Error Models for Reservoir Simulation

Alannah O'Sullivan and Mike Christie

Corresponding author: Mike Christie
Institute of Petroleum Engineering
Heriot-Watt University
Edinburgh EH14 4AS
Phone: +44 131 451 8063
Fax: +44 131 451 3127
Email: mike.christie@pet.hw.ac.uk

The general problem of accurate parameter estimation is common through many scientific disciplines, weather forecasting, contaminant transport etc. The accuracy of the estimated parameters determines the accuracy of a forecast. Calibration of results is key in producing confident forecasts. This study examines the role of a simulation error model in determining accurate parameter estimation with regard to model error within the oil industry.

Bayesian analysis is used in the history matching process to home in on accurate parameter estimates for use in the forward prediction process. This process is performed by comparing observed data with data simulated over a range of possible parameter values. Assuming Gaussian distributed errors, the likelihood is determined by the exponential of the misfit function. Standard practice in the oil industry is to use a least squares definition for the misfit function, which takes into account uncorrelated data errors only, making it inappropriate for time dependent simulation errors. Also, the standard least squares approach can result in biased estimates for parameters.

An important factor to consider when using the Bayesian framework for parameter estimation is the degree of detail to use. An accurate forecast requires accurate parameter estimation from using a fine grid. However the detail in a fine grid model causes the time taken to solve the system to be large and impractical. The advantage to using a coarse grid over a fine grid lies in speed, and the disadvantage is that coarse grid models can produce biased estimates. The time taken to run a model becomes an increasingly important issue as the dimension space for the parameter becomes large. In addition to dimensionality, the number of realisations simulated plays a role in increasing the total time taken to produce an estimate. As the exact permeability data is unknown, a set of realisations is used rather than a single realisation, gaining an average result with an associated degree of uncertainty. On a workstation, this can turn run time for a large model for a small number of simulations from of the order hours to days.

The motivation for using an error model in estimating parameter values is to combine the accuracy of a fine grid solution with the speed of a coarse grid solution. The basic framework of the error model requires both fine grid and coarse grid data. However the amount of fine grid data used in error model construction is small in comparison with using a fine grid model alone for parameter estimation. Starting with a broad range for a possible parameter space, a small number of base points are chosen within the range. A number of realisations are used to solve the system on a fine grid for the chosen base points. Similarly, the same is done on a coarse scale. Statistical analysis of the difference between results from both scales forms the basic structure of the error model. The difference between the two solutions is the error gained in using a coarse model assuming the fine grid is truth. Interpolation is used to estimate the statistics of the differences at intermediate points within the range of parameter space.

In practice the error model improves performance of parameter estimation when incorporated into the misfit function. Including the mean error in the misfit calculation is shown to drastically improve prediction performance than by using the difference between fine and coarse models alone. Covariance also adds value to the estimation by giving a realistic spread about the mean result.

A 2D example problem is used to demonstrate how an error model works. Injection of a more mobile gas into an oil reservoir displaces the more viscous oil. The interface between the two fluids is unstable and fingers of gas channel through the oil, causing early breakthrough of the injected gas, thus reducing recovery. This problem is a fluid dynamic instability analogous to the Rayleigh Taylor instability.

For this example, we chose 3 base viscosities (5,10,15) to use in the fine and coarse grid simulations. For this example problem, a simple one parameter model, known as Todd & Longstaff, has been used in lieu of an actual coarse grid model to increase speed-up. We chose 20 realisations for the fine grid model simulation and linear interpolation to estimate the intermediate points.

For the fine grid solution, a numerical simulator was run for each of the 20 realisations, producing a range of possible solutions to the problem for each of the three base viscosities. From the range of results, the mean and covariance of concentration is calculated. For the coarse solution, as mentioned earlier the one parameter Todd & Longstaff model will be used. This model is an empirical model used in the oil industry to describe the mean behaviour of unstable fluid flow. It is analogous to two-phase models used to describe the effective behaviour of the Rayleigh Taylor instability. Empirical models of this sort are frequently used in large scale calculations since the computational cost of full resolution in a 3D model is still prohibitive. The averaged difference between the fine grid and coarse grid results, with the variance of this difference, forms the error model. A more detailed error model is determined by interpolation.

The standard least squares approach to the misfit for this problem will be compared with the error model incorporated into the misfit definition. This means the misfit function is extended to include the average error for a given viscosity when comparing fine and coarse solutions. It will also use the covariance matrix rather than a single (time independent) value for variance. Both definitions will be compared with a third definition of misfit, which has a characteristic of each of the two previous definitions. This third misfit will use the mean error from the error model in its calculation, but will use the variance from the least squares definition rather than the full covariance matrix. In this way the benefits of both mean error and time dependent variance can be examined independently. A calibration plot will be produced, showing that using the full error model produces accurate, non-biased results.