Ensemble Kalman Filter

J. Jaime Gómez-Hernández
Research Institute of Water and Environmental Engineering
Universitat Politècnica de València
Camino de Vera, s/n
46022 Valencia
Spain

March 29, 2021

Definition

The ensemble Kalman filter (EnKF) is an evolution of the Kalman filter for its application to non-linear state-transition systems with a further extension to serve as a powerful parameter inversion method. Its main purpose is to improve the estimates of the system state as observation are acquired. As the Kalman filter, the EnKF is based on two steps, forecasting and updating (or filtering). During the forecasting step, the state of the system is forecasted on the basis of the latest state estimate; then, forecasts are compared with observations and a correction is made to the state (and possibly to the parameters of the state equation) that will serve as the basis for the next forecast.

The Kalman filter was developed by Kalman (1960) for the purpose of improving the trajectory estimation of spacecrafts and missiles and one of its very first implementations was in the Apollo program. The Kalman filter, as developed by Kalman, had a great success when it was first proposed since it provides optimal results for dynamic systems that are controlled by a linear state-transition equation; then, it went into oblivion because it did not work so well when dealing with non-linear systems. New formulations such as the extended Kalman filter were postulated with limited success, but it was not until the work by Evensen (1994), who introduced the EnKF as a mean to deal with the non-linearities, that the Kalman filter came back into the spotlight. This rebirth of the Kalman filter not only allowed a better modeling of non-linear systems but also proved that it can be used as a very

efficient inverse modeling approach for the identification of parameters of the state-transition equations.

The Kalman filter

Consider a dynamic system evolving in time according to a state-transition equation. There are two ways to predict the state system at time t: (i) By dead reckoning, that is, the state-transition model is run until time t given some known state at time 0, or (ii) by observing the state of the system using measuring probes. Neither way will provide an exact description of the state of the system at time t, the first one, because the state-transition model is always an approximation of the real system, and the second one, because the probes always have measurement errors and, generally, only sample sparsely the state itself. Kalman found a way to combine the two predictions into a better one in his formulation of his filter.

Kalman formulated his filter for a linear dynamic system

$$x^t = \mathcal{L}\left(x^{t-1}\right) \tag{1}$$

where x is the state of the system, and $\mathcal{L}$ the linear state-transition equation that gives the state of the system at time t given its state in a previous instant $t-1$. Since most of the models used in the geosciences are built on discretized versions of the real system, a matrix notation can be used to describe the evolution of the system

$$\mathbf{x}^t = \mathbf{A} \cdot \mathbf{x}^{t-1} \tag{2}$$

where $\mathbf{x}$ is a column vector of size $N \times 1$ with the values of the state at the centers of the N voxels that discretize the study area, and $\mathbf{A}$ is a state-transition matrix of size $N \times N$.

Adopting a random function model, we can interpret each element of $\mathbf{x}$, at any time t, as a random variable, and the entire vector, at any time t, as a random function. As a consequence, the expected value of the random function propagates in time according to

$$E\left\{\mathbf{x}^t\right\} = \mathbf{A} \cdot E\left\{\mathbf{x}^{t-1}\right\}, \tag{3}$$

and its covariance, according to

$$C\left(\mathbf{x}^t\right) = \mathbf{A} \cdot C\left(\mathbf{x}^{t-1}\right) \cdot \mathbf{A}', \tag{4}$$

where C is a covariance matrix of size $N \times N$ containing all pairwise covariances between any two elements in the vector state and the $'$ symbol is used to denote transpose.

Now, consider that the state has been observed at M locations (generally $M \ll N$). These observations are represented by vector $\mathbf{y}$. For the sake of simplicity, assume that these observations have been taken at the centers of the voxels that discretize the study area and that measure directly the state of the system not a proxy. This simplification implies that the observation matrix $\mathbf{H}$ that will be used later should be made up of just zeroes and ones and it serves to extract the specific rows and columns from the vectors and matrices used in the formulation that follows corresponding to the observation locations. Consider also that these observations have uncorrelated measurement errors with zero mean; the error covariance $\mathbf{R}$ is, therefore, diagonal. Under these considerations, the conditional expectation of the state at time t given the observations taken at that same time is the optimal estimate in a least-square sense of the state of the system and is given by the following expression:

$$E\left\{\mathbf{x}^t \mid \mathbf{y}^t\right\} = E\left\{\mathbf{x}^t\right\} + \mathbf{K} \cdot \left(\mathbf{y}^t - \mathbf{H} \cdot E\left\{\mathbf{x}^t\right\}\right) \tag{5}$$

where $\mathbf{K}$ is the so-called Kalman gain matrix of size $N \times M$, the expression of which is

$$\mathbf{K} = C(\mathbf{x}^t) \cdot \mathbf{H}' \cdot \left(\mathbf{H} \cdot C(\mathbf{x}^t) \cdot \mathbf{H}' + \mathbf{R}^t\right)^{-1}. \tag{6}$$

The analysis of expression (5) shows that the conditional expectation is equal to the unconditional one corrected by a factor that is proportional to the discrepancy between the observations and the predicted expected data at time t obtained with Eq. (3). The proportionality term is itself proportional to the predicted state covariance at time t, and inverse proportional to the measurement error covariance and to the covariance between the locations at which the state has been observed.

The conditional covariance is given by

$$C\left(\mathbf{x}^t \mid \mathbf{y}^t\right) = (\mathbf{I} - \mathbf{K} \cdot \mathbf{H}) \cdot C\left(\mathbf{x}^t\right); \tag{7}$$

it is the covariance associated with the conditional expectation.

The Kalman filter consists of two steps. A forecast step, given by Eq. (3) and an update step given by Eq. (5). Alongside, the covariances are also forecasted and updated (Eqs. (4) and (7), respectively). The predictions are based on the linear state transition equation and the updates are the result of an optimization in some least squares sense. Once the updates are made, they become the basis for the next forecast and update from t to $t+1$.

As described, the method is straightforward to implement and apply. One of its first applications was to help the guidance of the Apollo spacecrafts and it was coded in the Apollo computer, a computer with only 2 kilobytes of RAM, 36 kilobytes of ROM, and a chip running at 100 MHz.

The Extended Kalman Filter

Given its success with linear dynamic systems, an attempt to expand the Kalman filter to deal with non-linear state transition equations was done with the Extended Kalman filter (Jazwinski, 1970). When the system does not evolve linearly

$$x^t = \Phi\left(x^{t-1}\right) \tag{8}$$

with Φ being a non-linear operator, a Taylor expansion can be applied to get a linear approximation

$$\mathbf{x}^t \approx \mathbf{A} \cdot \mathbf{x}^{t-1} + \dots \tag{9}$$

and

$$C\left(\mathbf{x}^t\right) \approx \mathbf{A} \cdot C\left(\mathbf{x}^{t-1}\right) \cdot \mathbf{A}' + \dots, \tag{10}$$

with

$$\mathbf{A} = \frac{\partial \Phi}{\partial \mathbf{x}^t} \tag{11}$$

being the Jacobian of the state transition equation. When only the first term of the expansion is retained, the formulation defaults to the Kalman filter formulation, which is then applied. The main problem with this approach is that, even if the Jacobian is re-evaluated after each update step, the covariance estimate deteriorates quickly in time when Eq. (10) is used, yielding very poor results.

The Ensemble Kalman Filter

An improved method to compute the state covariances for non-linear systems was proposed by Evensen (1994) circumventing the main problem of the extended Kalman filter. A random function can be defined by the statistics of the random variables that conform it, or as a collection of realizations. Evensen replaced working directly with the expected value and covariance of the random function by working with an ensemble of realizations, which are forecasted and updated as per Kalman filter theory. Then, he proposes

to compute, experimentally, the expected value and covariance from the forecasted realizations. There is no need to approximate the forecast of neither the mean nor the covariance as in Eqs. (9) and (10); rather, the ensemble of realizations are forecasted in time using the non-linear transfer function of Eq. (8) and the expected values and covariances are computed afterwards from the set of forecasted realizations.

Consider an ensemble of K realizations from the random function $\mathbf{x}$, at time $t-1$, $\left\{\mathbf{x}_1^{t-1}, \mathbf{x}_2^{t-1}, \ldots, \mathbf{x}_K^{t-1}\right\}$, each realization composed of N components corresponding to the N state values associated to the discretization of the study area. They propagate in time according to the non-linear transfer function Φ:

$$\left\{\mathbf{x}_1^t, \mathbf{x}_2^t, \ldots, \mathbf{x}_K^t\right\} = \Phi\left\{\mathbf{x}_1^{t-1}, \mathbf{x}_2^{t-1}, \ldots, \mathbf{x}_N^{t-1}\right\}. \tag{12}$$

The expected value and the covariance of the state random function can be approximated from this ensemble of forecasted realizations. Let $\mathbf{X}$ be an array with N rows and K columns containing all states for all realizations. The expected value is given by

$$E\left\{\mathbf{x}^t\right\} \approx \frac{1}{K}(\mathbf{X}^t \cdot 1_{K\times 1}) \tag{13}$$

where $1_{K\times 1}$ is a column vector of K ones, and $\mathbf{X}^t$ contains all the forecasted realizations at time t. Likewise, the covariance is computed as

$$C\left(\mathbf{x}^t\right) \approx \frac{1}{K-1}\left(\mathbf{X}^t - E\left\{\mathbf{x}^t\right\} \cdot 1_{1\times N}\right) \cdot \left(\mathbf{X}^t - E\left\{\mathbf{x}^t\right\} \cdot 1_{1\times N}\right)' \tag{14}$$

where $1_{1\times N}$ is a row vector of N ones. The Kalman gain is computed using Eq. (6) and all realizations are updated according to

$$\mathbf{X}^t|\mathbf{y}^t = \mathbf{X}^t + \mathbf{K} \cdot \left((\mathbf{y}^t + \epsilon^t) \cdot 1_{1\times K} - \mathbf{H} \cdot \mathbf{X}^t\right) \tag{15}$$

where ϵ^t is a vector of random observation errors drawn from the diagonal covariance matrix $\mathbf{R}$. The conditional expectation and the conditional covariance could be computed, if needed, using $\mathbf{X}^t|\mathbf{y}^t$ in Eqs. (13) and (14) in replacement of $\mathbf{X}^t$. The conditional expectation would the best estimate, in some least squares sense, of the state of the system at time t, and the conditional covariance a measure of the estimation uncertainty.

The ensemble of updated realizations becomes the new starting ensemble of realizations and the forecast and update steps are repeated starting at Eq. (12).

The algorithm for the ensemble Kalman filter is quite simple and applicable to any dynamic system; its steps would be:

1. Generate an ensemble of realizations of the system state (these realizations can be generated by randomizing the parameters of the state equation and solving it, or by perturbing with noise a given solution of the state equation),

2. Forecast the state, realization by realization, until the next observation time (Eq. (8)) ,

3. Compute the expected value and covariance of the forecasted realizations (Eqs. (9) and (10)),

4. Compute de Kalman gain (Eq. (6)),

5. Update the forecasted realizations (Eq. (15)),

6. Computed the expected value and covariance of the updated realizations, if necessary (Eqs. (9) and (10)),

7. Make the updated realizations the new starting realizations and return to the forecast step (Eq. (8)).

The Ensemble Kalman Filter With Augmented State

The EnKF as proposed by Evensen solved the problem of the approximate calculation of the covariance matrices using an alternative (and better) approximation based on the statistical analysis of an ensemble of realizations, but its application had the same purpose as the original filter: the update of the system state each time new observations were obtained. But, it was realized that the method could be used to update the parameters defining the system, too. In modeling, for instance, an aquifer, these parameters could be material parameters such as permeability or porosity, forcing terms such as recharge or contaminant source location, or boundary conditions such as lateral inflows/outflows or leakages. The idea consists of considering the parameters as part of an augmented state vector (Wen and Chen, 2007), in which the parameters are updated each time a new observations are made.

The system state x is now augmented to include the last update of the model parameters α. The system state continues evolving in time according to a state-transition model Φ, but the model parameters do not evolve.

Forecast equation (3) becomes

$$\left\{ \begin{array}{c} \alpha^t \\ x^t \end{array} \right\} = \left\{ \begin{array}{c} \alpha^{t-1} \\ \Phi\left(x^{t-1}\right) \end{array} \right\}. \tag{16}$$

Using z to represent the augmented state, a new forecast equation is defined

$$z^t = \Psi\left(z^{t-1}\right), \tag{17}$$

and the EnKF algorithm can be applied once a random function model is adopted for z. Each ensemble realizations will contain a realization of the parameters and of the states associated to that parameter realization. The forecast will be performed using Eq. (17) on each realization, and the rest of the steps will follow the algorithm described in the previous section. Notice that new parameter observations can be included at the same time as new state observations.

This application of the EnKF becomes a powerful inverse modeling approach since, as observations are assimilated, improved estimates of the model parameters are obtained. The main criticism of this approach is the potential lack of consistency between the updated parameters and the updated state values. After all, the states should abide the principles under which the state-transition model is built, for instance, in the case of an aquifer modeling, the preservation of mass. This possible inconsistency resulted in the proposal of a variant of the filter, the restart EnKF (Xu and Gómez-Hernández, 2018), in which the system forecast is always made from time 0, rather than from the last updated estimate of the system state:

$$\left\{ \begin{array}{c} \alpha^t \\ x^t \end{array} \right\} = \left\{ \begin{array}{c} \alpha^{t-1} \\ \Phi\left(x^{0}\right) \end{array} \right\}. \tag{18}$$

The parameters keep being updated after each observation event, but the states are always recalculated from the state at time 0, x^0, with the latest update of the parameters. This procedure, obviously increase the running time of the algorithm, but ensures that the state values are always consistent with the latest estimate of the model parameters.

The Normal-Score Ensemble Kalman Filter

There was still a pitfall in the application of the EnKF to non-linear systems. The fact that all updates are made using covariances imbue a Gaussian flavor to the updated realizations, which, after several updates, have clear

Gaussian attributes, starting by their histogram. In geosciences, many of the attributes have non-Gaussian characteristics, like, for instance, the permeabilities of a sand-shale aquifer have a histogram closer to a mixture of two lognormal distributions than to a Gaussian one. Even if the starting parameter realizations display the desired non-Gaussian characteristics, such as a bi-modality of the histograms, the final histograms of the parameter realizations get close to normal. To avoid this problem, the normal-score EnKF (NS-EnKF) was proposed by Zhou et al. (2011).

In the NS-EnKF, all the EnKF steps are performed in variables that are marginally-Gaussian. The augmented state vector z is transformed into another one with marginal Gaussian distributions using a normal-score transform $\phi(\cdot)$, which is a monotonic transformation that will yield a resulting variable with a standardized Gaussian histogram, then, given the normal-score transforms u^t and u^{t-1} of the augmented states z^t and z^{t-1}

$$\begin{aligned} u^{t-1} &= \phi^{t-1}\left(z^{t-1}\right) \\ u^{t} &= \phi^{t}\left(z^{t}\right) \end{aligned}$$

and given the state transition equation (17), a new forecast equation can be written for the normal-score variates

$$u^t = \Gamma\left(u^{t-1}\right), \tag{19}$$

with

$$\Gamma = \phi^t \cdot \Psi \cdot \left(\phi^{t-1}\right)^{-1}. \tag{20}$$

The NS-EnKF proceeds as the EnKF but in normal-score space. At any time, the augmented state values z can be obtained by using the inverse of the normal-score transform $\phi^{-1}(\cdot)$.

Kalman Filter and Kriging

The observant reader will find streaking similarities between the update equation (5) and the simple kriging equation, with the Kalman gain being the kriging weights.

Summary and Conclusions

The ensemble Kalman filter is a variant of the Kalman filter capable to deal with dynamic non-linear systems. It provides optimal estimates of

the state of the system as well as of the parameters defining the state-transition model, which are updated sequentially in time as new observations are acquired. The rationale is simple as is its implementation, and could be used with virtually any model that allows predicting the state of the system in the future as a function of the state in the present. Its version using an augmented state vector has proven a very powerful and efficient inverse modeling approach.

Cross-References

Forward and Inverse Models
Inversion Theory
Inversion Theory in Geosciences
Kriging
Monte-Carlo Method
Multivariate Analysis
Random function
Random Variable
Realizations
Simulation
Spatial Statistics

References

Evensen G (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research 99(C5):10143–10162

Jazwinski AH (1970) Stochastic Processes and Filtering Theory. Academic Press

Kalman RE (1960) A new approach to linear filtering and prediction problems. Journal of Basic Engineering 82:35–45

Wen XH, Chen WH (2007) Some practical issues on real-time reservoir model updating using ensemble kalman filter. SPE Journal 12(02):156–166

Xu T, Gómez-Hernández JJ (2018) Simultaneous identification of a contaminant source and hydraulic conductivity via the restart normal-score ensemble kalman filter. Advances in Water Resources 112:106–123

Zhou H, Gómez-Hernández J, Hendricks Franssen H, Li L (2011) An approach to handling non-gaussianity of parameters and state variables in ensemble kalman filtering. Advances in Water Resources 34(7):844–864