

ASSESSMENT OF LOCAL HYDRAULIC PARAMETERS BY ENKF DATA ASSIMILATION IN REAL AQUIFERS: A CASE STUDY IN DOWNTOWN PADOVA (ITALY)

Matteo Camporese, Elena Crestani and Paolo Salandin

University of Padova
Department of Civil, Environmental, and Architectural Engineering
Via Loredan 20, Padova, 35131 Italy
e-mail: matteo.camporese@unipd.it

Key words: ensemble Kalman filter, groundwater modeling, data assimilation

Summary. The calibration of natural aquifer model parameters in real world cases is a challenging goal that requires a careful tuning work, whose complexity increases with the number of available data. Paradoxically, instead of making easier the subsurface model numerical assessment, a high spatial density of data highlights the local scale heterogeneity effects, complicating the calibration procedure when an accurate reproduction of the local water table variations is requested.

This is the case of the subsurface in the “Eremitani” area in downtown Padova (Italy), where an extensive monitoring was recently carried out to assess the water table alteration due to the possible realization of important underground works close to the “Scrovegni Chapel”, a renowned historical monument. Due to the implications for the foundation stability of the monumental building, the delicate structural equilibrium of the famous Giotto’s frescoes may be altered by the variations of the water table.

For this reason, a relatively large number of piezometers and wells (16) were drilled in an area of approximately 8 ha, within two aquifers characterizing the subsurface medium down to a depth of 30 m. Measurements were collected for a relatively long period and some pumping tests, as well as a field experiment involving the controlled variation of the Piovego Canal – the watercourse crossing the study area –, were realized for monitoring the corresponding response of the water table in each observation well.

To overcome the difficulties related to the calibration of a fully 3D finite element model solving the Richards’ equation, a data assimilation procedure was developed by integrating the groundwater model with the ensemble Kalman filter (EnKF) and the augmented state technique. The objective of this study is to retrieve the most relevant subsurface parameters (e.g., hydraulic conductivity and specific storage coefficient) by assimilating the piezometric data collected by the monitoring network, thus assessing the local scale heterogeneity due not only to the vertical stratification, but also to the horizontal spatial variability characterizing the area under investigation.

1 INTRODUCTION

One of the most challenging issues in groundwater numerical modeling for operational purposes in the field is the calibration of soil hydraulic properties (hydraulic conductivity, specific storage coefficient, and porosity, not to mention the retention curve parameters in the unsaturated zone). While typical field methods such as pumping or infiltration tests can give local estimates (at different scales), the spatial distribution of these properties can only be derived through inverse modeling. For this reason, many efforts have been devoted by the scientific community to develop inverse methods characterized by increased levels of complexity and performance, thanks also to the rapid advances and widespread availability of computational resources. Notwithstanding the large choice of inverse modeling approaches available in the literature (for a review, see, e.g., Vrugt et al.¹), relatively few studies exist in which these methods have been applied to complex real systems at the aquifer scale. The main difficulties stem from the large number of uncertain parameters that often characterize real three-dimensional problems in natural aquifers and thus from the high ratio between the degrees of freedom and the number of data available to constrain the system of interest. Among the few cases in which real groundwater inverse problems have been investigated, it is worth mentioning those for which Bayesian data assimilation methods such as the ensemble Kalman filter (EnKF) have been used as inverse modeling frameworks. Liu et al.², for instance, assimilated hydraulic head and concentration data to obtain the hydraulic conductivity spatial distribution at the MADE site, while Hendricks Franssen et al.³ implemented EnKF in a real operational framework for combined state and parameter estimation in the aquifer supplying water to the city of Zurich.

The objective of the present study is to estimate through the assimilation of pressure head data by EnKF the hydraulic properties of a complex groundwater system located in the city of Padova (Italy) and consisting of both an unconfined and a confined sandy aquifers, divided by a semipermeable layer of silty soils and interacting with a watercourse crossing the area. While similar works published so far (including the two cited above) focused on the hydraulic conductivity, our research takes into account also the uncertainty of the specific storage, which in this case is an important parameter that needs to be estimated in order to simulate the subsurface dynamics with sufficient accuracy.

2 METHODS

2.1 Groundwater model

Groundwater dynamics is simulated using the subsurface module of the model CATHY⁴ (CATchment HYdrology), which solves the three-dimensional Richards' equation for variably saturated porous media by means of the linear tetrahedral finite element method with backward Euler and Picard iterations for the nonlinear system.

The mathematical model can be written as:

$$S_w(\psi)S_s\frac{\partial\psi}{\partial t} + \phi\frac{\partial S_w(\psi)}{\partial t} = \text{div}[K_s K_r(S_w(\psi))(\nabla\psi + \eta_z)] + q_s, \quad (1)$$

where S_w is water saturation, S_s is the aquifer specific storage coefficient [L^{-1}], ψ is pressure head [L], t is time [T], ϕ is the porosity or saturated moisture content, K_s is the saturated hydraulic conductivity tensor [L/T], K_r is the relative hydraulic conductivity function [-], $\eta_z = (0, 0, 1)^T$, with z the vertical coordinate directed upward [L], and q_s represents distributed source or sink terms [L^3/L^3T]. The model is completed by appropriate boundary and initial conditions and by the retention curves $S_w(\psi)$ and $K_r(S_w)$, which can be given following any relationship available in literature (e.g., van Genuchten⁵).

2.2 Ensemble Kalman filter

Data assimilation techniques such as the ensemble Kalman filter allow model simulations to be updated with observation data and provide a framework for incorporating model and data errors and for quantifying prediction uncertainties.

The key idea of EnKF⁶ is to use a Monte Carlo approach to sample the state and measurement error covariance matrices needed by the update formula of the classic Kalman filter. To this aim, an ensemble of realizations is constructed from the nonlinear dynamics of the model by perturbing soil parameters, initial conditions, and boundary conditions.

For the implementation details of EnKF in the CATHY model, see Camporese et al.⁷. The main difference with the EnKF implementation used in this study is that here hydraulic conductivities and specific storage coefficients are updated by means of the state augmentation technique, i.e., soil parameters are added to the state vector in order to be updated together with the system state at each assimilation step.

3 STUDY AREA

The study area (Figure 1) is located in the city of Padova (northeastern Italy) and includes the renowned “Scrovegni Chapel”, which houses a remarkable cycle of frescoes completed in 1305 by Giotto.

A large and ambitious project for a new auditorium is currently under consideration by the town government. The area designated to host the auditorium is just a few hundred meters away from the Chapel, across the Piovego Canal, and will be subject to extensive dewatering in order to facilitate the foundation works. The main concern of the local authorities is that the water table lowering in the “Boschetti” area (Figure 1) might affect also the groundwater dynamics under the Chapel, jeopardizing the preservation of the inestimable frescoes due to possible soil differential settlements.

In order to forecast with a reasonable accuracy the water table response under the Chapel, a monitoring network was installed to collect pressure head data in 16 observation wells (Figure 1) and to use these data to calibrate/validate a groundwater model for simulating a number of future scenarios.

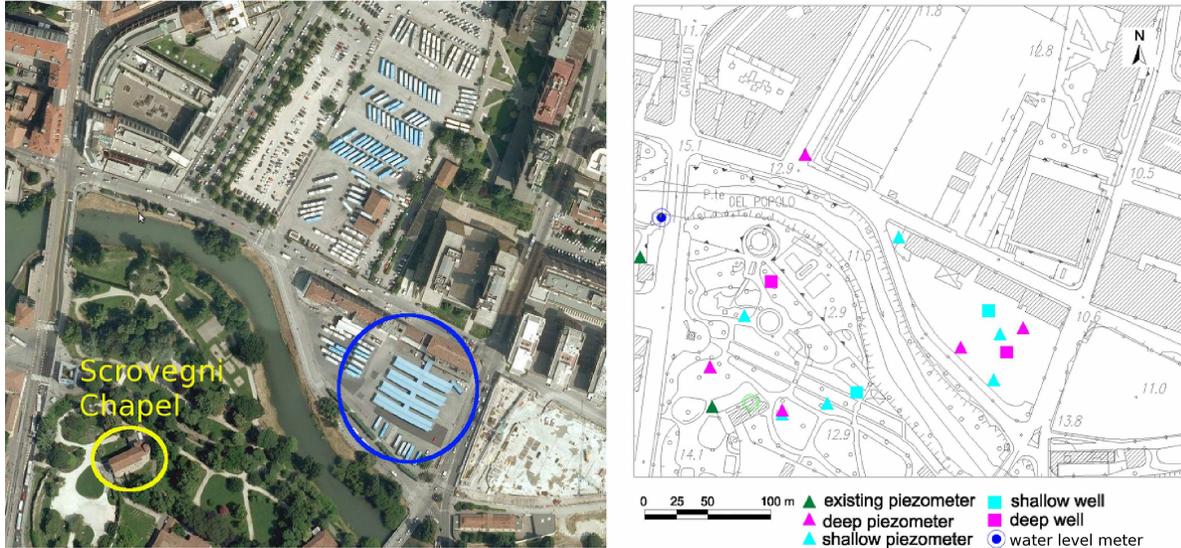


Figure 1: Map of the study area, (left) with the location of the Scrovegni Chapel and the zone (in blue) designated to house the new auditorium, (right) with the location of the observation wells.

A series of pumping tests was carried out to estimate the average hydraulic conductivity and specific storage coefficient and an experiment was performed involving the controlled variation of the Piovego Canal water level, to better understand the interactions between the groundwater and the watercourse.

Based on 12 preliminary cone penetration tests, 8 borehole logs, and the field experiments, it can be hypothesized that the subsurface system consists of two sandy aquifers: one shallow (approximately from 4 to 12 m depth) unconfined aquifer (formation b) and one deeper (approximately from 15 to 28 m depth) confined aquifer (formation d), connected by a semipermeable layer of silty soils (formation c). From the surface down to 4 m depth the soil is composed mostly by filling material (formation a).

4 SIMULATION SETUP

The domain was discretized by an unstructured surface grid (Figure 2) replicated vertically to yield a total number of 153016 nodes and 881640 tetrahedral elements. Moreover, the surface was assumed as divided in seven different soil regions, which, combined with the vertical formations described in the previous section, give a total of $4 \times 7 = 28$ material zones in which the soil parameters are assumed spatially homogeneous. The resulting mesh has coarse elements close to the boundaries, which are far enough from the main zone of interest in the centre (where the grid is refined), so that the hypothesized boundary conditions of no flow have a negligible impact on the solution in the middle of the domain. No flow is imposed also to the bottom of the domain. We simulated the experiment of controlled variation of the Piovego Canal water level, so that boundary

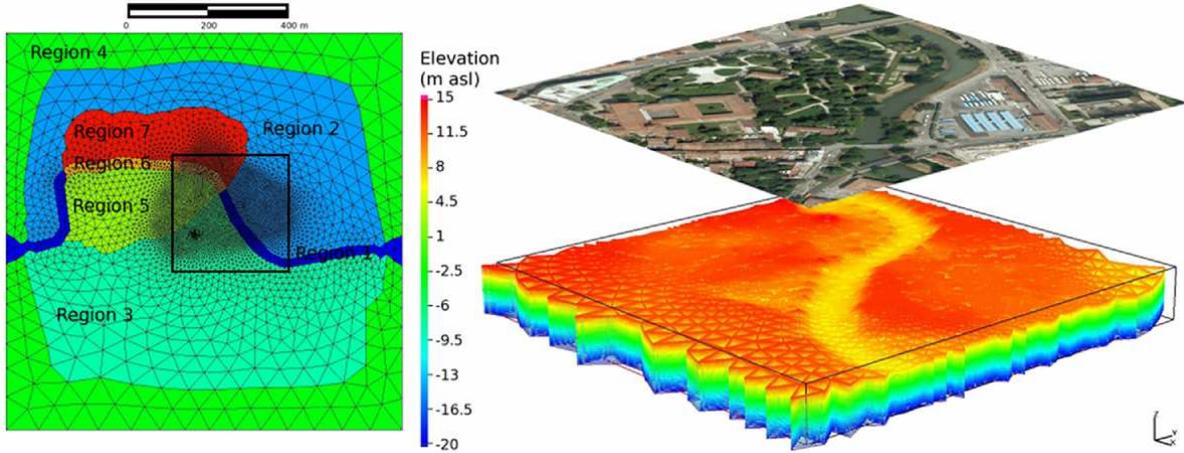


Figure 2: (left) Two-dimensional aerial view of the finite element mesh for the entire domain; (right) three-dimensional view of the same mesh zoomed in the zone of main interest (black rectangle on the left).

conditions at the surface consist of imposed pressure head for the nodes corresponding to the watercourse and atmospheric fluxes (measured rainfall or estimated evapotranspiration) for the remaining nodes. The model was initialized with a 7-day period of measured Piovego Canal levels and measured atmospheric forcing, while the duration of the actual experiment was three days. An ensemble of 50 realizations was generated by perturbing the nominal values of hydraulic conductivity and specific storage coefficient with a 200% coefficient of variation noise, while the uncertainty of atmospheric boundary conditions was set at 20%. Starting from day 7 of the simulation, pressure head data measured in the observation wells were assimilated every 6 hours at the 16 corresponding nodes, updating both pressure head distribution and soil parameters (K_s and S_s) at each assimilation step. Table 1 summarizes the main characteristics of the simulation.

5 RESULTS

Figure 3 shows some examples of how the parameters are updated during the data assimilation run. Not surprisingly, K_s exhibits larger corrections (on average) in the regions and formations where more assimilation points are available, while S_s cannot be easily corrected, with uncertainties that are not reduced by the updates, resulting in a larger residual standard deviation at the end of the simulation. This denotes a scarce sensitivity of S_s to the assimilation of pressure head data. Overall, the final estimated values of K_s and S_s are consistent with the soil classification obtained from log data and pumping test results.

Note that EnKF can potentially identify soil heterogeneity only according to the material zones previously determined. This is clearly a limitation of the present study, being

Grid information	
# of nodes in the surface grid	4936
# of triangles in the surface grid	9796
# of nodes in the 3D grid	153016
# of tetrahedral elements in the 3D grid	881640
Prior soil parameters	spatially homogeneous log-normally distributed
Saturated hydraulic conductivity	$E[K_s] = 1.0 \times 10^{-5}$ m/s, $CV_{K_s} = 200\%$
Specific storage	$E[S_s] = 1.0 \times 10^{-4}$ m ⁻¹ , $CV_{S_s} = 200\%$
Simulation period	7 (initialization) + 3 = 10 days
Atmospheric forcing	log-normally distributed ($CV = 20\%$) and time correlated ($\tau = 1$ h)
Measurements	
Pressure head ψ	normally distributed ($\sigma = 0.01$ m)
Ensemble size N	50

Table 1: Model discretization and parameter values for the simulation. Symbols CV , τ , and σ denote coefficient of variation, characteristic time scale of correlation, and standard deviation, respectively.

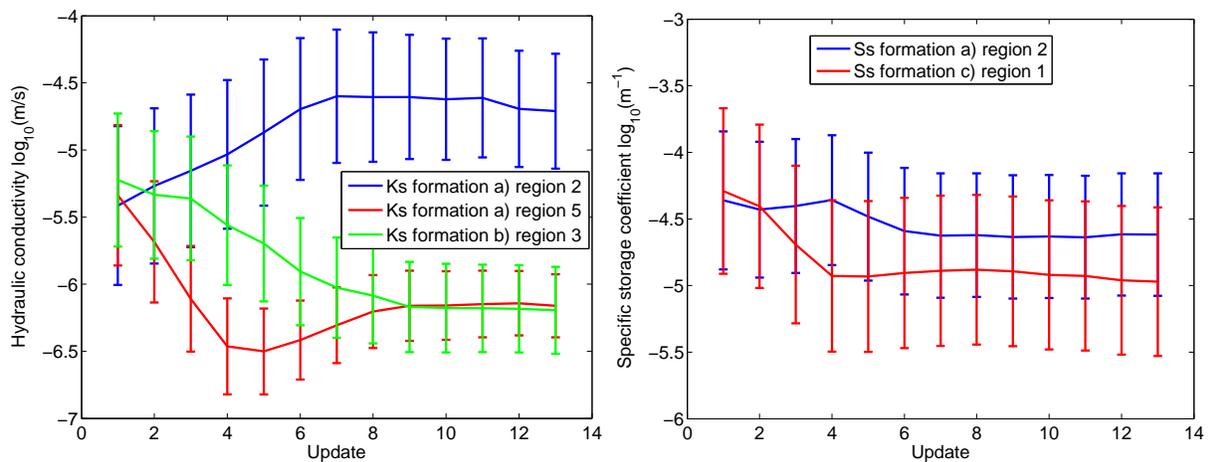


Figure 3: Updates of (left) K_s and (right) S_s for some of the material zones in which the domain is subdivided. Formations are defined in Section 3, while regions are shown in Figure 2. Vertical bars represent \pm one standard deviation.

in our case the EnKF performance potentially affected by a not suitable choice of the planimetric facies distribution.

We define the optimal estimate of the soil parameters as the ensemble average obtained by the data assimilation run at the last update. To evaluate the goodness of this set of K_s and S_s , the same simulation (i.e., the controlled variation of the Piovego Canal water level)

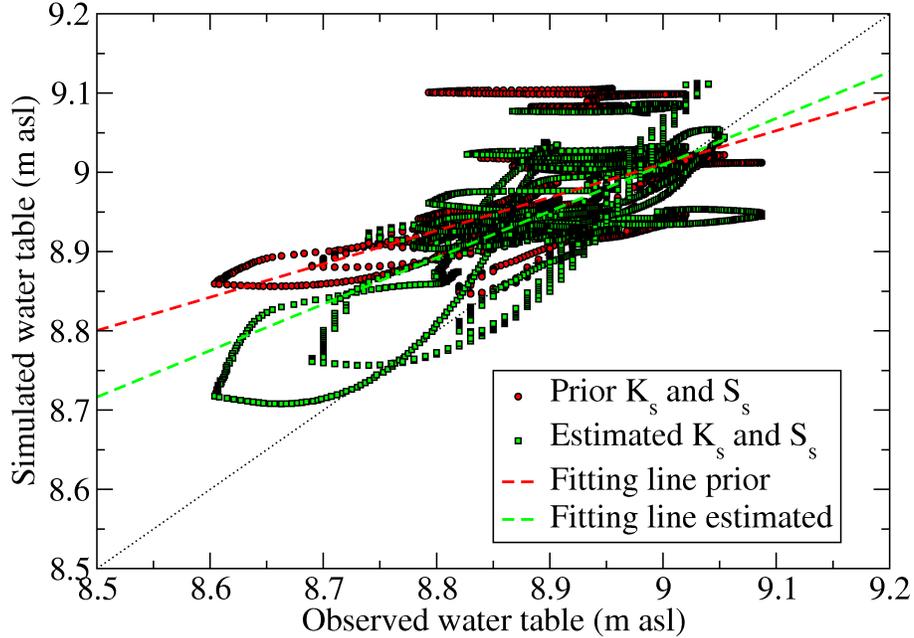


Figure 4: Comparison between water table levels observed and simulated using prior and updated sets of parameters.

Set of parameters	R	$RMSE$ (m)	$NRMSE$ (%)
Prior	0.529	0.089	18.5
Updated	0.678	0.072	14.9

Table 2: Correlation coefficient (R), root mean square error ($RMSE$), and normalized root mean square error ($NRMSE$) of the simulated against measured water table data.

was repeated from the beginning using just one realization with the estimated parameters and compared with the results obtained with the prior distribution of parameters (both K_s and S_s homogenous and equal to 1.0×10^{-5} m/s and 1.0×10^{-4} m $^{-1}$, respectively).

Figure 4 compares the water table levels measured in all the 16 observations wells with those simulated using prior and estimated distributions of K_s and S_s . The simulation run with estimated parameters yields better water table predictions than the run with prior parameters. This is shown by the slope of the fitting lines: the green one represents a clear improvement with respect to the red one, being the former closer to the 1:1 line (average perfect match) than the latter. Moreover, the green simulated data exhibit less spreading than the red data, as confirmed also by Table 2, which reports for the estimated set of parameters a higher value of correlation coefficient (28% more) and a smaller value of root mean square error and normalized root mean square error (19% less) than the prior parameters.

6 CONCLUSIONS

An ensemble Kalman filter (EnKF) was implemented in a finite element model of groundwater flow in variably saturated porous media for the combined estimation of states and parameters (hydraulic conductivity K_s and specific storage coefficient S_s). Using the state augmentation technique, it is possible to update soil hydraulic properties by assimilating pressure head data obtained by a number of available observation wells. The methodology was applied to a natural aquifer in the city of Padova (Italy), for which the hydrogeological characterization is of particular importance for the preservation of one of the Giotto's masterpieces, the fresco cycle contained in the Scrovegni Chapel.

Preliminary results show that the technique is suitable to give a first realistic estimate of the distribution of K_s , while S_s is difficult to correct. However, as implemented so far, EnKF does not allow the identification of different facies, which are currently constrained by the modeler's choice. For this reason, the next step in the inverse modeling procedure development will be the use of EnKF together with a facies generator, to allow a more flexible characterization of heterogeneity at the aquifer scale.

REFERENCES

- [1] J.A. Vrugt, P.H. Stauffer, T. Wöhling, B.A. Robinson, and V.V. Vesselinov. Inverse modeling of subsurface flow and transport properties: a review with new developments. *Vadose Zone J.*, **7**, 843–864, (2008).
- [2] G. Liu, Y. Chen, and D. Zhang. Investigation of flow and transport processes at the MADE site using ensemble Kalman filter. *Adv. Water Resour.*, **31**(7), 975–986, (2008).
- [3] H.J. Hendricks Franssen, H.P. Kaiser, U. Kuhlmann, G. Bauser, F. Stauffer, R. Müller, and W. Kinzelbach. Operational real-time modeling with ensemble Kalman filter of variably saturated subsurface flow including stream-aquifer interaction and parameter updating. *Water Resour. Res.*, **47**, W02532, (2011).
- [4] M. Camporese, C. Paniconi, M. Putti, and S. Orlandini. Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data. *Water Resour. Res.*, **46**, W02512, (2010).
- [5] M.T. van Genuchten. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.*, **44**, 892–898, (1980).
- [6] G. Evensen, *Data assimilation: The ensemble Kalman filter*, Springer, New York, (2007).
- [7] M. Camporese, C. Paniconi, M. Putti, and P. Salandin. Ensemble Kalman filter data assimilation for a process-based catchment scale model of surface and subsurface flow. *Water Resour. Res.*, **45**, W10421, (2009).