An innovative use of observations to alleviate weighted-residual asymmetry

Gilbert Barth, Ph.D.
S.S. Papadopulos and Associates, gbarth@sspa.com, Boulder, CO USA

ABSTRACT

Estimation of flow or transport parameters often involves observations such as flow or concentrations that may be significant over multiple orders of magnitude, typically dictating the use of weights inversely proportional to the product of a coefficient of variation multiplied by the observation. This allows observations of considerably different magnitude to have similar importance in terms of guiding the parameter-estimation process. However, if a simulated value is of smaller absolute magnitude than the observed, the weighted residual will be limited to the inverse of the coefficient of variation, while weighted residuals of simulated values greater than the observed are unbounded. This produces asymmetry in the potential contribution to the objective function from simulated values whose magnitudes are less-than and greater-than the observed. This work uses a simple transformation to create a supplemental observation set. The combined observation set is used to demonstrate an objective function with improved characteristics: including the transformed observations results in a balanced set of objective-function contributions from simulated values smaller-than and greater-than the observation magnitude. The potential for improving the process of parameter estimation, generating sensitivities, as well as precautions for interpreting final statistics, are considered.

INTRODUCTION

Successful parameter estimation relies on many factors including a set of observations that provide information sufficient to determine appropriate parameter-value adjustments that improve the fit between simulated and observed values. For a valid regression, weights need to be proportional to one divided by the variance of the observation error, σ_i^2, [*Draper and Smith*, 1998, p. 222] in order to reflect the uncertainty of the measurement. Proper weighting allows observations of different units, phenomena, and a wide range of magnitude to be combined, providing simultaneous feedback from the entire set of observations on changes in the fit between simulated and observed values.

Weights can typically be lumped into one of two categories: fixed-value or variable-value weights, FVW and VVW, respectively. These terms reflect the absolute and relative nature of the weights. FVWs typically reflect the uncertainty of a measurement device and are not considered in this work. VVWs are relative to the observation: they are proportional to a value, typically the observed value. Observed values significant over a limited range, such as groundwater heads, typically use FVW. Observations requiring VVW include concentrations and flows, when they are significant over multiple orders of magnitude.

This work provides an approach for addressing one of the issues commonly associated with VVW applied to observations significant over multiple orders of magnitude: the weights result in a limited magnitude of weighted residual for positive residuals (*observed - simulated* > 0). As a result, an automated parameter estimation routine will tend to adjust parameters to improve for any negative residuals with little, if any regard, for the positive residuals. Examined individually, residuals from any observation would provide an indication of the proper parameter adjustment to improve the fit between the simulated and observed values. However, it is quite possible for a few large weighted residuals, from negative residuals, to dominate the objective function so that the parameter estimation algorithm accepts the limited penalty associated with many positive residuals that produce a relatively minor contribution to the objective function. Previous work [*Barth and Hill*, 2005a and b] explored this and other issues associated with VVW and demonstrated precautions to improve potential for successful parameter estimation and sensitivity analysis when observations are significant over multiple orders of magnitude.

This work goes beyond previous efforts by proposing the use of a simple transformation to create additional observations for supplementing the original observation set. Using a simple one-observation example this work demonstrates the potential benefits of the supplemented observation set: the supplemented observation set results in a more balanced set of weighted residuals with respect to positive versus negative residuals, can be used to provide a more appropriate set of sensitivities, and eliminates one source of parameter-estimate bias. A small hypothetical set of observed and simulated values are also used to examine the benefits of supplementing the observation set. Finally, some considerations regarding observation redundancy are discussed.

METHODS

Parameter Estimation and Weighting

For this work it is assumed that parameters are estimated by quantifying the fit between observed and simulated values, minimizing with respect to the parameter values using a weighted least-squares objective function, $S(\underline{b}_r)$, Equation (1). The weight matrix, ω, and observed- and simulated-value vectors, $\underline{y}$ and $\underline{y}^*(\underline{b}_r)$, respectively, include terms for all the observations, $\underline{b}$ is the parameter vector, and the subscript r indicates the parameter estimation iteration number. This work assumes that the objective function is minimized using a modified Gauss-Newton method [e.g., *Seber and Wild*, 1989; *Sun*, 1994] as described in *Hill* [1998], however the approach of supplementing observations should benefit a wide range of objective-function-minimization approaches.

$$S(\underline{b}_r) = \left[\underline{y} - \underline{y}^*(\underline{b}_r)\right]^T \underline{\omega}\left[\underline{y} - \underline{y}^*(\underline{b}_r)\right] \quad (1)$$

When using VVWs a common approach is to make the weight inversely proportional to the observation multiplied by a coefficient of variation (Eqn. 2). In (2) ND is the number of observations with weights proportional to their value, Cv_i is the coefficient of variation, and y_i should be the true value of the observation. Several issues must be considered when applying this approach, including incorporation of detection limits to avoid very small observations, and that y_i is typically unknown and approximated using observed [*e.g., Keider and Rosbjerg*, 1991] or simulated values [*e.g., Wagner and Gorelick*, 1986]. The simple, one-observation example does not require a detection limit, but a detection limit was used in the example dataset examined in this work. This work uses observed values for y_i in (2).

$$\omega_i = \frac{1}{(Cv_i y_i)^2} \quad where\ i = 1...ND \quad (2)$$

In this work the residual is defined as the observed minus the simulated value. The weighted residual on the i[th] observation (WRy_i) is defined in (3). Assuming normalized values ranging from 0 to 1.0, when either the simulated or observed value is considerably larger than the other, the absolute value of the residual approaches 1.0 and the weighted residual approaches the square root of the weight (4). As a result, for observations significant over N orders of magnitude, there would typically be a difference of about N orders of magnitude for positive versus negative weighted residuals [*Barth and Hill*, 2005a].

$$WR_{y_i} = (y_i - y_i^*)\omega_i^{1/2} = \frac{(y_i - y_i^*)}{Cv_i y_i} \quad (3)$$

$$\left|WR_{y_i}\right| \approx \omega_i^{1/2} = \frac{1}{Cv_i y_i} \quad (4)$$

For this work the hypothetical examples include flows and concentrations ranging over multiple orders of magnitude. For example, flow observation would be the total volume that flowed in a stream over the irrigation stress period, e.g., 10^9 ft^3. Values could cover a wide range depending on the type, units, and other factors. For simplicity the first example, a single-observation example, considers the situation where the observed value has been normalized so that it equals 1.0. Using (2) and Cv = 0.1, Figure 1 shows that the magnitude of *WRy** ranges from about 10, for positive weighted residuals where *y** is less than about 0.1, to an unbounded amount for negative weighted residuals, as *y** grows larger than 10.

The use of a log scale is important to the interpretation of the results depicted in Figure 1. A linear scale would simply show an identical slope for both the negative and positive weighted residuals. However, the observation is significant over multiple orders of magnitude and as a result, the linear scale will tend to mask the fact that positive weighted residuals do not produce similar increases in weighted-residual magnitude for each additional order of magnitude deviation from the observed value. The log scale, having order of magnitude increments, has the capacity to correctly display and convey this issue.

Supplemental Observations

To eliminate the asymmetry depicted in Figure 1 this work proposes a simple transformation of the original observations, and appending the transformed values to the original observation set. The transformation is simply taking the multiplicative inverse of each observation, y_i, to create additional observations, z_i: $z_i = 1/(y_i)$. These new observations, referred to as inverse observations, are appended to the existing set, resulting in a set of 2**ND* observations. The implications of using observations twice are discussed below.

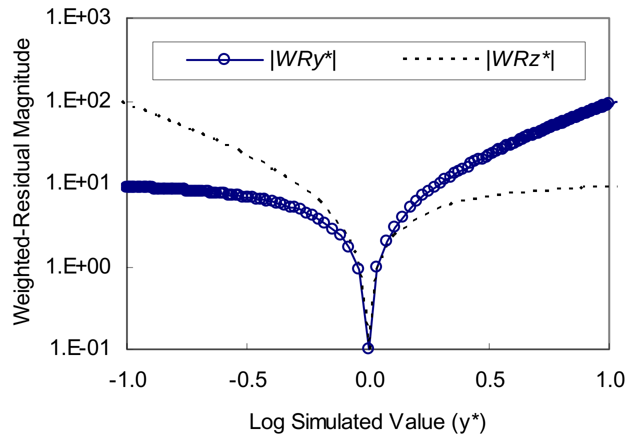

Figure 1. Weighted residual magnitudes using simulated, |WRy*|, and inverse simulated |WRz*| values, where y = 1, Cv = 0.1

Weights on z_i are simply calculated in the same manner as y_i. The weighted residuals, WRz_i in Equation (5), are quite similar to WRy_i. Comparison of (3) and (5) reveals that the net result of adding inverse observations is to produce additional weighted residuals that are opposite in sign, and are inversely proportional to the simulated value, as opposed to the observed value. As discussed below, being inversely proportional to the simulated value is one of the fundamental benefits of using the inverse observations.

$$WR_{z_i} = \left(\frac{1}{y_i} - \frac{1}{y_i^*}\right)\frac{1}{Cv_i \dfrac{1}{y_i}} = \frac{y_i^* - y_i}{Cv_i y_i^*} \qquad (5)$$

RESULTS AND DISCUSSION

Figure 1 also shows the weighted residual from the inverse observations (*WRz**), as a function of simulated value. For inverse observations and $Cv = 0.1$, negative weighted residuals have a maximum value of 10, while positive weighted residuals increase without bound. The original and inverse observation sets complement each other: the original observations produce significant weighted residuals for negative residuals, while the inverse-observation weighted residuals are significant for positive residuals. While this does not eliminate the issue of redundant use of observations, it at least suggests that the contribution of the original and inverse observations have only minor overlap for any given simulated value.

Figure 2 provides an example from a hypothetical-scenario dataset that includes flows and concentrations. Figure 2a shows the simulated, observed, and weights for the original set and the inverse set. Observations 1 – 5 are flow in a drain over the course of 5 stress periods with values varying between values of ~10^3 to 10^{12}, representing flow from the drain during the successive non-irrigation and irrigation seasons. The simulated values fluctuate between 10^7, and 10^9, attempting to represent low flows in the non-irrigation season, and high flows during the irrigation season. Observations 6 – 10 represent concentrations, in the range from 10^{-5} to 10^{-10}. The simulated equivalents are relatively consistent around 10^{-7} or 10^{-8}. Note that, since weights are inversely proportional to the values, the weights for the original observations are adjacent to the inverse observations and the weights for the inverse observations are adjacent to the observations.

Figure 2b shows the residual magnitudes from the original and inverse observations, and how they roughly track each other.

The weighted residuals in Figure 2c provide a good example of the combined potential of the original and inverse observations. Whenever *WRy* is small due to positive weighted residual, *WRz* becomes large, and vice versa. When both *WRy* and *WRz* are small then the simulated value really is getting close to the observed value. In this way, *y* and *z* complement each other, providing more balanced feedback through the objective function regardless of whether *y** and *z** are too large or too small.

Generating Dimensionless Scaled Sensitivities

While the combination of original and inverse observations provides improved potential for parameter estimation, this work advocates generating dimensionless scaled sensitivities (dss) [*Hill*, 1998] from the inverse observations alone. Equation (5) demonstrates that inverse observations have weight residuals inversely proportional to the simulated value so that dss of the inverse observations have the same benefits as simulated value weights (SVW) [*Barth and Hill*, 2005a]. The dss_{SVW} provide a better representation of the impact of changes in the parameter on the simulated values [*Barth and Hill*, 2005a]. Using only the inverse observations to generate sensitivities also eliminates any potential issues of data redundancy for the sensitivities.

An Approach for Using the Inverse Observations

Previous work [*Anderman and Hill*, 1999] has demonstrated that weights based on observations may produce a biased set of parameter estimates while weights based on the simulated values produce unbiased parameter estimates. *Anderman and Hill* [1999] point out that using simulated values from the start of a parameter estimation run could be problematic, and adopted an approach that started with observed-value weighting, and switched to simulated-value weighting as parameter estimate changes decreased. This approach is documented in the UCODE manual [*Poeter and Hill*, 1998]. Switching allows the parameter estimation to initially work with a consistent set of values and weights when parameters are likely to change significantly. As the per-iteration parameter-value changes decrease, SVWs are used to eliminate the potential source of bias.

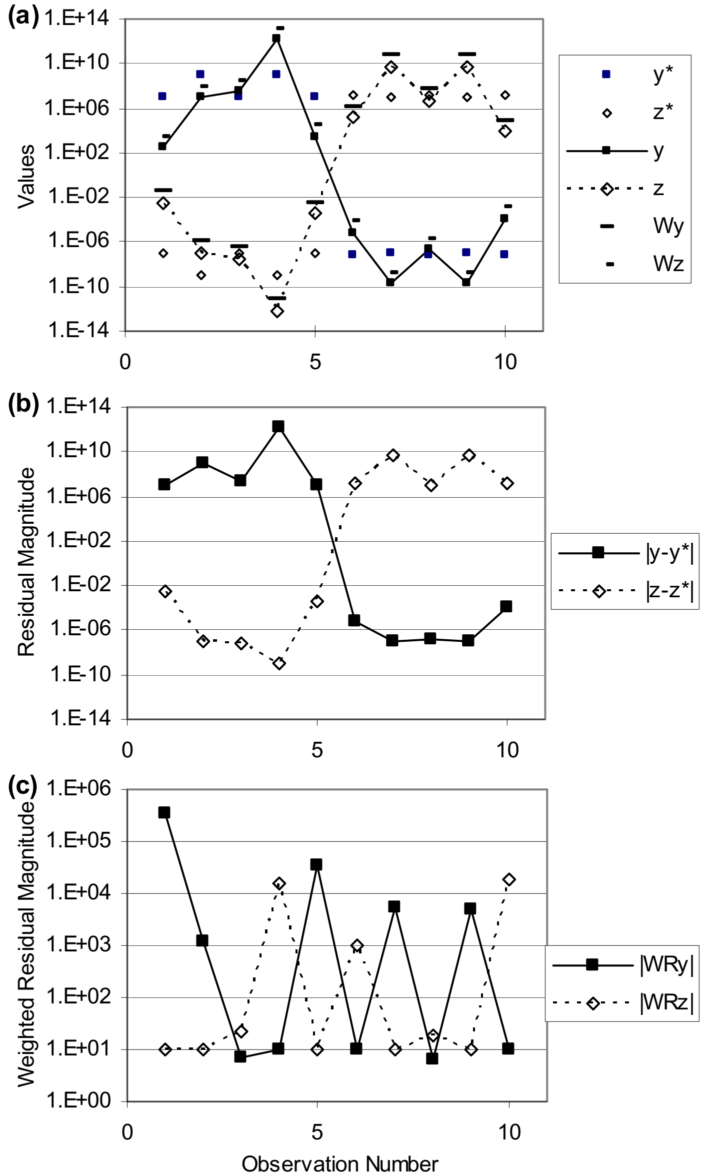

Figure 2. (a) Simulated (y* and z*), observed (y and z), and weights (Wy and Wz), (b) residuals, (c) weighted residuals.

Together, the asymmetry of weighted residuals demonstrated in this work and bias potential identified by *Anderman and Hill* [1999] suggest the need for modifying the approach to observations and weighting. The proposed approach is as follows

- Start with the supplemented observation set, which produces symmetric weighted residuals about the observed values.
- Drop the original observations, leaving only the inverse observations which do not have the potential for a weighting-induced bias of the parameter estimates, based on one of the following decision mechanisms:
    - the per-iteration changes in parameter values have become small, or
    - the residual magnitudes have become small.

This approach incorporates the balanced weighted-residual feedback of the supplemented observation set during initial iterations and then eliminates the potential parameter-estimate bias by using only the inverse observations for the final iterations. In addition, switching to inverse observations, as opposed to simulated values, for the final iterations avoids the need to generate new weights at each parameter-estimation iteration.

Ongoing research focuses on developing intuitive test cases for demonstrating inverse-observation impact and testing the two decision mechanisms: per-iteration parameter-estimate changes or residual magnitudes. Additional work includes refinement of the residual-magnitude decision mechanism, comparing outcomes using (1) the maximum residual, (2) a norm of residuals, (3) residuals of individual observations, or (4) comparison of positive and negative residuals.

CONCLUSIONS

The demonstrated benefits of supplementing the observation set with inverse observations include (1) symmetric weighted residuals, (2) an alternative method of producing a more appropriate set of dimensionless scaled sensitivities, and (3) the possibility of unbiased parameter estimates. These benefits advocate the development of a complete test case and refinement of the decision mechanism, which is part of the ongoing research. While observation-data redundancy is still a potential issue, it seems clear that there are ways to incorporate the benefits of the inverse observations, and still generate valid statistical summaries when needed.

ACKNOWLEDGEMENTS

Comments and suggestions from Steffen Mehl are greatly appreciated.

REFERENCES

Barth, G.R., Hill, M.C., 2005a. Numerical methods for improving sensitivity analysis and parameter estimation of virus transport simulated using sorptive–reactive processes. *J. Contam. Hydrol.* 76, 251– 277.

Barth, G.R., Hill, M.C., 2005b. Parameter and observation importance in modeling virus transport in saturated porous media – investigations in a homogeneous system, *J. Contam. Hydrol.* 80, 107–129.

Draper, N.R., Smith, H., 1998. Applied Regression Analysis, 3rd ed. John Wiley and Sons, New York.

Hill, M.C., 1998. Methods and guidelines for effective model calibration, U.S. Geological Survey Water Resources Investigation Report, 98–4005, 90 p.

Keider, A., Rosbjerg, D., 1991. A comparison of four inverse approaches to groundwater flow and transport parameter identification. *Water Resour. Res.* 27 (9), 2219– 2232.

Poeter, E. P. and M. C. Hill, 1998. Documentation of UCODE, a computer code for universal inverse modeling, USGS Water Resources Investigations Report 98-4080, 123 p.

Seber, G.A.F., Wild, C.J., 1989. Nonlinear Regression. Wiley, New York, NY. 768 pp.

Sun, N-Z, 1994. Inverse Problems in Groundwater Modeling. Kluwer Academic Publishers, The Netherlands.

Wagner, B.J., Gorelick, S.M., 1986. A statistical methodology for estimating transport parameters: theory and applications to one-dimensional advective–dispersive systems. *Water Resour. Res.* 22 (8), 1303–1315.