



A novel methodological approach for land subsidence prediction through data assimilation techniques

Laura Gazzola¹ · Massimiliano Ferronato¹ · Matteo Frigo¹ · Carlo Janna¹ · Pietro Teatini¹ · Claudia Zoccarato¹ · Massimo Antonelli² · Anna Corradi² · Maria Carolina Dacome² · Stefano Mantica²

Received: 6 April 2020 / Accepted: 10 May 2021
© The Author(s) 2021, corrected publication 2021

Abstract

Anthropogenic land subsidence can be evaluated and predicted by numerical models, which are often built over deterministic analyses. However, uncertainties and approximations are present, as in any other modeling activity of real-world phenomena. This study aims at combining data assimilation techniques with a physically-based numerical model of anthropogenic land subsidence in a novel and comprehensive workflow, to overcome the main limitations concerning the way traditional deterministic analyses use the available measurements. The proposed methodology allows to reduce uncertainties affecting the model, identify the most appropriate rock constitutive behavior and characterize the most significant governing geomechanical parameters. The proposed methodological approach has been applied in a synthetic test case representative of the Upper Adriatic basin, Italy. The integration of data assimilation techniques into geomechanical modeling appears to be a useful and effective tool for a more reliable study of anthropogenic land subsidence.

Keywords Numerical model · Data assimilation · Subsidence prediction · Uncertainties reduction

1 Introduction

It is well known that fluid withdrawal from the subsurface involves the compaction of the reservoir rock and the consequent propagation of the deformation from underground up to the land surface. The corresponding loss of land elevation, i.e., land subsidence, may cause severe environmental problems, especially in coastland and low-land areas, such as the increase of the risk of flooding and the salinization of fresh shallow groundwater, and generally affect several vulnerable human activities from both a social and economical viewpoint, e.g., [22, 48, 49]. For this reason, a reliable prediction of land motion is of considerable importance, and numerical models can play a key role to estimate the amount of land subsidence caused by human activities and simulate the possible related consequences.

As a matter of fact, uncertainties and approximations are implicit in any mathematical and numerical modeling activity of real-world processes, and this can cause lack of solution quality and reliability. In land subsidence simulation due to underground fluid withdrawal, such uncertainties can be related to a lot of different issues, such as the geometry and the lithology of the porous medium, the actual pore pressure variation and distribution in space and time, the mechanical behaviour of the deforming rock, the representativeness of the governing physical processes by the selected mathematical model, and so on. Traditionally, in practical numerical computations these uncertainties are accounted for by defining a number of different scenarios, with the resulting deterministic modeling outcomes combined to define confidence intervals. In recent years, however, there is an increasing availability of monitoring data, such as satellite data on the surface motion with the aid of GPS and InSAR technology [8, 12, 42, 46], or in-situ measurements of the formation compaction, e.g., by deep extensometers or the radioactive marker technique [21, 26, 32]. This has significantly enlarged the amount of available pieces of information, which can be used to monitor the subsidence occurrence in both space and time. Including all the available information in the simulation process in order

✉ Laura Gazzola
laura.gazzola.l@phd.unipd.it

¹ Department of Civil, Environmental and Architectural Engineering, University of Padova, Padova, Italy

² Eni S.p.A., Milan, Italy

to improve the prediction reliability represents a natural evolution of the land subsidence modeling activity. In this context, the use of data assimilation (DA) techniques can be a robust strategy to quantify and reduce the uncertainties in the subsidence prediction by incorporating measurements and model outcomes in a proper assimilation scheme.

Originally developed especially for atmospheric and oceanographic simulations, DA strategies have been recently used also in underground applications, such as hydrology, geomechanics and petroleum engineering [4, 7, 16, 24, 29, 43, 54, 56]. Among the most popular DA techniques used in geomechanical applications, here we mention three approaches that have proved to be particularly suitable for land subsidence analyses. The focus is specifically on Red Flag (RF), Ensemble Smoother (ES), and its variation in form of the Multiple Data Assimilation (MDA-ES). However, for a detailed review of DA techniques in geosciences the reader can for instance refer to [1, 9, 44], just to cite a few recent relevant works. The RF technique inexpensively provides the probability of an event following a purely Bayesian approach, with no update of the prior model [34]. This is a fast method for a preliminary analysis of an ensemble of model realizations, which can provide the most probable scenario given a number of uncertain parameters. The ES can be used to update both state variables and model parameters in a 4-D acquisition (space and time), and evaluate the related uncertainties. The ES is an ensemble-based approach that uses an approximate sampling of the posterior distribution by a finite ensemble of Monte Carlo realizations [19, 51]. For the straightforward implementation and adaptability to different fields of application, this ensemble-based technique and its recent variations are increasingly employed. The ES has been initially used in petroleum engineering, especially for history matching purposes [10, 11, 39], with recent applications also for subsidence predictions [4, 24, 56]. In [23] and [56], the ES technique has been used to calibrate the geomechanical parameters in a one-way coupled model. In [2] and [4] the efficiency of ES has been demonstrated for the calibration of geomechanical parameters such as Poisson coefficient, ratio between horizontal and vertical elastic modulus, ratio between horizontal and vertical shear modulus and ratio between I and II cycle compressibility. In [55, 56] and [58] ES has been used in real-world applications to constrain a geomechanical model through land surface and/or compaction measurements. Finally, the MDA-ES algorithm should help improve the ES outcome, especially when the relation between model solution and parameters is non-linear. In [16] the available observations have been assimilated multiple times with an inflated measurement covariance matrix. This technique has been combined

with ES approach (MDA-ES) in [17] and used in three synthetic history matching problems. A MDA-ES algorithm was also applied by [24] to estimate the reservoir compaction coefficient and the subsurface basement elastic modulus, thus improving the ES effectiveness in non-linear problems.

The objective of this paper is to develop a robust step-by-step workflow for the prediction of land subsidence above producing hydrocarbon fields, integrating well-established DA techniques with an appropriate uncertainty evaluation. In order to overcome the main limitations of traditional deterministic approaches, we aim at looking at the geomechanical modelling of land subsidence from a more comprehensive perspective. To this purpose, the proposed methodological approach automatically combines the available measurements with the mathematical and numerical description of the main undergoing processes. The workflow is based on successive steps of DA application, such as RF, ES and MDA-ES, requiring an increasing level of complexity and computational effort, but providing more detailed and refined outcomes. Prediction uncertainties are progressively reduced as new measurements become available, with the model gradually trained and updated as the knowledge of the actual phenomenon improves. The final result is a substantial increase of the prediction reliability.

In order to investigate the capabilities and possible drawbacks of the proposed approach, an extensive numerical experimentation has been carried out by considering a synthetic test case, which is, however, representative of an offshore real-world hydrocarbon reservoir. The effectiveness of the selected DA techniques has been tested in different possible scenarios of subsidence modelling, both changing the combinations of uncertain geomechanical parameters and constraining the model outcomes with a different number and type of synthetic measurements. The numerical results allow to characterize each step of the workflow and demonstrate that the integration of DA approaches in subsidence problems can be effective to both avoid an overconfidence in deterministic model applications and reduce the prediction uncertainties. In addition, they can suggest the possibility of easily extending and generalizing the application of the proposed methodology to other settings and modelling targets.

The paper is organized as follows. First, the forward numerical model used to analyse land subsidence due to a compacting reservoir is briefly described, and the investigated DA methods are introduced. Then, numerical results for a synthetic test case with two different constitutive laws are presented and discussed. On the basis of the numerical results, a tentative novel workflow for land subsidence modelling is introduced, thus closing the paper.

2 Methodological approach

The methodological approach is based on the idea of building a model for land subsidence management that: (i) is progressively trained by the assimilation of new data, (ii) reduces the prediction uncertainty as the production program proceeds, and (iii) is able to react almost in real-time to modifications in the reservoir development.

The candidate DA approaches require the generation through the forward model of an ensemble of realizations allowing to propagate the uncertainties from the input parameters to the problem state variables, e.g., land surface vertical displacements. The augmented ensemble, including both state variables and input parameters, is used with the DA techniques, in order to reduce the uncertainties and make the model results more reliable.

In the next sections, the geomechanical model is first introduced in general, and then each DA technique included in the workflow is described. The model presented herein represents a consistent option to investigate the efficiency and feasibility of the workflow for land subsidence analysis, but neither the only, nor necessarily the best, choice among the possible alternatives. The proposed approach can be generalized to other contexts as well, using different models, sources of uncertainties and types of available measurements.

2.1 Numerical forward model

Land subsidence due to fluid withdrawal is classically modeled by means of Biot’s theory of consolidation [5], where variations of source of strength reflect on the pore fluid movement and consequently on stress and displacement field of solid skeleton. To study the anthropogenic land subsidence due to compaction of a deep hydrocarbon reservoir, a flow and a geomechanical model are necessary.

The flow model controls the dynamics of the hydrocarbon reservoir and the connected aquifer system. It is based on Darcy’s law coupled with the continuity equation. The generalized multiphase flow model is:

$$\nabla \cdot \left[\frac{\kappa \rho_\alpha}{\mu_\alpha} \nabla p_\alpha(\mathbf{x}, t) \right] = \frac{\partial}{\partial t} (\phi S_\alpha \rho_\alpha) \tag{1}$$

where $\nabla \cdot$ and ∇ are the divergence and gradient operator, respectively, κ is the permeability of the porous medium, ρ_α the density of the fluid phase α and μ_α its viscosity, p_α is the fluid pore pressure that is a function of the position vector \mathbf{x} and the time t , ϕ is the porosity and S_α the saturation index for the fluid phase α .

Land settlements are computed by a geomechanical model that is based on the solution of the equilibrium equation of a

3D porous volume with Terzaghi’s effective stress principle [6, 50]:

$$\nabla \cdot (\underline{\underline{\sigma}}(\mathbf{x}, t) - b \tilde{p}(\mathbf{x}, t) \underline{\underline{1}}) = 0 \tag{2}$$

where $\underline{\underline{\sigma}}(\mathbf{x}, t)$ and $\underline{\underline{1}}$ are the effective stress and the identity tensor, respectively, b is the Biot coefficient and $\tilde{p}(\mathbf{x}, t)$ is the equivalent pore pressure:

$$\tilde{p} = \sum_\alpha S_\alpha p_\alpha \tag{3}$$

The system of Eq. 2, supplemented with appropriate boundary conditions, is numerically solved by means of a finite element formulation. The effective stress tensor $\underline{\underline{\sigma}}(\mathbf{x}, t)$ is generally linked to the displacement field by an appropriate constitutive relationship:

$$\underline{\underline{\sigma}}(\mathbf{x}, t) = \hat{\mathbf{D}} \left[\underline{\underline{\varepsilon}}(\mathbf{x}, t) \right] : \underline{\underline{\varepsilon}}(\mathbf{x}, t) \tag{4}$$

where $\hat{\mathbf{D}}$ is the rank-four constitutive tensor, $\underline{\underline{\varepsilon}}(\mathbf{x}, t)$ is the strain tensor and ‘:’ indicates the inner product. The strain tensor depends on the displacement field $\mathbf{u}(\mathbf{x}, t)$ according to the small strain hypothesis:

$$\underline{\underline{\varepsilon}}(\mathbf{x}, t) = \nabla^s \mathbf{u}(\mathbf{x}, t) \tag{5}$$

where ∇^s is the symmetric gradient operator. Introducing Eq. 3 to Eq. 5 into Eq. 2 and prescribing the appropriate boundary conditions leads to the definition of the geomechanical model in the main unknown $\mathbf{u}(\mathbf{x}, t)$. The selection of the generally non-linear operator $\hat{\mathbf{D}}$ is the key for the model outcome. In this application, we will consider two different non-linear constitutive models, which can allow for a reliable description of the mechanical behavior of geological formations: (i) the elasto-plastic modified Cam-Clay (MCC) and (ii) the visco-elasto-plastic (VEP) Vermeer-Neher law. For a thorough discussion on these constitutive models, the reader is referred to the relevant literature, e.g. [13, 52]. The governing material parameters that appear in the definition of $\hat{\mathbf{D}}$ for each law will be introduced in Section 3.1.

A one-way coupled approach is considered in our simulations, as is fully warranted by the time and space scale of interest. In particular, such an approach is usually employed in the context of real-world oil and gas reservoir engineering applications, where the mechanics-to-flow coupling is usually weak, e.g. [28, 37, 53]. As a consequence, the distribution of the pore pressure variation in the reservoir and the connected aquifer is firstly predicted by the flow model (1) and then the results are introduced in the geomechanical model. In this work, an in-house developed code is used. Details on the numerical formulation and implementation

are given in [27, 31, 41]. This code has been widely validated and used in several real-world applications, e.g., [22, 31, 46, 47, 55].

2.2 Data assimilation techniques

While the outcome of numerical models is obtained in a deterministic way, every computational tool describing real-world phenomena is based on a number of assumptions and approximations. Integration of DA techniques allows to consider and reduce the uncertainties implicitly introduced into numerical models constraining them with a set of measurements of the simulated process. In this work, we focus on the geomechanical model and assume the pore pressure $\tilde{p}(\mathbf{x}, t)$ to be known as the result of a history-matching process. Though the pressure field has its own uncertainty, in-situ well measurements coupled with the history-matching process typically reduce the size of the confidence interval, such that we can consider \tilde{p} as a deterministic value with respect to other uncertain parameters.

DA techniques are included into the structural model to account for and reduce uncertainties related both to the definition of the rock constitutive laws and the geomechanical parameters. After a preliminary analysis of the ensemble, denoted as model diagnostic, three main approaches are investigated and combined: RF, ES and MDA-ES. Each method is briefly reviewed in the next paragraphs focusing on the problem that follows. Generally, the vector of the true state variables in both space and time $\psi^t \in \mathbb{R}^{n_\psi}$ can be defined as a function of the vector of the true model parameters $\theta^t \in \mathbb{R}^{n_\theta}$ through the forward operator \mathcal{G} :

$$\psi^t = \mathcal{G}(\theta^t)$$

with n_ψ and n_θ the number of states and parameters, respectively. Focusing on the land subsidence problem, the operator \mathcal{G} is the geomechanical simulator that relates the true state variables, such as land settlements, to the true model parameters, for example the geomechanical parameters that control the rock behavior. The vector ψ^t is related to the vector of noisy empirical measurements $\mathbf{d} \in \mathbb{R}^{n_d}$, with n_d the number of measurements, through the generally non-linear relationship:

$$\mathbf{d} = \mathbf{d}^t + \varepsilon_d = \mathbf{H}[\psi^t] + \varepsilon_d \tag{6}$$

where \mathbf{H} is the measurement operator mapping from ψ^t to the true observable vector $\mathbf{d}^t \in \mathbb{R}^{n_d}$ and $\varepsilon_d \sim \mathcal{N}(0, \mathbf{C}_d) \in \mathbb{R}^{n_d}$ is the measurement error, with $\mathbf{C}_d \in \mathbb{R}^{n_d \times n_d}$ the covariance matrix of measurement error. The matrix \mathbf{C}_d depends on the kind and accuracy of the available measurements.

The vector of the model state $\psi \in \mathbb{R}^{n_\psi}$ is related to ψ^t by the relation:

$$\psi = \psi^t + \varepsilon_\psi \tag{7}$$

where $\varepsilon_\psi \sim \mathcal{N}(0, \mathbf{C}_\psi) \in \mathbb{R}^{n_\psi}$ is the unknown error in the model states with $\mathbf{C}_\psi \in \mathbb{R}^{n_\psi \times n_\psi}$ the covariance matrix of the model state.

In the following, $\varphi \in \mathbb{R}^{n_\varphi \times n_{mc}}$ denotes the matrix of augmented ensemble of states $\psi \in \mathbb{R}^{n_\psi}$ and parameters $\theta \in \mathbb{R}^{n_\theta}$ and $\mathbf{C}_\varphi \in \mathbb{R}^{n_\varphi \times n_\varphi}$ the related covariance matrix, with n_φ the sum of n_ψ and n_θ , and n_{mc} the size of the Monte Carlo ensemble. The superscripts *prior* and *update* designate the forecast and update ensembles, respectively.

2.2.1 Model diagnostic

A model diagnostic procedure can be initially carried out to validate the ensemble of model realizations with respect to the synthetic measurements. The main objectives of the validation are to verify if the assumptions are violated, if the uncertainties are underestimated and if too much confidence is given to the ‘a priori’ model [45]. A way to validate the forecast ensemble is the χ^2 -test, as proposed in [24]. This approach is based on the mismatch between model results and observations, that is [36]:

$$\mathcal{J}(\theta) = \frac{1}{2} (\theta - \theta^{prior})^T \mathbf{C}_\theta^{-1} (\theta - \theta^{prior}) + \frac{1}{2} (\psi(\theta) - \mathbf{d})^T \mathbf{C}_d^{-1} (\psi(\theta) - \mathbf{d}) \tag{8}$$

For linear problems, the minimum of $\mathcal{J}(\theta)$ allows for a χ^2 distribution with degrees of freedom equal to the number of measurements n_d [45]. Consequently, it is well recognized that an ensemble ensures better results in DA applications when

$$\chi^2 = \frac{\mathcal{J}(\theta)}{n_d} \simeq 1 \tag{9}$$

This could be used as a general guideline to define a range for the acceptable size of $\mathcal{J}(\theta)$ in non-linear inverse problems [11]. Generally, the data mismatch part of the cost function in Eq. 8 dominates the magnitude of the total function $\mathcal{J}(\theta)$, so the first contribution of Eq. 8 can be neglected [11].

2.2.2 Red Flag

RF, as introduced in [34], is a statistical technique that computes the probability of an event by combining prior information with the likelihood of the measurements. In this work, RF approach is used as a preliminary method to define the probability of every Monte Carlo realization in the prior ensemble. This allows to evaluate ‘a priori’ the model uncertainties without solving the inverse problem, thus

providing a first estimate of the combination of parameters which is most likely to occur. The Bayesian probability $P(\psi_k|\mathbf{d})$ of a particular realization k of the ensemble is:

$$P(\psi_k|\mathbf{d}) = \frac{P(\psi_k) P(\mathbf{d}|\psi_k)}{\sum_{j=1}^{n_{mc}} P(\psi_j) P(\mathbf{d}|\psi_j)} \tag{10}$$

where $P(\psi_k)$ is the prior probability of the realization, $P(\mathbf{d}|\psi_k)$ is the associated likelihood of the measurements, and the denominator is a normalizing factor [34]. To define the likelihood, the measurements must be compared with the model outcome calculated at each time and space location for the realization k . Considering a Gaussian distribution for the likelihood and introducing a cost function I_k , the likelihood reads [34]:

$$P(\mathbf{d}|\psi_k) = e^{-I_k} \tag{11}$$

$$I_k = \frac{1}{2} \mathbf{q}^T \mathbf{C}_d^{-1} \mathbf{q} \tag{12}$$

where \mathbf{q} is the vector of the differences between measurements and model results in both space and time.

2.2.3 Ensemble smoother

ES is a non-sequential DA algorithm originally proposed by [51]. It is a variance-minimizing estimator that combines prior information, measurements, and the solution of the forward model to update model states and parameters. With ES, the solution of the inverse problem can be written in a matrix form as:

$$\varphi^{update} = \varphi^{prior} + \mathbf{K} \left(\mathbf{D} - \mathbf{H} \left[\varphi^{prior} \right] \right) \tag{13}$$

where $\mathbf{D} \in \mathbb{R}^{n_d \times n_{mc}}$ is the matrix of measurements and $\mathbf{K} \in \mathbb{R}^{n_\varphi \times n_d}$ the Kalman gain, calculated as:

$$\mathbf{K} = \left(\mathbf{H} [\mathbf{C}_f] \right)^T \left(\mathbf{H} \left[\left(\mathbf{H} [\mathbf{C}_f] \right)^T + \mathbf{C}_d \right] \right)^{-1} \tag{14}$$

With a slight notation abuse, here \mathbf{H} denotes the same mapping operator from model to the observational space as defined in Eq. 6, but extended to the vector of state variables augmented by the parameters.

The analyses that follow are carried out with the ES implementation according to the approach introduced in [18]. Results are optimal when the probability distribution of uncertain parameters is Gaussian [51]. The quality of the outcome of the ES application can be evaluated by comparing the prior and update ensemble through two performance indices, the average absolute error (AE) and the average ensemble spread (AES) [29]:

$$AE = \frac{1}{n_{mc} n_d} \sum_{k=1}^{n_{mc}} \sum_{i=1}^{n_d} |\varphi_{i,k} - \varphi_i^{true}| \tag{15}$$

$$AES = \frac{1}{n_{mc} n_\psi} \sum_{k=1}^{n_{mc}} \sum_{i=1}^{n_\psi} |\varphi_{i,k} - \varphi_i^{mean}| \tag{16}$$

where $\varphi_{i,k}$ is either the prior or the posterior value of the k -th realization for the i -th observation or state, φ_i^{true} is the true reference value and φ_i^{mean} is the ensemble mean. These metrics were already used in other works as well, e.g., [56, 59]. The AE index is a measure of the algorithm capability to approach the truth, as it compares the model outcome with the true reference value. This index can be computed only in a synthetic case where the true reference is known. The AES index accounts for the deviation of the model results from the ensemble mean. Hence, it provides an indication of the spread of the distribution, i.e., it is a measure of the confidence in the predicted value. Generally, results of the assimilation are satisfactory when AE and AES of the update ensemble decrease with respect to the corresponding indices of the forecast ensemble. For this reason, the relative variation J of these indices from forecast to update is also computed:

$$J = \frac{\zeta^{prior} - \zeta^{update}}{\zeta^{prior}} \tag{17}$$

where ζ is either the AE or AES index.

2.2.4 Multiple data assimilation

MDA technique was originally introduced by [16] to improve the results of history-matching problems and was later combined with ES approach [17]. Basically, in MDA-ES the ES analysis is repeated for a definite number of ‘iterations’ to improve the results of a single assimilation. The covariance matrix of the measurement error \mathbf{C}_d is multiplied by an inflation coefficient $\alpha_k \geq 1$ for every ‘iteration’ k to avoid an over-confidence in the available measurements. The choice of α_k is such that:

$$\sum_k \frac{1}{\alpha_k} = 1 \tag{18}$$

This condition makes MDA-ES equivalent to ES for Gaussian-linear problems [15]. The selection of α_k is still a matter of debate, with several different proposals currently advanced by some researches, e.g. [14, 38].

3 Test case analyses and results

In this section, the test case is introduced and numerical results are presented. Analyses are performed on a synthetic test case, which is, however, based on the structural and mechanical properties representative of an off-shore real-world hydrocarbon reservoir buried in a sedimentary basin, such as the Northern Adriatic, Italy. Among the several

different uncertainties possibly affecting a land subsidence model, we focus on those associated to the constitutive behavior and the governing deep rock mechanical parameters, which most affect the expected ground motion. We consider feasible ranges for each uncertain parameter based on the physical experimental evidence and compute ensembles taken from the parameter distributions for two different constitutive laws. Then, such uncertainties are propagated on the state variables of the problem through the forward geomechanical model. The augmented ensemble, including both state variables and model parameters, is used with the DA techniques. To constrain the model, we use two different types of observations: (i) surface vertical displacements, and (ii) deep compaction measurements at the reservoir depth. In particular, we suppose the former to be recorded by a GPS station, while the latter derive from the radioactive marker technique.

First, geometry and properties of the synthetic model are described, with the outcome obtained for the combination of parameters we assume as the ‘true’ configuration. Then, an extensive numerical experimentation is carried out and discussed.

3.1 Synthetic test case

We consider a synthetic test case representative of an off-shore hydrocarbon reservoir buried in a sedimentary basin. To make the model configuration as much realistic as possible, geometry and properties typical of the Northern Adriatic basin, Italy, have been used. Two different possible constitutive laws, namely a MCC model [13] and a VEP model based on Vermeer-Neher theory [52], are associated to the reservoir and the hydraulically connected aquifer. The first constitutive law describes an elasto-plastic rate-independent model, the latter is rate-dependent and accounts for viscous effects. These two models are both non-linear and representative of the possible behavior of a deep porous volume subject to pressure variations [30, 41]. Since the behavior of underburden and overburden does not significantly affect surface movements [22, 46], for the sake of simplicity, they are characterized by a simple and deterministic linear elastic constitutive law.

The model domain covers an area of 50 km by 50 km and extends down to about 5 km depth (Fig. 1). The reservoir is located in a central position at a depth interval between 1038 m and 1075 m. The domain is discretized into a 3D finite element mesh, which consists of 71,734 nodes and 410,030 tetrahedrons. A part of the mesh showing the aquifer and reservoir is presented in Fig. 1a. We assume a homogeneous Poisson coefficient equal to 0.30 everywhere and a unitary Biot coefficient. The vertical compressibility c_m is assumed to follow the hysteretic law vs the effective vertical stress

σ_z developed by [3, 20] for the Northern Adriatic basin, Italy:

$$c_m^{Icycle} = 1.0044 \cdot 10^{-2} \cdot \sigma_z^{-1.1347} \quad (19)$$

$$c_m^{IIcycle} = 2.9087 \cdot 10^{-4} \cdot \sigma_z^{-0.4315} \quad (20)$$

where c_m and σ_z are in [MPa⁻¹] and [MPa], respectively. Homogeneous null boundary conditions are prescribed for displacements, on the lateral and bottom boundaries, and for stress, on the top surface.

The pore pressure variation due to hydrocarbon extraction is applied in the reservoir and aquifer. The pressure history shown in Fig. 2 is prescribed on reservoir elements to simulate a possible program of hydrocarbon production. In order to take into consideration a scenario as realistic and general as possible, we simulate two production phases (the first three years and from year 5 to year 7) divided by a temporary stop of the extraction. At the end of the production program, a natural pressure recovery is supposed. The pressure variation propagates from the reservoir to the hydraulically connected aquifer and is considered deterministic, as discussed above.

Since we consider a synthetic case, we run the forward model with a set of parameter values, that are considered as the ‘true’ configuration, to define hypothetical measurements. In order to simulate real recordings and avoid a trivial application, the outcome of the geomechanical model at the ‘measurement’ location is properly perturbed to get the assimilation data, that are for example the dots in Fig. 3. Characteristic size and correlations of the synthetic measurement error will be clarified in Section 3.1.1.

The material parameters required by both the MCC and VEP constitutive laws are the modified compression index λ^* and the modified swelling index k^* , while the modified creep index μ^* and the geotechnical initial overconsolidation ratio R are needed only by the VEP model. Specifically, their physical meaning is as follows:

- λ^* is the slope of the normal consolidation profile in the plot of volumetric strain vs axial stress in natural logarithmic scale, i.e. it is a parameter that mainly controls the rock behavior in I-cycle conditions;
- k^* is the slope of the unloading profile in the same plot as λ^* , thus it mainly impacts on II-cycle conditions;
- μ^* represents the slope of the volumetric strain profile vs time in natural logarithmic scale, so its value is a measure of the delay between the pressure variation and the related reservoir and land deformation;
- R is the ratio $p_{c,r,0}/p_{c,0}$, where $p_{c,r,0}$ is a function of the maximum volumetric stress ever experienced by the material before loading and $p_{c,0}$ depends on the volumetric stress state at initial conditions. The

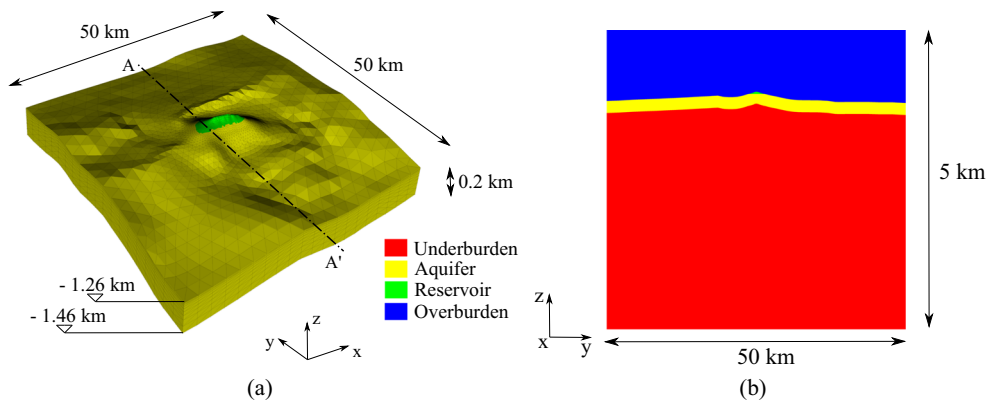


Fig. 1 Views of the discretized domain with the distinction of different layers: 3D view of the partial mesh of the aquifer and reservoir where axis Z is scaled by a factor 30 (a) and vertical section A-A' with coordinates Z multiplied by a factor 10 (b)

parameter R is required to describe implicitly the yield surface evolution in the VEP model.

$$k^* = 0.000700$$

$$\mu^* = 0.00015$$

$$R = 1.50$$

All material parameters listed above are dimensionless. For more details on the model properties and implementations, see [30, 35, 41].

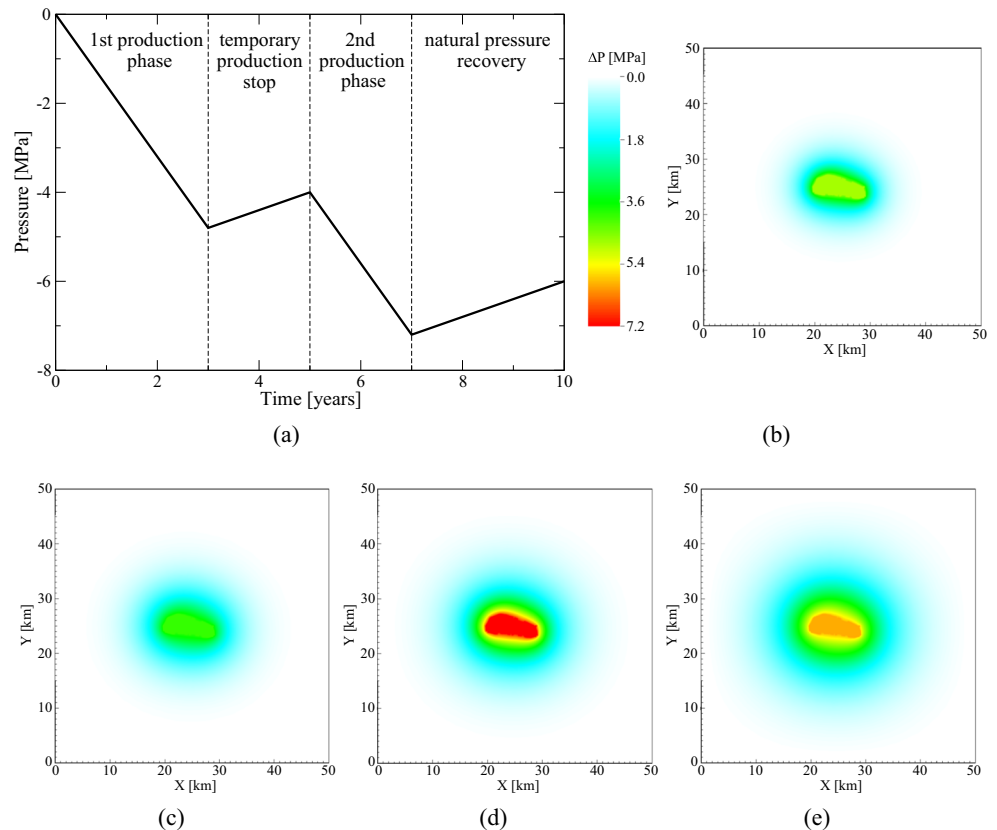
3.1.1 The parameter space

The ‘true’ configuration is selected as follows:

$$\lambda^* = 0.004992$$

Figure 3 shows the vertical displacements in time for the ‘true’ configuration, where the maximum vertical motion is recorded. As it can be easily seen, the difference between the MCC and VEP behavior is not only related to the displacement values, but also to their evolution in time. The VEP model is typically characterized by a delay between the variations of fluid production rates and the resulting land

Fig. 2 Prescribed pore pressure variation in time, $\bar{p}(\mathbf{x}, t)$, on reservoir elements (a) and its propagation [MPa] from reservoir to the hydraulically connected aquifer after three (b), five (c), seven (d) and ten (e) years



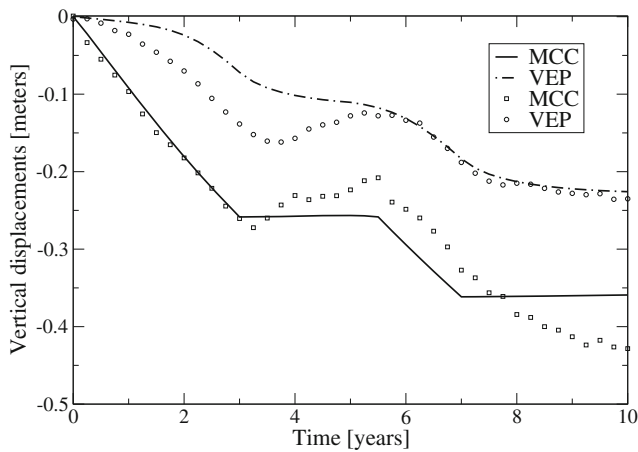


Fig. 3 Maximum vertical displacements in time resulting from the geomechanical model run with the ‘true’ parameter configuration. Squares and circles are an example of such outcomes perturbed to get the synthetic measurements, respectively for the MCC and the VEP constitutive law

settlement [30], which is not accounted for with a traditional MCC approach. These displacements, that are the result of the geomechanical model, have been perturbed with an error $\epsilon \sim \mathcal{N}(0; \mathbf{C}_d)$ to get the synthetic measurements assimilated in the procedure. A detailed description of the construction of \mathbf{C}_d is presented in the lines that follow.

To take into consideration the uncertainties that affect the geomechanical model, we define a variability range for the parameters above. Uncertainty on the first two parameters is related to the confidence interval of c_m in Eqs. 19 and 20, as described in [3], while the ranges of μ^* and R are chosen according to typical values reported in the literature [30, 52]. Thus, the following statistical distributions of the parameters are used in our tests:

$$\ln(\lambda^*) \sim \mathcal{N}(-4.9363; 0.33295)$$

$$\ln(k^*) \sim \mathcal{N}(-6.7271; 0.50975)$$

$$\mu^* \sim \mathcal{U}(0.0001; 0.0002)$$

$$R \sim \mathcal{U}(1.4; 1.8)$$

i.e., λ^* and k^* are Gaussian parameters in a log-scale, while μ^* and R have a uniform distribution. We assume the parameters to be constant over the portion of the domain where a pressure variation takes place. The reason for this assumption is twofold: (i) the variability in space of the geomechanical parameters is usually much lower than that of the hydraulic parameters. In particular, in sedimentary basins such a variability is expected to take place more along the vertical direction than in a horizontal plain. However, in this application the reservoir and aquifer thickness is relatively small to present a significant geomechanical heterogeneity in λ^* , k^* and μ^* . Nonetheless, we recall that the uniaxial compressibility c_m is not constant and is varying with depth as a function of the vertical effective

stress σ_z ; (ii) the available measurements over off-shore reservoirs usually reduce to displacement time series over a small (1 or 2) number of points, and this is not sufficient to effectively characterize the spatial heterogeneity of the uncertain parameters [58].

We suppose that a single GPS station located at the reservoir center collects measurements of the vertical displacement in time, while compaction measurements at the reservoir depth are recorded by radioactive markers. We assimilate displacement measurements every three months and compaction measurements every three years. The covariance matrix of measurement error \mathbf{C}_d has to be defined. Uncertainties related to observations can be computed as the sum of a measurement and an idealisation noise, according to [33]. The first component depends on the accuracy of the tool used for the measurements and the representativeness of the mathematical model for reproducing the observations, and affects only the diagonal entries of \mathbf{C}_d . The idealisation noise affects both the diagonal and the extra-diagonal entries according to the spatial and temporal correlation among measurements. In subsidence modeling, any deformation caused by sources other than the deep compacting layers should be considered as noise. Therefore, all signal components in geodetic observations that are not related to the signal of interest are treated as a noise and included in the idealisation part of \mathbf{C}_d . Following the fractional Brownian motion approach [33], the idealisation noise terms are computed as:

$$(\mathbf{C}_d)_{ii} = \zeta^2 t_i^p \tag{21a}$$

$$(\mathbf{C}_d)_{ij} = \frac{1}{2} \zeta^2 (t_i^p + t_j^p - |t_i - t_j|^p) \tag{21b}$$

where t_i is the time of the measurement i , ζ the standard deviation and p the Hurst index. In real-world studies, parameters ζ and p derive from specific studies on the measurement quality. Here, for the vertical displacements recorded by the GPS station, they have been computed through a fitting process of the variogram of the model outcome run with the ‘true’ parameters (Fig. 3), as described in [33]. For the selected constitutive laws, such parameters read:

$$p_{MCC} = 1.672$$

$$\zeta_{MCC}^2 = 1.112 \cdot 10^{-3} \text{ m}^2/\text{yr}^p$$

$$p_{VEP} = 1.895$$

$$\zeta_{VEP}^2 = 4.158 \cdot 10^{-4} \text{ m}^2/\text{yr}^p$$

The measurement noise is defined from a normal distribution with zero mean and standard deviation equal to 1.5 mm [33]. On the contrary, compaction measurements are considered uncorrelated. To define \mathbf{C}_d , the measurement noise is computed from a normal distribution with zero mean and standard deviation equal to 1.0 cm.

To analyze the role of each parameter, we first investigate the potential of the proposed DA techniques to reduce the uncertainty on each single parameter distribution individually, then we consider some of their combinations. Table 1 lists all the analyzed configurations. A total number of 12 sets is considered, 3 referring to MCC constitutive law and 9 to VEP behavior.

3.2 Model diagnostic

The ensemble of model realizations is initially validated with respect to the synthetic measurements with the model diagnostic procedure. Here, the model state ψ of Eq. 8 includes the synthetic displacement measurements resulting from the geomechanical model at the GPS location. The magnitude of $\mathcal{J}(\theta)$ is calculated for every realization of the displacement ensemble. Each ensemble includes 50 realizations created running the forward model with the probability distributions for the uncertain parameters described in paragraph Section 3.1. Then an average value of χ^2 for the entire ensemble is defined. The resulting values are shown in Table 1.

The χ^2 -test provides only a qualitative analysis of the ensemble and it is not possible to identify a clear trend. Most of the values are close to one. Despite of this, when the number of uncertain parameters grows, the χ^2 value is higher, as in set # 3 for the MCC constitutive law and in the last five sets of Table 1 for the VEP behavior. Higher values of χ^2 denote a probable increasing difficulty in constraining the model with DA techniques. Vice versa, if χ^2 is smaller than one, the variance associated to the ensemble is lower than the variance associated to the measurements, meaning that DA techniques cannot provide improvements in the update of the ensemble and is therefore likely that DA processes are ineffective.

Table 1 Analyzed configuration sets with their associated uncertain parameters and constitutive law, and χ^2 values

Set	Parameters	Law	χ^2
1	λ^*	MCC	1.22
2	k^*	MCC	0.95
3	$\lambda^* k^*$	MCC	1.32
4	λ^*	VEP	1.47
5	k^*	VEP	1.41
6	μ^*	VEP	0.93
7	R	VEP	1.03
8	$\lambda^* k^*$	VEP	4.35
9	$\lambda^* R$	VEP	1.91
10	$\lambda^* k^* \mu^*$	VEP	1.85
11	$\lambda^* k^* R$	VEP	1.79
12	$\lambda^* k^* \mu^* R$	VEP	1.65

Generally, the constitutive law that describes the behavior of the porous medium is ‘a priori’ unknown. However, a guess could be obtained from available information, e.g., laboratory tests, in-situ measurements or data provided by other reservoirs in the same basin. The χ^2 -test can be also used as a preliminary tool to select the most appropriate constitutive law for the geomechanical model, that is the law which appears to be more representative from the available measurements. As an example, we consider cases with the uncertain modified compression index λ^* and modified swelling index k^* , with the related forecast ensembles. Results are shown in Table 2. We compute the χ^2 value when the observations are obtained with the same constitutive law of the ensemble and when they are computed with a constitutive law different from the one used to generate the ensemble. The latter test allows to mimic the case where the choice of the constitutive law in the modeling workflow is not consistent with the actual physical governing process.

As expected, the χ^2 value associated to the ensemble created with the same constitutive law of the measurements is always closer to 1 than the inconsistent one. Consequently, χ^2 -test can provide a useful and cheap preliminary screening of the generated forecast ensembles.

3.3 Red flag

After the ensemble diagnostic, the RF approach is carried out for every configuration shown in Table 1. Every realization of the forecast ensembles is characterized by its own probability using Eq. 10, where $P(\psi_k)$ is derived from the joint probability of the parameter set used to define the realization. Table 3 provides for each case listed in Table 1 the realization characterized by the largest probability

Table 2 Outcome of χ^2 -test used for the choice of the most appropriate constitutive law for the forecast ensemble

Estimate parameters	Measurements constitutive law	Ensemble constitutive law	χ^2
λ^*	MCC	MCC	1.22
λ^*	MCC	VEP	1.38
λ^*	VEP	VEP	1.47
λ^*	VEP	MCC	6.53
k^*	MCC	MCC	0.95
k^*	MCC	VEP	1.49
k^*	VEP	VEP	1.41
k^*	VEP	MCC	5.29
$\lambda^* k^*$	MCC	MCC	1.32
$\lambda^* k^*$	MCC	VEP	1.73
$\lambda^* k^*$	VEP	VEP	4.35
$\lambda^* k^*$	VEP	MCC	9.51

Table 3 RF approach: combination of parameters with maximum probability of occurrence

Set	$\max (P (\psi_k \mathbf{D}))$ [%]	λ^*	k^*	μ^*	R
1	7.11	0.005591			
2	2.84		0.001189		
3	8.27	0.004990	0.000958		
4	20.33	0.005113			
5	5.95		0.000843		
6	2.58			0.000151	
7	3.90				1.5032
8	27.52	0.004990	0.000958		
9	22.65	0.005410			1.4655
10	13.82	0.005520	0.001004	0.000165	
11	41.83	0.005365	0.000930		1.5278
12	25.29	0.005437	0.001142	0.000150	1.5157

of occurrence, along with the corresponding parameter values. The configuration with the largest probability is not always the one closest to the ‘true’ parameter set. The reason is that we compare the state ensemble with the vertical displacements, which are the ‘effect’ through the geomechanical model of the parameter selection and not directly parameters themselves. In other words, we are considering the image of θ_k through the forward model \mathcal{G} and the RF approach computes a ‘score’ for each of such images. In some cases, more than one set provides displacements close to the observation and consequently has a high probability. This could suggest a possible over-parameterization of the problem, i.e. different combinations of the parameter values can be able to reproduce similarly the available observations. This is usually an unfortunate situation, where DA is expected to have a low effectiveness on the parameter characterization.

Despite this, in most cases, there is an important difference between the largest and the smallest probability. Therefore, RF can help to preliminary reduce the uncertainties by neglecting the realizations characterized by the smallest probability of occurrence.

3.4 Ensemble smoother

ES has been employed to find a tentative solution to the inverse model and reduce the initial uncertainties on the model prediction. Both state and parameter ensembles are updated by constraining the model with available measurements. The quality of the outcome is evaluated through the indices AE (Eq. 15) and AES (Eq. 16) and their variation (Eq. 17). Note that in the computation of such indices, the matrix of updated state variables refers to the solution of the forward model by using the updated parameter vector.

3.4.1 Parameter constraints

For the test cases of Table 1, the outcome of the ES approach is provided in Table 4. A positive value of J indicates a restriction around the true configuration of the updated ensemble with respect to the prior. It is clear that the assimilation effectiveness strongly depends on the uncertain parameter set.

The choice of the set of uncertain parameters has a significant impact on the model in relation to the geometry of the domain, the constitutive law, the observed data, the boundary conditions and the imposed external load. For example, the uncertainty associated to the modified compression index λ^* and the modified swelling index k^* has a similar range of variability, as defined in Section 3.1. Nevertheless, the corresponding state ensembles for the MCC constitutive law (set # 1 and # 2) have very different forecast AES. A variation of k^* in the model does not provide significant changes in the resulting state in term of land vertical motion, u_z . As a matter of fact, k^* is a parameter mainly controlling the rock behavior in II-cycle conditions, so it appears to play a secondary role for the selected production program (Fig. 2a). Consequently, in this case ES cannot help constrain k^* because variations related to the state ensemble are lower than the errors associated to measurements. By distinction, variations in λ^* produce a significant change in the forecast state ensemble that provide enough information to constrain the model outcome close to the observations, with a reduction of the ‘a priori’ parameter uncertainties.

ES has been applied considering uncertainty related to each parameter individually and some of their combinations to evaluate ES capability of conditioning the model in different situations with the same available observations.

Table 4 ES approach: AE and AES for the forecast ensembles and the relative variation J of AE and AES from the update ensembles. The larger J, the more effective is the assimilation approach on the parameter or state (vertical displacements, u_z) ensemble, that is indicated in column 2

Set	Ensemble	$AE_{forecast}$	J [%]	$AES_{forecast}$	J [%]
1	λ^*	$4.159 \cdot 10^{-1}$	51	$2.429 \cdot 10^{-1}$	16
	u_z	$1.432 \cdot 10^{-1}$	62	$9.260 \cdot 10^{-2}$	42
2	k^*	$6.488 \cdot 10^{-1}$	-29	$3.706 \cdot 10^{-1}$	-125
	u_z	$6.331 \cdot 10^{-3}$	-37	$3.348 \cdot 10^{-3}$	-150
3	λ^*	$3.996 \cdot 10^{-1}$	49	$2.563 \cdot 10^{-1}$	30
	k^*	$6.488 \cdot 10^{-1}$	-20	$3.706 \cdot 10^{-1}$	-109
4	u_z	$1.454 \cdot 10^{-1}$	64	$1.000 \cdot 10^{-1}$	53
	λ^*	$4.159 \cdot 10^{-1}$	71	$2.429 \cdot 10^{-1}$	51
5	u_z	$3.530 \cdot 10^{-2}$	69	$2.001 \cdot 10^{-2}$	45
	k^*	$6.488 \cdot 10^{-1}$	46	$3.706 \cdot 10^{-1}$	6
6	u_z	$5.440 \cdot 10^{-2}$	65	$3.772 \cdot 10^{-2}$	49
	μ^*	$2.503 \cdot 10^{-5}$	43	$2.485 \cdot 10^{-5}$	42
7	u_z	$2.513 \cdot 10^{-2}$	42	$2.504 \cdot 10^{-2}$	42
	R	$1.193 \cdot 10^{-1}$	47	$9.400 \cdot 10^{-2}$	34
8	u_z	$4.026 \cdot 10^{-2}$	38	$3.092 \cdot 10^{-2}$	19
	λ^*	$3.996 \cdot 10^{-1}$	54	$2.563 \cdot 10^{-1}$	31
9	k^*	$6.488 \cdot 10^{-1}$	53	$3.706 \cdot 10^{-1}$	24
	u_z	$4.650 \cdot 10^{-2}$	44	$4.519 \cdot 10^{-2}$	42
10	λ^*	$3.900 \cdot 10^{-1}$	-87	$2.483 \cdot 10^{-1}$	-190
	R	$1.348 \cdot 10^{-1}$	21	$9.623 \cdot 10^{-2}$	-10
11	u_z	$6.137 \cdot 10^{-2}$	15	$3.444 \cdot 10^{-2}$	-57
	λ^*	$4.333 \cdot 10^{-1}$	38	$2.851 \cdot 10^{-1}$	7
12	k^*	$6.089 \cdot 10^{-1}$	43	$3.501 \cdot 10^{-1}$	5
	μ^*	$1.325 \cdot 10^{-5}$	-165	$8.957 \cdot 10^{-6}$	-292
13	u_z	$4.879 \cdot 10^{-2}$	8	$4.490 \cdot 10^{-2}$	-2
	λ^*	$4.083 \cdot 10^{-1}$	-34	$2.351 \cdot 10^{-1}$	-117
14	k^*	$6.351 \cdot 10^{-1}$	43	$3.380 \cdot 10^{-1}$	-3
	R	$1.276 \cdot 10^{-1}$	6	$9.295 \cdot 10^{-2}$	-28
15	u_z	$5.208 \cdot 10^{-2}$	-23	$4.976 \cdot 10^{-2}$	-32
	λ^*	$3.992 \cdot 10^{-1}$	-87	$2.308 \cdot 10^{-1}$	-206
16	k^*	$6.144 \cdot 10^{-1}$	24	$3.927 \cdot 10^{-1}$	-12
	μ^*	$1.323 \cdot 10^{-5}$	-9	$9.152 \cdot 10^{-6}$	-35
17	R	$5.423 \cdot 10^{-2}$	-145	$5.373 \cdot 10^{-2}$	-145
	u_z	$5.621 \cdot 10^{-2}$	-90	$4.716 \cdot 10^{-2}$	-140

Bold entries are used to emphasize the negative values

The results of Table 4 point out that an over-parameterized problem can prevent from constraining the model to the measurements, as in the last four parameter sets. For the VEP behavior, this can be seen when the set of uncertain parameters includes both λ^* and the ratio R (set # 9). In this case measurements are not enough to constrain the model, while results of set # 4 and # 7, in which these parameters are separately considered into the ES approach, provide a satisfactory reduction of the ensembles spread around the true configuration.

3.4.2 Prediction of land subsidence

Table 4 suggests that the most satisfactory results can be obtained in the configurations with uncertain λ^* , sets # 1 and # 4 for MCC and VEP behavior, respectively. For these cases, we evaluate the ES capabilities also in a predictive sense by assimilating 3, 5 and 7 years of displacement measurements only, with respect to the assimilation of observation data during all the simulation period, i.e. 10 years. This is to investigate the use of ES

during the productive life of the reservoir, introducing into the DA framework the up-to-date set of measurements. The resulting forecast and update ensembles are shown in Fig. 4 together with the variation J with respect to the forecast ensemble of the state variable, i.e. vertical displacements.

As the number of assimilated measurements increases, the effectiveness of the ES algorithm grows. Using a MCC model, the assimilation of few observation data enables the ES to reduce significantly the model uncertainties. In the case of VEP behavior, the assimilation of three years of measurements (Fig. 4b) is not enough for constraining the model, because the spread of the forecast state ensemble ($AES = 0.020$) is already smaller than the one for MCC ($AES = 0.092$) and so the model characterization is intrinsically more difficult. Nevertheless, seven years of observation data appear to be enough to reduce uncertainties and improve the model predictive capabilities. In other words, the continuous assimilation of new land displacement observations in time appears to be effective for 'training' automatically the model, thus increasing its reliability in the prediction capability.

3.4.3 Influence of measurement covariance matrix

The results presented in Table 4 and Fig. 4 are obtained with the covariance matrix of the measurement error C_d described in Section 3.1. Reduction in observation errors, i.e. increase in measurement reliability, usually implies a better performance of the DA algorithms [4]. To evaluate the importance of the measurement error, we reduce the standard deviation ζ in the computation of C_d , i.e. ζ^2 equal to $1.112 \cdot 10^{-5} \text{ m}^2/\text{yr}^p$ for MCC constitutive law and to $4.158 \cdot 10^{-6} \text{ m}^2/\text{yr}^p$ for VEP model. Results provided in Fig. 5 show the outcome of the ES algorithm for the estimate of the modified compression index λ^* by assimilating displacement measurements for the first five years of the simulation process. These results, compared to those obtained with the original C_d shown in Fig. 4c, d, 5a and d, point out a clear improvement in the solution of inverse problem with an increase of J by over 61%.

3.4.4 Assimilation of compaction measurements

Synthetic measurements of the formation compaction, ideally measured for instance by the radioactive marker technique, have been assimilated. We suppose to place a marker borehole near the GPS location, where the reservoir is approximately 30 m-thick. Since the initial space between two adjacent radioactive markers is approximately 10 m, a maximum number of three compaction measurements are available for the assimilation. To evaluate the influence of these additional measurements, we consider four significant configuration sets of uncertain parameters (set # 1, # 4,

6 and # 12). Figure 6 shows the variation J of the AE and AES indices from the forecast to the update ensemble. It is the result of ES application when both displacement and compaction measurements are assimilated. Comparing Fig. 6 with the results in Table 4 allows to understand when adding more accurate measurements with a different nature could be useful.

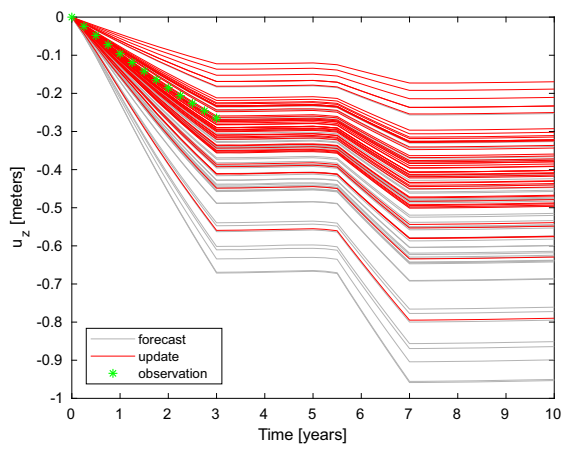
First, set # 1 and # 4 are considered, with uncertain λ^* for the MCC and VEP law, respectively. In these cases, the assimilation of only GPS measurements allows to constrain the model. Adding other observations does not provide a relevant improvement in the ES application. Then, we consider set # 6, when the modified creep index μ^* with the VEP model is uncertain. This parameter has a low range of variability. In this case, the assimilation of both displacement and compaction measurements with the covariance matrix as described in Section 3.1 proves a significant variation in the ES results, with an improvement of J of about 107% with respect to the assimilation of vertical displacements only. Finally, we evaluate the case of an over-parameterized problem, i.e. set # 12 of Table 1. Results obtained with the assimilation of both GPS and marker measurements cannot be considered thoroughly satisfactory, because of the small, or even negative, values of J . Nevertheless, they are better than results in Table 4 for the assimilation of surface vertical displacements only, at least for parameter λ^* , ratio R and displacement u_z .

3.5 Multiple data assimilation

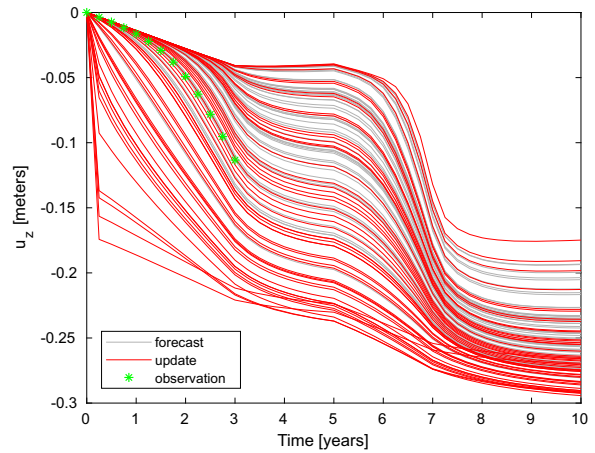
MDA-ES can help improve the effectiveness of the ES algorithm, especially when there is a strongly non-linear relationship between the uncertain parameters and the state variables [16]. To avoid an over confidence in the observation data, the covariance matrix of measurements error is multiplied by an inflation coefficient α_k at every assimilation step k . Different algorithms have been recently developed to define the sequence of α_k values, e.g. in [14]. Here, we consider only the simplest option, i.e. a constant inflation coefficient equal to the total number of successive assimilations.

3.5.1 Influence of measurement covariance matrix

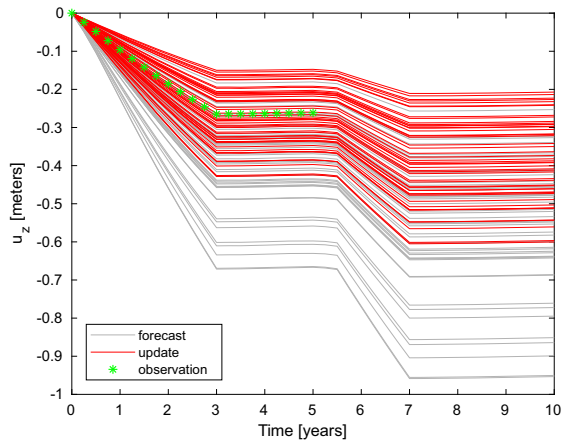
Application of MDA-ES for the test cases of Table 4 improves the parameters estimation with respect to ES only when uncertainty is associated to a single parameter. For the sake of brevity, here these outcomes are not provided. Better results have been achieved considering a reduced covariance of the measurement error. Two groups of simulations are carried out to point out the influence of the measurement error in the MDA-ES approach. The



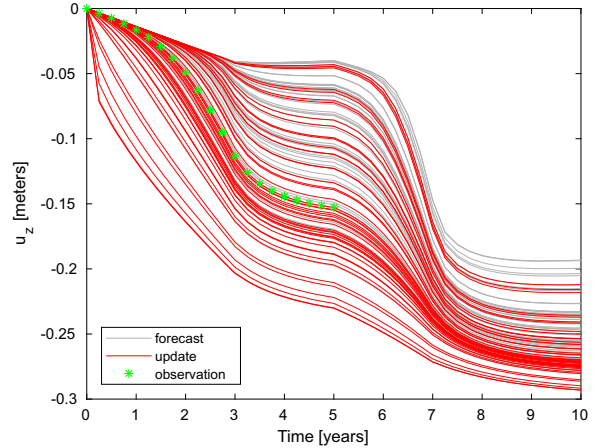
(a) $J(AE)=56\%$ $J(AES)=36\%$



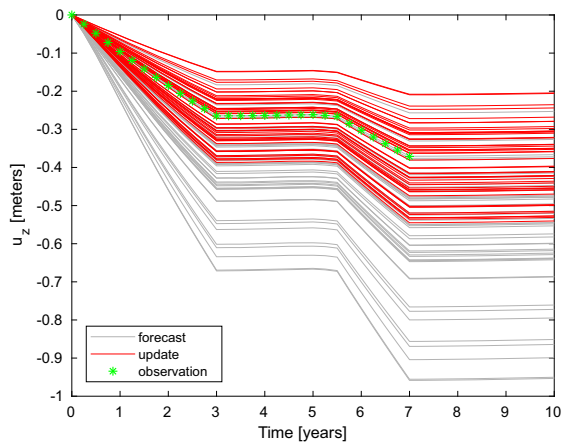
(b) $J(AE)=-6\%$ $J(AES)=-92\%$



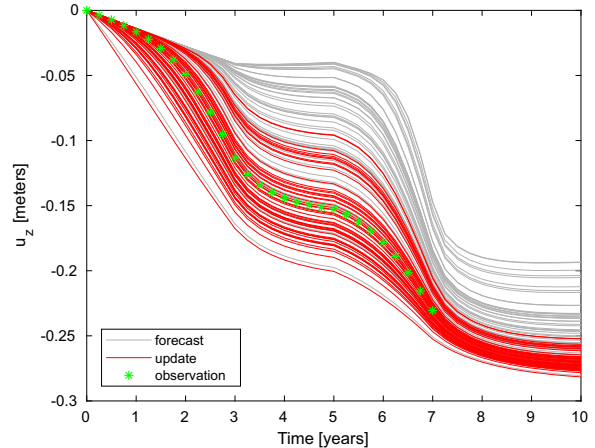
(c) $J(AE)=54\%$ $J(AES)=30\%$



(d) $J(AE)=30\%$ $J(AES)=-29\%$



(e) $J(AE)=61\%$ $J(AES)=40\%$

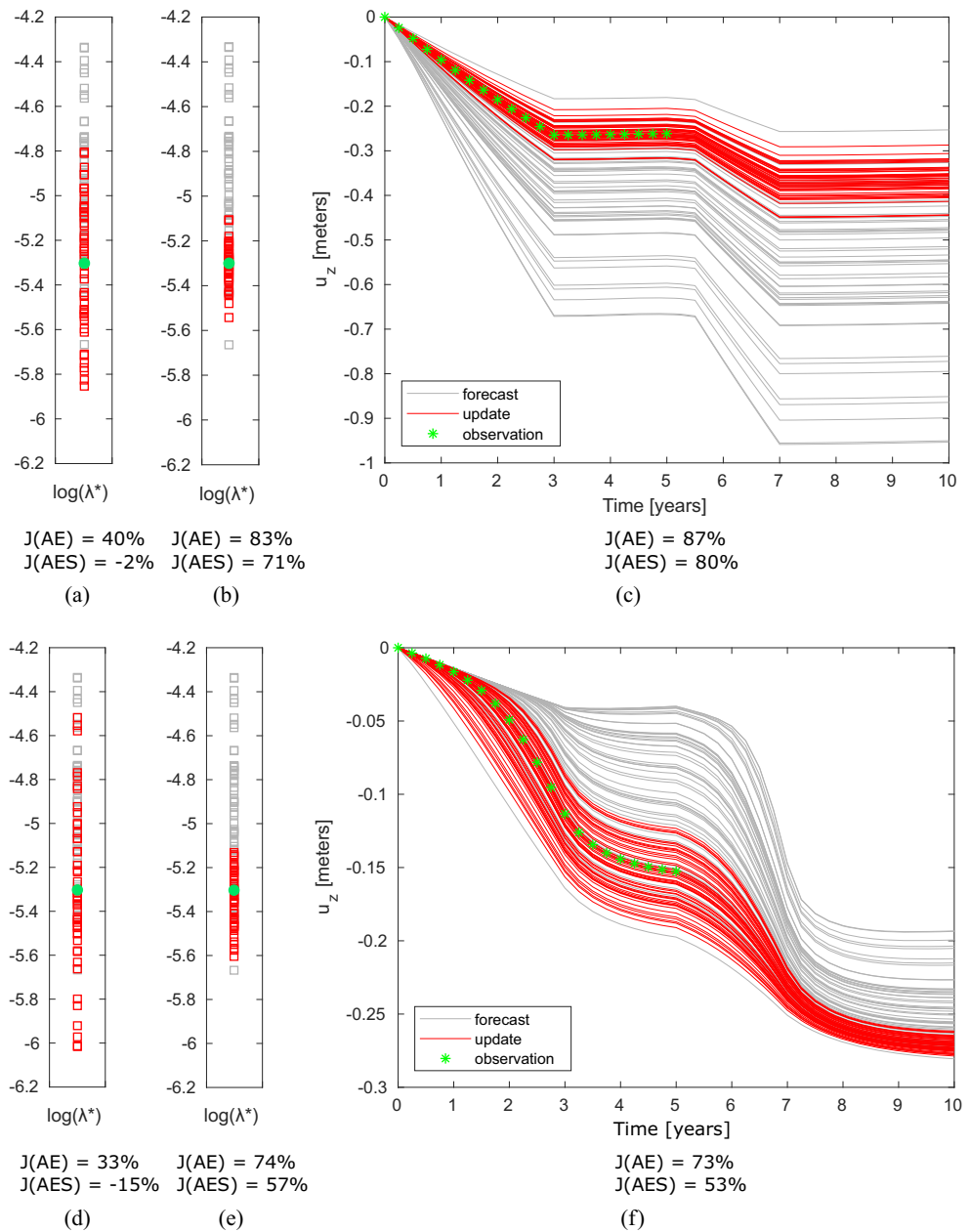


(f) $J(AE)=64\%$ $J(AES)=37\%$

Fig. 4 State ensembles resulting from ES application: maximum land subsidence in time for the MCC (a, c, e) and the VEP constitutive laws (b, d, f) by assimilating 3, 5, and 7 years of displacement measurements, respectively. The forecast and update ensembles are grey

and red, respectively, while green stars denote the exact outcome of the geomechanical model used to get the perturbed observations that are shown in Fig. 3. Variations J of AE and AES with respect to the forecast state ensemble are reported

Fig. 5 Ensembles resulting from ES application: parameter ensembles with the original covariance matrix \mathbf{C}_d (a, d) and the reduced one (b, e) with the MCC and VEP constitutive law, respectively; state ensembles with the reduced \mathbf{C}_d for the MCC (c) and the VEP law (f). Forecast and update ensembles are grey and red, respectively. The AE and AES improvements J with respect to the forecast ensembles are reported

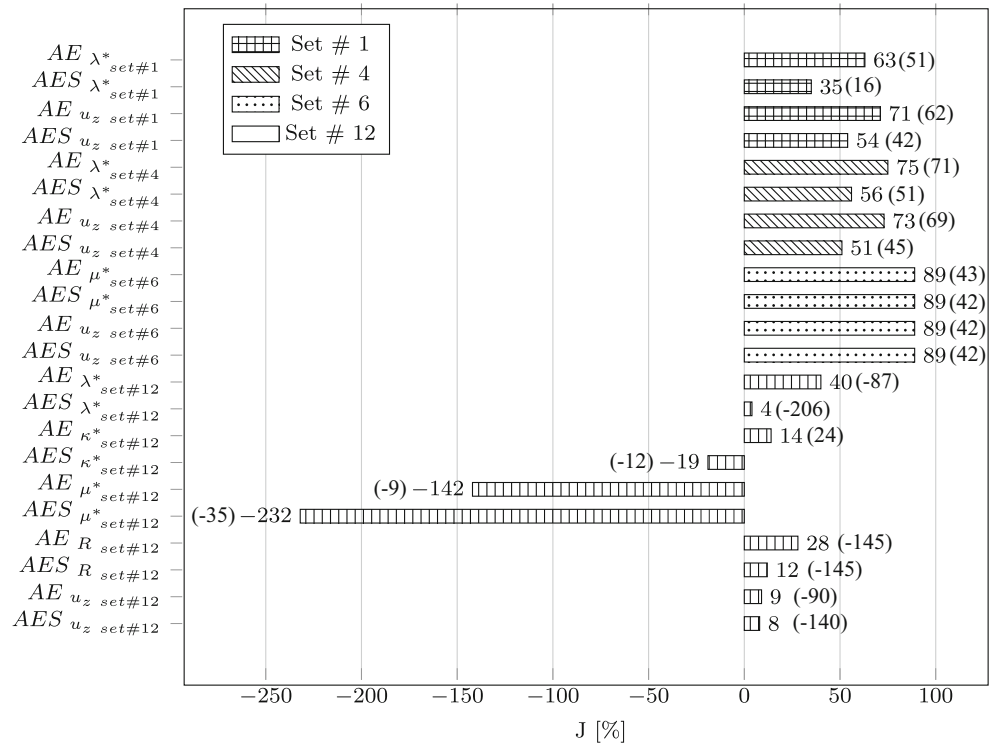


first with the measurement covariance matrix as described before in Section 3.1 and the second with a reduced one. In the latter, we consider again Eq. 21 with the parameters properly changed to obtain a much smaller final error. In particular, the entries of the measurement noise are sampled from a normal distribution with zero mean and standard deviation equal to 1 mm, while the terms of the idealisation noise have been calculated considering Hurst index p equal to 1.67 and variance ζ^2 equal to $2 \cdot 10^{-5} m^2/yr^p$.

Figure 7 shows a comparison between ES and MDA-ES for the estimation of parameter set # 1 considering the original covariance matrix, while in Fig. 8 the reduced one is used. The forecast and update parameter ensembles for

a single ES application and three MDA-ES assimilation of vertical displacements are represented. In the first case (Fig. 7a and b), i.e. for the original covariance matrix, the final update ensembles of the two approaches appear quite similar, except for a few outliers. Conversely, for the reduced covariance matrix (Figs. 8a and b), progressive improvements at every successive assimilation are registered and the final ensemble is more clustered and centered around the true value than the one derived from a single ES assimilation. The latter behavior is the one expected in applying the MDA-ES approach. In the numerical experimentations that follow, the reduced measurement covariance matrix is used.

Fig. 6 Variation J of AE and AES indices with respect to the forecast ensembles reported in Table 4 when both displacement and compaction measurements are assimilated for some significant sets of uncertain parameters, i.e. sets # 1, # 4, # 6 and # 12. In brackets, values of the variation J of AE and AES indices when only the displacement measurements are assimilated (the same as in Table 4) are reported



3.5.2 Comparison between ES and MDA-ES for parameter constraint

In this section, MDA-ES is compared to ES with a reduced measurement covariance matrix, as explained in

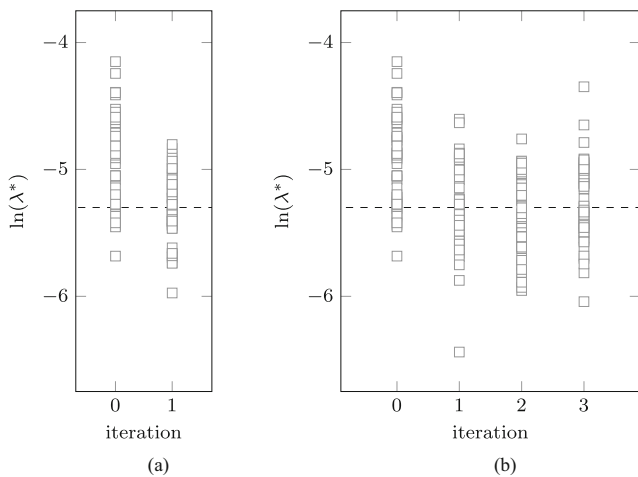


Fig. 7 Parameter ensembles resulting from the ES (a) and MDA-ES (b) applications for the estimate of the modified compression index λ^* with the MCC constitutive law (set # 1) and the original measurement covariance matrix. Dashed line corresponds to true value

Section 3.5.1. Set # 9 formed by uncertain λ^* and the ratio R and an additional set, set # 13 with λ^* and μ^* uncertain, are considered. The sets of Table 4 including the modified swelling index k^* have been discarded because of its low influence for this specific reservoir problem, as discussed

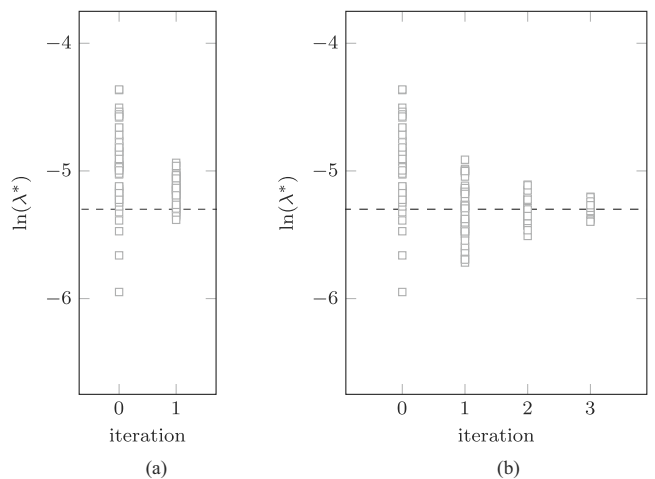


Fig. 8 The same as Fig. 7 with the reduced covariance matrix of measurement error

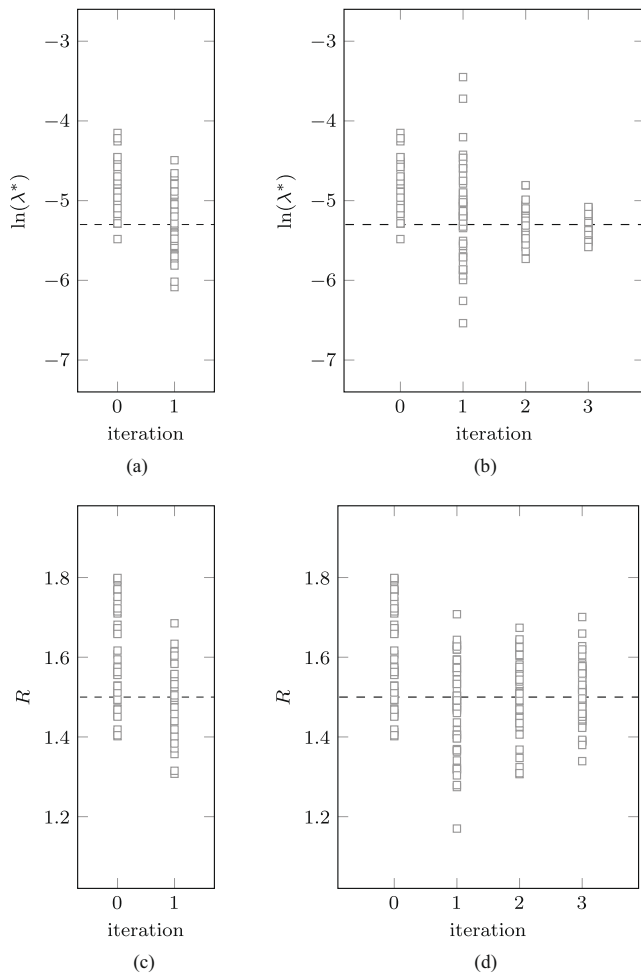


Fig. 9 Parameter ensembles resulting from the ES (a, c) and MDA-ES (b, d) applications for the estimate of the modified compression index λ^* and the ratio R (set # 9) with the reduced measurement covariance matrix. Dashed line corresponds to true value

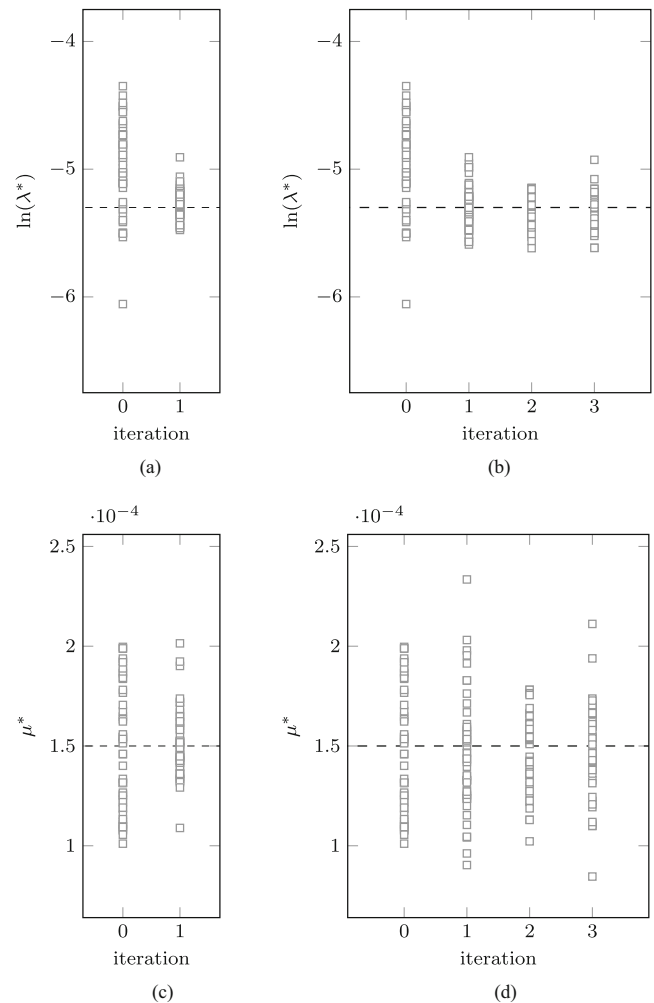


Fig. 10 The same as Fig. 9 for set # 13, i.e. for the estimation of the modified compression index λ^* (a, b) and the modified creep index μ^* (c, d)

in paragraph 3.4.1. Also sets of three and four uncertain parameters are not taken into consideration to avoid an over-parameterized solution.

Figures 9 and 10 show the comparison between ES and MDA-ES for set # 9 and set # 13, respectively. Similarly to the outcome presented in Fig. 8, Fig. 9 shows that during the MDA-ES process progressive improvements are achieved and the final result is better than the one of a single ES assimilation for the constraining of λ^* . By distinction, if the set of parameters is formed by λ^* and μ^* (Fig. 10), MDA-ES does not provide any improvement with respect to the ES results. This is due to the difficulty in the μ^* estimate by using vertical land displacements only, as already observed in Table 4. Similar results are therefore expected for the sets with three and four uncertain parameters.

4 Discussion and conclusion

In this work, we analyze the effectiveness of the integration of DA techniques into numerical models for land subsidence prediction above producing hydrocarbon reservoirs. The aim is to define a modern methodological approach able to take into consideration, quantify and reduce uncertainties in land subsidence prediction by a progressive ‘training’ of the forward numerical model through the assimilation of the available pieces of information. The potential and drawbacks of the different DA approaches have been investigated in a synthetic test case, representative of a real-world hydrocarbon producing reservoir buried in a sedimentary basin. Selected mechanical behaviors and parameter ranges are typical of the Northern Adriatic

basin, Italy. On the basis of the numerical experimentation carried out in the previous sections, the following main considerations are worth summarizing.

1. *Preliminaries.* Land subsidence models, as any other numerical model of real-world processes, are affected by a number of different sources of uncertainties. The first preliminary step consists of recognizing the most influential uncertain factors and, among these, the ones that can be explicitly included as stochastic variables in the construction of the model ensembles. For instance, uncertainties in the subsurface geometry or in the mathematical description of the governing processes cannot be easily quantified, while it is often convenient, and more supported by available information, to treat some material parameters as stochastic variables. Uncertainties in the geometry or the mathematical modeling can be implicitly accounted for by artificially inflating the errors associated to assimilated measurements. Then, the appropriate selection of the set of uncertain material parameters is a fundamental aspect for the application success. A necessary condition for an effective DA is that the uncertain material parameter has a significant impact on the monitored model outcome, i.e. ground motion in this case, while including parameters with a low relevance is often detrimental for the quality of the overall assimilation process. For this reason, a preliminary sensitivity analysis on the relative influence of each material parameter on the model outcome, along with the identification of feasible variation ranges, is of paramount importance. For instance, such sensitivity analysis can be quantitatively performed by mean of Sobol's indices [40], as recently done in [25, 57, 59].
2. *Model diagnostic.* Once the most significant governing parameters are defined, a model diagnostic analysis is useful to preliminarily evaluate the quality of the forecast ensemble. This step is important especially to help identify the most appropriate constitutive behaviour of deep rocks and recognize the actual representativeness of the selected ranges for the uncertain parameter set. To this aim, here we propose the use of the χ^2 -test, which is a simple and inexpensive technique based on the mismatch between model results and measurements. The numerical experiments show that χ^2 -test can be helpful for choosing the most appropriate constitutive law as the one better fitting the observation data, i.e. χ^2 value closest to one. It should be recalled, however, that the χ^2 -test has mainly a qualitative meaning. In case the outcome of the χ^2 -test is not definitely clear, as it can often happen when the quantity and/or quality of the available measurements is limited, it is useful to keep more than one ensemble and select the most appropriate one as new pieces of information come in.
3. *Red Flag.* The RF technique provides a fast analysis that characterizes every Monte Carlo realization with its own probability in a purely Bayesian framework. This can be regarded as an intermediate step between the model diagnostic and the actual assimilation process. In our tests, there is usually an important difference between the highest and the lowest probability, with a relatively small number of realizations with a relevant probability of occurrence. Some exceptions to this outcome has been found for the cases in which χ^2 is lower than one, i.e. when the ensemble should have been rejected a priori in the diagnostic stage. RF provides a preliminary idea of the most likely parameter combination and allows to exclude some realizations that can be considered too unrealistic. In this way, a refinement of the feasible ranges for the uncertain parameter set can be performed, with the possibility of building more representative model ensembles. Of course, the most likely parameter combination along with the expected ranges can change as new observations become available, hence such a refinement should be done with some caution, especially when the amount and quality of such data is limited.
4. *Assimilation.* The assimilation stage allows to incorporate the available measurements and progressively train the geomechanical model as the monitoring of the ongoing process proceeds. The outcome of this stage is a new updated ensemble with a progressively smaller uncertainty in the model prediction. To this aim, we have elected to employ the ES and MDA-ES techniques. ES is a non-sequential DA approach that provides an update ensemble where both the state variables, i.e. surface motion in this case, and the material parameters are constrained by the observation data. The ES application mainly depends on three different aspects: (i) the set of uncertain parameters, (ii) the measurements, and (iii) the error associated to the observations, i.e. ultimately the definition of the measurement covariance matrix. First, the uncertain parameter set has to be actually relevant for the observed process without leading to over-parametrization, i.e. multiple combinations can provide similar results with respect to the measurements. Generally, if the choice of the uncertain parameter set is consistent with the available observations, the ES algorithm appears to be especially suitable for subsidence predictions, with a progressive model improvement as the quantity of assimilated measurements increases. Of course, the effectiveness of ES application improves with the decrease of the error associated to the observations. Hence, the definition

of the covariance matrix of the measurement error, which might be also artificially inflated to account for other sources of uncertainty, plays a fundamental role in the application of the DA algorithm. Anyway, the results obtained in the tests investigated herein suggest ES to be a promising tool to quantify and reduce the uncertainties in land subsidence predictions. MDA-ES uses multiple ES applications with an inflated covariance matrix of the measurement error. In principle, it can be used to improve the ES outcome, especially in case of a strongly non-linear relationship between state variables and uncertain parameters. In our numerical experiments, MDA-ES does not always provide better results than ES, despite the higher computational cost, and appears to be strongly influenced by the selection of the uncertain parameter set and the covariance matrix of the measurement error. Its use in real-world reservoir applications should be better investigated in additional test cases.

On the basis of the considerations above, a tentative workflow for modern land subsidence simulations above producing hydrocarbon fields can be defined. In the context of a synthetic test case, the investigated procedure, consisting of points 1 to 4 above, allows to take into consideration most of the uncertainties that affect the numerical modeling of land subsidence above producing hydrocarbon reservoirs, and then quantify and reduce them as the reservoir development proceeds and new monitoring data are incorporated in the overall model. Further experiences are necessary by investigating real-world producing reservoirs and assimilating real observation data. These experiences, which are currently underway, will allow for refining the steps 1-4 above and help define a ‘protocol’ to be followed for state-of-the-art and reliable predictions of land subsidence.

Funding Open access funding provided by Università degli Studi di Padova within the CRUI-CARE Agreement.

Declarations

Conflict of Interests The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Aanonsen, S.I., Naevdal, G., Oliver, D.S., Reynolds, A.C., Vallès, B.: The ensemble Kalman filter in reservoir engineering—a review. *SPE J.* **14**, 393–412 (2009)
2. Baù, D., Alzraiee, A., Zoccarato, C., Gambolati, G., Ferronato, M., Bottazzi, F., Mantica, S., Teatini, P.: Testing a data assimilation approach to reduce geomechanical uncertainties in modelling land subsidence. *Environmental Geotechnics* **3**(6), 386–396 (2016). <https://doi.org/10.1680/envgeo.15.00005>
3. Baù, D., Ferronato, M., Gambolati, G., Teatini, P.: Basin-scale compressibility of the Northern Adriatic by the radioactive marker technique. *Geotechnique* **52**, 605–616 (2002)
4. Baù, D., Ferronato, M., Gambolati, G., Teatini, P., Alzraiee, A.: Ensemble Smoothing of land subsidence measurements for reservoir geomechanical characterization. *Int. J. Numer. Anal. Methods Geomech.* **39**, 207–228 (2015). <https://doi.org/10.1002/nag.2309>
5. Biot, M.A.: General theory of three – dimensional consolidation. *J. Appl. Phys.* **12**, 155–164 (1941). <https://doi.org/10.1063/1.1712886>
6. Bishop, A.W.: The principle of effective stress. *Teknisk Ukeblad* **106**(39), 859–863 (1959). Text of lecture to NGI 1955
7. Candela, T., Osinga, S., Ampuero, J.P., Wassing, B., Pluymaekers, M., Fokker, P., Wees, J., De Waal, H., Muntendam-Bos, A.G.: Depletion-induced seismicity at the groningen gas field: coulomb rate-and-state models including differential compaction effect. *J. Geophys. Res. Solid Earth* **124**, 7081–7104 (2019). <https://doi.org/10.1029/2018JB016670>
8. Carlson, G., Shirzaei, M., Ojha, C., Werth, S.: Subsidence-Derived Volumetric Strain Models for Mapping Extensional Fissures and Constraining Rock Mechanical Properties in the San Joaquin Valley, California. *J. Geophys. Res. Solid Earth* **125**(9). <https://doi.org/10.1029/2020JB019980>. <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JB019980>. E2020JB019980 2020JB019980 (2020)
9. Carrassi, A., Bocquet M., Bertino, L., Evensen, G.: Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *WIREs Climate Change* **9**(5), e535 (2018). <https://doi.org/10.1002/wcc.535>. <https://onlinelibrary.wiley.com/doi/abs/10.1002/wcc.535>
10. Chen, Y., Oliver, D.: Ensemble randomized maximum likelihood method as an iterative ensemble smoother. *Math. Geosci.* **44**. <https://doi.org/10.1007/s11004-011-9376-z> (2012)
11. Chen, Y., Oliver, D.: Levenberg-Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification. *Comput. Geosci.* **17**, 689–703 (2013). <https://doi.org/10.1007/s10596-013-9351-5>
12. Comola, F., Janna, C., Lovison, A., Minini, M., Tamburini, A., Teatini, P.: Efficient global optimization of reservoir geomechanical parameters based on synthetic aperture radar-derived ground displacements. *Geophysics* **81**(3), M23–M33 (2016). <https://doi.org/10.1190/geo2015-0402.1>
13. De Souza Neto, E., Peric, D., Owen, D.: Computational methods for plasticity, theory and application. Wiley, UK (2008)
14. Emerick, A.: Analysis of the performance of ensemble-based assimilation of production and seismic data. *J. Pet. Sci. Eng.* **139**, 219–239 (2016). <https://doi.org/10.1016/j.petrol.2016.01.029>
15. Emerick, A.: Deterministic ensemble smoother with multiple data assimilation as an alternative for history-matching seismic data. *Comput. Geosci.* **22**, 1175–1186 (2018). <https://doi.org/10.1007/s10596-018-9745-5>
16. Emerick, A., Reynolds, A.: History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilation. *Comput. Geosci.* **16**(3), 639–659 (2012). <https://doi.org/10.1007/s10596-012-9275-5>

17. Emerick, A., Reynolds, A.: Ensemble smoother with multiple data assimilation. *Comput. Geosci.* **55**, 3–15 (2013). <https://doi.org/10.1016/j.cageo.2012.03.011>
18. Evensen, G.: The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dyn.* **53**, 343–367 (2003). <https://doi.org/10.1007/s10236-003-0036-9>
19. Evensen, G., Eikrem, K.S.: Conditioning reservoir models on rate data using ensemble smoothers. *Comput. Geosci.* **22**(5), 1251–1270 (2018). <https://doi.org/10.1007/s10596-018-9750-8>
20. Ferronato, M., Castelletto, N., Gambolati, G., Janna, C., Teatini, P.: II cycle compressibility from satellite measurements. *Geotechnique* **63**, 479–486 (2013)
21. Ferronato, M., Frigo, M., Gazzola, L., Teatini, P., Zoccarato, C.: On the radioactive marker technique for in-situ compaction measurements: a critical review. *Proceedings of the International Association of Hydrological Sciences* **382**, 83–87 (2020). <https://doi.org/10.5194/piahs-382-83-2020>. <https://piahs.copernicus.org/articles/382/83/2020/>
22. Ferronato, M., Gambolati, G., Janna, C., Teatini, P.: Geomechanical issues of anthropogenic CO₂ sequestration in exploited gas fields. *Energy Convers. Manag.* **51**, 1918–1928 (2010)
23. Fokker, P., Wassing, B., van Leijen, F., Hanssen, R., Nieuwland, D.: Data assimilation of PS-InSAR movement measurements applied to the Bergermeer gas field EAGE. <https://doi.org/10.3977/2214-4609.20131971> (2013)
24. Fokker, P., Wassing, B., van Leijen, F., Hanssen, R., Nieuwland, D.: Application of an ensemble smoother with multiple data assimilation to the Bergermeer gas field, using PS-InSAR. *Geomechanics for Energy and the Environment* **5**, 16–28 (2016)
25. Formaggia, L., Guadagnini, A., Imperiali, I., Lever, V., Porta, G., Riva, M., Scotti, A., Tamellini, L.: Global sensitivity analysis through polynomial chaos expansion of a basin-scale geochemical compaction model. *Comput. Geosci.* **17**(1), 25–42 (2013). <https://doi.org/10.1007/s10596-012-9311-5>
26. Gambolati, G., Ferronato, M., Teatini, P.: Anthropogenic land subsidence. *Encycl. Hydrol. Sci.* 2443–2459 (2005)
27. Gambolati, G., Pini, G., Ferronato, M.: Numerical performance of projection methods in finite element consolidation models. *Int. J. Numer. Anal. Methods Geomech.* **25**, 1429–1447 (2001)
28. Gambolati, G., Teatini, P., Baù, D., Ferronato, M.: Importance of poroelastic coupling in dynamically active aquifers of the Po river basin, Italy. *Water Resour. Res.* **36**, 2443–2459 (2000)
29. Hendricks Franssen, H., Kinzelbach, W.: Real time groundwater flow modeling with ensemble Kalman filter: joint estimation of states and parameters and the filter inbreeding problem. *Water Resour. Res.* **44**. <https://doi.org/10.1029/2007WR006505> (2008)
30. Isotton, G., Teatini, P., Ferronato, M., Janna, C., Spiezia, N., Mantica, S., Volontè, G.: Robust numerical implementation of a 3d rate-dependent model for reservoir geomechanical simulations. *Int. J. Numer. Anal. Methods Geomech.* **43**(18), 2752–2771 (2019). <https://doi.org/10.1002/nag.3000>. <https://onlinelibrary.wiley.com/doi/abs/10.1002/nag.3000>
31. Janna, C., Castelletto, N., Ferronato, M., Gambolati, G., Teatini, P.: A geomechanical transversely isotropic model of the Po River basin using PSInSAR derived horizontal displacement. *Int. J. Rock Mech. Mining Sci.* **51**, 105–118 (2012)
32. Macini, P., Mesini, E.: Measuring reservoir compaction through radioactive marker technique. *J. Energy Res. Technol.* **124**(4), 269–275 (2002). <https://doi.org/10.1115/1.1506700>
33. NAM: Ensemble based subsidence application to the Ameland gas field - long term subsidence study part two (LTS II)(2017). Overview Report EP201701217189. <https://nam-feitenencijfers.data-app.nl/download/rapport/2a2da56c-fae4-4453-a8a7-c64c43b418a6?open=true>
34. Nepveu, M., Kroon, I., Fokker, P.: Hoisting a red flag: an early warning system for exceeding subsidence limits. *Mathematical geosciences* **42**, 187–198 (2010). <https://doi.org/10.1007/s11004-009-9252-2>
35. Nguyen, S.K., Volontè, G., Musso, G., Brignoli, M., Gemelli, F., Mantica, S.: Implementation of an elasto-viscoplastic constitutive law in Abaqus/Standard for an improved characterization of rock materials. *Proceedings of the 51st U.S Rock Mechanics/Geomechanics Symposium*. DOI paper. ARMA–2017–1043 (2017)
36. Oliver, D.S., Alfonzo, M.: Calibration of imperfect models to biased observations. *Comput. Geosci.* **22**, 145–161 (2018). <https://doi.org/10.1007/s10596-017-9678-4>
37. Pearse, J., Singhroy, V., Samsonov, S., Li, J.: Anomalous surface heave induced by enhanced oil recovery in northern Alberta: InSAR observations and numerical modeling. *J. Geophys. Res. Solid Earth* **119**(8), 6630–6649 (2014). <https://doi.org/10.1002/2013JB010885>. <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JB010885>
38. Rafiee, J., Reynolds, A.C.: Theoretical and efficient practical procedures for the generation of inflation factors for ES-MDA. *Inverse Problems* **33**(11), 115003 (2017). <https://doi.org/10.1088/1361-6420/aa8cb2>
39. Skjervheim, J., Evensen, G.: An ensemble smoother for assisted history matching. *SPE Reservoir Simulation Symposium* **15**. <https://doi.org/10.2118/141929-MS> (2011)
40. Sobol', I.: Sensitivity estimates for nonlinear mathematical models. *Matematicheskoe Modelirovanie* **2**, 112–118 (1993). 1990 in Russian translated in English in Sobol'
41. Spiezia, N., Ferronato, M., Janna, C., Teatini, P.: A two-invariant pseudoelastic model for reservoir compaction. *Int. J. Numer. Anal. Methods Geomech.* **41**, 1870–1893 (2017). <https://doi.org/10.1002/nag.2704>
42. Staniewicz, S., Chen, J., Lee, H., Olson, J., Savvaidis, A., Reedy, R., Breton, C., Rathje, E., Hennings, P.: InSAR reveals complex surface deformation patterns over an 80,000 km² oil-producing region in the Permian Basin. *Geophys. Res. Lett.* **47**(21), e2020GL090151 (2020). <https://doi.org/10.1029/2020GL090151>. <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL090151>
43. Sun, A.Y., Morris, A., Mohanty, S.: Comparison of deterministic ensemble Kalman filters for assimilating hydrogeological data. *Adv. Water Res.* **32**(2), 280–292 (2009). <https://doi.org/10.1016/j.advwatres.2008.11.006>. <http://www.sciencedirect.com/science/article/pii/S0309170808002108>
44. Sun, N.Z., Sun, A.: *Model Calibration and Parameter Estimation: For Environmental and Water Resource Systems*. Springer, New York (2015). <https://doi.org/10.1007/978-1-4939-2323-6>
45. Tarantola, A.: *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society of Industrial and Applied Mathematics, Philadelphia (PA) (2005)
46. Teatini, P., Castelletto, N., Ferronato, M., Gambolati, G., Janna, C., Cairo, E., Marzorati, D., Colombo, D., Ferretti, A., Bottazzi, F.: Geomechanical response to seasonal gas storage in depleted reservoirs: a case study in the Po river Basin, Italy. *J. Geophys. Res.* **116**. <https://doi.org/10.1029/2010JF001793> (2011)
47. Teatini, P., Ferronato, M., Gambolati, G., Baù, D., Putti, M.: Anthropogenic Venice uplift by seawater pumping into a heterogeneous aquifer system. *Water Resour. Res.* **46**. <https://doi.org/10.1029/2010WR009161> (2010)
48. Teatini, P., Ferronato, M., Gambolati, G., Bertoni, W., Gonella, M.: A century of land subsidence in Ravenna, Italy. *Environ. Geol.* **47**, 831–846 (2005)
49. Teatini, P., Ferronato, M., Gambolati, G., Gonella, M.: Groundwater pumping and land subsidence in the Emilia-Romagna coastland, Italy: Modeling the past occurrence and the future trend. *Water Resour. Res.* **42**. <https://doi.org/10.1029/2005WR004242> (2006)
50. Terzaghi, K.: *Erdbaumechanik auf Bodenphysikalischer Grundlage*. F. Düticke, Vienna (1925)

51. Van Leeuwen, P., Evensen, G.: Data assimilation and inverse methods in terms of a probabilistic formulation. *Mon. Weather. Rev.* **124**, 2898–2913 (1996)
52. Vermeer, P., Neher, H.: A soft soil model that accounts for creep. *Proceedings of the International Symposium “Beyond 2000 in Computational Geotechnics”*: 249–261 (1999)
53. Wang, H.F.: *Theory of Linear Poroelasticity with Applications to Geomechanics and Hydrogeology*. Princeton University Press, Princeton (2000). <http://www.jstor.org/stable/j.ctt1jktrr4>
54. Wilschut, F., Peters, E., Visser, K., Fokker, P., Hooff, P.V.: Joint history matching of well data and surface subsidence observation using the ensemble Kalman filter: A field study. *Reservoir Simulation Symposium* 732–743 (2011)
55. Zoccarato, C., Baù, D., Bottazzi, F., Ferronato, M., Gambolati, G., Mantica, S., Teatini, P.: On the importance of the heterogeneity assumption in the characterization of reservoir geomechanical properties. *Geophys. J. Int.* **207**, 47–58 (2016). <https://doi.org/10.1093/gji/ggw259>
56. Zoccarato, C., Baù, D., Ferronato, M., Gambolati, G., Alzraiee, A., Teatini, P.: Data assimilation of surface displacements to improve geomechanical parameters of gas storage reservoirs. *J. Geophys. Res. Solid Earth* **121**. <https://doi.org/10.1002/2015JB012090> (2016)
57. Zoccarato, C., Ferronato, M., Franceschini, A., Janna, C., Teatini, P.: Modeling fault activation due to fluid production: Bayesian update by seismic data. *Comput. Geosci.* **23**, 705–722 (2019). <https://doi.org/10.1007/s10596-019-9815-3>. In press
58. Zoccarato, C., Ferronato, M., Teatini, P.: Formation compaction vs land subsidence to constrain rock compressibility of hydrocarbon reservoirs. *Geomechanics for Energy and the Environment* **13**, 14–24 (2018). <https://doi.org/10.1016/j.gete.2017.12.002>
59. Zoccarato, C., Gazzola, L., Ferronato, M., Teatini, P.: Generalized polynomial chaos expansion for fast and accurate uncertainty quantification in geomechanical modelling. *Algorithms* **13**, 156 (2020). <https://doi.org/10.3390/a13070156>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.