Short Note on Dynamic Data Integration in Characterizing Permeability and Flow Fields

Amir H. Hosseini and Clayton V. Deutsch

Integration of data from different data sources is very common when characterizing permeability and flow fields in hydrogeology and reservoir engineering applications. In this short note, incremental value of additional static and dynamic data in understanding the permeability field and pressure response is investigated. In a steady-state mode, unconditional and conditional permeability realizations are flow simulated and compared to the reference permeability field and associated flow field. Conditioning of the permeability realizations to the pressure head data is performed by Sequential Self-Calibration (SSC) approach and a post-processing step. The results of this study show that (1) as expected, conditioning to both static and dynamic data improves the estimation of reference permeability and flow fields; (2) the first few pressure measurements that are incorporated into the model significantly reduce the uncertainty in the permeability field; and (3) a post-processing can be quite useful to remove unrepresentative realizations after conditioning to dynamic data. This short note is presented for a groundwater application. However, all the conclusions equally apply to the case of single phase flow in reservoir engineering applications.

Synthetic aquifer and multiple scenarios

A two-dimensional 1000 m × 700 m synthetic aquifer with no-flow boundary conditions on the west and east boundaries and constant pressure boundary conditions on the north and south boundaries is considered (Figure 1). The constant pressure head boundary conditions have been set to 5 m and 2 m at north and south boundaries, respectively. Incremental addition of information is investigated in the context of sixteen scenarios. An anisotropic reference permeability field with a Gaussian distribution is considered. The reference permeability and the associated pressure head response are shown in Figure 2. Permeability and head values were sampled from the reference field at 15 locations in total. Table 1 shows the number of the sampled data and their data types that have been used in different scenarios. Each set of three wells (numbered uniquely in Figure 1) are added to the data set from one scenario to the next. Depending on the scenario considered, the initial permeability fields are conditioned to the static data (permeability measurements) by Sequential Gaussian Simulation (Deutsch and Journel 1998) and/or conditioned to the dynamic data (pressure head measurements) by Sequential Self-Calibration (Gomez-Hernandez et al. 1997).

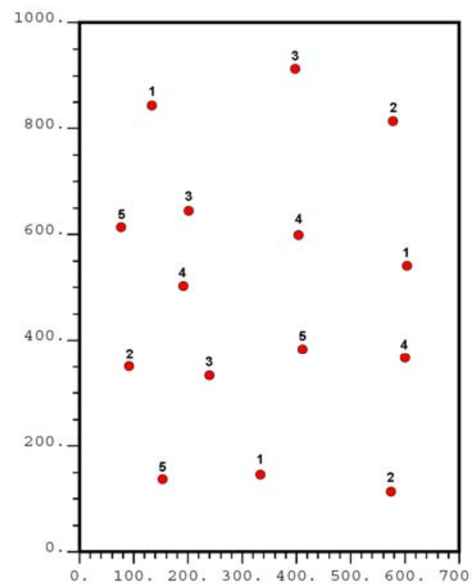

Figure-1: Modeling domain and the location map of the wells with sampled data. Each set of three wells (numbered uniquely) are added to the data set from one scenario to the next.

Post-processing for SSC

Sequential Self-Calibration approach uses a gradient based optimization algorithm and may well be prone to get stuck in local minima, particularly when there is a decent number of conditioning dynamic data. To investigate the significance of this problem, 100 realizations are conditioned to all permeability and head data (scenario K+H_5), using SSC. In implementation of SSC, the maximum number of outer iterations, the minimum tolerance, the minimum difference of objective function in two consecutive iterations and the maximum number of times that the difference of objective function in two consecutive iterations is smaller than the specified value are set to 25, 0.05, 10^{-2} and 5, respectively.

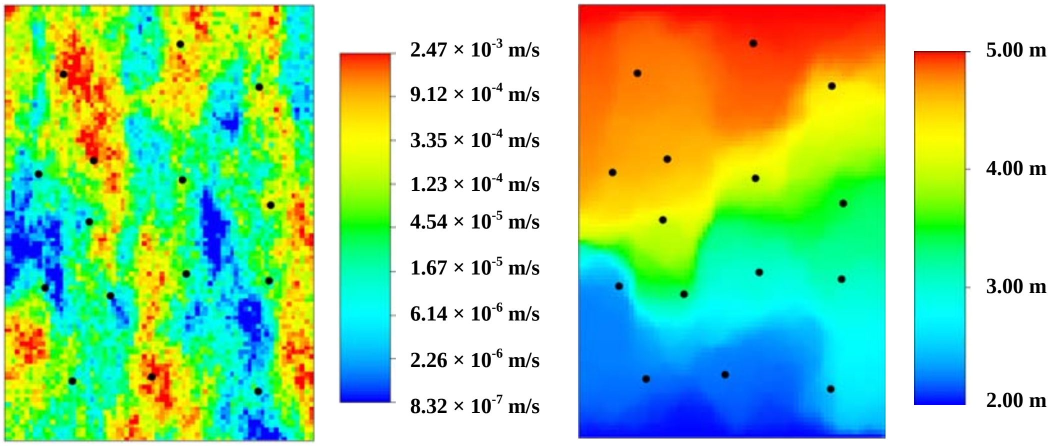

Figure-2: Reference permeability and reference head field and sampling points

Scenario	# of K data	# of H data	Scenario	# of K data	# of H data	Scenario	# of K data	# of H data
No Data	0	0	HEAD_1	0	3	K+H_1	3	3
PERM_1	3	0	HEAD_2	0	6	K+H_2	6	6
PERM_2	6	0	HEAD_3	0	9	K+H_3	9	9
PERM_3	9	0	HEAD_4	0	12	K+H_4	12	12
PERM_4	12	0	HEAD_5	0	15	K+H_5	15	15
PERM_5	15	0						

Table-1: Number of data and different data types for different scenarios

The weights used in construction of the objective function in the SSC algorithm are equal to the measurement error associated with pressure heads. The objective function is defined by:

$$O.F. = \sum_{i=1}^{N_{obs}} W_i \left(h_i^{obs} - h_i^{cal} \right)^2 \qquad [1]$$

where, h_i^{obs}, h_i^{cal} and W_i are observed and calculated pressure heads, and the associated optimization weights, respectively. N_{obs} is the number of observation locations. The weights can be written in terms of standard deviation of observation error σ_i, that is:

$$W_i = \frac{1}{\sigma_i} \qquad [2]$$

In this study, error variance associated with all observations σ_i is set equal to 0.1 m. When the objective function and the associated errors are defined by equations [1] and [2], 'calculated error variance' and 'standard error of the regression' are defined as s^2 and s, respectively and given by:

$$s^2 = \frac{O.F.}{N_{obs}} \qquad [3]$$

For non-linear regression to be successful, standard error of the regression, s, must be close to one (Hill and Tiedeman 2007). If it is considerably larger than one, it means that the regression can not reach to the optimal state. This occurs when either parameterization of the model is poor or the sampling data does not adequately support the state of the system being modeled or model is stuck in a local minimum. One way of checking this is by plotting the histogram of calculated error variance for different realizations. Ideally, we want this distribution to have a normal distribution with a mean of one and a small standard deviation.

The histograms in Figure 3 show the distribution of standard error for 100 realizations, before and after ranking based on the value of standard error of regression. In fact, histogram to the left shows the distribution of standard error for the first 100 realizations that are conditioned by SSC. The histogram to the right shows the distribution of the best 100 realizations out 400 realizations that were conditioned by SSC after ranking based on the value of standard error of regression.

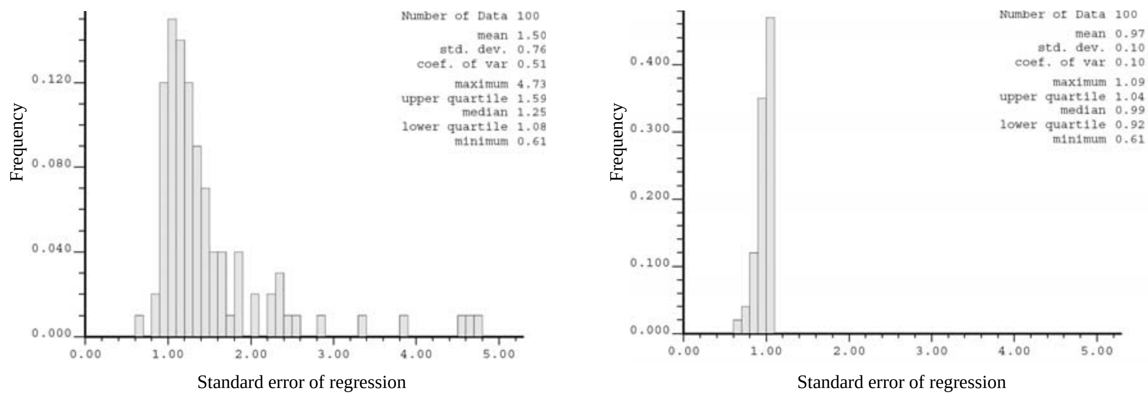

Figure-3: The distribution of standard error for the first 100 realizations conditioned by SSC (left), and the distribution of standard error for the best 100 out of 400 realizations conditioned by SSC (right).

This study shows that (1) not all realizations conditioned by the SSC fully honor the pressure head observations and their associated errors and (2) a post-processing by ranking can remove unrepresentative realizations.

Incremental value of additional data

The value of additional data is studied based on two evaluation criteria. The first criterion is the average absolute error (AAE) which is calculated for both ln-K and pressure head and is given by:

$$\mathrm{AAE}(\mathrm{X}) = \frac{1}{N}\sum_{i=1}^{N}\left|\overline{\mathrm{X}}_{\mathrm{SIM},i} - \overline{\mathrm{X}}_{\mathrm{REF},i}\right| \qquad [4]$$

Where, N is the number of grid cells and X represents either natural log permeability or steady – state pressure head. The other evaluation criterion is average ensemble standard deviation given by:

$$\mathrm{AESD}(\mathrm{X}) = \frac{1}{N}\sum_{i=1}^{N}\sigma_{X_i} \qquad [5]$$

where, σ_{X_i} is the ensemble standard deviation at a given node. Figures 4, 5, 6 show the calculated AAE for natural log permeability and hydraulic head, and AESD for natural log permeability, respectively. The results are presented for different scenarios indicated in Table 1. Also, the values of AAE and AESD are standardized so that each of AAE and AESD is equal to 100 for scenario with no conditioning data.

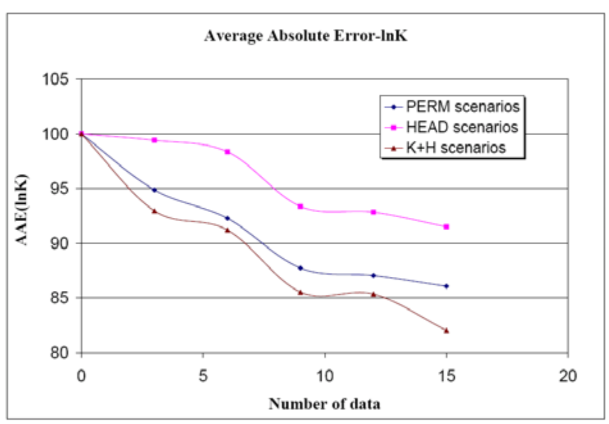

Figure-4: Reduction of AAE for ln-K with inclusion of more static, dynamic, and static+dynamic data.

The results show that as expected conditioning to both static and dynamic data improves the estimation of permeability and flow fields. The importance of conditioning to dynamic data in reducing the uncertainty in permeability field is clearly observed in Figure 6.

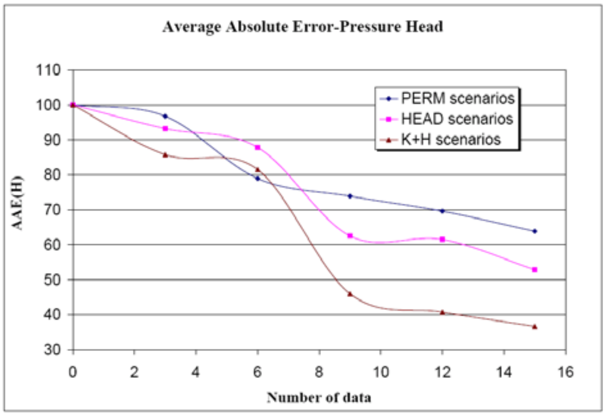

Figure-5: Reduction of AAE for pressure head with inclusion of more static, dynamic and static+dynamic data.

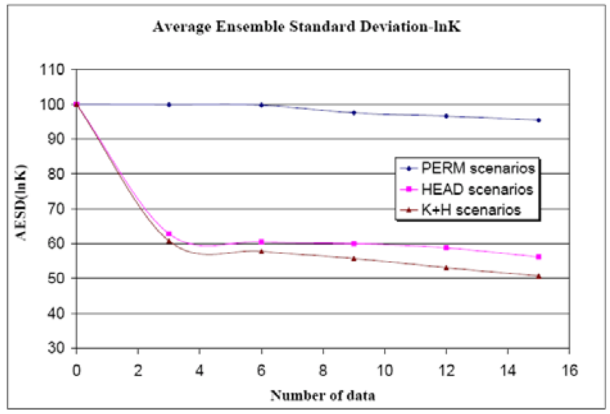

Figure-6: Reduction of AESD for ln-K with inclusion of more static, dynamic, and static+dynamic data.

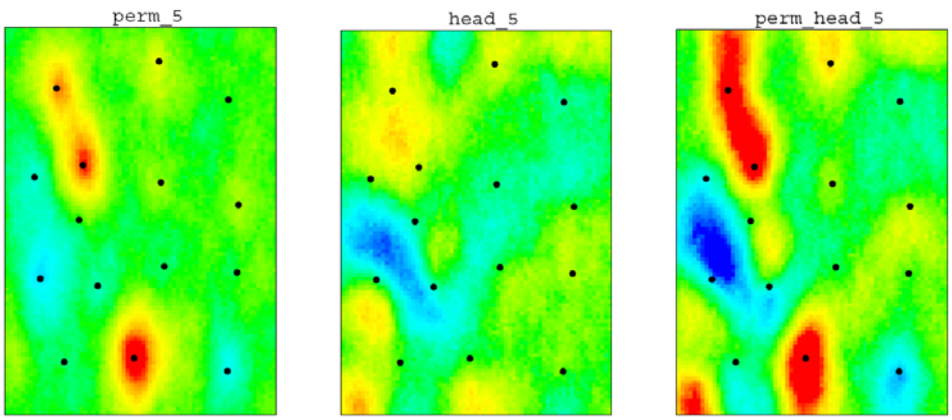

Figure-7: E – type mean for ensemble of realizations conditioned by static data only (left), dynamic data only (middle) and static and dynamic data (right). It is clearly shown that inclusion of pressure head data (dynamic data) significantly improves the prediction of large scale features and their connectivity. From left to right, these E-type maps correspond to scenarios PERM_5, HEAD_5 and K+H_5 in Table 1.

References

Deutsch, C.V. and Journel, A.G. 1998, *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, New York, 2nd Ed.

Gomez-Hernandez, J.J., Sahuquillo, A., Capilla, J.E. 1997, Stochastic Simulation of transmissivity fields conditioned to both transmissivity and piezometric data, 1.Theory. *Journal of Hydrology*, 203(1- 4): 163 – 174.

Hill, M.C. and Tiedeman, C.R. 2007, *Effective Groundwater Model Calibration*, John Wiley & sons, New Jersey, 1st Ed.