UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Ingegneria Civile, Edile e Ambientale (ICEA)

SCUOLA DI DOTTORATO DI RICERCA IN:

SCIENZE DELL'INGEGNERIA CIVILE ED AMBIENTALE

XXVIII CICLO

DATA ASSIMILATION IN GEOMECHANICS: CHARACTERIZATION OF HYDROCARBON RESERVOIRS

Direttore della Scuola: Ch.mo Prof. Stefano Lanzoni

Relatore: Ch.mo Prof. Giuseppe Gambolati

Correlatore: Dott. Ing. Pietro Teatini

Dottorando: Claudia Zoccarato

27 GENNAIO 2016

UNIVERSITÀ DEGLI STUDI DI PADOVA

Abstract

Department of Civil, Architectural and Environmental Engineering

Doctor of Philosophy

Data Assimilation in Geomechanics: Characterization of Hydrocarbon Reservoirs

Claudia ZOCCARATO

The prediction of the stress field distribution induced by the pore pressure change in deep hydrocarbon reservoirs and the consequent compaction of the porous rock formation is modeled with the aid of a Finite-Element (FE) geomechanical model. Despite the reliability of the model, which has been tested in several previous applications, many sources of uncertainty may affect the model outcome in terms of ground surface displacements. The uncertainty are mainly related to the mathematical model itself, that is an approximation reproducing a real and complex system, the initial and boundary conditions, the forcing terms, and the model parameters. The latter are the physical properties of the reservoir that are usually a-priori poorly known. A proper estimation of these parameters using a deterministic approach is discouraged as several parameters combinations may equally reproduce the observed data. Instead, the reservoir characterization is here performed by establishing a stochastic approach providing also for the quantification of the uncertainties affecting the parameter calibration. For this purpose, an ensemble-based data assimilation algorithm, i.e., the Ensemble Smoother, is elected among the available literature approaches. The methodology is investigated and tested in both synthetic cases and in real case applications by assimilating the available observations from in-situ measurements of ground-surface displacements. The characterization of the reservoir rock properties is provided for an Underground Gas Storage (UGS) reservoir and an offshore producing gas reservoir. Different set of parameters are estimated depending on the available information on the different fields. The parameters of a transversely isotropic model are calibrated using horizontal and

vertical displacements from Persistent Scatterer Interferometry (PSI) measured above the UGS field, while vertical displacements from a time-lapse bathymetry are used to calibrate the uniaxial vertical compressibility of an isotropic constitutive law characterizing the behaviour of the offshore gas reservoir. Generally, it is obtained a satisfactory estimation of the geomechanical parameters with a significant spread reduction of the prior probability distributions when synthetic measurements, i.e., the displacements generated by an independent model run, are assimilated. However, more difficulties are encountered using real observations. This study gives indications on the main factors influencing the geomechanical characterization when assimilating movements of the land surface. The numerical results underline the importance of the consistency between the forward model and the assimilated measurements with an appropriate selection of data necessary to eliminate potential biases of the measurements and/or the modeling procedure.

UNIVERSITÀ DEGLI STUDI DI PADOVA

Sommario

Dipartimento di Ingegneria Civile, Edile e Ambientale

Dottorato di Ricerca

Data Assimilation in Geomechanics: Characterization of Hydrocarbon Reservoirs

Claudia ZOCCARATO

Lo stato tensionale indotto dalla variazione di pressione in giacimenti profondi e la conseguente compattazione delle formazioni geologiche sono simulati con l'ausilio di un modello geomeccanico agli Elementi Finiti (FEM). Nei decenni passati, il citato modello è stato utilizzato in molteplici applicazioni e, tuttavia, le incertezze introdotte nella modellazione sono numerose e possono influire significativamente sulla risposta del modello, in termini di spostamenti superficiali. Le incertezze sono principalmente legate alla semplificazione intrinseca nel processo di modellazione, alle scarsamente note condizioni iniziali e al contorno, alle forzanti esterne e ai parametri del modello, e cioè le proprietà fisiche del giacimento, solitamente non conosciute a-priori. La stima di questi ultimi è ottenuta, in questo lavoro di tesi, attraverso lo sviluppo e l'implementazione di metodologie di tipo probabilistico che permettono di quantificare anche il grado di incertezza associato alla stima dei parametri del modello. Per questo scopo viene utilizzato il cosiddetto Ensemble Smoother, un particolare algoritmo di data assimilation basato su un approccio di tipo Monte Carlo. La metodologia proposta è stata applicata e testata sia su casi sintetici che su casi reali assimilando dati di spostamento superficiale misurati in-situ. I parametri geomeccanici sono stati stimati in due specifici giacimenti. Nel primo caso, si tratta di un sito per lo stoccaggio di gas metano mentre, il secondo caso, riguarda un sito offshore utilizzato per l'estrazione di gas. Nei due casi, per descrivere il comportamento geomeccanico del giacimento, sono state utilizzate leggi costitutive differenti, sulla base delle osservazioni disponibili nei due campi di interesse. In un caso, i parametri di un modello trasversalmente isotropo

sono stati stimati usando misure interferometriche satellitari di spostamento superficiale sia orizzontale che verticale disponibili sul sito di stoccaggio. Nell'altro caso, una legge costitutiva più semplice di tipo isotropo è stata calibrata nel sito offshore dove le osservazioni a disposizione forniscono solo la componente verticale dello spostamento, stimata da una mappa differenziale di batimetria. Nei test sintetici, è stato dimostrato che la metodologia permette di valutare in modo soddisfacente i parametri geomeccanici con una riduzione notevole dell'incertezza inizialmente ipotizzata per i parametri in gioco. Tuttavia, la stima degli stessi è più difficile nei casi reali dove la discrepanza tra il risultato del modello FEM e le misure assimilate può suggerire una preliminare selezione delle misure disponibili per eliminare potenziali evidenti errori nelle misure stesse.

Acknowledgements

This thesis would not have been possible without the collaboration and support from many people, so it is time to say *thank you*.

I first wish to express my special thanks to my advisor Prof. Giuseppe Gambolati as well as to my co-advisor Pietro Teatini to whom I am grateful for the constant supervision and guidance during the past three years.

I would like to thank the whole research group 'ex DMMSA' at the University of Padova. In particular, I thank Massimiliano Ferronato for all valuable advices, discussions, and for the time we spent at conferences.

I also wish to thank Domenico Baù for his kind support during my visiting period at the Colorado State University and for the ongoing collaboration, which always help me to improve my research.

I would like to acknowledge Ayman Alzraiee for the interesting talks about Data Assimilation and Damiano Pasetto for encouraging me to start working on gPCE. I am also grateful to Noemi Friedman for our collaboration started during the past year.

A special thank goes to Omar Tosatto for his help, friendship and shared days in the office, Andrea Pellizzon who always helped me with IT issues, and the *Sala Calcolo* people.

The last thank is to the support and love of my parents, family, friends Elena, Caterina, Andrea, Giorgia, Cecilia, Carlotta, Emma and flatmates Elisa and Sergio.

Most importantly, thank you to Michele to whom I am indebted for his patience and love. I dedicate this work to him.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis is the outcome of a three-years research project, which was the initiative of the Department of Civil, Architectural and Environmental Engineering of the University of Padova (Italy), the Department of Civil and Environmental Engineering of the Colorado State University (Colorado, USA), and the Italian multinational oil and gas company eni S.p.A. The main objective of the '*Data Assimilation of SAR-Based Measurements for Geomechanical Characterization*'(DAG) project was to develop and implement a Data Assimilation (DA) based computational framework for characterizing the geomechanical parameters and the constitutive relationship of deep gas bearing formations of the Po river Plain and the Upper Adriatic sedimentary basin, Italy.

Over the last decades, the geomechanical modeling of deep reservoirs has attracted a continuous interest with a growing need for the development of reliable models to forecast and prevent the possible environmental impacts due to the oil and gas activities, e.g., anthropogenic land subsidence/heave and induced seismicity. Typically much of uncertainty affects the geomechanical modeling of real porous media. This is mainly related to the initial and boundary conditions prescribed to the mathematical models, the forcing terms, the system parameters or the model itself that may be unable to capture the full behavior of the occurrence. For this reason, the uncertainty quantification plays an important role for a reliable assessment of the model outcome and great efforts are therefore devoted to the implementation of stochastic approaches to adapt the classical deterministic workflows and improve the computation of the uncertainties of the model solutions.

Evaluating and reducing the uncertainties related to the geomechanical modeling of deep reservoirs is the main goal of this thesis with the final target to improve the

geomechanical characterization of the porous rock formations. In particular, the aim is at reducing the uncertainty of the unknown model parameters using a DA approach. DA consists of the combination of two main sources of information, i.e., the forecast solution from the numerical model and the observations of the physical system. In other words, this approach looks for an update of the model variables and parameters inferring the response using available measurements. Although several DA algorithms are available in literature, the investigation is here restricted to the so-called ensemble based or Monte Carlo methods to stochastically simulate the model response. In particular, the Ensemble Smoother (ES) algorithm is implemented to solve the inverse problem by incorporating all available observations in only one analysis step.

A different approach is also preliminarily investigated thanks to the collaboration with the Institute of Scientific Computing of the Technische Universitat Braunschweig (Germany). In fact, due to the slow convergence of the Monte Carlo methods and the high computational cost required for large and complex reservoir models, the generalized polynomial chaos (gPCE) expansion may provide a valuable alternative. gPCE is a non-sampling-based method to evaluate the propagation of the prior uncertainty through the forward problem using a surrogate model based on the projection of the outputs on a basis of stochastic polynomials. The comparison of this approach with the ensemble-based methods will be briefly discussed.

1.1 Problem Statement and Study Objectives

Consider a deep hydrocarbon reservoir and the processes of fluid production from the subsurface. The ensuing compaction of the porous rock formation is a well-known consequence which in turn causes the deformation of the reservoir overburden and the settlement of the ground surface. A mathematical model to simulate this physical system aiming at predicting the anthropogenic land subsidence can generally be written as:

$$\mathbf{u}(\mathbf{x}, t) = \mathcal{G}(\psi(\mathbf{x}, t), \alpha) \tag{1.1}$$

where $\mathbf{u}(\mathbf{x}, t)$ is the model outcome, i.e., the land surface displacement, function of the space $\mathbf{x}$ and time t, $\mathcal{G}$ is the model operator, $\psi(\mathbf{x}, t)$ is the model state, i.e., the pore

fluid pressure, and, α is a vector of the model parameters which can possibly vary in space and time. The forward problem relies on the solution of the poro-elastic partial differential equations (PDEs) to simulate the reservoir response together with the initial and boundary conditions and the prior knowledge of the model parameters. On the contrary, the inverse problems seeks the best estimate of α by matching the model predictions $\mathbf{u}(\mathbf{x}, t)$ to observed data $\mathbf{d}(\mathbf{x}, \mathbf{t})$.

This work aims at developing a probabilistic framework to account for and quantify the uncertainty within the geomechanical analysis of hydrocarbon reservoirs. To deal with this problem, the workflow of Fig. 1.1 is used. The physical process of anthropogenic land subsidence/uplift is described with a mathematical model which implies a few simplifications and assumptions. The snake arrow connecting the physical process to the model indicates the uncertainty introduced into the problem description. Other sources of uncertainty affect the prior knowledge of the system, e.g., the parameters α. The solution of the PDEs provides the forecast model solution, e.g., ground surface displacement, which in turn are affected by the above mentioned uncertainty. On the other hand, measurements of the surface movement are provided by ad-hoc monitoring surveys and with the aid of technologically advanced techniques. However, although accurate as much as possible, observations are intrinsically affected by errors. Combining information from the model forecast and the available data from the field through a data assimilation technique, the final goal is to constrain and update the prior knowledge of the parameters thus improving the geomechanical characterization of the reservoir.

The investigation is primarily tested on simple test cases but the main part of the thesis will address the more challenging task of the application to real cases. The capability of the ES to constrain the geomechanical parameters is tested with the aid of a Finite-Element (FE) geomechanical model. Two different reservoir activities are simulated. A transversely isotropic model is employed to forecast the behaviour of a UGS field experiencing a seasonal change of the stress field. By distinction, the deformation caused by gas extraction from a hydrocarbon reservoir is computed using an isotropic geomechanical model. The choice of the models is motivated by the available measurements in the gas fields. PSI observations of land motion are available over the

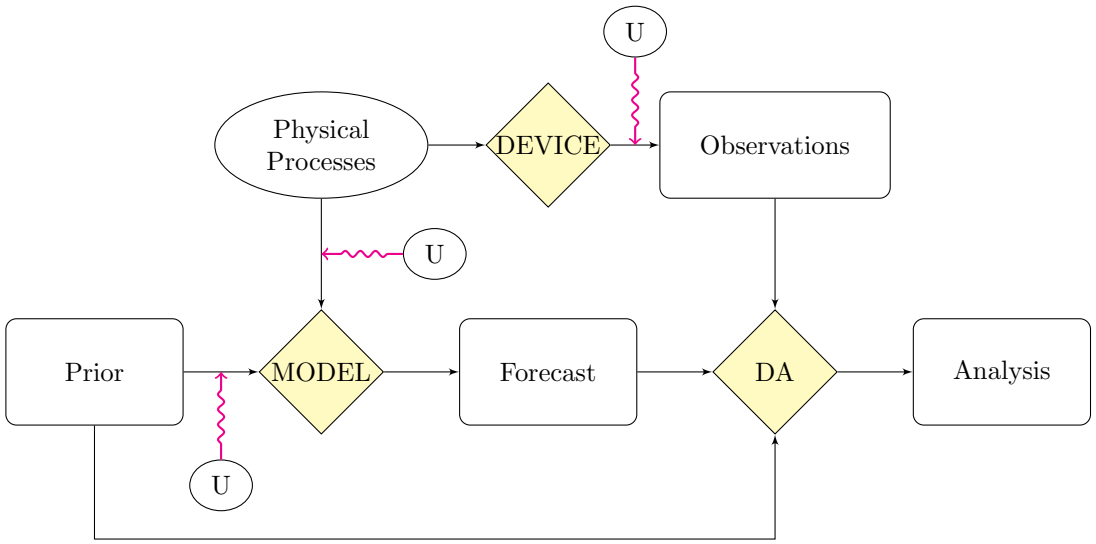

FIGURE 1.1: General workflow for the geomechanical modeling and parameter estimation through DA analysis. U indicates uncertainty.

UGS field and consist of both the horizontal (West-East) and the vertical components of land surface displacement. On the other hand, for the offshore gas reservoir measurements obtained with a time-lapse bathymetric survey have been provided. Thus, the only vertical component is available and this implies the use of an isotropic model because there is no information on the horizontal component required to calibrate a transversely isotropic model.

Although the main focus of this thesis in on the ES, the comparison of this methodology with other available techniques, e.g., the gPCE-based ES, is also made to provide a preliminary evaluation of the pros and cons of these techniques.

1.2 State of the Art

Over the last decade, the importance of uncertainties and risks quantification in reservoir management has been widely recognized. It has been shown that the deterministic calibration of reservoir models does not usually suffice to forecast the future behavior of such complex systems. Indeed, they may give a wrong impression of confidence as several parameters combinations might be able to reproduce equally well the observed data. On the contrary, stochastic methods provide an understanding of how much confidence one should place in any particular model, although such a confidence may not

be high due to data limitations. In this respect, ensemble-based data assimilation methods have been largely adopted to deal with reservoir history matching [*Lorentzen et al.*, 2003; *Nævdal et al.*, 2003; *Gu & Oliver*, 2005; *Skjervheim et al.*, 2011; *Emerick & Reynolds*, 2013; *Iglesias et al.*, 2013a], i.e., to identify parameter (e.g., permeability and porosity) distributions that make the model match satisfactorily in situ pore pressure measurements. An interesting review of history matching in petroleum reservoir engineering is given in *Oliver & Chen* [2011].

Data assimilation and inverse problem methods rely on conditioning the model variables and parameters based on observed data. In *Evensen* [2009], ensemble-based methods are defined as a class of so-called particle methods which use 'a Monte Carlo or ensemble representation for the probability distribution functions (PDFs), an ensemble integration using stochastic models to model the time evolution and a scheme for conditioning the predicted PDFs given the observations'. The Ensemble Kalman Filter (EnKF) has been the most widely used approach. A complete review of its applications in reservoir engineering is provided by *Aanonsen et al.* [2011] and references therein.

Although reservoir geomechanics for land subsidence estimate dates back to the late 1960s [*Geertsma*, 1966, 1973a,b], petroleum industry only recently has started to integrate the stochastical geomechanical modeling into the general framework of reservoir management. For example, the aim is at predicting the compaction of the pumped rock formation or the related drop of permeability or the displacement at the land surface. Moreover, the probabilistic framework implies the assessment of the uncertainty of the reservoir forecast which is important to estimate the reliability and the level of confidence of the model forecast for the future prediction.

To date data assimilation in geomechanical modeling has been seldom used. Reservoir compaction is estimated through a Bayesian inversion scheme from land subsidence observations in *Muntendam-Bos & Fokker* [2009]. In *Wilschut et al.* [2011] subsidence data measured over nine leveling campaigns are used with static pressure data to estimate fault trasmissibilities via the EnKF. In *Fokker et al.* [2013] an Ensemble Smoother (ES) algorithm with multiple data assimilation has been developed to constrain the reservoir parameters such as the compaction coefficient and the subsurface basement elastic modulus by assimilating ascending and descending Line-of-Sight

(Los) displacement observations provided by Persistent Scatterer Interferometry (PSI). The applicability and effectiveness of ES algorithm to constrain parameters uncertainties is investigated in *Baù et al.* [2014, 2015]. The ES is applied for a disc-shaped reservoir in a semi-infinite homogeneous porous medium by jointly assimilating horizontal and vertical land surface displacements. A significant improvement in the history match by incorporating PSI surface data is also obtained in *Katterbauer et al.* [2014] where an EnKF-based InSAR history matching framework is established, by integrating PSI surface displacement measurements with production data. An ES estimator with multiphase flow and geomechanics simulator as forward models are used by *Jha et al.* [2015] to assimilate pressure and PSI data. The inversion scheme is applied assuming a one-way coupling approach thus discarding the effect of the rock deformation on the fluid flow. By distinction, an inversion methodology is proposed in *Iglesias & McLaughlin* [2012] where geodetic and production data are jointly used in the more general framework of a fully coupling between geomechanics and subsurface flow. The numerical tests are promising with accurate estimation of the field properties. However, it is pointed out that estimating a set of parameters increases the ill-posedness of the problem with respect to the single-parameter case.

1.3 Summary

This introductive Chapter which gives a brief outline of the problem statement, study objectives, and state of the art is followed by the main body of the thesis that is organized in Chapters as follows.

- **Chapter 2** presents the FE geomechanical model used for the prediction of the land displacements as a major consequence to both UGS and gas production activities. The parametrization of the processes is described by isotropic and transversely isotropic constitutive models. The uniaxial vertical compressibility law implemented into the model is also briefly addressed.

- **Chapter 3** addresses the DA method. In particular, a general outlook of the ensemble-based algorithms is provided with a more detailed derivation of the

ES formulation. The description of the gPCE-based ES is also given as an alternative to the ensemble-based methods. The pros and cons of the two approaches are discussed.

- **Chapter 4** discusses the outcome from applying the methodology to UGS test applications and a real case. The numerical results are preceded by a wide description of the UGS reservoir and the available measurements from satellite radar images.

- **Chapter 5** investigates the calibration of the rock compressibility characterizing a production gas reservoir using measurements from time-lapse bathymetric surveys. The experiments are run for both the test and the real applications and the results from the ES and gPCE-based ES approaches are compared.

- **Chapter 6** concludes the thesis by emphasizing the major outcomes and the expected from future developments.

Chapter 2

Reservoir Geomechanics

The prediction of the subsurface compaction in producing hydrocarbon reservoirs is an important issue within the general reservoir management framework. Negative impacts such as casing deformations and wellbore failures [*Hilbert et al.*, 1999; *Fredrich et al.*, 2000; *Sayers et al.*, 2006] must be prevented to reduce significant economical risks and ensure the maximum safety of the drilling operations. Moreover, the forecast of land subsidence caused by compaction of the rock formation can be of major importance. Indeed, the anthropogenic settlements can cause the platform sinking, e.g., the Ekofisk field in the North Sea [*Kristiansen & Plischke*, 2010], pipeline excessive deformations, and large environmental impact especially in coastal areas [*Gambolati & Teatini*, 2015]. Although the predominant displacement is vertical, horizontal movements also occur giving rise to a fully three-dimensional (3D) process. Large horizontal displacements can have devastating effects on pipelines and other extensive surface structures unless they are designed to accommodate the strain [*Doornhof et al.*, 2006].

Geomechanical simulators are becoming more and more important to predict and monitor the risks during the field life and after its field abandonment as well. This thesis is concerned with two different reservoir activities that are simulated, i.e., the seasonal ground deformations due to the UGS activity and the land subsidence due to gas extraction from a deep hydrocarbon reservoir. Generally speaking, a reservoir is a formation where hydrocarbons, e.g. natural gas, are trapped within a porous rock matrix and prevented from the migrating by the top and bottom sealings. Figure 2.1 shows a sketch of a typical anticline reservoir structure with accumulation of oil and gas. This common geological structure blocks the hydrocarbon upward movement thus facilitating the accumulation on the top. Hydrocarbon reserves are mainly found

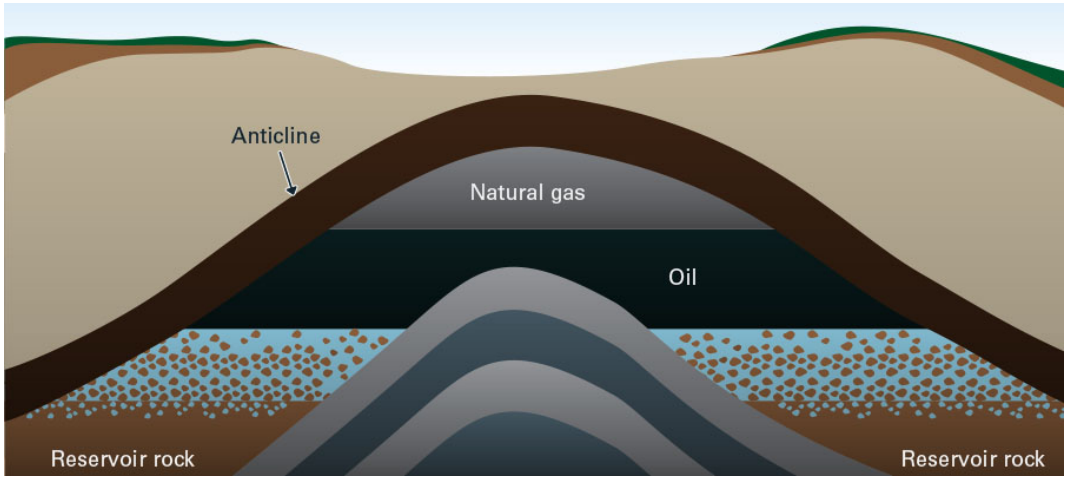

FIGURE 2.1: Typical structure of an anticline reservoir. The hydrocarbons are trapped by the impervious cap rocks (Source: Tortoise Capital Advisors).

in sedimentary rocks of clastic and chemical origin. In particular, in Italy, biogenic gas accumulates in the Plio-Quaternary foredeep which includes both the Adriatic Sea and the Po river Plain (Figure 2.2), where the reservoirs of the present case studies are located. The formation of this clastic sedimentary basin was enhanced by the interaction of the European and the African tectonic plates arising in the orogenesis of the Alpine and Apennine chains [*Encyclopaedia of Hydrocarbons*, 2005; *AAPG*, 2008, and references therein].

The geomechanical reservoir modeling aims at computing the stress and displacements fields generated by the pore pressure changes in space and time due to the reservoir activities, i.e., fluid extraction from or injection into the subsurface. The numerical solution of the governing flow and the structural partial differential equations (PDEs) are required to simulate the deformation up to the land surface. A Finite-Element model developed at the Department of Civil, Architectural and Environmental Engineering, University of Padova (Italy) is employed to compute the change of the stress/strain fields in the porous rock formations. The theory underlying the FE code, hereafter GEPS3D, is reviewed in this Chapter with main reference to the paper by *Gambolati et al.* [2001]. A detailed description of the implemented constitutive laws is also provided.

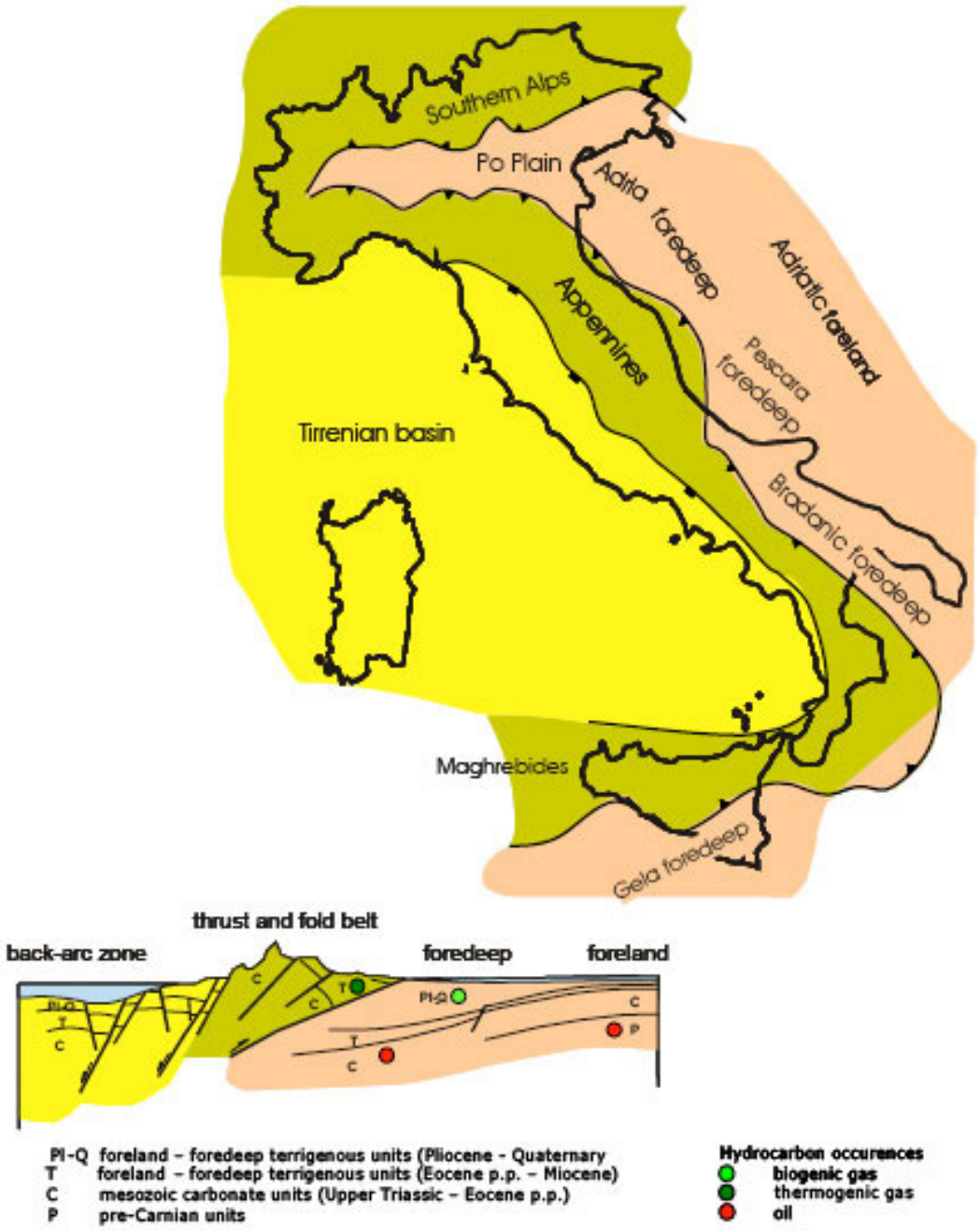

FIGURE 2.2: Hydrocarbon occurrences in different structural and tectonic settings in Italy (*after* AAPG *[2008]*).

2.1 Theory Overview

The process of settlement due to a load application within a porous matrix fully saturated with water is commonly referred to as *soil consolidation*. The generalization of the 1D consolidation theory [Terzaghi, 1923] is based upon the work by *Biot* [1941], who extended the theory to a 3D problem with time variable load. The aim is at describing the stress field distribution of a rock formation containing interconnected fluid-saturated pores by coupling the deformation of the soil matrix and the fluid flow. Two main phenomena of *solid-to-fluid* and *fluid-to-solid* couplings describe the behavior of the porous material. The change in applied stress to the material causes a change in the fluid pressure and, at the same time (quasi-static approximation), its variation may produce a change of the volume of the solid skeleton.

Let $\boldsymbol{u} = \{u_x, u_y, u_z\}$ be the displacement vector in the $\{x, y, z\}$ Cartesian reference system and $\epsilon = \epsilon_x + \epsilon_y + \epsilon_z$ the volume strain with $\epsilon_x, \epsilon_y, \epsilon_z$ the 1D strains along $\{x, y, z\}$. The incremental pore pressure is denoted with p. Within a gas reservoir, p takes the gas and water pressure components, namely p_g and p_w, into account according to *Bishop* [1959]:

$$p = S_g p_g + S_w p_w \tag{2.1}$$

with S_g and S_w the saturation index of water and gas, respectively. The equilibrium equation governing the phenomenon of consolidation reads

$$\nabla \cdot \boldsymbol{\sigma} - \alpha \nabla p = \boldsymbol{b} + \rho \boldsymbol{g} \tag{2.2}$$

with ∇ the gradient operator, $\boldsymbol{\sigma} = \{\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \tau_{xy}, \tau_{yz}, \tau_{zx}\}^T$ the effective stress vector, ρ the fluid density, $\boldsymbol{g}$ the gravity acceleration, $\boldsymbol{b}$ the external body forces and α the Biot coefficient defined as

$$\alpha = 1 - \frac{c_{br}}{c_{bm}} \tag{2.3}$$

where c_{bm} and c_{br} are the volumetric compressibility of the medium (the bulk) and the solid grains, respectively.

The water and gas dynamics in the porous material is usually described by the so-called *Black Oil Model* (e.g., *Peaceman* [1977]):

$$\nabla \cdot \left(\frac{\boldsymbol{k} k_{r,i}}{B_i \mu_i} (\nabla p_i - \rho_i g \nabla z) \right) + \frac{q_i}{\rho_{S,i}} = \frac{\partial}{\partial t} \left(\frac{\phi S_i}{B_i} \right) \tag{2.4}$$

where the subscript i refers to the water and gas phases with p_i denoting the respective pressure. $\boldsymbol{k}$ is the hydraulic conductivity tensor, ϕ the medium porosity, $k_{r,i}$ the relative permeability, S_i the saturation index, B_i is the volumetric deformation factor, μ_i is the viscosity, ρ_i and $\rho_{S,i}$ are the densities in the actual and standard conditions, respectively, q_i is the mass injection or production rate per unit reservoir volume, g the gravity acceleration and t is the time. In this study, the Equation (2.4) is solved with the aid of the commercial simulator $Eclipse^{TM}$ developed by Schlumberger. However, the aquifers hydraulically connected with the reservoir are not usually taken numerically into account in production forecast and the simulation extends only within the gas field boundary. To allow for the simulation of the lateral and bottom aquifers connected with the reservoir, a non-linear 3-D FE model is used thus numerically simulating the pressure propagation outside the so-called *analytical* boundaries. The continuity of the flow due to ground water movements is governed by the classic flow storage equation [*Verruijt*, 1969]:

$$\frac{1}{\gamma} \nabla (k \nabla p) = \phi \beta \frac{\partial p}{\partial t} + \frac{\partial \epsilon}{\partial t} + f \tag{2.5}$$

where γ and β are the specific weight and compressibility of water, respectively, ϕ is the porosity, k the saturated hydraulic conductivity, f a forcing function including possible flow sources and sinks, ∇ the gradient operator, and t the time. Equation (2.5) can be simply generalized to the case of deformable solid grains with compressibility c_{br} by [*Gambolati et al.*, 2000]:

$$\frac{1}{\gamma} \nabla (k \nabla p) = [\phi \beta + c_{br} (\alpha - \phi)] \frac{\partial p}{\partial t} + \alpha \frac{\partial \epsilon}{\partial t} + f \tag{2.6}$$

In this thesis, the uncoupled formulation or one-way coupling is used thus assuming that the mechanics-to-fluid coupling is negligibly weak. A sketch of the general

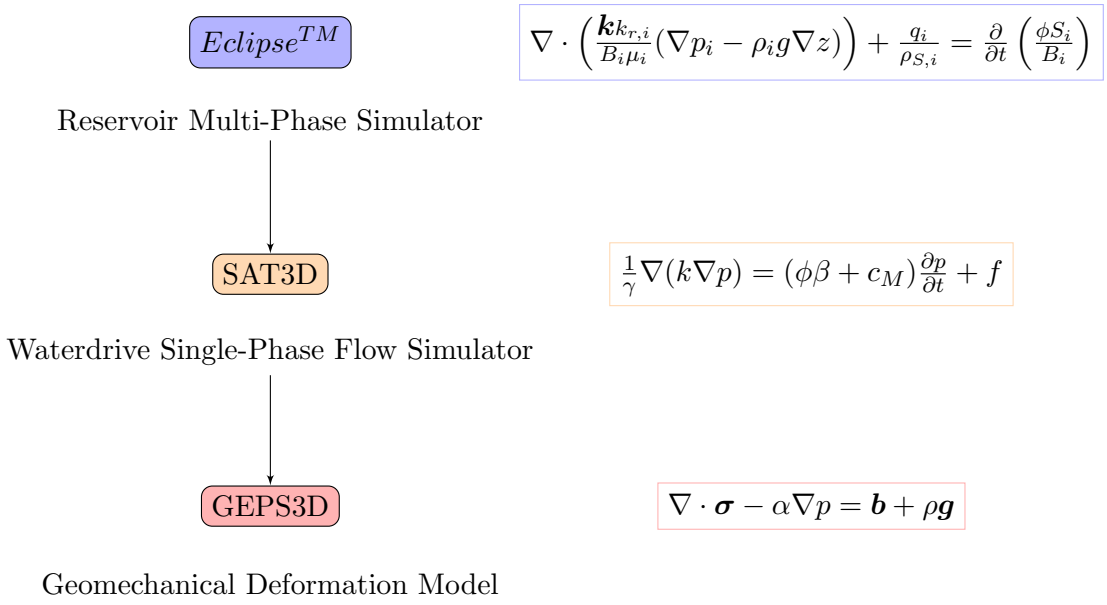

FIGURE 2.3: Sketch of the procedure used to model the Reservoir-Geomechanical system.

procedure implemented to model the reservoir-geomechanical system is shown in Figure 2.3. The pore pressure increment in space and time is first solved for the fluid flow dynamics and then used as an external source of strength in Equations (2.2) to solve the equilibrium equations. This simplified approach is generally acceptable over the time scale of practical interest [*Gambolati et al.*, 2000] and particularly justified for gas reservoirs, where the solid-to-flow coupling is usually weak because gas compressibility is much larger than rock compressibility [*Wang*, 2000]. For the purpose of this thesis, c_{br} is assumed to be negligibly small, thus dropping $c_{br}(\alpha - \phi)$ and implying $\alpha = 1$ in Equation (2.6). Following the uncoupled formulation, Equation (2.2) still holds and Equation (2.5) maybe shown to become

$$\frac{1}{\gamma} \nabla (k \nabla p) = (\phi \beta + c_M) \frac{\partial p}{\partial t} + f \tag{2.7}$$

where c_M is the uniaxial vertical compressibility representing the ratio between the relative compaction of a rock sample and the vertical load (in oedometric conditions). The coefficient $S_s = \phi \beta + c_M$ is the specific elastic storage that, in the uncoupled formulation, takes into account through the parameter c_M the pointwise deformation and discards the contribution from the remainder of the medium [*Gambolati et al.*, 2000].

2.2 Finite-Element Discretization

The equilibrium of a saturated porous volume Ω with surface Σ is prescribed by minimizing the total potential energy of the system. The necessary and sufficient conditions for equilibrium are given by the virtual work theorem, namely:

$$\int_{\Omega} \delta\boldsymbol{\epsilon}^T(\boldsymbol{\sigma} - \alpha \mathbf{i} p)\mathrm{d}\Omega = \int_{\Omega} \delta\boldsymbol{u}^T \boldsymbol{b}\mathrm{d}\Omega + \int_{\Sigma} \delta\boldsymbol{u}^T \hat{\boldsymbol{t}}\mathrm{d}\Sigma \tag{2.8}$$

where $\boldsymbol{\epsilon}^T = \{\epsilon_{xx}, \epsilon_{yy}, \epsilon_{zz}, \gamma_{xy}, \gamma_{yz}, \gamma_{zx}\}$, $\boldsymbol{\sigma}^T = \{\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \tau_{xy}, \tau_{yz}, \tau_{zx}\}$, and $\boldsymbol{u^T} = \{u_x, u_y, u_z\}$ are the strain, effective stress, and displacement vectors, p is the incremental pore pressure, α is the Biot coefficient, $\boldsymbol{i}^T = \{1, 1, 1, 0, 0, 0\}$ is the Kronecker delta in vectorial form, $\boldsymbol{b}$ and $\hat{\boldsymbol{t}}$ denotes the body forces and the total external forces acting per unit surface. The subscripts $\{x, y, z\}$ denotes the axes of a Cartesian reference system and the symbol δ indicates the virtual variables of displacement and strain. The integral formulation (2.8) must be supplemented with appropriate Dirichlet and Neumann boundary conditions:

$$\boldsymbol{u} = \hat{\boldsymbol{u}} \quad on \quad \Gamma_d$$

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = \hat{\boldsymbol{t}} \quad on \quad \Sigma$$

In the FE framework, displacements and pore pressure can be written as:

$$\boldsymbol{u} = \boldsymbol{N_u}\tilde{\boldsymbol{u}} \tag{2.9}$$

$$p = \boldsymbol{N_p}\boldsymbol{p} \tag{2.10}$$

where $\tilde{\boldsymbol{u}}$ and $\boldsymbol{p}$ are the nodal component vectors of displacements and pore pressure, and

$$\boldsymbol{N_u} = \sum_e \boldsymbol{N_u}^{(e)} \tag{2.11}$$

$$\boldsymbol{N_p} = \sum_e \boldsymbol{N_p}^{(e)} \tag{2.12}$$

are the matrices of basis functions with $\boldsymbol{N_u}^{(e)}$ and $\boldsymbol{N_p}^{(e)}$ the matrices of the nodal functions referred to a single element e. For example, using linear tetrahedral elements and

denoting by i, j, k, and l the vertices of each tetrahedron e, $\boldsymbol{N_u}^{(e)}$ and $\boldsymbol{N_p}^{(e)}$ are:

$$\boldsymbol{N_u}^{(e)} = \begin{bmatrix} \xi_i^{(e)} & 0 & 0 & \xi_j^{(e)} & 0 & 0 & \xi_k^{(e)} & 0 & 0 & \xi_m^{(e)} & 0 & 0 \\ 0 & \xi_i^{(e)} & 0 & 0 & \xi_j^{(e)} & 0 & 0 & \xi_k^{(e)} & 0 & 0 & \xi_m^{(e)} & 0 \\ 0 & 0 & \xi_i^{(e)} & 0 & 0 & \xi_j^{(e)} & 0 & 0 & \xi_k^{(e)} & 0 & 0 & \xi_m^{(e)} \end{bmatrix} \tag{2.13}$$

$$\boldsymbol{N_p}^{(e)} = \begin{bmatrix} \xi_i^{(e)} & \xi_j^{(e)} & \xi_k^{(e)} & \xi_m^{(e)} \end{bmatrix} \tag{2.14}$$

with the basis function $\xi_i^{(e)}$ referred to the i-node and e-element:

$$\xi_i^{(e)} = \frac{a_i + b_i x + c_i y + d_i z}{6|V_e|} \tag{2.15}$$

where

$$\begin{gathered} a_i = det \begin{bmatrix} x_j & y_j & z_j \\ x_k & y_k & z_k \\ x_l & y_l & z_l \end{bmatrix}, b_i = -det \begin{bmatrix} 1 & y_j & z_j \\ 1 & y_k & z_k \\ 1 & y_l & z_l \end{bmatrix} \\ c_i = det \begin{bmatrix} 1 & x_j & z_j \\ 1 & x_k & z_k \\ 1 & x_l & z_l \end{bmatrix}, d_i = -det \begin{bmatrix} 1 & x_j & y_j \\ 1 & x_k & y_k \\ 1 & x_l & y_l \end{bmatrix} \end{gathered} \tag{2.16}$$

and V_e is the element volume (positive or negative depending on nodal numeration):

$$V_e = \frac{1}{6} det \begin{bmatrix} 1 & x_i & y_i & z_i \\ 1 & x_j & y_j & z_j \\ 1 & x_k & y_k & z_k \\ 1 & x_l & y_l & z_l \end{bmatrix}$$

The basis functions for $\xi_j^{(e)}$, $\xi_k^{(e)}$, and $\xi_l^{(e)}$ are obtained from a proper index permutation. The strain components are related to the displacements according to:

$$\boldsymbol{\epsilon} = \boldsymbol{L}\boldsymbol{u} = \boldsymbol{L}\boldsymbol{N_u}\tilde{\boldsymbol{u}} = \boldsymbol{B}\tilde{\boldsymbol{u}} \tag{2.17}$$

with $\boldsymbol{L}$ an appropriate first-order linear operator. Replacing equations (2.11) through

(2.17) into Equation (2.8), and canceling the virtual displacements, yields the FE formulation of the elastic equilibrium for a saturated porous body:

$$\left(\int_{\Omega} \boldsymbol{B^T}\boldsymbol{\sigma} d\Omega\right) - \left(\int_{\Omega} \alpha \boldsymbol{B^T} \mathbf{i} \boldsymbol{N_p} d\Omega\right) \boldsymbol{p} = \int_{\partial\Omega} \boldsymbol{N_u^T}\boldsymbol{b} d\Omega + \int_{\Sigma} \boldsymbol{N_u^T}\hat{\boldsymbol{t}} d\Sigma \tag{2.18}$$

which is generally valid for any stress-strain relationship. In this study, a non-linear constitutive model is used where the incremental relation in terms of deformations and tension is given by

$$d\boldsymbol{\sigma} = \boldsymbol{D_t} d\boldsymbol{\epsilon} \tag{2.19}$$

where $\boldsymbol{D_t}$ is the tangent constitutive matrix. The equilibrium Equation (2.18) can be rewritten as

$$\int_{\Omega} \boldsymbol{B^T}\boldsymbol{\sigma} d\Omega = \boldsymbol{f} \tag{2.20}$$

with $\boldsymbol{\sigma}$ depending on the solution $\tilde{\boldsymbol{u}}$. This non-linear equation is solved using the Newton-Raphson method where the Jacobian matrix is:

$$\frac{d}{d\tilde{\boldsymbol{u}}}\left(\int_{\Omega} \boldsymbol{B^T}\boldsymbol{\sigma}(\tilde{\boldsymbol{u}}) d\Omega\right) = \int_{\Omega} \boldsymbol{B^T} \frac{d\boldsymbol{\sigma}(\tilde{\boldsymbol{u}})}{d\tilde{\boldsymbol{u}}} d\Omega = \left(\int_{\Omega} \boldsymbol{B^T}\boldsymbol{D_t}\boldsymbol{B} d\Omega\right) d\tilde{\boldsymbol{u}} \tag{2.21}$$

2.3 Constitutive Models

This section provides a description of the constitutive laws implemented into the FE discretization. The stress-strain relationship is given for a isotropic and transversely isotropic model. In fact, depending on the data available in the field, different models are employed. The isotropic formulation is deemed more adequate when available measurements are restricted to vertical displacements. By distinction, when land surface displacements are available for both the horizontal and vertical components, the transversely isotropic law is employed. In this case, the geomechanical properties vary along the two directions thus the parameter calibration is expected to enhance the match of the horizontal and vertical component of the displacement.

2.3.1 Isotropic Model

For an isotropic porous material, the constitutive relationship between the incremental effective stress $\boldsymbol{\sigma}$ and strain $\boldsymbol{\epsilon}$ vectors reads:

$$d\boldsymbol{\epsilon} = \boldsymbol{C} d\boldsymbol{\sigma} \Rightarrow \begin{Bmatrix} d\epsilon_{xx} \\ d\epsilon_{yy} \\ d\epsilon_{zz} \\ d\gamma_{xy} \\ d\gamma_{yz} \\ d\gamma_{zx} \end{Bmatrix} = \begin{bmatrix} \frac{1}{E} & -\frac{\nu}{E} & -\frac{\nu}{E} & 0 & 0 & 0 \\ \frac{\nu}{E} & \frac{1}{E} & -\frac{\nu_v}{E} & 0 & 0 & 0 \\ -\frac{\nu}{E} & -\frac{\nu}{E} & \frac{1}{E} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2(1+\nu)}{E} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{2(1+\nu)}{E} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{2(1+\nu)}{E} \end{bmatrix} \begin{Bmatrix} d\sigma_{xx} \\ d\sigma_{yy} \\ d\sigma_{zz} \\ d\tau_{xy} \\ d\tau_{yz} \\ d\tau_{zx} \end{Bmatrix} \tag{2.22}$$

where $\boldsymbol{C}$ is the deformation matrix, d$\boldsymbol{\epsilon}$ and d$\boldsymbol{\sigma}$ are the incremental strain and stress vectors, respectively, and $\boldsymbol{D}$ is constitutive matrix, i.e., the inverse of $\boldsymbol{C}$:

$$\boldsymbol{D} = \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} \begin{bmatrix} 1 & \frac{\nu}{1-\nu} & \frac{\nu}{1-\nu} & 0 & 0 & 0 \\ \frac{\nu}{1-\nu} & 1 & \frac{\nu}{1-\nu} & 0 & 0 & 0 \\ \frac{\nu}{1-\nu} & \frac{\nu}{1-\nu} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} \end{bmatrix} \tag{2.23}$$

with E and ν the Young and Poisson moduli, respectively. The coefficient $[E(1-\nu)]/[(1+\nu)(1-2\nu)]$ is the inverse of c_M. Medium isotropy is assumed whenever available observations for the parameter calibration consist of vertical displacements and subsurface compaction. However, anisotropy can be easily introduced as is described in the next section.

2.3.2 Transversely Isotropic Model

Transverse isotropy is more appropriate for the simultaneous match of both the vertical and horizontal displacements. For reference, the transversely isotropic implementation is described in details in *Janna et al.* [2012]. The stress-strain relationship for a

transversely isotropic medium reads:

$$d\boldsymbol{\epsilon} = \boldsymbol{C} d\boldsymbol{\sigma} \Rightarrow \begin{Bmatrix} d\epsilon_{xx} \\ d\epsilon_{yy} \\ d\epsilon_{zz} \\ d\gamma_{xy} \\ d\gamma_{yz} \\ d\gamma_{zx} \end{Bmatrix} = \begin{bmatrix} \frac{1}{E_h} & -\frac{\nu_h}{E_h} & -\frac{\nu_v}{E_h} & 0 & 0 & 0 \\ -\frac{\nu_h}{E_h} & \frac{1}{E_h} & -\frac{\nu_v}{E_h} & 0 & 0 & 0 \\ -\frac{\nu_v}{E_h} & -\frac{\nu_v}{E_h} & \frac{1}{E_v} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2(1+\nu_h)}{E_h} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{G_v} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{G_v} \end{bmatrix} \begin{Bmatrix} d\sigma_{xx} \\ d\sigma_{yy} \\ d\sigma_{zz} \\ d\tau_{xy} \\ d\tau_{yz} \\ d\tau_{zx} \end{Bmatrix} \tag{2.24}$$

where E_h, E_v, ν_h, and ν_v are Young moduli and Poisson ratios along the horizontal and vertical direction, respectively, and G_v is the shear modulus along the vertical direction. Thus, the constitutive law for a transversely isotropic medium depends upon five independent parameters, i.e., E_h, E_v, ν_h, ν_v, G_v. The following definitions are introduced:

- $\beta = E_h/E_v$, ratio of the horizontal-to-vertical Young modulus;
- $\gamma = \nu_v/\nu_h$, ratio of the horizontal-to-vertical Poisson ratio;
- $\theta = G_h/G_v$, ratio of the horizontal-to-vertical shear modulus;

where G_h is the horizontal shear modulus given by

$$G_h = \frac{E_h}{2(1+\nu_h)} \tag{2.25}$$

As a result, the constitutive matrix $\boldsymbol{D} = \boldsymbol{C}^{-1}$ is a function of

$$\boldsymbol{D} = \boldsymbol{D}(E_v, \beta, \nu_v, \gamma, \theta) \tag{2.26}$$

Assuming an isotropic Poisson ratio, i.e., ν_h=ν_v=ν, the parameter $\gamma = 1$ and the model calibration reduces to four parameters:

$$\boldsymbol{D} = \boldsymbol{D}(E_v, \beta, \nu, \theta) \tag{2.27}$$

The isotropy condition of ν is introduced following a sensitivity analysis on the parameter γ. A disc-shaped reservoir benchmark (Figure 2.4) is used for the purpose with

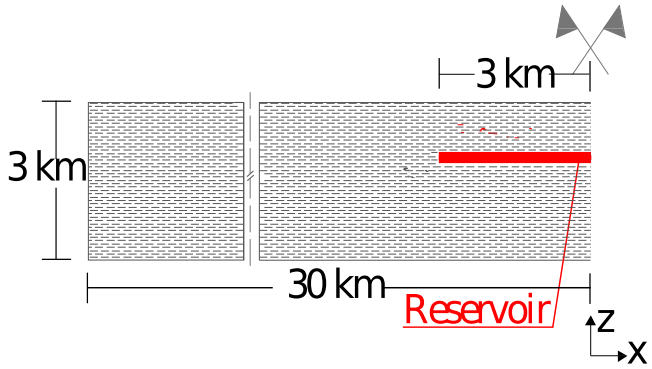

FIGURE 2.4: Geometry of the benchmark used for the sensitivity analysis of the transversely isotropic material.

TABLE 2.1: Geometric properties of the reservoir in Figure 2.4.

Radius	r	3000 m
Thickness	h	100 m
Depth	H	1000 m

the characteristic values provided in Table 2.1 and applying a linear time-dependent decreasing pore pressure. The effect of the γ values on the vertical and horizontal movements of the ground-surface is shown in Figure 2.5. Clearly, the influence of γ for the considered range of variation is not significant in terms of model response allowing to assume ν-isotropy in the real case application.

The values on the four parameters E_v, β, ν and θ are constrained by a set of thermodynamic conditions, which require the matrix $\boldsymbol{C}$ in Equation (2.24) to be positive

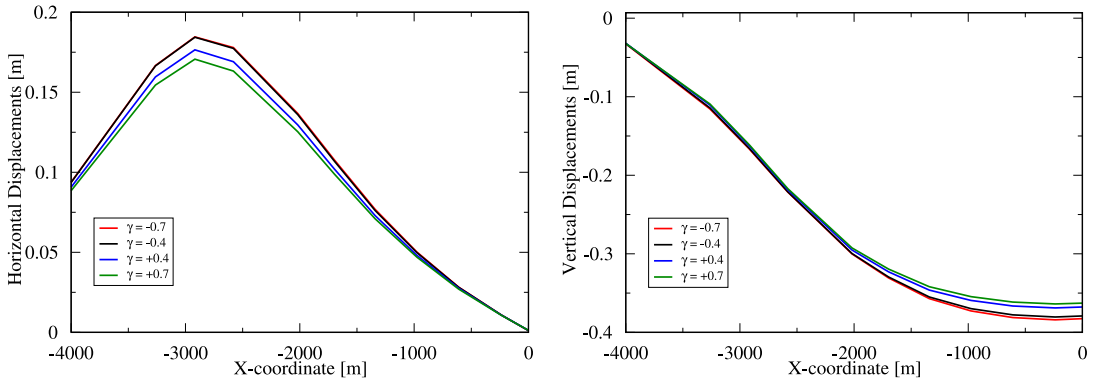

FIGURE 2.5: Variation of the horizontal (left) and vertical (right) displacements according to the variation of the parameter γ.

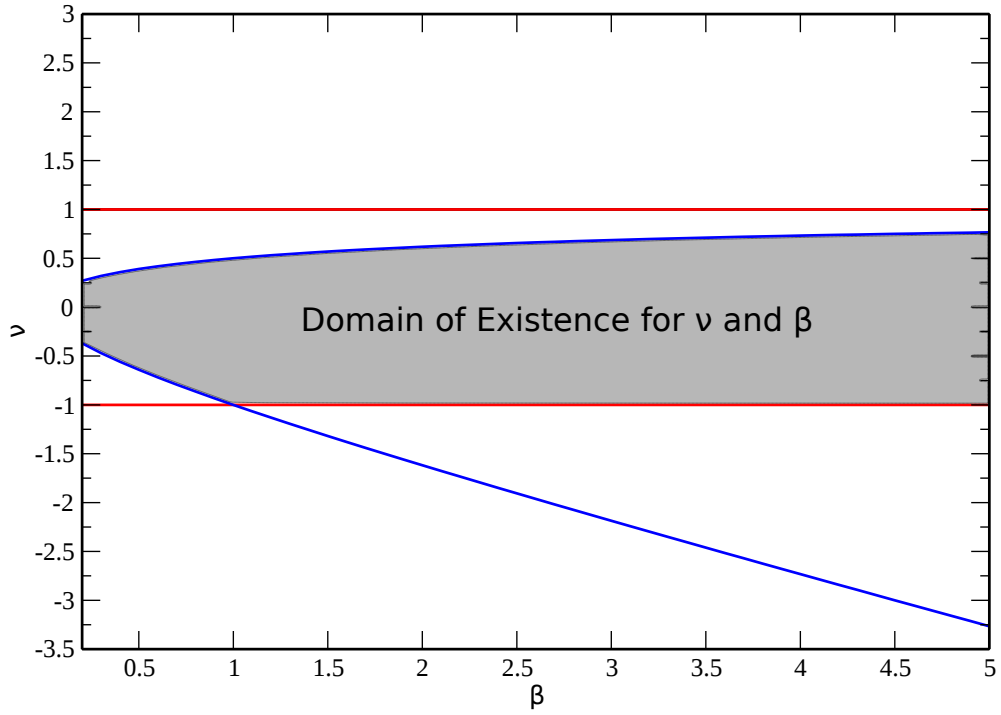

FIGURE 2.6: Domain of existence for the Poisson ratio ν as a function of the ratio β for the case in which the porous medium is considered isotropic with respect to the Poisson ratio itself.

definite. This yields:

$$E_h, E_v, G_v > 0 \tag{2.28a}$$

$$1 - \nu^2 > 0 \tag{2.28b}$$

$$1 - \nu - 2\frac{\nu^2}{\beta} > 0 \tag{2.28c}$$

Equation (2.28b) leads to the constraint:

$$-1 < \nu < 1 \tag{2.29}$$

whereas Equation (2.28c) can be rearranged to provide the interval:

$$(\beta/4)[-1 - \sqrt{1 + 8/\beta}] < \nu < (\beta/4)[-1 + \sqrt{1 + 8/\beta}] \tag{2.30}$$

Figure 2.6 shows the domain of existence for the Poisson ratio as a function of β. At

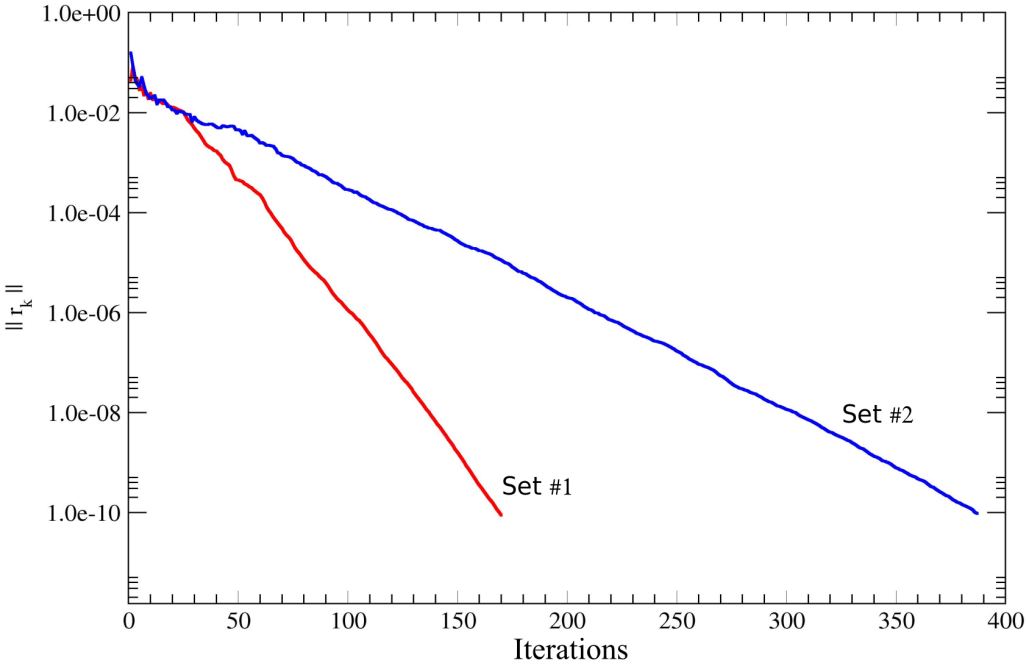

FIGURE 2.7: Convergence profiles with two input parameter sets.

TABLE 2.2: Parameter sets used to obtained the convergence profiles of Figure 2.7.

	β	ν	θ	ν_{min}	ν_{max}
Set #1	0.66	0.23	0.55	-0.76	0.43
Set #2	1.66	0.52	1.94	-1.00	0.58

this point, it is interesting to note the convergenge profiles of the FE model varying the parameter input. Indeed, a slower convergence is obtained if the parameters are close to the limiting values of Equations (2.28). As an example, Figure 2.7 shows the convergence profiles of the Preconditioned Conjugate Gradient (PCG) method employed as inner solver for the solution of Equation (2.20) and related to the parameters sets given in Table 2.2. These are two sets in terms of β, ν, and θ selected from an ensemble of realizations that are generated and further used in Chapter 4. ν_{min} and ν_{max} correspond to the lower and upper boundaries of the domain of existence in Figure 2.6 for the specific value of β. The closer the value of ν to ν_{min} or ν_{max}, the slower the convergence to the solution.

2.4 Reservoir compaction and expansion

Reservoir compaction can be conventionally described using the vertical uniaxial compressibility c_M [*van Hasselt*, 1992; *Baù et al.*, 2002; *Hueckel et al.*, 2005]. It represents the ratio between the relative compaction of a rock sample and the vertical load (in oedometric conditions):

$$c_M = \frac{1}{h}\frac{\Delta h}{\Delta \sigma_z} \tag{2.31}$$

where h represents the height of the compacting sample and Δ stands for finite increment. Numerically, this value is obtained as an approximation of $\frac{d\epsilon_z}{d\sigma_z}$, with ϵ_z and σ_z the vertical components of strain and stress, respectively. If isotropy conditions are assumed, c_M is related to the elastic parameters $\boldsymbol{D}$ in Equation (2.23) by the following relationship:

$$c_M = \frac{1}{E} \cdot \frac{(1+\nu)(1-2\nu)}{(1-\nu)} \tag{2.32}$$

By distinction, for transversely isotropy stress-strain relationship, c_M reads:

$$c_M = \frac{1}{E_v} \cdot \left(1 - \frac{2\nu_v^2}{1-\nu_h}\frac{E_v}{E_h}\right) \tag{2.33}$$

This can be easily derived from the computation of a sample theoretical vertical compaction with prevented lateral expansion under axial load. Writing Equation (2.24) in a cylindrical reference system:

$$d\boldsymbol{\epsilon} = \boldsymbol{C} d\boldsymbol{\sigma} \Rightarrow \begin{Bmatrix} d\epsilon_r \\ d\epsilon_\theta \\ d\epsilon_z \\ d\gamma_{\gamma_{rz}} \end{Bmatrix} = \begin{bmatrix} \frac{1}{E_h} & -\frac{\nu_h}{E_h} & -\frac{\nu_v}{E_h} & 0 \\ -\frac{\nu_h}{E_h} & \frac{1}{E_h} & -\frac{\nu_v}{E_h} & 0 \\ -\frac{\nu_v}{E_h} & -\frac{\nu_v}{E_h} & \frac{1}{E_v} & 0 \\ 0 & 0 & 0 & \frac{1}{G_v} \end{bmatrix} \begin{Bmatrix} d\sigma_r \\ d\sigma_\theta \\ d\sigma_z \\ d\tau_{\gamma_{rz}} \end{Bmatrix} \tag{2.34}$$

and prescribing the uniaxial oedometric test conditions ($d\epsilon_r = d\epsilon_\theta = d\gamma_{\gamma_{rz}} = 0$), the following holds:

$$d\tau_{\gamma_{rz}} = 0$$

$$d\sigma_r = d\sigma_\theta = \frac{\nu_v}{1-\nu_h} d\sigma_z \tag{2.35}$$

Substitution of Equation (2.35) into the expression of $d\epsilon_z$ in (2.34) gives $\frac{d\epsilon_z}{d\sigma_z} = c_M$ of Equation (2.33).

An overview of the methodologies used to estimate c_M is given in *Ferronato et al.* [2004, 2013]. Currently, the geomechanical parameter c_M can be evaluated from either laboratory tests and *in situ* field investigations. In the former category, core sample analysis are carried out through compaction devices which hopefully replicate the loading conditions on the reservoir specimens. However, such conditions can be hardly reproduced in the laboratory with the major consequence of a possible lack of reliability of the reservoir rock properties. On the other hand, in the 1990s and 2000s the radioactive-marker technique (RMT) was developed for the estimate of the reservoir compaction and rock properties [*Cassiani & Zoccatelli*, 2000; *Baù et al.*, 2002; *Kristiansen & Plischke*, 2010]. RMT evaluates the compaction by measuring the distance among a set of weakly radioactive bullets originally placed in a vertical well at fixed positions. Despite the promising results, the interpretation of the marker measurements is affected by a number of uncertainties and some warnings on their use are provided in *Ferronato et al.* [2003, 2004]. For example, the radioactive markers should be installed in non-producing wellbores because the presence of a horizontal pressure gradient can cause an underestimation of c_M. It is also shown that the knowledge of a detailed lithostratigraphy of the reservoir is needed to locate the radioactive markers and correctly interpret the corresponding measurements. The c_M values can be also indirectly assessed from the observation of ground movements using, for example, interferometric satellite measurements [*Ferronato et al.*, 2013].

For the Po river sedimentary sequence, the constitutive relationship c_M versus the vertical effective stress σ_z has been developed by *Baù et al.* [2002] and later improved by *Ferronato et al.* [2013] by using both RMT data and surface PSI measurements. For virgin loading conditions (I cycle) the following relationship is found (Figure 2.8):

$$c_{M,Icycle} = 1.0044 \cdot 10^{-2} \sigma_z^{-1.1347} \tag{2.36}$$

where σ_z is in MPa and c_M in MPa^{-1}. Note that compressive stress is taken with the positive sign. The validity range of Equation (2.36) can be defined from $\sigma_z = 10$ MPa to

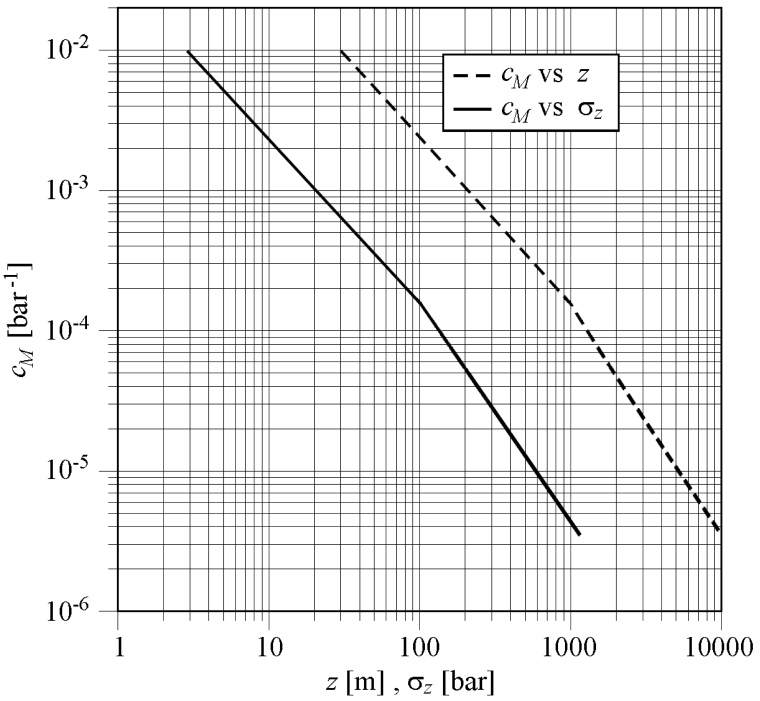

FIGURE 2.8: Vertical uniaxial compressibility law developed for the Northern Adriatic Basin, Italy (after [*Baù et al.*, 2002]).

$\sigma_z = 80$ MPa that is, approximately for z between 900 m and 7000 m before the inception of pumping. However, UGS mainly involves operations in unloading/reloading conditions (II cycle) because this activity is often performed at the end of the primary production, i.e., when the vertical stress is smaller than the maximum value ever experienced by the reservoir. Since during the II cycle the reservoir rock typically exhibits an elastic behavior, the additional parameter s is introduced to account for the ratio between c_M in I cycle and II cycle at the loading inversion:

$$s = \frac{c_{M,I}}{c_{M,II}} \exp[f_1(\sigma_z) - f_2(\sigma_z)] \tag{2.37}$$

where $f_1(\sigma_z)$ and $f_2(\sigma_z)$ come out of the integration of the one-dimensional compressibility, defined as

$$c_M = \frac{1}{1+e}\frac{\mathrm{d}e}{\mathrm{d}\sigma_z} \tag{2.38}$$

with e the void ratio of the medium. The value of s, which varies with the vertical effective stress σ_z, has been evaluated by *Baù et al.* [2002] using RMT data and recently confirmed by *Ferronato et al.* [2013] using InSAR data. It is worth noting that although

β, ν, θ and s are constant in the geomechanical simulations presented in this thesis, the geomechanical parameters E_h, E_v, G_h and G_v are heterogeneous and vary according to the dependency of c_M on σ_z. From the analysis of the in-situ expansion measurements from RMT, *Baù et al.* [2002] obtained the following relationship for $c_{M,II}$ in loading/unloading conditions (II cycle):

$$c_{M,IIcycle} = 2.9087 \cdot 10^{-4} \sigma_z^{-0.4315} \tag{2.39}$$

The validity range is from $\sigma_z = 37$ MPa to $\sigma_z = 65$ MPa, that is for z in between 3200 and 6800 m in undisturbed conditions. Note that s is defined at the load inversion point, i.e., when σ_z equals the largest stress $\sigma_{z,max}$ ever experienced in the reservoir. The FE implementation of the constitutive matrices (2.23) and (2.24) employies the non-linear compressibility law of Equation (2.36) and the parameter s in Equation (2.37) to obtain $c_{M,II}$ at the load inversion point, that is when $\sigma_z = \sigma_{z,max}$.

Chapter 3

Data Assimilation

Data Assimilation (DA) methods aim at constraining the forecast solution of a mathematical model based on spatio-temporal observations collected from the response of a dynamical system. The approximation of the true system state is achieved including the information from the available data and using the prior knowledge of the system, e.g, boundary and initial conditions, physical parameters constraints, etc. Figure 3.1 shows a schematic example of a DA analysis applied to the solution of a dynamical model. The forecast, i.e., the prediction of the future state of the system, is corrected by minimizing a penalty or misfit function at the times where observations are available. In this framework, three main ingredients are fundamental for the analysis or update:

- a mathematical model to simulate the physical process of interest;
- a set of observation data;
- a suitable algorithm to incorporate these data into the model response.

DA was first developed for dynamic meteorology and physical oceanography in the early 1980s. The main literature references on this topic area are provided in *Daley* [1991]; *Lorenc* [1986]; *Ghil* [1989] with an interesting review by *Bouttier & Courtier* [1999]. Successively, DA has been adopted in other disciplines of the geosciences such as hydrology and geology. In petroleum engineering, DA gained attention in the early 2000s for reservoir *history matching* to estimate the unknown parameters such as porosity and permeability fields in oil and gas reservoir applications [*Nævdal et al.*, 2003; *Lorentzen et al.*, 2003; *Aanonsen et al.*, 2011].

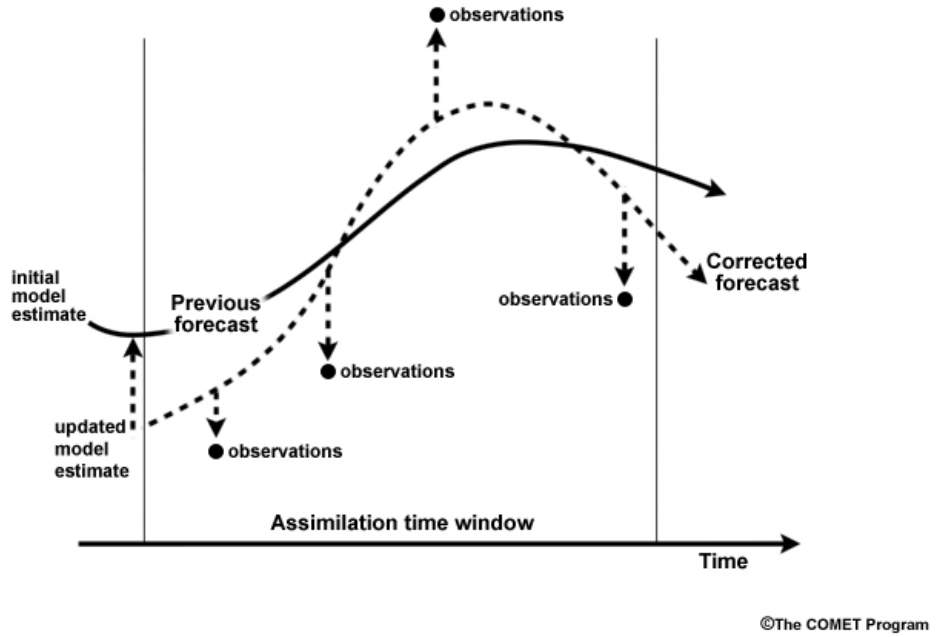

FIGURE 3.1: General analysis scheme adopted in Data Assimilation. The forecast state is updated using the available observations and producing a corrected forecast.

Several DA methods have been developed and are available in literature. A major classification of these methods usually distinguishes between *sequential* and *non-sequential* updates. The first class considers the real-time assimilation of observations with only past and present data used to update the model state. The Kalman Filter (KF) [*Kalman*, 1960] belongs to this class and the updated solution is optimal in the mean square error (MSE) sense assuming linear dynamics and Gaussian unbiased priors [*van Leeuwen & Evensen*, 1996]. When non-linear dynamics are modeled the approximate linearized equations are employed in the Extended version of the Kalman Filter (EKF). Another sequential method for nonlinear applications is the ensemble Kalman Filter (EnkF) [*Evensen*, 1994] where the error statistics are approximated based on an ensemble approach.

By distinction, the complete set of available observations is assimilated within the *non-sequential* DA framework. This class includes the generalized inverses or smoothers which can be derived as variational methods. These *non-sequential* schemes are usually

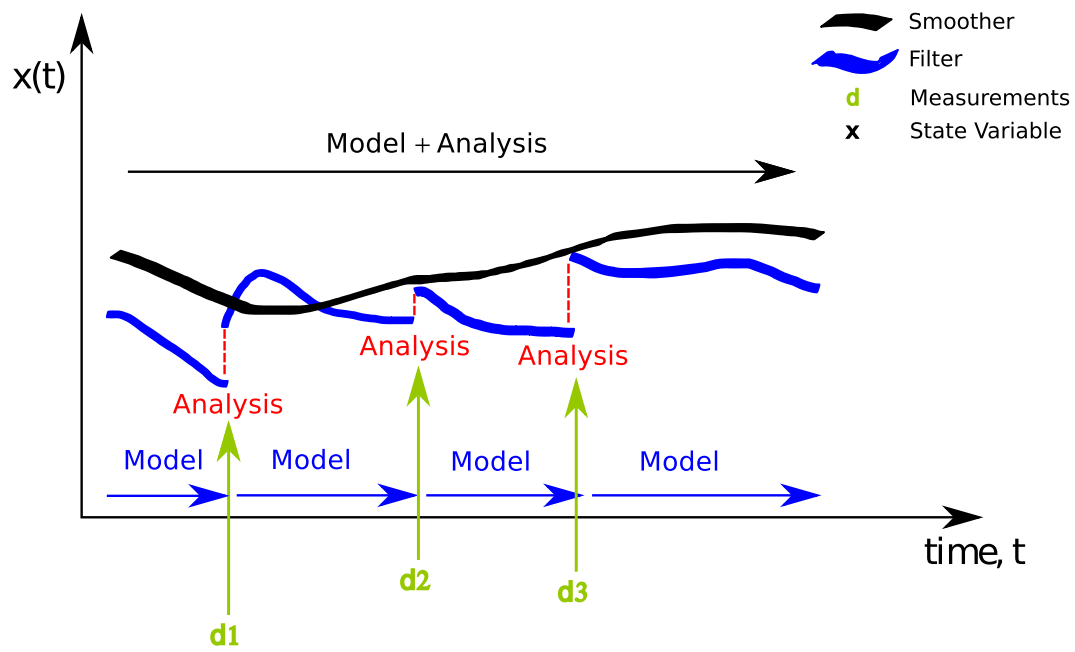

FIGURE 3.2: *Sequential* and *non-sequential* analysis schemes.

applied to estimate the state variables including the information from all collected observations. In addition, the estimation of parameters and boundary conditions of a problem are more suitable pursued with the *non-sequential* approaches because filter solutions are discontinuous in time at data points while smoothers have at most a discontinuous time derivative at data points [*van Leeuwen & Evensen*, 1996]. An explicative example of the different approaches to obtain the filter and smoother solutions are shown in Figure 3.2. The *Model* indicates the forward problem that allows to make predictions of the state of the system by solving a system of partial differential equations (PDEs) discretized in space and time. By distinction, the *Analysis* solves an inverse problem by updating the state of the system and/or jointly estimating the model parameters. In the *sequential* approach, the analysis equation is solved as many times as the number of the available observations (filtering) while the *non-sequential* DA solves the problem just once in space and time (smoothing). In this thesis, the focus is at solving an inverse problem in terms of parameter estimation hence the smoothing approach is preferred.

In this chapter, the general solution of an inverse problem for parameter estimation is first derived as a variational optimization problem, i.e, minimizing an appropriate cost function, and showing that the solution corresponds to the Best Linear Unbiased

Estimator (BLUE). This preliminary review is given to introduce the methodology used in this thesis for joint parameter and state estimation following the formulation of *van Leeuwen & Evensen* [1996] and *Evensen* [2009]. The Ensemble Smoother (ES) algorithm is derived in a Bayesian framework assuming linear model and Gaussian statistics. Moreover, some basics on the gPCE-based ES is provided in the last section.

3.1 Least-square analysis equations

Traditionally, the *inverse problem* is meant as a particular case of *parameter inverse problem* where one seek the estimated model parameters that are as close as possible to the prior values and which result in a model solution which is as close as possible to a set of measurements [*Evensen*, 2009]. Thus the inverse modeling problem consists of using a set of measurements to infer the values of the parameters that characterize the system [*Tarantola*, 2005]. A general relationship between the model variables $\boldsymbol{m}$ and the observable variables or data $\boldsymbol{d}$ can be expressed as [*Oliver et al.*, 2008]

$$\boldsymbol{d} = g(\boldsymbol{m}) \tag{3.1}$$

As stated before, the solution of the system of equations resulting from the discretization of Equation (3.1) is called the *forward* problem. On the other hand, if measurements $\boldsymbol{d_o}$ are available, the *inverse* problem seeks a solution for $\boldsymbol{m}$:

$$\boldsymbol{d_o} = g(\boldsymbol{m}) + \epsilon_o \tag{3.2}$$

where ϵ_o is the measurement error (white noise).

Assuming the case of linear inverse problems, Equation (3.1) relating the calculated data $\boldsymbol{d}$ and the model variables $\boldsymbol{m}$ can be written as

$$\boldsymbol{d} = \boldsymbol{G}\boldsymbol{m} \tag{3.3}$$

where $\boldsymbol{m}$ is a M-dimensional column vector, $\boldsymbol{d}$ is a N_d-dimensional column vector, and $\boldsymbol{G}$ is a $M \times N_d$ matrix. Thus, the inverse problem is

$$\boldsymbol{d_o} = \boldsymbol{G}\boldsymbol{m} \tag{3.4}$$

Some problems of this form has no solution if there exists $\boldsymbol{d_o}$ for which the equation has no solution. In this case, it is reasonable to seek a least-square type solution. Therefore, the inversion problem is solved as a variational minimization problem defining the objective function J as

$$J(\boldsymbol{m}) = \frac{1}{2}||(\boldsymbol{d} - \boldsymbol{d_o})||_2^2 = \frac{1}{2}||(\boldsymbol{G}\boldsymbol{m} - \boldsymbol{d_o})||_2^2 \tag{3.5}$$

where $||\,||_2$ is the l^2-norm or Euclidean norm. The minimization of the misfit J between predicted and observed data is obtained applying the gradient operator:

$$\nabla_{\boldsymbol{m}} J = \nabla_{\boldsymbol{m}} \left(\frac{1}{2}||\boldsymbol{G}\boldsymbol{m} - \boldsymbol{d_o}||_2^2 \right) = \nabla_{\boldsymbol{m}} \left(\frac{1}{2}(\boldsymbol{G}\boldsymbol{m} - \boldsymbol{d_o})^T(\boldsymbol{G}\boldsymbol{m} - \boldsymbol{d_o}) \right) \tag{3.6}$$

where T is the transpose. From simple linear algebra it follows that

$$\nabla_{\boldsymbol{m}} J = \boldsymbol{G}^T\boldsymbol{G}\boldsymbol{m} - \boldsymbol{G}^T\boldsymbol{d_o} \tag{3.7}$$

To minimize J the gradient has to be imposed equal to zero thus yielding to the linear system

$$\nabla_{\boldsymbol{m}} J = 0 \rightarrow \boldsymbol{G}^T\boldsymbol{G}\boldsymbol{m} = \boldsymbol{G}^T\boldsymbol{d_o} \tag{3.8}$$

with the solution for $\boldsymbol{m}$ given by the *Normal Equations*:

$$\boldsymbol{m} = (\boldsymbol{G}^T\boldsymbol{G})^{-1}\boldsymbol{G}^T\boldsymbol{d_o} \tag{3.9}$$

However, if $\boldsymbol{G}^T\boldsymbol{G}$ is singular, the least-square solution of Equation (3.9) is undefined.

In this sense, regularization is needed to improve the well-posedness of the problem and reduce the degrees of freedom of the system. This can be achieved introducing

the available prior information on the parameters $\boldsymbol{m_{prior}}$ and the structure of the uncertainty distribution in terms of:

- $\boldsymbol{C_d}$ covariance matrix of the error observation vector;
- $\boldsymbol{C_m}$ covariance matrix of the prior model parameters.

We assume Gaussian distributions for the observation error ϵ_o with zero mean and standard deviation σ_{ϵ_o}. The observation are uncorrelated so that the expectation $\mathbb{E}[\epsilon_{o,i}, \epsilon_{o,j}]$ is equal to 0. Introducing in Equation (3.6) the covariance terms and the regularization matrix defined as

$$\boldsymbol{R}(\boldsymbol{m}) = \frac{1}{2}(\boldsymbol{m} - \boldsymbol{m_{prior}})^T {\boldsymbol{C_m}}^{-1}(\boldsymbol{m} - \boldsymbol{m_{prior}}) \tag{3.10}$$

the misfit function reads:

$$J(\boldsymbol{m}) = \frac{1}{2}(\boldsymbol{Gm} - \boldsymbol{d_o})^T {\boldsymbol{C_d}}^{-1}(\boldsymbol{Gm} - \boldsymbol{d_o}) + \frac{1}{2}(\boldsymbol{m} - \boldsymbol{m_{prior}})^T {\boldsymbol{C_m}}^{-1}(\boldsymbol{m} - \boldsymbol{m_{prior}}) \tag{3.11}$$

Again, the minimization of $J(\boldsymbol{m})$ is obtained imposing the gradient equal to 0 or, equivalently, solving the Newton's method:

$$\mathcal{H}(J(\boldsymbol{m}))\Delta\boldsymbol{m} = -\nabla J(\boldsymbol{m}) \tag{3.12}$$

where $\Delta\boldsymbol{m} = (\boldsymbol{m} - \boldsymbol{m_{prior}})$ and ∇ and $\mathcal{H}$ are the Gradient and Hessian operators given by:

$$\nabla J(\boldsymbol{m}) = \boldsymbol{G}^T {\boldsymbol{C_d}}^{-1}(\boldsymbol{Gm} - \boldsymbol{d_o}) + {\boldsymbol{C_m}}^{-1}\Delta\boldsymbol{m} \tag{3.13}$$

$$\mathcal{H}(J(\boldsymbol{m})) = \boldsymbol{G}^T {\boldsymbol{C_d}}^{-1}\boldsymbol{G} + {\boldsymbol{C_m}}^{-1} \tag{3.14}$$

Substituting Equations 3.13 and 3.14 in Equation 3.12 leads to:

$$\left(\boldsymbol{G}^T {\boldsymbol{C_d}}^{-1}\boldsymbol{G} + {\boldsymbol{C_m}}^{-1}\right)\Delta\boldsymbol{m} = -\left(\boldsymbol{G}^T {\boldsymbol{C_d}}^{-1}(\boldsymbol{Gm_{prior}} - \boldsymbol{d_o})\right) \tag{3.15}$$

which is solved for the vector $\boldsymbol{m}$ as follows:

$$\boldsymbol{m} = \boldsymbol{m_{prior}} + \left(\boldsymbol{G}^T {\boldsymbol{C_d}}^{-1}{\boldsymbol{C_m}}^{-1}\right)^{-1} \boldsymbol{G}^T {\boldsymbol{C_d}}^{-1}(\boldsymbol{d_o} - \boldsymbol{Gm_{prior}})$$

$$= \boldsymbol{m_{prior}} + \boldsymbol{C_m G}^T(\boldsymbol{G C_m G}^T + \boldsymbol{C_d})^{-1}(\boldsymbol{d_o} - \boldsymbol{G m_{prior}}) \tag{3.16}$$

Notice that the solution obtained as variational minimization problem is equivalent to the Best, i.e., lower variance, Linear Unbiased Estimate (BLUE) analysis, that is, the optimal least-squares estimator, see proof in [*Bouttier & Courtier*, 1999].

Defining $\boldsymbol{x}^a$ and $\boldsymbol{x}^b$ the analysis and background model states, respectively, $\boldsymbol{y}$ is the observations vector, the optimal least-squares estimator is defined by the following analysis equation:

$$\boldsymbol{x}^a = \boldsymbol{x}^b + \boldsymbol{K}(\boldsymbol{y} - \boldsymbol{H}[\boldsymbol{x}^b]) \tag{3.17}$$

where $\boldsymbol{H}$ is the observation operator and $\boldsymbol{K}$ the Kalman gain, i.e., the weight matrix, given by:

$$\boldsymbol{K} = \boldsymbol{B H}^T(\boldsymbol{H B H}^T + \boldsymbol{R})^{-1} \tag{3.18}$$

where $\boldsymbol{B}$ and $\boldsymbol{R}$ are the covariance matrices of background errors and observation errors. The analysis of Equation (3.17) is optimal, i.e., it is a minimum variance estimate, if the following assumptions hold [*Bouttier & Courtier*, 1999]:

- linearity for the observation operator $\boldsymbol{H}$;
- $\boldsymbol{B}$ and $\boldsymbol{R}$ are s.p.d. matrices;
- unbiased errors, i.e., expectation equal to zero;
- uncorrelated background and observations errors;
- linear analysis;

The covariance matrix $\boldsymbol{C}$ of the analysis errors reads:

$$\boldsymbol{C} = (\boldsymbol{I} - \boldsymbol{K H})\boldsymbol{B}(\boldsymbol{I} - \boldsymbol{K H})^T + \boldsymbol{K R K} \tag{3.19}$$

3.2 Ensemble-based algorithms: the Ensemble Smoother

The smoothing approach estimates a set of state variables conditioned to the whole set of available measurements with the updating algorithm run only once after the model ensemble simulation. The ES algorithm founds its basis in the work of *van Leeuwen &*

Evensen [1996]. In this thesis, the ES is derived following the Bayesian formalism formulated by *Evensen & van Leeuwen* [2000] as a linear variance minimization analysis. The Bayesian approach thus subsumes the classical approach (see Section 3.1) to inversion while providing rigorous quantification of uncertainty of predictions, given clear assumptions on the prior and noise probabilities [*Iglesias et al.*, 2013b].

The generalized inverse problem provides an estimation of the uncertain model state, $\boldsymbol{\psi}$(**x**,t), and parameters, $\boldsymbol{\alpha}$(**x**), conditioned to a set of available measurements, $\boldsymbol{d}$(**x**,t), where **x** represents the coordinate position vector, and t is the time.

For the sake of brevity, in the following formulation the dependence of the variables on **x** and t is omitted except where strictly necessary. All implied variables are aleatory variables characterized by a PDF.

Consider a nonlinear dynamical model operator $\mathcal{G}$ defined in the spatial domain $\mathcal{D}$ bounded by $\delta\mathcal{D}$. The model equation can be written as:

$$\frac{\delta \boldsymbol{\psi}}{\delta t} = \mathcal{G}(\boldsymbol{\psi}, \boldsymbol{\alpha}) + \eta \tag{3.20}$$

with initial conditions

$$\boldsymbol{\psi}(\boldsymbol{x}, t_0) = \boldsymbol{\psi_0}(\boldsymbol{x}) + a(\boldsymbol{x}) \tag{3.21}$$

and boundary conditions

$$\boldsymbol{\psi}(\boldsymbol{x},t)|_{\delta\mathcal{D}} = \boldsymbol{\psi_b}(\boldsymbol{x},t) + b(\boldsymbol{x}, t) \tag{3.22}$$

where η, a and b are the random errors associated with the dynamical model (3.20), the initial conditions (3.21), and the boundary conditions (3.22), respectively.

When observations are available, the measurement equation takes the form

$$\boldsymbol{d} = \mathcal{H}(\boldsymbol{\psi^t}) + \boldsymbol{\epsilon} \tag{3.23}$$

where $\boldsymbol{\epsilon}$ is the observations error and $\mathcal{H}$ is the observation function relating the true model state $\boldsymbol{\psi^t}$ to the observations.

The ES equation is derived using Bayes' rule and the concept of conditional probability:

$$f[(\boldsymbol{\psi}, \boldsymbol{\alpha})|\boldsymbol{d}] = \frac{f[(\boldsymbol{\psi}, \boldsymbol{\alpha})]f[\boldsymbol{d}|(\boldsymbol{\psi}, \boldsymbol{\alpha})]}{f(\boldsymbol{d})} \tag{3.24}$$

where $f\,[(\boldsymbol{\psi}, \boldsymbol{\alpha})|\boldsymbol{d}]$ is the joint PDF for the model state, $\boldsymbol{\psi}$, and parameters, $\boldsymbol{\alpha}$, given the observation vector $\boldsymbol{d}$; $f\,[\boldsymbol{d}|(\boldsymbol{\psi}, \boldsymbol{\alpha})]$ is the PDF of the data given the model evolution, i.e., likelihood function; $f[(\boldsymbol{\psi}, \boldsymbol{\alpha})]$ is the joint probability function of the model evolution and parameters. In practice, the denominator $f(\boldsymbol{d})$ in Equation (3.24) serves as a normalization factor in the expression. Using Bayes' theorem again, the joint PDF $f[(\boldsymbol{\psi}, \boldsymbol{\alpha})]$ is

$$f[(\boldsymbol{\psi}, \boldsymbol{\alpha})] = f(\boldsymbol{\alpha})f(\boldsymbol{\psi}|\boldsymbol{\alpha}) \tag{3.25}$$

and Equation (3.24) can be rewritten as:

$$f[(\boldsymbol{\psi}, \boldsymbol{\alpha})|\boldsymbol{d}] \propto f(\boldsymbol{\alpha})f(\boldsymbol{\psi}|\boldsymbol{\alpha})f[\boldsymbol{d}|(\boldsymbol{\psi}, \boldsymbol{\alpha})] \tag{3.26}$$

where the symbol $\propto$ indicates proportionality between the two equation members.

Following the notation of *Evensen* [2009], the above formulation is discretized in time with $\boldsymbol{\psi}$(t) at time interval t_i represented as $\boldsymbol{\psi}_i$. Assuming that the model and the measurement processes are first-order Markov processes, i.e., the state at the current time depends only on the state at the previous time in the model temporal discretization, we can define the PDFs for the model evolution in time and the observations as:

$$f(\boldsymbol{\psi}_1, ..., \boldsymbol{\psi}_K, \boldsymbol{\alpha}) \propto f(\boldsymbol{\alpha}) \prod_{i=1}^{K} f(\boldsymbol{\psi}_i|\boldsymbol{\psi}_{i-1}, \boldsymbol{\alpha}) \tag{3.27}$$

$$f[\boldsymbol{d}|(\boldsymbol{\psi}, \boldsymbol{\alpha})] = \prod_{j=1}^{J} f[\boldsymbol{d}_j|(\boldsymbol{\psi}_{i(j)},\ \boldsymbol{\alpha})] \tag{3.28}$$

where K is the total number of time steps in which the model state is defined and J is the number of discrete times where measurements are available, with J a subset of K. Thus $\boldsymbol{\psi}_{i(j)}$ represents the state variable at time t_i corresponding to the time t_j from the observation subset.

Under first-order Markov process hypothesis, Equation (3.26) can be written as

$$f[(\psi_1, ..., \psi_k, \boldsymbol{\alpha})|\boldsymbol{d}] \propto f(\boldsymbol{\alpha}) \prod_{i=1}^{K} f(\psi_i|\psi_{i-1}, \boldsymbol{\alpha}) \prod_{j=1}^{J} f[\boldsymbol{d}_j|(\psi_{i(j)}, \boldsymbol{\alpha})] \tag{3.29}$$

Assuming Gaussian PDFs for the right-hand side terms of Equation (3.29), the conditional PDF (3.26) can be written as

$$f[(\psi, \boldsymbol{\alpha})|\boldsymbol{d}] \propto \exp\left(-\frac{1}{2}\mathcal{J}[\psi, \boldsymbol{\alpha}]\right) \tag{3.30}$$

where $\mathcal{J}$ is the penalty function as defined in Equation (3.11) with $\boldsymbol{\alpha} = \boldsymbol{m}$, $\boldsymbol{d} = \boldsymbol{d_o}$ and $\psi = \boldsymbol{Gm}$. . In this case, the maximization of the conditional joint density $f[(\psi, \boldsymbol{\alpha})|\boldsymbol{d}]$ is equivalent to the minimization of $\mathcal{J}[\psi, \boldsymbol{\alpha}]$, which gives also the maximum likelihood solution for ψ and $\boldsymbol{\alpha}$ [*Evensen*, 2009]. In this thesis, although the forward model is not linear with respect to the model parameters, the ES approach which is based on the Gaussian assumption is used to approximate the posterior distribution of the model parameters. However, the weak non linearity of the forward solver and the limited number of estimated parameters suggests that the ES may provide an accurate solution to the inverse problem as well.

The maximum likelihood estimate is defined by the minimum of the penalty function $\mathcal{J}$. Defining $\boldsymbol{u}^a$ and $\boldsymbol{u}^f$ the augmented state and parameter vectors of the analysis and forecast steps, respectively, the minimizing solution for state and parameter estimation can be written as

$$\boldsymbol{u}^a = \boldsymbol{u}^f + \mathbf{C}_{uu}^f H^T \left(\mathbf{H}\mathbf{C}_{uu}^f \mathbf{H}^T + \mathbf{C}_{\epsilon\epsilon}\right)^{-1} \left(\boldsymbol{d} - \mathbf{H}\boldsymbol{u}^f\right) \tag{3.31}$$

where $\mathbf{C}_{\epsilon\epsilon}$ is the covariance of the measurement error ϵ and $\mathbf{C}_{uu}$ is the forecast covariance of the augmented state and parameters. In this case the observation operator $\mathbf{H}$ is linear. In matrix form, Equation (3.31) reads:

$$\boldsymbol{A}_t^a = \boldsymbol{A}_t^f + \boldsymbol{K}_t \cdot (\boldsymbol{D}_t - \boldsymbol{H} \cdot \boldsymbol{A}_t^f) \tag{3.32}$$

where $\boldsymbol{A}_t^f$ and $\boldsymbol{A}_t^a$ are the forecast and update ensemble matrices for the joint state and

parameters from time t_0 to t, respectively, and $\boldsymbol{H}$ is a matrix that maps the measurement locations into the model grid so that the product $\boldsymbol{H} \cdot \boldsymbol{A}_t^f$ holds the ensemble of the model results at the measurement locations. The matrix $\boldsymbol{D}_t$ holds the measurements perturbed using an ensemble of Gaussian noises. At the right-hand side of Equation (3.32), the residual $\boldsymbol{D}_t - \boldsymbol{H} \cdot \boldsymbol{A}_t^f$ defines the deviation between the forecast state and measured state at the measurement locations. This residual forms the basis for correcting the forecast ensemble [*Baù et al.*, 2014]. The degree of correction depends upon the uncertainty of both the forecast ensemble and the measurement data, which is contained in the Kalman gain matrix $\boldsymbol{K}_t$:

$$\boldsymbol{K}_t = \boldsymbol{C}_f \boldsymbol{H}^T (\boldsymbol{H} \boldsymbol{C}_f \boldsymbol{H}^T + \boldsymbol{R})^{-1} \tag{3.33}$$

where $\boldsymbol{C}_f$ is the ensemble forecast error covariance matrix and $\boldsymbol{R}$ is the measurement error covariance matrix.

The implementation of the ES algorithm used here follows the methodology presented by *Keppenne* [2000]. Rearranging Equation (3.32) as:

$$\boldsymbol{A}_t^a - \boldsymbol{A}_t^f = \boldsymbol{C}_f \boldsymbol{H}^T \cdot \boldsymbol{B} \tag{3.34}$$

where the matrix $\boldsymbol{B}$ is defined as:

$$\boldsymbol{B} = (\boldsymbol{H} \boldsymbol{C}_f \boldsymbol{H}^T + \boldsymbol{R})^{-1} \cdot (\boldsymbol{D}_t - \boldsymbol{H} \cdot \boldsymbol{A}_t^f) \tag{3.35}$$

each column of matrix $\boldsymbol{B}$ is calculated by solving the following linear system many times as the number of ensemble members:

$$(\boldsymbol{H} \boldsymbol{C}_f \boldsymbol{H}^T + \boldsymbol{R}) \boldsymbol{b} = (\boldsymbol{d}_t - \boldsymbol{H} \cdot \boldsymbol{u}_t^f) \tag{3.36}$$

where $\boldsymbol{b}$, $\boldsymbol{d}_t$ and $\boldsymbol{u}_t^f$ are the i_{th} columns of $\boldsymbol{B}$, $\boldsymbol{D}_t$ and $\boldsymbol{A}_t^f$, respectively.

3.3 Generalized Polynomial Chaos Expansion (gPCE)

DA via Monte Carlo based simulations requires a large number of realizations to effectively reduce the sampling error leading to high time consuming methods for large physical models [*Li & Xiu*, 2009]. By contrast, gPCE-based ES allows to compute the solution of the stochastic equations by gPC expansion and afterward an ensemble of realizations is sampled to be updated in the ES assimilation scheme of Equation (3.32). The assimilation method is illustrated assuming the case in which the stochastic model solution depends on a single random variable and multiple random variables.

Without loss of generality let us write the forward problem that solves this general equation:

$$\mathcal{G}(\boldsymbol{\psi}(\boldsymbol{x}, t), \boldsymbol{\alpha}(\boldsymbol{x})) = 0 \tag{3.37}$$

for the displacement field, $\boldsymbol{\psi}$. $\mathcal{G}$ is a nonlinear operator mapping from the pore pressure field to the displacements field, and $\boldsymbol{\alpha}$ is the model parameter vector. $\boldsymbol{\psi}$ depends on the space $\boldsymbol{x}$ and time t while $\boldsymbol{\alpha}$ is only a function of $\boldsymbol{x}$. The application is here focused on the prediction of $\boldsymbol{\psi}$ but, for a high fidelity model, the parameter $\boldsymbol{\alpha}$ should be inferred from observations. Unfortunately, from a given measurement of $\boldsymbol{\psi}$, the explicit form of $\boldsymbol{\alpha}$ is not straightforward. Therefore, the problem is first set in a probabilistic framework treating the model parameter as a random variable with an a-priori PDF. Equation (3.37) becomes:

$$\mathcal{G}(\boldsymbol{\psi}(\boldsymbol{x}, t, \omega), \boldsymbol{\alpha}(\boldsymbol{x}, \omega)) = 0 \tag{3.38}$$

where ω is the actual realization from the stochastic space. Naturally, in this way the field $\boldsymbol{\psi}$ is also a random variable, depending on the actual realization of ω. Considering a homogeneous field for the parameter $\boldsymbol{\alpha}$, i.e., a single random variable, and defining a mapping, so that everything can be written as a function of a random variable $\xi(\omega)$ with PDF f_ξ. With this simplification, the abstract description modifies to:

$$\mathcal{G}(\boldsymbol{\psi}(\boldsymbol{x}, t, \xi), \boldsymbol{\alpha}(\xi)) = 0 \tag{3.39}$$

To update this prior knowledge, one needs a probabilistic description of the vertical

displacements, a prediction from the prior knowledge to be compared with the measurement. Instead of sampling from $\boldsymbol{\alpha}$ and running the deterministic solver for each sample point (Monte Carlo approach), a surrogate model is used to approximate the response $\boldsymbol{\psi}$ in function of ξ with an M degree general PCE:

$$\psi(\boldsymbol{x}, t, \xi) \approx \tilde{\psi}_M(\boldsymbol{x}, t, \xi) = \sum_{i=0}^{M-1} \psi_i(\boldsymbol{x}, t)\Phi_i(\xi), \tag{3.40}$$

in which Φ_i are the polynomial basis functions.

The random variable ξ is uniformly distributed thus in this problem Φ_i are the univariate Legendre polynomials, which serves as an orthogonal basis with respect to the underlying probabilistic space defined by ξ:

$$\mathbb{E}[\Phi_j(\xi)\Phi_k(\xi)] = \int_{\Re} \Phi_j(\xi)\Phi_k(\xi) f_\xi \mathrm{d}\xi = h\delta_{jk}$$

where δ_{ij} is the Kroenecker delta, and h is the norm of the polynomials:

$$h = \mathbb{E}[\Phi_i(\xi)\Phi_i(\xi)]$$

In reality, for the application of Chapter 5, the simplification of supposing homogeneous $\boldsymbol{\alpha}$ proves to be unrealistic. When $\boldsymbol{\alpha}$ is a random field, and not a random variable, the problem is somewhat more complicated. However, with the help of the Karhunen-Loeve Expansion (KLE), $\boldsymbol{\alpha}$ can be written as a product of random variables and spatial functions, so that the stochastic and the spatial domain can be separated. Using the KLE, the computation boils down to the same problem, but with the number of the random variables equal to L, the non-truncated (spatial) eigenfunctions, and accordingly the gPCE, will take the same form as given in (3.40). However, the basis functions of the gPCE are multivariate polynomials given as a function of $\boldsymbol{\xi}(\omega) : \Omega \to \Re^L$, and $|i|$ is the multi-index (see details in [*Xiu*, 2010]):

$$\psi(\boldsymbol{x}, \boldsymbol{\xi}) \approx \tilde{\psi}_M(\boldsymbol{x}, \boldsymbol{\xi}) = \sum_{|i|<M} \psi_i(\boldsymbol{x})\Phi_i(\boldsymbol{\xi}), \tag{3.41}$$

The coefficient of the approximating gPCE can be calculated in different ways.

Here, we focus on non-intrusive methods, where the forward model does not have to be changed, and can be used in a black-box fashion. Among the possible non-intrusive methods we calculate here the coefficient with the pseudo spectral projection described as follows.

The target here is to find a best approximation of ψ in the form of (3.40) in the L_2 sense, that is, to minimize the L_2 norm of the error

$$\mathbb{E}\left[\left(\psi - \tilde{\psi}_M\right)^2\right].$$

This can be done by orthogonally projecting to the subspace spanned by the approximating polynomials Φ_i, which gives the coefficients:

$$\psi_i = \mathbb{E}[\psi(\boldsymbol{x}, \xi)\Phi_i(\xi)] = \int_{\Re} \psi(\boldsymbol{x}, \xi)\Phi_i(\xi) f_\xi \mathrm{d}\xi$$

The integral term can be approximated with a quadrature rule, which gives:

$$\psi_i \approx \sum_{i=1}^{Q} \psi(\boldsymbol{x}, \xi_i)\Phi_i(\xi_i) w_i$$

Here $\xi_i = \xi(\omega_i)$ is one realization of the random variable, i.e., the integration point, and w_i is the corresponding weight. ξ_i and w_i are defined by the quadrature rule determined by the weighting function f_ξ. In this case, the Gauss-Legendre rule is used. When an M degree gPCE approximation of ψ is to be calculated, then the polynomials Φ_i are at most of degree M, so the integrand is at most of degree $2M$. Accordingly, it is necessary to use a ($Q = M + 1$)-point quadrature rule. This way, determination of the proxi model is reduced to evaluating the forward model and the polynomials Φ_i at the integration points ξ_i, where the later is computationally very cheap. To give an example, for a 3 degree gPCE approximation, the proxi model coefficients can be evaluated from 4 deterministic solver calls, by evaluating ψ at $\boldsymbol{\alpha}(\xi_i)$.

Chapter 4

Underground Gas Storage Reservoirs[1]

4.1 Introduction

Natural gas is an importance source of energy to cope with the worldwide growing demand. At January 2014, the worldwide reserves are of about 200.576 billion cubic meters (bcm) as estimated and published by Cedigaz, the international association for natural gas. Underground Gas Storage (UGS) reservoirs are crucial facilities for different purposes. *Cedigaz* [2013] points out two different trends in the gas storage market. In emerging markets, UGS projects are still developing as in the past with UGS plants mostly dedicated to balance seasonal and peak needs. On the other hand, in mature natural gas countries, UGS serves new functions in addition to the operational ones. In fact, the new storage frontiers are related to trading activity and to using natural gas as back–up of intermittent renewable energy sources in electricity generation. Natural gas in storage serves also as strategic reserve for any occuring emergency. At the beginning of 2013, 688 UGS plants were operating worldwide and 256 projects were under construction or planned. The overall working gas capacity is of about 377 bcm with the higher percentage of working gas volume operated in North America [*Cedigaz*, 2013]. Depleted natural gas and oil fields, aquifers, and salt caverns are good candidates for storing natural gas. Worldwide, gas and oil reservoirs account for about 74% of these storage sites. Among all the possible operational activities of UGS, the seasonal and

[1]The topics covered in this chapter have also been partially presented in [*Zoccarato et al.*, 2016, under review].

peak demands are a major concern in this study. The natural gas is injected into the UGS fields during low-demand seasons and delivered during high-demand seasons or daily peaks. The overall impact on the land surface displacements may be significant and correlated with the seasonal management. In particular, injecting a fluid into the subsurface leads to the increase of the pore pressure and, consequently, induces, an expansion of the porous rock formation. In turn, the reservoir expansion yields a upheaval of the land surface, where horizontal and vertical displacements can be measured. Vice versa, the gas withdrawal leads to the contraction of the deep-gas bearing pools causing land settlement.

The seasonal changes of the stress and displacement fields should be predicted and monitored to avoid damages to the wells and/or the structures and infrastructures on the land surface. The accurate prediction of the amount and areal extent of settlement and upheaval is crucial to i) reduce the management uncertainties related to the UGS fields and ii) mitigate the environmental impacts [*Doornhof et al.*, 2006].

A geomechanical, transversely-isotropic, 3D finite-element (FE) model has been calibrated with a trial-and-error procedure by *Teatini et al.* [2011] to match both the vertical and horizontal seasonal displacements above the *Lombardia* UGS field situated in the Po River sedimentary basin (Italy). The available observations consist of vertical and horizontal movements measured by PSI using Radarsat scenes retrieved between 2003 and 2008 over the UGS field. In this thesis, the same *Lombardia* case study is used to investigate the capability and effectiveness of the ES to reduce the uncertainty on the knowledge of the deep geomechanical parameters. Initially, synthetic observations collected from a virtual observation-simulation system experiment (OSSE) are assimilated for algorithm testing purposes. Then, the ES is used to assimilate the real vertical and horizontal displacement measurements provided by PSI over the *Lombardia* UGS field.

This chapter is organized as follows. Sections 4.2 and 4.3 summarize the major properties of the *Lombardia* UGS site and the measurements available from PSI. The generation of the parameter ensemble and the Monte Carlo simulations used to generate the forecast ensemble is described in Section 4.4. The results obtained with the assimilation of both synthetic and real measurements are presented and discussed in Sections 4.5 and 4.6, respectively. Section 6 ends this chapter with some comments

about the further developments.

4.2 The *Lombardia* gas reservoir

The *Lombardia* gas field is situated in the Po river basin (north of Italy). The gas reservoir was developed in the period 1981-1986 and was used afterward for underground storage of methane gas imported from eastern European countries. The burial depth is between 1050 m and 1350 m below mean sea level. In a typical gas-storage-and-recovery (GSR) cycle, gas is injected from April to November and extracted from November until April (Figure 4.1). The evolution in time and space of the pressure change has been accurately predicted using a multi-phase fluid-dynamic model accounting for both the gas-bearing formations and the hydraulically connected aquifers [*Teatini et al.*, 2011]. The model has been calibrated over the primary production phase (1981-1986) and the following 1986-2008 UGS phase using traditional history matching techniques with pressure data from 32 wells. Due to the large amount of available information for the history matching of the reservoir model, the uncertainty on the flow field is assumed negligible compared with the uncertainty related to the geomechanical parameters. Thus, the pore pressure behavior in time and space is a deterministic input to the model. By distinction, a significant variability is allowed for the geomechanical model parameters, as further discussed in Section 4.4. For more details about the reservoir characteristics and its production model, see *Teatini et al.* [2011]. Using an explicitly (one way) coupled approach, the pressure changes have been used as strength sources within the geomechanical model.

The geomechanical model domain of the *Lombardia* gas field extends over a 60 km $\times$ 50 km horizontal area and down to a 10 km depth below the land surface. Boundary conditions are prescribed such that no displacement is allowed on the lateral and the bottom boundaries. The top of the domain, which represents the land surface, is modeled as a traction-free boundary. The simulation period spans from 1981 to 2008, thus including the primary production period (1981-1986), and the following UGS cycles (1987-2008). Figure 4.2 depicts the 3D finite element grid of the UGS reservoir adopted in the geomechanical model.

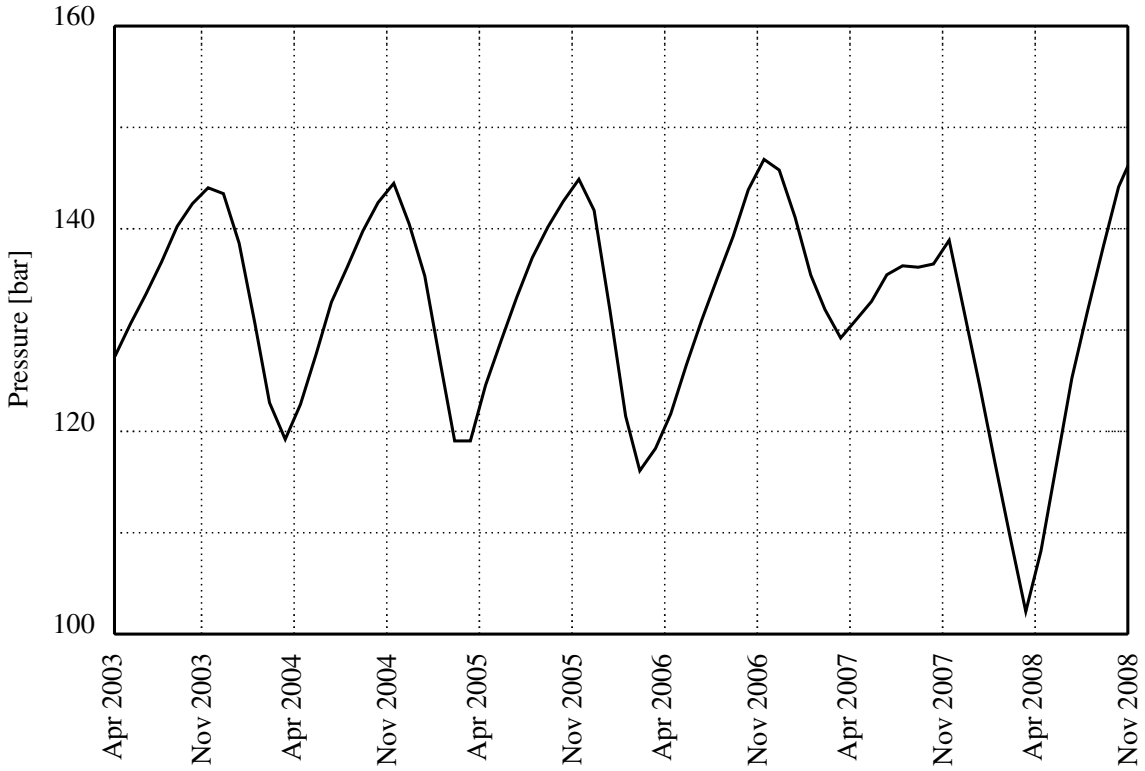

FIGURE 4.1: Average pore pressure versus time in *Lombardia* reservoir over the period April 2003-November 2008.

4.3 PSI measurements

PSI techniques can provide accurate measurements of the land surface movements over large areas and over long periods of time and can be effectively used to help characterize the development of deep oil and gas reservoirs [*Tamburini et al.*, 2010].

Over the *Lombardia* gas field, satellite radar images were acquired by RADARSAT-1 from March 2003 to October 2008 over ascending and descending orbits. These data have been processed to determine the spatial distribution as well as the magnitude and the timing of time-lapse land displacements both in the vertical and in the West-East horizontal directions over periods of about 6 months, corresponding to the injection and production phases of the reservoir. Further details on the methodology used to acquire, process and analyze the satellite radar images are provided in *Teatini et al.* [2011].

Figure 4.3 shows the displacement maps of over the *Lombardia* reservoir obtained by interpolating PSI records, which are provided at the locations represented by the blue

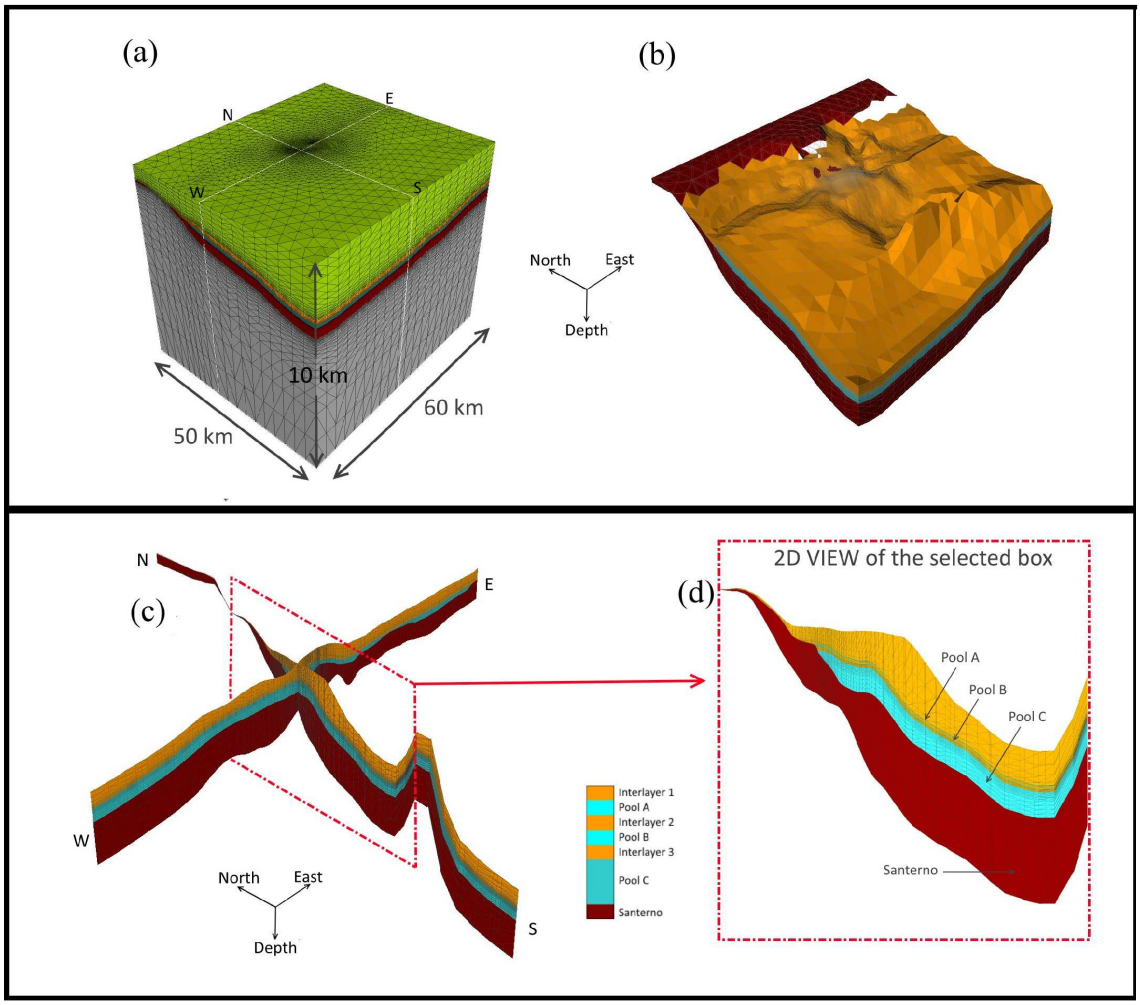

FIGURE 4.2: (a) Axonometric view of the 3-D FE grid used in the geomechanical model, (b) axonometric view of the productive layers of the reservoir, (c) North-South and West-East vertical cross sections of the FE grid and (d) 2D view of the geological units and pools of the North-South cross section. The vertical exaggeration is 5 in (a)-(c) and 10 in (d).

dots in Figure 4.4. The sub-panels in the left and right columns of Figure 4.3 provide the horizontal and vertical displacements over two representative injection and production cycles. Figure 4.3 highlights the effects of the expansion experienced by land during warm seasons when gas is stored into the reservoir, and the effects of contraction during the cold season, when gas is in turn recovered. Vertical displacements, u_v, are assumed positive during injection cycle, i.e. land surface uplift, and negative during the extraction cycle, i.e. subsidence. The horizontal displacement, u_h, is assumed negative if directed westward ($u_{h,W}$), and positive if directed eastward ($u_{h,E}$). In this study, the number of ensemble members is greater than the number of observations. This choice prevents the matrix ($\boldsymbol{HC}_f\boldsymbol{H}^T + \boldsymbol{R}$) of Equation (3.33) from being rank deficient [*Keppenne*, 2000]. On the other hand, because of the nature of the PSI records, the whole set of available measurements is characterized by a certain noise so that selecting a subset of measurements is generally appropriate to provide a set of more reliable data actually consistent with the simulated physical process. The selection methods are investigated with the results discussed in Section 4.5.3.

4.4 Model Set-Up

4.4.1 Generation of Parameter Ensembles

The goal of this study is to demonstrate the ability of the proposed data assimilation algorithm to reduce the uncertainty on the β, ν, θ, and s parameters by assimilating vertical and horizontal displacement measurements at a number of points on the land surface. To assess the assimilation potentiality, uniform PDFs are used to address the parameter uncertainty. According to a uniform PDF, the only conditions prescribed are the minimum and maximum values that a parameter can take on and each value in between has the same probability of occurrence. These limits are chosen based upon typical literature values or available information of the study area. Moreover, the values on these four parameters are constrained by the thermodynamic conditions given in Equations 2.28-2.30.

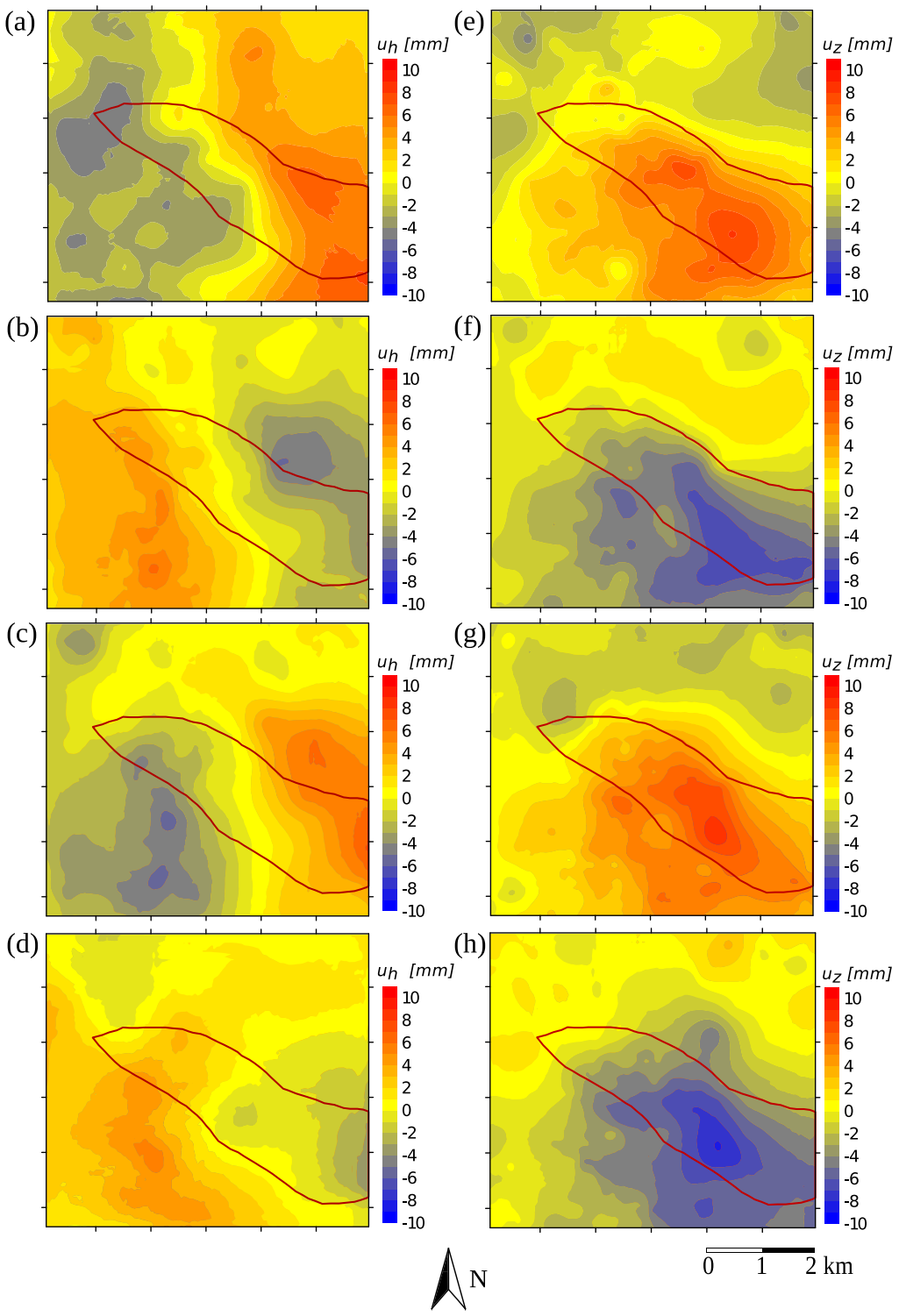

FIGURE 4.3: Maps of time-lapse displacements over the *Lombardia* reservoir obtained using the PSI technique over the period April 2003-April 2005. The subpanels a), b), c) and d) indicate horizontal displacements, u_h, at time intervals T1, T2, T3 and T4, respectively (see Table 4.1); whereas e), f), g) and h) indicate vertical displacements, u_z at same time intervals. The trace of the UGS field is marked by the dark-red line.

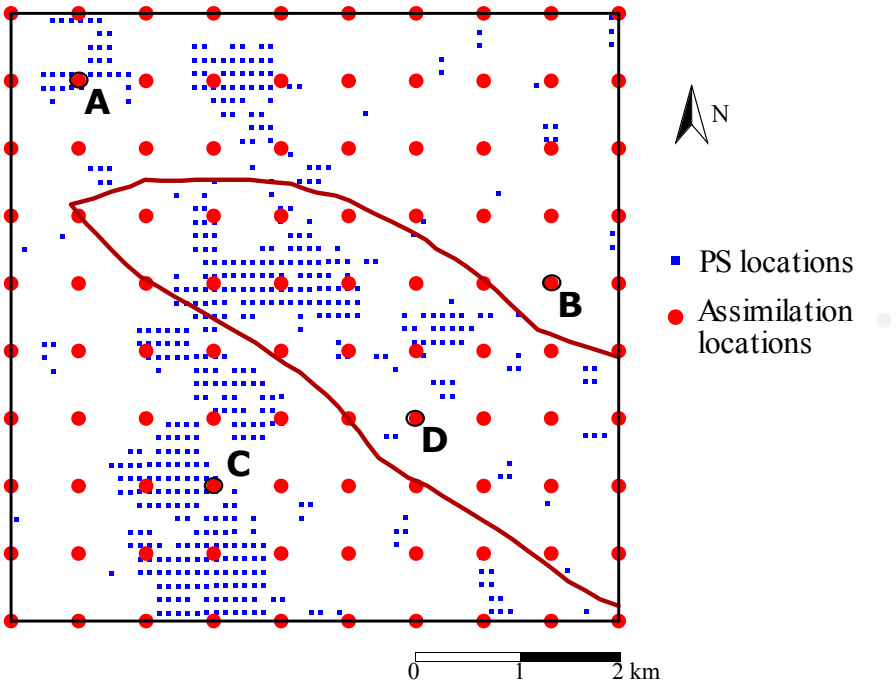

FIGURE 4.4: Map of the PSI record locations (blue dots), measurement grid (i.e. assimilation locations) adopted in the assimilation of surface displacement data (red dots), and the four representative locations A, B, C and D.

The parameters β, θ and s are stochastically independent and sampled from the following uniform PDFs:

$$\beta = \frac{E_h}{E_v} \in U[0.2, 5] \tag{4.1a}$$

$$\theta = \frac{G_h}{G_v} \in U[0.5, 2] \tag{4.1b}$$

$$s = \frac{c_{M,loading}}{c_{M,unloading}} \in U[1, 10] \tag{4.1c}$$

The sampling interval for ν is defined by the inequalities (2.29) and (2.30), which depends on the sampling of parameter β. Figure 2.6 shows the domain of existence for the Poisson ratio as a function of β. The chosen ensemble size, equal to 1000, is deemed adequate to accurately reproduce the selected statistics for each parameter, leading to linearly shaped CDFs, except for ν. Indeed, for the Poisson ratio the CDF is not linear since Equation (2.30) implies the statistical dependence of ν on β, such that only one of the two parameters can fit a uniform distribution. However, the CDF of ν is in essence linear, with some slight non-linearity observed only in the upper tail of the distribution.

By selecting uniform PDFs, the Gaussian hypothesis required by the ES formulation is not fulfilled. To comply with this hypothesis, an *ad hoc* transformation of the prior distribution of β, ν, θ and s is performed. The Normal Score Transform (NST) is used for this purpose. With the NST, each sampled state variable ensemble is first transformed to a local cumulative distribution by sorting its values in ascending order ($\boldsymbol{x_1 < x_2 < ... < x_{n_{MC}}}$) and calculating the associated sample CDF value by the Hazen formula [*Hahn & Shapiro*, 1967]:

$$CDF(x_i) = \frac{i - 0.5}{n_{MC}} \quad i = 1, ..., n_{MC} \tag{4.2}$$

where n_{MC} is the number of Monte Carlo realizations of the ensemble. Then, Equation (4.2) is transformed to a standard normal distribution using the inverse of the standard normal CDF, ϕ:

$$z_i = z(x_i) = \phi^{-1}\left(\frac{i - 0.5}{n_{MC}}\right) \quad i = 1, ..., n_{MC} \tag{4.3}$$

The transformed state variable ensemble z_i ($i = 1, ..., n_{MC}$) is then updated using the ES scheme and back transformed to the x state space using ϕ and the inverse of the

prior sample CDF of the update ensembles of parameters.

4.4.2 Model prediction by FE geomechanical simulations

Each realization of the parameters is used as an input in the FE geomechanical model to obtain the corresponding model forecast in terms of displacements. Note that, in each simulation, the spatial and temporal distribution of the fluid pressure is the same. Although the observation data are available over the period 2003-2008, the simulation time spans the whole 27 year period (1981-2008) of the field production life, in order to have a consistent stress state for the year 2003, i.e., the beginning of the data assimilation process. The loading step of the simulation is set equal to six months according to the time interval between injection and extraction in the UGS field. Figure 4.5 and Figure 4.6 summarize the outcome of the model forecast simulations. The two figures display the time series for the horizontal and vertical displacement, respectively, of the land surface at the four representative locations A, B, C, and D given in Figure 4.4, over time lapses T1, T2, T3, and T4, i.e., time-lapse of six months from April 2003 to April 2005 (Table 4.1). Four time-lapse displacements were considered in the assimilation of the hypothetical data: a) T1 (April 2003 - November 2003), b) T2 (November 2003 - April 2004), b) T3 (April 2004 - November 2004), and d) T4 (November 2004 - April 2005). Each realization of the forecast ensemble, the mean, and the associated standard deviation of the forecast ensemble are shown in the plots. Each sub-panel provides also the values of the corresponding observed PSI component of land surface displacement. Note that the scale of the displacements is different for each location. For example, at time lapse T1 the forecast mean plus and minus the standard deviation of the horizontal displacements (Figure 4.5) are equal to -5.2 ± 30.7 mm, $+13.3 \pm 88.4$ mm, -15.8 ± 98.1 mm, and -4.7 ± 41.0 mm at locations A, B, C, and D, respectively. It is worth mentioning that the presence of a few outliers in the forecast ensemble (not shown in Figure 4.5) due to simulations with particular combinations of parameters leads to large values of the standard deviation. Comparably large standard deviations are also computed for time lapses T2, T3 and T4. Notice also that the forecast mean is larger (approximately by a factor 3) at locations B and C with respect to A and D. These

TABLE 4.1: Summary of the performed synthetic tests.

Test	Time Intervals	Assimilation Times
1	T1+T2	(April03-November03)+(November03-April04)
2	T2+T3	(November03-April04)+(April04-November04)
3	T3+T4	(April04-November04)+(November04-April05)
4	T1+T2+T3+T4	(April03-November03)+(November03-April04)+ (April04-November04)+(November04-April05)

distributions are strongly non-Gaussian, as clearly revealed by large values of skewness and kurtosis, and thus the mean of these distributions are located towards the tails of the ensemble. The same considerations hold true for the vertical displacements (Figure 4.6) with means and standard deviations equal to 0.11 ± 0.67 mm, 10.50 ± 20.4 mm, 14.1 ± 63.1 mm, and 35.3 ± 224.9 mm at time lapse T1 above locations A, B, C, and D, respectively. The largest average vertical displacements are computed at location D, right above the gas pools, and the lowest at location A. Intermediate values are calculated at locations B and C. Note that there is a direct correspondence between the locations where the largest displacements and ensemble spreads are obtained and the regions where the largest fluid pressure excursion occurs. At location A, the PSI observations at all time lapses T1-T4 are found at the fringes of the forecast ensemble and particularly so for vertical displacements, meaning that in this region the model output does not reproduce well the observed data.

4.5 Numerical results

4.5.1 Synthetic Tests

The assimilation of the real displacement measurements was preceded by a series of preliminary tests aiming at assessing the ability of the ES algorithm to effectively perform the update of, and therefore reducing the prior uncertainty on, the geomechanical transversely isotropic parameters β, ν, θ, and s. In these tests, the displacements of the land surface were obtained from the result of a separate simulation using as "true reference values" β=3.62, ν= 0.19, θ= 0.96 and s= 4.82. The horizontal and vertical components of the displacement obtained from this simulation constitute virtual OSSE measurements sampled over a regular grid made up by 10×10=100 locations that

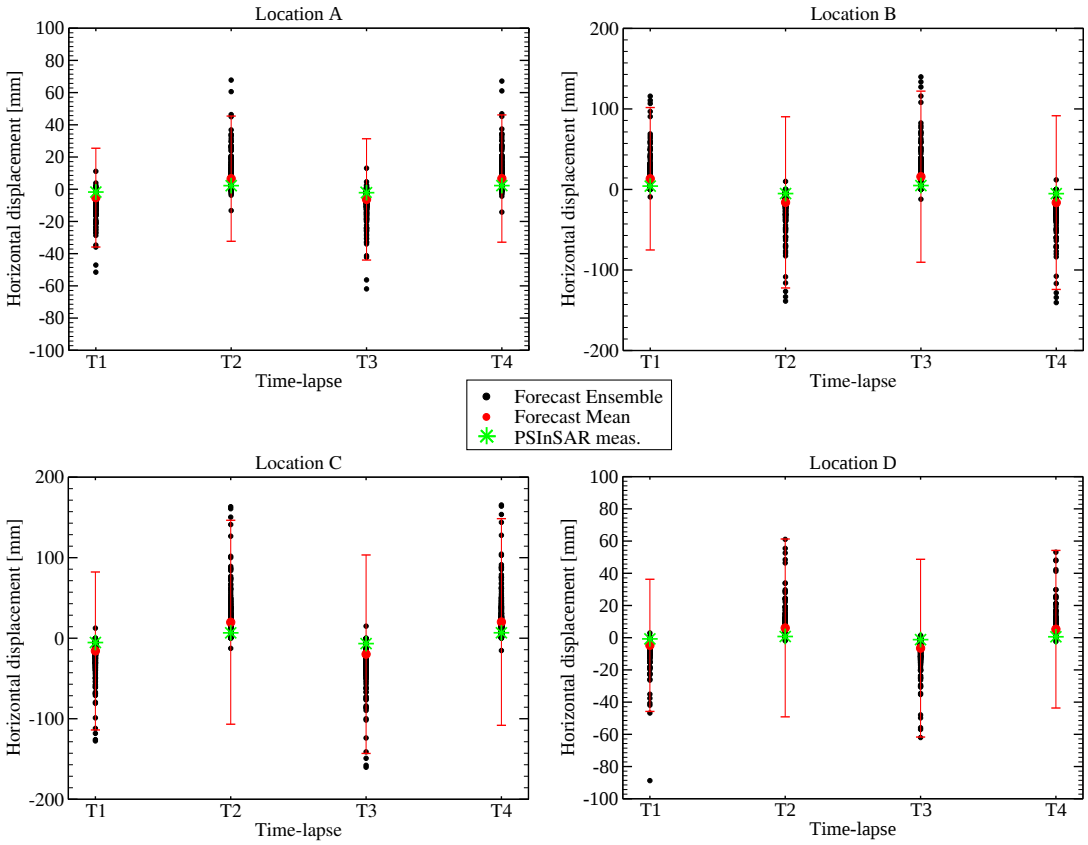

FIGURE 4.5: Ensemble of the horizontal land surface displacement at four representative locations (A, B, C, and D) in the areas above the reservoirs and over the time lapses T1, T2, T3, and T4 (see Table 4.1). Each sub-panel shows also the ensemble mean, the ensemble standard deviation and the observed PSI data. Note the different value scale on the y-axis of the various plots depending on the location point with respect to the reservoir. Large standard deviations are due to a few outliers located outside the y-axis ranges of the sub-panels.

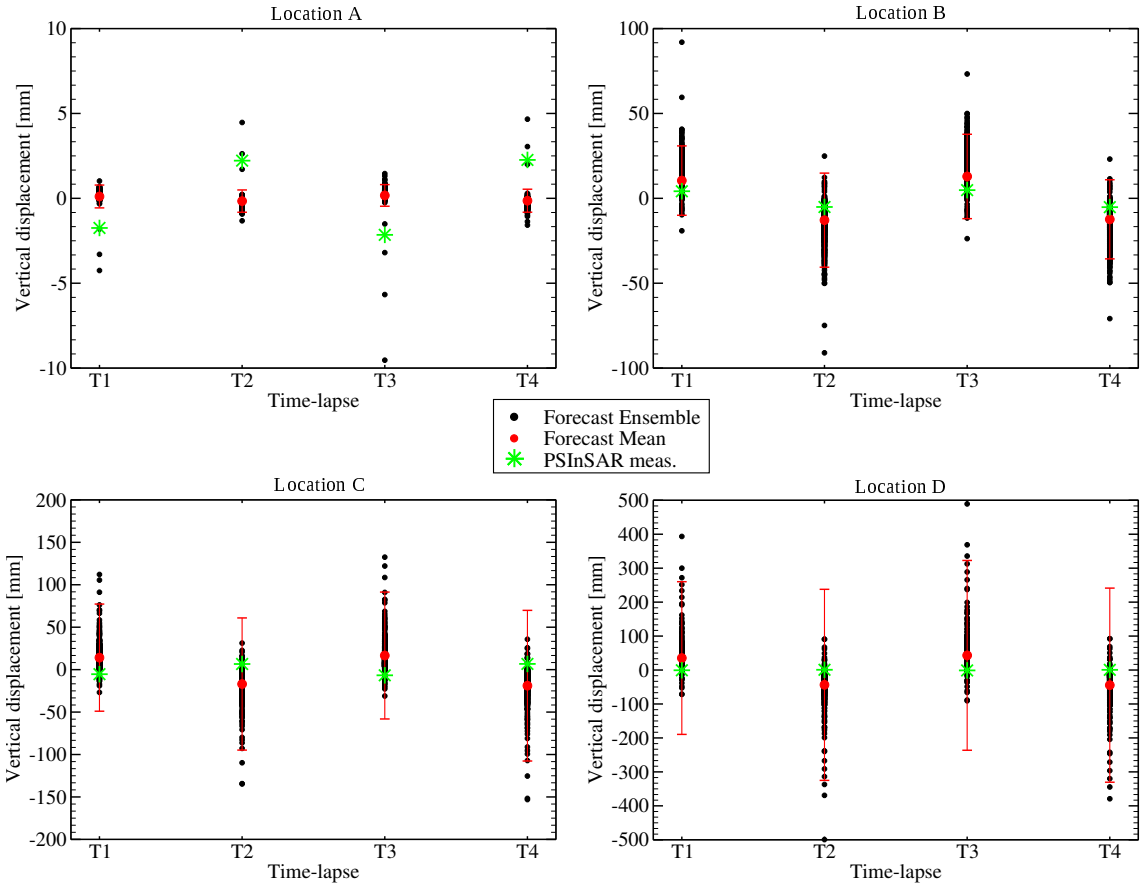

FIGURE 4.6: Ensemble of the vertical land surface displacement at four representative locations (A, B, C, and D) in the areas above the reservoirs and over the time lapses T1, T2, T3, and T4 (see Table 4.1). Each sub-panel shows also the ensemble mean, the ensemble standard deviation and the observed PSI data. Note the different value scale on the y-axis of the various plots depending on the location point with respect to the reservoir. Large standard deviations are due to a few outliers located outside the y-axis ranges of the sub-panels.

is deemed adequate to provide an accurate representation of the surface observations (Figure 4.4).

A first set of experiments addresses the joint assimilation of horizontal and vertical displacements collected during 1 year of UGS, i.e., one injection and one extraction phase. In Tests 1, 2, and 3 time lapses (T1+T2), (T2+T3), and (T3+T4) are assimilated, respectively. The preliminary test also aims at investigating the sensitivity of the data assimilation performance to the number of observation periods; then, in a further test (Test 4), two complete gas storage cycles are assimilated (T1+T2+T3+T4). Refer to Table 4.1 for the description of the performed tests.

The results are summarized in Figure 4.7, which shows the CDFs of the parameters β, ν, θ and s prior to and after the assimilation of the displacement data. Figure 4.7 provides evidence that a significant reduction in parameter uncertainty is achieved by data assimilation, as the spread of the posterior CDFs is significantly decreased with respect to the prior CDFs and the updated parameter ensembles tend to approach their respective reference value in all four test cases. The results also indicate that the posterior distributions are quite similar if measurements are extracted from either two (Tests 1, 2, 3) or four (Test 4) observation periods. This is an important finding as it shows that many data are actually redundant in time and can be safety dropped, thus allowing for a significant reduction of the size of the state matrices which directly affects the computational cost required by the ES algorithm. In order to analyze in detail the performance of the applied methodology and to compare the results from the four test cases, two performance parameters are introduced. The updated model state/parameters and the prior state/parameters are compared through the Average Absolute Error (AE) and Average Ensemble Spread (AES) [*Hendricks Franssen & Kinzelbach*, 2008]:

$$AE = \frac{1}{n_{MC} \cdot n} \sum_{j=1}^{n_{MC}} \sum_{i=1}^{n} |x_{i,j} - x_{i,true}| \tag{4.4}$$

$$AES = \frac{1}{n_{MC} \cdot n} \sum_{j=1}^{n_{MC}} \sum_{i=1}^{n} |x_{i,j} - \overline{x}_i| \tag{4.5}$$

where n_{MC} and n are the number of Monte Carlo realizations and the total number of measurement locations, respectively, $x_{i,j}$ is either the prior or the posterior value

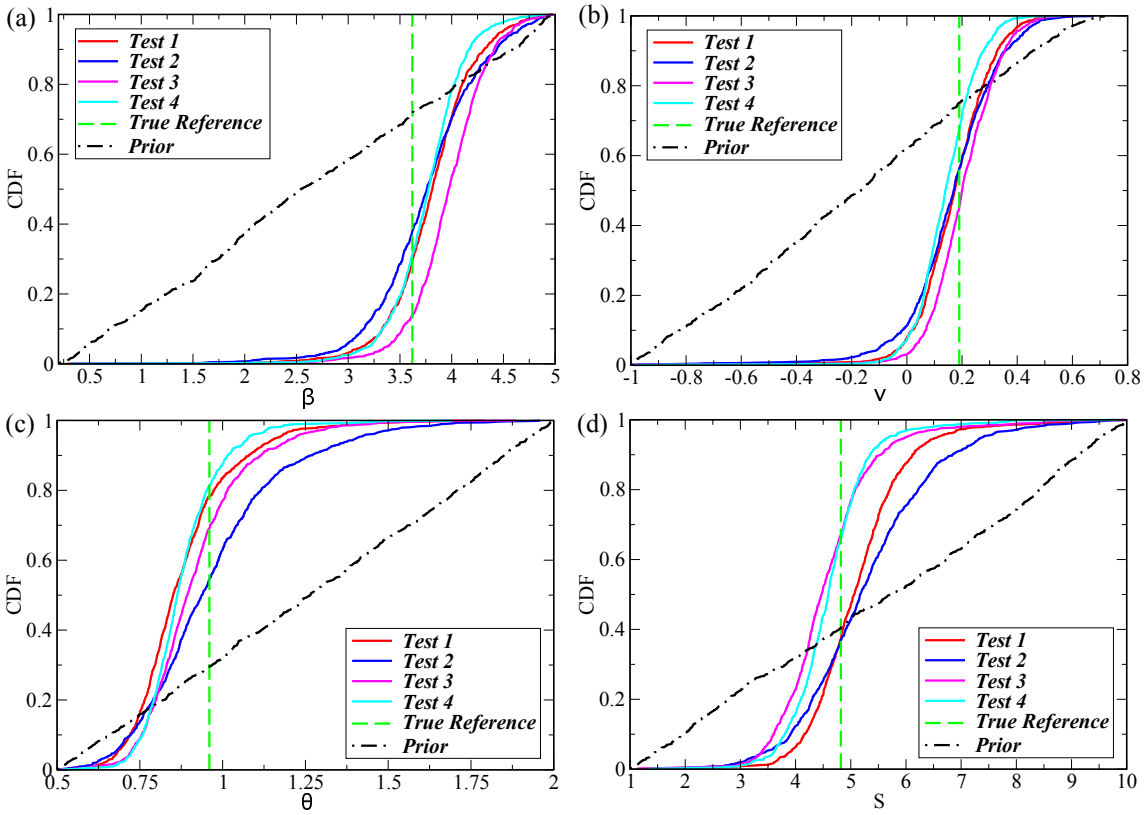

FIGURE 4.7: CDFs of the parameters (a) β, (b) ν, (c) θ, and (d) s prior to and after the assimilation of displacement data. Each sub-panel shows also the "true" parameter value for the synthetic reference system used to simulate data collection. Updated CDFs are graphed for test cases 1 (T1+T2), 2 (T2+T3), 3 (T3+T4) and 4 (T1+T2+T3+T4).

TABLE 4.2: Performance indexes AE and AES for parameters β, ν, θ and s prior and posterior the assimilation of synthetic measurements. The four tests of Table 4.1 are compared.

	AE_{prior}	AES_{prior}	$AE_{posterior}$				$AES_{posterior}$			
			Test1	Test2	Test3	Test4	Test1	Test2	Test3	Test4
β	1.41	1.19	0.34	0.41	0.42	0.29	0.31	0.39	0.28	0.26
ν	0.48	0.39	0.10	0.12	0.09	0.09	0.10	0.12	0.09	0.08
θ	0.43	0.37	0.14	0.16	0.12	0.12	0.11	0.16	0.11	0.08
s	2.39	2.31	0.65	0.95	0.67	0.53	0.63	0.91	0.61	0.49

at the i_{th} measurement location of the j_{th} ensemble realization, $x_{i,true}$ and $\overline{x}_i$ are the "reference true" value and the ensemble mean at the i_{th} measurement location. A similar formula holds for each individual parameter with n=1. The AE takes into account every value in each ensemble model, comparing the model values to the true reference value at each location in the model domain; AES accounts for the deviation of the model values from the ensemble mean at each location, providing an indication of the spread of the distribution [*Bailey & Baù*, 2010]. Hence, posterior AE, compared to prior AE, is a measure of the algorithm capability to predict the true values of the state variables, i.e., the horizontal and vertical displacements, and parameters β, ν, θ and s. AES is a measure of the confidence in the predicted values. Quite obviously, the smaller AE and AES, the better the assimilation outcome. However, Equation (4.4) cannot be used when applied to real observations because the true state is unknown. Table 4.2 summarizes the AE and AES values for the geomechanical parameters β, ν, θ, and s prior to and after the assimilation of data in test cases 1-4. Test case 4 performs better for all parameters compared to tests 1, 2 and 3 with an AE and AES reduction of about 80%, even though the significantly higher computational cost does not justify the small gain. Indeed, in tests 1, 2 and 3, the updating of the parameters gains about 70% for both AE and AES. This outcome suggests that the performance of the ES algorithm is little influenced by the observation periods chosen in the assimilation and that the ES algorithm is able to update effectively the parameters of the transversely isotropic geomechanical model by assimilating vertical and horizontal displacements.

TABLE 4.3: Three set of parameters β, ν, θ and s to obtain OSSE#1, OSSE#2, OSSE#3, respectively.

	β	ν	θ	s
OSSE#1	3.60	0.18	0.96	4.80
OSSE#2	2.90	-0.29	1.02	9.50
OSSE#3	0.66	0.23	0.55	1.10

TABLE 4.4: Prior mean and AES of the parameters β, ν, θ and s

	β	ν	θ	s
mean	2.62	-1.77	1.26	5.6
AES	1.19	0.39	0.37	2.31

4.5.2 Convergence of the ES algorithm

In the previous section, a testing assimilation is carried out using a OSSE obtained through the parameter set with β=3.62, ν= 0.19, θ= 0.96 and s= 4.82. However, choosing a different set of parameters to simulate the OSSE may result in performance indices of the assimilation (Equations (4.4) and (4.5)), which differs from those previously obtained. In this section, the assimilation of three different OSSE are compared. Table 4.3 reports the parameter sets used to simulate the OSSEs in terms of horizontal and vertical displacements u_h and u_v. OSSE#1 is the one used in the previous tests (Section 4.5.1).

As previously discussed (Section 4.4.1) the prior PDFs of the parameters β, ν, θ and s are uniformly distributed with prior mean and AES provided in Table 4.4. The prior CDFs are shown in Figure 4.7 (black dashed line). Dealing with synthetic test cases, one expect that the ES performs quite similar when assimilating displacements from different OSSEs. However, some discrepancies are found in these tests. Table 4.5 provides the assimilation results in term of the AES index posterior to the assimilation. Comparing the three assimilations, it is evident that the spread of the updated ensembles do not vary from case to case (except for θ in OSSE#3 that is much lower

TABLE 4.5: AES posterior to the joint assimilation of horizontal and vertical displacements for the OSSE#1, OSSE#2, and OSSE#3.

	β	ν	θ	s
OSSE#1	0.34	0.07	0.10	0.35
OSSE#1	0.34	0.07	0.12	0.35
OSSE#3	0.34	0.07	0.01	0.35

TABLE 4.6: AE prior and posterior to the joint assimilation of horizontal and vertical displacements for the OSSE#1, OSSE#2, and OSSE#3.

	AE_{prior}				$AE_{posterior}$			
	β	ν	θ	s	β	ν	θ	s
OSSE#1	1.00	0.36	0.30	0.82	0.06	0.05	0.11	0.15
OSSE#2	0.24	0.14	0.24	3.90	0.52	0.07	0.12	0.41
OSSE#3	1.96	0.41	0.70	4.5	1.04	0.15	0.05	0.82

than in OSSE#1 and OSSE#2). To evaluate the estimation of the parameters, i.e., if the mean of the updated ensemble tends to approach the true reference value of Table 4.3, the AE prior and posterior to the assimilation are reported in Table 4.6. The overall reduction of AE from prior to posterior distributions is higher within the assimilation of OSSE#1. For example, the AE index for β is equal to 0.52 and 1.04 in OSSE#2 and OSSE#3 while in OSSE#1 is much lower and equal to 0.06, indicating less accurate estimation in OSSE#2 and OSSE#3.

These tests suggest the possibility that the parameter space may lack in sampling density in some regions of the domain. Note that a number of 1000 realizations are sampled. Figure 4.8 shows the 4-dimensional parameter space with β, ν and θ plotted in the x-axis, y-axis, and z-axis, respectively, and s by a colour scale. In the same figure, the three parameter sets used to simulate OSSE#1, OSSE#2 and OSSE#3 are highlighted. The effectiveness of the assimilation, differing from one to another set of parameters, may be due to the random sampling technique where the number of samples are not enough to describe the whole parameter domain. By distinction, assuming that the parameter space is evenly sampled, it maybe be instructive look at the model outcome from the forecast ensembles of u_h and u_v (Figures 4.9 and 4.10). The ensembles are plotted against the parameter β for representation convenience. For values of β approaching the lower boundary of the prior uniform distribution it is seen a general increase of variability of the model response for both u_h and u_v. Generally, when the output is at the edge of the ensemble, i.e., OSSE#2 and OSSE#3, the ES algorithm provides less accurate parameter estimation. A sensitivity analysis on the role played by the parameters and their combinations in the geomechanical model is provided in the next section.

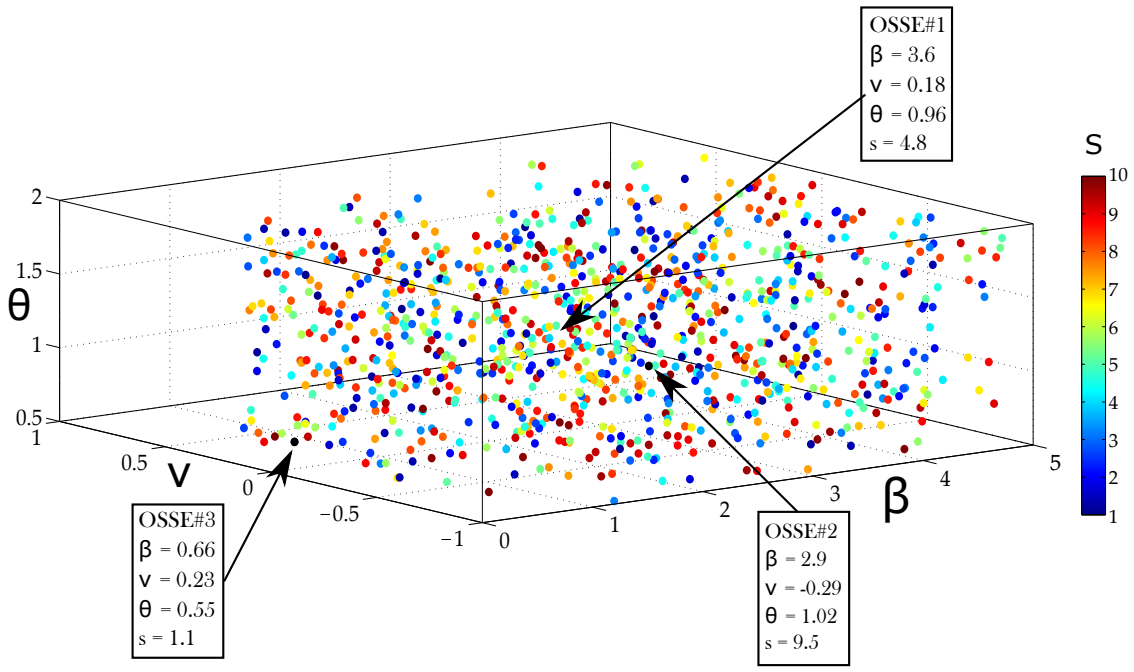

FIGURE 4.8: Parameter space from the Monte Carlo sampling of the parameters β, ν, θ, and s (refer to Section 4.4.1 for the parameter generation).

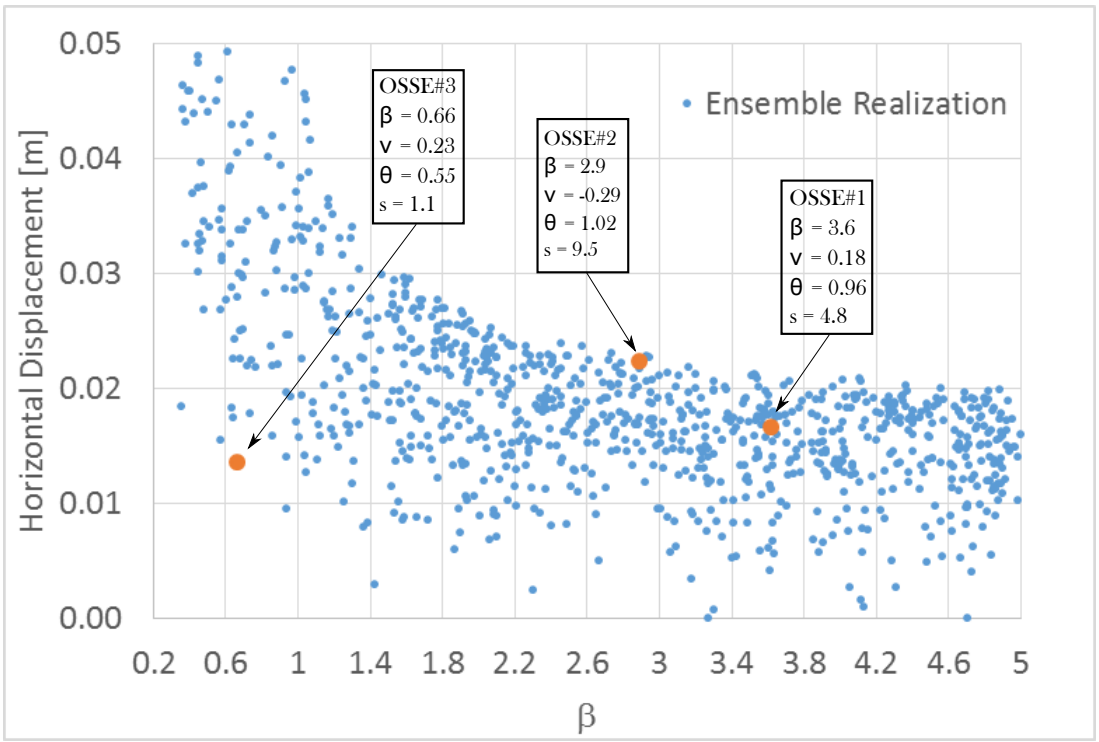

FIGURE 4.9: Forecast ensemble of horizontal displacements u_h obtained with the parameter space of Figure 4.8 plotted against the parameter β.

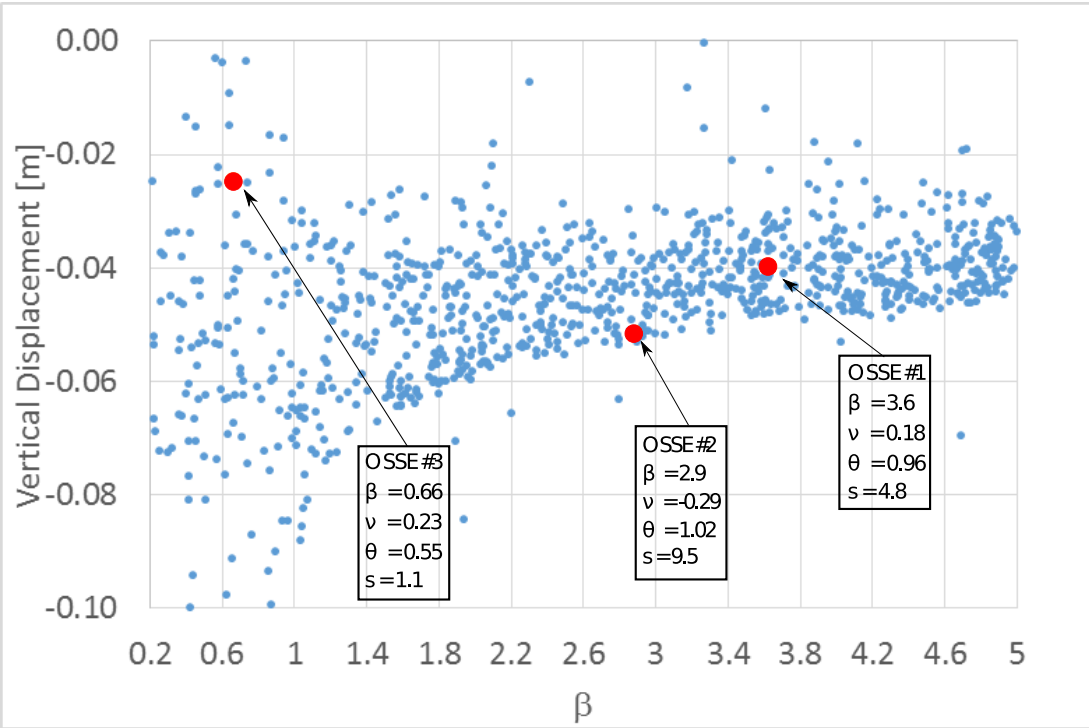

FIGURE 4.10: Forecast ensemble of vertical displacements u_v obtained with the parameter space of Figure 4.8 plotted against the parameter β.

4.5.3 Sensitivity analysis of the geomechanical parameters

A preliminary sensitivity analysis is performed to investigate the relative influence of the geomechanical parameters on the model response, in particular on the relative distribution of the horizontal to the vertical components of surface displacements ($\rho = |u_h|/|u_v|$), and ultimately to revise the prescribed prior distribution of the model parameters.

This analysis is performed on a 3D homogeneous, linearly elastic, isotropic and axisymmetric subsurface system embedding a disk-shaped reservoir characterized by a uniform pore pressure decline and depth-to-radius ratio equal to 0.34, which is comparable with that of *Lombardia* gas field. The sketch of the reservoir is shown in the inset of Figure 4.11. In order to quantify the relative sensitivity of surface displacement components to each of the parameters β, ν and θ, a series of simulations is carried out by varying one parameter at a time. The effect on ρ of the parameter s can be neglected since the horizontal and vertical components of the land displacements are scaled up almost proportionally when varying the value of s. Hence, this investigation is not

addressed by this sensitivity analysis.

The results of the simulations are presented in Figure 4.11. The top row of Figure 4.11 summarizes the dependence of horizontal (left sub-panel) and vertical (right sub-panel) surface displacements on β. As β increases the horizontal displacement is reduced, whereas the vertical displacement grows. It is worth nothing that for values of β below 0.5, the vertical displacement is positive, i.e. uplift. Although theoretically possible, this behavior does not seem to be realistic. The second row of Figure 4.11 presents horizontal and vertical displacement profiles obtained by varying the Poisson ratio ν. It may be observed that increasing ν induces a decrease of either horizontal and vertical displacements. The effect of the ν on the ratio ρ is not significant thus ρ rather insensitive to this parameter. On the other hand, β has a significant influence on horizontal-to-vertical surface displacement ratio, which is shown to decrease for increasing values of β. Finally, the horizontal and vertical displacement profiles obtained by varying θ within the range 0.5-2.0 are shown in the bottom panels of Figure 4.11. These profiles indicate that the land surface displacement is not very sensitive to this parameter.

The results of the sensitivity tests described above provide a benchmark for the following assumptions. First, since the shear moduli ratio does not seem to have much influence on the surface displacement profiles, θ can be considered known deterministically and equal to 1, which practically reduces the size of the inverse problem of one degree of freedom. Another important assumption is made on the value of ν, which is assumed to be strictly positive, a condition that is generally used in rock mechanics [*Gercek*, 2007]. The lower bound of the parameter β is increased from 0.2 to 0.5 to avoid unrealistic model output.

4.5.4 Assimilation of PSI measurements

This section presents the results from assimilating PSI measurements collected over the *Lombardia* reservoir (Figure 4.3 and 4.4) into the geomechanical model response. Figure 4.12 shows the prior distribution of the geomechanical parameters β, ν and s (black dashed lines) based on the result of Section 4.5.3. The prior uniform distributions are transformed to Gaussian distributions (see Section 4.4.1) using the NST. The

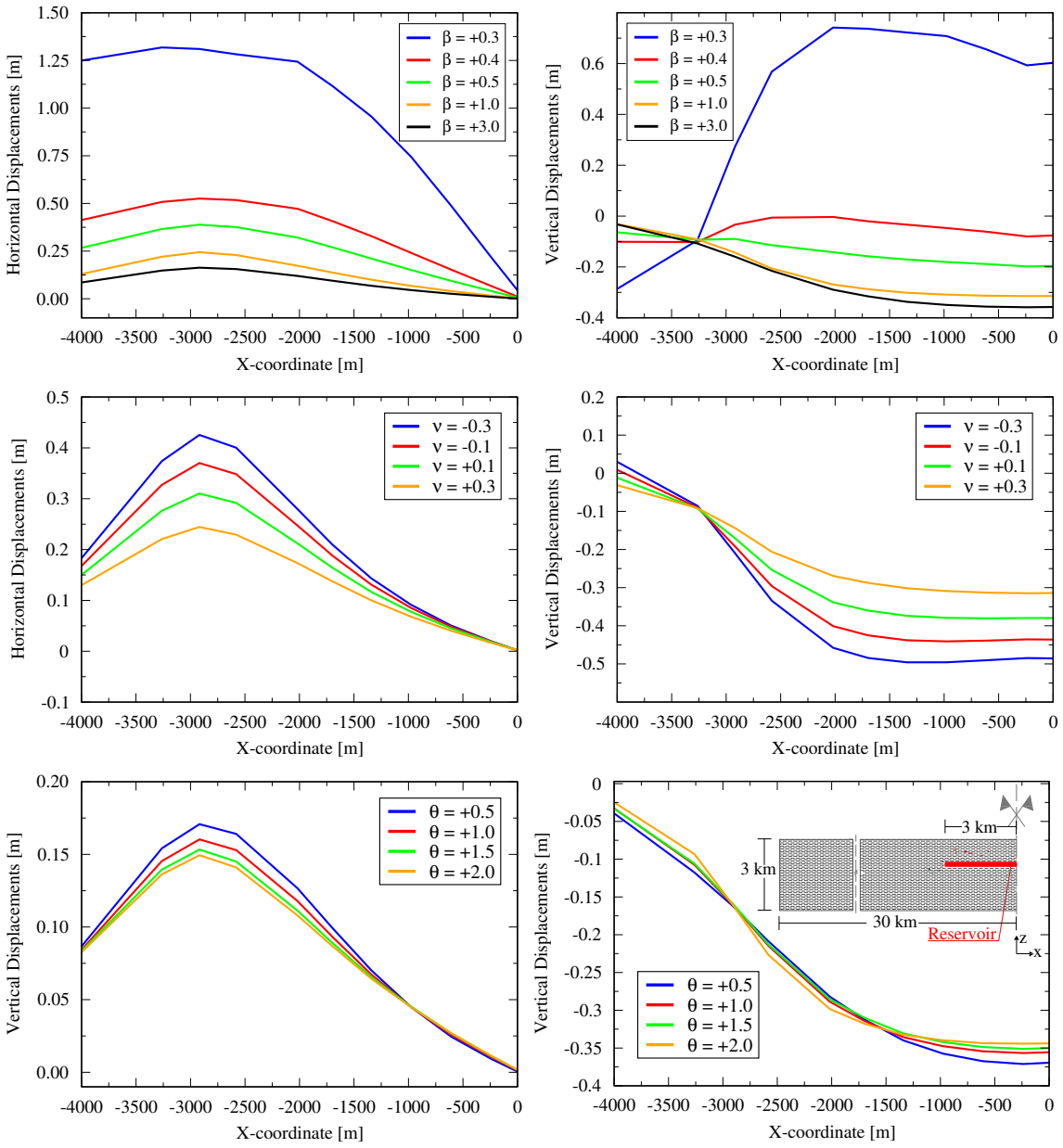

FIGURE 4.11: Radial distribution of the horizontal (left sub-panels) and vertical (right sub-panel) component of the land surface displacement. The three rows (from top to bottom) depict the effect of variation of the coefficient β, the Poisson ratio ν and the coefficient θ.

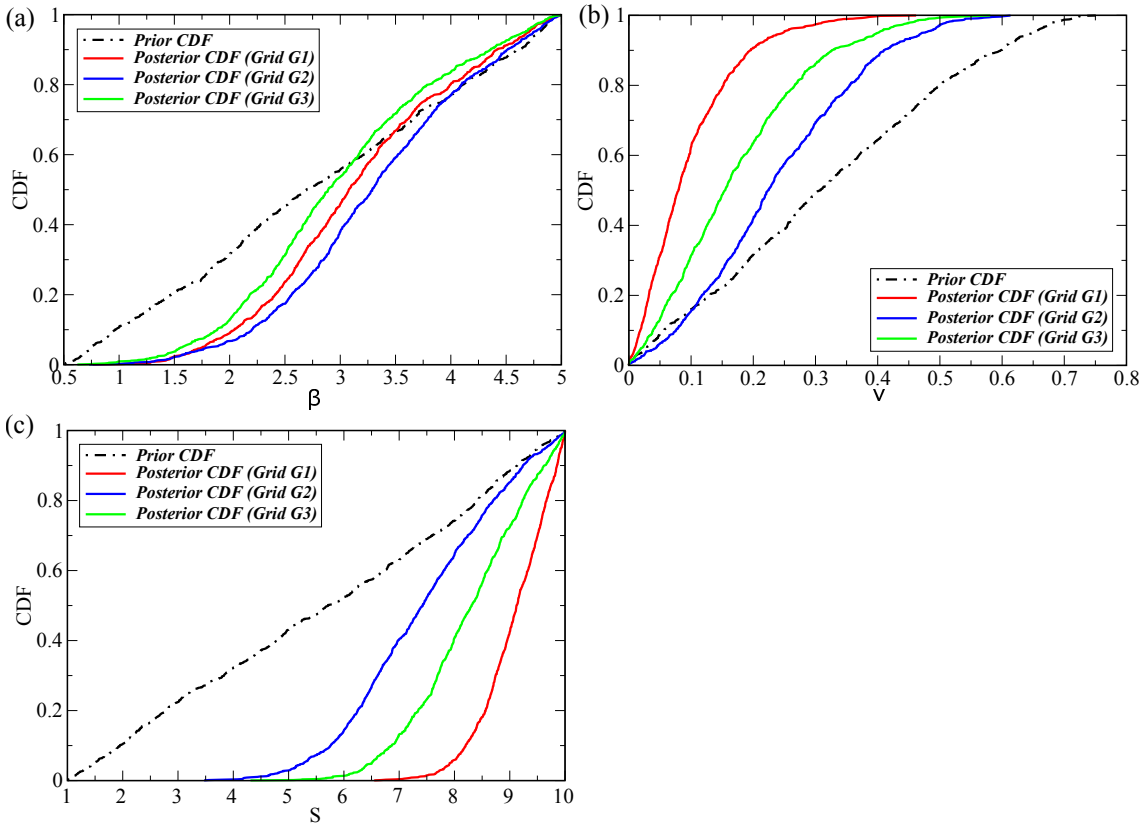

FIGURE 4.12: Prior and posterior CDFs of the parameters (a) β, (b) ν and (c) s. Data assimilation is based on PSI measurements over the time lapses T2 and T3.

assimilation of two time lapse displacements (one complete UGS cycle) is used to update the geomechanical parameters, based on the outcome of the synthetic case tests (Section 4.5.1). The surface displacement data sets for the periods November 2003-April 2004 (T2) and April 2004-November 2004 (T3) are deemed adequate for testing the performance of the assimilation algorithm.

Preliminary tests (not shown here) reveal that the number of the locations used for data assimilation may significantly impact on the update results; hence, three measurement location grids are taken into consideration. The first grid G1 (Figure 4.4) is the same as the one used in the synthetic case with 100 assimilation locations uniformly distributed over the *Lombardia* model domain. The values of horizontal and vertical displacements at each location are taken from the maps of Figure 4.3. One of the challenges of selecting surface displacement data for assimilation is to provide a realistic value for the error to assign to PSI derived observations. The measurement error is here set equal to the kriging standard deviation calculated from the variograms used to obtain the displacement maps of Figure 4.3 [*Teatini et al.*, 2011]. For time lapses T2 and T3, the spatial average of the standard deviation of the vertical displacement, $\sigma_{v,meas}$, is equal to 1.8 mm. The observation random error associated with each measurement is sampled from a Gaussian distribution with zero mean and standard deviation $\sigma_{v,meas}$. Since the horizontal component of the land displacement is generally noiser than the vertical one, the associated measurement error is set equal to $\sigma_{v,meas}$ multiplied by a factor 1.5, i.e. $\sigma_{h,meas} = 2.7$, according to the error analysis of the PSI measurements discussed by *Teatini et al.* [2011].

The results from the assimilation for grid G1 are shown in Figure 4.12, which provides the posterior CDFs of parameters β, ν, and s. The ES update routine actually reduces the prior uncertainty associated with the parameters by jointly assimilating measurements of horizontal and vertical surface displacement. However, the 50th percentiles, i.e., the median value, of the parameters ν and s tend to approach the lower and upper upper bound of their prior distribution, respectively. Thus, the ES seems to suggest that a better match of the measurements might be obtained with a set of parameters that are outside of the prior hypothesized intervals.

The investigation of a potential bias in the measurements and/or the geomechanical model is carried out by assimilating data selected from grid G2, which is obtained by removing the 30 northernmost data points from grid G1. The choice of these nodes is accounted for their likely low level of confidence because few PS are located in this portion of the study area. The standard deviation of the measurement error is the same as in grid G1 since measurements are derived from the same interpolated maps (Figure 4.3). The assimilation results are again plotted in Figure 4.12. The posterior CDF of the parameter β is close to that obtained with grid G1. By distinction the posterior CDF of ν and s are shifted toward the center of the investigated prior intervals. Hence, the effect of excluding the 30 upper-north locations from the assimilation is an increase of the median value of ν and a reduction of the median value of s. Grid G2 uses a set of measurements yielding updated parameters more constrained within the prior intervals. The comparison of the results produced by grids G1 and G2 suggests the possibility of a bias in the PSI derived surface displacements and/or in the model setting for the *Lombardia* reservoir. Indeed, a different update ensemble of parameters is obtained when assimilating two different measurement grids due to the exclusion, the case of grid G2, of a set of observed data that are not resembled by the forecast ensemble of displacements (refer to Figure 4.6 - location A). On the other hand, the exclusion of data points in the northern reservoir leads to updated parameters β, ν and s that in any case could not account for the model behavior in the northern part of the domain.

Figure 4.13 shows the updated ensembles of vertical u_v and horizontal u_h displacements at the four locations A, B, C and D of Figure 4.4, obtained by assimilation of data from grids G1 and G2. The prior ensembles of vertical and horizontal displacements are also displayed. In general, the spread of the updated ensembles of grid G1 is narrower than that of grid G2. At location A, in upper-north part of the domain, the updated u_v values are overestimated for grids G1 and G2 and the observation falls outside the updated ensemble. The PSI vertical measurement is well matched at location B for both grids even though the spread of the updated ensemble of grid G1 is narrower than that of grid G2. At location C, the PSI observation is underestimated with both grids, but the ensemble mean obtained with grid G1 is closer to the observed value. At location D, above the gas field, the updated ensemble of grid G2 better resembles the

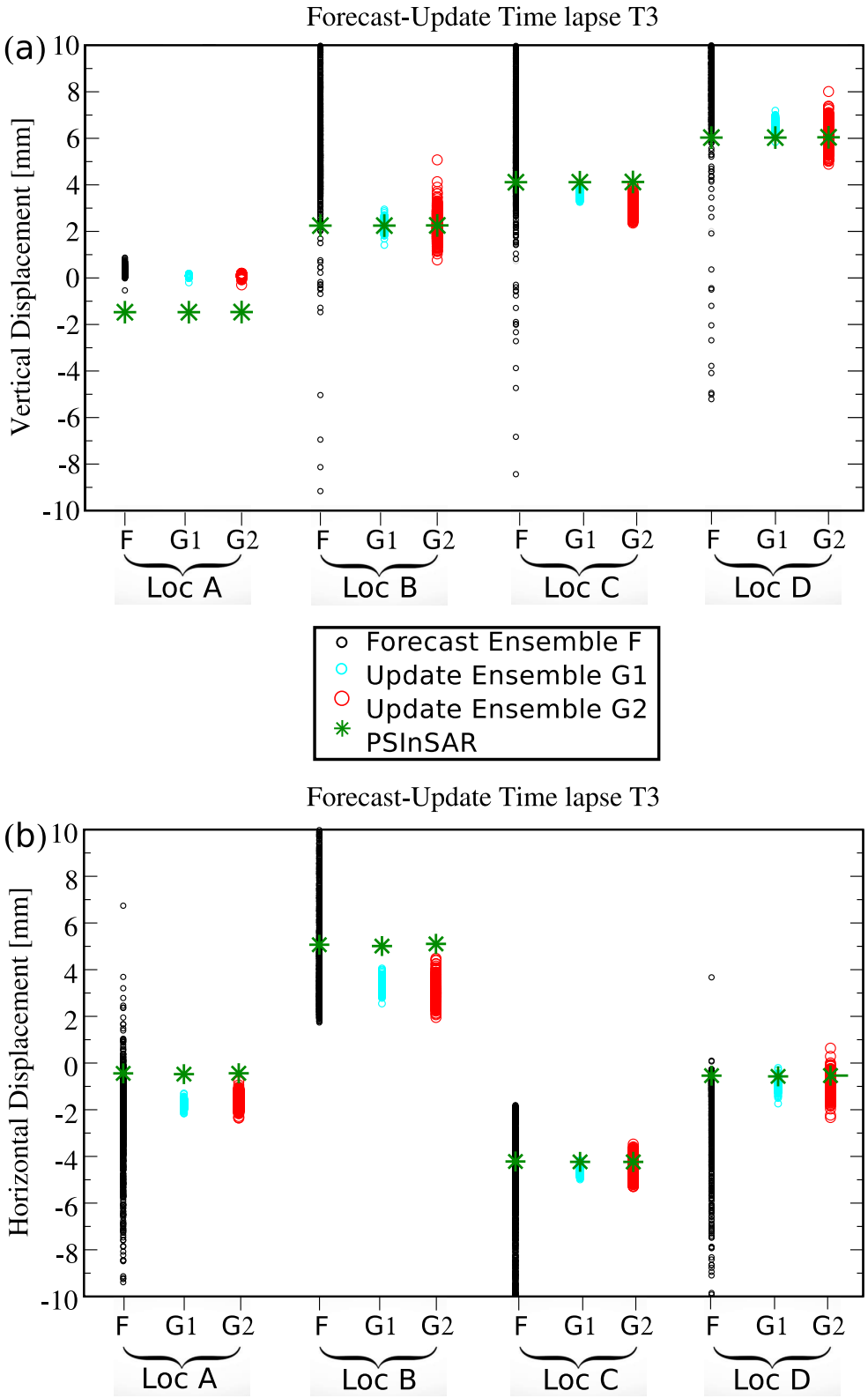

FIGURE 4.13: Updated ensemble of (a) vertical and (b) horizontal displacements by assimilation of data points of grids G1 (cyan circles) and G2 (red circles), at four locations A, B, C and D (Figure 4.4). The prior ensembles of vertical and horizontal displacements are also displayed (black circles). The PSI measurements are plotted with green stars at the four locations.

PSI, while the mean updated ensemble obtained with grid G1 slightly overestimates the observation.

The mean values of the updated ensembles of the horizontal displacement u_h are similar in the assimilation of grids G1 and G2. The difference between the two assimilation is in the spread of the updated ensembles, which are narrower with grid G1 at the four locations. At locations A the PSI observations are outside the updated ensembles whereas the horizontal measurement are well resembled at locations C and D.

Excluding the uppermost locations from grid G1 leads to an increase of the spread of the posterior ensembles since less constraints are prescribed on the assimilation. However, it appears that the method cannot significantly improve the matching of the PSI observations, in particular at locations north of the reservoir for vertical displacements and east of the reservoir for horizontal displacements. This suggests that a possible bias in the model and/or in the observations is not be completely eliminated by dropping the uppermost data points from the three top rows of grid G1.

To further address these aspects, a third grid G3 is investigated for the assimilation. Grid G3, shown in Figure 4.14, totals 70 locations, i.e., the same number as grid G2, distributed all over the model domain, which cover also areas situated north of the reservoir footprint. These locations are not uniformly distributed as is for grids G1 and G2. Actually, the 70 data points are chosen directly as a subset of the available measurements from the 772 PSI locations (blue dots of Figure 4.4). In this case, the interpolation of the PSI data is avoided with the "raw" PSI displacements directly used in the assimilation. Moreover, in grid G3, horizontal and vertical displacements are selected so as to be consistent with the expected contraction/expansion sequences associated with the pressure variations in the gas pools. The standard deviation of the measurement error is calculated on the basis of the new available data set. Indeed, a new experimental variogram is built and the average standard deviation of the measurements computed from the kriging using that variogram is assumed as the standard deviation of the measurement error. Over time lapses T2 and T3, the vertical standard deviation is equal to 1.5 mm, whereas the horizontal one is again increased by a factor of 1.5 with respect to the vertical measurement error. The updated CDFs of the reservoir parameters obtained with grid G3 are shown in Figure 4.12. The CDF of the parameter β is similar

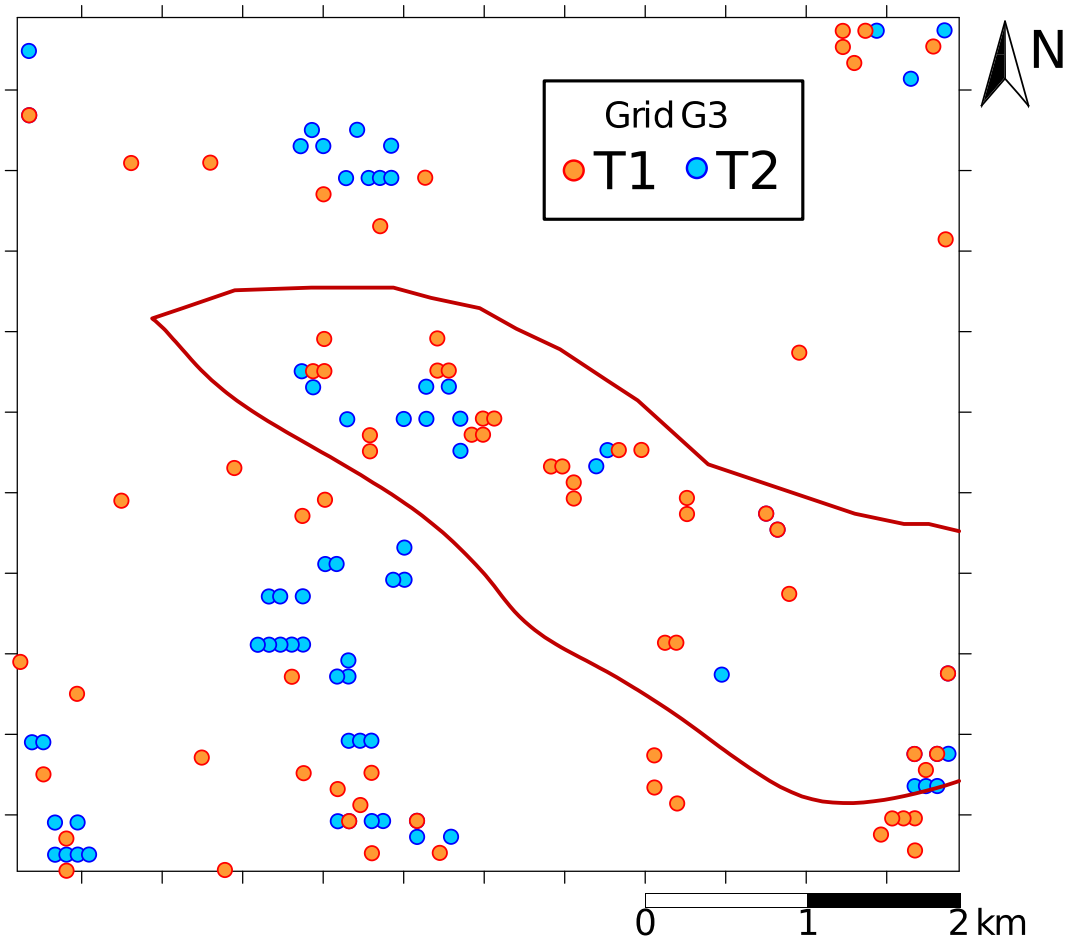

FIGURE 4.14: Map of the measurement grid G3, i.e., assimilation locations, adopted in the assimilation of surface displacement data. Note that the assimilation locations used for assimilation of time lapses T2 and T3 are not the same. The trace of the UGS field is marked by the dark-red line.

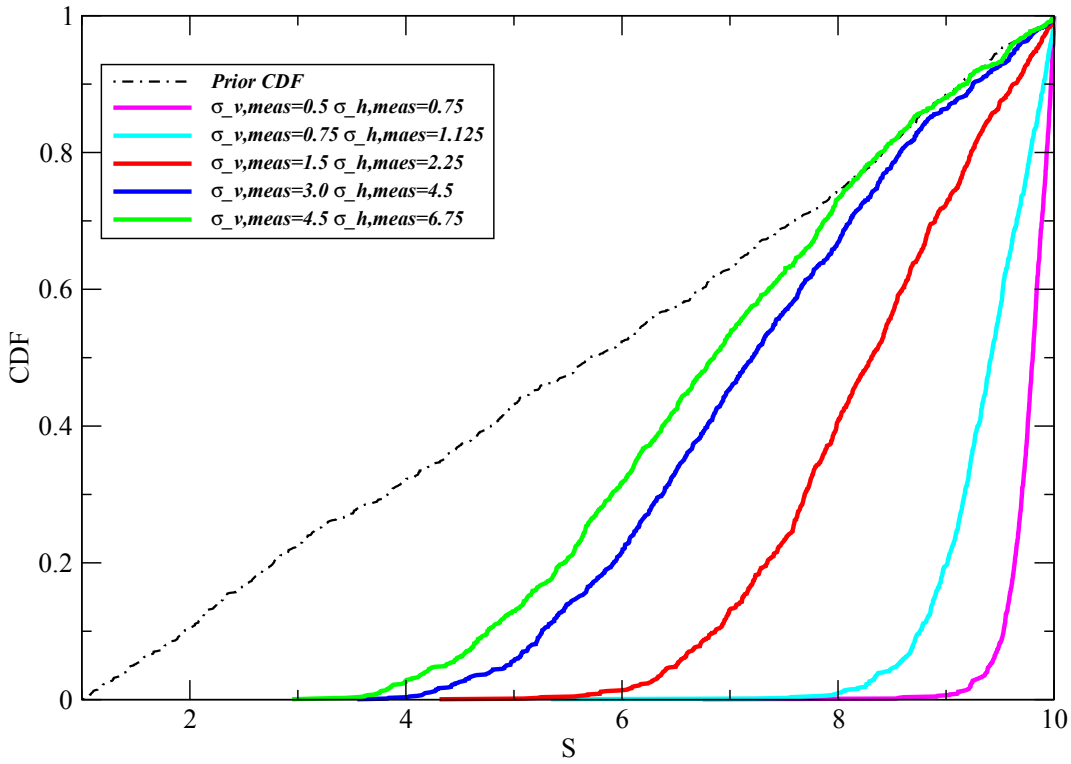

FIGURE 4.15: Prior (dashed line) and posterior CDFs (solid lines) of the parameters s with respect to varying values of error standard deviation, i.e $\sigma_{v,meas}$ and $\sigma_{h,meas}$, associated with the PSI data points of grid G3 (see Figure4.14).

to the ones from grids G1 and G2, while the CDFs of ν and s lies in between the corresponding CDFs of grids G1 and G2. These results can be explained as follows. Grid G3 contains a smaller measurement bias with respect to grid G1, hence the shift of ν and s CDF's toward the interval bounds is clearly reduced. On the other hand, since grid G3 accounts for observations in the northern part of the domain, the CDF's of ν and s are shifted to the left and to the right, respectively, of those obtained with grid G2, which instead misses these observations.

The effect of the measurement error on the smoothing performance is also investigated. The results of a sensitivity analysis are shown in Figure 4.15 with reference to grid G3. The profiles of posterior CDF are given only for the parameter s as similar considerations hold for β and ν. Figure 4.15 points out that the posterior CDF depends on $\sigma_{v,meas}$ and $\sigma_{h,meas}$=1.5 $\cdot$ $\sigma_{v,meas}$. Notice that the proportionality factor between $\sigma_{h,meas}$ and $\sigma_{v,meas}$ is kept equal to 1.5. The increase of $\sigma_{v,meas}$ (i.e. more uncertain observations) obviously reduces the ES performance. Conversely, the spread of the

posterior CDF reduces when $\sigma_{v,meas}$ decreases, at the expense of a CDF's translation toward the boundaries of the prior interval assigned to the parameter. This shift could again be seen as the effect of a measurement and/or model biases: more accurate measurements are hardly reproduced by the model using the prior forecast ensemble of the parameters. Thus, the more constraint is given to the assimilation, the higher is the effect of the biases on the ES update. Hence, in the present case study the most appropriate values of $\sigma_{h,meas}$ and $\sigma_{v,meas}$ are those derived from the previous error analysis of the PSI measurements (Figure 4.12).

The outcomes of the assimilation procedure is further investigated by performing posterior Monte Carlo geomechanical simulations using the update ensembles of the parameters β, ν and s as obtained with the three reference grids. The measurement error is assigned based on the results of Figure 4.12. Figure 4.16a, 4.16b and 4.16c show horizontal displacement maps for the period April 2004-November 2004 (T3) for grids G1, G2 and G3, respectively; Figure 4.16d, 4.16e and 4.16f show the corresponding vertical maps. The maps show the updated ensemble mean, that is, the average of the posterior land surface displacement ensemble. To compare the model updated mean and the observations, PSI data are also displayed with circles. Note that the color scale is different for the vertical and the horizontal displacements.

In general, the updated ensemble mean of vertical displacements obtained on grids G1, G2 and G3 exhibits an overall good match of the PSI data above the UGS field, with relatively small differences among them. The distribution of vertical displacements obtained with grid G3 (Figure 4.16f) shows the smaller estimated subsidence bowl as compared to grids G1 and G2 (Figure 4.16d and 4.16e), which better fits the PSI observations ($R^2_{\text{G3}} = 0.83$ with $R^2_{\text{G1}} = 0.81$ and $R^2_{\text{G2}} = 0.80$, where R^2 is defined as $1 - \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 / \sum_{i=1}^{n}(y_i - \bar{y}_i)^2$ with y_i and $\bar{y}_i$ the observation values and the respective mean, $\hat{y}_i$ the predicted values and n the total number of observations).

Apart from the simulation of surface displacement over the gas reservoir, difficulties are encountered when fitting observations over the area north of the *Lombardia* field, which turn out to be generally overestimated regardless of the measurement grids used for the assimilation. Indeed, in this part of the domain the pore pressure variations due to UGS operations are considerably lower than in the reservoir, meaning that

the forecast ensemble of land surface displacements contains a prior information that is less significant and the assimilation algorithm is less sensitive to data collected over that area of the domain.

Figure 4.16a, 4.16b and 4.16c show an even better match between the simulated and observed horizontal displacements of the land surface ($R^2_{\text{G1}} = 0.88$, $R^2_{\text{G2}} = 0.89$ and $R^2_{\text{G3}} = 0.89$). The maps of horizontal displacements obtained with grids G1, G2, and G3 well represents the PSI observations in the eastern part of the domain, while the data noise is higher in the western part of the UGS field and the model response is significantly smoother than the observations. Although the differences are relatively small, the overall characterization obtained with grid G3 (Figure 4.16c and 4.16f) is the most satisfactory both for horizontal and vertical displacements.

4.6 Discussion

A number of papers published over the last decade demonstrate the capability of PSI to capture the evolution in time of vertical and horizontal displacements above producing and UGS reservoirs. PSI provides significant datasets for reservoir history matching and subsurface characterization (e.g., *Stancliffe & van der Kooij* [2001]; *Ketelaar et al.* [2007]; *Klemm et al.* [2010]; *Tamburini et al.* [2010]; *Vasco et al.* [2010]; *Teatini et al.* [2011]; *Janna et al.* [2012]). On the other hand, data assimilation methods have been increasingly adopted for parameter estimation problems in petroleum engineering (*Nævdal et al.* [2003], *Gu & Oliver* [2005], *Skjervheim et al.* [2011]; *Iglesias et al.* [2013a]). However, researchers have tried to combine data assimilation algorithms and PSI to improve the reservoir characterization only very recently. Generally speaking, it has been found that this combination performs effectively for synthetic/realistic test cases, while a few difficulties are still met in real field cases. Usually, a complex and specific tuning is required, with the focus (i.e., the uncertain parameters to be constrained) of the assimilation that changes from case to case. The results presented in this thesis confirm these findings.

Wilschut et al. [2011] demonstrate that reservoir compartmentalization can be properly investigated by assimilating different sources of information, i.e., production and

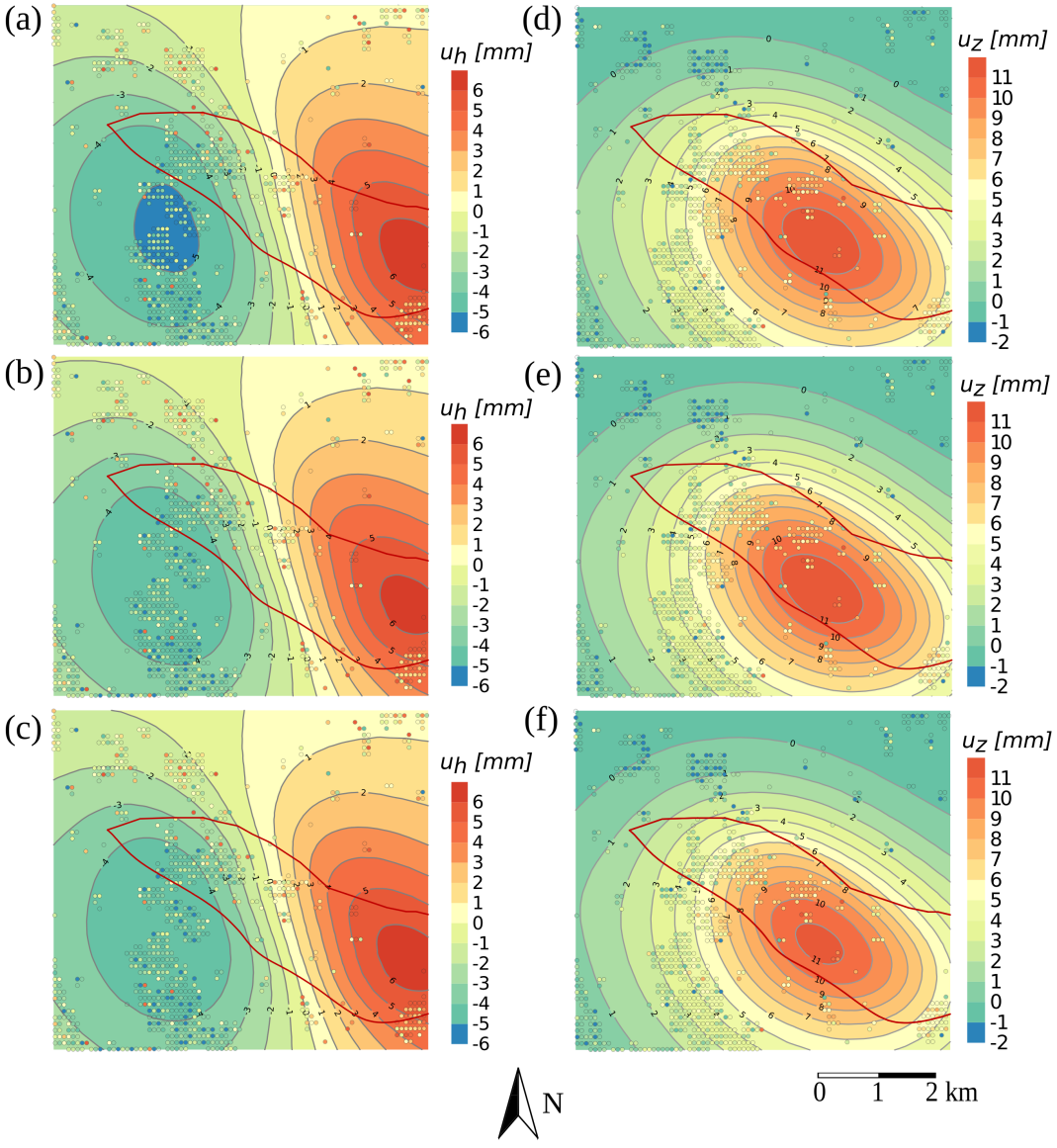

FIGURE 4.16: Maps for (a)-(c) horizontal and (d)-(f) vertical time-lapse land-surface displacements above the *Lombardia* reservoir during the time-lapse April 2004-November 2004 (T3). The updated ensemble means of the posterior Monte Carlo simulations are displayed from top to bottom, on the basis of the update parameters of Figure 4.12 referring to grids G1, G2, G3, respectively. The trace of the UGS field is marked by the dark-red line.

subsidence data, through EnKF. However, their application to the Roswinkel field (The Netherlands) shows that a satisfactory characterization is not easy to achieve because the initial forecast ensemble provided by the model is not consistent with a portion of the available measurements. *Fokker et al.* [2013] make use of an elastic geomechanical model for the reservoir characterization of the Bergermeer gas field (The Netherlands). They employ an ES scheme by assimilating the ascending and descending PSI measurements along the line-of-sight (LoS) geometry. The outcome highlights that the measurement effectiveness in constraining the reservoir compressibility can be quite different for the different reservoir compartments, likely depending on the pressure change history and the influence of other surrounding reservoirs. An ES estimator with multiphase flow and geomechanics simulator as forward models is implemented by *Jha et al.* [2015] to assimilate pressure and PSI data for reducing uncertainty in the prior distributions of rock properties of a UGS gas reservoir in northern Italy. The results indicate that the spread of the posterior displacement ensemble is significantly reduced after assimilation of fluid pressure and surface displacement data. However, there remains a significant discrepancy between the simulated and the observed land displacements. This is an evidence that the uncertainty has been smoothed out, however the calibration is not entirely accurate, with the ES converging to parameter sets that cannot clearly reproduce the observations and the posterior ensemble narrowing significantly but far from the available measurements.

Recent works by *Baù et al.* [2014, 2015] investigate the capability of the ES algorithm to reduce the parameter uncertainty for a transversely isotropic geomechanical model by assimilating both vertical and horizontal ground-surface displacements. A schematic axi-symmetric configuration is tested. It is found that the quality of the estimation mostly depends on the number of the assimilated measurements and the magnitude of the measurement errors. The same transversely isotropic geomechanical model is herein used in a real case, i.e., the *Lombardia* reservoir. From the comparison of the prior and posterior PDFs of the geomechanical parameters, an appreciable reduction of the assigned uncertainty is easily obtained when synthetic measurements, i.e., the displacements generated by an independent model run, are assimilated. More

difficulties are encountered with real PSI data. This is mainly due to the limited consistency between the forward model and a portion of the available measurements rather than a poor cross-correlation among the forward model and the parameters, i.e., the capability of the model to reproduce the behaviour of the system. This result has been detected by choosing three different grids of assimilation points. Excluding from the assimilation a subset of low-confidence measurements in the northern part of the domain (grid G2), or selecting a subset of measurements strictly consistent with the expected geomechanical behavior associated with the pressure variation (grid G3) yields an improvement in constraining the model parameters. Notice that the large variability of the PSI displacements, in the northern part of the domain, even over a very short distance precludes a bias in the model setting linked to a possible uncertainty in the pressure change distribution.

The consistency between the displacement records and the physical process represented by the forward model is surely a key point for a satisfactory implementation of the assimilation algorithms. It is worth emphasizing that, apart from its intrinsic accuracy, PSI provides the cumulative motion of scattered radar targets that may be subject to different independent processes. In the majority of the applications, such processes are characterized by variable time and space scales, e.g., tectonics, fluid withdrawal and injection, target instability, differently contributing to the measured values [*Tosi et al.*, 2009]. Hence, it is preferable to use a subset of 'reliable' records, in the sense that they show a displacement clearly related to the investigated process at hand, rather than a larger dataset with a higher noise level, as is done, for example, by *Jha et al.* [2015]. In the case of deep reservoirs, we have elected to assimilate movements decomposed along the vertical and easting directions instead of original LoS displacements as the decomposition procedure filters out local components most likely unrelated to hydrocarbon production and storage.

4.7 Conclusions

The Ensemble Smoother (ES) algorithm is employed to characterize the geomechanical parameters of a UGS field situated in the Upper Adriatic sedimentary basin (Italy). A

transversely-isotropic, 3D FE model is used to simulate the deformation field caused by the contraction and expansion of the UGS reservoir subjected to seasonal pressure variation due to extraction/injection of gas from/into the subsurface. The model calibration is provided by jointly assimilating horizontal and vertical movements of land surface displacements observed from space. The available measurements consist of PSI data acquired by RADARSAT-1 from March 2003 to October 2008 over the domain of interest.

The results show that ES is an effective inversion procedure to improve the a-priori estimate of the parameters of transversely isotropic geomechanical models for deep hydrocarbon reservoirs provided that horizontal and vertical PSI displacements are used. The joint assimilation of components helps to increase significantly the reliability associated with the prior uncertain parameters. The posterior Monte Carlo simulations provide evidence that an overall good match of the PSI measurements may be obtained.

More detailed insights are the following. An appropriate selection of the observations is important, however, is not straightforward especially where hundreds to thousands measurements provided by SAR-based algorithms are available. The updated parameters are dependent to some extent on different data sets and this indicates the presence of a possible bias in the measurements and/or in the model forecast. A selection of observations with the smallest data noise has yielded the updated parameters producing the most satisfactory match.

Moreover, the error associated with the measurements also plays an appreciable role when assimilating real observations. In synthetic test cases, the measurement error does not affect the ES smoothing performance as the model forecast is unbiased with respect to the virtual system of observations. On the other hand, the assimilation of real measurements requires a more accurate estimate of the observation error that is preferably physically-based, i.e. linked to the precision of the monitoring technique.

In geomechanical problems as the one addressed in the present study, i.e. with a relatively weak parameter nonlinearity, the use of an ES algorithm is deemed adequate for its low computational cost and its possible "off-line" application, that avoids the simulation restart. Possible improvements will be investigated. The parameter estimation will be addressed by assimilation approaches more appropriate for nonlinear

problems, e.g. the Iterative Ensemble Smoother *Chen & Oliver* [2011]. Furthermore, the bias effects of measurements and modeling will be properly accounted by implementing the algorithm by *Pauwels et al.* [2013].

Chapter 5

Gas Producing Reservoirs[1]

5.1 Introduction

In this Chapter, the geomechanical modeling of a producing hydrocarbon reservoir is presented. The aim is at estimating the vertical uniaxial compressibility c_M by assimilation of ground surface displacements. This investigation originates from the need for properly addressing and explaining the seafloor displacements observed over an offshore gas field, hereafter the *Maja* reservoir. Although dealing with a real case application, the geographic reference of the reservoir location is omitted in this thesis for confidentiality reasons with the oil company managing the reservoir operations.

Sections 5.2 and 5.3 are dedicated to describe the fluid pore pressure distribution in the three main gas pools of the *Maja* gas field and the measurements collected above the reservoir.

Before the presentation of the real case, some numerical tests are carried out with the main objective of comparing and discussing two different conceptual models for c_M. In particular, the first conceptual model considers a c_M variable with depth z and effective vertical stress σ_z, that is, heterogeneity due to litostratigraphic variability according to *Baù et al.* [2002] as discussed in Section 2.4. However, due to the strong compartmentalization of the *Maja* reservoir, partitioned into various blocks by faults and thrusts, the second scenario addresses a possible heterogeneity within the horizontal plane, that is, c_M is assumed to be spatially distributed within the stratigraphic unit. To authors' knowledge, this is the first attempt to calibrate c_M based on the spatial distribution of the parameter within a compartmentalized reservoir. The presence of

[1]Part of the results presented in this chapter has been published in [*Zoccarato et al.*, 2015, submitted].

sealing faults and thrusts suggests a block heterogeneous c_M for the *Maja* field with the number of uncertain parameters equal to the the number of reservoir blocks.

The constitutive law of c_M versus σ_z is assumed to be *a priori* known and obtained from previous basin characterizations from RMT surveys (Figure 2.8). The uncertainties are introduced into the model by a multiplier f_{c_M} scaling c_M. Indeed, this study is not aimed to derive a new compressibility law, rather to establish a probabilistic framework to infer the geomechanical property of the rock using data from seafloor observations above a compartmentalized geological structure.

Moreover, in the test cases, two methodologies for parameter estimation are compared. The ES based on Monte Carlo simulation (Section 5.4) and a gPCE-based ES (Section 5.5) are employed to estimate both the case with one random variable, that is f_{c_M} uniform in the reservoir domain, and the case with multiple random variable, i.e., spatially variable f_{c_M} in the horizontal plane.

The real observations are assimilated with the results presented in Section 5.6 where the tests on uniform and spatially variable c_M are discussed and summarized.

5.2 The *Maja* Gas Field

The *Maja* gas field is an offshore hydrocarbon reservoir that was developed over ten years. As mentioned above, the field location cannot be disclosed for confidentiality reasons. The gas production is obtained from the main gas pools A, B, and C which are not hydraulically connected (Figure 5.1). The fluid pore pressure distribution is predicted by the reservoir multiphase simulator $Eclipse^{TM}$ through a history matching of the measured wellbore fluid pressures and production. The pressure change ΔP at the end of the ten-years production life is illustrated in Figure 5.2. The reservoir compartmentalization, namely the accumulation of gas into a number of individual fluid/pressure compartments [*Jolley et al.*, 2010], can be derived from the distribution of pressure change caused by the field development. The largest ΔP is equal to -75 bar and is experienced in the intermediate pool B at the end of the field production. The waterdrive is compartmentalized is divided into three main blocks with $\Delta P \sim$ -20 bar, -45 bar, and -30 bar, respectively. Pool A shows a different compartmentalization with

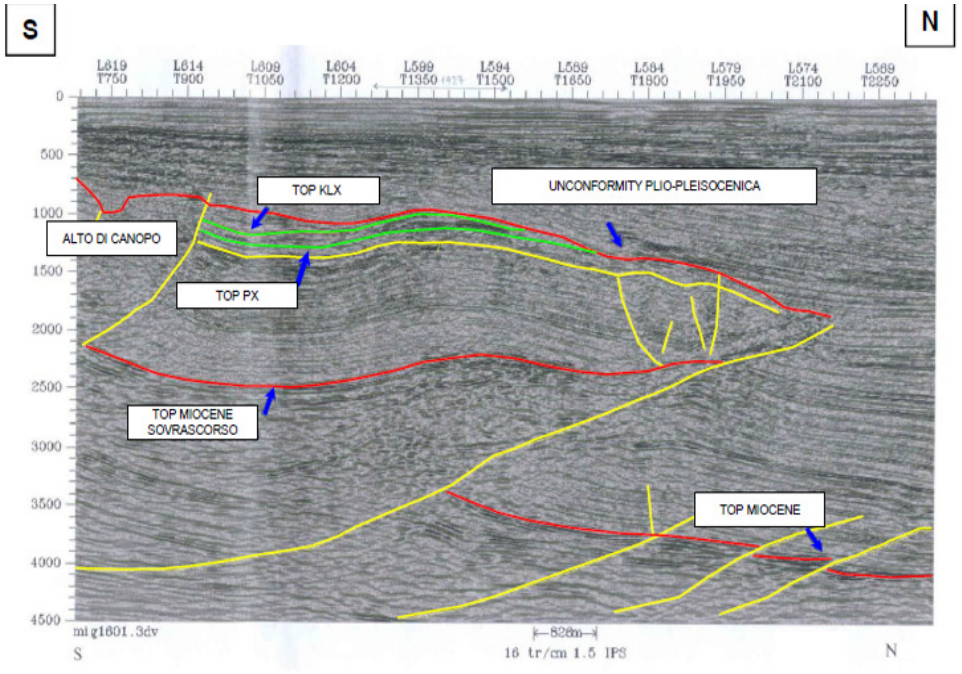

FIGURE 5.1: South-North structural cross-section through the *Maja* reservoir.

ΔP varying from a -45 bar to -18 bar. Negligible ΔP is found in the deeper gas pool C except for the central block. Average reservoir porosity ϕ ranges from 15% to 29%. In pools A and B, the range of the horizontal permeability k is 30-250 mD and 13-40 mD, respectively. In pool C, k equals 13 mD.

The three-dimensional grid of Figure 5.3(b) is made up from 320,901 nodes and 1,824,768 tetrahedral elements to cover a domain of 52 km x 49 km x 5 km. The total number of elements that undergoes pressure variation is 54,720 in the reservoir layers of Figure 5.3(a). Zero-displacements conditions are prescribed on the lateral and bottom boundaries while the top of the domain, i.e., the seafloor, is traction-free.

5.3 Measurements from bathymetric survey

The bathymetry is the measurement of the depth of a water body and corresponds to the topography on the sea bottom. The difference between two bathymetric profiles recorded at different times provides a measure of the displacement of the sea bottom within the time-lapse of the two surveys. A multi-beam echosounder system is employed to acquire the bathymetry above the *Maja* gas field and a large sorrounding

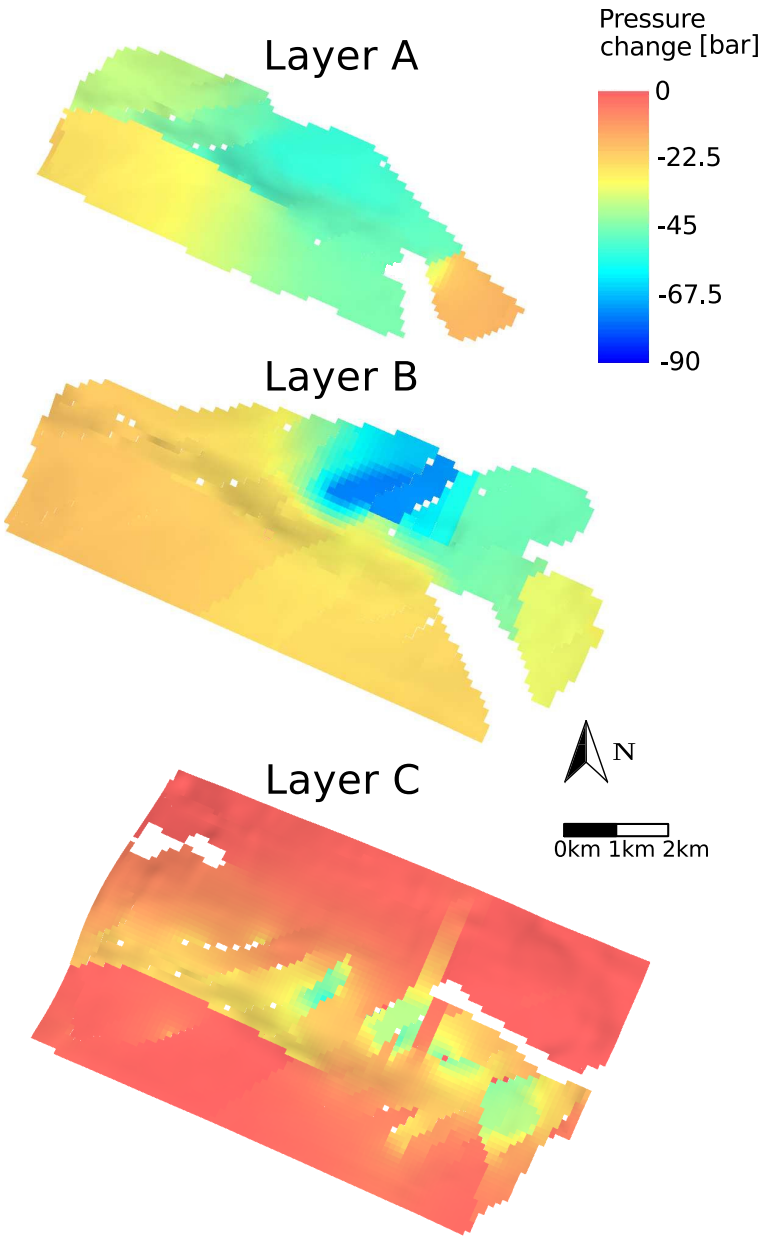

FIGURE 5.2: Maps of the pressure change within *Maja* gas layers A, B, C (including waterdrive). The pore pressure change ΔP is experienced over ten years of production.

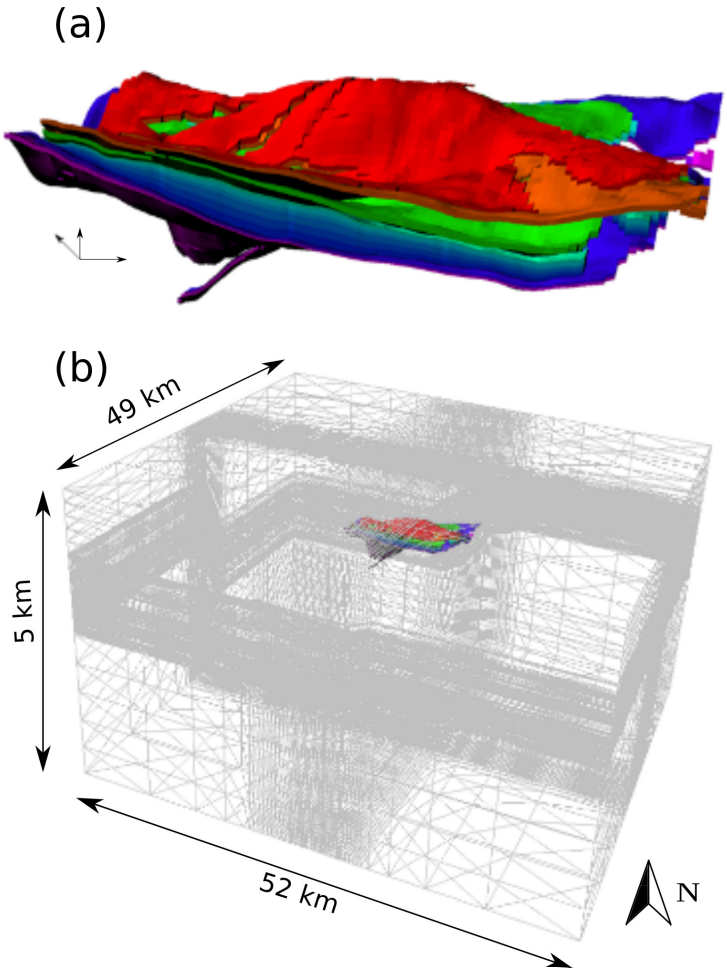

FIGURE 5.3: (a) 3D Reservoir production model and (b) 3D Finite Element (FE) geomechanical model of the *Maja* gas field with the coloured elements corresponding to the productive units. The colours in (a) are representative of the reservoir layers.

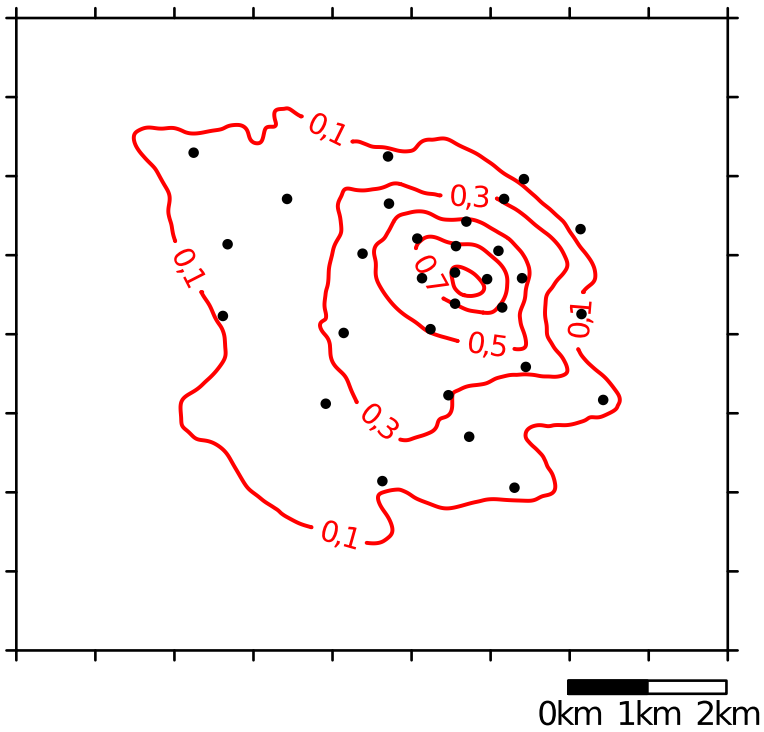

FIGURE 5.4: Subsidence contour-lines derived from multi-beam bathymetric surveys above the *Maja* gas field. The measured values are normalized to the value of the maximum displacement, u_{max}. The black dots are the assimilation data points as discussed in Section 5.6.

area. The multi-beam technique is based on a sound wave that is emitted from a ship. The measurement of the travel time that the wave takes to bounce off the seabed and return back to the receiver is proportional to the water depth. In the area above the *Maja* gas field, multi-beam acquisitions at the seafloor depth before and after the reseroir development provides a map of the seafloor subsidence caused by gas production.

Figure 5.4 shows the vertical displacements of the seafloor observed in the time-lapse of ten years. For confidentiality reasons, the contour lines are normalized to the measured peak value, u_{max}.

5.4 Test case: Ensemble Smoother for c_M estimation

Geomechanical simulations of the *Maja* reservoir are performed employing the same finite-element (FE) poro-elasto-plastic numerical model used for the previous UGS field [*Gambolati et al.*, 2001]. In this study, an isotropic stress-strain constitutive law is used with the vertical uniaxial compressibility c_M that depends on the stress state according

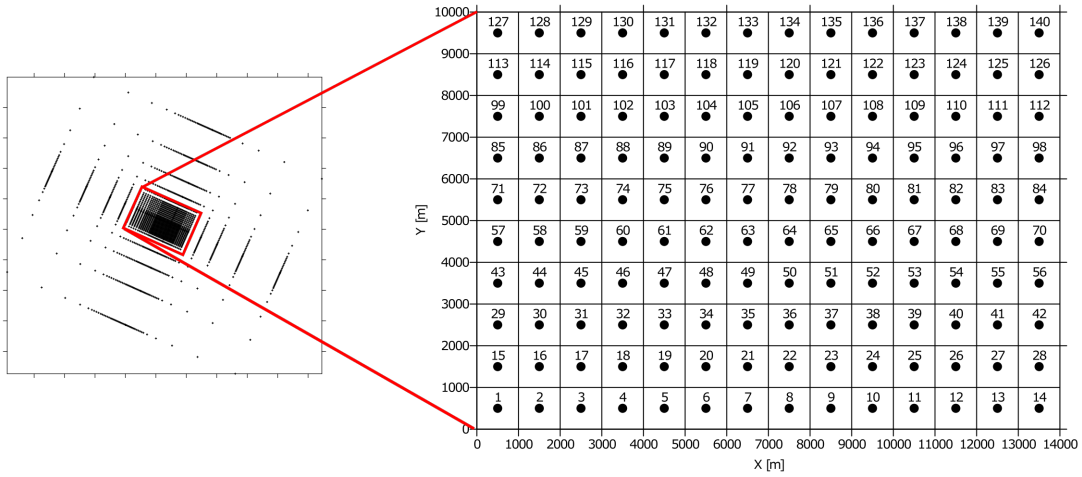

FIGURE 5.5: Two-dimensional view of the geomechanical FE grid of Figure 5.3. The enlargement view within the red rectangle refers to the area affected by the pore pressure change. In the synthetic simulations the area has been subdivided in 140 sub-zones.

to the hypo-plastic hysteretic model developed by *Baù et al.* [2002] and *Ferronato et al.* [2013]. The uncertain compressibility c_M is calibrated by introducing a spatially variable multiplicative factor f_{c_M}, which allows scaling c_M values in the regions where fluid pressure changes occur. Poisson's coefficient is assumed to be known and equal to 0.3. The complete description of the isotropic constitutive law used in these simulation is provided in Section 2.4.

A preliminary set of simulations has been carried out with representative data to test the proposed methodology. The simulations span one year when the reservoir experiences fluid extraction. The pressure data are obtained from the history matching of the real case where the reservoir production spans ten years. The pressure data are taken as the pressure variation during the first year. The reservoir domain is shown in Figure 5.3.

5.4.1 Prior Distribution of Uncertain Parameters

In this investigation, c_M via f_{c_M} is assumed to be the only uncertain geomechanical parameter. Because the ES relies on a Monte Carlo approach, a prior probability distribution function (PDF) is needed to sample the prior ensemble of the multiplicative

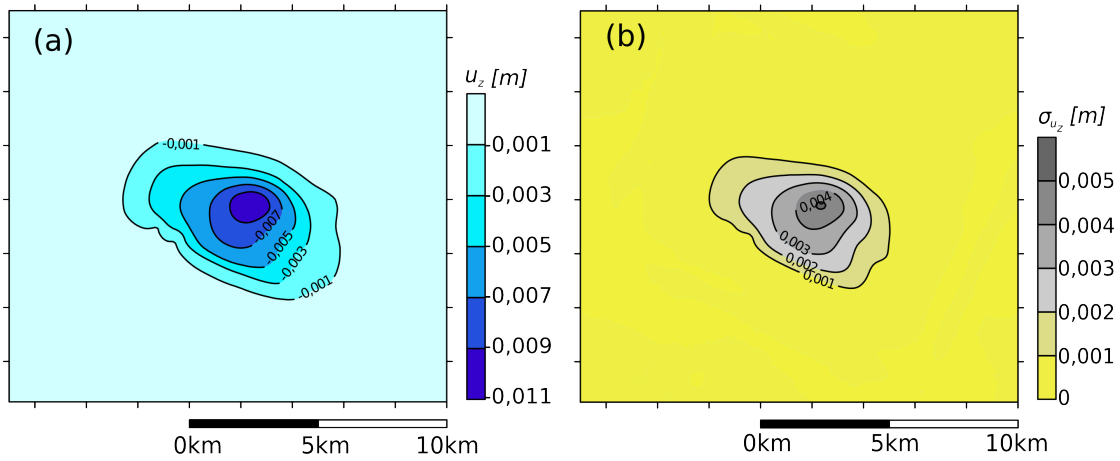

FIGURE 5.6: Homogeneous case: (a) mean and (b) standard deviation, σ, from the forecast ensemble of the vertical displacements, u_z.

calibration factor f_{c_M}. In this section, the generation of the prior PDFs in two conceptual models is described: (1) f_{c_M} is uniform within the reservoir, and (2) f_{c_M} is spatially distributed. In the latter, the heterogeneity on f_{c_M} occurs only in the area shown in the enlargement of Figure 5.5 where the pressure changed due to gas production.

Homogeneous Compressibility - One random variable

The calibration factor f_{c_M} is assumed to be spatially uniform within the area of Figure 5.5. The values of the prior f_{c_M} ensemble are randomly sampled from a uniform PDF spanning the interval between 1 and 10:

$$f_{c_M} \sim U[1, 10] \tag{5.1}$$

The selected range is based on the outcome of a sensitivity analysis carried out to investigate the possible interval of the f_{c_M} variation. Figure 5.6 shows the spatial distribution of the mean and the standard deviation (σ) of the vertical displacements (u_z) from the forecast ensemble obtained by performing 100 Monte Carlo geomechanical simulations using the prior f_{c_M} ensemble.

Heterogeneous Compressibility - Multiple random variables

The f_{c_M} is spatially distributed within the red area of Figure 5.5. This area is 14 $\times$ 10 km^2 wide and subdivided into 140 square cells. Each cell is potentially characterized by a different value of f_{c_M}. A categorical indicator algorithm, that creates random realizations of a heterogeneous f_{c_M} field according to a given covariance model, is used. The f_{c_M} values are drawn from a discrete uniform distribution with ten prescribed categories ranging from 1 to 10. Each category has an equal unconditional probability (1/10). To account instead for the spatial statistical dependence between f_{c_M} values on different cells, a stationary correlation model is introduced. According to this model, the probability of observing a f_{c_M} category in the discrete domain [1,10] at any given cell depends on the distance from the value of f_{c_M} at surrounding cells. The correlation between two grid cells is based on an exponential isotropic function depending on the distance between these cells and a prescribed correlation length, λ. The λ value has a direct influence on the degree of heterogeneity assigned to the spatial distribution of f_{c_M}. Consistent with the size of the cells and also based on the outcome of preliminary simulations, λ = 4000 m is assumed. Figure 5.7 shows one of the 100 realizations of the generated prior ensemble of the f_{c_M} field. Obviously, the mean over the ensemble in each grid block is equal to 5. Similar to the homogeneous case, the mean and the standard deviation of the u_z forecast ensemble are shown in Figure 5.8.

5.4.2 Synthetic Land Subsidence Data

Ground-surface displacements data are used to infer the model state and the geomechanical parameters. The observations are taken from the land subsidence map of Figure 5.10, which is obtained from a geomechanical reference simulation with a prescribed and 'known' compressibility distribution (see Table 5.1 and Figure 5.9). The f_{c_M} field is assigned on the basis of a plausible reservoir partition derived from the presence of faults and thrusts (Figure 5.10). The synthetic data locations are uniformly distributed over the reservoir (Figure 5.10) in the area where the mean u_z (Figure 5.6a and 5.8a) is significant, i.e., $|u_z|$ greater than 0.001 m. In the same area, the standard deviation σ is also approximately greater than 0.001 m (Figure 5.6b and 5.8b).

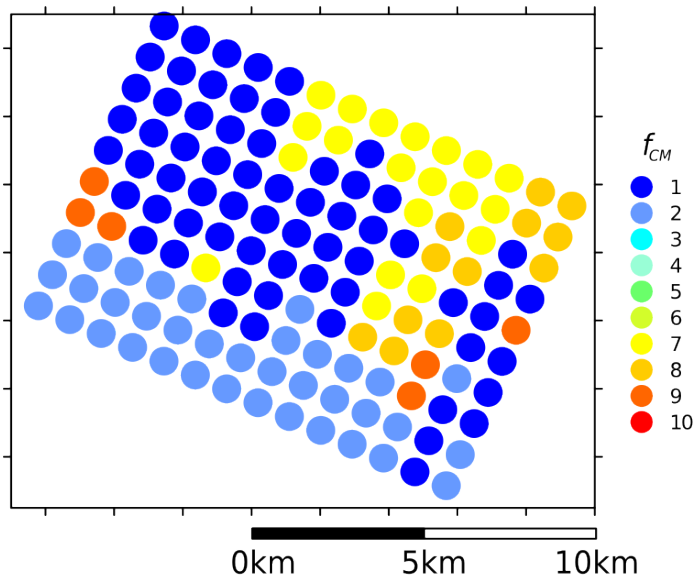

FIGURE 5.7: Heterogeneous case: one realization of the 2D f_{c_M} field ensemble. The compressibility varies in the area corresponding to the enlargement of Figure 5.5

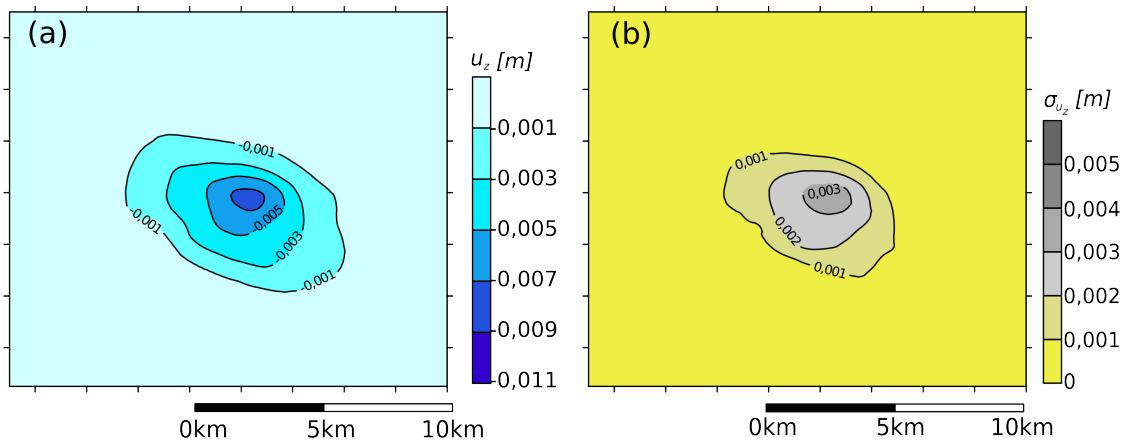

FIGURE 5.8: Heterogeneous case: (a) mean and (b) standard deviation, σ, from the forecast ensemble of the vertical displacements, u_z.

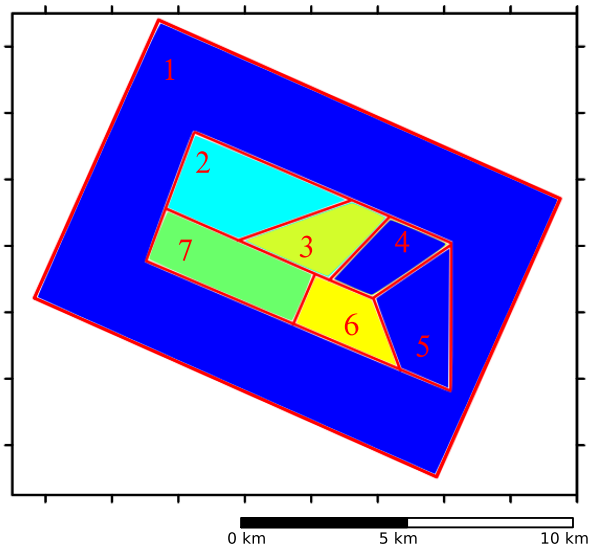

FIGURE 5.9: Prescribed f_{c_M} distribution for the simulation of the synthetic data provided in Figure 5.10

TABLE 5.1: True reference values of f_{c_M} used to simulate the spatial distribution of the synthetic land subsidence.

Reservoir block #	f_{c_M}
1	1.0
2	3.0
3	6.0
4	1.0
5	1.0
6	7.0
7	5.0

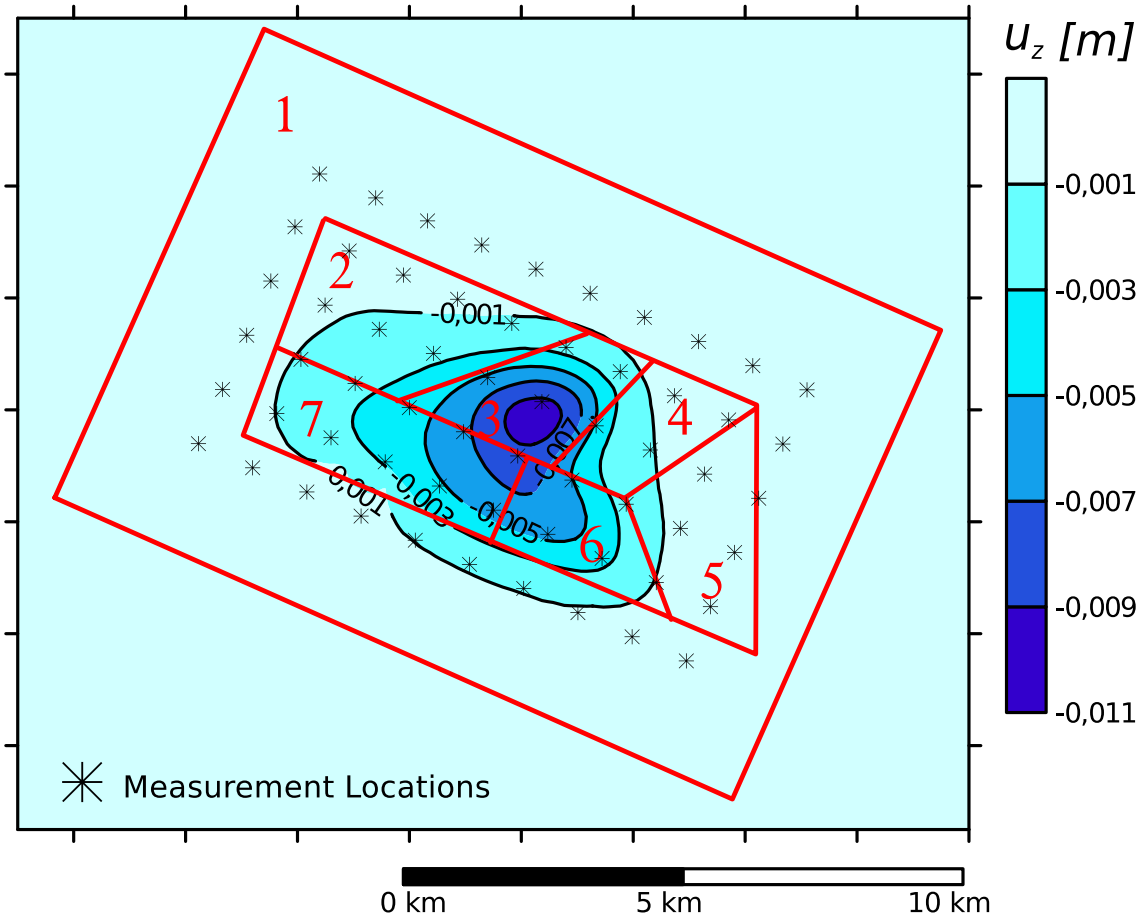

FIGURE 5.10: Spatial distribution of the synthetic land subsidence data, u_z. Assimilation data are collected at the 60 measurements locations displayed in the map. The red lines represent the trace of the faults partitioning the reservoir into seven blocks. The prescribed reference f_{c_M} values of the blocks are reported in Table 5.1.

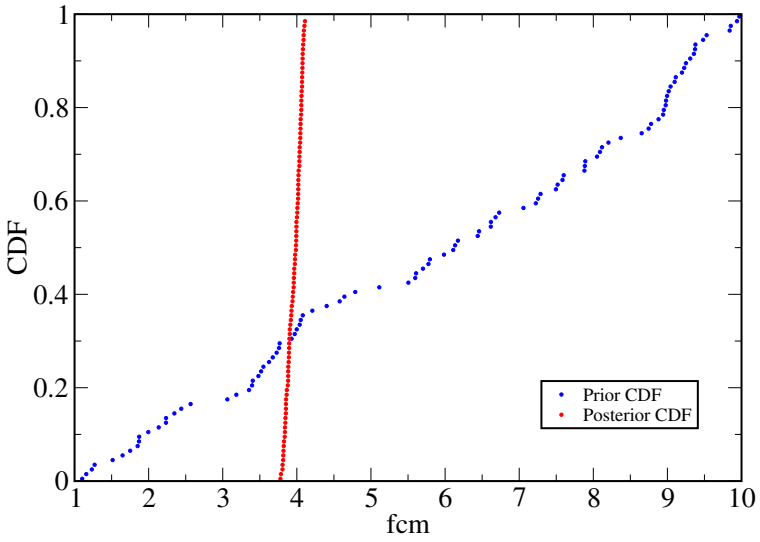

FIGURE 5.11: Homogeneous case: prior and posterior CDFs from the f_{c_M} update ensemble.

5.4.3 Results and Discussion

Update of f_{c_M}

In case of a single random variable, a significant reduction of the uncertainty associated with the prior f_{c_M} distribution is obtained, as shown in Figure 5.11 comparing the prior and the posterior f_{c_M} cumulative distribution functions (CDFs). Using the definition of the ensemble spread (AES) given in Equation (4.5), the prior AES, i.e., before the assimilation of data, equals 2.41 and reduces by about 93% after assimilation. Therefore, the posterior CDF is highly constrained compared to the prior CDF with the most probable estimate value for f_{c_M} equal to 4.0, corresponding to the mean, or expected value, of the updated ensemble.

If a multiple random variable is used, the ES performance is evaluated by calculating the AES index over each grid block of Figure 5.5. The prior AES index ranges between 2.0 and 2.9 over the domain. After assimilation, the spread of the updated ensemble significantly reduces over the area where data points are collected, while higher AES values are found in the surrounding area (see Figure 5.12b). A sensitivity analysis reveals that collecting data over the entire domain does not improve the assimilation outcome. Indeed, the previous observations cannot yield enough information to infer the model parameters because the deep reservoir experiences small pressure variation

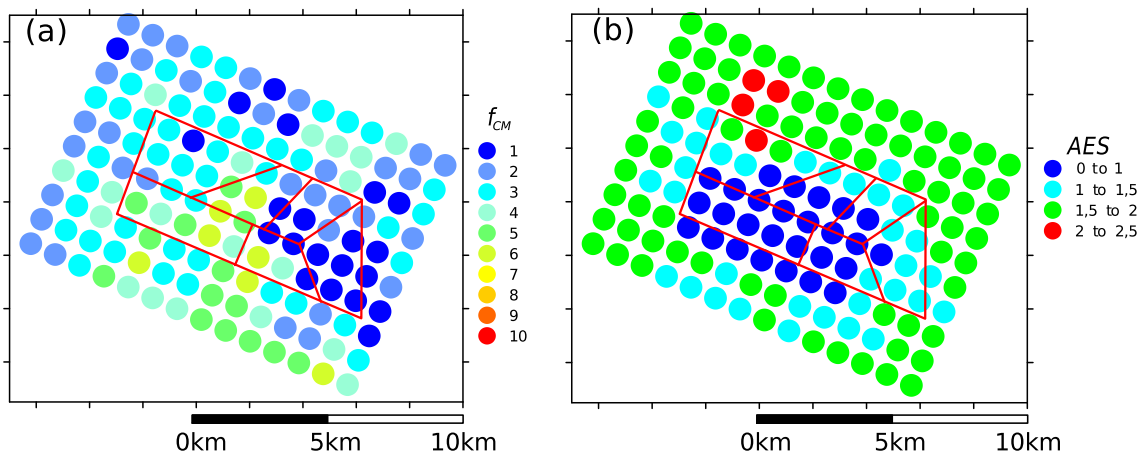

FIGURE 5.12: Heterogeneous case: spatial distribution of (a) the mean and (b) the performance index AES from the update f_{c_M} ensemble after data assimilation.

in the outer zone with a negligible influence on the land surface deformation. The average AES over all grid blocks reduces by about 33% after assimilation. Despite a lower relative reduction of AES compared to test 1, the inverse problem outcome is much better verified in terms of ground surface displacements as shown below. Figure 5.12a depicts the spatial distribution of the mean of the update f_{c_M} ensemble.

Forecast of Ground Surface Displacement

The quality of the parameter estimation is validated by performing the posterior geomechanical simulations for both uniform and spatially variable f_{c_M}. The f_{c_M} model input is the mean of the updated f_{c_M} ensemble from the outcome of test cases with one and multiple random variables. The results of these simulations are compared with the synthetic land surface data of Figure 5.10 using the Normalized Root Mean Square Error (NRMSE), which represents the standard deviation of the differences between the simulated values and the observations, calculated as

$$\frac{\sqrt{\sum_{i=1}^{N}\left(u_{z_{i,sim}} - u_{z_{i,obs}}\right)^2}}{\left|\left(u_{z_{obs,max}} - u_{z_{obs,min}}\right)\right|} \tag{5.2}$$

where $u_{z_{i,sim}}$ and $u_{z_{i,obs}}$ are the simulated and observed land ground displacement at the i_{th} assimilation location, respectively; $u_{z_{obs,max}}$ and $u_{z_{obs,min}}$ are the maximum and minimum observation values, respectively, N and is the total number of assimilation

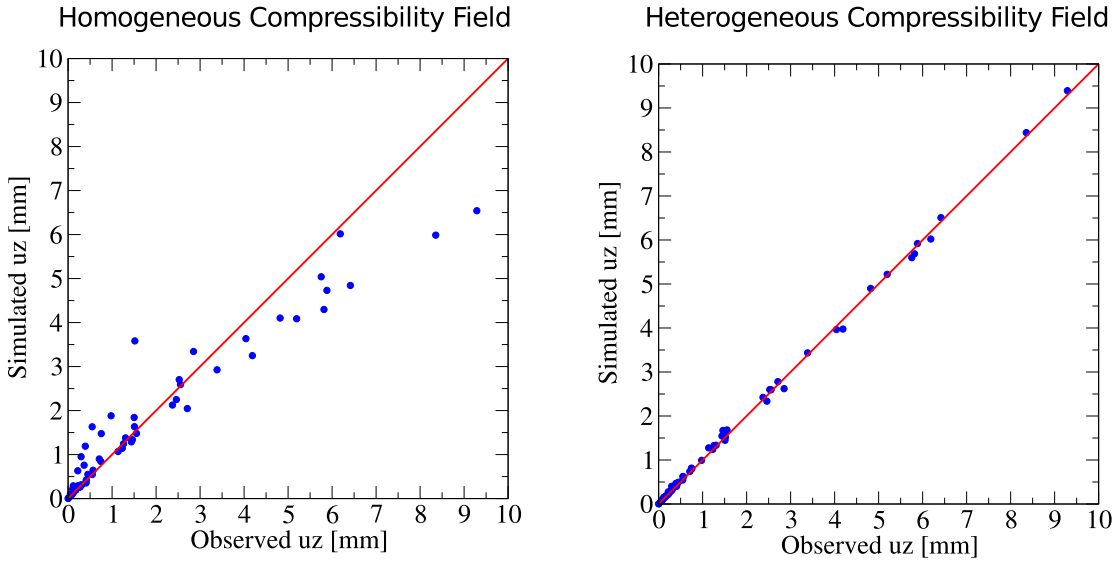

FIGURE 5.13: Simulated vs. observed values of u_z in the homogeneous and heterogeneous test cases.

locations, i.e., 60 data points. Single and multiple random variables tests give NRMSE values equal to 61% and 7%, respectively. Hence, the predicted vertical displacements provides a far better data match in heterogeneous case as shown by comparing the measured and observed displacements (Figure 5.13). The fitting of the model predictions to the reference data is very accurate with the heterogeneous c_M while underestimation and overestimation values are found in the homogeneous test case for the larger and smaller measured displacements, respectively.

5.5 Test case: gPCE-based Ensemble Smoother for c_M estimation

In this section, the calibration of the parameter f_{c_M} is obtained using the generalized Polynomial Chaos Expansion (gPCE) combined with the ES (Section 3.3). The aim is at reducing the overall computational cost of the assimilation avoiding the Monte Carlo simulations. For the homogeneous test case with only one random variable the application of the methodology is straightforward. The uncertain parameter f_{c_M} is uniformly distributed as in the ensemble-based approach from the PDF given in Equation (5.1). The model output in term of vertical displacements depends on the space $\boldsymbol{x}$ and on the

TABLE 5.2: Integration points n_i used to run the geomechanical simulations with the order of the gPCE expansion k equal to 3.

	n_1	n_2	n_3	n_4
f_{c_M}	1.6249	3.9701	7.0299	9.3751

random input f_{c_M}. The gPCE allows to write the model output $\boldsymbol{u}(\boldsymbol{x}, f_{c_M})$ as

$$\boldsymbol{u}(\boldsymbol{x}, f_{c_M}) = \sum_{\alpha=0}^{M} \boldsymbol{u}_\alpha(\boldsymbol{x}) \cdot \Psi_\alpha(f_{c_M}) \tag{5.3}$$

where $\Psi_\alpha(f_{c_M})$ are the basis functions of the expansion, i.e., polynomials. In this case, the Legendre polynomials are used because a uniform PDF is elected to describe the distribution of f_{c_M}. M is the number of the expansion coefficients $\boldsymbol{u}_\alpha$ that can be calculated by the following formula:

$$M = \frac{(N+k)!}{N!k!} - 1 \tag{5.4}$$

with N the number of random variables and k the order of the gPCE expansion. The univariate case implies $N = 1$ and chosing, for example, $k = 3$ the number of expansion coefficients M is equal to 3. To determine the number of model runs corresponding to the number of integration points Q it is sufficient to notice that n points integrate exactly polynomials up to order $2n - 1$ in the univariate case. Thus using a gPCE order 3, the polynomial degree of $\boldsymbol{u}_\alpha$ is equal to 6 and Q is equal to 4 (see Section 3.3). It is necessary to run four independent geomechanical simulations to get the coefficients $\boldsymbol{u}_\alpha$ of Equation (5.3). The four points are given in Table 5.2. Then, the response in term of vertical displacements from the geomechanical simulations is extracted at the measurement locations highlighted in Figure 5.10. Thus the assimilation points are the same used in the ES test case. The estimation of the parameter f_{c_M} is achieved with a ES scheme where the forecast matrix is built from sampling the realizations from the solution obtained with the gPCE. Table 5.3 summarizes the results of the f_{c_M} updating providing the mean and AES (Equation 4.5) reduction prior and posterior to the assimilation. The updating is performed for different ensemble size n_{MC} equal to 10, 100, 1000, and 10.000 realizations. It is clear that the number of realizations do not affect the solution of the updating scheme with a mean value for f_{c_M} equal to 4.7. The

TABLE 5.3: Mean and AES reduction of the updating f_{c_M} with varing number of realizations.

n_{MC}	mean f_{c_M}	% AES
10	4.71	99.1%
100	4.70	99.1%
1000	4.69	99.1%
10.000	4.70	99.1%

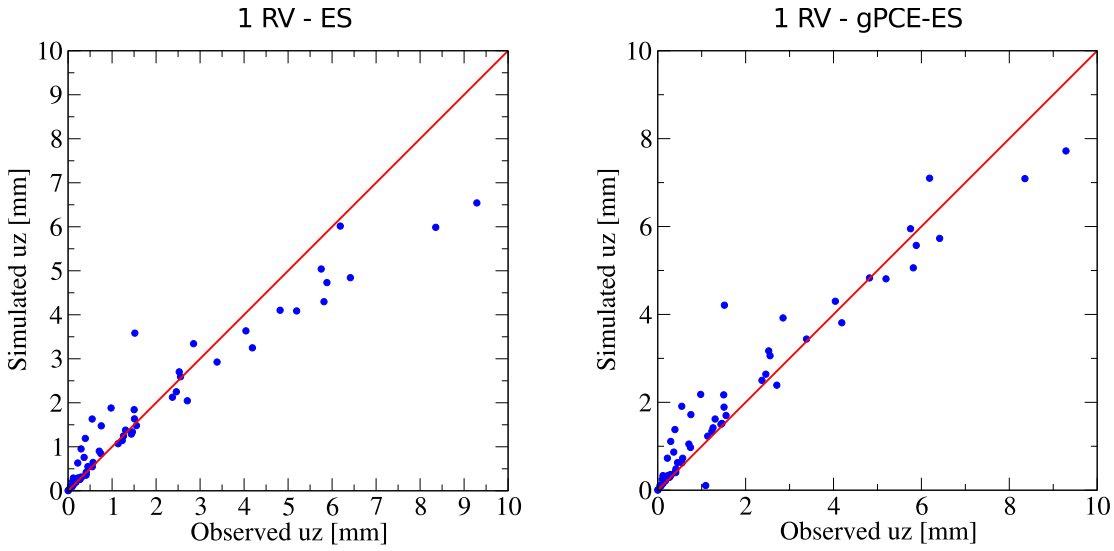

FIGURE 5.14: Simulated vs. observed values of u_z in the homogeneous case, i.e., one random variable, obtained by ES and gPCE-based ES algorithms.

reduction of the ensemble spread is very high an equal to 99.1%.

The comparison of these results with those obtained in Section 5.4.3 demonstrates the effectiveness of the two approaches for f_{c_M} estimation assimilating measurements of vertical displacements. However, the mean of the updating f_{c_M} slightly differs in the two cases. In particular, the ES scheme with Monte Carlo realizations of the geomechanical model converge to a mean value of about 4.0 (Figure 5.11) while with sampling from a gPCE expansion of the model solution with order 3 gives a value of 4.7. It is interesting to observe the result in term of displacements by using f_{c_M} from the calibration obtained with the two methodologies. Figure 5.14 shows the comparison between observed and simulated vertical displacements at the 60 data points used for the assimilation. The NRMSE (Equation (5.2)) is equal to 61% and 53 % in the two cases thus a slightly better match of the true reference is obtained with the gPCE based approach. However, the large discrepancy between observed and simulated data is not

entirely removed introducing this new approach for the calibration.

The gPCE-based ES is also tested in the case of heterogeneous f_{c_M} field where a stochastic random process in space is modeled. The stochastic random field of Section 5.4 is such that 140 random variables are subjected to calibration. For application with the gPCE method this is quite unfeasible for large-scale forward models. Recalling Equation (5.4), the number of coefficients M of the gPCE expansion are almost equal to 10^6 with $N = 140$ and $k = 3$. Thus a proper dimension reduction technique is used to reduce the number of random variables although maintaining enough approximation accuracy.

The Karhunen-Loève (KL) expansion is a widely used technique for representation of random processes as series expansions. It allows to write a random field $r(x, \omega) \in L_2(\mathcal{D}) \otimes L_2(\Omega)$ with mean $\bar{r(x)}$ and covariance function $C(x, y) = cov_r(x, y)$ as a infinite sum

$$r(x, \omega) = \sum_{i=0}^{\infty} \sigma_i r_i(x) \xi_i(\omega) \tag{5.5}$$

where $\sigma_0 \geq \sigma_1 \geq \sigma_2 \geq ...$ is a decreasing and non-negative sequence of real numbers with limit point 0, r_i are the shape functions defining an orthonormal system in $L_2(\mathcal{D})$, and ξ_i are random variables in $L_2(\Omega)$. $\mathcal{D}$ and Ω denote the spatial domain of the stochastic process and the space of the random events. $\xi_i(\omega)$ are uncorrelated random variables of unit variance:

$$\mathrm{E}[\xi_i] = 0, \quad \mathrm{E}[\xi_i, \xi_j] = \delta_{ij} \tag{5.6}$$

which depend on the probability distribution function attributed to the stochastic process. The first term of the expansion, i.e., $i = 0$, corresponds to σ_0, $r_0(x) = \bar{r}(x)$, i.e., the mean function of the stochastic field, and $\xi_0(\omega) \equiv 1$. The eigenfunctions r_i and the corresponding eigenvalues σ_i^2 are the solution to the following eigenvalue problem:

$$\int_{\mathcal{D}} C(x_1, x_2) r_i(x_2) dx_2 = \sigma_i^2 r_i(x_1), \quad x_1 \in \mathcal{D} \tag{5.7}$$

In practice, the sum in Equation (5.5) is truncated to a fintite number of terms L providing a representation of the random field with a countable number of random variables

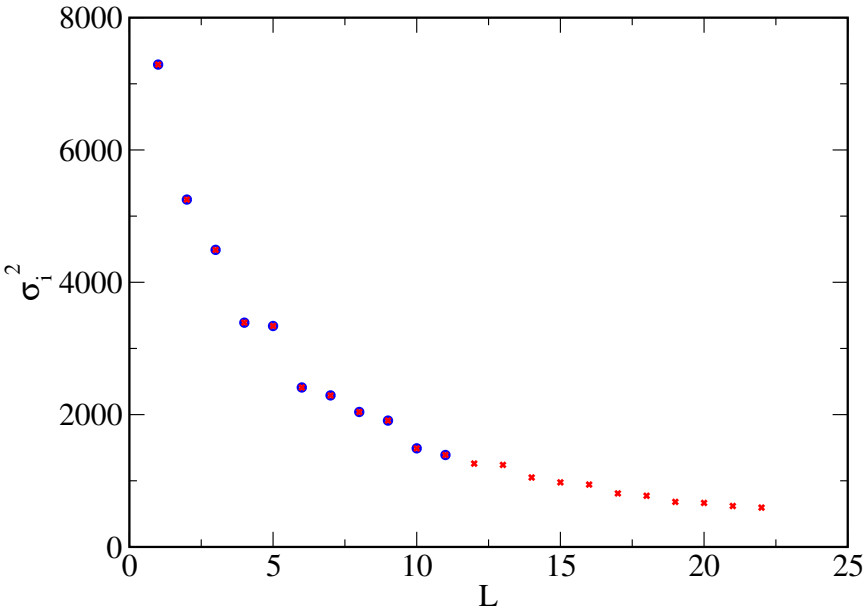

FIGURE 5.15: Eigenvalues σ_i^2 with truncation of the KLE equal to $L = 11$ and $L = 22$.

[*Zander*, 2013]:

$$r(x,\omega) = \sum_{i=0}^{L} \sigma_i r_i(x)\xi_i(\omega) \tag{5.8}$$

The number of terms of retain in the expansion mostly depends on the decay of the eigenvalues σ_i^2 with increasing index i. The Matérn covariance function $C(d)$ is here employed to describe the stastical covariance between points in the domain $\mathcal{D}$:

$$C(d) = \sigma^2 \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(\sqrt{2\nu}\frac{d}{\rho}\right)^{\nu} K_\nu \left(\sqrt{2\nu}\frac{d}{\rho}\right) \tag{5.9}$$

where Γ is the Gamma function, K_ν is the Bessel function of the second kind, d is the distance between two points, ρ and ν are function parameters. Choosing a value for ν equal to 2 and a Gaussian process the realizations of the stochastic field would be one time continuous differentiable. The correlation length is maintained equal to 4000 m as in the previous test, i.e., where the stochastic field is generated with a categorical indicator algorithm. Figure 5.15 shows the eigenvalues obtained solving the problem of Equation (5.7) varying the truncation of the expansion series with covariance from Equation (5.9). Obviously, the decay of the eigenvalues increases with increasing of the number of expansion terms L. The eigenfunction corresponding with the first 11

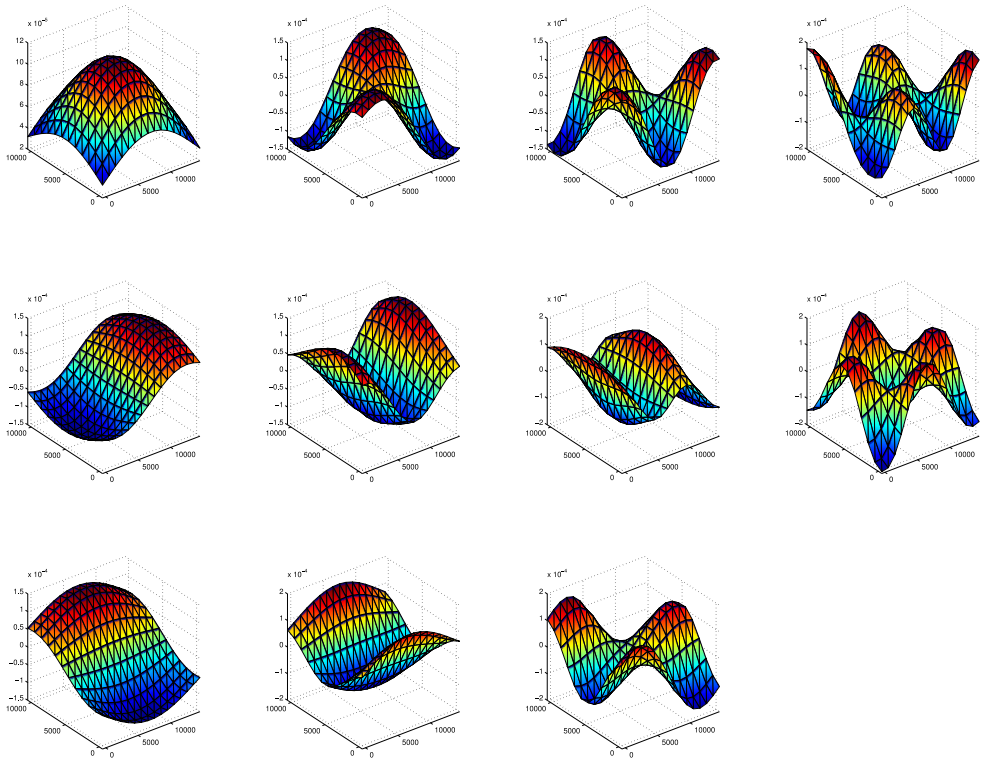

FIGURE 5.16: Eigenfunctions $r_i(x)$ corresponding to the eigenvalues σ_i^2 of the first $L = 11$ terms plotted in Figure 5.15.

eigenvalues are plotted in Figure 5.16 within the reservoir domain (Figure 5.5). Notice that if the random field has a constant variance σ^2, the error in truncating the KLE can be calculated as [*Zander*, 2013]:

$$||r_L - r||^2_{L_2(\mathcal{D})\otimes L_2(\Omega)} = |\mathcal{D}|\sigma^2 - \sum_{i=1}^{L} \sigma_i^2 \tag{5.10}$$

In the investigated case the error in variance is equal to 11.11%.

The KL expansion is used to discretize the compressibility random field, i.e, the scaling factor f_{c_M}, to reduce the number of random variables. Then, the dependent stochastic process (the displacement field) is represented via gPCE. The gPC order for the expansion of the solution is assumed equal to 2 for these preliminary tests. Thus, the number of integration points, i.e., the number of model runs, is equal to 23 from the use of a sparse grid and a gPCE order 2 with the KLE truncated to $L = 11$ terms. The assimilation results in terms of parameter estimation of f_{c_M} is shown in Figure 5.17 using an ensemble size of 100 realizations. The updated stochastic field gives a pre-

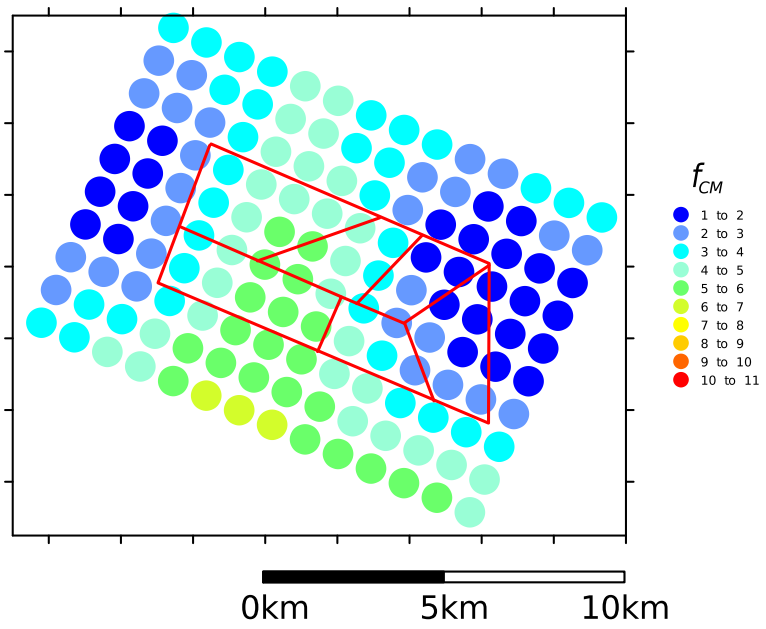

FIGURE 5.17: Spatial distribution of the updated f_{c_M} random field using the gPCE-based ES.

liminary understanding on the effectiveness of the proposed methodology. It actually produce an updated picture that is comparable with the true reference field prescribed in order to obtained the reference assimilation measurements (see Figure 5.10 and Table 5.1). Indeed, in the central portion of the domain, i.e., where the pressure depletion is actually relevant, the gPCE-based ES reduces the prior uncertainty associated with the stochastic field. This is again evaluated through the AES performance index which is reduced by about 35% after assimilation of data. The proposed gPCE-based ES seems a promising methodology because the uncertainty reduction is comparable with that obtained with the ES but at much lower computational. Indeed, neglecting the cost of the assimilation which is much lower than the cost of the forward model runs, the gPCE-ES allows to reduce the total number of geomechanical model runs from 100 to 23 realizations. Obviously, further investigations are necessary to investigate the effect of the gPCE order and the KLE truncation on the paramter update.

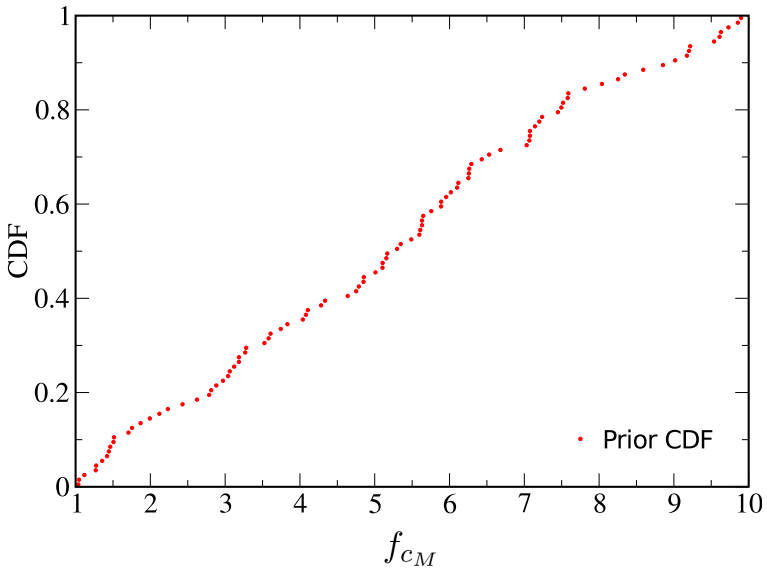

FIGURE 5.18: Prior sample Cumulative Didtribution Function (CDF) of the multiplier f_{c_M}.

5.6 The *Maja* Case Study

5.6.1 Conceptual model 1

In the real case, the conceptual model 1 considers only one prior unknown parameter f_{c_M}. The *a priori* uncertainty for f_{c_M} is modeled by randomly sampling from a uniform PDF within the range from 1 to 10:

$$f_{c_M} \sim U[1, 10] \tag{5.11}$$

The use of a uniform PDF is such that each member is equally probable sampled, reflecting the lack of information on f_{c_M}. Preliminary sensitivity analysis has been performed to select the f_{c_M} range to include the measured peak value of land subsidence within the forecast ensemble. The cumulative distribution function (CDF) of the prior ensemble for f_{c_M} is shown in Figure 5.18. The sample CDF is roughly uniform and departure from uniformity is due to the finite ensemble size, n_{MC}, of 100 members.

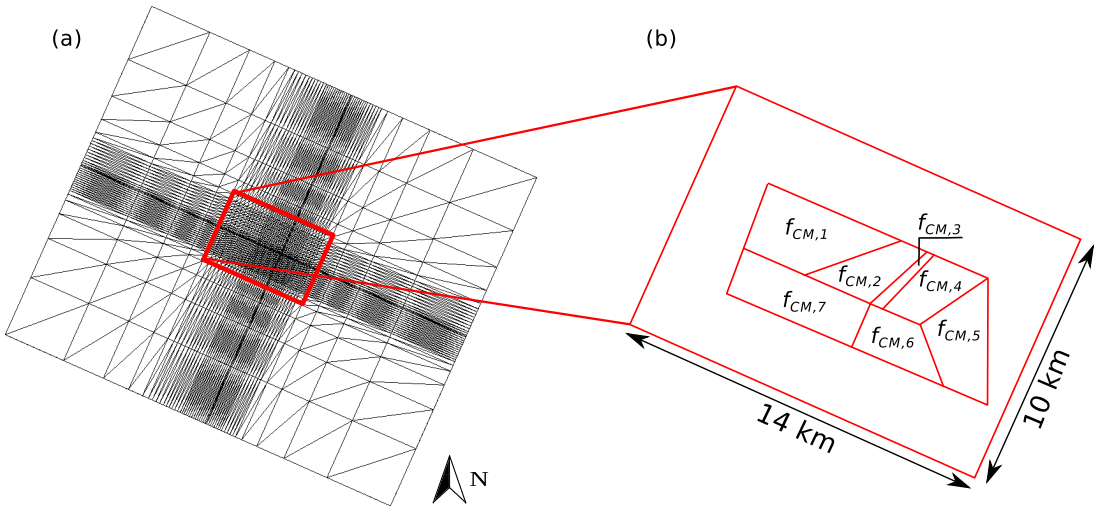

FIGURE 5.19: (a) 2D view of the geomechancial grid (see Figure 5.3(b)) and (b) sketch of the faults distribution. Each reservoir block is assigned with a different multiplier, $f_{c_{M,i}}$.

5.6.2 Conceptual model 2

The conceptual model that follows is intended to better match the observed data by accounting for the geological structure of the *Maja* reservoir. The multiplier f_{c_M} is variable in space within the reservoir. This is suggested by the strong compartmentalization of the reservoir. Heterogeneous distribution of f_{c_M} is assumed within the production model domain where the depletion occurs, that is, within the red rectangle of the 2D view of the grid domain in Figure 5.19(a). This subdomain is partitioned into seven blocks which represent a plausible compartmentalization based on the distribution of sealing faults and thrusts detected from the 3D seismic investigation (Figure 5.19b). Thus, the seven blocks are identified and each of them can assume a different f_{c_M} value. The parameters $f_{c_{M,i}}$ $(i = 1, ..., 7)$ are assumed to be uncorrelated and sampled from a uniform PDF within the interval between 1 and 10:

$$f_{c_{M,i}} \sim U[1, 10] \tag{5.12}$$

The CDFs of the model parameters $f_{c_{M,i}}$ $(i = 1, ..., 7)$ are similar to the one of Figure 5.18. The mean and standard deviation of the prior ensembles are of about to 5.5 and 2.6, respectively. Overburden, underburden and sideburden are modeled with a linear elastic material. A set of 100 Monte Carlo simulations n_{MC} are used as in the prevoius

case.

5.6.3 Results

In this section, the forecast of the vertical displacements obtained with the Monte-Carlo geomechanical simulations and the parameter updating via the ES algorithm are described in detail for both conceptual models 1 and 2. Furthermore, the updating of the seabed subsidence is carried out with the calibrated multiplier f_{c_M} as constrained in the analysis step.

Conceptual model 1 - Homogeneous f_{c_M}

The Monte-Carlo simulations of the geomechanical model are carried out by running the prior ensemble of f_{c_M} realizations. An overview of the forecast stochastic simulations is provided in Figure 5.20(a) and Figure 5.20(b), where the mean, $\mu_{v,prior}$, and the coefficient of variation, $C_{v,prior}$, of the displacement field are respectively shown over the reservoir domain at the end of the ten-years production life. Note that $\mu_{v,prior}$ is normalized to u_{max}. $C_{v,prior}$ provides an estimate of the variability of the ensemble in relation to the mean $\mu_{v,prior}$. The standard deviation is often used in literature to quantify the ensemble dispersion. However, in this case, the coefficent $C_{v,prior}$ is preferred beacuse it avoids the use of the data units. The comparison between the measurements of Figure 5.4 and $\mu_{v,prior}$ as provided in Figure 5.21 shows that the peak value of the $\mu_{v,prior}$ field is slightly shifted to the west and the computed subsidence bowl generally overestimates the areal extent of the observed settlement bowl. The maximum $\mu_{v,prior}$ is found above the reservoir areas where the greater pressure depletion occurs. A $C_{v,prior}$ of about 24% is found over the major area of the reservoir domain with a progressively decrease as the distance from the peak of the computed $\mu_{v,prior}$ increases, i.e., toward the outer region of the domain. The parameter updating is performed by assimilation of data provided by the bathymetric survey. The parameter estimation is carried out with the subset of 30 observations shown in Figure 5.4. The data points are chosen such that the interpolated values over the domain reproduce the measured subsidence (Figure 5.4) with acceptable accuracy. Note that each data point is associated with the nearest node of the FE grid. The ES updating algorithm requires that an error

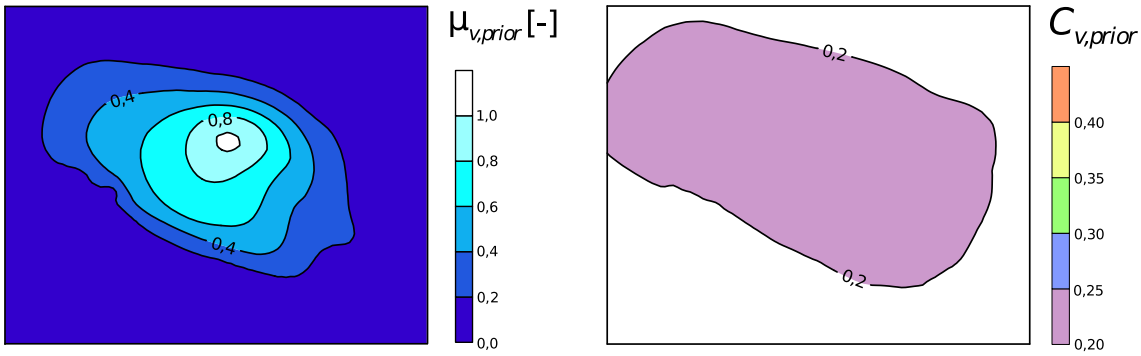

FIGURE 5.20: Model 1: forecast ensemble of the seabed subsidence over the reservoir domain in terms of (a) mean, $\mu_{v,prior}$, normalized to u_{max}, and (b) coefficient of variation, $C_{v,prior}$.

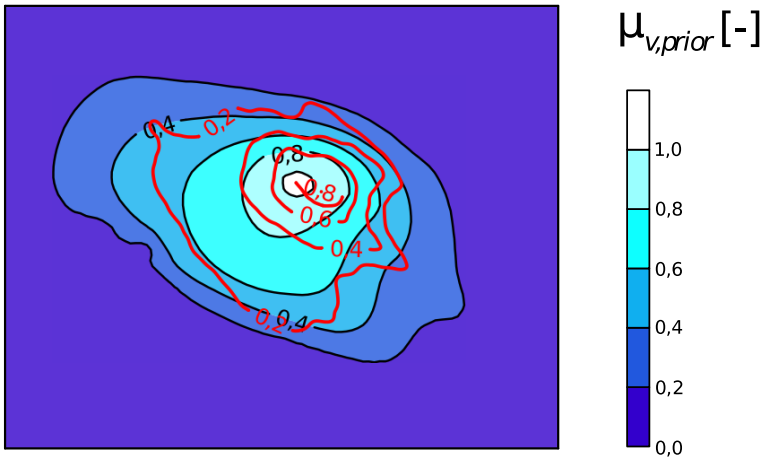

FIGURE 5.21: Comparison between the measured subsidence (red contour-lines) and the mean of the forecast subsidence, $\mu_{v,prior}$, (black contours).

must be assigned to each collected observation. The measurement perturbation ϵ_D is randomly sampled from a Gaussian PDF of the measurement error with zero mean and variance σ_ϵ^2, that is, the so-called Gaussian white noise. Two scenarios A and B are investigated in this Section. They depend on the σ_ϵ which is assigned to the measurement error. In scenario A, all data assume the same $\sigma_A = 0.0625$ based on the accuracy of the bathymetric measurements, whereas in scenario B σ_B changes within the domain. σ_B varies depending on the distance between the location of the measurement and that where u_{max} occurs. In particular, σ_B ranges between 0.025 at the subsidence peak and 0.125 at the most far location, i.e., where the subsidence is smaller and, therefore, the measurements more uncertain. The other data points are characterized by a σ_B value as obtained by a linear interpolation between the two extremes depending on the relative distance. Consequently, in scenario A, all observations have the same weights in the assimilation procedure while, in scenario B, measurements away from the centre of the subsidence bowl are characterized by a lower weight than those closer to the displacement peak. In Figure 5.22 the measurement error PDFs are plotted for three representative data points P1, P2 and P3 which are located at the centre of the displacement bowl and at growing distances from point P1, respectively, with σ_B increasing from P1 (center of the bowl) to P3 (farthest point from P1).

Figure 5.23 shows the calibration outcome for the multiplier f_{c_M} in both scenario A and B. The posterior and prior CDFs are compared and a significant reduction of the parameter uncertainty is found by assimilating vertical displacements from the time-lapse bathymetric map. Indeed, in both scenarios A and B, the spread of the updated CDFs is drastically reduced with respect to the prior distribution. On the other hand, one could observe that the two scenarios slightly differ and the updated CDF in B is translated to the right compared to A. This result is the consequence of σ_A and σ_B assigned to the measurement error. In scenario A, the ES assigns the same weights to all assimilated data and, thus, produces an updated ensemble for f_{c_M} that attempts to match and mediate between higher and lower displacement data. On the other hand, in scenario B a higher weight is given to the larger vertical displacement and the ES output from the updating provides a higher f_{c_M} value, which imply a larger compressibility to match the higher subsidence observations.

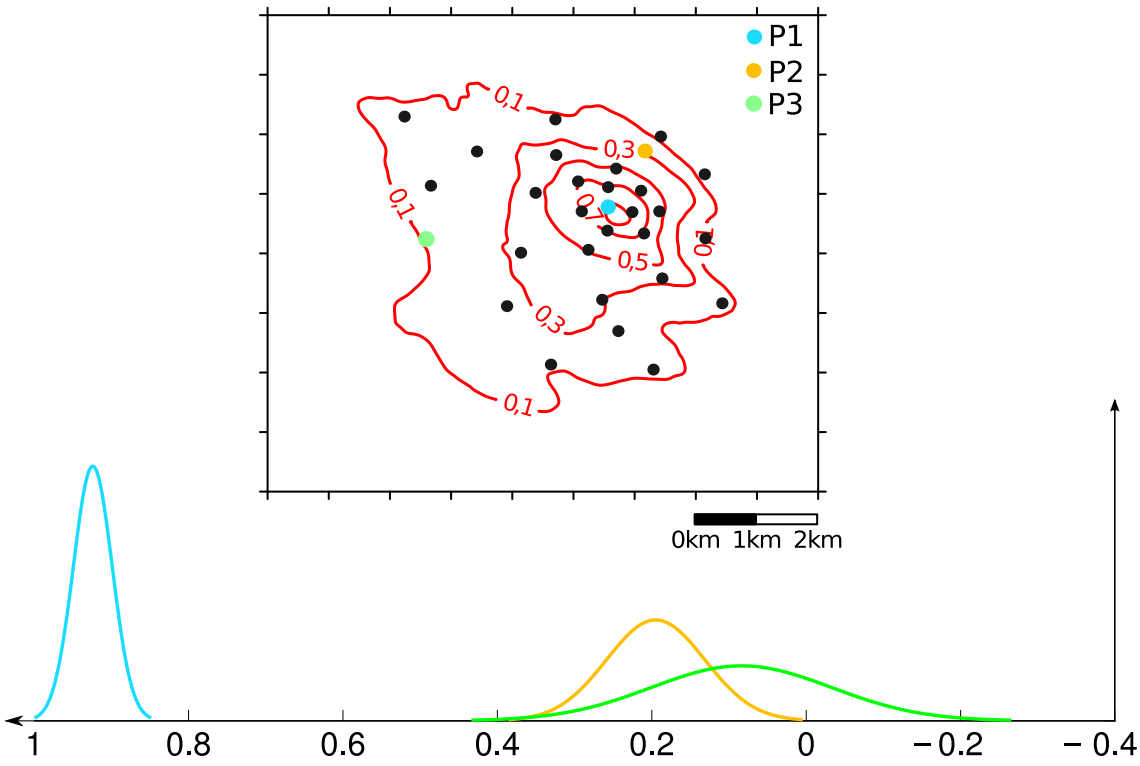

FIGURE 5.22: Measurement error PDFs at data points P1, P2 and P3: σ_B grows as the point distance increases from the centre of the subsidence bowl.

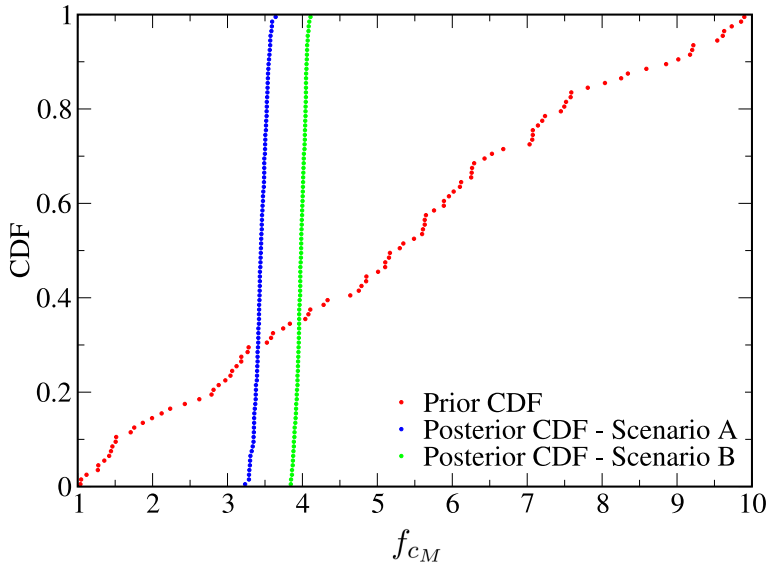

FIGURE 5.23: Model 1: prior and posterior CDFs of the updating model parameter f_{c_M} for scenario A and B.

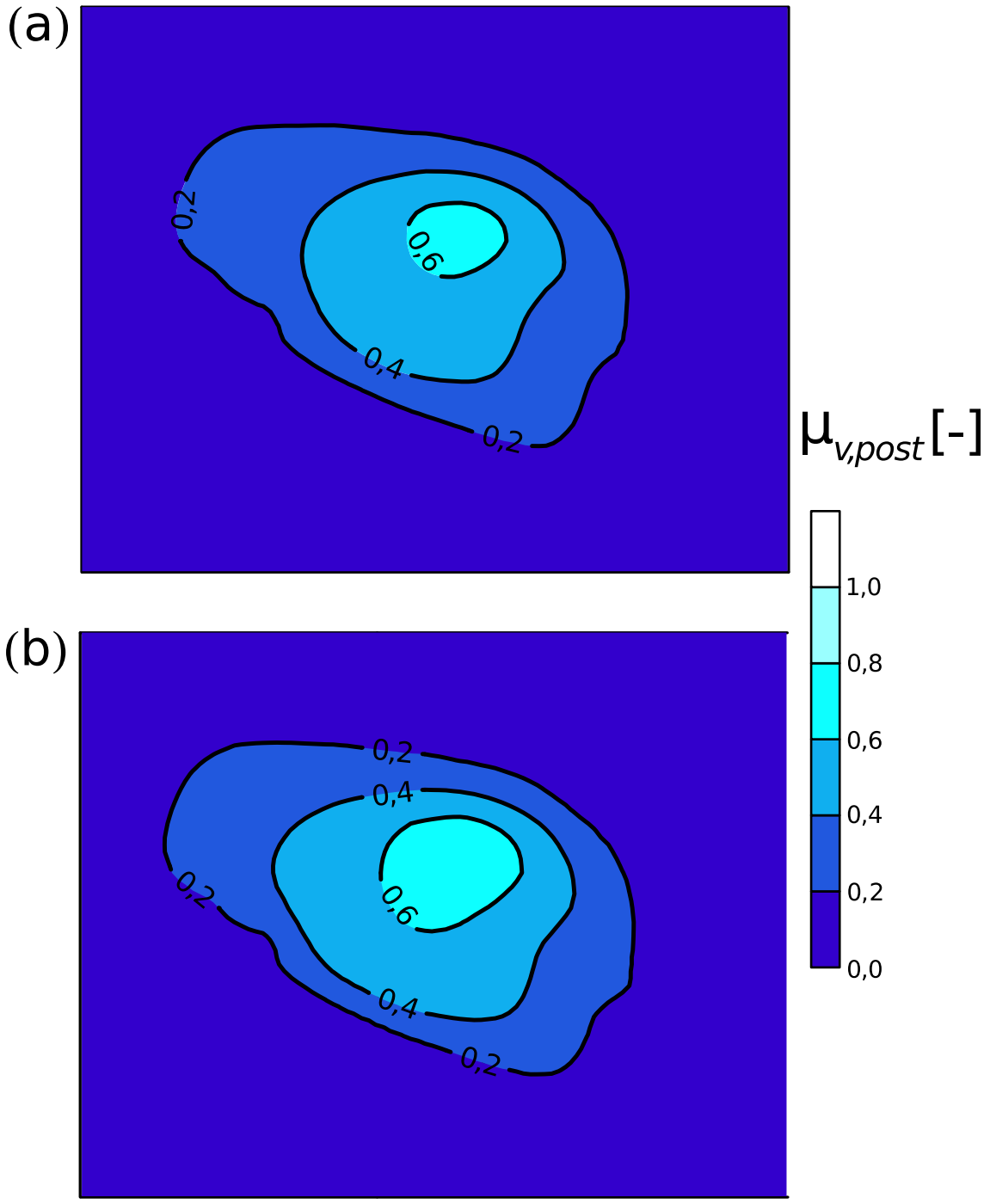

FIGURE 5.24: Model 1: updated subsidence, $\mu_{v,post}$, obtained with the f_{c_M} median after ten-year of gas extraction from scenario A (a) and scenario B (b) normalized to u_{max}.

The updated medians of the f_{c_M} ensemble, that is, the 50^{th} percentile drawn from the updated CDFs of Figure 5.23 are used to run a posterior simulation of the geomechanical model. The updated f_{c_M} median equals 3.6 and 4.0 in scenario A and B, respectively. Figure 5.24(a) and 5.24(b) show the spatial distribution of the vertical displacement of the seafloor after ten-years of gas production as computed by the updated model for the two scenarios. In scenario A, the maximum underestimation and overestimation of the time-lapse bathymetry are equal to 34% and 25%, respectively. In scenario B, the maximum underestimation is smaller than in scenario A and equals 27%. Instead, the maximum overestimation is equal to 28%, slighlty larger than in scenario A. As expected, increasing the value of the multiplier f_{c_M} from 3.6 (scenario A) to 4.0 (scenario B) yields:

1. a reduction of the maximum subsidence underestimate;
2. an enlargement of the subsidence bowl.

These contrasting effects suggest that a better match of the observations cannot be achieved assuming f_{c_M} as a single random variable within the whole model. These results point to the use of a spatial variability of f_{c_M}, as previously discussed in Section 5.6.2.

Conceptual model 2 - Heterogeneous f_{c_M}

The prior ensemble of f_{c_M}, as generated in Section 5.6.2, is run to build the forecast ensemble of the vertical displacements. The simulation spans the ten-years of production and the ensemble size consists of 100 realizations. Figure 5.25 provides the mean and the coefficient of variation from the forecast ensemble obtained with the forward model. The peak subsidence value from the mean forecast ensemble is equal to $1.01u_{max}$. It is found that conceptual model 1 and 2 provides a similar outcome in terms of forecast mean of seabed subsidence (Figure 5.20(a) and Figure 5.25(a)). By contrast, the coefficient of variation takes on higher values compared to the case of homogeneous f_{c_M} and is distributed over the domain with a pattern which resembles the f_{c_M} distribution (see Figure 5.19). It is seen that higher ΔP provides greater variations of the ensemble spread from the forecast displacements. Figure 5.25(b) shows a relative

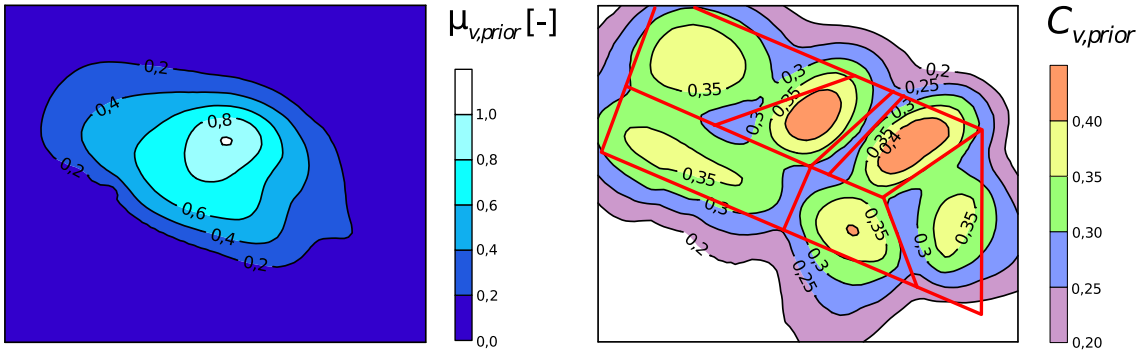

FIGURE 5.25: Model 2: forecast ensemble of the seabed subsidence over the reservoir domain in terms of (a) mean and (b) coefficient of variation. The mean values are scaled to the maximum measured vertical displacement.

TABLE 5.4: Statistics from the posterior $f_{c_{M,i}}$ ensembles.

$\#f_{c_M}$	mean μ	median m	std dev σ^2	C_v
1	3.05	2.91	0.89	0.29
2	3.59	3.58	0.28	0.08
3	9.67	9.57	1.14	0.12
4	2.02	2.08	0.64	0.31
5	2.92	2.77	1.65	0.56
6	2.69	2.66	0.50	0.18
7	1.46	1.47	0.51	0.35

variability of the forecast ensemble of about 43% and 45% in compartments 2 and 4, respectively, whereas in compartments 1, 5, 6 and 7 is slightly higher than 35%. The analysis step of the ES algorithm is carried out similarly to conceptual model 1, except that a set of seven parameters are here estimated using the information from the same data points of Figure 5.4.

The standard deviation of the measurement error σ_ϵ is uniformly distributed for all measurements and is equal to 0.0625. The results from the updating scheme are provided in Figure 5.26 and in Table 5.4. It appears that the posterior CDFs shrink for all updated f_{c_M} ensembles and the prior $\sigma_{f_{c_M}}$ reduces after the assimilation of the observations. The higher $\sigma_{f_{c_M}}$ is observed for parameter $f_{c_{M,5}}$, relative to the most eastward compartment (see Figure 5.19). The pressure variation in this reservoir block is small and do not contribute significantly to the measured vertical displacements on the seabed-surface. Thus, the information from the data points on the bathymetric map is not significant to constrain the parameter in this block. The updated $\mu_{f_{c_M}}$ are

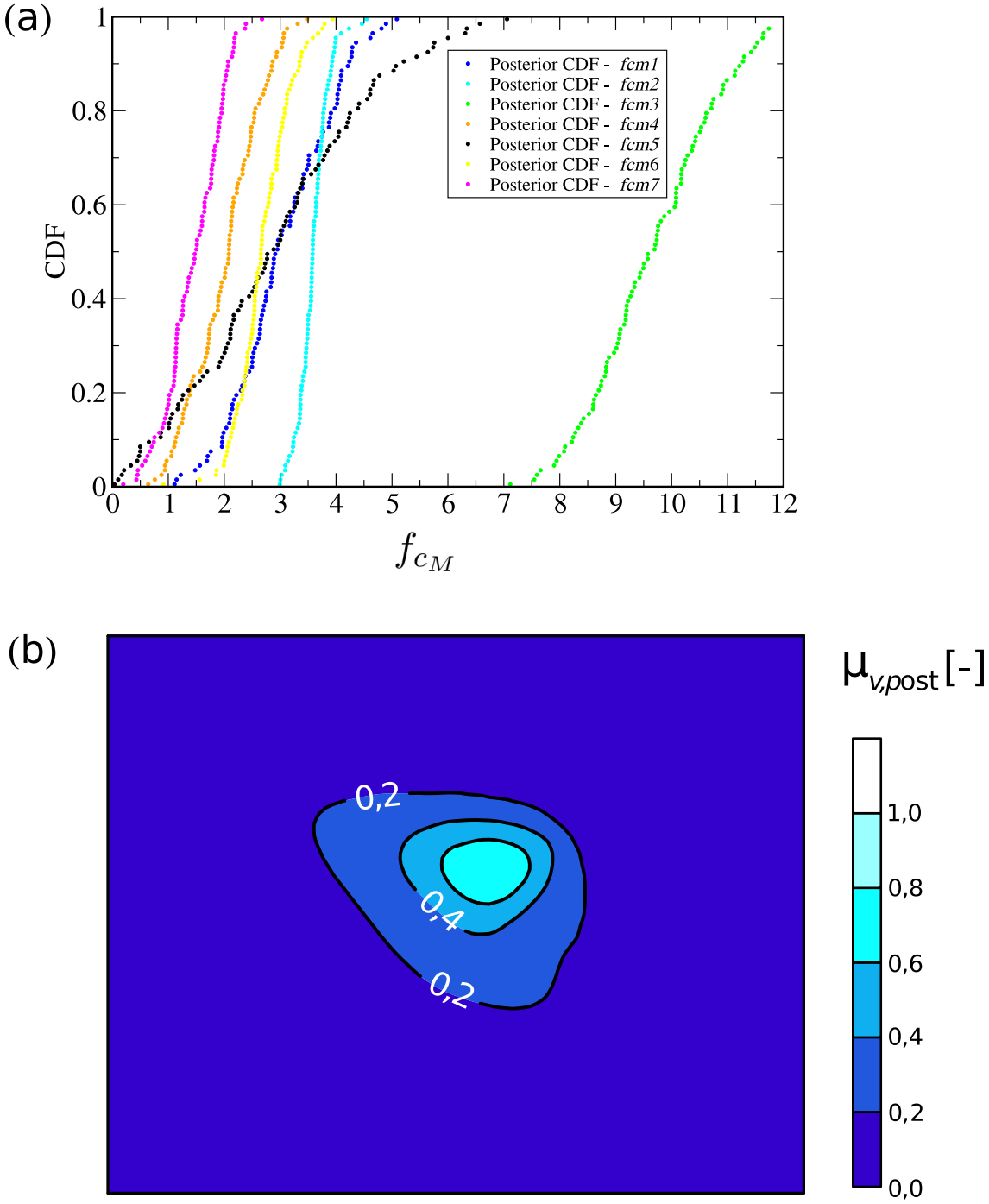

FIGURE 5.26: Model 2: (a) Posterior CDFs of the random variables $f_{c_{M,i}}$ after assimilation of vertical displacements data and (b) updated seabed subsidence obtained with the median after ten-year of gas production and normalized to u_{max}.

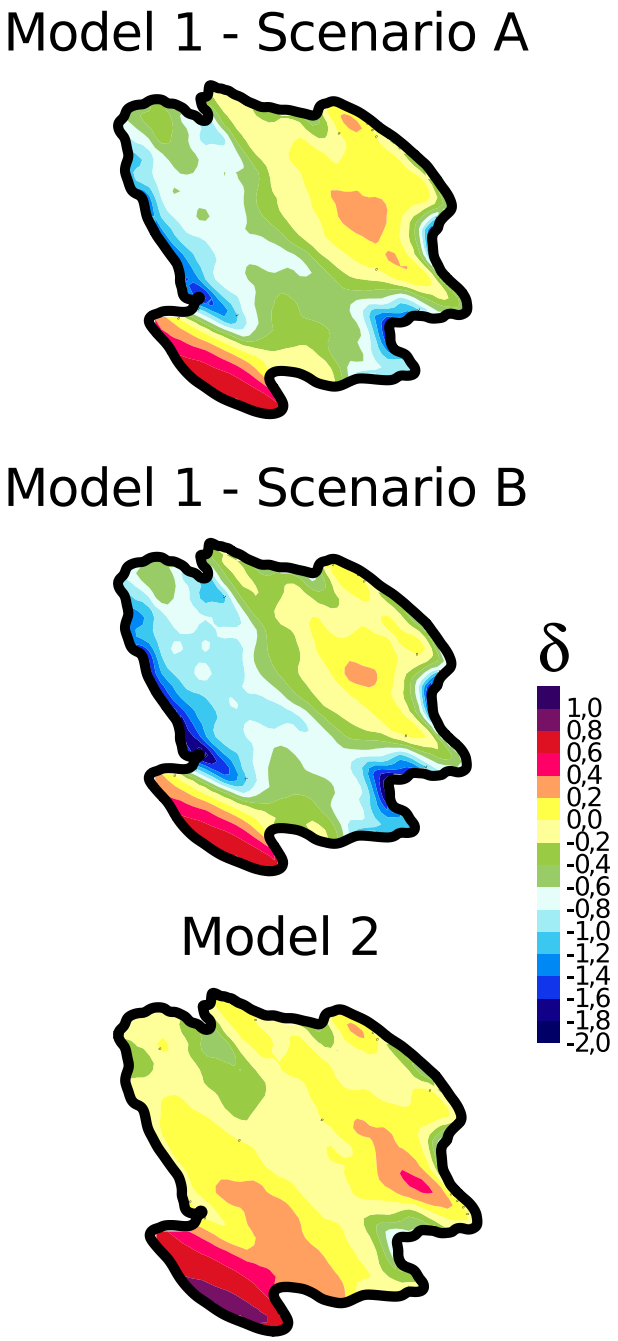

FIGURE 5.27: Distribution of the percentage error δ over the measurement area. Positive and negative values are representative of the model underestimation and overestimation, respectively.

lower than the prior mean 5.5 in the whole reservoir, except for compartment 3 where it achieves a value equal to 9.67. This almost exceeds the upper bound of the prior uniform PDF, namely 10. Thus, the ES analysis suggests that higher values for $f_{c_{M,3}}$ are required to match the measured seafloor settlement. The largest spread reduction of the prior CDF is obtained in compartment 2 with a coefficient of variation, C_v, equal to 0.08. Table 5.4 suggests that the updated parameters are simmetrically distributed as the medians of the updated CDFs are close to the mean values.

The updated median *m* of Table 5.4 are used to run a posterior geomechanical simulation. The outcome of the computation after ten-years of gas production is shown in Figure 5.26(b). The bowl of the seabed subsidence appears to be more restricted as compared to the subsidence maps of Figure 5.24. The improved outcome of the calibration procedure with model 2 compared to model 1 is shown in Figure 5.27. The percent error, defined as $\delta = 100 \times (u_{meas} - u_{sim})/|u_{meas}|$ where u_{meas} and u_{sim} are the measured and simulated vertical displacements respectively, is plotted over the reservoir domain for scenario A and B (model 1), and model 2. Positive and negative values denote the area where the models underestimate and overestimate the measurements, respectively. Scenario A and B from model 1 show a similar distribution of δ over the domain with an underestimation of about 34% and 27 % at the peak measured data point. By contrast, model 2 allows to reduce the underestimation up to 23%. The average percentage error, δ_{av}, is also evaluated over the whole domain and is defined as $100 \times (\sum_{i=1}^{N} |(u_{sim} - u_{meas})|/|u_{meas}|)/N$, where N is the total data points falling into the grid of the contour map of Figure 5.27. Model 2 provides a δ_{av} of about 20%, while scenarios A and B of model 1 give values of about 45% and 57%, respectively. Thus, the parametrization of the problem in the form of model 2, i.e., seven random variables, provides a better match of the seabed subsidence data.

5.6.4 Conclusions

The ES algorithm provides an efficient tool for reservoir geomechanical parameter estimation using observations of seafloor subsidence, i.e., vertical displacements measured through time-lapse bathymetric surveys. These data may represent an important, although indirect, information on the rock formation properties. In particular, the ES

allows for the characterization of the reservoir vertical uniaxial compressibiltiy, c_M, namely the geomechanical parameter that mostly controls the reservoir compaction due to the pore pressure depletion during fluid production. The method is herein tested on a real offshore gas reservoir with a highly complex distribution of sealing faults and thrusts that are also involved in the ES algorithm. The major conclusions can be summarized as follows:

1. data of seafloor displacements can be helpful to estimate the mechanical properties of a gas reservoir;

2. weighting the observations according to their reliability affects the outcome of the updating scheme;

3. using information on the actual reservoir geologic structure may improve the reservoir characterization;

4. the ES constrains the prior PDF of the heterogeneous geomechanical parameters where enough data are available;

5. the assumption of a heterogeneous parametrization for the compressibility, i.e., a different value of c_M, in each reservoir block may provide a better matching of the seafloor subsidence compared to the case of a uniform c_M.

Further improvements will focus on the validation of the above results using different data sources including compaction measurements from RMT. Moreover, the compressibility law could be revisited in light of the fact the only one law for the whole reservoir may not prove fully satisfactory to address the local reservoir geomechanical behaviour.

Chapter 6

Conclusions

In this thesis, a Bayesian framework is proposed for the geomechanical characterization of deep porous rock formations. The aim is at integrating, within an automated procedure, field observations into the response of subsurface simulation models, thus achieving the dual goal of (i) reducing the uncertainty in model predictions and (ii) improving the characterization of the geomechanical properties of the porous media. Specifically, the ES (Ensemble Smoother) algorithm based on Monte Carlo simulations is chosen to assimilate ground surface deformations and infer the model parameters of a FE (Finite Element) geomechanical model.

The ES inversion scheme addresses the assimilation of all available measurements in space and time within one updating step. This allows for (i) an *offline* update of the model parameters avoiding the ensemble re-run at each time step as in the widely used EnKF (Ensemble Kalman Filter), and (ii) a unique solution over time for the estimated parameters.

The focus is on the investigation of the applicability of the proposed methodology in real case applications where large systems are simulated with approximately one million of degree of freedom. Obviously the computational cost of each simulation highly affects the efficiency of the methodology based on Monte Carlo simulations. In fact, these methods converge to the solution with an increasing number of model realizations which requires a tradeoff between the ES convergence and a reasonable computational cost for the whole procedure.

Preliminary numerical tests on synthetic cases are employed to evaluate the effectiveness of the ES for the parameter estimation. The tests are built such that the reservoir model is the actual real model and the measurements used in the assimilation

are synthetic as derived from the geomechanical model. Two reservoir activities are modeled using a transversely isotropic model to forecast the behavior of a UGS (Underground Gas Storage) field and an isotropic geomechanical model to simulate the deformation caused by gas extraction from a hydrocarbon reservoir.

Different parameters are estimated in the two cases. The ratio of the horizontal-to-vertical Young's modulus, the Poisson ratio, the ratio of the horizontal-to-vertical shear modulus, and the ratio of the c_M in I cycle-to-II cycle, are estimated using a transversely isotropic constitutive law. By distinction, in the second case, the parameter c_M is estimated employing an isotropic law because of lack of horizontal displacements data. It is observed that the assimilation of the land surface motion is important to characterize the subsurface parameters and reduce their prior uncertainty. Although the model is slightly non linear and the ES hypothesis of Gaussian PDFs is not fulfilled, satisfactory ES performances are obtained in terms of spread reduction of the posterior PDFs after assimilation of data with the mean values closer to the true reference simulation. However, it is also shown that when the true reference is too far from the mean of the forecast solution the convergence of the algorithm is more difficult to achieve. This suggests that proper prior PDFs for the parameters must be chosen considering the fact that enlarging the prior PDFs may help include the observations within the forecast ensemble. However, this may lead to run a too large number of Monte Carlo realizations to adequately sample the posterior PDFs.

Homogeneous parameters are estimated in the UGS field while, in the producing reservoir, the case of a heterogeneous distribution of the parameter c_M is also investigated. The motivation for this assumption is based on the geological structure of the studied reservoir, with the gas field highly compartmentalized by a complex distribution of sealing faults and thrusts. As expected the highest reduction of the prior parameters uncertainty is obtained in the areas where the forcing term, i.e., the pressure variation, is maximum. Moreover, the ES provide a satisfactory estimate of the c_M field used as reference to test the updating performance. Comparing the posterior simulations using the c_M from homogenous and heterogeneous fields, the misfit between measurements and model prediction is much lower when less constraints are prescribed on the model, i.e., in the heterogeneous case. In fact, the higher variation

of the parameters in the model domain allows the model to better adapt to the land surface movements.

Of more interest are the applications where real measurements of land displacements are assimilated into the ES scheme. In this case, the parameter estimation is not straightforward. In particular, in the UGS fields vertical and horizontal displacements from PSI are used. The updated parameters are dependent to some extent on different data sets and this indicates the presence of a possible bias in the measurements and/or in the model forecast. Indeed, the consistency between the displacement records and the physical process represented by the forward model is a key point for the parameter estimation. An appropriate selection of observations with the smallest data noise yields the most satisfactory match.

In the offshore gas reservoir, land subsidence data collected from a time-lapse bathymetric map highly impact on the c_M estimation. Seven independent variables for c_M are chosen to describe the possible heterogeneous behavior of the reservoir. It is found that the prior uncertainty reduction of the parameter in each block depends on the magnitude of the available data. Thus the higher pressure variation in the block the higher the influence of the parameter on the model outcome and the effectiveness of the ES scheme. The impact of a parameter on the model solution is crucial for estimation purposes. In fact, it is seen that the variation of c_M in a small block does not provide a significant coefficient of variation in terms of vertical displacements and the reduction of the prior ensemble spread is smaller than in the other reservoir blocks.

The ongoing work on the parameter updating via gPCE (generalized Polynomial Chaos Expansion) based ES is quite promising to reduce the overall computational cost of the ensemble based approaches such as the one used in this thesis. Satisfactory results are obtained with c_M described by one random variable (i.e., homogeneous case) while further numerical tests are needed to better understand the capability of the gPCE combined with the KLE (Karhunen Loève Expansion) to address the heterogeneity of the compressibility and improve the match of the observed land subsidence.

Bibliography

Aanonsen, S., G. Nævdal, D. Oliver, A. C. Reynolds, & B. Vallès (2009), The Ensemble Kalman Filter in Reservoir Engineering – A Review, *SPE Journal*, *14*(3), 393–412.

AAPG, American Association of Petroleum Geologists (2008), *AAPG-ER Newsletter*,*3*, http://www.aapg.org/europe/newsletters/index.cfm

Bailey, R.T. & D. Baù (2010), Ensemble Smoother assimilation of hydraulic head and return flow data to estimate hydraulic conductivity distribution *Water Resour. Res.*, *46*, W12543, doi:10.1029/2010WR009147.

Baù D., M. Ferronato, G. Gambolati & P. Teatini (2002), Basin scale compressibility of the Northen Adriatic by the radioactive marker technique *Geotechnique*, *52*, 605–616.

Baù, D., M. Ferronato, G. Gambolati, P. Teatini, & A. Alzraiee (2014), Ensemble Smoothing of land subsidence measurements for reservoir geomechanical characterization *International Journal for Numerical and Analytical Methods in Geomechanics*, *39*, 207–228, doi 10.1002/nag.2309.

Baù, D., A. Alzraiee, C. Zoccarato, G. Gambolati, M. Ferronato, F. Bottazzi, S. Mantica, & P. Teatini (2015), Testing a data assimilation approach to reduce geomechanical uncertainties in modelling land subsidence, *Environmental Geotechnics*, 15.00005.

Biot, M. A., 1941. General theory of three dimensional consolidation. *Journal of Applied Physics*, *12(2)*, 155–164.

Bishop, A. W. (1959), The principle of effective stress, *Tek. Ukebl.*, *39*, 859–863.

Bouttier, F., & Courtier, P., (1999), Data assimilation concepts and methods, *Meteorological training course lecture series. ECMWF*, Available online at: http://www.ecmwf.int

Burgers, G., van Leeuwen, P. J., & Evensen, G., 1998. Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review*, em 126, 1719–1724.

Cassiani, G., & Zoccatelli, C., 2000. Subsidence risk in Venice and nearby areas, Italy, owing to offshore gas fields: A stochastic analysis. *Environmental & Engineering Geoscience*, *6*, 115–128.

Cedigaz (The International Association for Natural Gas), Underground Gas Storage in the World, 2013 (Fifth Edition), 346 pp.

Chen, Y. & Dean S. Oliver, (2012), Ensemble Randomized Maximum Likelihood Method as an Iterative Ensemble Smoother *Math. Geosci*, *44*, 1–26, doi:10.1007/s11004-011-9376-z

Cosme, E., J. Verron, P. Brasseur, J. Blum, & D. Auroux (2012), Smoothing Problems in a Bayesian Framework and Their Linear Gaussian Solutions, *Mon. Wea. Rev.*, *140*, 683–695, doi:10.1175/MWR-D-10-05025.1.

Daley, R., (1991), Atmospheric data analysis, *Cambrige Atmospheric and Space Science Series*, Cambrige University Press, ISBN 0-521-38215-7, 457 pp.

Doornhof, D., T. G. Kristiansen, N. B. Nagel, P. D. Pattilo, & C. Sayers (2006), Compaction and Subsidence, *Oilfield Review*, *8*(3), 50–68.

Emerick, A. A., & Reynolds, A. C., (2013), Ensemble smoother with multiple data assimilation, *Computers & Geosciences*, *55*, 3–15.

Encyclopaedia of Hydrocarbons: Exploration, production and transport (2005), *ENI - Istituto della Enciclopedia Italiana Fondata da Giovanni Treccani*, *Vol. 1*, Rome.

Evensen, G. (1994), Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, *99(C5)*, 10143–10162.

Evensen, G., & van Leeuwen, P. J., 2000. An Ensemble Kalman Smoother for Nonlinear Dynamics. *Monthly Weather Review*, *128*, 1852–1867.

Evensen, G., 2003. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*, 343–367.

Evensen, G. (2009), Data Assimilation. The Ensemble Kalman Filter, Second Edition. *Springer–Verlag Berlin Heidelberg*, 2009.

Evensen, G., & P. J. van Leeuwen (2000), An Ensemble Kalman Smoother for Nonlinear Dynamics, *Mon. Wea. Rev.*, *128*, 1852–1867.

Ferronato, M., Gambolati, G., Teatini, P., & Baù, D., (2003) Interpretation of radioactive marker measurements to evaluate compaction in the Northern Adriatic gas fields. *SPE Journal of Reservoir Evaluation and Engineering*, *6*, 401–411.

Ferronato, M., Gambolati, G., Teatini, P., & Baù, D. (2004) Radioactive Marker Measurements in Heterogeneous Reservoirs: Numerical Study, *Int. J. Geomech.*, *4(2)*, 79–92.

M. Ferronato, N. Castelletto, G. Gambolati, C. Janna & P. Teatini (2013), II cycle compressibility from satellite measurements, *Geotechnique*, *63*, 479–486.

Fokker, P. A., B. B. T. Wassing, F. J. van Leijen, R. F. Hanssen, & D. A. Nieuwland (2013), Data Assimilation of PS-InSAR Movement Measurements Applied to the Bergermeer Gas Field, *International Workshop on Geomechanics and Energy*, doi:10.3997/2214-4609.20131971.

Fredrich, J. T., Arguello, L. G., Deitrick, G. L., & de Rouffignac, E. P., 2000. Geomechanical modeling of reservoir compaction, surface subsidence, and casing damage at the Belridge diatomite field. *SPE Reserv. Eval. Eng.*, *3*, 348–359.

Gambolati, G., P. Gatto, & G. Ricceri (1986), Land subsidence due to gas-oil removal in layered anisotropic soils by finite elements, in *Land Subsidence - Proceedings of the Third International Symposium on Land Subsidence*, edited by A. I. Johnson et al., IAHS Publ, *151*, 29–41.

Gambolati, G., P. Teatini, D.Baù, & M. Ferronato (2000), The importance of poroelasting coupling in dynamically active aquifers of the Po river basin, Italy, *Water Resour. Res.*, *36(9)*, 2443–2459.

Gambolati, G., Ferronato, M., Teatini P., Deidda, R., & Lecca, G., 2001. Finite element analysis of land subsidence above depleted reservoirs with pore pressure gradient and total stress formulations. *International Journal for Numerical and Analytical Methods in Geomechanics*, *25*, 307–327.

Gambolati, G., Teatini, P., 2015. Geomechanics of subsurface water withdrawal and injection. *Water Resources Research*, *51*, 3922–3955.

Geertsma, J. (1966), Problems of rock mechanics in petroleum production engineering, *In Proc. 1st Cong. Int. Soc. of Rock Mechanics*, Lisbon, 585–594.

Geertsma, J. (1973a), A basic theory of subsidence due to reservoir compaction: the homogeneous case, *Verhandelingen Kon. Ned. Geol. Mijnbouwk Gen.*, *28*, 43–62.

Geertsma, J. (1973b), Land subsidence above compacting oil and gas reservoir, *J. Petrol. Technol.*, *25*, 734–744.

Gercek H. (2007), Review – Poisson's ratio values for rocks, *International Journal of Rock Mechanics and Mining Sciences*,*44*, 89–101.

Ghil, M., 1989, Meteorological data assimilation for oceanographers. Part I: description and theoretical framework, *Dyn. of Atmos. Oceans*, *13*, 171–218.

Gu, Y., Oliver, D. S., (2005), History matching of the PUNQ-S3 reservoir model using the Ensemble Kalman Filter, *SPE J*, *10*(2), 51–65.

Hahn, G. J., Shapiro S. S.(1967), Statistical Models in Engineering, *John Wiley & Sons: New York*.

Hatchell, P., Bourne, S., 2005. Rocks under strain: Strain-induced time-lapse time shifts are observed for depleting reservoirs. *The Leading Edge*, *24*(12), 1222–1225.

Hendricks Franssen, H. J., Kinzelbach W. (2008), Real-time groundwater flow modeling with the ensemble Kalman filter: joint estimation of states and parameters and the filter inbreeding problem, *Water Resources Research*, *44*, W09408, doi: 10.1029/2007WR006505.

Herwanger, J.V. Horne, S. A., 2009. Linking reservoir geomechanics and time-lapse seismics: Predicting anisotropic velocity changes and seismic attributes. *Geophysics*, *74*(4), W13–W33.

Hilbert L. B., Gwinn R. L., Moroney T. A., & Deitrick G. L., 2005. Field-scale and wellbore modeling of compaction-induced casing failures. *SPE Journal of Drilling & Completion 1999*, *14*(2), 92–101.

Hueckel, T., Cassiani, G., Prévost, J. H. & Walters, D. A., 2005. Field derived compressibility of deep sediments of Northern Adriatic. *Proceedings of Seventh International Symposium on Land Subsidence*, Shanghai, China, Special Volume, F.B.J. Barends et al. (eds.), 35–50.

Holm, E. V. (2003), Lecture notes on assimilation algorithm, *Educational Material, ECMWF*.

Iglesias, M. A., D. McLaughlin (2012), Data inversion in coupled subsurface flow and geomechanics models, *Inverse Problems*, *28*, 115009 (47pp).

Iglesias, M. A., Law, K. J. H., & A. M. Stuart (2013a), Ensemble Kalman methods for inverse problems, *Inverse Problems*, *29(4)*, 045001.

Iglesias, M. A., Law, K. J. H., & A. M. Stuart (2013b), Evaluation of gaussian approximations for data assimilations in reservoir models (2013), *Comput Geosci*, *17*, 851–885.

Janna, C., N. Castelletto, M. Ferronato, G. Gambolati, & P. Teatini (2012), A geomechanical transversely isotropic model of the Po River basin using PSInSAR derived horizontal displacement, *International Journal of Rock Mechanics and Mining Sciences*, *51*, 105–118.

Jha B., F. Bottazzi, R. Wojcik, M. Coccia, N. Bechor, D. McLaughlin, T. Herring, B. H. Hager, S. Mantica & R. Juanes (2015), Reservoir characterization in an underground gas storage field using joint inversion of flow and geodetic data, *International Journal for Numerical and Analytical Methods in Geomechanics*, *39*, 1619–1638.

Jolley, S. J., Fisher, Q. J. & Ainsworth, R. B., 2010. Reservoir compartmentalization: an introduction. *Geological Society, London, Special Publications*, *347*, 1–8.

Kalman, R. E., 1960. A new approach to linear filtering and prediction problems. *Transaction of ASME - Journal of Basic Engineering*, *82* (Series D), 35–45.

Katterbauer, K., I. Hoteit, & S. Sun (2014), Data Assimilation of InSAR Measurements for Large Scale Reservoirs, *SPE Saudi Arabia Section Technical Symposium and Exhibition* (SPE–172211–MS), doi:10.2118/172211-MS.

Keppenne, C. L. (2000), Data Assimilation into a Primitive-Equation Model with a Parallel Ensemble Kalman Filter, *Monthly Weather Review*, *128*, 1971–1981.

Ketelaar, G., F. van Leijen, P. Marinkovic, & R. Hanssen (2007), Multi-track PS-InSAR datum connection and reliability assessment, *International Geoscience and Remote Sensing Symposium (IGARSS), Montreux, Switzerland*, 2481–2484.

Klemm, H., I. Quseimi, F. Novali, A. Ferretti, & A. Tamburini (2010), Monitoring horizontal and vertical surface deformation over a hydrocarbon reservoir by PSInSAR, *First break*, *28*.

Kristiansen, T.G. Plischke, B., (2010), History matched full field geomechanics model of the Valhall Field including water weakening and re-pressurization. *SPE*, 131505.

Li, J., D. Xiu (2009), A generalized polynomial chaos expansion based ensemble Kalman filter with high accuracy, *Journal of Computational Physics*, *228*, 5454–5469.

Lorenc, A., 1986, Analysis methods for numerical weather prediction, *Quart. J. Roy. Meteor. Soc.*, *112*, 1177–1194.

Lorentzen, R. J., Nævdal, G., & Lage A. C. V. M., (2003), Tuning of parameters in a two-phase flow model using an Ensemble Kalman Filter, *International Journal of Multiphase Flow*, *29*(8), 1283–1309.

Mezghani, M., Fornel, A., Langlais, V., & Lucet, N., 2004. History matching and quantitative use of 4D seismic data for an improved reservoir characterization. *SPE Annual Technical Conference and Exibition, 26-29 September, Houston, Texas*.

Muntendam-Bos, A. G., P. A. Fokker (2009), Unraveling reservoir compaction parameters through the inversion of surface subsidence observations, *Comput Geosci*, *13*, 43–55, doi:10.1007/s10596-008-9104-z

Nævdal, G., L. M. Johnsen, S. I. Aanonsen, & E. Vefring, (2003), Reservoir monitoring and continuous model updating using the ensemble Kalman filter, *SPE Annual Technical Conference and Exhibition* (SPE 84372).

Oliver, D.S., Y. Chen (2011), Recent progress on reservoir history matching: a review, *Comput Geosci*, *15*, 185–221, doi:10.1007/s10596-010-9194-2.

Oliver, D.S., A. C. Reynolds, & N. Liu (2008), Inverse theory for Petroleum Reservoir Characaterization and History Matching, *Cambrige University Press*.

Peaceman, D. W., (1977), Fundamentals of numerical reservoir simulation, *Dev. Pet. Sci., vol. 6*, Elsevier, Amsterdam.

Pauwels, V. R. N., G. J. M. De Lannoy, H.-J. Hendricks Franssen, & H. Vereecken Simultaneous estimation of model state variables and observations and forecast biases using two-stage hybrid Kalman filter (2013), *Hydrol. Earth Syst. Sci.*, *17*, 3499–3521, doi:10.5194/hess-17-3499-2013.

Sayers, C., den Boer, L., Hooyman, P., & Lawrence, R., 2006. Predicting reservoir compaction and casing deformation in deepwater turbidities using a 3D mechanical earth model. *SPE*, 103926, DOI: 10.2118/103926-MS.

Skjervheim, J.–A., G. Evensen, J. Hove, & J. G. Vabø(2011), An Ensemble Smoother for assisted History Matching, *SPE Reservoir Simulation Symposium*, (SPE 141929), doi:10.2118/141929-MS.

Stancliffe, R. P. W., van der Kooij, M. W. A. (2001), The use of satellite-based radar interferometry to monitor production activity at the Cold Lake heavy oil field, Alberta, Canada, *AAPG Bulletin*, *5*(85), 781–793.

Tamburini, A., M. Bianchi, C. Giannico, & F. Novali (2010), Retrieving surface deformation by PSInSARTM technology: A powerful tool in reservoir monitoring, *International Journal of Greenhouse Gas Control*, *4*, 928–937.

Tarantola, A. (2005), Inverse problem theory and methods for model parameter estimation, *Society for Industrial and Applied Mathematics*, Philadelphia (PA), ISBN: 978-0-89871-572-9.

Teatini, P., N. Castelletto, M. Ferronato, G. Gambolati, C. Janna, E. Cairo, D. Marzorati, D. Colombo, A. Ferretti, A. Bagliani, & F. Bottazzi (2011), Geomechanical response to seasonal gas storage in depleted reservoirs: A case study in the Po River basin, Italy, *Journal of Geophysical Research - Earth Surface*, *116*, F02002, doi:10.1029/2010JF001793.

Tosi, L., P. Teatini, L. Carbognin, & G. Brancolini (2009), Using high resolution data to reveal depth-dependent mechanisms that drive land subsidence: The Venice coast, Italy, *Tectonophysics*, *474*, 271-284.

van Leeuwen, & P. J., G. Evensen (1996), Data Assimilation and Inverse Methods in Terms of a Probabilistic Formulation, *Mon. Wea. Rev.*, *24*, 2898–2913.

van Hasselt, J. P., 1992. Reservoir compaction and surface subsidence resulting from oil and gas production. *Geol. Mijnbouw*, *71*, 107–118.

Vasco, D. W., A. Rucci, A. Ferretti, F. Novali, R. Bissell, P. Ringrose, A. Mathieson, & I. Wright (2010), Satellite-based measurements of surface deformation reveal fluid flow associated with the geological storage of carbon dioxide, *Lawrence Berkeley National Laboratory*, http://escholarship.org/uc/item/2bd7g4bt.

Verruijt A. (2010), Elastic storage aquifers. In R. De Wiest (Ed.), *Flow Through Porous Media*, New York, 331–376. Academic Press.

Wang, H. F. (2000), Theory of Linear Poroelasticity with Applications to Geomechanics and Hydrogeology, *Princeton University Press, Princeton, N. J.*.

Wilschut, F., E. Peters, K. Visser, P. A. Fokker, & P. M. E. van Hooff (2011), Joint History Matching of Well Data and Surface Subsidence Observations Using the Ensemble Kalman Filter: A Field study, *SPE Reservoir Simulation Symposium*, (SPE–141690–PP), doi:10.2118/141690-MS

Xiu, D., 2010. Numerical methods for stochastic computations. A spectral method approach. *Princeton University Press*, Princeton, New Jersey.

Zander, E., K., (2013), Tensor Approximation methods for stochastic problems. *Shaker Verlag*, Aachen, 2013.

Zienkiewicz, O., C., R. L. Taylor (2000), The Finite Element Method, 5th ed., *Oxford, UK: Butterworth-Heinemann*

Zoccarato, C., Baù D., Bottazzi, F., Ferronato, M., Gambolati, G., Mantica, S., & Teatini, P., 2015. Estimate of spatially variable reservoir compressibility by assimilation of ground surface displacement data. *Proc. IAHS*, *372*, 351–356, doi:10.5194/piahs-372-351-2015.

Zoccarato, C., Baù D., Ferronato, M., Gambolati, G., Alzraiee, A., & Teatini, P., 2016. Data assimilation of surface displacements to improve geomechanical parameters of gas storage reservoirs. Submitted to *JGR: Solid Earth*.