



## Estimate of a spatially variable reservoir compressibility by assimilation of ground surface displacement data

C. Zoccarato<sup>1</sup>, D. Bau<sup>2</sup>, F. Bottazzi<sup>3</sup>, M. Ferronato<sup>1</sup>, G. Gambolati<sup>1</sup>, S. Mantica<sup>3</sup>, and P. Teatini<sup>1</sup>

<sup>1</sup>Department of Civil, Architectural and Environmental Engineering, University of Padova, Padova, Italy

<sup>2</sup>Department of Civil & Structural Engineering, The University of Sheffield, Sheffield, UK

<sup>3</sup>Development, Operations & Technology, eni S. p. A., San Donato Milanese, Italy

*Correspondence to:* C. Zoccarato (claudia.zoccarato@dicea.unipd.it)

Published: 12 November 2015

**Abstract.** Fluid extraction from producing hydrocarbon reservoirs can cause anthropogenic land subsidence. In this work, a 3-D finite-element (FE) geomechanical model is used to predict the land surface displacements above a gas field where displacement observations are available. An ensemble-based data assimilation (DA) algorithm is implemented that incorporates these observations into the response of the FE geomechanical model, thus reducing the uncertainty on the geomechanical parameters of the sedimentary basin embedding the reservoir. The calibration focuses on the uniaxial vertical compressibility  $c_M$ , which is often the geomechanical parameter to which the model response is most sensitive. The partition of the reservoir into blocks delimited by faults motivates the assumption of a heterogeneous spatial distribution of  $c_M$  within the reservoir. A preliminary synthetic test case is here used to evaluate the effectiveness of the DA algorithm in reducing the parameter uncertainty associated with a heterogeneous  $c_M$  distribution. A significant improvement in matching the observed data is obtained with respect to the case in which a homogeneous  $c_M$  is hypothesized. These preliminary results are quite encouraging and call for the application of the procedure to real gas fields.

### 1 Introduction

Fluid extraction from aquifer systems and hydrocarbon reservoirs are among the most frequent causes of anthropogenic land subsidence. In the framework of a sustainable development of energy resources, the availability of numerical models able to reproduce the monitoring data and to predict the future development of the land settlement is nowadays of paramount importance. In this study, an ensemble-based DA method is used to infer the geomechanical parameters characterizing the rock formation of a deep gas reservoir, thus reducing the prior uncertainties of the geomechanical model response. The DA framework essentially requires three main ingredients: (i) a model to simulate the physical process of interest, (ii) a set of observation data and (iii) a suitable algorithm to incorporate these data into the model response. In this work, a 3-D finite-element (FE) geomechanical model is used to predict the land surface displacements above a gas field where displacement observations are available. The cal-

ibration focuses on the uniaxial vertical compressibility  $c_M$ , which mostly influences the occurrence of land subsidence. Partitioning the reservoir into blocks by faults provides a basis for assuming a homogeneous  $c_M$  within each block, and a heterogeneous  $c_M$  from one block to another. The effectiveness of the DA algorithm to reduce the parameter uncertainty associated with the block-to-block  $c_M$  heterogeneity is evaluated for a synthetic test case. Significant improvements are obtained with respect to the assumption of a homogeneous  $c_M$  field throughout the model domain. In this work, an Ensemble Smoother (ES) is the DA algorithm used to estimate the compressibility by inversion of land surface displacement data. The ES is deemed adequate for this purpose because its implementation (i) avoids the simulation restart necessary to apply other more common data assimilation techniques, such as the Ensemble Kalman Filter (EnKF); (ii) significantly reduces the overall computational cost required by geomechanical model and the inversion procedure. In this paper, the description of the geomechanical model and its implementation

are presented in Sect. 2, whereas Sect. 3 reviews the basic concepts of the ES approach. In Sects. 4 and 5, the generation of the prior uncertain parameters and the synthetic case results are presented and discussed with both a homogeneous and a heterogeneous  $c_M$ .

## 2 Geomechanical model

### 2.1 Model description

The subsurface deformation results from the pore pressure change in space and time due to fluid injection into, or extraction from, deep reservoirs. In order to simulate the deformation up to the ground surface, we need to solve both the governing flow and the structural partial differential equations (PDEs). Geomechanical simulations are performed using a finite-element (FE) poro-elasto-plastic numerical model (e.g. 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 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  $\nu$  is assumed to be known and equal to 0.3.

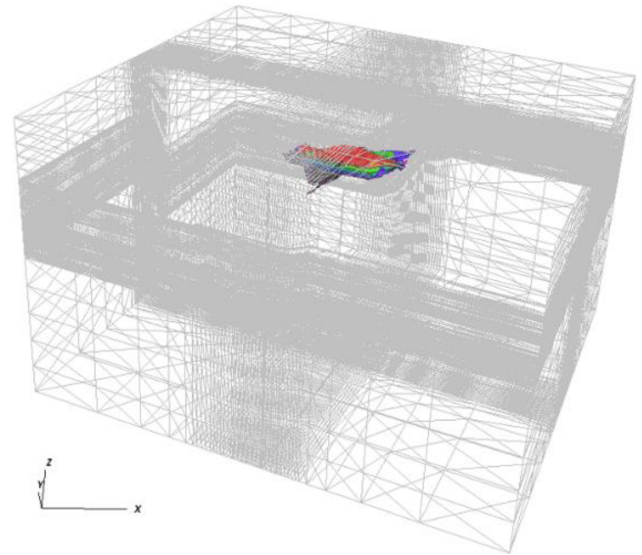
### 2.2 Model setup

A one-way coupling approach is followed with the flow model first run and the outcome subsequently used as input data for the geomechanical model. The geomechanical FE grid comprises 320 901 nodes and 1 824 768 tetrahedral elements. The model domain covers an area of 52 km  $\times$  49 km and extends to about 5 km depth. Zero-displacements conditions are prescribed on the lateral and bottom boundaries. The top of the domain is a traction-free boundary. The simulations span one year, during which the reservoir experiences a fluid extraction. The pressure data are synthetic. Figure 1 shows the reservoir domain and the grid used for the geomechanical simulations.

## 3 Ensemble smoother

The ES algorithm consists of Monte Carlo stochastic simulations based extension to nonlinear problems of the classic Kalman Filter (KF) (Kalman, 1960). The ES algorithm follows a two-step forecast-update process. The forecast step involves simulating an ensemble of model states  $\mathbf{X}$  based upon the solution of the geomechanical FE model  $\Phi$ , which depends upon uncertain parameters  $\mathbf{P}$  and forcing terms  $\mathbf{q}$  (e.g. pore-pressure):

$$\mathbf{X} = \Phi(\mathbf{P}, \mathbf{q}). \quad (1)$$



**Figure 1.** Axonometric view of the 3-D FE grid of the geomechanical model. The reservoir is embedded within the grid with different colors distinguishing the producing layers. The vertical exaggeration is 5.

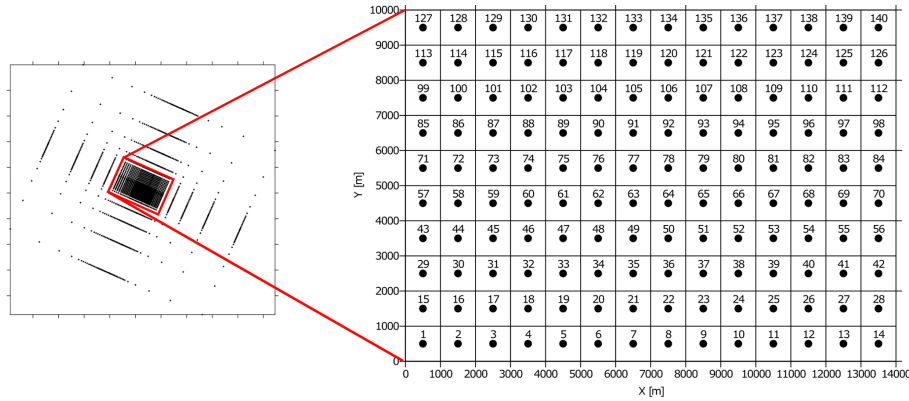
In these simulations, each model state  $\mathbf{X}$  is represented by land surface displacements from a subsidence distribution map. Each realization of the ensemble is run forward in time using random sets of the uncertain geomechanical parameters  $\mathbf{P}$ , thus creating an ensemble  $\mathbf{X}^f$  of model states  $\mathbf{X}$ . The model results at any given location in the simulated domain are spread over a range of values, representing the uncertainty in the surface displacement prediction. In the update step, the set of measurements  $\mathbf{z}$  collected to-date, i.e. point measurements of land displacement at a number of locations, is perturbed to account for measurement errors and assimilated into the forecast system state  $\mathbf{X}^f$  to produce the updated state ensemble:

$$\mathbf{X}^u = \mathbf{X}^f + \mathbf{K} \cdot (\mathbf{d} - \mathbf{H} \cdot \mathbf{X}^f). \quad (2)$$

Matrix  $\mathbf{d}$  includes the perturbed measurements, and the matrix  $\mathbf{H}$  contains binary constants (0 or 1) that map model results at measurement locations. The matrix  $\mathbf{d} - \mathbf{H} \cdot \mathbf{X}^f$  incorporates the residual at these locations between the measured and the predicted data. The matrix  $\mathbf{K}$  is called the “Kalman Gain” matrix (Kalman, 1960), and has the following structure:

$$\mathbf{K} = \mathbf{C}^f \mathbf{H}^T \cdot (\mathbf{H} \mathbf{C}^f \mathbf{H}^T + \mathbf{R})^{-1} \quad (3)$$

where  $\mathbf{C}^f$  is the forecast error covariance matrix associated with the model forecast  $\mathbf{X}^f$ , and  $\mathbf{R}$  is the measurement error covariance matrix associated with the perturbed measurements  $\mathbf{d}$ . The matrix  $\mathbf{K}$  plays the dual role of: (a) spreading information from measurement locations to adjacent areas,



**Figure 2.** 2-D view of the geomechanical FE grid of Fig. 1. The enlargement view within the red rectangle refers to the area affected by the pore pressure change.

allowing for the measurements to correct the predicted values throughout the model domain; and (b) acting as a weight that scales the correction terms according to model and measurement errors. As  $\mathbf{R}$  approaches zero, which means low-error measurements, the influence of  $\mathbf{K}$  increases and the residual is weighted more heavily, so that the model forecast approaches the measurements. In contrast, as  $\mathbf{C}^f$  approaches zero, which indicates a relative agreement among model realizations, the influence of  $\mathbf{K}$  decreases, and the residual is weighted less heavily, so that the model forecast receives little or no correction from measurements.

Within the ES algorithm, any variable incorporated into the system state matrix  $\mathbf{X}^f$  can be corrected by assimilating measurement data if a spatial correlation exists between these variables and the data (van Leeuwen, 2001). In Eq. (1), since the geomechanical parameters  $\mathbf{P}$  dictate the behavior of the model response  $\Phi$ , all uncertain parameters used in the forecast step can be included into the system state matrix  $\mathbf{X}^f$ , and conditioned by land surface movements in Eq. (3). This conditioning may provide updates for the geomechanical parameters  $\mathbf{P}$  that should approach those of the rock formation. Compared to other techniques for characterizing subsurface systems, the ES algorithm is quite attractive because of its low computational burden and the ability to run entirely independent of the simulation model (Bailey and Baù, 2010a, b).

#### 4 Prior distribution of uncertain parameters

In the present study,  $c_M$  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 calibration factor  $f_{CM}$ . In this section, the generation of the prior PDFs in two test cases is described: (1)  $f_{CM}$  is uniform within the reservoir, and (2)  $f_{CM}$  is spatially distributed. In the latter, the heterogeneity on  $f_{CM}$  occurs only

in the area shown in the enlargement of Fig. 2 where the pressure has changed due to fluid extraction.

##### 4.1 Homogeneous compressibility (test case 1)

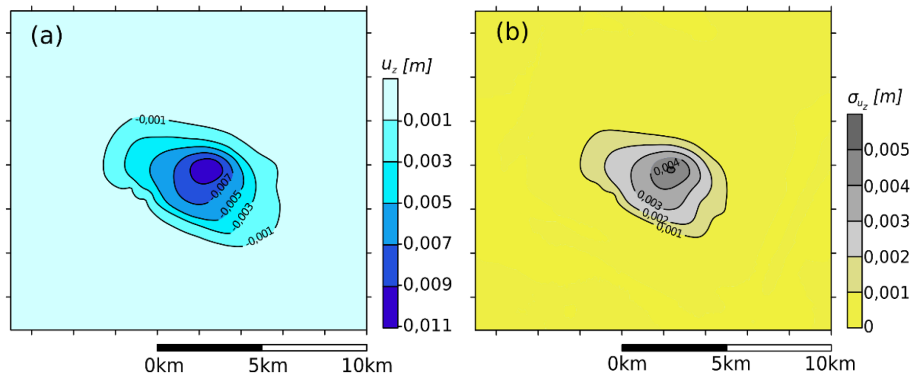
The calibration factor  $f_{CM}$  is assumed to be spatially uniform within the area of Fig. 2. The values of the prior  $f_{CM}$  ensemble are randomly sampled between 1 and 10 from a uniform PDF:

$$f_{CM} \in U[1, 10]. \tag{4}$$

The selected range is based on the outcome of a sensitivity analysis (not shown here) carried out to investigate the possible interval of the  $f_{CM}$  variation. Figure 3 shows the spatial distribution of the mean and the standard deviation ( $\sigma_z$ ) of the vertical displacements ( $u_z$ ) from the forecast ensemble obtained by performing 100 Monte Carlo geomechanical simulations using the prior  $f_{CM}$  ensemble.

##### 4.2 Heterogeneous compressibility (test case 2)

In test case 2, the  $f_{CM}$  is spatially distributed within the same area used in test case 1. This area is  $(14 \times 10) \text{ km}^2$  wide and subdivided into 140 square cells. Each cell is characterized by a different  $f_{CM}$ . A categorical indicator algorithm, that creates random realizations of a heterogeneous  $f_{CM}$  field according to a given covariance model, is used. The  $f_{CM}$  values are drawn from a discrete uniform distribution with the prescribed ten categories ranging from 1 to 10. Each category has an equal unconditional probability (1/10). To account instead for the spatial statistical dependence between  $f_{CM}$  values on different cells, a stationary correlation model is introduced. According to this model, the probability of observing a  $f_{CM}$  category in the discrete domain  $[1, 10]$  at any given cell is conditional to the presence of the same category at surrounding cells. The correlation between two grid cells is based on an exponential isotropic function law depending



**Figure 3.** Test case 1: (a) mean and (b) standard deviation ( $\sigma_z$ ) from the forecast ensemble of the vertical displacement ( $u_z$ ).

on the distance between these cells and a prescribed correlation length,  $\lambda$ . The  $\lambda$  value has a direct influence on the degree of heterogeneity assigned to the spatial distribution of  $f_{CM}$ . Preliminary simulations have suggested the choice of  $\lambda = 4000$  m. Figure 4 shows one of the 100 realizations of the generated prior ensemble of the  $f_{CM}$  field. Obviously, the mean over the ensemble in each grid block is equal to 5. As for test 1, the mean and the standard deviation of the  $u_z$  forecast ensemble are shown in Fig. 5.

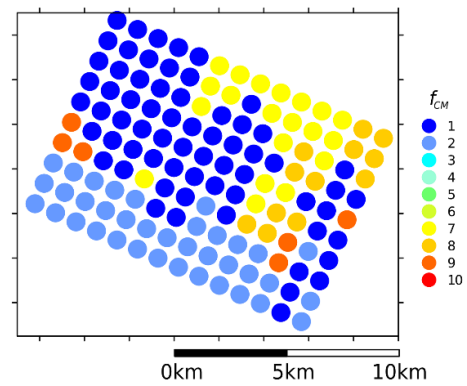
### 5 Synthetic land subsidence data

Ground-surface displacements data are used to infer the model state and the geomechanical parameters. The observations are collected from the land subsidence map shown in Fig. 6, obtained from a geomechanical reference simulation with a prescribed and “known” compressibility distribution. The  $f_{CM}$  field is assigned on the basis of a plausible reservoir partition derived from the presence of faults and thrusts (Fig. 6). The synthetic data locations are uniformly distributed over the reservoir (Fig. 6) in the area with  $\sigma_z$  greater than 0.001 (see Figs. 3b and 5b). This is to account for the reservoir behavior where pressure variation occurs causing land surface deformation.

## 6 Results and discussion

### 6.1 Update of $f_{CM}$

In test case 1, a significant reduction of the uncertainty associated with the prior  $f_{CM}$  distribution is evident by comparing the prior and the posterior cumulative distribution functions (CDFs) plotted in Fig. 7. Defining the ensemble spread as the average absolute difference between the ensemble members and the ensemble mean, the Average Ensemble Spread (AES) for  $n_{MC}$  Monte Carlo realizations is calculated as:

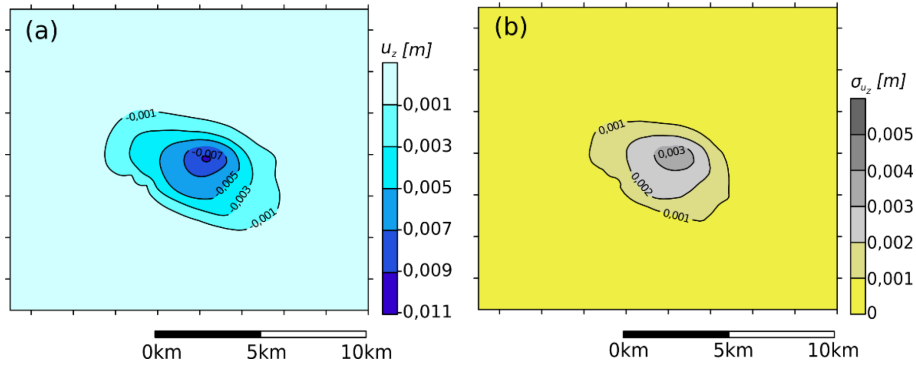


**Figure 4.** Test case 2: one realization out of the 2-D  $f_{CM}$  field ensemble. The compressibility varies in the area corresponding to the enlargement of Fig. 2.

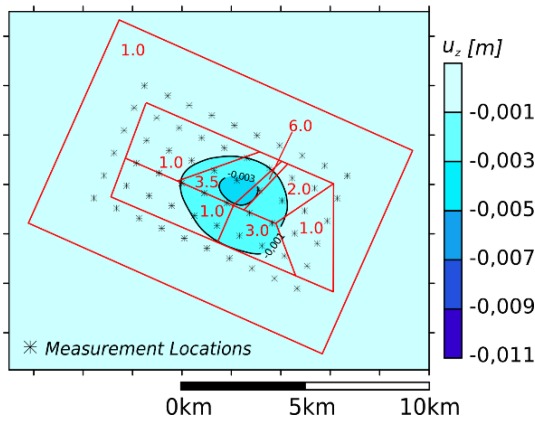
$$AES = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} |x_i - \bar{x}| \quad (5)$$

where  $x_i$  and  $\bar{x}$  are the  $f_{CM}$  value of the  $i$ th ensemble member and the ensemble mean, respectively. 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  $f_{CM}$  equal to 1.89, corresponding to the mean, or expected value, of the updated ensemble.

In test case 2, the ES performance is evaluated by calculating the AES index over each grid block of Fig. 2. 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 Fig. 8b). Although data are collected only over a limited portion of the domain, a sensitivity analysis (not shown here) reveals that collecting data over the entire domain does not improve the assimilation outcome. Indeed, these observa-



**Figure 5.** Test case 2: (a) mean and (b) standard deviation ( $\sigma_z$ ) from the *forecast* ensemble of the vertical displacement ( $u_z$ ).

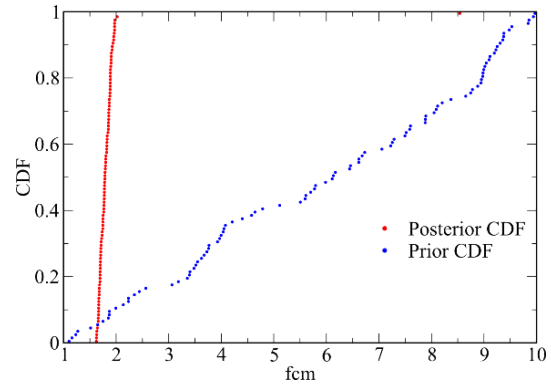


**Figure 6.** Spatial distribution of the synthetic land subsidence data ( $u_z$ ). Assimilation data are collected at the 60 measurement locations displayed in the map. The red lines represent the trace of the faults subdividing the reservoir into blocks with the prescribed reference  $f_{CM}$  values.

tions cannot yield enough information to infer the model parameters because the deep reservoir experiences small pressure variation with a negligible influence on the land surface deformation of outer regions. 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 (Sect. 6.2). Figure 8a depicts the spatial distribution of the mean from the updated  $f_{CM}$  field.

### 6.2 Forecast of ground surface displacement

The quality of the parameter estimation is validated by executing the posterior geomechanical simulations for both uniform and spatially variable  $c_M$ . The  $f_{CM}$  model input is the mean of the updated  $f_{CM}$  ensemble from the outcome of test cases 1 and 2. The results of this simulation are compared with the synthetic land surface data of Fig. 6 using the Nor-

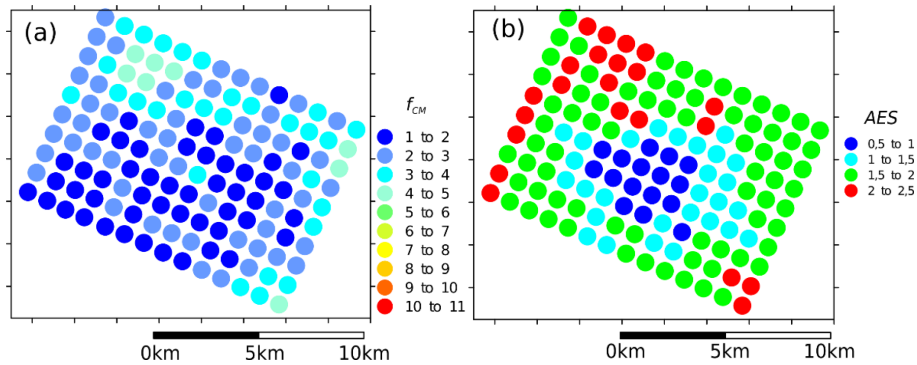


**Figure 7.** Test case 1: prior and Posterior CDFs from the updated  $f_{CM}$  ensemble.

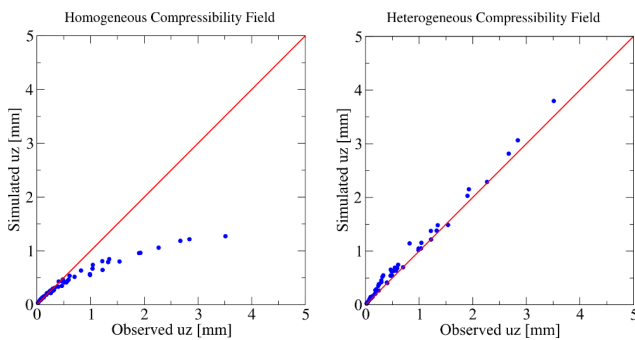
malized Root Mean Square Error (NRMSE), which represents the standard deviation of the differences between the simulated values and the observations, calculated as

$$NRMSE = \frac{\sqrt{\sum_{i=1}^N (u_{zi, sim} - u_{zi, obs})^2}}{(u_{zobs, max} - u_{zobs, min})} \quad (6)$$

where  $u_{zi, sim}$  and  $u_{zi, obs}$  are the simulated and observed land ground displacement at the  $i$ th assimilation location, respectively;  $u_{zobs, max}$  and  $u_{zobs, min}$  are the maximum and minimum observation values, respectively, and  $N$  is the total number of assimilation locations, i.e. 60 data points. Test case 1 and 2 gives NRMSE values equal to 15 and 3 %, respectively. Hence, the predicted vertical displacements provides lower data mismatch in test case 2 and the assumption of a heterogeneous compressibility field allows for a better fit of the synthetic observations (Fig. 9). The fitting of the model predictions to the reference data is very accurate with the heterogeneous  $c_M$ , while a relevant underestimation is found in the homogeneous test case for the largest subsidence values, i.e.  $u_z$  greater than 1 mm.



**Figure 8.** Test case 2: spatial distribution of (a) the mean and (b) the performance index AES from the updated  $f_{CM}$  ensemble after data assimilation.



**Figure 9.** Simulated vs. observed values of  $u_z$  in test case 1 (homogeneous  $c_M$ ) and test case 2 (heterogeneous  $c_M$ ).

## 7 Conclusions

In this study, an ensemble-based DA approach, i.e. the ES, is used to infer the compressibility of the geomechanical model of a producing hydrocarbon reservoir by assimilating ground surface displacements. The methodology is applied herein to investigate two different conceptual models assuming (i) a homogeneous  $c_M$  within the reservoir and (ii) a heterogeneous reservoir compressibility. The latter assumption is made to account for the reservoir compartmentalization due to the presence of a complex fault system. Thus,  $c_M$  spatially varies within the reservoir and calls for the calibration of a heterogeneous  $c_M$  field. A test case is used to assess the validity of the proposed methodology on a reservoir with a synthetic pressure variation. The assimilation of synthetic ground-surface displacements, i.e. simulated data obtained with a priori known  $c_M$  distribution, provides satisfactory results. In particular, the heterogeneous  $c_M$  field gives a much better fitting than the homogeneous  $c_M$ . Hence, a more satisfactory application to a real case is expected by a DA approach with a spatially variable  $c_M$ .

**Acknowledgements.** This research was supported by eni S. p. A. within the DAG (DA in Geomechanics) project.

## References

- Bailey, R. T. and Baù, D.: Ensemble Smoother assimilation of hydraulic head and return flow data to estimate hydraulic conductivity distribution, *Water Resour. Res.*, 46, W12543, doi:10.1029/2010WR009147, 2010a.
- Bailey, R. T. and Baù, D.: Estimating spatially-variable first-order rate constants in groundwater reactive transport systems, *J. Contam. Hydrol.*, 122, 104–121, doi:10.1016/j.jconhyd.2010.11.008, 2010b.
- Baù D., Ferronato, M., Gambolati, G., and Teatini, P.: Basin scale compressibility of the Northern Adriatic by the radioactive marker technique, *Geotechnique*, 52, 605–616, 2002.
- Ferronato, M., Castelletto, N., Gambolati, G., Janna, C., and Teatini, P.: II cycle compressibility from satellite measurements, *Geotechnique*, 63, 479–486, 2013.
- Gambolati, G., Ferronato, M., Teatini, P., Deidda, R., and Lecca, G.: Finite element analysis of land subsidence above depleted reservoirs with the pore pressure gradient and the total stress formulations, *Int. J. Num. Anal. Meth. Geomech.*, 25, 307–327, 2001.
- Kalman, R. E.: A new approach to linear filtering and prediction problems, *Trans. ASME*, 82, 35–45, 1960.
- van Leeuwen, P. J.: An ensemble smoother with error estimates, *Mon. Weather Rev.*, 129, 709–728, 2001.