EFFICIENT DATA ASSIMILATION TOOL IN CONJUNCTION WITH TOUGH2 FOR CO_2 MONITORING

Judith Yue Li, Sivaram Ambikasaran, Peter K. Kitanidis, Eric Darve

Stanford University
161 Y2E2, 473 Via Ortega,
Stanford, CA, 94305, Santa Clara
e-mail: yuel@stanford.edu

ABSTRACT

We address the problem of monitoring CO_2 in real time to provide early warning of possible leakage. The new data assimilation approach we develop takes advantage of the temporal resolution of continuously acquired seismic survey data, and has been tested on a TOUGH2 synthetic CO2 injection model. This Kalman Filter (KF) based method is sufficiently fast for online data processing and provides risk-based estimates. We also discuss the validity of describing the unknown system dynamics using a random walk model in the context of CO_2 monitoring.

INTRODUCTION

Subsurface monitoring is essential for risk management during many field operations, such as CO_2 injection during Carbon Capture and Sequestration (CCS) and Enhanced Oil Recovery (EOR), hydraulic fracturing. The key challenge in subsurface monitoring is the data assimilation (DA)/ filtering problem of reconstructing an unobservable dynamic process (such as CO_2 plume) using both static data (well-log, seismic data) and dynamic data (such as 4-D seismic data, or well production data).

Time-lapse seismic surveys have been widely adopted for monitoring subsurface fluid flow (Lumley 2001). Recent developments in 4D seismic and continuous active-source seismic monitoring (CASSM) has increased the data output frequency to the order of minutes (Daley et al. 2007). Thus, there is a growing need for a fast data assimilation solution that can update the parameter based on the most recently acquired observations, i.e. real-time monitoring.

On the other hand, practical implementation of data assimilation methods for subsurface monitoring must also address the challenge of limited memory and processing time. Solving the spatio-temporal process jointly incorporating all data sets is time-consuming, whereas sequential filtering only requires storing solutions from the last step and processing the measurement from the current step, which makes the problem tractable for real-time subsurface monitoring purposes.

The general starting point for sequential data assimilation is the Bayesian framework, which finds the posterior pdf of the parameters and model state, i.e., observation-conditioned parameter estimates at each time instance, given a set of measurements and a dynamical model with uncertainties. The dynamic model is assumed to be a Markov model, which holds that state variables at one time-instant are dependent only on the states of the previous time-instant. This condition is normally satisfied for most dynamic models, such as flow simulators. Another assumption is that measurement errors are uncorrelated in time. These two assumptions allow the Bayesian theorem to be written in a recursive fashion. The resulting sequential Bayesian filter obviates the need for reprocessing previous data, while still preserving the influence from measurements at all previous time steps during model updating.

Several Bayesian filters have been applied to subsurface dynamic imaging problems. For linear dynamic problems, the Kalman Filter gives a linear, unbiased, and minimal error-variance estimate of the unknown state (Kalman 1960). The Extended Kalman Filter (EKF), an adaptation of the Kalman Filter for solving nonlinear problems through linearization (Anderson and

Moore 1979), has been applied to field environmental monitoring (Nenna, Pidlisecky and Knight 2011). However, there are two major drawbacks associated with both KF and EKF: (1) advancing a covariance matrix in time is computationally expensive when the number of unknowns is large, and (2) linearization of nonlinear equations by removing high-order terms results in the evolution of the approximation error (Evensen 2003).

The Ensemble Kalman Filter (EnKF), developed by Evensen (1994) provides an alternative solution for data assimilation of large-scale nonlinear problems. A Monte Carlo version of KF, EnKF uses a group of random samples (ensemble) to represent the distribution of the unknown states. For a linear dynamical model, EnKF will converge exactly to KF with larger ensemble size. Another sequential Monte Carlo (SMC) based filter similar to EnKF is the particle filter. However, the particle filter (Doucet, De Freitas and Gordon 2001) requires a large number of realizations to converge, which makes it computationally unrealistic for applications to geophysical problems.

In this paper, we form the monitoring problem as a linear dynamical problem, and compare the performance and computational cost of the Basic Kalman Filter, Fast Kalman Filter, and EnKF. We propose an algorithm that can incorporate fast linear algebra methods to accelerate the statistical filter, in order to cope with real-time high-dimensional problems in subsurface monitoring. We also provide a numerical example of its application to CO_2 monitoring.

DATA ASSIMILATION THEORY

Linear State-Space Model and Random Walk Transition Model

When observations are collected in rapid succession, we can exploit "continuity over time" to combine measurements from different time steps. The intrinsic state s being monitored is a dynamic process governed by a physical model f. We can define a state space model (SSM) that consists of two time series, an unobservable intrinsic state x_k , $k \geq 1$ and an observable intrinsic state $y_k, k \geq 1$ —their relationship is described by $y_i = h(s_i, v_i)$. In the CO_2 monitoring example, we are interested in tracking the CO_2 plume location, so the intrinsic state to be estimated is CO_2 saturation, and the observation is seismic survey data. As CO_2 pushes its way into the pore space, CO_2 saturation changes with time, and the seismic data we record changes accordingly.

We first consider a linear dynamical system governed by the Linear Gauss-Markov model:

$$s_i = Fs_{i-1} + u_i$$
$$y_i = Hs_i + v_i$$

where $F \in R^{m \times m}$ is the state transition matrix; it represents our knowledge about the temporal behavior of s (change in s after one time step); the current state depends only on the previous state. The vector u_i is the zero-mean Gaussian process noise, with $u_i \sim N\ (0, Q)$, Q is the model error covariance matrix. The measurements (given data) are denoted d_i,,,,,,,, , and are related to x_i through linear operator H; the measurement noise is modeled as white noise $v_i \sim N\ (0, R)$, where R is a diagonal matrix.

Assuming that the state s_k is a random variable, whose first two moments, mean $\hat{s}_k$ and error covariance P_k, are known, then the mean and covariance propagate as follows:

$$\bar{s}_k = \bar{s}_{k-1}$$
$$P_k = P_{k-1} + Q_k$$

For Gaussian random variables, the probability distribution of s is completely characterized by its first two moments. Without any observations, the linearized forward model can predict the next state with large uncertainty. Extra information can be extracted from online measurements to improve our knowledge of the states and reduce the uncertainty in the system. For a linear state-space model, KF will give LMMSE estimates of the state given observations.

One common state transition model, the random walk model, is usually adopted when data is acquired rapidly, at a rate faster than discernible change in the system, in the absence of a valid physical model for state evolution (Nenna et al. 2011). This random-walk-model assumption

allows us to write the forecast step in a simple linear form, as follows,

$$s_{k+1} = s_k + u_i$$

Given that the dynamics of a complex system (e.g. subsurface fluid flow) often depends on prior knowledge of parameters with high uncertainties, the random walk model is useful in the absence of a reliable physical dynamical model. Moreover, the computational cost of Kalman Filtering can be greatly reduced given this simple model.

Basic Kalman Filter

The KF provides Linear Minimum Mean Square Error (LMMSE**)** estimates for the linear state-space model above. The Kalman filter algorithm enables predicting the most likely state s_i compatible with the probabilistic model for the evolution model (state transition matrix F) and the measurements (matrix H). There are two steps in this algorithm: (1) Forecast Step: projecting forward the current state and error covariance to obtain an a priori estimate for the next time step; (2) Analysis Step: incorporating the new measurements into the a priori estimate to obtain an improved a posteriori estimate. The Kalman gain K tells you how much weight to put on measurements versus the temporal prediction.

Forecast Step

$$s_f^k = Fs_a^{k-1}$$
$$P_f^k = FP_a^kF^T + Q$$

Analysis Step

$$K = P_fH^T\left(HP_fH^T + R\right)^{-1}$$
$$s_a = s_f + K(y - Hs_f)$$
$$P_a = (I - KH)P_f$$

k: time step index f: forecast, a: analysis
s: state to be estimated , F: state transition matrix, P: state error covariance matrix, Q: forecast error covariance matrix, R: measurement error covariance matrix

For numerical implementations, a more efficient algorithm for the analysis step, in which the covariance update is at the cost of $O(m^2)$, is as follows:

$$P_{yy} = HP^fH^T\ ,\ P_{xy} = P^fH^T$$
$$\left(P_{yy} + R\right)B = y - Hs_f$$
$$s_k^a = s_k^f + MB$$
$$P^a = P^f - P_{xy}(C + R)^{-1}P_{xy}{}^T$$

Ensemble Kalman Filter (EnKF)

The EnKF is a Monte Carlo approximation of the KF developed to approximate the LMMSE estimates when the state dimension is too large for KF to propagate the state error covariance in a reasonable amount of time, or the linearization of the forecast model is not possible.

Consider an ensemble matrix

$$X = [x_1, x_2, \dots, x_N], \in R^{m\times N},$$

where each ensemble member x_i is a realization of a known vector. The ensemble mean denoted with a bar is given by

$$\bar{x} = \frac{1}{N}\sum_1^N x_i\ ,$$

and the ensemble covariance matrix is given by

$$P_e = \frac{1}{N-1}\sum_1^N(x_i - \bar{x})(x_i - \bar{x})^T,$$

The ensemble perturbation matrix is $X' = X - \bar{X}$, thus

$$P_e = X'X'^T/(N-1).$$

Generating an initial state ensemble S and integrating each ensemble member s_i forward in time according to a dynamical model, we have:

$$s_i^f = As_i + u_i \ \text{ or } \ s_i^f = G(s_i, u_i)$$

Given N realizations of the state vector s and observation vector y to form the state perturbation matrix $S' \in R^{m\times N}$ and observation perturbation matrix $Y' \in R^{n\times N}$, we have:

$$P_e^f = S'^f\left(S'^f\right)^T/(N-1)$$
$$R_e = Y'^f\left(Y'^f\right)^T/(N-1)$$

Rewriting the analysis equation in ensemble form:

$$K_e = P_e^fH^T\left(HP_f^eH^T + R_e\right)^{-1}$$
$$S^a = S^f + K_e(Y - HS^f)$$

With the perturbation form of the EnKF update, the error covariance P^f does not need to be constructed explictly, only the low rank approximation $P_e^f = S'S'^T$ is used. The ensemble approximated Kalman Gain thus can be computed at the cost of $O(m)$, which makes EnKF suitable for high-dimensional geophysical problems. The perturbation of observation is necessary for the EnKF to maintain a sufficient spread in the ensemble and prevent filter divergence (Houtekamer and Mitchell 1998).

Fast Kalman Filter

The Kalman Filter requires recursively updating the state and error covariance. If we denote the state dimension using m, then the computational cost of state update is $O(m)$, while the cost of covariance update, according to the Riccati equation, is $O(m^2)$. In geophysical problems, state dimension N is typically $\sim 10^4 - 10^8$, which makes updating and storing the state error covariance matrix of size N^2 at every time step computationally intractable. An alternative idea is to compute and store only the covariance matrix for measurement prediction $P_{yy} = HP_fH^T \in R^{n\times n}$ and cross covariance $P_{xy} = P^fH^T \in R^{m\times n}$. Under the random walk forecast model assumption (and state transition matrix F is an identity matrix), we show that it is possible to circumvent the state error covariance update, thus reducing the computational cost from $O(m^2)$ to $O(m)$.

This new algorithm consists of two parts: initialization and update. As a first step, we initialize the cross covariance as QH^T, which is the product of model error covariance Q and the transpose of the linear measurement operator. Normally, the cost of computing this product is $O(m^2,)$; however, because Q and H have unique structures, we can use fast linear algebra methods such as the Fast Multipole Method (FMM) hierarchical matrix solver as a blackbox to compute the cross covariance at the cost of $O(m)$.

During the update step, we recursively update the cross covariance at the cost of $O(m)$. Then Kalman Gain is computed as a function of cross covariance, which is then used to update the state at the cost of $O(m)$. The overall cost of the Kalman Filter is now $O(m)$, which means that the computational cost grows linearly instead of quadratically with the state dimension.

Reduce sampling error in EnKF with covariance tapering

From the Monte Carlo theory, the convergence rate is $1/\sqrt{N}$ regardless of the state dimension m. Without any approach for reducing the sampling error, the EnKF is both costly and inaccurate compared to KF. However, covariance tapering or localization can improve the covariance estimates using small sample sizes. It reduces the variance in the estimator, but it does introduce a bias to the solution. We can choose a covariance taper C with a similar structure as P_f so that $P_f^e = C \circ P_f^e$. C can be selected as an identity matrix or a distance-based covariance matrix. $C = \theta \exp\left(\frac{\|x\|}{l}\right)$.

Because we need to run the forward model for each ensemble member at every time step, (which is costly), we want to reduce the size of the ensemble as much as possible. However, a small ensemble size may yield spurious correlations that, when used in the Kalman Gain, can cause changes in the state variables over which the data has no real influence. Also, in the standard EnKF, the posterior ensemble is the linear combination of the prior ensemble, whose rank is less than or equal to N_e, yielding a rank-deficiency problem. To reduce both the sampling error and rank-deficiency problem, we can introduce a distance-dependent localization (Houtekamer and Mitchell 2001). The localized covariance matrix is the Schur product of the correlation function ρ and the original sample covariance, $\rho \circ C_{ss}$. The small error covariances associated with remote observation are removed (Aanonsen et al. 2009) and the new EnKF update step is given by

$$K_e = \left[(\rho H^T)\circ(A_sA_y^T)\right]\left[(H\rho H)^T\circ(A_yA_y^T) + A_dA_d^T\right]^{-1}$$

We use the ray operator H to project the correlation function ρ from $R^{N_s\times N_s}$ to $R^{N_s\times N_d}$. After applying localization, the posterior ensemble is

not restricted to the subspace spanned by the prior ensemble. If the sample covariance is positive semi-definite, with positive values on the diagonal, and the localization matrix is positive definite, then from the Schur product theorem, the resulting localized covariance matrix is guaranteed to be of full rank and positive definite.

Covariance localization reduces the variance of the estimator, yet introduces a bias to the solution. Ideally, the covariance taper C should have a structure similar to the P_f, so that $P_f^e \approx C \circ P_f^e$.

NUMERICAL EXPERIMENTS

The purpose of these numerical experiments is to demonstrate the performance of KF and EnKF in reconstructing the dynamic 2-D object (such as a CO_2 plume) from tomographic measurement (seismic travel time) when no deterministic model for the state evolution is given (transition matrix $F = I$).

We use, as data, synthetic results produced by TOUGH2 (Pruess, Oldenburg and Moridis 1999). This model was implemented to reconstruct the CO_2 injection process at the FRIO II site using real field parameters as constraints (Daley, Ajo-Franklin and Doughty 2011). In the FRIO pilot project, investigators injected, at a depth of 1657 m, about 380 tons of CO_2 at a rate of 76T/day into a 17 m thick blue sand brine aquifer, with a dip of 18 degrees, about 30% porosity, and permeability of 1 to over 4 darcy. The breakthrough time for CO_2 at the observation wells (~30 m apart) was around 2 days. Based on well logs and core measurements, a realistic model was constructed with layered permeability and porosity distribution.

Introducing a petrophysical relationship between CO_2 saturation and seismic properties, we can obtain a time-varying map of seismic velocity and attenuation, and then simulate the CASSM response—which can then be compared with the CASSM field measurement. As previously observed, CO_2 is heterogeneously distributed in brine water, so that the relation between CO_2 saturation and the change in the seismic velocity can be described with a patchy petrophysical model better than with the more classic Gassmann model. The change in velocity and attenuation as a function of CO_2 is shown in Figure 3 of the previously cited work. The change in velocity is captured by CASSM. Field CASSM is equipped with one piezotube source located at 1657 m depth generating 1 kHz seismic signals, and 13 receivers located at varying depths. Travel-time measurements are acquired every 15 minutes beginning 2 hours after injection (Daley et al. 2007). The effects of buoyancy are particularly pronounced. Simulation and data suggest that CO_2 flow is straight up near the injection point until it hits the cap rock; then, it continues upwards along the dip of the cap rock.

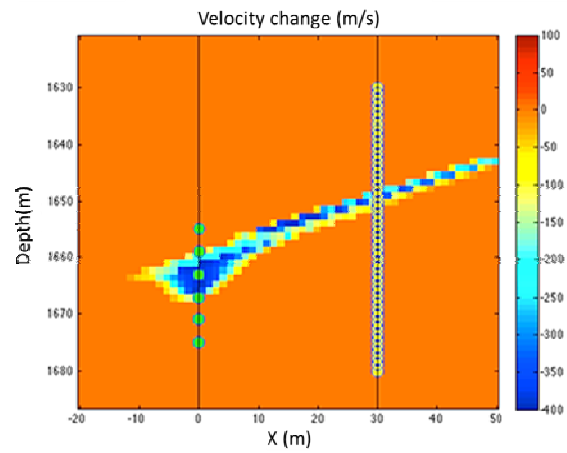

Figure 1. Illustration of general seismic crosswell setup; a transmitter array is installed at the injection well (here: at X-location 0 m) and a receiver array is installed at a nearby monitoring well (here: at X-location 30 m). The CO_2 plume shows up as a decrease in velocity.

The synthetic CASSM survey design is shown in Figure 1. Six sources and 48 receivers are deployed and remain the same over the course of CO_2 monitoring. The background velocity has been removed; only changes in velocity after CO_2 injection are demonstrated in the graph. White Gaussian noise has been added to each measurement, resulting in a 65 dB signal-to-noise ratio (SNR) defined as

$$\mathrm{SNR} = 10 \log_{10} \frac{\|Hs\|_2^2}{\|v\|_2^2}$$

Before injection begins, we assume the CO_2 induced velocity perturbation is zero, so it is reasonable to assume $s_0 = 0$ and $P_0 = 0$, true model and inversion results from KF and EnKF are shown in Figure 2. Pre-injection change in velocity is zero. Three hours after injection, a low-velocity zone began to form around the injection point. Another low velocity zone formed below the cap rock and extended laterally over time, which is predicted well by the KF result. The time interval between adjacent pictures is 30 hours. Here, our target domain to be imaged is between the monitoring wells; thus in the figure, only relevant parts are shown. EnKF gives a larger relative estimation error than the fast KF.

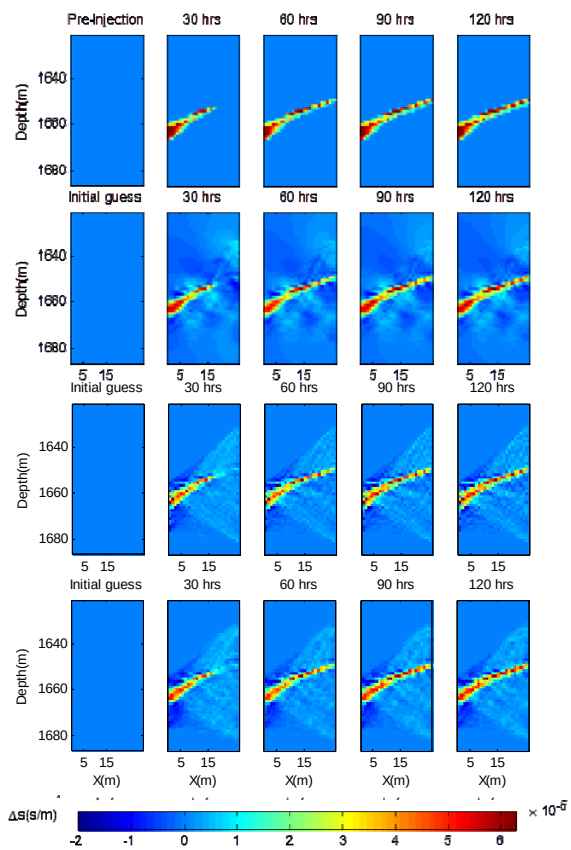

Figure 2. Reconstruct CO_2 induced velocity perturbation (slowness in 10^{-5} s/m) on a 55×59 grid. Only the inversion result between the well is shown in the graph. Case I: True slowness change; Case II: slowness change given by KF; Case III: slowness change given by EnKF with linear covariance tapering; Case VI: slowness change given by EnKF with exponential covariance tapering.

Computational and storage cost is summarized in Table 1. Generally, initialization of cross covariance exhibits quadratic growth with state dimension. However, the Fast Multipole Method (FMM) can reduce the cost to $O(N)$. The cost of state update in 41 time frames also grows linear with state dimension, as well as the storage cost. Thus, we dramatically reduce the computational cost by updating covariance directly with $O(N^2)$ to $O(N)$.

Table 1. Computational and storage cost of fast KF

State Dimension	Initialize QH^T	Initialize QH^T with FMM	State update (41 frames)	Storage Cost
3245 (55x59)	26.37s	31.37s	1.91s	7.5MB
12753 (117x109)	400.9s	91.76s	5.84s	29.4MB
50778 (237x217)	6360.96s	280.07s	21.76s	117MB
250000 (500x500)	N/A	1157.6s	N/A	576MB

CONCLUSIONS

This paper addresses the problem of subsurface monitoring by formulating the problem using a linear state-space model. We showed that the Kalman Filter-based data assimilation tool can reconstruct dynamic objects recursively with good quality and computational efficiency. This is ideal for real-time monitoring applications in which data are collected at a rate faster than a discernible change in state. Fast KF with FMM significantly reduces the computational cost of the Kalman Filter update, from $O(m^2)$ to $O(m)$. Localized EnKF also reduces the computational cost to $O(m)$, yet introduces a bias into the estimation.

ACKNOWLEDGMENT

We want to acknowledge Stanford GCEP for funding this project. The synthetic TOUGH2 results were obtained through the courtesy of Dr. Jonathan Ajo-Franklin at LBNL and his colleagues. We want to thank their general support of this project.

REFERENCE

Aanonsen, S. I., G. Naevdal, D. S. Oliver, A. C. Reynolds & B. Valles (2009) The Ensemble Kalman Filter in Reservoir Engineering-a Review. *Spe Journal,* 14**,** 393-412.

Anderson, B. D. O. & J. B. Moore. 1979. *Optimal filtering*. Prentice-hall Englewood Cliffs, NJ.

Daley, T. M., J. B. Ajo-Franklin & C. Doughty (2011) Constraining the reservoir model of an injected CO2 plume with crosswell CASSM at the Frio-II brine pilot. *International Journal of Greenhouse Gas Control,* 5**,** 1022-1030.

Daley, T. M., R. D. Solbau, J. B. Ajo-Franklin & S. M. Benson (2007) Continuous active-source seismic monitoring of injection in a brine aquifer. *Geophysics,* 72**,** A57-A61.

Doucet, A., N. De Freitas & N. Gordon. 2001. *Sequential Monte Carlo methods in practice*. Springer Verlag.

Evensen, G. (2003) The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics,* 53**,** 343-367.

Houtekamer, P. L. & H. L. Mitchell (1998) Data Assimilation Using an Ensemble Kalman Filter Technique. *Monthly Weather Review,* 126**,** 796-811.

--- (2001) A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review,* 129**,** 123-137.

Kalman, R. E. (1960) A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME – Journal of Basic Engineering***,** 35-45.

Lumley, D. E. (2001) Time-lapse seismic reservoir monitoring. *Geophysics,* 66**,** 50-53.

Nenna, V., A. Pidlisecky & R. Knight (2011) Application of an extended Kalman filter approach to inversion of time-lapse electrical resistivity imaging data for monitoring recharge. *Water Resour. Res.,* 47**,** W10525.

Pruess, K., C. Oldenburg & G. Moridis. 1999. TOUGH2 user's guide, version 2.0. Report LBNL-43134, Lawrence Berkeley National Laboratory, Berkeley, Calif.