



Appendix 6. Parameter sensitivity and identifiability

In the history-matching process, the change of observation values in response to changes in parameters, known as the sensitivity of observations to parameters, is used to define the magnitude and directions of changes in candidate parameter values from iteration to iteration. The Jacobian matrix is a matrix of these sensitivity values relating each observation to each parameter. The Jacobian matrix is defined as:

$$\frac{\partial y_i}{\partial p_j} \cong \frac{g(\mathbf{p}_j + \Delta \mathbf{p}) - g(\mathbf{p}_j)}{\Delta p}$$

where

\mathbf{y} is an observation

\mathbf{p} is a parameter value

$g(\mathbf{x})$ is the model outcome corresponding to observation time and location \mathbf{y}

i is a dummy variable in [1,...,NOBS];

j is a dummy variable in [1,...,NPAR];

NOBS is the number of observations;

NPAR is the number of parameters; and

$\Delta \mathbf{p}$ is a small (10%) increment in parameter value.

The result of evaluating this equation for all combinations of parameters and observations is the NOBS * NPAR Jacobian matrix \mathbf{X} . Beyond its use in the history-matching algorithm, sensitivity can provide insight into which parameters have the greatest impact on forecasts made by the model. In fact, sensitivity can be considered as one metric of parameter importance. A challenge is encountered in that parameters are often correlated with one another, so sensitivity can be misleading if thought of as the only

measure of parameter importance. Correlation is evaluated in pairs of parameters so with hundreds or thousands of parameters it is impractical to deconvolve sensitivity from correlation. Identifiability addresses this issue because it is based on singular value decomposition (SVD) (described in appendix 3), which splits information content among correlated parameters. For this reason, we focus on identifiability here.

By virtue of using SVD for parameter estimation, it is also possible to calculate identifiability (Doherty and Hunt, 2009). This is a qualitative metric that indicates how much information from the observations (taken as a whole set) is projected onto the parameters in the history-matching process. Based on a singular value cutoff, the identifiability value calculates the projection of information from observations onto the calibration space. The remaining information is projected onto the null space, meaning parameter values can change, potentially by large amounts, without impacting the model outputs of interest.

Parameters with high identifiability are interpreted to be well-informed by the observations, while parameters with low identifiability are interpreted to be ill-informed by the observations and are generally held relatively constant in the history-matching process. Identifiability values are normalized to a maximum value of 1.0. In this report, we explore the identifiability of parameters for the steady-state model. The reason for this is that the transient model has few parameters whereas the steady state represents the tradeoffs among hydraulic conductivity, recharge, and stream parameters. The lower number of parameters in the transient model is due to holding hydraulic conductiv-

ity and stream parameters constant in the transient calibration. Figure 1 presents the identifiability of all parameters with values greater than 0.5. In the figure, the overall height of each bar indicates its identifiability on a normalized scale of 0.0, where the parameter is not informed by the observations, to 1.0, where the parameter is fully informed by the observations.

By definition, if all singular values are considered, all identifiability values will be equal to 1.0. The identifiability value quantifies the projection of observation information onto the solution space of the parameters as defined by an SVD cutoff. As more singular values are considered to be part of the solution space, fewer are in the null space, and the amount of projection onto solution space increases. The SVD cutoff for the parameters shown in figure 1 is 80, which is the number of SVD-Assist (SVDA) super parameters used in this model. In addition to the 80 singular-value cutoff, a stability criterion was used to enforce a threshold on how many of those super parameters were considered part of the solution space. For each parameter, higher bars on figure 1 indicate the most identifiable parameters. On each bar, however, the contribution to identifiability due to each singular value is displayed which helps to distinguish parameters with similar total identifiability. Hotter colors (reds, oranges, and yellows) indicate identifiability attributable to most-informative singular values (1, 2, and so on), whereas cooler colors (greens and blues) indicate identifiability attributable to the least-informative singular values (78, 79, 80). This spectrum is displayed to highlight that no particular singular-value cutoff is the perfect definition of separating the solution



space from the null space, but bars that are made up almost entirely of hot colors can be qualitatively interpreted to have overall higher identifiability than those containing cooler colors. To avoid plotting bars for all parameters, only those with identifiability (at 80 singular values) greater than 0.5 are displayed on figure 1.

Identifiability is a qualitative metric and should not be over-interpreted; it can, however, provide insight into important model behavior. In a

qualitative sense, parameters with high identifiability can be interpreted as important controls on model behavior. The two parameters with the highest identifiability are rm1 and sfrc, which are the recharge multiplier and streambed hydraulic conductivity, respectively. These two parameters are the main controls on overall water balance into the model and the connection to streams. The Little Plover River flow observations are the most important and highly weighted observations in the model, so it fol-

lows that the parameters controlling water exchange with the stream and overall water available in the model should be the most important. The remaining parameters with identifiability > 0.5 are horizontal hydraulic conductivity parameters—a mixture of homogeneous zones and pilot points. This highlights the importance of lateral distribution of groundwater to supply the streams and also to maintain a head surface honoring head observations. Vertical hydraulic conductivity values are less identifi-

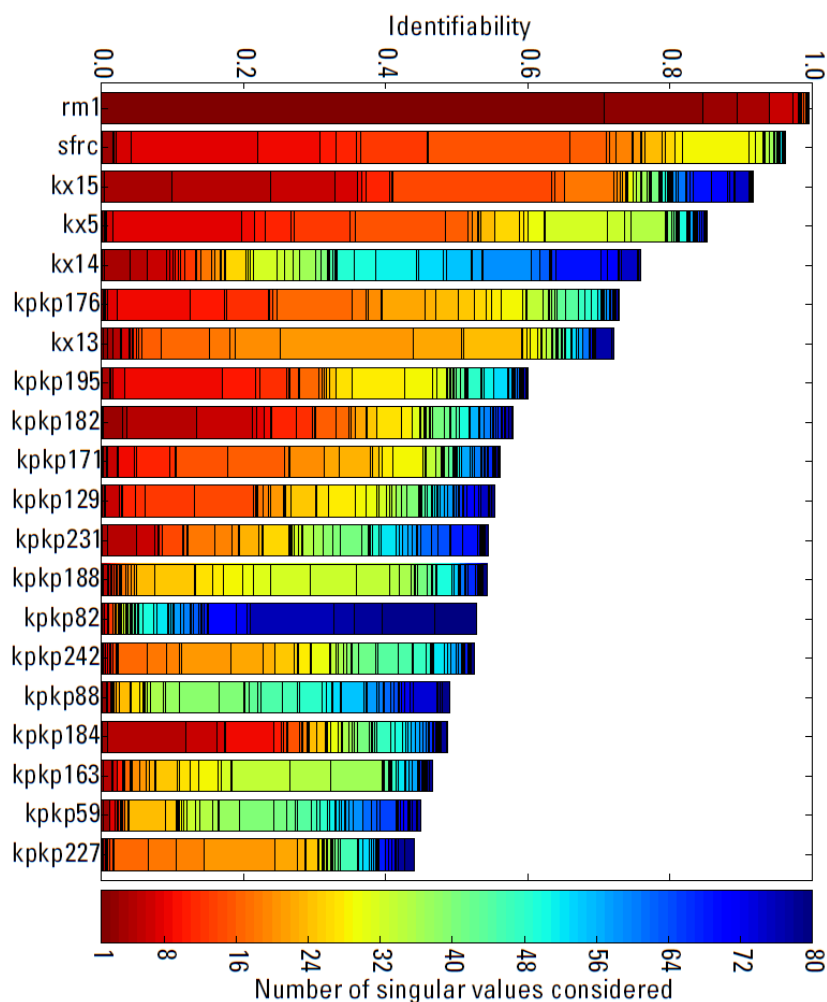


Figure 1. Identifiability of parameters. Set includes only those with identifiability (at 80 singular values) greater than 0.5. The color scale indicates identifiability for each singular value.



able, indicating that they control the solution less and are informed less by the observations.

Figures 2 and 3 display the identifiability along with the estimated K_x and K_v parameter values. Identifiability is presented using the 80 singular value cutoff, corresponding with the values at the top of the bars in figure 1 of the K_x and K_v pilot points. In figure 2, the pilot points

with the highest identifiability are those located near the Little Plover River and driving the high-hydraulic conductivity anomaly near the Cambrian sandstone subcrops discussed in the main report. In figure 3, the highest identifiability values are lower (around 0.3) and meaningful patterns are not easily discernible, indicating that vertical hydraulic con-

ductivity is less important to model results than is horizontal hydraulic conductivity.

Reference

Doherty, J., and Hunt, R.J., 2009, Two statistics for evaluating parameter identifiability and error reduction: *Journal of Hydrology*, v. 366, no. 1–4, p. 119–127.

Identifiability of K_x pilot points

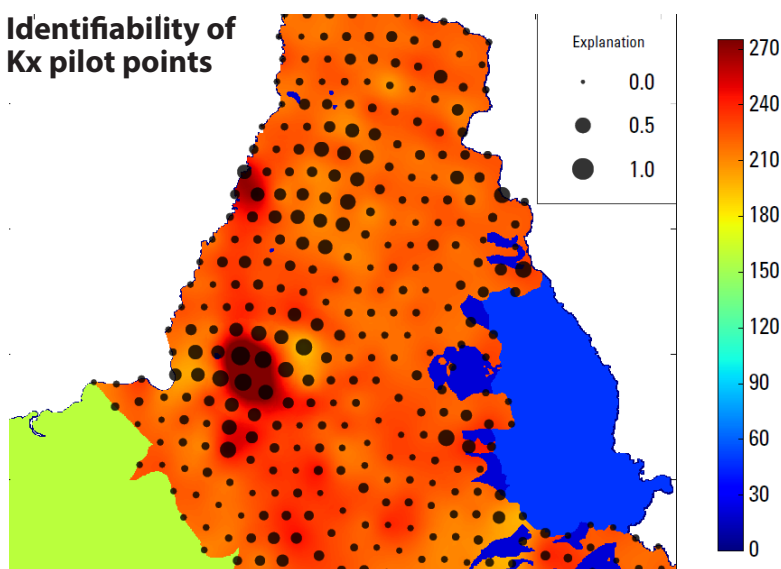


Figure 2. Identifiability of horizontal hydraulic conductivity (K_x) pilot points overlain on the K_x optimal values. The color scale shows K_x in feet per day and the size of the circles indicates identifiability.

Identifiability of K_v pilot points

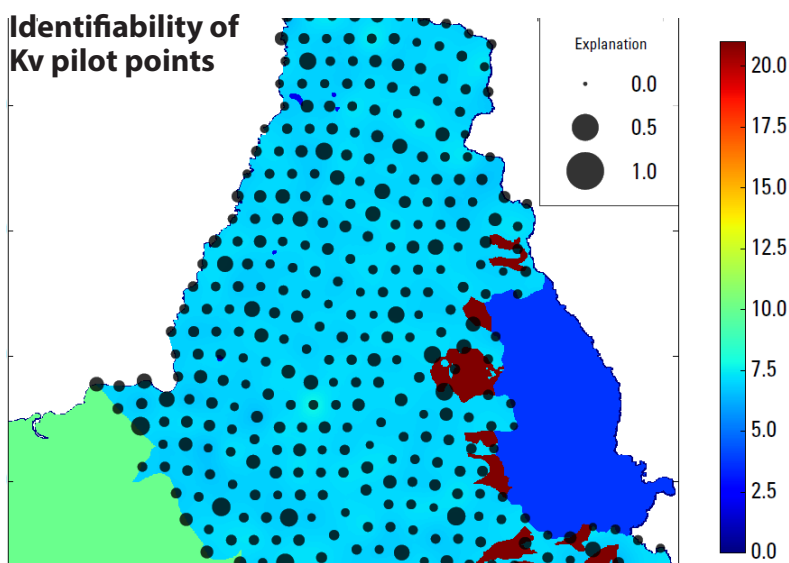


Figure 3. Identifiability of vertical hydraulic conductivity (K_v) pilot points overlain on the K_v optimal values. The color scale shows K_v in feet per day and the size of the circles indicates identifiability.