ground water

Forward Model Nonlinearity versus Inverse Model Nonlinearity

by Steffen Mehl[1,2]

Abstract

The issue of concern is the impact of forward model nonlinearity on the nonlinearity of the inverse model. The question posed is, "Does increased nonlinearity in the head solution (forward model) always result in increased nonlinearity in the inverse solution (estimation of hydraulic conductivity)?" It is shown that the two nonlinearities are separate, and it is not universally true that increased forward model nonlinearity increases inverse model nonlinearity.

Introduction

Coping with model nonlinearities is a routine practice for ground water modelers. The problems caused by nonlinearities differ depending on whether forward modeling or inverse modeling is attempted. Nonlinearities in the forward model can cause convergence difficulties for numerical solvers and/or limit the applicability of results based on principles of superposition. In the extreme, unique solutions are not guaranteed. Inverse model nonlinearity can cause convergence difficulties in regression procedures used for model calibration and/or can limit the applicability of statistical measures of uncertainty based on linear theory. The issue of concern for this work is how the forward model nonlinearity affects the inverse model nonlinearity. The inverse model is based on the forward model, so one might assume that nonlinearity in the forward model will cause additional nonlinearity in the inverse model. This work uses a simple test case to examine the relationship between these two nonlinearities and demonstrates that increasing the forward model nonlinearity does not necessarily increase the inverse model nonlinearity. The issue of nonuniqueness in the forward model and its affects on the inverse model results, such as parameter estimates and predictions, is not addressed in this article.

[1]Corresponding author: U.S. Geological Survey, 3215 Marine Street, Boulder, CO 80303; (303) 541-3097; (303) 447-2505; swmehl@usgs.gov

[2]Currently at California State University, Chico, Department of Civil Engineering, Chico, CA 95929.

Received March 2007, accepted July 2007.

doi: 10.1111/j.1745-6584.2007.00372.x

Methods

The simple model considered in this work is shown in Figure 1. It is a three-node, cross-sectional model with specified-head boundaries on the left and right nodes. It really is a one-node problem in that the only unknown head is at the central node. For both confined and unconfined conditions, the forward and inverse equations for this system are derived, examined, and compared.

The Nonlinearity in the Forward Model

The forward model solves for the unknown head, given the boundary conditions and flow system properties. For the simple test case considered here, this can be obtained by performing a steady-state mass balance on the central cell, which results in:

$$q_L - q_R = 0 \tag{1}$$

where q_L = flow per unit width into the central cell through the left face, and q_R = flow per unit width out of the central cell through the right face.

The flow across the cell faces can be expressed in terms of conductances and head differences, as follows:

$$q_L = C_L \cdot (h_L - h) \tag{2a}$$

$$q_R = C_R \cdot (h - h_R) \tag{2b}$$

where C_L = the conductance of the aquifer material between the left node and the central node, C_R = the

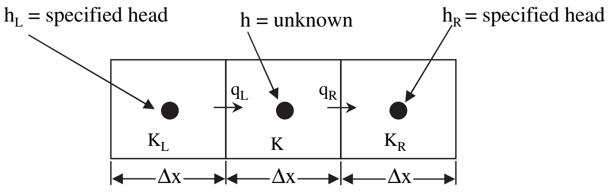

Figure 1. Schematic of a three-node block-centered cross-sectional model with uniform grid spacing and specified-head boundaries at the left and right nodes, (h_L and h_R, respectively).

conductance of the aquifer material between the right node and the central node, h_L = head specified as a boundary condition at the left node, h_R = head specified as a boundary condition at the right node, and h = head at the central node (unknown).

Substituting Equation 2 into Equation 1 and gathering terms that multiply h results in:

$$h \cdot (-C_L - C_R) = -C_L \cdot h_L - C_R \cdot h_R \quad (3)$$

Solving for h results in:

$$h = (-C_L \cdot h_L - C_R \cdot h_R)/(-C_L - C_R) \quad (4)$$

For a confined aquifer, conductance for a unit width is calculated as follows:

$$C_L = (2 \cdot K_L \cdot b_L \cdot K \cdot b)/(K_L \cdot b_L \cdot \Delta x + K \cdot b \cdot \Delta x) \quad (5a)$$

$$C_R = (2 \cdot K_R \cdot b_R \cdot K \cdot b)/(K_R \cdot b_R \cdot \Delta x + K \cdot b \cdot \Delta x) \quad (5b)$$

where K_L = the hydraulic conductivity of the left cell, K_R = the hydraulic conductivity of the right cell, K = the hydraulic conductivity of the central cell, b_L = the saturated thickness of the left cell, b_R = the saturated thickness of the right cell, b = the saturated thickness of the central cell, and Δx = the grid spacing or cell dimension.

When these conductances are substituted into Equation 4, head (h) can be solved for explicitly, and therefore the forward model is linear with respect to head.

For an unconfined aquifer, *conductance is a function of head* (this is where the nonlinearity in the forward model comes from) and for a unit width is calculated as follows:

$$C_L = (2 \cdot K_L \cdot h_L \cdot K \cdot h)/(K_L \cdot h_L \cdot \Delta x + K \cdot h \cdot \Delta x) \quad (5c)$$

$$C_R = (2 \cdot K_R \cdot h_R \cdot K \cdot h)/(K_R \cdot h_R \cdot \Delta x + K \cdot h \cdot \Delta x) \quad (5d)$$

The only difference between the confined and unconfined conductances is the saturated thickness; in the unconfined case, it is equal to the head in the aquifer (the datum for head is the bottom of the cell). Thus, in the unconfined case, the head solution (Equation 4) is nonlinear in head. That is, the head is a function of terms that are themselves functions of head. Specifically, for unconfined aquifers, the conductances depend on the head, as shown in Equations 5c and 5d. Because we do not know the correct head yet, in the unconfined case, we do not know the correct value of conductance, and so Equation 4 must be solved iteratively.

The Nonlinearity in the Inverse Model

The nonlinearity of the inverse model differs from that of the forward model because the function of interest is not the head solution but the derivative of the head solution with respect to some parameter (these derivatives are also called sensitivities). The derivatives and observations are used to update parameter values during model calibration. In the inverse model, the nonlinearity is with respect to parameters; in the simple model used here, the parameter is hydraulic conductivity. The nonlinearity of the inverse model with respect to hydraulic conductivity arises directly from Darcy's Law where head changes ($h - h_0$) are proportional to distance (ΔL) and flow rate (Q) and inversely proportional to hydraulic conductivity (K) and cross-sectional area (A):

$$h = h_0 - \Delta L \cdot Q/(K \cdot A) \quad (6)$$

Thus, the derivative of head with respect to K results in:

$$\partial h/\partial K = \Delta L \cdot Q/(K^2 \cdot A) \quad (7)$$

as also shown by Hill and Tiedeman (2007, 12–13). This is a nonlinear function of K and arises from inverse proportionality of K with respect to h in Darcy's Law. Equation 7 also shows a dependence on the flow rate, which indicates that the flow system properties play a role in nonlinearity.

For the three-node model presented, the derivative of h (Equation 4) with respect to K can be calculated. For the confined case, making the simplification that the saturated thickness, b, is uniform throughout, the sensitivity coefficient simplifies to:

$$\partial h/\partial K = [(h_L - h_R) \cdot K_L \cdot K_R \cdot (K_L - K_R)]/[2 \cdot K_L \cdot K_R + K \cdot (K_L + K_R)]^2 \quad (8)$$

For the unconfined case, the chain rule must be invoked to calculate the derivative of h with respect to K because Equation 4 is a function of both K and h, and h is itself a function of K. Designating the right-hand side of Equation 4 as "f" and using the chain rule for the derivatives results in a sensitivity coefficient of:

$$\partial h/\partial K = \partial f/\partial h \cdot \partial h/\partial K + \partial f/\partial K \quad (9)$$

Solving for $\partial h/\partial K$ results in

$$\partial h/\partial K = \partial f/\partial K/(1 - \partial f/\partial h) \quad (10)$$

where

$$\partial f/\partial K = [h \cdot h_L \cdot (h_L - h_R) \cdot h_R \cdot K_L \cdot K_R \cdot (h_L \cdot K_L - h_R \cdot K_R)]/[2 \cdot h_L \cdot h_R \cdot K_L \cdot K_R + h \cdot K \cdot (h_L \cdot K_L + h_R \cdot K_R)]^2 \quad (11a)$$

$$\partial f/\partial h = [K \cdot h_L \cdot (h_L - h_R) \cdot h_R \cdot K_L \cdot K_R \cdot (h_L \cdot K_L - h_R \cdot K_R)]/[2 \cdot h_L \cdot h_R \cdot K_L \cdot K_R + h \cdot K \cdot (h_L \cdot K_L + h_R \cdot K_R)]^2 \quad (11b)$$

The second derivative of h with respect to K is a measure of curvature and can be used to indicate the degree of nonlinearity. The curvature indicates how much the first derivative (sensitivity) changes with changes in K. For the confined case, the curvature is:

$$\partial^2 h/\partial K^2 = -[2 \cdot (h_L - h_R) \cdot K_L \cdot K_R \cdot (K_L - K_R) \cdot (K_L + K_R)]/[2 \cdot K_L \cdot K_R + K \cdot (K_L + K_R)]^3 \quad (12)$$

For the unconfined case, the chain rule is again invoked as for Equation 9 to obtain:

$$\partial^2 h/\partial K^2 = [(\partial^2 f/\partial h^2 \cdot \partial h/\partial K + 2 \cdot \partial^2 f/\partial K \partial h) \cdot \partial h/\partial K + \partial^2 f/\partial K^2]/(1 - \partial f/\partial h) \quad (13)$$

where

$$\partial^2 f/\partial K^2 = -[2 \cdot h^2 \cdot h_L \cdot (h_L - h_R) \cdot h_R \cdot K_L \cdot K_R \cdot (h_L \cdot K_L - h_R \cdot K_R) \cdot (h_L \cdot K_L + h_R \cdot K_R)]/[2 \cdot h_L \cdot h_R \cdot K_L \cdot K_R + h \cdot K \cdot (h_L \cdot K_L + h_R \cdot K_R)]^3 \quad (14a)$$

$$\partial^2 f/\partial h^2 = -[2 \cdot K^2 \cdot h_L \cdot (h_L - h_R) \cdot h_R \cdot K_L \cdot K_R \cdot (h_L \cdot K_L - h_R \cdot K_R) \cdot (h_L \cdot K_L + h_R \cdot K_R)]/[2 \cdot h_L \cdot h_R \cdot K_L \cdot K_R + h \cdot K \cdot (h_L \cdot K_L + h_R \cdot K_R)]^3 \quad (14b)$$

$$\partial^2 f/\partial K \partial h = [h_L \cdot (h_L - h_R) \cdot h_R \cdot K_L \cdot K_R \cdot (h_L \cdot K_L - h_R \cdot K_R) \cdot (2 \cdot h_L \cdot h_R \cdot K_L \cdot K_R - h \cdot K \cdot (h_L \cdot K_L + h_R \cdot K_R))]/[2 \cdot h_L \cdot h_R \cdot K_L \cdot K_R + h \cdot K \cdot (h_L \cdot K_L + h_R \cdot K_R)]^3 \quad (14c)$$

Results and Discussion

For the confined case, the derivative of h with respect to K (Equation 8) is a function that does not depend on h and is a nonlinear function of K only. This can be compared to the unconfined case where the derivative of h with respect to K (Equations 10 and 11) is a nonlinear function of both h and K. In general, the unconfined case is more difficult to solve because the forward model has nonlinearities and the inverse model must calculate additional derivative terms that result from the nonlinearity in the forward model. This is perhaps where confusion arises in terms of nonlinearity of the forward model producing more nonlinearity in the inverse model. The additional derivative terms do not necessarily increase the nonlinearity because, depending on their sign relative to the other terms, they can cause cancellation and decrease the nonlinearity.

In both the confined and the unconfined case, changes in K change the head solution. This is shown in Figure 2, which plots the head solution (Equation 4 using Equation 5a and 5b or 5c and 5d) for both cases in relation to K. If either function were linear in K, it would plot as a straight line. The curvature of the graphs in Figure 2 shows that both the confined and unconfined solutions are nonlinear with respect to K. For different values of K, the slope of the lines and therefore the nonlinearity vary. For example, for larger values of K, the graphs are more linear.

Examining Equations 12, 13, and 14 provides some insights regarding the characteristics of the nonlinearity. If $h_L = h_R$, then in both cases the derivative (Equations 8 and 11) and curvature (Equations 12 and 14) are zero and there is no nonlinearity. Physically, when $h_L = h_R$, there is no flow in the system, and therefore changes in K do not change head. While a trivial example, it clearly illustrates that boundary conditions affect the nonlinearity in the inverse model; therefore, flow system characteristics beyond hydraulic parameters affect nonlinearity. If $K_L = K_R$ in the confined case, then the same conductance links the central node to the left and right specified-head boundary conditions. The head at the central node is bound along a straight line between the two constant heads independent of K at the central node and there is no nonlinearity. In the unconfined case, because of the head dependence, the conductances are not the same and the system is nonlinear. This trivial example demonstrates that in some situations, the nonlinearity in the forward solution can cause more nonlinearity in the inverse solution. These results are an artifact of a single node bound between two specified-head nodes and indicate, as before, that flow system characteristics can influence the nonlinearity. In more general cases, numerical measures such as modified Beale's measure and measures of total and

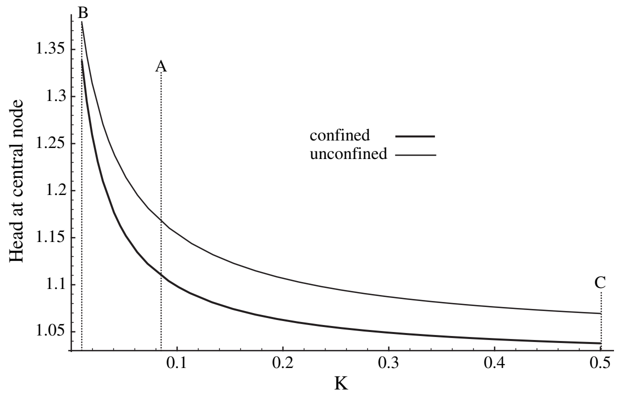

Figure 2. Head at the central node with respect to K for the confined and unconfined cases. h_L=1, h_R=2, K_L=0.5, K_R=0.01, and K is varied on the x-axis from 0.01 to 0.5. Points A, B, and C relate to Figure 3. Referring to Figure 3, for K less than 0.0845 (point A), the curvature in the head solution is greater for the confined case compared to the unconfined case.

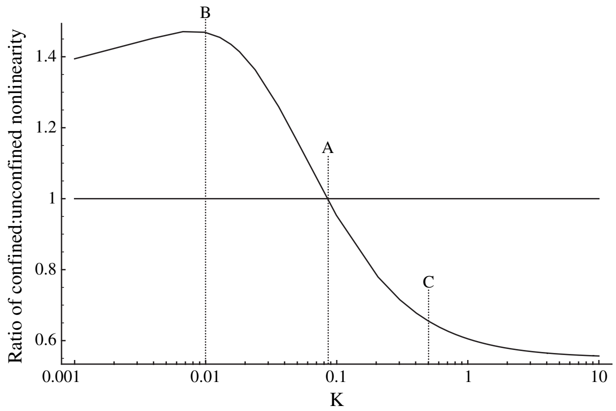

Figure 3. Ratio of confined to unconfined second derivatives with respect to *K*. Second derivatives are a measure of curvature (nonlinearity). Values greater than 1 (left of point A) indicate the confined case has more nonlinearity and values less than 1 (right of point A) indicate the unconfined case has more nonlinearity. Points A, B, and C correspond to those in Figure 2.

intrinsic nonlinearity (Cooley 2004, 85–87; Poeter et al. 2005, 214) can be used.

The question remains of how the inverse model nonlinearities of the confined and unconfined cases compare. This is demonstrated in Figure 3 using a plot of the ratio of the second derivatives for the confined and unconfined cases. In this plot, the flow system is identical to that of Figure 2 with $h_L=1$, $h_R=2$, $K_L=0.5$, and $K_R=0.01$. K varies from 0.001 to 10. Values greater than 1 indicate more curvature (more nonlinearity) in the inverse model for the confined case. Values less than 1 indicate more curvature in the inverse model for the unconfined case.

Figure 3 clearly shows that the heterogeneity affects the degree of nonlinearity and that for some values of K, the confined case has more nonlinearity than the unconfined case. The saturated thickness is fixed for the confined case and results not reported here show different trends depending on this fixed value. Therefore, Figures 2 and 3 cannot be used to make a general statement regarding how the relative nonlinearity varies with K. Nonetheless, in all simulations investigated, there were K values where the nonlinearity in the confined cases was greater than the unconfined case.

Conclusions

Forward model nonlinearity does not necessarily make the inverse model more nonlinear. The two nonlinearities are separate and do not always combine in a way to make the inverse model more nonlinear. In some cases, it is more nonlinear; in other cases, it is not. Nonlinearity in the inverse model depends on many factors that influence the flow system. How these factors combine to create nonlinearities is difficult to assess without using formal numerical measures.

Acknowledgments

This work was inspired by conversations with Randy Hunt and improved through discussions with Mary Hill. Review comments by Tracy Nishikawa, Howard Reeves, Claire Tiedeman, and Tim Durbin improved this manuscript.

References

Cooley, R.L. 2004. A theory for modeling ground-water flow in heterogeneous media. USGS Professional Paper 1679. Reston, Virginia: USGS.

Hill, M.C., and C.R. Tiedeman. 2007. Effective groundwater model calibration: With analysis of data, sensitivities, predictions, and uncertainty. Hoboken, NJ: John Wiley and Sons.

Poeter, E.P., M.C. Hill, E.R. Banta, S. Mehl, and S. Christensen. 2005. UCODE_2005 and six other computer codes for universal sensitivity analysis, calibration, and uncertainty evaluation. USGS Techniques and Methods 6-A11. Reston, Virginia: USGS.