	Surface	-0.006
Chianca	693 975	4 495 246	Subaerial	Surface	-0.002
Chiatona	670 975	4 489 060	Subaerial	Surface	-0.007
Citro	692 046	4 485 479	Submarine	Deep	-0.175
Citrello	692 080	4 485 360	Submarine	Deep	-0.175
Citro Galeso	690 784	4 485 400	Submarine	Deep	-0.750
Citro Le Copre	696 338	4 484 792	Submarine	Deep	-0.650
Fontana Canza	669 550	4 491 757	Subaerial	Surface	-0.009
Galeso	690 222	4 486 158	Subaerial	Deep	-0.500
Gravaglione	662 734	4 488 538	Subaerial	Surface	-0.001
Lavandaia	690 444	4 485 791	Subaerial	Deep	-0.030
Marangio	690 731	4 486 032	Subaerial	Deep	-0.010
Matrice	665 143	4 492 763	Subaerial	Surface	-0.015
Patemisco	678 100	4 489 879	Subaerial	Deep	-0.075
Riso	698 439	4 483 210	Subaerial	Deep	-0.100
Sant'Andrea	661 843	4 487 362	Subaerial	Surface	-0.005
Tara	682 082	4 488 695	Subaerial	Deep	-3.360
Tre Fontane	694 839	4 494 459	Subaerial	Surface	-0.001
Visciolo	665 565	4 492 444	Subaerial	Surface	-0.002

Table 6.1: Characteristics of freshwater springs shown in Figure 6.13.

Chapter 7

Numerical model for the multi-layered aquifer system

In this chapter, the application of codes and methods described in Part I to the Taranto aquifer system will be presented.
Firstly, a numerical charaterization with YAGMod and the CMM will allow to get a preliminary identification of the areas affected by the saltwater intrusion phenomenon and to highlight the most challenging difficulties about the estimation of the source terms and the effects of certain boundary conditions. As a consequence, MODFLOW, UCODE and the CMM itself will be used in order to address many of these problems and get a deeper insight into the assessment of the groundwater dynamics. Finally, the saltwater intrusion phenomenon will be simulated with SEAWAT and some likely scenarios will be presented.

7.1 Model setup

As already assessed in section 6.1, the aquifer system under exam was discretized by using a grid consisting of 147×100 (along x and y directions, respectively) square cells with $500\ \mathrm{m}$ side length in the horizontal plane. Regarding the vertical discretization, the model extends from the DTM surface (see Figure 5.5) to $500\ \mathrm{m}$ below msl and consists of 5 layers with variable thicknesses, corresponding to the HUs defined in section 6.1.
From the hydraulic point of view, isotropic conditions were assumed and the conductivities of both the Cont&Terr and GCalc HUs were set to $10^{-4}\ \mathrm{m/s}$, while those of the less permeable sediments of the SubCl and Gal HUs were set to $10^{-6}\ \mathrm{m/s}$. Instead, the hydraulic conductivity of the deepest layer (AltLim HU) was estimated by using the CMM, as explained in section 7.2.
Regarding the source terms, the areal recharge due to rain infiltration (see Figure 6.9) was applied to the uppermost active layer, according to the map of the outcropping HUs (Figure 5.6), while the three extraction terms (see Figures 6.10, 6.11 and 6.12) were located within the deepest layer. Instead, the freshwater springs were located in the uppermost (Cont&Terr HU) or in the deepest (AltLim HU) layer, according to which aquifer they are fed from (see Table 6.1), and their hydraulic conductances were set

to the value of their mean flow rates (Table 6.1), decreased by an order of magnitude.
Finally, due to the lack of information about monitoring activities of the hydraulic head over time, the numerical model was developed under steady-state conditions.

7.2 Numerical model with YAGMod

The application of YAGMod for the characterization of the coastal aquifer under exam required some modifications to equation 4.4 for computing the saturated thickness within each cell of the domain. To achieve this purpose, equation 4.4 was so modified, in order to take into account the saltwater intrusion phenomenon, as defined by Ghyben and Herzberg (see subsection 2.4.4). The saturated thickness was thus defined as:

$$\vartheta_{i,j,k} = \min(h_{i,j,k}; Top_{i,j,k}) - \max(-\delta h_{i,j,k}; Bot_{i,j,k}), \tag{7.1}$$

where the meaning of symbols is obvious from section 4.1.1 and 37 was used as a value for δ.
Notice that $\vartheta_{i,j,k} < 0$ in the following circumstances:

- whenever $h_{i,j,k} < 0$ (i.e., the hydraulic head is below the msl);
- if $Top_{i,j,k}$, $Bot_{i,j,k}$ and $h_{i,j,k}$ are above the msl, but $h_{i,j,k} < Bot_{i,j,k}$;
- if $Top_{i,j,k}$ and $Bot_{i,j,k}$ are below the msl, while $h_{i,j,k}$ is above the msl and $Top_{i,j,k} < -\delta h_{i,j,k}$ (i.e., the cell is fully saturated with saltwater).

In the first and in the last cases, a warning message is displayed and the involved cell is declared as inactive (e.g., its label is changed to *E*, see section 3.1). On the other hand, the second case is automatically handled by YAGMod by involving the corresponding dry cell into a horizontal or vertical balance (see subsection 4.2.2).

Boundary conditions (BCs)

A no-flow BC was imposed along the western and eastern boundaries of the domain, while a Dirichlet BC was used along the coastline in all layers and the northern boundary in the deepest layer. Regarding head values in constant-head cells, the contact with the sea was supposed to occur in the shallow aquifer, so that $h = 0$ m with respect to the msl was imposed along the coastline for the upper layer (Cont&Terr HU). On the other hand, in layers two to four (SubCl to Gal HUs), head values in *D* cells were set according to head initialization, i.e., the highest value between the reference hydraulic head (see Figure 6.7) and the height of the barycenter of the cell. Such an initialization was used also for internal cells in the uppermost layer: in this way, values higher than the cells' bottom were chosen all over the domain, as suggested by Valota et al. [133]. Similarly, interpolated values (see Figure 6.7) were used both to initialize the hydraulic head field and to define Dirichlet BC along the coastline and the northern boundary in the deepest layer.

Source terms

While the recharge due to rain infiltration and the withdrawals for irrigation, drinking and industrial purposes were treated as fixed and distributed source terms, the outflows from freshwater springs were simulated with equation 4.7, where we assumed: $F_1 = F_2 = 0$; $K_2 = 0$ and K_1 is equal to the average flow rate assigned to each spring decreased by an order of magnitude (see Table 6.1 and section 7.1); $H^{cal} = H^{act}$ are both set to $0.5\ \mathrm{m}$ above the bottom surface of the cell for springs fed by shallow aquifers, and to $0.5\ \mathrm{m}$ below the top surface of the cell for springs fed by the deep aquifer. A check was made, in order to avoid that the springs were dry at the beginning of the simulation (i.e., H^{act} was below the starting-head value in the corresponding cell).

Calibration with the CMM

Due to the lack of information (e.g., the reference hydraulic head) for layers one to four (Cont&Terr to Gal HUs), the CMM method (see subsection 4.2.5) was applied just to the deepest layer, so that the hydraulic conductivity field was estimated for the karstic AltLim HU. The large spatial variability of this parameter did not allow to consider a homogeneous starting conductivity field. Therefore, the hypothetical hydraulic conductivity of the deep aquifer was assumed to be represented by two regions with different values: $10^{-3}\ \mathrm{m/s}$ inland, where the hydraulic gradients are high, and $10^{-2}\ \mathrm{m/s}$ along the coastline (Figure 7.1(a)). Then, by applying the integral approach of the CMM (equation 4.12), the estimated field for the AltLim HU was obtained and filtered to smooth out some irregularities related to the ill-conditioning of the inverse problem. The final result is shown in Figure 7.1(b).
Such a conductivity field, indeed, shows a large spatial variability, especially along the coastline, with values ranging over four orders of magnitude. The lowest values (of the order of $10^{-4}\ \mathrm{m/s}$) were calculated inland and along the extremely western Ionian coast, while the highest ones (of the order of $10^{-1}\ \mathrm{m/s}$) can be observed along the coast.

Results

The color maps in Figure 7.2 show the modeled hydraulic head for the five layers, obtained from the resolution of the forward problem by using the hydraulic conductivity field shown in Figure 7.1(b) and the other input parameters listed above.
Differences between the modeled (Figure 7.2(e)) and the reference (Figure 6.7) hydraulic heads for the deep aquifer are less than $6\ \mathrm{m}$ and the highest values are found in the northern, central part of the domain, along the vertical strip between Laterza and Castellaneta, and near the coast south of San Giorgio Jonico: at this stage, the model cannot correctly simulate the high hydraulic gradients in those areas, nor the maximum of the reference hydraulic head in the north-eastern part of the domain, where the simulated head is more than $30\ \mathrm{m}$ lower than the reference one. However, since the contour lines map shown in Figure 6.7 was obtained by interpolation of information deduced from official maps, where no indication about the location of monitoring points is provided, there is no way to estimate the accuracy

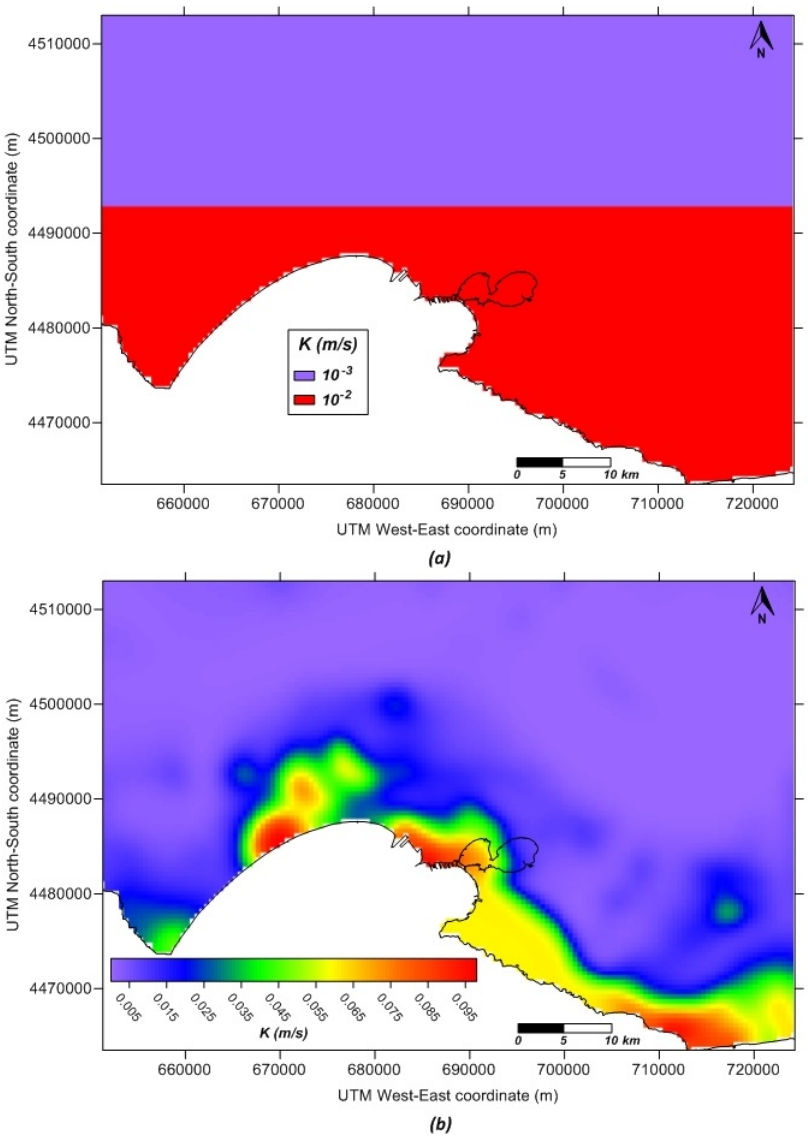


Figure 7.1: (a) Starting conductivity field for the AltLim HU as an input for the CMM; (b) estimated conductivity field for the AltLim HU after the application of the integral approach of the CMM.

of these values.

Regarding the simulation of the saltwater intrusion phenomenon, the application of the Ghyben-Herzberg approximation (see equations 2.30 and 7.1) allowed to identify the cells where this phenomenon would occur (white cells inside the domain in Figures 7.2(c)-(e)), i.e., an area extending for more than $90\ \mathrm{km}^2$ inland along the south-western coastline in the deepest three layers (GCalc, Gal and AltLim HUs) and the cell where the Citro Le Copre freshwater spring is located in the fifth layer (AltLim HU). However, the use of the Ghyben-Herzberg approach could be rather crude, so that the estimates of groundwater flow in the shallow layers and the quantification of vertical water exchanges in these areas remain rather uncertain. For this reason, at this stage, the simple application of the Ghyben-Herzberg approximation would suggest the necessity to address additional geophysical surveys in order to get a deeper insight about this issue.

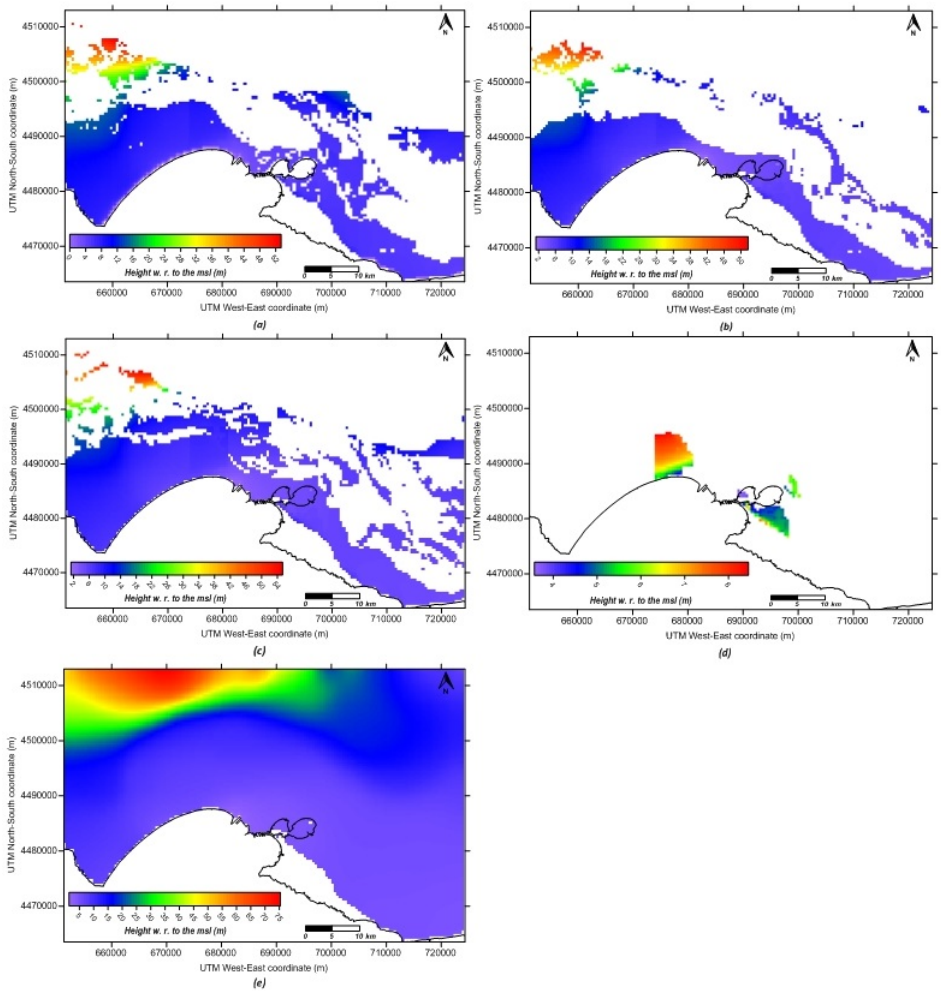


Figure 7.2: Modeled hydraulic head for the (a) Cont&Terr; (b) SubCl; (c) GCalc; (d) Gal; (e) AltLim HUs.

As a further result, Table 7.1 reports a quantitative estimation of the groundwater balance terms involved in this analysis. Regarding the source terms quantified in Table 7.1, the outflows related to the freshwater springs are underestimated by $3.77\ \mathrm{m}^3/\mathrm{s}$ with respect to the estimate obtained with the conceptual model (see Table 6.1 and Figure 6.14). Moreover, the model permits to quantify the groundwater fluxes occurring in the vertical direction and in the horizontal plane, through the boundaries of the domain, for each layer. In particular, the deep aquifer hosted in the AltLim HU would be supplied by the Murge plateau with a net inflow of about $69.89\ \mathrm{m}^3/\mathrm{s}$ and would discharge a net flow of about $71.79\ \mathrm{m}^3/\mathrm{s}$ through the Ionian coast. These outcomes are consistent with those obtained from the conceptual model from a qualitative point of view.

Finally, it is important to stress that some specific numerical precautions would be necessary because the multi-layered aquifer system is characterized by "aquitard lenses", namely interruptions or discontinuities of aquifers, which can produce numerical difficulties, analogously to what was observed by Valota et al. [133] for "aquitard windows". This is the case of some areas (western coast), where the Gal HU has non-zero thicknesses and trivial applications of standard approaches produce numerical artifacts

Layer (HU)	Rain infiltration	Extractions (industry)	Extractions (irrigation)	Extractions (drinking water)	Freshwater springs
Cont&Terr	5.82	-	-	-	$-4.92 \cdot 10^{-3}$
SubCl	0.28	-	-	-	-
GCalc	2.21	-	-	-	-
Gal	-	-	-	-	-
AltLim	10.09	-0.83	-1.31	-0.25	-2.52
Total	**18.40**	**-0.83**	**-1.31**	**-0.25**	**-2.52**

Table 7.1: Modeled balance terms for the aquifer system under exam (values in m^3/s; negative values denote outflows).

in the modeled hydraulic head of the overlying layers. However, these artifacts could be related to an improper estimate of the infiltration coefficients: the occurrence of these aquitard lenses, indeed, was found to disappear as soon as the recharge term is decreased by an order of magnitude. This suggests that the model response is highly sensitive to this term which, therefore, requires special consideration in successive calibration phases.

7.3 Calibration with joint direct and indirect methods

The numerical model developed with YAGMod acts as a solid base for further investigations about parameters uncertainty. In fact, as stated in section 7.2, it allowed to detect the most challenging weaknesses regarding both the conceptual model and the model setup.
In this framework, a new methodology is proposed to explore model adequacy by evaluating *alternative models*, i.e., models that differ conceptually and mathematically. This is done by merging direct and indirect calibration methods, performed by the CMM and UCODE, respectively, so taking advantage of their strenghts.
The aim of using alternative models is to produce more defensible modeling results, while understanding the consequences of considering different system processes and characteristics. Furthermore, the idea to perform a joint calibration arised from recognising that the hydraulic conductivity field is one of the most uncertain parameters to be estimated, especially when dealing with carbonatic aquifers. This is mostly due to a general inconsistency between the model scale and the field scale at which the measurements are done. For this reason, in most MODFLOW applications (see, e.g., Copty et al. [35], Foglia et al. [51, 53], Hyndman et al. [76], Rubin & Dagan [109]) different zones with different K values are defined based on prior information (e.g., hydrogeological maps) and then calibrated using indirect methods. Instead, the advantage of using direct inverse techniques is that a K distribution all over the domain can be obtained. Then, the idea is to scale such distribution through a multiplier, in order to condition it on head and flow observations through indirect inverse modeling.
This methodology was applied to five alternative models of the aquifer system under exam, developed to understand the role of boundary conditions imposed along the coastline on the groundwater dynamics. In fact, the selection of realistic boundary conditions is one of the most critical and difficult aspects, be-

cause they should be determined (or assumed) based on field information, which is not always available and reliable, and then expressed in mathematical forms. Therefore, to avoid a failed modeling effort, effects of different boundary conditions on model response should be critically evaluated while testing different hypotheses about systems conceptualization and the involved processes [55]. However, despite this topic has been widely treated in a general framework (e.g., in Bear [14] and Rushton & Redshaw [110]), and from the standpoint of groundwater flow modeling (e.g., in Das & Datta [38], Franke & Reilly [54], Franke et al. [55], Lu et al. [79], Padilla & Cruz [97], Werber & Simmons [140]), practical applicatons cannot be found in the scientific literature.
Of course, the alternative models developed need to go through evaluation of their adequacy in properly representing real, natural systems. Here this purpose is achieved identifying model uncertainties and using the proposed methodology to condition parameter values on observed, available data. In this regard, it is worth mentioning that, whatever the efficiency of the methodology used, thorough evaluation of model adequacy and/or accuracy necessarily depends on the availability of dense and good-quality field data.

Alternative models

Five alternative models were developed by setting different kinds of BCs along the coast for layers two (SubCl HU) to five (AltLim HU).
With respect to what stated in section 7.2 about the model setup, the alternative models share an imposed-flow BC along the northern boundary, instead of a Dirichlet BC. This choice was made because, according to the model balance estimated with YAGMod (see section 7.2), the flow terms related to Dirichlet BCs dominate over the other ones, in particular over the recharge term and the outflows from freshwater springs, which are among the most uncertain terms on which the calibration is focused. For this reason, an inflow of about $1.13\ \mathrm{m}^3/\mathrm{s}$ through the northern boundary was estimated, by multiplying the mean annual rainfall (about $600\ \mathrm{mm/y}$) and the estimated recharge area identified by using the topographic map [117] and measurements of water levels under static conditions [120] beyond the northern boundary of the study area. This total inflow from the Murge plateau was equally distributed among 75 cells in the central part of the northern boundary, where the inflow was estimated to occur, and applied to the layer corresponding to the AltLim HU, which outcrops in that zone. No flow was imposed at the remaining cells of the same border.
Furthermore, all the alternative models, share Dirichlet BCs along the coastline in the uppermost layer, with the hydraulic head equal to $0\ \mathrm{m}$ msl (so the contact with the sea occurs in this layer), except for some cells where the bottom surface is above the msl and, in that case, no flow was imposed. Hereinafter, differences in BCs along the coast for layers two to five are highlighted for the five alternative models.
In *model A*, Dirichlet BCs were used along the coastline in layers two to five. Due to the lack of information about the offshore position where the contact with the sea occurs, for the hydraulic head in constant-head cells, values adopted for the YAGMod model (i.e., the highest value between the reference hydraulic head and the height of the barycenter of the cell, see section 7.2) were used for all layers.

In *model B*, the contact with the sea was assumed to occur along the coastline for all layers, so that Dirichlet BCs were set as in model A, but the imposed head values are equal to $0\ \mathrm{m}$ msl in layers two to five.

BCs along the coast in models C to E are justified by geological profiles found in literature between the Mar Grande and Mar Piccolo basins, as that shown in Figure 7.3: in this profile, the presence of a normal fault in the zone of the submarine "citri" would obstruct the groundwater flow in the deep aquifer, at least at shallow depths, because a direct contact between the AltLim HU and the poorly permeable Subappennine clays would occur. Based on that, for models C to E, no-flow BCs were assigned along the coastline.

In particular, for *model C* Dirichlet BCs were still used for the layer corresponding to the SubCl HU, as done in model B, while no flow was imposed in the underlying layers. In *model D*, no-flow BCs were set along the coastline also in layer two. Finally, *model E* is similar to model D except for a discharge of $4.8\ \mathrm{m}^3/\mathrm{s}$ imposed along the Mar Grande coastline in the deepest layer to make some "citri" to discharge towards the sea and get a better match between the measured (see section 6.3) and modeled outflows from springs (see following sections).

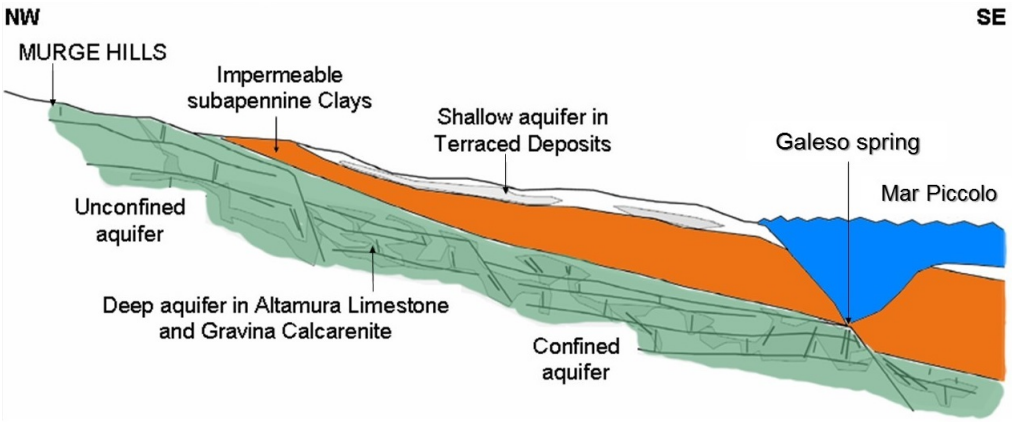


Figure 7.3: Schematic hydrogeological sketch near the Galeso spring (modified after Margiotta et al. [85]).

Direct inverse modeling with the CMM

For the aims of this calibration phase, the CMM was performed again with the same setup described in section 7.2. This means that Dirichlet BCs were still applied along the northern boundary of the domain, but such choice was not considered to be inconsistent with the remarks about the model balance obtained with YAGMod (see section 7.2), because the methodology proposed here requires a further calibration of the conductivity field with an indirect method.

What is new here, with respect to the calibration presented in section 7.2, is the computation of the saturated thickness in each cell of the domain for the deepest layer (AltLim HU). In fact, the joint calibration aims at understanding the system behaviour with respect to the groundwater flow dynamics,

so no saltwater intrusion is considered in this phase. On the other hand, in section 7.2 the estimated conductivity field shown in Figure 7.1(b) was obtained by using the Ghyben-Herzberg approximation in the computation of the saturated thickness (see equations 2.30 and 7.1). To avoid an inconsistency in this regard, the CMM was performed again with saturated thickness calculated as in equation 4.4. The resulting conductivity field is shown in Figure 7.4 and was obtained by using the same starting conductivity field shown in Figure 7.1(a).
With respect to results shown in Figure 7.1(b), the conductivity field still shows a great spatial variability (over two orders of magnitudes), but here values are smaller all over the domain, especially along the central-eastern coast where the saturated thicknesses have now higher values.

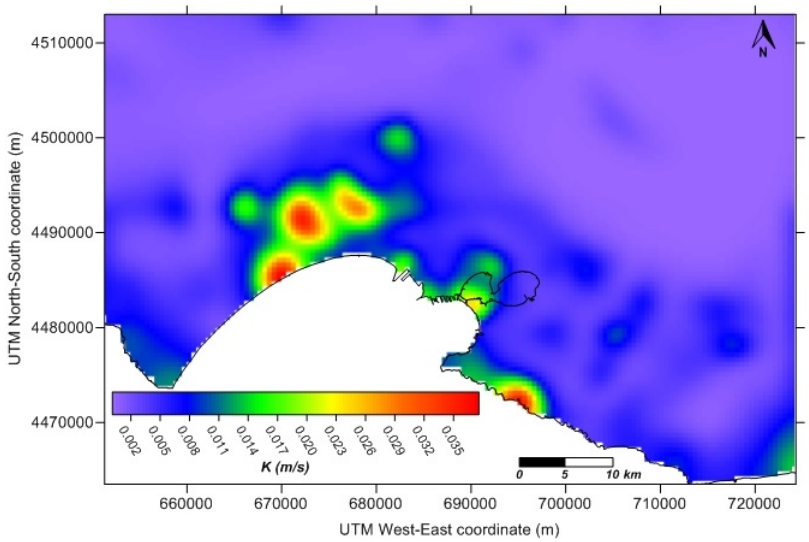


Figure 7.4: Estimated conductivity field for the AltLim HU after the application of the integral approach of the CMM and obtained ignoring the effects of the saltwater intrusion on the computation of the saturated thickness.

Indirect inverse modeling with UCODE

Before performing the inversion with the UCODE software, the five alternative models were built with the MODFLOW GUI ModelMuse. With respect to the model configuration set for YAGMod, here the new BCs mentioned above were introduced and the distribution shown in Figure 7.4 was used for the hydraulic conductivity of the AltLim HU. Furthermore, the RCH and WEL Packages were used to simulate the recharge and extraction terms, while the freshwater springs were represented as drains and simulated with the DRN Package. The ModelMuse GUI was used also to export input files for UCODE and to perform groundwater flow simulations during the updating of the parameter values.
The model parameters which were considered to be the most uncertain and for which the regression was performed are listed below:

- the hydraulic conductivity of the uppermost layer (Cont&Terr HU), K_{UP}, whose initial value was set to 10^{-4} m/s;

- the hydraulic conductivity of the third layer (GCalc HU), K_1, whose initial value was set to 10^{-4} $\mathrm{m/s}$;
- the hydraulic conductivity of the less permeable layers (SubCl and Gal HUs), K_2, whose initial value was set to 10^{-6} $\mathrm{m/s}$;
- a multiplier applied to the hydraulic conductivity distribution shown in Figure 7.4, $MultK$, whose initial value was set to 1;
- two coefficients for vertical anisotropy, for the uppermost ($VANI_{UP}$) and for the remaining layers ($VANI_{DOWN}$), whose initial values were set to 1;
- the hydraulic conductances for drains, whose initial values were set to one tenth of the corresponding measured average flow rates in $\mathrm{m^3/s}$ (see section 6.3);
- a multiplier applied to the recharge distribution due to rain infiltration, $RCHMult$, whose initial value was set to 1;
- a multiplier applied to each cell where the inflow from the Murge plateau occurs, $Mult_{IN}$, whose initial value was set to 1;
- a multiplier applied to each cell where the outflow through the coast near the Mar Grande in model E occurs, $Mult_{OUT}$, whose initial value was set to 1.

To perform the modified Gauss-Newton regression (see subsection 4.4.3) both hydraulic head and flow observations were used. In particular, the available head observations refer to well data recorded in the deep layer (dots in Figure 5.5) and used to develop the conceptual model (see chapter 6). Recalling the remark at the beginning of this section, it is worth stressing that head observations located along the coastline (red dots in Figure 5.5) are believed to be more accurate, due to a better agreement with official maps, rather than those located inland (blue dots in Figure 5.5). For this reason and because the research activity within the RITMARE project focuses on the area nearby the Mar Piccolo basin, the highest weights were assigned to head observations along the coastline during the calibration phase, so that they counted more in the computation of the objective function. In this regard, a standard deviation of 0.1 m was used for head observations located along the coastline, while for head values inland we somehow took into account the mismatch between the observed value and the reference one (conceptual model, see Figure 6.7) at the same location. In particular, we assigned the standard deviation as one half of the absolute difference between measured and reference heads. Anyway, for those points where the absolute difference was less than 0.2 m, the standard deviation was assigned as the absolute difference between measured and reference head, so that it was always greater than 0.1 m, which is the value we assigned to observations located along the coastline. This choice is equivalent to assume that the differences between the observed and the reference heads correspond to the 95% confidence interval of a normal distribution with zero mean, when they are greater than 0.2 m, and to the 68% confidence interval, when they are lower than 0.2 m. On the other hand, average flow rates of freshwater springs

were used as flow observations (Table 6.1) and we used standard deviation provided by Zuffianò et al. [152] for measured values and 50% of the flow rate for values taken from literature. Then, we calculated weights for residuals as the inverse of standard deviation.
Composite Scaled Sensitivity values (see subsection 4.4.2) were used to discriminate the most sensitive parameters. In this regard, we expect that variations of hydraulic properties of all layers will affect the hydraulic heads, according to Darcy's law 2.5, and consequently outflows from drains, according to equation 4.5. But, since the available observations are mostly located in the deepest layer, the parameter $MultK$ is expected to be more sentitive than K_{UP}, K_1, K_2, $VANI_{UP}$ and $VANI_{DOWN}$, whose variations would affect heads and outflows in the AltLim HU indirectly, through variations in vertical exchanges. Similar comments arise when dealing with hydraulic conductances of freshwater springs and parameters $RCHMult$, $Mult_{IN}$ and $Mult_{OUT}$. By taking into account hydraulic conductances of drains, we expect that heads and outflows will be mostly affected by variations of conductances of the major springs, mostly fed by the deep aquifer. Furthermore, since the AltLim HU widely outcrops inland, where many head observations are located, parameters $RCHMult$ and $Mult_{IN}$ are expected to be very sensitive. Analogous remarks can be made for $Mult_{OUT}$ in model E, due to the presence of head and flow observations near the Mar Grande.
Based on css values, indeed, parameters $MultK$, $RCHMult$ and $Mult_{IN}$ (and $Mult_{OUT}$ for model E) are the most sensitive ones, while the hydraulic conductances of some springs fed by the shallow aquifer, often dry over the year, were not included in the regression because their css values are more than two orders of magnitude lower than the highest ones.
Furthermore, Parameter Correlation Coefficients were analysed to assess correlations among the parameters listed above. Extreme correlations (i.e., high pcc in absolute value) were found to occur among $MultK$, $RCHMult$ and $Mult_{IN}$ (and $Mult_{OUT}$ for model E). In these cases, it was necessary to break such correlations by fixing some of these parameters.
Furthermore, log-transformation was used for most of the parameters as those related to hydraulic conductivities of layers and hydraulic conductances of drains, which may vary over several orders of magnitude. Moreover, sensitivities were calculated using the forward-difference method with a 5% perturbation of parameter values, also during regression, to reduce computation time. Finally, regarding the convergence criteria used to stop regression, an amount of 0.01 was defined as a maximum change of parameter values and objective function through consecutive iterations.
Hereinafter, the most important outcomes of calibration for models A to E are shown and compared.

Results and comparison

Table 7.2 reports the optimal parameter values estimated for all the alternative models. As explained before, parameters with css lower than the highest one by more than two orders of magnitudes were not included in the regression (insensitive parameters in Table 7.2) and their reference values were used to run the models. Furthermore, values of some of the most sensitive parameters were fixed to break correlations, while less sensitive K_{UP} and K_1 were usually modified looking for an optimal combination

Parameter	Reference value	Calibrated value (model A)	Calibrated value (model B)	Calibrated value (model C)	Calibrated value (model D)	Calibrated value (model E)
K_{UP}	10^{-4}	10^{-2}	Insensitive	10^{-2}	10^{-2}	10^{-3}
K_1	10^{-4}	Insensitive	Insensitive	Insensitive	10^{-5}	10^{-5}
K_2	10^{-6}	Insensitive	Insensitive	$7 \cdot 10^{-7}$	$2 \cdot 10^{-6}$	10^{-7}
$MultK$	1.00	Fixed	Fixed	1.01	1.72	0.88
$VANI_{UP}$	1.00	Insensitive	Insensitive	Insensitive	Insensitive	Insensitive
$VANI_{DOWN}$	1.00	Insensitive	Insensitive	Fixed	0.47	Fixed
$C_{Battentieri}$	0.0200	0.0063	0.0065	0.0112	0.0067	0.0052
C_{Galeso}	0.0530	0.0211	0.0225	0.0340	0.0287	0.0258
C_{Riso}	0.0100	0.0046	0.0050	0.0056	0.0045	0.0144
$C_{Marangio}$	0.0010	0.0004	0.0005	0.0004	0.0004	0.0003
$C_{CitroGaleso}$	0.0750	0.0183	0.0184	0.0668	Fixed	0.0529
C_{Burago}	0.0214	Dry	Dry	Dry	Dry	Insensitive
C_{Tara}	0.3360	0.4724	Fixed	Fixed	Fixed	Fixed
$C_{Patemisco}$	0.0075	0.0018	0.0020	0.0012	0.0014	0.0006
$C_{FontanaCanza}$	0.0009	Insensitive	Dry	Insensitive	Insensitive	Insensitive
$C_{CitroLeCopre}$	0.0650	0.0122	0.0114	Fixed	0.0628	0.0162
$C_{CitroCitrello}$	0.0350	0.0157	0.0168	0.0170	Fixed	0.0129
$RCHMult$	1.00	Fixed	Fixed	Fixed	Fixed	0.15
$Mult_{IN}$	1.00	Fixed	Fixed	12.46	15.24	14.69
$Mult_{OUT}$	1.00	-	-	-	-	Fixed

Table 7.2: Comparison between reference and calibrated values for each parameter defined in models A to E. Values of K_{UP}, K_1 and K_2 are in $\mathrm{m/s}$; values of hydraulic conductances of drains are in $\mathrm{m^2/s}$.

to avoid flood in the uppermost layer without worsen the goodness of the model fit, as measured by the objective function.

To comment values reported in Table 7.2, it is worth analysing models fits with respect to both head and flow observations.

Color maps in Figure 7.5 refer to the hydraulic head field simulated for the deep layer (to which head observations are referred) in models A to E. In each sub-figure, circles size is proportional to residuals on head observations (e_i terms in equation 4.20). As mentioned before, the highest residuals on head observations are found inland (largest circles in Figure 7.5), due to the lack of good-quality data.

In Table 7.3, simulated flow rates for active drains (i.e., those which do not dry during the simulation) are compared to the average measured values (see Table 6.1). From this Table we can infer that flow observations are overall underestimated in models A and B, with constant-head BCs along the coast, and overestimated in models C to E, with imposed-flow BCs along the coast. As shown in Figure 7.5, this is clearly related to an increase of the hydraulic head along the coastline when setting no-flow BCs (models C to E). As shown in Figure 7.3, indeed, such BCs would interrupt outflow towards the sea, thus increasing vertical exchanges. On the other hand, specific BCs in model B cause hydraulic head to decrease, so that total outflow from drains is lower than what estimated with the conceptual model.

The discrepancy in simulated outflows from drains, arising between the two different model settings (i.e., with Dirichlet BCs and with no-flow BCs), is mostly due to the Citro Galeso and Citro Le Copre submarine springs (located within the Mar Piccolo), whose simulated flow rates are one order of magnitude higher than the observed ones, especially in models C and D. Hence the decision of imposing a

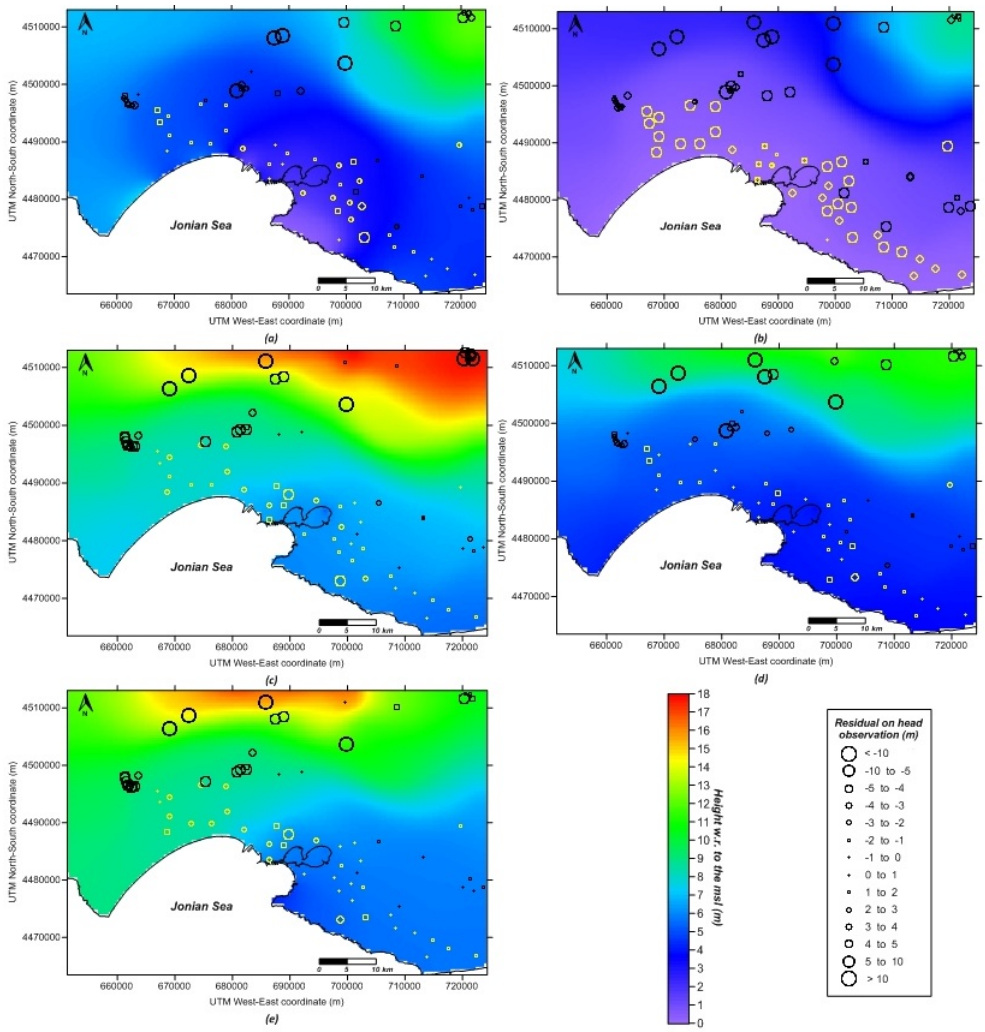


Figure 7.5: Modeled hydraulic head for the deepest layer (color maps) and residuals on head observations (circles) for models (a) A, (b) B, (c) C, (d) D and (e) E. Black circles refer to data provided by SOGESID S.P.A. & Regione Puglia [120] (blue dots in Figure 5.5); yellow circles refer to data kindly granted by Ente Irrigazione (red dots in Figure 5.5).

discharge of $4.8\ \mathrm{m}^3/\mathrm{s}$ through the Mar Grande coastline in the deepest layer: in this way, the flow rate of the Citro Le Copre drain is significantly reduced, while the same conclusion is not valid for the Citro Galeso spring, whose calibrated conductance, however, is higher than what estimated in models A and B (see Table 7.2).

Table 7.4 reports goodness of fit, as expressed by the objective function $S(\mathbf{b})$. Here, both components of $S(\mathbf{b})$ referred to head and flow observations are explicited and we can notice that head component plays a major role in the calculation of total $S(\mathbf{b})$. This can probably be an indicator of possible inconsistencies between head and flow observations, i.e., head and flow datasets may refer to different periods over the year, so that a good match of flow observations does not necessarily causes a good match of head observations. The same conclusion can probably be derived for the two head datasets (blue and red

Drain name	Measured flow rate	Simulated flow rate (model A)	Simulated flow rate (model B)	Simulated flow rate (model C)	Simulated flow rate (model D)	Simulated flow rate (model E)
Battentieri	-0.200	-0.195	-0.188	-0.389	-0.220	-0.180
Galeso	-0.530	-0.518	-0.506	-0.980	-0.762	-0.734
Riso	-0.100	-0.098	-0.094	-0.141	-0.102	-0.348
Marangio	-0.010	-0.010	-0.010	-0.012	-0.010	-0.008
Citro Galeso	-0.750	-0.730	-0.697	-2.942	-3.138	-2.310
Burago	-0.214	Dry	Dry	Dry	Dry	-0.002
Tara	-3.360	-2.858	-1.185	-3.531	-2.696	-3.598
Patemisco	-0.075	-0.074	-0.072	-0.052	-0.058	-0.027
Fontana Canza	-0.009	-0.002	Dry	-0.004	-0.002	-0.005
Citro Le Copre	-0.650	-0.604	-0.540	-3.414	-3.185	-0.854
Citro Citrello	-0.350	-0.340	-0.328	-0.440	-0.827	-0.329
Total	**-6.248**	**-5.428**	**-3.619**	**-11.905**	**-11.001**	**-8.396**

Table 7.3: Simulated flow rates for active (i.e., not dry) drains in models A to E and comparison with the average measured values (values in m^3/s).

	model A	model B	model C	model D	model E
$S(\mathbf{b})$ **- head obs.**	15 066.6	88 430.5	23 122.8	9 360.9	15 259.9
$S(\mathbf{b})$ **- flow obs.**	2.8	2.4	184.6	132.9	154.9
Total $S(\mathbf{b})$	15 069.4	88 432.9	23 307.4	9 493.8	15 414.8

Table 7.4: Values of the objective function ($S(\mathbf{b})$, dimensionless), for models A to E, referred to head and flow observations. Total $S(\mathbf{b})$ is simply the sum of its components.

dots in Figure 5.5) to justify discrepancies in residuals (see Figure 7.5). As a more obvious remark, head and flow components of $S(\mathbf{b})$ fit worse or better in different alternative models, reflecting what shown in Figure 7.5 and reported in Table 7.3. In this regard, a comparison among the models is provided hereinafter.

With reference to Figure 7.5(a) and Table 7.3, outcomes from model A show an overall discrete agreement with head and flow observations. Hydraulic head residuals along the coast and nearby the Mar Piccolo, which is the target area of the RITMARE project, are generally lower than $4\ \mathrm{m}$ (absolute values). With BCs set in model A, also flow observations are underestimated, with a total extracted flow rate of about $0.8\ \mathrm{m}^3/\mathrm{s}$ lower than the measured one (see Table 7.3). In this case, the value of the objective function, calculated as the sum of weighted squared residuals, is about $15\,069$ (see Table 7.4). Notice that, in principle, weights on both head and flow observations were calculated as the inverse of the standard deviations on measurements [72]. Then, weights were modified for observations with the highest residuals (largest circles in Figure 7.5), so that they were less important in the calculation of the objective function, and this was done in the same way for all models so that the objective function values could be used to compare goodness of models fits. Finally, from color map in Figure 7.5(a), we can infer that some inflow occurs through the coast, especially at the western side, where high hydraulic

Flow term (m^3/s)	model A	model B	model C	model D	model E
Inflow from the coast	39.5	-	-	-	-
Inflow from recharge	18.6	18.6	18.6	18.6	2.9
Inflow from the Murge plateau	1.1	1.1	14.0	17.1	14.7
Outflow from the coast	-51.4	-13.7	-18.3	-22.3	-1.9
Outflow from wells	-2.4	-2.4	-2.4	-2.4	-2.4
Outflow from drains	-5.4	-3.6	-11.9	-11.0	-8.4
Outflow from the Mar Grande	-	-	-	-	-4.8

Table 7.5: Alternative models balance.

gradients were obtained by interpolation (see contour lines in Figure 6.7), based on information from official maps. Furthermore, the simulated aquifer water budget indicates that inflow and outflow through the coast dominate over the remaining flow terms (e.g., inflow and outflow through the coast are more than two times higher than inflow due to rain infiltration, see Table 7.5).

In model B (see Figure 7.5(b)), no inflow through the coast occurs, because null values were imposed in constant-head cells. For this reason, outflow through the coast is mostly balanced by the recharge term, which is higher with respect to the flow imposed through the northern boundary (the difference between them is of the order of $10\ \mathrm{m}^3/\mathrm{s}$, see Table 7.5). Among all developed models, model B fits worse for head observations. In fact, imposing that specific BC along the coast makes the head values to be much lower than in other models, so residuals are the highest over the whole domain. For this reason, also outflows from drains are underestimated more than in model A and the residual on total simulated outflow is about $2.6\ \mathrm{m}^3/\mathrm{s}$ (see Table 7.3). The objective function has a value of about $88\,433$ (see Table 7.4).

In model C (see Figure 7.5(c)), the no-flow BC along the coast makes head values to increase with respect to models A and B, so that the simulated head field fits generally better along the coastline (with residuals lower than $3\ \mathrm{m}$ in absolute value), but worse within the industrial and military area and some inland locations (in the western-central part of the domain and in the upper-right corner). On the other hand, the rising of the hydraulic head makes the total outflow from drains to be much higher than the measured values (the difference is about $5.7\ \mathrm{m}^3/\mathrm{s}$ and the springs with the highest flows are the Citro Galeso and Citro Le Copre, as mentioned before, see Table 7.3). The rising of the hydraulic head along the coastline and the necessity to fix the recharge term for correlation issues, make the model to require more inflow from the Murge plateau to balance the increased outflow from drains ($Mult_{IN}$ increased by one order of magnitude with respect to initial estimate, see Table 7.5). In this case, the objective function has a value of about $23\,307$ (see Table 7.4).

The same remarks made for model C apply also to model D (see Figure 7.5(d)) which, however, fits better with residuals on head observations generally lower than $2\ \mathrm{m}$ (in absolute value) along the coastline, while the total outflow from drains is still overestimated by about $4.8\ \mathrm{m}^3/\mathrm{s}$ (the calibrated conductances are comparable in models C and D - see Table 7.3). The overall goodness of model D is confirmed by the objective function, which has a value of about $9\,494$ (see Table 7.4).

Finally, similar conclusions are valid also for model E (see Figure 7.5(e)), which shows similar issues

related to residuals on head observations, as stated in model C (i.e., a worse fit within the industrial and military area and in the western part of the domain). On the contrary, simulated flow rates for drains fit much better than in models C and D: by imposing an outflow of $4.8\ \mathrm{m^3/s}$ through the Mar Grande coastline, the total extracted flow rate is overestimated by about $2.1\ \mathrm{m^3/s}$ (see Table 7.3) and the overall objective function is about $15\,415$ (see Table 7.4).

7.4 Saltwater intrusion assessment

This section takes into account the occurrence of the saltwater intrusion as a natural- and human-induced phenomenon. The numerical code SEAWAT was applied to develop a large-scale variable-density flow model for the aquifer system under exam with both technical-scientific and management-oriented aims.

Model setup

With respect to the groundwater flow models described in previous sections, all the processings involved to simulate the variable-density flow required including more cells within the active domain, i.e., those along the coastline located into the sea, where a Dirichlet BC was set to specify seawater concentration at $35\ \mathrm{g/L}$. Kriging interpolation of data presented in section 5.4 was used to get values of the hydrodynamic properties of each model layer and to define aquifer geometry at the added cells. For the vertical discretization, instead, the AltLim HU was discretized in 10 layers (hereinafter referred to as AltLim HU_1, AltLim HU_2 and so on, from top to bottom, see Figure 7.6). This choice is due to the thickness of such HU within the study area (from about 300 m along the coastline to about 1000 m inland, where it widely outcrops, see section 6.1): as saltwater intrusion particularly affects this HU and dispersion mechanism occurs at a finer scale with respect to the advection one [66], a finer vertical discretization is needed to guarantee a proper representation of this phenomenon.
Regarding the groundwater flow component, alternative models B and C were resumed to test possible effects of different flow boundary conditions on the transport process. With respect to what described in section 7.3, here the hydraulic conductivity shown in Figure 7.1 was assigned to the AltLim HU and related model layers, as it takes into account the occurrence of the saltwater intrusion in the calculus of the saturated thickness. Furthermore, conductances of drains fed by the deep aquifer were further reduced by $1/4$ to avoid hydraulic head to become negative within the Mar Piccolo basin in model B.
For the transport component, an implicit finite-difference scheme was adopted, i.e., the flow and transport equations were solved several times within the same time-step until the maximum difference in fluid density between consecutive iterations is less than a user-specified tolerance, here set to $10^{-6}\ \mathrm{kg/m^3}$.
To solve for the diffusion term, the molecular diffusion coefficient was set to $10^{-6}\ \mathrm{m^2/s}$ for each model layer, while for the dispersivity coefficients a value of $50\ \mathrm{m}$, i.e., one tenth of the spatial horizontal discretization [57], was set for the longitudinal dispersivity (α_L), and 0.1 and 0.01 were assumed as values for the ratios between α_L and the horizontal and vertical transverse dispersivities, respectively [148].
To calculate the sinks/sources term, concentration must be defined for each sinks/sources defined for the

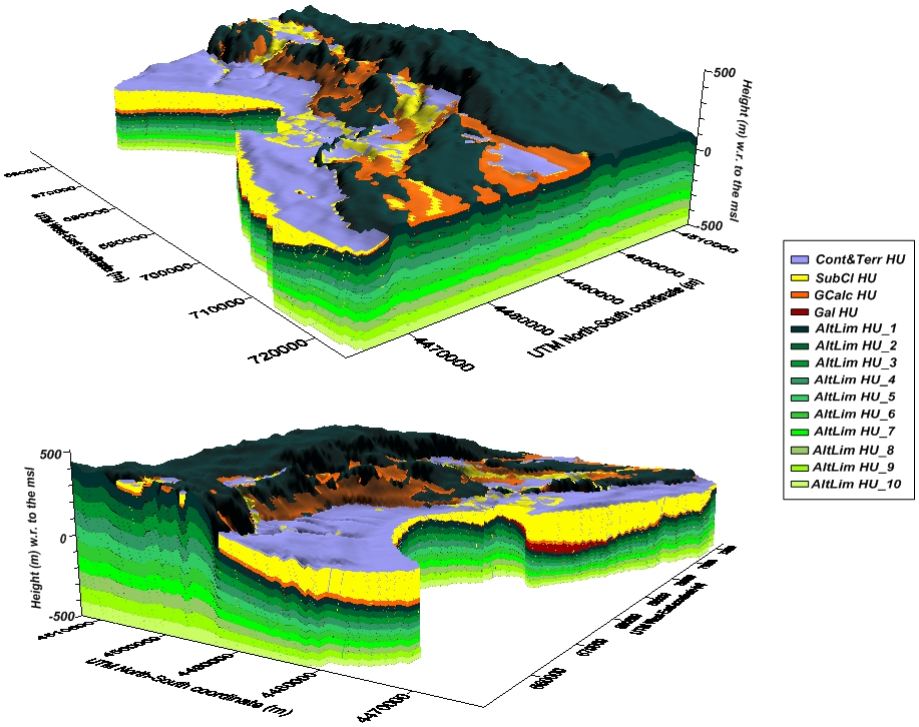


Figure 7.6: Vertical discretization of the aquifer system - (top) east-side view; (bottom) west-side view.

groundwater flow model. In this case, null concentration was specified for the recharge due to rainfall (Dirichlet BC), while in the remaining sinks/sources cells concentration could vary freely, according to what calculated through the transport equation. As already stated, Dirichlet BCs were defined along the coastline for all layers by specifying a concentration of $35\ \mathrm{g/L}$ to represent seawater.
The groundwater flow model was run over six months at steady-state using three stress periods with time-steps of two months. As mentioned in Guo & Langevin [66], there is no option to run the transport model as steady-state in SEAWAT, so the transport component is run in transient conditions, even if the flow part is specified as steady-state. This is because the transport solution has either stability constraints and/or accuracy requirements that are more restrictive than those for the flow solution, so that each time-step is further divided into smaller transport-steps, during which heads are constant.
Again, to reduce the model complexity and avoid many dry cells and numerical difficulties, the simulation was run with all layers under confined conditions.

Results

The two models developed with two different configurations of the flow component (alternative models B and C) were used to test possible effects of flow boundary conditions on the transport model solution. In this regard, no remarkable differences were found to occur in the simulated concentration field: as an example, in Figure 7.7 the $0.5\ \mathrm{g/L}$ salinity contour lines are drawn for both configurations (red line for model B and black line for model C) and for different model layers.

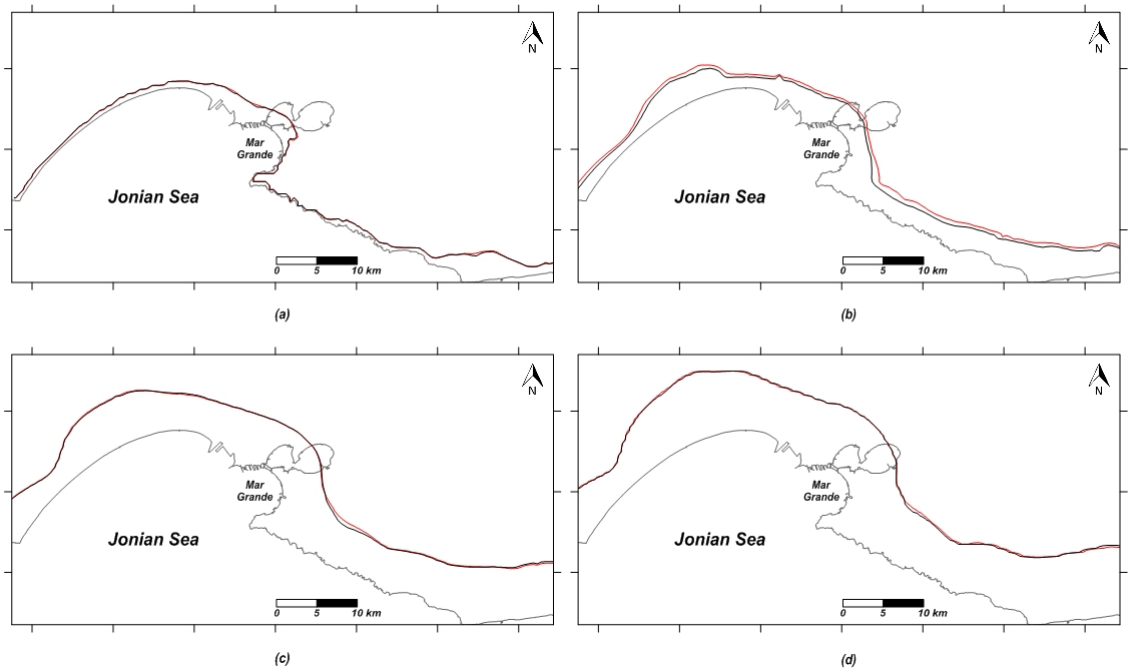


Figure 7.7: 0.5 g/L salinity contour lines simulated with configurations set in model B (red lines) and model C (black lines) and for the following model layers: (a) GCalc HU, (b) AltLim HU_1 , (c) AltLim HU_5, (d) AltLim HU_10. Zoom over the southern part of the domain.

Considering the contour line of 0.5 g/L as a reference threshold for salinity in groundwater [36], we can assess that the saltwater intrusion phenomenon would affect also porous aquifers hosted in the GCalc HU. Focusing on the deep aquifer hosted in the AltLim HU, we can infer that the spreading of intrusion gradually advances towards inland, from about 4 km from the coast in AltLim HU_1 to about 7 km from the coast in AltLim HU_10.

As further results, Figure 7.8 reports three salinity profiles along three vertical grid columns at three different locations near the coast, while Figure 7.9 shows two cross-sections, where the locations of some drains are highlighted.

The salinity profiles shown in Figure 7.8(b) are evaluated at three locations: the first one inside the industrial and military area (Location 1 in Figure 7.8(a)), the second one where the Tara subaerial spring is located (Location 2 in Figure 7.8(a)) and the third one in correspondence to the Citro Galeso spring (Location 3 in Figure 7.8(a)). According to these profiles, there are different saline levels in groundwater causing intense salinity stratification, with lower values in the upper model layers (from Cont&Terr HU to Gal HU) and concentration increasing gradually from AltLim HU_1 to AltLim HU_10. This conclusion is in accordance with previous works (see Cotecchia et al. [36] and references therein).

The cross-sections shown in Figure 7.9 can help estimate the extent of the saline plume through the model layers and with respect to the coastline, where concentration is specified. The locations of some submarine springs located within the Mar Piccolo are also highlighted to draw conclusions about the quality of the outflowing freshwater. In this regard, the salt concentration simulated by the model at

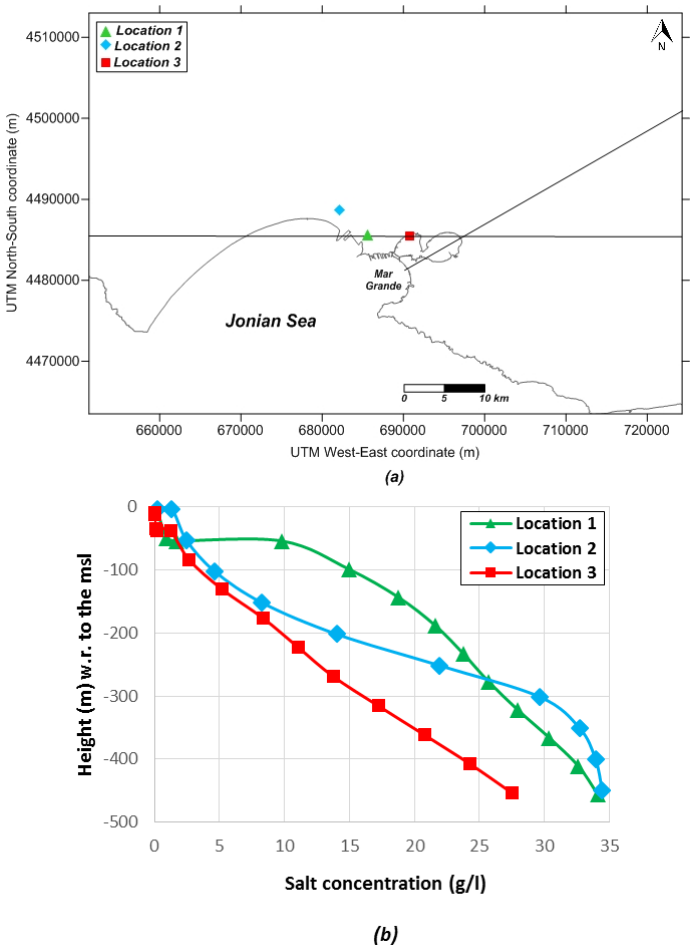


Figure 7.8: Map in (a) reports the profiles along which cross-sections shown in Figure 7.9 are drawn and locations 1, 2 and 3 where the salinity profiles shown in (b) are evaluated (values simulated in model C).

the elevation of the Citro Galeso spring is about 1.2 g/L, which is within the interval estimated by Umgiesser et al. [132]. As a further check, we calculated also the fraction of seawater at the elevation of the Citro Galeso spring itself, as defined in Zuffianò et al. [152], i.e., the relative concentration of chloride ions (mol/L) with respect to that of freshwater and saltwater. We obtained a value of 3.4%, which, considering the model scale at which the analysis presented in this thesis was carried on, is in good agreement with 5.0% calculated by Zuffianò et al. [152] based on field measurements and laboratory analysis.

Sensitivity analysis

Finally, a sensitivity analysis was performed to assess the effects of changing stresses on the concentration field. In particular, two scenarios were defined, decreasing rainfall recharge by 10% and increasing irrigation withdrawals by 25%, according to forecasts of the European Environment Agency [48] for the 21^{st} century. Hereinafter we will refer to the scenario presented in the previous paragraph by using the expression “undisturbed conditions”, while the other two scenarios will be identified as “decreased

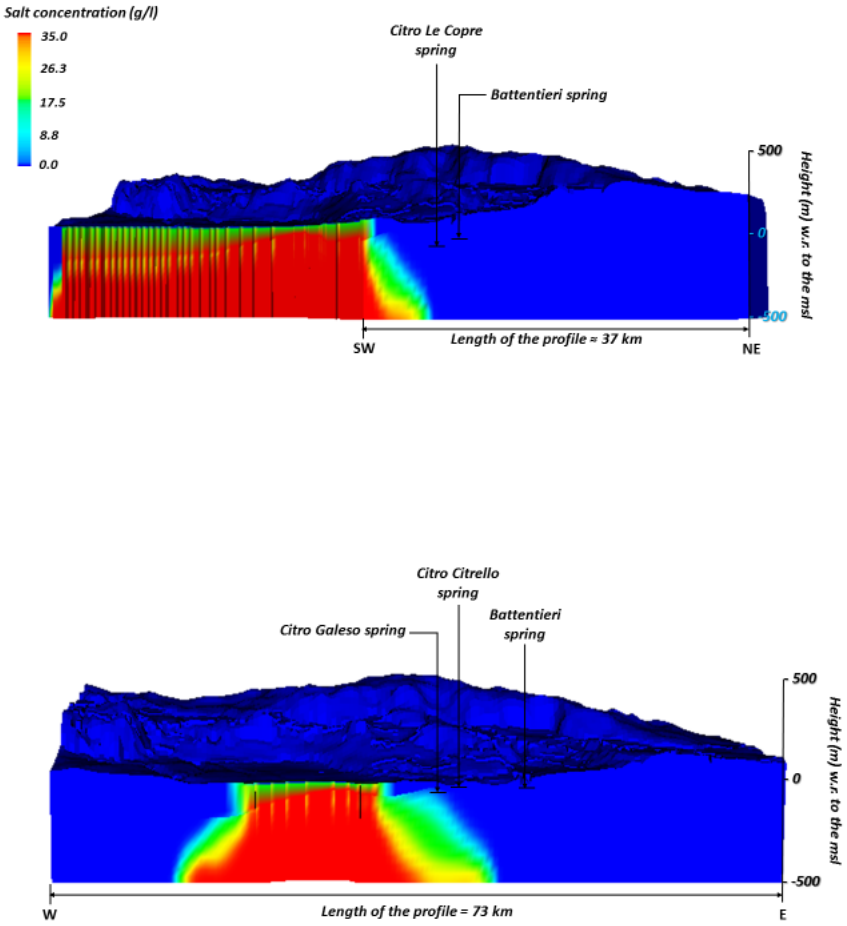


Figure 7.9: Cross-sections drawn along profiles reported in Figure 7.8(a); (top) SW-NE section; (bottom) W-E section (same color scale; 20-fold vertical exaggeration; values simulated in model C).

recharge" and "increased pumping".

Since the differences in the simulated concentration fields are difficult to show with color maps, in Figure 7.10 the contour lines corresponding to a concentration of $10\ \mathrm{g/L}$ and $30\ \mathrm{g/L}$ are drawn for AltLimHU_1 and AltLim HU_7, respectively, for both models B and C, as an example to show the effects of decreased recharge and increased pumping.

The main differences in the simulated concentration field can be observed in areas where the hydraulic conductivity of the AltLim HU (see Figure 7.1) is higher, i.e., where more recirculation can occur. Moreover, such differences are more evident for model B, due to the particular flow boundary condition which allows to further enhance this mechanism.

The salinity profiles shown in Figure 7.11 are evaluated at location 1 (see Figure 7.8(a)) and refer to the three different scenarios.

Figure 7.11 clearly shows that the effects of decreasing recharge and increasing pumping are opposite above and below a certain depth around $200\ \mathrm{m}$ below the msl. In fact, from the uppermost layer to the

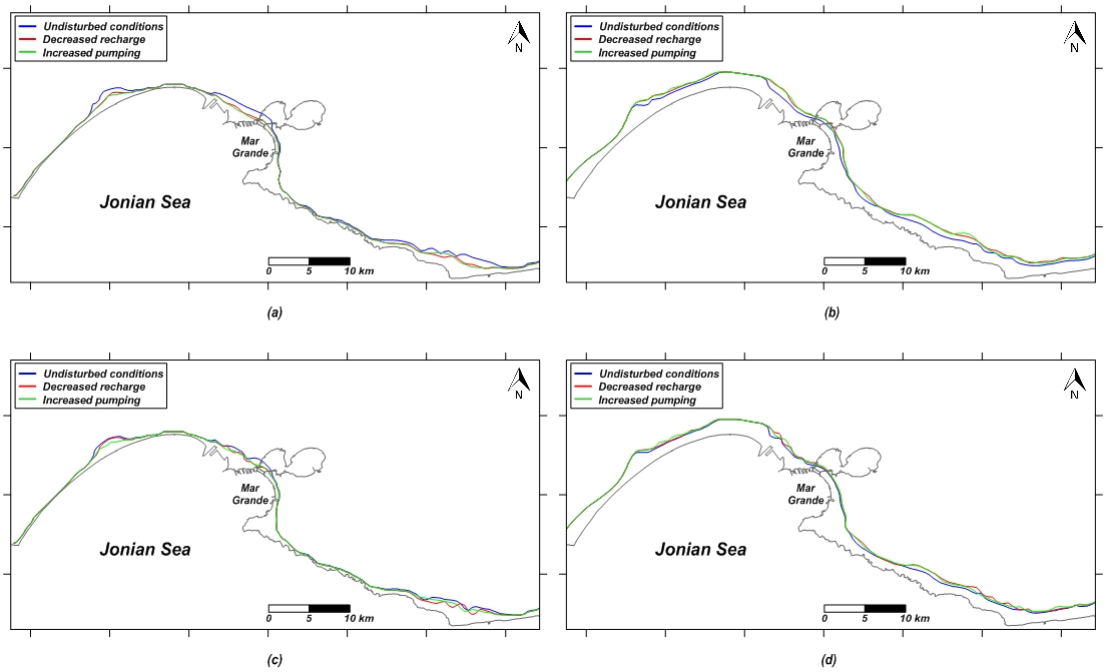


Figure 7.10: Salinity contour lines simulated for the three different scenarios: (a) $10\ \mathrm{g/L}$ for AltLim HU_1 and model B; (b) $30\ \mathrm{g/L}$ for AltLim HU_7 and model B; (c) $10\ \mathrm{g/L}$ for AltLim HU_1 and model C; (d) $30\ \mathrm{g/L}$ for AltLim HU_7 and model C. Zoom over the southern part of the domain.

AltLim HU_7 (in model B) or AltLim HU_8 (in model C), the salinity levels simulated after decreasing recharge and increasing pumping are lower than those estimated under undisturbed condition. The major differences can be found within the AltLim HU_1 layer, probably due to the presence of drains, whose flow rates are lower due to a decrease in the hydraulic head. As an example, the salt concentration calculated in model C for the two "perturbed" scenarios is about $0.9\ \mathrm{g/L}$ at the location of the Citro Galeso spring, with respect to $1.2\ \mathrm{g/L}$ simulated under undisturbed conditions. At greater depths, instead, the effects of decreasing recharge or increasing pumping cause the concentration values to increase and, as stated before, differences with respect to the undisturbed conditions are more appreciable in model B. Furthermore, the effects of decreased recharge and increased pumping are reverse according to depth and flow boundary conditions: in model B, higher concentration values are simulated with decreased recharge at shallow depths and with increased pumping in the deepest layers and vice versa for model C. This is mostly due to the effects of different boundary conditions on the increased pumping scenario: concentration values simulated with decreased recharge (see red profiles in Figure 7.11), indeed, do not show appreciable differences between models B and C, while remarkable differences arise for the increased pumping scenario. This is reasonable, as increased pumping has major effects than decreased recharge on the simulated processes in the deep aquifer (AltLim HU_1 through AltLim HU_10), considering that rainfall infiltration occurs through a thick ($\approx 100\ \mathrm{m}$) impermeable layer (SubCl HU, see Figure 6.2).

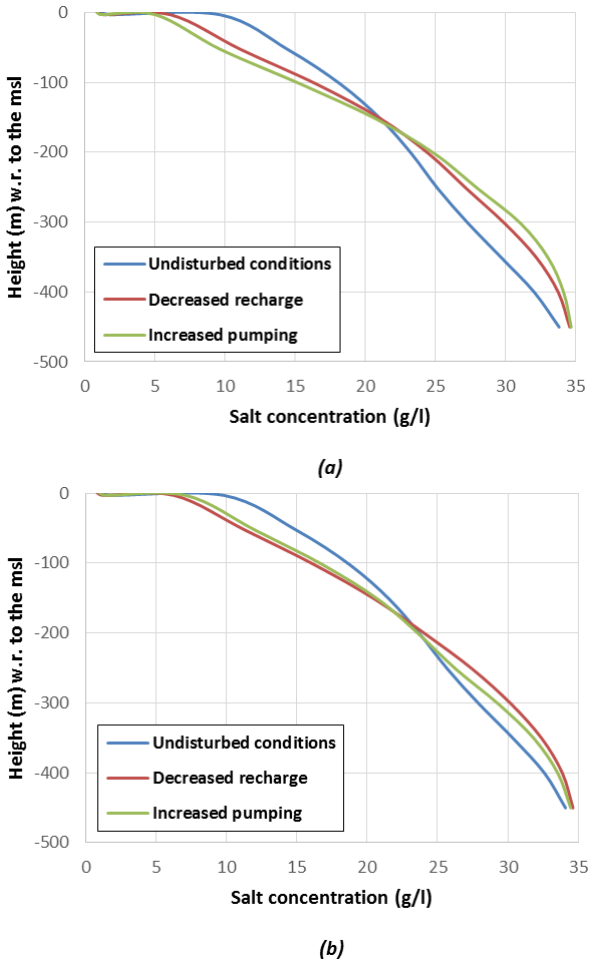


Figure 7.11: Salinity profiles along vertical grid columns at location 1 shown in Figure 7.8(a). Values simulated with (a) model B and (b) model C for the three different scenarios.

Chapter 8

Conclusions

In Mediterranean areas, groundwater stored in karst aquifers is a valuable resource in satisfying demand for several human activities. Anyway, the very fragile nature of carbonatic systems makes them sensitive to climate changes and human disturbance, especially in areas where surface water is scarce. Furthermore, very often these aquifers have hydraulic links with the sea resulting in seawater intrusion into the aquifer through karst conduits [50, 135].
Protection of carbonatic aquifers thus poses many scientific and practical challenges which require a specific approach. Anyway, comprehensive management approaches and land-use planning strategies that take into account the high vulnerability of karst are often lacking [104]. For this reason, sound understanding of the functioning of groundwater systems and their interactions with numerous and interlinked external factors is an indispensable basis for informed management. Numerical groundwater models are among the most important predictive tools used in decision support systems, as they allow to predict how the aquifer might respond to changes in pumping and climate [58]. However, the results obtained from these models sometimes cannot be considered complete or totally accurate because of the inherent uncertainty of the models themselves [145].
Within this framework, the carbonatic, coastal aquifer of the Ionian area of the Taranto gulf (southern Italy) has been taken into account for the characterization of groundwater flow and transport processes. The multi-layered aquifer system of the Taranto area proved to be an ideal case study, as it is part of an ecosystem severely affected by human impact, thus motivating the efforts of several scientific initiatives, like the RITMARE project. On the other hand, such aquifer system was particularly suitable to prove the scientific value of numerical tools and methodologies proposed to represent groundwater flow and transport, despite several difficulties arised due to the non-linearity of the involved processes, the complex geometry of the aquifer system, the lack of good-quality data.
A thorough characterization of the aquifer system under exam for management purposes required the following activities.
As preliminary steps towards the numerical characterization, the hydrostratigraphic setup and the conceptual model for the multi-layered aquifer were developed. Different datasets were integrated to get a discrete representation of the aquifer geometry and an estimation of the source and sink terms affecting the groundwater dynamics. As a result, some hydrostratigraphic sections were drawn to show the

complexity of the underground architecture and the hydraulic head map for the deep aquifer allowed to describe the overall groundwater dynamics and to detect some areas within the domain which are worth of attention, due to the occurrence of high hydraulic gradients. Then, a preliminary groundwater budget indicated a positive balance for the aquifer system, with recharge due to rainfall supplying it with a surplus of about $10\ \mathrm{m}^3/\mathrm{s}$ over the estimated sink terms.

The hydrostratigraphic and conceptual models were used as a basis to set up the numerical model, firstly with the YAGMod code for the simulation of the groundwater flow. This code was modified to introduce the Ghyben-Herzberg approximation in the calculus of the saturated thickness of the aquifer, thus obtaining an estimate of the sharp interface between freshwater and saltwater. According to what simulated, the saltwater intrusion was found to affect the deep aquifer, as well as the itermediate porous sediments of the Gravina calcarenites, along the coast south of the Mar Piccolo. A further important result is represented by the estimated conductivity field of the Altamura limestone by using the CMM to solve the inverse problem with a direct approach, so allowing to take into account the heterogeneity of the karst medium. The model developed with YAGMod played a major role in the identification of the main numerical difficulties, hence the necessarity to revise the conceptual model (in particular the estimation of the recharge term due to rain infiltration) and the model setup. The simulated budget terms showed that it was necessary to pay particular attention to the definition of boundary conditions at the northern boundary of the domain and along the coast.

Then the model was revised, mainly by varying the boundary conditions. Imposed-flow boundary conditions along the northern boundary were assigned, instead of Dirichlet boundary conditions, by detecting the potential infiltration zone and estimating the inflow from the Murge plateau thanks to head data found in literature. The model was then calibrated with respect to the most uncertain model parameters, i.e., the recharge term due to rain infiltration, the inflow from the Murge plateau, the hydraulic conductivity of the carbonatic, deep aquifer and the conductances of the major freshwater springs. Calibration involved also testing the effects of different kinds of boundary conditions along the coast: five different alternative models were set up and calibrated to check for suitability of Dirichlet or no-flow boundary conditions. Such suitability was quantified through a measure of the model fit with respect to the available head and flow data, the latter referring to measured outflows from freshwater springs.

For the aims of this calibration phase, a new methodology was proposed and adopted, which consists in merging the results of either the direct and indirect approaches: the CMM was used to evaluate the heterogeneity of the deep aquifer, by adopting the standard approach to calculate the saturated thickness of the aquifer itself. A multiplier was then used to scale such distribution of the hydraulic conductivity field and allowed to further calibrate it with an indirect approach performed with UCODE. As an overall comparison among the alternative models, with respect to head observations, computed residuals were much higher inland than along the coastline, as expected due to a higher uncertainty of the available data, and the goodness of model fit was case-dependent. On the other hand, Dirichlet boundary conditions proved to be more effective than no-flow boundary conditions in matching flow observations, as was evidenced by the objective function.

Besides the particular results obtained for each model, several numerical difficulties arised due to the

high non-linearity of the simulated process and the occurrence of extreme correlations among the most sensitive parameters (specifically, the hydraulic conductivity of the deep aquifer, the recharge term and the inflow/outflow across boundaries). Moreover, conclusions were drawn about possible inconsistency among the available data, which influenced the measure of the model fits goodness and made it difficult a proper evaluation of pros and cons of such methodology.

Saltwater intrusion was finally assessed by using SEAWAT to apply a dispersive solute-transport approach. Also for the simulation of such process, the effects of Dirichlet and no-flow boundary conditions for flow simulation were tested by using two among the alternative models developed for joint calibration. However, no remarkable differences were found to occur depending on flow boundary conditions. Saltwater intrusion was found to affect both the shallow and deep aquifers and the saline plume was evaluated in the AltLim HU at different depths and was found to extend inland up to $7\ \mathrm{km}$ from the coast at the model bottom. Intense saline stratification, described in literature [36], was further assessed through three salinity profiles and two cross-sections; particular attention was given to the Citro Galeso spring, whose salt concentration value was in accordance with measurements found in literature.

Finally, a sensitivity analysis allowed to assess the effects of decreased recharge and incresed pumping on the saline plume spreading. Major differences among the undisturbed and perturbed scenarios were found where the hydraulic conductivity of the deep aquifer is higher, so more recirculation occurs. Moreover, major differences with respect to undisturbed conditions were highlighted for the model with Dirichlet boundary conditions, which allow recirculation mechanism to be enhanced, and for the increased pumping scenario, which affects the deep aquifer much more than what decreased recharge does.

The research activities presented in this thesis can provide a valuable contribution to the scientific literature dealing with groundwater modeling and sustainable management issues. In fact, all the numerical tools proposed here proved to contribute to a sound understanding of the involved processes in a complex aquifer system which is facing several challenges and criticalities. Furthermore, the thesis sets the stage for future improvements of the calibration methodology adopted here; such an approach has been so far unexplored and its pros and cons has not yet been fully analysed.

Chapter 9

Appendix

9.1 YAGMod program structure

The general structure of the YAGMod code is arranged so that the MAIN program calls several subroutines, each of which accomplishes a particular task. The flow chart of YAGMod is drawn in Figure 9.1. First of all, the user must create the directory `input_data`, where all input files will be stored.

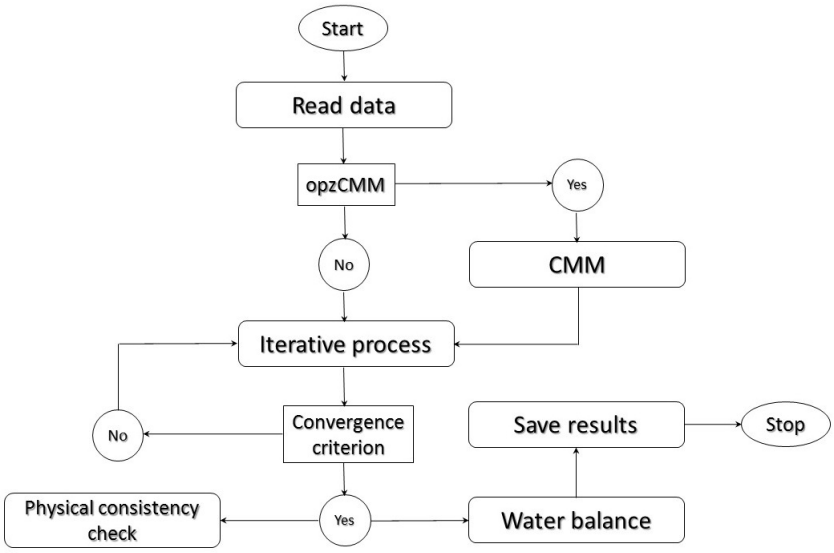


Figure 9.1: Flow chart of the general structure of the YAGMod code.

The main input file (`parameters_file.txt`), which has to be included in this directory, contains the main intructions required to characterize the aquifer system under exam, namely: discretization and geometry of the model domain, definition of internal, external and Dirichlet cells, source terms, initial distributions of hydraulic heads, hydraulic conductivities, details for solving the matrix equation 3.2 and for calibration with the CMM method, the convergence criterion and the maximum number of iterations to be performed (see Appendix 9.2 for details).
The subroutine `read_data` refers to the `parameters_file.txt` in order to read input files and

allocate array variables to store their input values.
Then, the subroutine `CMM` solving the 2D inverse problem is eventually performed, according to what is specified inside the `parameters_file.txt`. This subroutine, in turn, needs three more subroutines in order to calculate saturated thicknesses, absolute value of the hydraulic gradient and total source terms for every cell of the domain.
At this point, the iterative process is performed by the subroutine `solve`, in order to get the solution of the matrix equation 3.2 (forward problem). This subroutine calls three more subroutines to calculate saturated thicknesses, interblock transmittances and total source terms within each cell of the domain. In addition, the subroutine `SORM_iteration` performs the relaxation method at every iteration step (see subsection 4.2.4).
When the solution procedure is completed, the code starts the next procedure of checking the physical reliability of the solution reached, through the subroutine `connection`.
Regarding the output data, the model program calculates a volumetric water budget for the overall model (like MODFLOW does), with associated maximum local error inside a specific cell, in order to provide a summary of all inflows (sources) and outflows (sinks) of water to the flow system. In addition, net boundary fluxes are provided for each model layer in the vertical direction and on the horizontal plane. For furher checks, the code allows calculating net lateral fluxes across different parts of the boundary domain, as needed. For this aim, besides character codes (I, E or D) to define internal, external or Dirichlet cells, the user must distinguish also different parts of the boundary by using different integer codes. These calculations are accomplished by the subroutine `water_balance` and some of these outputs are provided also for the CMM, if performed. Moreover, additional outputs include saturated thicknesses and modeled heads in each layer after the resolution of the forward problem and some statistics about these latter (e.g., absolute average error, root mean square error and maximum absolute error, calculated with respect to the hydraulic head used to initialize the iteration procedure). After the (eventual) resolution of the inverse model, output files with the hydraulic head modeled with the CMM and the estimated hydraulic conductivity field for a specific layer are further provided.
Finally, the subroutine `save_results` allows creating output files with different extensions (e.g., .grd, .asc), in order to use several softwares for visualization and postprocessing.

9.2 Input files to YAGMod

Input files to YAGMod are read from the folder `input_data`, where all the input `.txt` files are stored, along with the file `parameters_file.txt` which contains the following information to run the model (refer to section 4.2):

- the name of the simulation to be performed;
- the name of the folder for output files;
- the number of model cells along the x, y and z directions (N_x, N_y and N_z, respectively);

- cells width (in meters) along the x and y directions (Δx and Δy, respectively);
- the maximum number of iterations to attempt the solution of the balance equation;
- the convergence parameter (ε);
- the relaxation coefficient (ω);
- the geographic coordinates (x, y) of the grid node $(1, 1, 1)$;
- the name of an external file containing N_z matrix dimensioned $N_x \times N_y$ with codes (I, E or D) to define the status of grid cells in each layer;
- a code for vertical discretization: **C** (i.e., **C**onstant) if model layers have constant thicknesses, **V** (i.e., **V**ariable) if model layers have variable thicknesses;
- if the previous code is **C**, the name of an external file is read: it contains $N_z + 2$ values, namely the heights of the topography, top surfaces of each layer and model bottom; if the previous code is **V**, the name of an external file is read: it contains $N_z + 2$ matrix dimensioned $N_x \times N_y$ with the heights of the topography, top surfaces of each layer and model bottom within each grid cell;
- a code for the presence of point variable source terms (e.g., drains): **Y** for **Y**es and **N** for **N**o;
- if the previous code is **Y**, the name of an external file is read: it contains (for each point variable source) the coordinates (x, y, z) of the source location (these are then converted in (i, j, k) inside the code) and the parameters F_1, F_2, K_1, K_2, $H^{(cal)}$ and $H^{(act)}$ used to define it;
- a code for the presence of screened wells: **Y** for **Y**es and **N** for **N**o;
- if the previous code is **Y**, the name of an external file is read: it contains (for each screened well) the coordinates (x, y, z_{min}, z_{max}) of the screened interval and the extracted flow rate (q_W);
- the number of fixed and areally distributed source terms;
- the previous number allows reading just as many lines which refer to external files: each of them contains N_z matrix dimensioned $N_x \times N_y$ with flow rates associated to the specific source term, in each layer and within each grid cell;
- the name of an external file containing N_z matrix dimensioned $N_x \times N_y$ with hydraulic conductivities of each layer within each grid cell;
- the name of an external file containing N_z matrix dimensioned $N_x \times N_y$ with the reference hydraulic head, in each layer and within each grid cell: these values are used both to initialize the iterative procedure for the solution of the forward problem and as the reference field for the CMM;
- the name of an external file containing N_z matrix dimensioned $N_x \times N_y$ with integer codes, in order to distinguish different parts of the boundary, in each layer and within each grid cell;

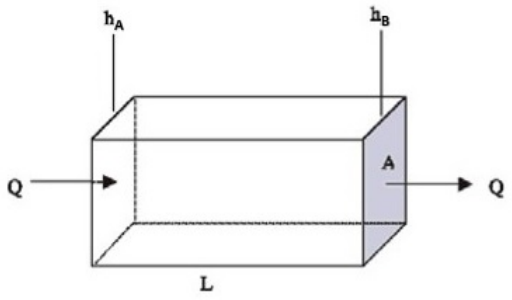


Figure 9.2: Prism of porous material illustrating Darcy's law. Modified after Harbaugh [69].

- a code for the activation of the CMM: **Y** for **Y**es and **N** for **N**o; if this code is **N**, the following items are blank;
- a code to chose the approach used to estimate the hydraulic conductivity field: **I** for **I**ntegral and **D** for **D**ifferential;
- a code for the initial estimate of the conductivity field: **F** if a constant value is used, **V** if the distribution previously read for K within each cell of the specific layer is to be used;
- if the previous code is **F**, the constant K value for the inizialization of the CMM is read;
- the maximum number of iterations to perform the CMM;
- three values for tolerance parameters (l_1 norm, l_2 norm and l_∞ norm), in order to stop CMM iterations if the difference between the reference hydraulic head and $h^{(CM)}$ field is small enough;
- a value for the coefficient c used to calculate the weight w.

9.3 Dependencies to discussion about MODFLOW

Some details about what stated in section 4.1 are provided here. For a full discussion the reader should refer to Harbaugh [69] and references therein.

Calculus of the interblock hydraulic conductance between adjacent cells

Hydraulic conductance is a combination of several parameters used in Darcy's law. Consider the formulation of Darcy's law given by equation 2.3 for a grid cell, namely a prism of porous material represented in Figure 9.2.
Darcy's law defines one-dimensional flow through the prism shown in Figure 9.2 as:

$$Q = \frac{KA}{L}(h_A - h_B),$$

where Q is the volumetric flow through the prism $[L^3T^{-1}]$, K is the hydraulic conductivity of the material in the direction of flow $[LT^{-1}]$, A is the cross-sectional area perpendicular to the flow path, L is the length

of the prism parallel to the flow path and $(h_A - h_B)$ is the head difference across the prism parallel to flow. According to the definition given in section 4.1, hydraulic conductance can be expressed as:

$$C = \frac{KA}{L}, \tag{9.1}$$

so that Darcy's law becomes:

$$Q = C(h_A - h_B). \tag{9.2}$$

Moreover, if W is the width of the prism and S its thickness, then $A = WS$ and K can be expressed in terms of *hydraulic transmissivity*, $T = KS$ $[L^2T^{-1}]$. So, hydraulic conductance can be expressed also as:

$$C = \frac{TW}{L}. \tag{9.3}$$

If a prism of porous material consists of two or more subprisms in series (aligned sequentially in the direction of flow) and the conductance of each subprism is known, a conductance representing the entire prism can be calculated. From equation 9.2, the equivalent conductance for the entire prism is the rate of flow in the prism divided by the head change across it:

$$C = \frac{Q}{h_A - h_B}. \tag{9.4}$$

Assuming continuity of head across each subprism gives the identity:

$$\sum_{i=1}^{n} \Delta h_i = h_A - h_B,$$

which, using equation 9.2 becomes:

$$\sum_{i=1}^{n} \frac{q_i}{c_i} = h_A - h_B,$$

where q_i is the volumetric flow across the i-th subprism and c_i its hydraulic conductance.
Since flow is one-dimensional and we assume no accumulation or depletion in storage, each q_i is equal to the total flow (Q), therefore

$$Q \sum_{i=1}^{n} \frac{1}{c_i} = h_A - h_B \Leftrightarrow \frac{h_A - h_B}{Q} = \sum_{i=1}^{n} \frac{1}{c_i}.$$

By comparing this latter with equation 9.4, we obtain:

$$\frac{1}{C} = \sum_{i=1}^{n} \frac{1}{c_i}, \tag{9.5}$$

namely for a set of conductances arranged in series, the inverse of the equivalent conductance equals the sum of the inverses of the individual conductances. When only two subprisms (i.e., adjacent grid cells) exist, the equivalent conductance reduces to:

$$C = \frac{c_1 c_2}{c_1 + c_2}, \tag{9.6}$$

which is nothing but the harmonic mean between individual conductances.

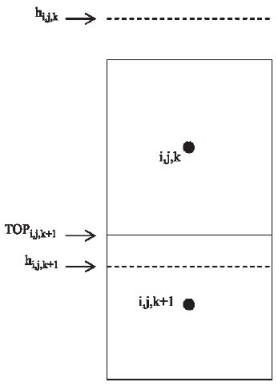


Figure 9.3: Partial desaturation of a cell located below a fully or partially saturated cell. Modified after Harbaugh [69].

Vertical exchange between a fully or partially saturated cell and an unconfined cell below

If head drops below the aquifer bottom, the cell is considered to be fully dewatered and is set to no-flow. However, the model also has provision for the re-saturation of fully dewatered cells. In this regard, the vertical flow calculation is modified if a cell is unconfined (i.e., head is below its top elevation), while the cell directly above is fully or partially saturated (Figure 9.3). This situation can occur, for example, when drawdown due to pumpage causes water level to fall, at least locally, below the top of a layer.
With reference to equation 4.2, the term that represents flow into the cell (i, j, k) through its lower face is:

$$Q_{i,j,k+1/2} = C_{i,j,k+1/2}(h_{i,j,k+1} - h_{i,j,k}). \tag{9.7}$$

Here, a positive value of $Q_{i,j,k+1/2}$ indicates flow into the cell (i, j, k), while a negative value indicates flow out of it. Equation 9.7 is based on the assumption that cells (i, j, k) and $(i, j, k + 1)$ are fully saturated.
If situation shown in Figure 9.3 occurs, in the upper cell head is simply $h_{i,j,k}$ while just below it unsaturated conditions prevail, so that the pressure sensed on the lower surface of the unsaturated zone is atmospheric (p_a). Thus, the head at the bottom of the upper cell is simply the elevation at that point, namely the elevation of the top of the lower cell ($Top_{i,j,k+1}$). Then the actual flow through the unsaturated zone is obtained by substituting $Top_{i,j,k+1}$ for $h_{i,j,k+1}$ in equation 9.7:

$$Q_{i,j,k+1/2} = C_{i,j,k+1/2}(Top_{i,j,k+1} - h_{i,j,k}). \tag{9.8}$$

Thus, the flow will be downward, from cell (i, j, k) to cell $(i, j, k + 1)$, but under this condition the flow will no longer be dependent on the water level ($h_{i,j,k+1}$) in the lower cell. The simplest approach to this problem in formulating the equation for cell (i, j, k) would be to substitute the flow expression of equation 9.8 into equation 4.2, in place of the expression given in equation 9.7. If we consider, however, the matrix of coefficients of the entire system of equations, this direct substitution would render this

matrix asymmetric, thus generating problems in the solution process. To avoid this condition, the flow term of equation 9.7 is allowed to remain on the left-hand side of equation 4.2 but, because the actual flow into cell (i, j, k) is given by equation 9.8, a correction term is introduced by subtracting equation 9.8 from equation 9.7:

$$q_c = C_{i,j,k+1/2}(h_{i,j,k+1} - Top_{i,j,k+1}).$$

This correction term should be added to the right-hand side of equation 4.2 to compensate for allowing the computed flow (equation 9.7) to remain at the left-hand side. This again introduces a problem of matrix asymmetry because q_c contains $h_{i,j,k+1}$. To circumvent this difficulty, q_c is actually computed using the value of $h_{i,j,k+1}$ from the preceeding iteration, rather than from the current one:

$$q_c^n = C_{i,j,k+1/2}(h_{i,j,k+1}^{n-1} - Top_{i,j,k+1}),$$

where the superscript is referred to iteration. As convergence is approached, the difference between $h_{i,j,k+1}^{n-1}$ and $h_{i,j,k+1}^{n}$ becomes progressively smaller, and q_c^n becomes more accurate. In equation 4.3, q_c^n is added to $RHS_{i,j,k}$. This correction applies whenever dewatering of a cell occurs below a cell that is fully or partially saturated.

Re-wetting of dry cells

When saturated thickness is zero, as defined by head being less than the bottom elevation of a cell (see last case of equation 4.4), it is clear that a variable-head (wet) cell converts to dry. The opposite (i.e., a dry cell converts to wet) occurs if the head at the adjacent cell equals or exceeds a certain threshold specified by user at the beginning of an iteration. Adjacent cell allowing such a conversion can be any of the four cells that are directly adjacent to the cell horizontally or the cell directly below. The head below can be a better wetting indicator than the head at horizontally adjacent cells when the head variations between adjacent horizontal cells are larger than the vertical head variations, which is frequently the case. The threshold value must be determined by trial-and-error: if it is too low, a cell may be incorrectly converted to wet (i.e., it may convert to wet and then reconvert to dry in later iterations because the head does not stay above the bottom elevation) preventing the solver from converging; on the other hand, larger values increase the non-uniqueness of the solution.
When a cell converts to wet, the initial estimate of head is established according to one of two following equations:

$$h_{i,j,k} = Bot_{i,j,k} + A(h_n - Bot_{i,j,k}), \tag{9.9}$$

$$h_{i,j,k} = Bot_{i,j,k} + A(thresh_{i,j,k}), \tag{9.10}$$

where $A \in [0, 1]$ is a constant specified by the user and h_n is the head at the neighboring cell that causes cell (i, j, k) to convert to wet. This initial estimate of head at a cell that converts to wet is theoretically unimportant, because the solver should calculate the correct value in subsequent iterations. In practice, however, it is often important for solver efficiency.
Another user-specified option is the solver iteration interval for attempting to wet cells.

The PCG solver

PCG method used to solve the matrix equation 3.2 supposes that the matrix **A** can be split into the sum of two matrices: $\mathbf{A} = \mathbf{M} + \mathbf{N}$, where **M** is called the *preconditioned form of A* and is a symmetric and positive-definite matrix. The goal is to define **M** so that it is easy to invert and resembles **A** as much as possible. **M** can be defined through the use of two *preconditioner*: the *modified incomplete Cholesky* (MICCG) or the *least-squares polynomial* (POLCG).
In MICCG preconditioner, $\mathbf{M} = \mathbf{U}^T\mathbf{D}\mathbf{U}$, where **U** is an upper-triangular matrix and **D** is a positive-diagonal matrix with $d_{ii} = 1/u_{ii}$. Components of matrix **U** are given by:

$$u_{ij} = a_{ij} \qquad for \qquad j > i,$$

$$u_{ii} = 0 \qquad for \qquad j < i,$$

$$u_{ii} = a_{ii} - \sum_{k=1}^{i-1} \frac{u_{ki}^2}{u_{kk}} - \alpha(\sum_{j=1}^{i-1} f_{ji} + \sum_{j=i+1}^{N} f_{ij}),$$

where the variable α is a user-defined relaxation parameter and

$$f_{ij} = \begin{cases} \sum_{k=1}^{i-1} \frac{u_{ki}^2}{u_{kk}} & \text{if } a_{ij} = 0 \\ 0 & \text{if } a_{ij} \neq 0 \end{cases}$$

In POLCG preconditioner, the inverse matrix $\mathbf{M}^{-1}$ is approximated by a truncated Neuman polynomial series: $\mathbf{M}^{-1} = c_0\mathbf{I} + c_1\mathbf{A} + c_2\mathbf{A}^2 + ... + c_n\mathbf{A}^n$, where the weighting coefficients (c_i) are chosen to optimize convergence.
Once **M** has been defined, the basic iterative equation is developed from equation 3.2 and the splitting of **A**, and can be written as:

$$\mathbf{M}\mathbf{h}^{k+1} = \mathbf{M}\mathbf{h}^k + \mathbf{b} - \mathbf{A}\mathbf{h}^k,$$

where k is the iteration index. Noting that ($\mathbf{b} - \mathbf{A}\mathbf{h}^k$) is the residual ($\mathbf{r}^k$) of the original set of equations at the k-th iteration, and setting $\mathbf{s}^k = \mathbf{h}^{k+1} - \mathbf{h}^k$ gives:

$$\mathbf{M}\mathbf{s}^k = \mathbf{r}^k \qquad or \qquad \mathbf{s}^k = \mathbf{M}^{-1}\mathbf{r}^k.$$

At first iteration ($k = 0$), $\mathbf{s}^0 = \mathbf{M}^{-1}\mathbf{r}^0 = \mathbf{M}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{h}^0)$ and the change in heads can be calculated as $\mathbf{h}^1 = \mathbf{s}^0 + \mathbf{h}^0$.
Then, for $k > 0$, the new changes in heads ($\mathbf{p}^k$) are calculated as $\mathbf{p}^k = \mathbf{s}^k + \beta^k\mathbf{p}^{k-1}$, where $\beta^k = \frac{(\mathbf{s}^k)^T\mathbf{r}^k}{(\mathbf{s}^{k-1})^T\mathbf{r}^{k-1}}$ and $\mathbf{p}^0 = \mathbf{s}^0$.
By introducing $\mathbf{p}^k$ values, we get $\mathbf{h}^{k+1} = \mathbf{h}^k + \alpha^k\mathbf{p}^k$ and $\mathbf{r}^{k+1} = \mathbf{r}^k - \alpha^k\mathbf{A}\mathbf{p}^k$, where $\alpha^k = \frac{(\mathbf{s}^k)^T\mathbf{r}^k}{(\mathbf{p}^k)^T\mathbf{A}\mathbf{p}^k}$.

References

[1] G. Alabiso, M. Cannalire, D. Ghionda, M. Milillo, G. Leone, O. Caciorgna (1997), "Particulate matter and chemical-physical conditions of an inner sea: the Mar Piccolo in Taranto. A new statistical approach", *Marine Chemistry*, 58, 373-388.

[2] M. Amodio, M. Caselli, B.E. Daresta, G. de Gennaro, P. Ielpo, C.M. Placentino, M. Tutino (2008), "Air quality impact for industrial area of Taranto city (South Italy): a miltivariate statistical analysis application", *Chemical Engeneering Transactions*, 16, 193-200.

[3] M. Amodio, E. Andriani, I. Cafagna, M. Caselli, B.E. Daresta, G. de Gennaro, A. Di Gilio, C.M. Placentino, M. Tutino (2010), "A statistical investigation about sources of PM in South Italy". *Atmospheric Research*, 98, 2-4. DOI:10.1016/j.atmosres.2010.03.023

[4] S. Anwar & M.C. Sukop (2009), "Lattice Boltzmann Models for Flow and Transport in Saturated Karst", *Ground Water*, 47/3, 401-413. DOI:10.1111/j.1745-6584.2008.00514.x

[5] APAT & Università degli Studi di Napoli "Federico II" (Dipartimento di Geofisica e Vulcanologia) (1999), "Hydrogeological Map of Southern Italy - Map II (Murge - Salento - Ionian Arc)", *Program INTERREG IIC*.

[6] Assessorato all'Assetto del Territorio della Provincia di Taranto (2008), "Piano Territoriale di Coordinamento Provinciale (PTCP) - Territorial plan for coordination of the Province", language: Italian.

[7] A. Aureli, J. Ganoulis, J. Margat (2008), "Groundwater Resources in the Mediterranean Region: Importance, Uses and Sharing", *UNESCO International Hydrological Programme (IHP)*, Paris, pp. 96-105.

[8] R. Auriemma, G. Mastronuzzi, P. Sansò (2004), "Middle to Late Holocene relative sea-level changes recorded on the coast of Apulia (Italy)", *Géomorphologie: relief, processus, environnement*, 1, 19-34.

[9] A.A. Ahmed (2009), "Using lithologic modeling techniques for aquifer characterization and groundwater flow modeling of the Sohag area, Egypt", *Hydrogeology Journal*, 17, 1189-1201. DOI:10.1007/s10040-009-0461-z

[10] Z. Bakalowicz (2015), "Karst and karst groundwater resources in the Mediterranean", *Environ Earth Sci*, 74, 5-14. DOI:10.1007/s12665-015-4239-4

[11] M. Bakker, F. Schaars, J.D. Hughes, C.D. Langevin, A.M. Dausman (2013), "Documentation of the Seawater Intrusion (SWI2) Package for MODFLOW", *U.S. Department of the Interior and U.S. Geological Survey*, Reston, Virginia.

[12] C. Balducci & A. Cecinato (2010), "Particulate organic acids in the atmosphere of Italian cities: Are they environmentally relevant?", *Atmospheric Environment*, 44/5, 652-659. DOI:10.1016/j.atmosenv.2009.11.015

[13] A.Y. Baye, M. Razack, T. Ayenew, E. Zemedagegnehu (2013), "Estimating transmissivity using empirical and geostatistical methods in the volcanic aquifers of Upper Awash Basin, central Ethiopia", *Environ Earth Sci*, 69, 1791-1802. DOI:10.1007/s12665-012-2011-6

[14] J. Bear (1979), "Hydraulics of groundwater", *McGraw-Hill*.

[15] G. Belluomini, M. Caldara, C. Casini, M. Cerasoli, L. Manfra, G. Mastronuzzi, G. Palmentola, P. Sansò, P. Tuccimei, P.L. Vesica (2002), "The age of Late Pleistocene Shorelines and tectonic activity of Taranto area, Southern Italy", *Quaternary Science Reviews*, 21, 525-547.

[16] M. Bentivenga, M. Coltorti, G. Prosser, E. Tavarnelli (2004a), "A new interpretation of terraces in the Taranto Gulf: the role of extensional faulting", *Geomorphology*, 60, 383-402. DOI:10.1016/j.geomorph.2003.10.002

[17] M. Bentivenga, M. Coltorti, G. Prosser, E. Tavarnelli (2004b), "Recent extensional faulting in the Gulf of Taranto area: implications for nuclear waste storage in the vicinity of Scanzano Ionico (Basilicata)", *Boll. Soc. Geol. It.*, 123, 391-404.

[18] G. Böhm, R. Brauchler, D.Y. Nieto, L. Baradello, A. Affatato, M. Sauter (2013), "A field assessment of site-specific correlations between hydraulic and geophysical parameters", *Near Surface Geophysics*, 11, 473-483. DOI:10.3997/1873-0604.2013034

[19] G.K. Brouwer, P.A. Fokker, F. Wilschut, W. Zijl (2008), "A direct inverse model to determine permeability fields from pressure and flow rate measurements", *Math Geosci*, 40, 907-920. DOI:10.1007/s11004-008-9183-3

[20] E. Bruno, D. Calcaterra, M. Parise (2008), "Development and morphometry of sinkholes in coastal plains of Apulia, southern Italy. Preliminary sinkhole susceptibility assessment", *Engineering Geology*, 99, 198-209. DOI:10.1016/j.enggeo.2007.11.017.

[21] A. Buccolieri, G. Buccolieri, N. Cardellicchio, A. Dell'Atti, A. Di Leo, A. Maci (2006), "Heavy metals in marine sediments of Taranto gulf (Ionian Sea, southern Italy)", *Marine Chemistry*, 99, 227-235. Doi:10.1016/j.marchem.2005.09.009

[22] F. Cangialosi, G. Intini, L. Liberti, M. Notarnicola, P. Stellacci (2008), "Health risk assessment of air emissions from a municipal solid waste incineration plant-a case study", *Waste Management*, 28/5, 885-95. DOI:10.1016/j.wasman.2007.05.006

[23] J. Cao & G. Jiang (2015), "Introductory editorial: karst contribution to global change", *Environ Earth Sci*, 74, 909-911. DOI:10.1007/s12665-015-4300-3

[24] N. Cardellicchio, A. Buccolieri, S. Giandomenico, L. Lopez, F. Pizzulli, L. Spada (2007), "Organic pollutants (PAHs, PCBs) in sediments from the Mar Piccolo in Taranto (Ionian Sea, Southern Italy)", *Mar Pollut Bull*, 55, 451-458. DOI:10.1016/j.marpolbul.2007.09.007

[25] A. Careghini, A.F. Mastorgio, S. Saponaro, E. Sezenna (2014), "Biosphenol A, nonylphenos, benzophenones, and benzotriazoles in soils, groundwater, surface water, sediments, and food: a review", *Environ Sci Pollut Res*. DOI:10.1007/s11356-014-3974-5

[26] J. Carrera (1988), "State of the art of the inverse problem applied to the flow and solute transport equations", Chapter: Groundwater Flow and Quality Modelling, *Volume 224 of the series NATO ASI Series* pp. 549-583. DOI:10.1007/978-94-009-2889-3_31

[27] J. Carrera, A. Alcolea, A. Medina, J. Hidalgo, L.J. Slooten (2005), "Inverse problems in hydrogeology", *Hydrogeol J*, 13, 206-222.

[28] L. Cattaneo, A. Comunian, G. De Filippis, M. Giudici, C. Vassena (2016), "Modeling Groundwater Flow in Heterogeneous Porous Media with YAGMod", *Computation*, 4/2. DOI:10.3390/computation4010002

[29] R.A. Cavallo & L. Stabili (2002), "Presence of vibrios in seawater and Mytilus galloprovincialis (Lam.) from the Mar Piccolo of Taranto (Ionian Sea)", *Water Research*, 36, 3719-3726.

[30] H.I. Chaminé (2015), "Water resources meet sustainability: new trends in environmental hydrogeology and groundwater engineering", *Environmental Earth Sciences*, 73, 2513-2520. DOI:10.1007/s12665-014-3986-y

[31] H.I. Chaminé & J.M. Carvalho (2015), "Meeting sustainability and water resources: a dialogue about groundwater science and hydrogeology practice", *Environ Earth Sci*, 73, 2531-2542. DOI:10.1007/s12665-014-3969-z

[32] Z. Chen, G. Huan, Y. Ma (2006), "Flows in fractured porous media", in *Computational Methods for Multiphase Flows in Porous Media*, Computational Science & Engineering.

[33] A. Cilumbriello, L. Sabato, M. Tropeano, S. Gallicchio, A. Grippa, P. Maiorano, G. Mateu-Vicens, C.A. Rossi, G. Spilotro, L. Calcagnile, G. Quarta (2010), "Sedimentology, stratigraphic architecture and preliminary hydrostratigraphy of the Metaponto coastal-plain sebsurface (Southern Italy)", *Mem. Descr. Carta Geol. d'It.*, XC, 76-84.

[34] Consea S.r.l. (2010), "Progetto definitivo e studio di impatto ambientale (SIA) - Final project and environmental impact study", language: Italian.

[35] N. Copty, Y. Rubin, G. Mavko (1993), "Geophysical-hydrological identification of field permeabilities through Bayesian updating", *Water Resour. Res.*, 29/8, 2813-2825.

[36] V. Cotecchia, D. Grassi, M. Polemio (2005), "Carbonate Aquifers in Apulia and Seawater Intrusion", *Giornale di Geologia Applicata*, 1, 219-231. DOI:10.1474/GGA.2005-01.0-22.0022

[37] E. Custodio (2010), "Coastal aquifers of Europe: an overview", *Hydrogeology Journal*, 18, 269-280. DOI:10.1007/s10040-009-0496-1

[38] A. Das & B. Datta (2001), "Simulation of seawater intrusion in coastal aquifers: Some typical responses", *Sadhana*, 26/4, 317-352.

[39] G. De Filippis, M. Giudici, S. Margiotta, F. Mazzone, N. Negri, C. Vassena (2013), "Numerical modeling of the groundwater flow in the fractured and karst aquifer of the Salento peninsula (Southern Italy)", *Acque Sotterranee*, 2, 17-28. DOI:10.7343/AS-016-013-0040

[40] G. de Marsily, J.P. Delhomme, A. Coudrain-Ribstein, A.M. Lavenue (2000), "Four decades of inverse problems in hydrogeology", *Geological Society of America Special Papers*, 348, 1-17. DOI:10.1130/0-8137-2348-5.1

[41] F. De Serio, D. Malcangio, M. Mossa (2007), "Circulation in a Southern Italy coastal basin: Modelling and field measurements", *Continental Shelf Research*, 27/6, 779-797. DOI:10.1016/j.csr.2006.11.018

[42] A. Dhar, S. Sahoo, S. Dey, M. Sahoo (2014), "Evaluation of Recharge and Groundwater Dynamics of a Shallow Alluvial Aquifer in Central Ganga Basin, Kanpur (India)", *Natural Resources Research*, 23/4. DOI:10.1007/s11053-014-9251-y

[43] D. Di Bucci, S. Coccia, U. Fracassi, V. Iurilli, G. Mastronuzzi, G. Palmentola, P. Sansò, G. Selleri, G. Valensise (2009), "Late Quaternary deformation of the southern Adriatic foreland (southern Apulia) from mesostructural data: preliminary results", *Ital.J.Geosci. (Boll.Soc.Geol.It.)*, 128/1, 33-46.

[44] R. Di Maio, S. Fabbrocino, G. Forte, E. Piegari (2014), "A three-dimensional hydrogeological-geophysical model of a multi-layered aquifer in the coastal alluvial plain of Sarno River (southern Italy)", *Hydrogeology Journal*, 22, 691-703. DOI:10.1007/s10040-013-1087-8

[45] Distretto Idrografico dell'Appennino Meridionale (2010), "Piano di Gestione Acque; All.3: Caratterizzazione geologica e idrogrologica - Identificazione degli acquiferi - Groundwater management; Att.3: Geological and hydrogeological characterization - Aquifer identification", language: Italian.

[46] P.A. Domenico & F.W. Schwartz (1990), "Physical and Chemical Hydrogeology", *John Wiley & Sons, Inc., Second edition*.

[47] M.A. El-Rawy, W. Zijl, O. Batelaan, G.A. Mohammed (2010), "Application of the double constraint method combined with MODFLOW", *In Proceedings of the Valencia IAHR International Groundwater Symposium*, September 22-24.

[48] European Environment Agency. http://www.eea.europa.eu/it

[49] F. Fiorillo & Z. Stevanović (2015), "Introductory editorial thematic issue: Mediterranean karst hydrogeology", *Environ Earth Sci*, 74, 1-3. DOI:10.1007/s12665-015-4548-7

[50] P. Fleury, M. Bakalowicz, G. de Marsily (2007), "Submarine springs and coastal karst aquifers: a review", *Journal of Hydrology*, 339, 79-92. DOI:10.1016/j.jhydrol.2007.03.009

[51] L. Foglia, S.W. Mehl, M.C. Hill, P. Perona, P. Burlando (2007), "Testing Alternative Ground Water Models Using Cross-Validation and Other Methods", *Ground Water*, 45/5, 627-641. DOI: 10.1111/j.1745-6584.2007.00341.x

[52] L. Foglia, M.C. Hill, S.W. Mehl, P. Burlando (2009), "Sensitivity analysis, calibration, and testing of a distributed hydrological model using error-based weighting and one objective function", *Water Resources Research*, 45. DOI: 10.1029/2008WR007255

[53] L. Foglia, S.W. Mehl, M.C. Hill, P. Burlando (2013), "Evaluating model structure adequacy: The case of the Maggia Valley groundwater system, southern Switzerland", *Water Resources Research*, 49, 260-282. DOI: 10.1029/2011WR011779

[54] O.L. Franke & T.E. Reilly (1987), "The effects of boundary conditions on the steady-state response of three hypothetical ground-water systems - Results and implications of numerical experiments", *U.S. Geological Survey*, Water-Supply Paper 2315, 19 pp.

[55] O.L. Franke, T.E. Reilly, G.D. Bennett (1987), "Definition of boundary and initial conditions in the analysis of saturated ground-water flow systems - An introduction", *U.S. Geological Survey*, Techniques of Water-Resources Investigations, Book 3, Chapter B5.

[56] C. Gariazzo, V. Papaleo, A. Pelliccioni, G. Calori, P. Radice, G. Tinarelli (2007), "Application of a Lagrangian particle model to assess the impact of harbour, industrial and urban activities on air quality in the Taranto area, Italy", *Atmospheric Environment*, 41/30, 6432-6444. DOI:10.1016/j.atmosenv.2007.06.005

[57] L.W. Gelhar, C. Welty, K.R. Rehfeldt (1992), "A critical review of data of field-scale dispersion in aquifers", *Water Resources Research*, 28, 1955-1974.

[58] V. Gholami, H. Aghagoli, A.M. Kalteh (2015), "Modeling sanitary boundaries of drinking water wells on the Caspian Sea southern coasts, Iran", *Environ Earth Sci*, 74, 2981-2990. DOI:10.1007/s12665-015-4329-3

[59] M. Giudici, material from "Earth Physics" course at University of Milan.

[60] M. Giudici, material from "Geophysical and environmental modeling" course at University of Milan.

[61] M. Giudici, G. Morossi, G. Parravicini, G. Ponzini (1995), "A new method for the identification of distributed transmissivities", *Water Resources Research*, 31/8, 1969-1988. DOI:10.1029/95WR01205

[62] M. Giudici (2001), "Development, calibration and validation of physical models. Geographic Information Systems and Environmental Modeling", edited by K. C. Clarke, B. O. Parks e M. C. Krane, 100-121, Prentice-Hall, Upper Saddle River (NJ).

[63] M. Giudici & C. Vassena (2006), "dsm.f90: A computer code for the solution of an inverse problem of ground water hydrology by the differential system method", *Computers & Geosciences*, 32/10, 1709-1719.

[64] M. Giudici (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*, XC, 113-120. ISSN:0536-0242

[65] M. Giudici, S. Margiotta, F. Mazzone, S. Negri, C. Vassena (2012), "Modelling hydrostratigraphy and groundwater flow of a fractured and karst aquifer in a Mediterranean basin (Salento peninsula, southeastern Italy)", *Environ Earth Sci*. DOI:10.1007/s12665-012-1631-1

[66] W. Guo & C.D. Langevin (2002), "User's guide to SEAWAT: a computer program for simulation of three-dimensional variable-density ground-water flow", *U.S. Geological Survey*, Tallahassee, Florida.

[67] R.T. Green, S.L. Painter, A. Sun, S. Worthington (2006), "Groundwater contamination in karst terranes", *Water, Air and Soil Pollution: Focus*, 6, 157-170. DOI: 10.1007/s11267-005-9004-3

[68] A.W. Harbaugh (1995), "Direct solution package based on alternating diagonal ordering for the U.S. Geological Survey modular finite-difference ground-water flow model", *U.S. Geological Survey*, Open-File Report 95-288, 46 p.

[69] A.W. Harbaugh (2005), "MODFLOW-2005, The U.S. Geological Survey Modular Ground-Water Model - The Ground-Water Flow Process", *U.S. Geological Survey*.

[70] H.J. Hendricks Franssen, A. Alcolea, M. Riva, M. Bakr, N. Van der Wiel, F. Stauffer, A. Guadagnini (2009), "A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments", *Advances in Water Resources*, 32/6, 851-872.

[71] M.C. Hill (1990), "Preconditioned conjugate-gradient 2 (PCG2), a computer program for solving ground-water flow equations", *U.S. Geological Survey*, Water-Resources Investigations Report 90-4048, 43 p.

[72] M.C. Hill & C.R. Tiedeman (2007), "Effective groundwater model calibration with analysis of data, sensitivities, predictions, and uncertainty", *Wiley & Sons*, New York, 455 p.

[73] D. Hillel (1998), "Environmental soil physics", *Academic Press*.

[74] K.W.F. Howard (2015), "Sustainable cities and the groundwater governance challenge", *Environmental Earth Sciences*, 73, 2543-2554. DOI:10.1007/s12665-014-3370-y

[75] http://en.wikipedia.org/wiki/File:Saltwater_Intrusion.gif

[76] D.W. Hyndman, J.M. Harris, S.M. Gorelick (1994), "Coupled seismic and tracer test inversion for aquifer property characterization", *Water Resour. Res.*, 30/7, 1965-1977.

[77] P.S. Huyakorn, B.H. Lester, J.W. Mercer (1983), "An efficient finite element technique for modeling transport in fractured porous media - 1. Single species transport", *Water Resources Research*, 19/3, 841-854.

[78] Y. Lan, M. Jin, C. Yan, Y. Zou (2015), "Schemes of groundwater exploitation for emergency water supply and their environmental impacts on Jiujiang City,China", *Environ Earth Sci*, 73, 2365-2376. DOI:10.1007/s12665-014-3586-x

[79] C. Lu, Y. Chen, J. Luo (2012), "Boundary condition effects on maximum groundwater withdrawal in coastal aquifers", *Ground Water*, 50/3, 386-393. doi:10.1111/j.1745-6584.2011.00880.x

[80] M. Maggiore & P. Pagliarulo (2004), "Circolazione idrica ed equilibri idrogeologici negli acquiferi della Puglia - Groundwater flow and hydrogeological equilibrium of aquifers in Apulia region", *Geologi e Territorio*, 13-35, language: Italian.

[81] W.E. Mahmod, K. Watanabe, A.A. Zahr-Eldeen (2013), "Analysis of groundwater flow in arid areas with limited hydrogeological data using the Grey Model: a case study of the Nubian Sandstone, Kharga Oasis, Egypt", *Hydrogeology Journal*, 21, 1021-1034. DOI:10.1007/s10040-013-0959-2

[82] A. Marangella & M. Parise (1989), "Caratteri geomorfologici e naturalistici delle gravine di Grottaglie - Geomorphological and natural features of Grottaglie ravines", *Atti Spelaion 2007*, Altamura, 7-9 Dicembre 2007, language: Italian.

[83] S. Margiotta & S. Negri (2004), "Alla ricerca dell'acqua perduta - Searching for lost water", *Congedo Editore*, language: Italian.

[84] S. Margiotta & S. Negri (2005), "Geophysical and stratigraphycal research into deep groundwater and intruding seawater in the mediterranean area (the Salento Peninsula, Italy)", *Natural Hazards and Earth System Sciences*, 5, 127-136.

[85] S. Margiotta, F. Mazzone, S. Negri (2010), "Stratigraphic revision of Brindisi-Taranto plain: hydrogeological implications", *Mem. Descr. Carta Geol. D'It.*, XC, 165-180.

[86] S. Martos-Rosillo, M. Rodríguez-Rodríguez, A. Pedrera, J.J. Cruz-SanJulián, J.C. Rubio (2013), "Groundwater recharge in semi-arid carbonate aquifers under intensive use: the Estepa Range aquifers (Seville, southern Spain)", *Environ Earth Sci*, 70, 2453-2468. DOI:10.1007/s12665-013-2288-0

[87] M. Masciopinto (2006), "Simulation of coastal groundwater remediation: the case of Nardò fractured aquifer in Southern Italy", *Environmental Modelling & Software*, 21, 85-97. DOI:10.1016/j.envsoft.2004.09.028

[88] G. Mastronuzzi & P. Sansò (2000), "Boulders transport by catastrophic waves along the Ionian coast of Apulia (southern Italy)", *Marine Geology*, 170, 93-103.

[89] G. Mastronuzzi & P. Sansò (2002a), "Pleistocene sea-level changes, sapping processes and development of valley networks in the Apulia region (southern Italy)", *Geomorphology*, 4, 19-34.

[90] G. Mastronuzzi & P. Sansò (2002b), "Holocene coastal dune development and environmental changes in Apulia (southern Italy)", *Sedimentary Geology*, 150, 139-152.

[91] M.G. McDonald & A.W. Harbaugh (1988), "A modular three-dimensional finite-difference groundwater flow model", *U.S. Geological Survey*, Techniques of Water-Resources Investigations, book 6, chap. A1, 586 p.

[92] D. McLaughlin & L.R. Townley (1996), "A reassessment of the groundwater inverse problem", *Water Resour. Res.*, 32/5, 1131-1161.

[93] Millennium Ecosystem Assessment - MEA (2005), "Ecosystems and Human Well-Being: Synthesis", *Island Press*, Washington. 155pp.

[94] K.A. Mogaji, H.S. Lim, K. Abdullah (2015), "Modeling of groundwater recharge using a multiple linear regression (MLR) recharge model developed from geophysical parameters: a case of groundwater resources management", *Environ Earth Sci*, 73, 1217-1230. DOI:10.1007/s12665-014-3476-2

[95] S.P. Neuman (1973), "Calibration of distributed parameter groundwater flow models viewed as a multiobjective decision process under uncertainty", *Water Resour Res*, 9, 1006-1021.

[96] T. Odeh, T. Rödiger, S. Geyer, M. Schirmer (2015), "Hydrological modelling of a heterogeneous catchment using an integrated approach of remote sensing, a geographic information system and hydrologic response units: the case study of Wadi Zerka Ma'in catchment area, north east of the Dead Sea", *Environ Earth Sci*, 73, 3309-3326. DOI:10.1007/s12665-014-3627-5

[97] F. Padilla & J. Cruz-Sanjuliàn (1997), "Modeling sea-water intrusion with open boundary conditions", *Ground Water*, 35/4, 704-712.

[98] M. Parise & M. Sammarco (2015), "The historical use of water resources in karst", *Environ Earth Sci*, 74, 143-152. DOI:10.1007/s12665-014-3685-8

[99] G. Parravicini, M. Giudici, G. Morossi, G. Ponzini (1995), "Minimal a priori assignment in a direct method for determining phenomenological coefficients uniquely", *Inverse Problems*, 11/3.

[100] E.P. Poeter, M.C. Hill, D. Lu, C.R. Tiedeman, S. Mehl (2014), "UCODE_2014, with new capabilities to define parameters unique to predictions, calculate weights using simulated values, estimate parameters with SVD, evaluate uncertainty with MCMC, and more", *IGWMC - International Groud Water Modeling Center*.

[101] M. Polemio (1996), "Le calamità idrogeologiche dell'inverno 1995-96 nel territorio Tarantino - Hydrogeological disasters occurred in the area of Taranto in winter 1995-96", *International congress "La prevenzione delle catastrofi idrogeologiche: il contributo della ricerca scientifica - Prevention of hydrogeological disasters: the contribution of scientific research"*, 5-7 November 1996, Alba (Italy).

[102] G. Ponzini & A. Lozej (1982), "Identification of aquifer transmissivities: the comparison model method", *Water Resources Research*, 18, 597-622.

[103] G. Ponzini & G. Crosta (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.

[104] N. Ravbar & G. Kovacic (2015), "Vulnerability and protection aspects of some Dinaric karst aquifers: a synthesis", *Environ Earth Sci*, 74, 129-141. DOI:10.1007/s12665-014-3945-7

[105] Regione Puglia & Autorità di Bacino della Puglia (2010), "Aggiornamento del Bilancio Idrogeologico dei Corpi Idrici Sotterranei della Regione Puglia; All.4.5: Caratteristiche della circolazione idrica negli acquiferi carbonatici pugliesi: deflussi a mare e scambi idrici tra idrostrutture confinanti - Updating the hydrogeological balance of groundwater basins of the Apulia region; Att.4.5: Characteristics of groundwater dynamics in carbonate aquifers of the Apulia region: outflows towards the sea and water exchanges among hydrogeological structures", language: Italian.

[106] G. Ricchetti, N. Ciaranfi, P. Pieri (1988), "Note alla Carta Geologica delle Murge e del Salento (Puglia Centromeridionale) - Geologic map of the Murge plateau and Salento peninsula (midsouthern Apulia)", *Memorie della Società Geologica Italiana*, 41/1, language: Italian.

[107] A. Romanazzi, F. Gentile, M. Polemio (2015), "Modelling and management of a Mediterranean karstic coastal aquifer under the effects of seawater intrusion and climate change", *Environ Earth Sci*, 74, 115-128. DOI:10.1007/s12665-015-4423-6

[108] M. Rossi (2011), "Hydrogeology outlines and seawater intrusion theory and modeling", *Technical workshop on monitoring seawater intrusion in coastal groundwater*, Dar es Salaam, Tanzania.

[109] Y. Rubin & G. Dagan (1992), "Conditional estimation of solute travel time in heterogeneous formations: Impact of transmissivity measurements", *Water Resour. Res.*, 28/4, 1033-1040.

[110] K.R. Rushton & S.C. Redshaw (1979), "Seepage and groundwater flow", *John Wiley & Sons, Inc.*, New York, 339 pp.

[111] D. Sauer, S. Wagner, H. Brückner, F. Scarciglia, G. Mastronuzzi, K. Stahr (2010), "Soil development on marine terraces near Metaponto (Gulf of Taranto, southern Italy)", *Quaternary International*, 222, 48-63. DOI:10.1016/j.quaint.2009.09.030

[112] B.R. Scanlon, R.E. Mace, M.E. Barrett, B. Smith (2003), "Can we simulate regional groundwater flow in a karst system using equivalent porous media models? Case study, Barton Springs Edwards aquifer, USA", *Journal of Hydrology*, 276, 137-158. DOI:10.1016/S0022-1694(03)00064-7

[113] S. Scarascia & G. Ponzini (1972), "An approximate solution for the inverse problem in hydraulics", *L'Energia Elettrica*, 49, 518-531.

[114] M.R. Senatore, G. Diplomatico, L. Mirabile, T. Pescatore, M. Tramutoli (1982), "Franamenti sulla scarpata continentale Pugliese del Golfo di Taranto (alto Ionio-Italia) - Landslides on the continental slope of the Gulf of Taranto (Apulia, Ionian Sea - Italy)", *Geologica Rom.*, 21, 497-510, language: Italian.

[115] O. Simone & A. Fiore (2014), "Five Large Collapse Dolines in Apulia (Southern Italy)-the Dolina Pozzatina and the Murgian Puli", *Geoheritage*, 6, 291-303. DOI:10.1007/s12371-014-0122-z

[116] B.B.S. Singhal & R.P. Gupta (2010), "Applied Hydrogeology of Fractured Rocks", *Springer Netherlands*, 2nd edition.

[117] SIT-Puglia. http://www.sit.puglia.it/

[118] SOGESID S.p.A. & Regione Puglia (2009a), "Piano di Tutela delle Acque; All.1.2: Caratterizzazione climatologica; Att.1.2: Climatological characterization", language: Italian.

[119] SOGESID S.p.A. & Regione Puglia (2009b), "Piano di Tutela delle Acque; All.4: Stima delle pressioni e degli impatti da fonti puntuali e diffuse; Att.4: Estimation of pressures and impacts due to point and distributed sources", language: Italian.

[120] SOGESID S.p.A. & Regione Puglia (2009c), "Piano di Tutela delle Acque; All.6.1: Archivio anagrafico ed analisi dei punti acqua censiti (pozzi e sorgenti) - Groundwater protection plan; Att.6.1: Data archive and analysis of groundwater surveys", language: Italian.

[121] Z. Stevanović, V. Ristić-Vakanjac, S. Milanović, L. Vasić, B. Petrović, M. Čokorilo-Ilić (2015), "Karstification depth and storativity as main factors of karst aquifer regimes: some examples from southern Alpine branches (SE Europe and Middle East)", *Environ Earth Sci*, 74, 227-240. DOI:10.1007/s12665-015-4046-y

[122] M.M. Storelli & G.O. Marcotrigiano (2000), "Polycyclic aromatic hydrocarbon distributions in sediments from the Mar Piccolo, Ionian Sea, Italy", *Bull. Environ. Contam. Toxicol.*, 65, 537-544. DOI:10.1007/s001280000157

[123] N.Z. Sun (1994), "Inverse Problems in groundwater modeling", *Kluwer Academic*, Dordrecht, 337 pp.

[124] N.Z. Sun (2004), "Inverse Problems in Groundwater Modeling", *Kluwer Academic*, Dordrecht.

[125] Sviluppo Italia S.p.A. (2010), "Annex 4 to the illustrative-technical report: Realizzazione del piano di caratterizzazione ambientale dell'area marino-costiera prospiciente il Sito di Interesse Nazionale di Taranto: Mar Piccolo - Environmental characterization plan of the coastal area near the Taranto site of National interest: Mar Piccolo", language: Italian.

[126] M.G. Trefry & C. Muffels (2007), "FEFLOW: A Finite-Element Ground Water Flow and Transport Modeling Tool", *Ground Water*, 45, 525-528. DOI:10.1111/j.1745-6584.2007.00358.x

[127] A. Trykozko, G.K. Brouwer, W. Zijl (2008), "Downscaling: a complement to homogenization", *International Journal of Numerical Analysis and Modeling*, 5, 157-170.

[128] A. Trykozko, G.A. Mohammed, W. Zijl (2009), "Downscaling: the inverse of upscaling", *in Conference on Mathematical and Computational Issues in the Geosciences*, SIAM GS 2009, June 15-18, Leipzig.

[129] Y.W. Tsang, C.F. Tsang, F.V. Hale, B. Dvertstorp (1996), "Tracer transport in a stochastic continuum model of fractured media", *Water Resour. Res.*, 32/10, 3077-3092.

[130] L. Tulipano (2003), "Overexploitation Consequences and management criteria in coastal karstic aquifers", *Conf. TIAC'03*, Alicante, Spain.

[131] L. Tulipano & G. Sappa (2012), "Lezioni di Idrogeologia Applicata - Lessons of applied hydrogeology, Cap.10 Gli acquiferi costieri - Coastal aquifers", http://ww2.mediateca.com/mediateca/index.php/it/educational/doc_download/213-capitolo-10.html

[132] G. Umgiesser, I. Scroccaro, G. Alabiso (2007), "Mass exchange machanism in the Taranto Sea", *Transitional Waters Bulletin*, 2, 59-71. DOI:10.1285/i1825229Xv1n2p59

[133] G. Valota, M. Giudici, G. Parravicini, G. Ponzini, E. Romano (2002), "Is the forward problem of ground water hydrology always well posed?", *Ground Water*, 40, 500-508.

[134] M. Usman, R. Liedl, A. Kavousi (2015), "Estimation of distributed seasonal net recharge by modern satellite data in irrigated agricultural regions of Pakistan", *Environ Earth Sci*, 74, 1463-1486. DOI:10.1007/s12665-015-4139-7

[135] A. Vallejos, J.M. Andreu, F. Sola, A. Pulido-Bosch (2015), "The anthropogenic impact on Mediterranean karst aquifers: cases of some Spanish aquifers", *Environ Earth Sci*, 74, 185-198. DOI:10.1007/s12665-014-3994-y

[136] C. Vassena, C. Durante, M. Giudici, G. Ponzini (2008), "The importance of observations on fluxes to constrain ground water model calibration", *Physics and Chemistry of the Earth*, 33, 1105-1110. DOI:10.1016/j.pce.2008.01.004

[137] C. Vassena, M. Rienzner, G. Ponzini, M. Giudici, C. Gandolfi, C. Durante, D. Agostani (2012), "Modeling water resources of an highly irrigated alluvial plain: coupling and calibrating soil and ground water models", *Hydrogeology Journal*. DOI:10.1007/s10040-011-0822-2

[138] V. Velasco, R. Gogu, E. Vázquez-Suñè, A. Garriga, E. Ramos, J. Riera, M. Alcaraz (2013), "The use of GIS-based 3D geological tools to improve hydrogeological models of sedimentary media in an urban environment", *Environ Earth Sci*, 68, 2145-2162. DOI:10.1007/s12665-012-1898-2

[139] C.I. Voss & A.M. Provost (2010), "SUTRA: A Model for Saturated-Unsaturated Variable-Density Ground-Water Flow with Solute or Energy Transport", *U.S. Geological Survey*, Reston, Virginia.

[140] A.D. Werner & C.T. Simmons (2009), "Impact of sea-level rise on sea water intrusion in coastal aquifers", *Ground Water*, 47/2, 197-204. doi: 10.1111/j.1745-6584.2008.00535.x

[141] A.D. Werner, M. Bakker, V. Post, A. Vanderbohede, C. Lu, B. Ataie-Ashtiani, C.T. Simmons, D.A. Barry (2013), "Seawater intrusion processes, investigation and management: recent advances and future challenges", *Advances in Water Resources*, 51, 3-26.

[142] S.R.H. Worthington (2003), "A comprehensive strategy for understanding flow in carbonate aquifer", *Speleogenesis and Evolution of Karst Aquifers*, 1/1.

[143] Z. Xun, S. Chao, L. Ting, C. Ruige, Z. Huan, Z. Jingbo, C. Qin (2015), "Estimation of aquifer parameters using tide-induced groundwater level measurements in a coastal confined aquifer", *Environ Earth Sci*, 73, 2197-2204. DOI:10.1007/s12665-014-3570-5

[144] W.G.W. Yeh (1986), "Review of parameter identification procedures in groundwater hydrology: the inverse problem", *Water Resour Res*, 22, 95-108.

[145] À. Yustres, V. Navarro, L. Asensio, M. Candel, B. Garcìa (2013), "Groundwater resources in the Upper Guadiana Basin (Spain): a regional modelling analysis", *Hydrogeology Journal*, 21, 1129-1146. DOI:10.1007/s10040-013-0987-y

[146] A. Zahid, M.Q. Hassan, K.M.U. Ahmed (2015), "Simulation of flowpaths and travel time of groundwater through arsenic-contaminated zone in the multi-layered aquifer system of Bengal Basin", *Environ Earth Sci*, 73, 979-991. DOI:10.1007/s12665-014-3447-7

[147] A. Zander, A. Fulling, H. Bruckner, G. Mastronuzzi (2006), "OSL dating of upper Pleistocene littoral sediments: a contribution to the chronostratigraphy of raised marine terraces bordering the Gulf of Taranto, South Italy", *Geogr. Fis. Dinam. Quat.*, 29, 33-50.

[148] C. Zheng & P.P. Wang (1999), "MT3DMS - A modular three-dimensional multispecies transport model for simulation of advection, dispersion and chemical reactions of contaminants in groundwater systems", *The University of Alabama*.

[149] H. Zhou, J.J. Gomez-Hernandez, L. Li (2014), "Inverse methods in hydrogeology: evolution and recent trends", *Advances in Water Resources*, 63, 22-37.

[150] W. Zijl, G.A. Mohammed, O. Batelaan, F. De Smedt (2010), "Constraining methods for direct inverse modeling", *XVIII International Conference on Water Resources*, CMWR, CIMNE, Barcelona.

[151] D.A. Zimmerman, G. de Marsily, C.A. Gotway, M.G. Marietta, C.L. Axness, R.L. Beauheim, R.L. Bras, J. Carrera, G. Dagan, P.B. Davies, D.P. Gallegos, A. Galli, J. Gòmez-Hernàndez, P. Grindrod, A.L. Gutjahr, P.K. Kitanidis, A.M. Lavenue, D. McLaughlin, S.P. Neuman, B.S. RamaRao, C. Ravenne, Y. Rubin (1998), "A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow", *Water Resources Research*, 34/6, 1373-1413.

[152] L.E. Zuffianò, P.P. Limoni, A. Basso, D. Casarano, V. Dragone, A. Romanazzi, F. Santaloia, M. Polemio (2015), "Coastal hydrogeological system of Mar Piccolo (Taranto, Italy)", *Environmental Science and Pollution Research*. DOI:10.1007/s11356-015-4932-6

Acknowledgements

The activities described in this thesis were funded by the Flagship Project RITMARE (la Ricerca Italiana per il Mare - the Italian Research for the Sea) coordinated by the National Research Council and funded by the Ministry for Education, University and Research within the National Research Program 2011-2013.
My tutor, Professor Mauro Giudici is kindly acknowledged for scientific contribution to the research activities described here and because he accomplished his tasks with high precision and professionalism. Dr. Laura Foglia from Institut für Angewandte Geowissenschaften at Technische Universität Darmstadt (Germany) and Professor Steffen Mehl from Department of Civil Engineering at California State University in Chico are gladly acknowledged for their scientific support to the tasks presented in this thesis. Professor Sergio Luigi Negri and Dr. Stefano Margiotta from DiSTeBA at Università del Salento in Lecce (Italy) are acknowledged as well, as members of the CINFAI Operative Unit to which I belonged, within the RITMARE Project.