Estimation of parameter uncertainty using inverse model sensitivities

Velimir V. Vesselinov[a]

[a]Hydrology, Geochemistry, and Geology Group, Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Forward model sensitivities are commonly applied to evaluate the uncertainty in model parameter estimates obtained through inverse analysis. In this case, the forward sensitivity (Jacobian) matrix is applied to compute an approximate representation of the covariance matrix of inverse parameter estimates. However, this approach can produce biased estimates of the covariance matrix because it does not account accurately for correlations between uncertainty of calibration targets and estimates. Typically, these correlations are non-linear and depend on the spatial and temporal structure of inverse targets and estimated parameters. A better but much more computationally intensive method to measure parameter uncertainty, which we call the inverse-sensitivity approach, directly evaluates the sensitivity of inverse estimates of model parameters with respect to the calibration targets. Further, we can evaluate the sensitivity of model predictions based on inverse model parameter estimates with respect to the calibration targets. The proposed methodology can also be applied to problems such as estimation of predictive uncertainty, optimization of data collection strategies, and design of monitoring networks. Its implementation can be performed efficiently through parallelization. Results based on a simple groundwater flow inverse problem are presented to illustrate the basis for the method.

1. INTRODUCTION

Inverse models are widely used in the field of hydrogeology [2–6,10]. One of the most important aspects in the inverse analysis is the evaluation of uncertainty in the estimated parameters. The commonly-used evaluation techniques are obtained from the existing vast body of parameter estimation literature [1] and are generally applicable when the number of calibration targets (observations) are significantly larger than the number of model parameters. However, we frequently deal with inverse solutions of groundwater flow models for which the number of calibration targets is slightly larger than the number of model parameters. It can be argued that these problems are also ultimately ill-posed, i.e., there is no unique inverse solution, and therefore it is very important to accurately assess the uncertainty in the model-predicted estimates. Further, the relationship between the spatial structure (locations) of calibration targets and the spatial structure (parameterization) of estimated model parameters can cause correlations between observation and estimation errors that might be very important to consider. In this paper, we analyze analytically simple test cases and compare the sensitivities and estimation uncertainties of model parameters using a traditional technique and an alternative method described below.

2. METHODOLOGY

Let us define a forward operator $\mathcal{F}$, which is a functional that maps a given set of model parameters $\mathbf{p}$ onto a set of model-predicted observations $\hat{\mathbf{o}}$:

$$\hat{\mathbf{o}} = \mathcal{F}(\mathbf{p}) \tag{1}$$

The corresponding inverse problem can be defined formally as solving (1) for $\hat{\mathbf{p}}$ given a set of observations (calibration targets) $\mathbf{o}$:

$$\hat{\mathbf{p}} = \mathcal{F}^{-1}(\mathbf{o}) \tag{2}$$

where $\mathcal{F}^{-1}$ is an inverse operator. There are various methods for solving this inverse problem [6]. The covariance matrix of estimation errors of model parameters $\mathbf{C}_{pF}$ are commonly computed using the following approximate expression [1]:

$$\mathbf{C}_{pF} = \left[\mathbf{J}_F^T \mathbf{C}_o^{-1} \mathbf{J}_F\right]^{-1} \tag{3}$$

where $\mathbf{J}_F$ is a sensitivity (Jacobian) matrix of forward model-predicted observations $\hat{\mathbf{o}}$ with respect to model parameters $\mathbf{p}$ ($\mathbf{J}_F = \partial\hat{\mathbf{o}}/\partial\mathbf{p}$), and $\mathbf{C}_o$ is a covariance matrix of observation errors. The expression in (3) is obtained by applying generalization of the Cramér-Rao inequality to the multivariate case [1], and, as a result, $\mathbf{C}_{pF}$ estimate is defining 'a lower bound' for the actual covariance matrix of estimation errors (i.e., the actual estimation-error variances and covariances are larger than or equal to the $\mathbf{C}_{pF}$ estimates). The derivation of (3) is also based on first-order error analysis [1].

An alternative approach to computing the estimation errors that is theoretically more accurate can be derived by considering the inverse model (2) as a "forward" model mapping $\mathbf{o}$ onto $\hat{\mathbf{p}}$. In this case, we can formally estimate the parameter uncertainties approximated up to the first order by using the definition of a covariance matrix [1]:

$$\begin{aligned}
\mathbf{C}_{pI} &= E\left[[\hat{\mathbf{p}}(\mathbf{o}) - E[\hat{\mathbf{p}}(\mathbf{o})]]\,[\hat{\mathbf{p}}(\mathbf{o}) - E[\hat{\mathbf{p}}(\mathbf{o})]]^T\right] \\
&= E\left[[\hat{\mathbf{p}}(\mathbf{o}) - \hat{\mathbf{p}}(\tilde{\mathbf{o}})]\,[\hat{\mathbf{p}}(\mathbf{o}) - \hat{\mathbf{p}}(\tilde{\mathbf{o}})]^T\right] \\
&= E\left[[\hat{\mathbf{p}}(\tilde{\mathbf{o}}) + \mathbf{J}_I[\mathbf{o} - \tilde{\mathbf{o}}] - \hat{\mathbf{p}}(\tilde{\mathbf{o}})]\,[\hat{\mathbf{p}}(\tilde{\mathbf{o}}) + \mathbf{J}_I[\mathbf{o} - \tilde{\mathbf{o}}] - \hat{\mathbf{p}}(\tilde{\mathbf{o}})]^T\right] \\
&= E\left[[\mathbf{J}_I[\mathbf{o} - \tilde{\mathbf{o}}]]\,[\mathbf{J}_I[\mathbf{o} - \tilde{\mathbf{o}}]]^T\right] \\
&= \mathbf{J}_I E\left[[\mathbf{o} - \tilde{\mathbf{o}}]\,[\mathbf{o} - \tilde{\mathbf{o}}]^T\right] \mathbf{J}_I^T \\
&= \mathbf{J}_I \mathbf{C}_o \mathbf{J}_I^T
\end{aligned} \tag{4}$$

where $\hat{\mathbf{p}}(\mathbf{o})$ is the set of inverse-model-predicted parameters given a set of observations (calibration targets) $\mathbf{o}$, $\tilde{\mathbf{o}}$ is the "expected value" for the observations $\mathbf{o}$ (i.e., the actually observed values), $\mathbf{J}_I$ is the sensitivity matrix of the inverse model representing the partial derivatives of parameter estimates $\widehat{\mathbf{p}}$ with respect to calibration targets $\mathbf{o}$ ($\mathbf{J}_I = \partial\widehat{\mathbf{p}}/\partial\mathbf{o}$). In (4), we make an assumption that the inverse-model-predicted estimates given $\widetilde{\mathbf{o}}$, $\widehat{\mathbf{p}}(\widetilde{\mathbf{o}})$, represent the "expected value" of $\widehat{\mathbf{p}}(\mathbf{o})$ (i.e., $\widehat{\mathbf{p}}(\widetilde{\mathbf{o}}) = E\left[\widehat{\mathbf{p}}(\mathbf{o})\right]$).

Note the difference in the way $\mathbf{J}_F$ and $\mathbf{J}_I$ are computed: the forward-sensitivity matrix $\mathbf{J}_F$ represents how the changes in model parameters impact observations predicted by the forward model; the inverse-sensitivity matrix $\mathbf{J}_I$ represents how the changes in calibration targets impact model parameters predicted by the inverse model. If the partial derivatives in $\mathbf{J}_I$ cannot be computed analytically, the numerical computation of $\mathbf{J}_I$ might require solving of multiple inverse problems for different sets of calibration targets.

We will define the two methods of computing the covariance matrix of estimation errors outlined in (3) and (4) as forward- and inverse-sensitivity approaches, respectively. The expressions obtained through both approaches are approximate since they are based on first-order analyses. However, there are important differences. In (3), the first-order approximation is applied to represent the dependency of model-predicted observations on model parameters ($\mathbf{J}_F$). In (4), the first-order approximation is applied to represent the dependency of inverse-model-predicted parameters on calibration targets ($\mathbf{J}_I$). How appropriate these approximations for both approaches are depends on the mathematical properties of the respective forward and inverse problems (1 and 2). However, since we are interested in the effect of observation errors on the parameter-estimation errors, the inverse-sensitivity approach is mathematically more suitable for this purpose. Even if the two first-order approximations are appropriate (linear models) or produce similar impacts on the respective covariance matrix estimates, the inverse-sensitivity approach can be expected to be superior to the forward-sensitivity approach because the $\mathbf{C}_{pF}$ values might be theoretically smaller than the actual error variances, as discussed above.

The differences between the two covariance matrices of estimation errors (3 and 4) will be further analyzed below for a simple groundwater flow system. We will also discuss the differences between the forward- and inverse-model sensitivity matrices and their implications.

3. SIMPLE 1D EXAMPLE

Let us consider a simple one-dimensional groundwater flow system (Figure 1). There are two zones with permeabilities k_1 and k_2 $[L/T]$. The constant velocity of groundwater flow passing through the system is q_f $[L/T]$, and the heads (pressures) are observed at four locations along the flow direction, h_f, h_1, h_2 and h_3 $[L]$; the observations are evenly distributed with separation distance l $[L]$. To solve the forward problem (1), we can use values of k_1, k_2, h_f, and q_f, and Darcy's law to compute estimates for $\widehat{h_1}$, $\widehat{h_2}$, and $\widehat{h_3}$:

$$
\begin{aligned}
\widehat{h_1} &= h_f + \frac{q_f l}{k_1} \\
\widehat{h_2} &= h_f + \frac{q_f l}{k_1} + \frac{q_f l}{k_2} \\
\widehat{h_3} &= h_f + \frac{q_f l}{k_1} + \frac{2 q_f l}{k_2}
\end{aligned} \tag{5}
$$

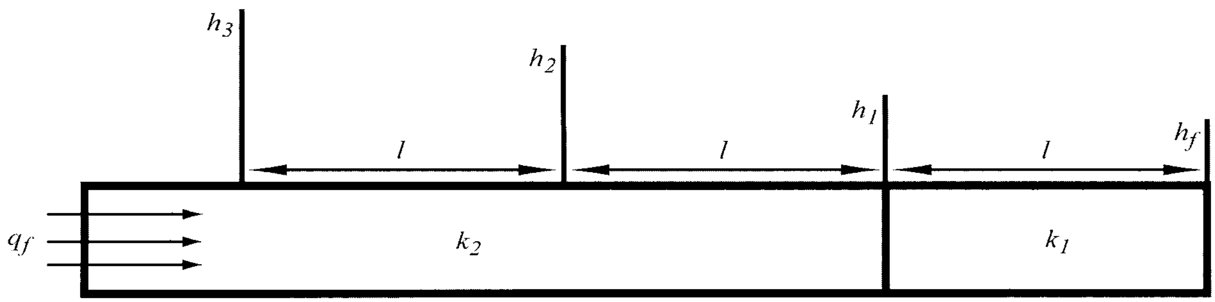

Figure 1. Schematic representation of a simple one-dimensional groundwater flow system.

Alternatively, we can solve the inverse problem (2) and estimate $\widehat{k_1}$ and $\widehat{k_2}$ based on our knowledge about q_f, h_f, h_1, h_2, and h_3:

$$\begin{aligned} \widehat{k_1} &= \frac{q_f l}{h_1 - h_f} \\ \widehat{k_2} &= \frac{q_f l}{2}\left[\frac{1}{h_2 - h_1} + \frac{1}{h_3 - h_2}\right] \end{aligned} \tag{6}$$

The values for q_f and h_f are assumed to be perfectly known, but h_1, h_2, and h_3 are considered uncertain with variances of observation errors equal to σ_{h1}^2, σ_{h2}^2, and σ_{h3}^2, respectively; further, the observation errors are considered uncorrelated, causing the matrix $\mathbf{C}_o$ to have a diagonal form. Based on the forward model equations (6), we compute the sensitivity matrix, $\mathbf{J}_F$, representing the partial derivatives of model-predicted observations ($\widehat{h_1}$, $\widehat{h_2}$ and $\widehat{h_3}$) with respect to model parameters (k_1 and k_2) as follows:

$$\mathbf{J}_F = q_f l \begin{bmatrix} -\frac{1}{k_1^2} & 0 \\ -\frac{1}{k_1^2} & -\frac{1}{k_2^2} \\ -\frac{1}{k_1^2} & -\frac{2}{k_2^2} \end{bmatrix} \tag{7}$$

The covariance matrix of estimation errors is then defined using (3):

$$\mathbf{C}_{pF} = \left[\mathbf{J}_F^T \mathbf{C}_o^{-1} \mathbf{J}_F\right]^{-1} = \frac{1}{q_f^2 l^2} \begin{bmatrix} \frac{\sigma_{h1}^2(4\sigma_{h2}^2+\sigma_{h3}^2)}{(\sigma_{h1}^2+4\sigma_{h2}^2+\sigma_{h3}^2)} k_1^4 & -\frac{\sigma_{h1}^2(2\sigma_{h2}^2+\sigma_{h3}^2)}{(\sigma_{h1}^2+4\sigma_{h2}^2+\sigma_{h3}^2)} k_1^2 k_2^2 \\ -\frac{\sigma_{h1}^2(2\sigma_{h2}^2+\sigma_{h3}^2)}{(\sigma_{h1}^2+4\sigma_{h2}^2+\sigma_{h3}^2)} k_1^2 k_2^2 & \frac{\sigma_{h2}^2\sigma_{h3}^2+\sigma_{h1}^2(4\sigma_{h2}^2+\sigma_{h3}^2)}{(\sigma_{h1}^2+4\sigma_{h2}^2+\sigma_{h3}^2)} k_2^4 \end{bmatrix} \tag{8}$$

Alternatively, for the inverse sensitivity approach, we compute the sensitivity matrix, $\mathbf{J}_I$, of model parameters estimates ($\widehat{k_1}$ and $\widehat{k_2}$) with respect to calibration targets (h_1, h_2, and h_3) based on (6):

$$\mathbf{J}_I = q_f l \begin{bmatrix} -\frac{1}{(h_1-h_f)^2} & 0 & 0 \\ \frac{1}{2(h_2-h_1)^2} & \frac{1}{2}\left(\frac{1}{(h_3-h_2)^2} - \frac{1}{(h_2-h_1)^2}\right) & -\frac{1}{2(h_3-h_2)^2} \end{bmatrix} \tag{9}$$

and compute an expression for the covariance matrix of estimation errors using (4):

$$\begin{aligned}\mathbf{C}_{pI} &= \mathbf{J}_I\mathbf{C}_o\mathbf{J}_I^T \\ &= q_f^2 l^2 \begin{bmatrix} \frac{\sigma_{h1}^2}{(h_1-h_f)^4} & -\frac{\sigma_{h1}^2}{2(h_2-h_1)^2(h_3-h_2)^2} \\ -\frac{\sigma_{h1}^2}{2(h_2-h_1)^2(h_3-h_2)^2} & \frac{1}{4}\left[\frac{\sigma_{h1}^2}{(h_2-h_1)^4} + \sigma_{h2}^2\left(\frac{1}{(h_3-h_2)^2} - \frac{1}{(h_2-h_1)^2}\right)^2 + \frac{\sigma_{h3}^2}{(h_3-h_2)^4}\right]\end{bmatrix}\end{aligned} \tag{10}$$

The forward-model sensitivity matrix suggests that all the model-predicted observations ($\widehat{h_1}$, $\widehat{h_2}$, and $\widehat{h_3}$) depend on the model parameter k_1 (7; matrix column 1). However, based on the inverse-model sensitivity matrix, we conclude that only the observation h_1 impacts the inverse estimate $\widehat{k_1}$. The model-predicted observation $\widehat{h_1}$ does not depend on k_2 (7; matrix element [1,2]), whereas the inverse estimate of $\widehat{k_2}$ depends on the calibration target h_1 (9; matrix element [2,1]). These comparisons demonstrate that if we want to estimate the importance of calibration targets in the parameter estimation, the analysis should be based on a inverse-model sensitivity matrix, not on a forward-model sensitivity matrix.

The disparity between the sensitivity matrices causes significant differences in the respective correlation matrices of estimation errors (8 and 10). The variance of estimation errors associated with $\widehat{k_1}$ based on forward-model sensitivities (8; matrix element [1,1]) depends on the errors of all the observations, even though observations h_2 and h_3 have no impact on the $\widehat{k_1}$ estimate (6). The variance of estimation errors associated with $\widehat{k_1}$ based on inverse-model sensitivities (10; matrix element [1,1]) depends only on the observation error of h_1 (σ_{h1}^2). The covariance (off-diagonal) terms in both matrices demonstrate the same discrepancy: the covariance of estimation errors between the two parameters should depend only on the observation error σ_{h1}^2 (10) and not on all the observation errors (8). Finally, the variance of estimation errors associated with $\widehat{k_2}$ has different expressions in the matrices (8 and 10; matrix elements [2,2]), but both of them are functions of all three observation errors, as expected. These comparisons demonstrate that the inverse-sensitivity approach is superior to the forward-sensitivity approach in estimating errors in inverse-model parameters.

It is important to note that the estimation uncertainty of $\widehat{k_1}$ derived from the inverse-sensitivity approach does not depend on the number of observations to the left of h_1 (Figure 1), nor on their respective observation errors; however, this is not the case with the forward-sensitivity approach. In addition, the estimation uncertainty in the inverse-sensitivity approach depends on the way we compute the model parameters. In the expression for $\widehat{k_2}$, should we average the gradients between the three observations differently than in (6), for example

$$\widehat{k_2} = \frac{q_f l}{2}\left[\frac{2}{h_3 - h_1} + \frac{1}{h_3 - h_2}\right] \tag{11}$$

The $\widehat{k_2}$ estimation uncertainty would be different as well. This demonstrates that the inverse-sensitivity approach allows us to take into account how the mathematical formulation of the inverse problem impacts the propagation of uncertainties from the observation space onto parameter space. Differences among alternative mathematical formulations of

Table 1
Estimation errors of model parameters using forward- and inverse-sensitivity approaches.

Case number	Observation errors			Estimation errors (forward)		Estimation errors (inverse)	
	$\sigma_{h1}^2[m^2]$	$\sigma_{h2}^2[m^2]$	$\sigma_{h3}^2[m^2]$	$\sigma_{k1}^2[m^2/d^2]$	$\sigma_{k2}^2[m^2/d^2]$	$\sigma_{k1}^2[m^2/d^2]$	$\sigma_{k2}^2[m^2/d^2]$
1	0.1	0.1	0.1	0.083	0.05	0.1	0.05
2	0.1	1.	0.1	0.098	0.05	0.1	0.05
3	0.1	0.1	1.	0.093	0.14	0.1	0.275
4	0.1	1.	1.	0.098	0.235	0.1	0.275
5	1.	0.1	0.1	0.333	0.14	1.	0.275
6	1.	1.	1.	0.833	0.5	1.	0.5

the inverse problem represent one type of conceptual model uncertainty that might be important to consider in our error analysis. Apparently, this conceptual model uncertainty cannot be the assessed by the forward-sensitivity approach.

We should also remark that if the number of observations is equal to the number of parameters and if the spatial distribution of observations and parameters is such that each parameter is directly associated with a single observation, both approaches produce mathematically equivalent expressions for the covariance matrices of estimation errors ($\mathbf{C}_{pF} \equiv \mathbf{C}_{pI}$). For our simple case in Figure 1, these conditions will be satisfied if the permeability zones k_i are encompassing the spaces between consecutive observation locations h_i (where $i = 1, \ldots, N$ and N is the number of observations/parameters); e.g., if observation h_3 is ignored or if one extra parameter k_3 for permeability of the zone between h_3 and h_2 is added.

To further demonstrate the differences between the two approaches, we evaluate the estimation errors for a series of examples. We set $l = 1\ m$, $q_f = 1\ m/d$, $h_f = 0\ m$, $h_1 = 1\ m$, $h_2 = 2\ m$, $h_3 = 3\ m$. The estimates of model parameters based on (6) are $\widehat{k_1} = \widehat{k_2} = 1\ m/d$. In Table 1, we display the variances of estimation errors calculated for different variances of observation errors. Overall, the forward-sensitivity estimates are smaller than the inverse-sensitivity estimates. The highest discrepancy (on order of 50%) is for Case 5. For the rest of the cases, both approaches produce equal or close values for the $\widehat{k_2}$ estimation uncertainty (σ_{k2}^2); however, the variance of $\widehat{k_1}$ estimation error (σ_{k1}^2) is systematically underpredicted by the forward-sensitivity approach. The numerical results also demonstrate the dependence of "forward-sensitivity" σ_{k1}^2 estimate on h_2 and h_3 observation errors (Table 1; e.g., Cases 1-5). Note that the "inverse-sensitivity" estimate of σ_{k2}^2 does not depend on σ_{h2}^2 (Table 1; e.g., Cases 1 and 2, Cases 3 and 4) even though it should (10); due to selected values of calibration targets we get $h_3 - h_2 = h_2 - h_1$, and the middle term of matrix element [2,2] in (10) containing σ_{h2}^2 gets canceled. That is, σ_{k1}^2 becomes σ_{k2}^2, σ_{h1}^2 becomes σ_{h2}^2, and σ_{h1}^2 becomes σ_{h2}^2.

4. DISCUSSION AND CONCLUSIONS

The assessment of estimation uncertainty using the inverse-sensitivity approach is generally superior to the commonly-applied forward-sensitivity approach. However, the proposed methodology requires the evaluation of inverse-model sensitivity matrix representing the dependency of inverse-model-predicted parameters on calibration targets. For the simple case presented in this paper, the evaluation can be done analytically. For much more complicated numerical models, the derivation might only be performed numerically (e.g., using a finite-difference approach), which can be a very computationally intensive task. Nonetheless, the matrix evaluation can be performed efficiently through parallelization; we have been successful in the numerical derivation of inverse-model sensitivity matrix elements for relatively large and complex inverse models [9]. Another way to substantially decrease the computational burden is to use approximate (reduced) representations of the forward model in the inverse process [7,8].

Special care should be taken when applying the forward-model sensitivity analysis of estimation errors for inverse models, especially when (a) the number of calibration targets and the number of model parameters are in the same order, or (b) the spatial structures of calibration targets and model parameters prompt dependency between observation and estimation errors.

The differences between the forward- and inverse-model sensitivity matrices derived for the simple test case demonstrate that the forward sensitivity analysis for evaluation of the importance of calibration targets and/or quality of inverse estimates might not always be accurate. For example, high forward-model sensitivity of model-predicted observations to the model parameters does not necessarily imply high importance of the respective targets in the calibration process. Also, model parameters that cause substantial changes in the model-predicted observations might not be estimated with high accuracy by the inverse model.

In contrast, inverse-sensitivity analyses address these potential deficiencies, and therefore may be useful for problems such as estimation of predictive uncertainty, optimization of data collection strategies, and design of monitoring networks.

REFERENCES

1. J. Bard. *Nonlinear Parameter Estimation.* Academic Press, New York, 1974.
2. J. Carrera, A. Medina, C. Axness, and T. Zimmerman. Formulations and computational issues of the inversion of random fields. In G. Dagan and S. P. Neuman, editors, *Subsurface Flow and Transport: A Stoshastic Approach*, pages 62–79, Cambridge UK, 1997. University Press.
3. J. Carrera and S. P. Neuman. Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood incorporating prior information. *Water Resources Research*, 22:199–210, 1986.
4. D. McLaughlin and L. R. Townley. A reassessment of the groundwater inverse problem. *Water Resources Research*, 32(5):1131–1161, May 1996.
5. S. P. Neuman. A comprehensive strategy of hydrogeologic modeling and uncertainty analysis for nuclear facilities and sites. Technical Report NUREG/CR-6805, U.S. Nuclear Regulatory Commission, Washington D.C., July 2003.

6. N. Z. Sun. *Inverse Problems in Groundwater Modeling.* Kluwer, Norwell, Mass., 1994.
7. D. W. Vasco and S. Finsterle. Numerical trajectory calculations for the efficient inversion of transient flow and tracer observations. *Water Resources Research*, 40:doi:10.1029/2003WR002362, 2004.
8. P. Vermeulen, A. Heemink, and C. T. Stroet. Reduced models for linear groundwater flow models using empirical orthogonal functions. *Advances in Water Resources*, 27(1):56–69, 2004.
9. V. V. Vesselinov and E. H. Keating. Inverse-inverse analysis: A general methodology for exploring sensitivity and uncertainty of model predictions. volume 83. AGU Fall Meeting, December 6-10 2002.
10. W. W.-G. Yeh. Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resources Research*, 22(2):95–108, Feb. 1986.