Chapter 2. THEORY

The modeling concepts and terminology described in the this section have been more thoroughly described by Cooley (2004) and Christensen and Cooley (2003).

A ground-water flow system can be characterised by hydrogeologic variables (termed system characteristics) such as hydraulic head, hydraulic conductivity, recharge, discharge, hydraulic heads and fluxes along internal and external boundaries, and point sources and sinks. In a model these system characteristics can be conceptualized as being discretely variable (in space and/or time) because discrete variation can be at as small a scale as desired (depending only on model discretization), allowing it to be virtually the same as continuous variation. All of these system characteristics can be assembled into an m-vector $\boldsymbol{\beta}$. Because $\boldsymbol{\beta}$ includes all scales of variation necessary to construct an accurate model, any model function of $\boldsymbol{\beta}$ is almost free of model error, assuming that the model accurately represents the physical processes.

Seen from a practical modeling perspective, vector $\boldsymbol{\beta}$ represents small-scale variability that cannot be explicitly represented in a model and larger scale variability (the drift) that can. In addition, vector $\boldsymbol{\beta}$ is unknown. Because $\boldsymbol{\beta}$ has much too large a dimension to be estimated from a small number of uncertain observations, the vector of model parameters to be estimated is reduced from one of large dimension, $\boldsymbol{\beta}$, to one of much smaller dimension, $\boldsymbol{\theta}_*$. The p-vector $\boldsymbol{\theta}_*$ represents the spatial and temporal average of $\boldsymbol{\beta}$ and has the same form as the drift $\overline{\boldsymbol{\theta}}$; it does not represent the small-scale variation. A model function $f(\boldsymbol{\beta})$, which contains large- and small-scale variation from $\boldsymbol{\beta}$, is represented in terms of $\boldsymbol{\theta}_*$ as $f(\boldsymbol{\gamma}\boldsymbol{\theta}_*)$, where $\boldsymbol{\gamma}$ is an $m \times p$ interpolation or averaging matrix.

In ground-water modeling we are interested in estimating $\boldsymbol{\theta}_*$ rather than $\overline{\boldsymbol{\theta}}$ because $\boldsymbol{\theta}_*$ characterizes the actual realization $\boldsymbol{\beta}$. An estimate of $\boldsymbol{\theta}_*$, $\hat{\boldsymbol{\theta}}$, can be found by minimizing

$$S(\boldsymbol{\theta}) = \left(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta})\right)' \boldsymbol{\omega} \left(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta})\right), \tag{1}$$

where $\boldsymbol{Y}$ is an n-vector of observations ($m>>n>p$), $\boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta})$ is the corresponding n-vector of values simulated using general vector $\boldsymbol{\gamma}\boldsymbol{\theta}$ instead of $\boldsymbol{\beta}$, $\boldsymbol{\omega}$ is an $n \times n$ weight matrix, and the prime indicates transpose.

Predictions can be made with the ground-water model. In practice, variables of interest $g(\boldsymbol{\beta})$ are predicted by $g(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}})$.

Different types of observations, $\boldsymbol{Y}$, can be used in the Parameter Estimation Process. Cooley (2004) does not distinguish among the several possible types of observations such as hydraulic heads, measured streamflow gains and losses, measured spring and well discharges, and measured direct (often called “prior”) information on model parameters. It is convenient to divide the types of information into two groups: (1) measured direct observations of parameters or linear combinations of them and (2) observations corresponding to model functions of the parameters such as hydraulic heads and fluxes. The various matrices and vectors pertaining to

observations have a subscript *d* for the partition of direct information and a subscript *m* for the partition for model-function information. Corresponding to this notation, the number of direct observations is designated n_d and the number of model-function observations is designated n_m, for a total of $n_d+n_m=n$ observations. The matrices and vectors are used in the partitioned form in an equation only when the form of the result for each partition is different. The difference is caused by the forms of the observation second-moment matrix, $\boldsymbol{\Omega}\sigma_\varepsilon^2 = E(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\theta}_*))(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\theta}_*))'$, and the weight matrix, $\boldsymbol{\omega}$, used as an approximation for $\boldsymbol{\Omega}^{-1}$ in the weighted regression analysis. As in standard regression analysis, σ_ε^2 is a generally unknown scalar multiplier that is estimated by the regression. We assume that the partition of the observation second-moment matrix for the model functions depends on an unknown model error and often may not be known, but that the partition for the direct information would be well estimated using the same techniques used to acquire the direct information. We also assume that the two partitions are not correlated. (This is a standard assumption dating back to Theil (1963). Should it prove to be inaccurate in any application, the correction factors computed by CORFAC-2k will be somewhat in error. If $n_d<<n_m$, the errors should be small.) Thus, the forms for $\boldsymbol{\Omega}$ and $\boldsymbol{\omega}$ are

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\Omega}_m & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Omega}_d \end{bmatrix}, \tag{2}$$

where $\boldsymbol{0}$ is a matrix of zeros, and

$$\boldsymbol{\omega} = \begin{bmatrix} \boldsymbol{\omega}_m & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Omega}_d^{-1} \end{bmatrix}. \tag{3}$$

The partitioning also can be used for the $n\times n$ matrix $\boldsymbol{R}$ defined as

$$\boldsymbol{R} = \boldsymbol{\omega}^{1/2}\boldsymbol{Df}\left(\boldsymbol{Df}'\boldsymbol{\omega}\boldsymbol{Df}\right)^{-1}\boldsymbol{Df}'\boldsymbol{\omega}^{1/2} = \begin{bmatrix} \boldsymbol{R}_m & \boldsymbol{R}_{md} \\ \boldsymbol{R}_{dm} & \boldsymbol{R}_d \end{bmatrix} \tag{4}$$

in which $\boldsymbol{Df}$ is the $n\times p$ sensitivity matrix of derivatives of $\boldsymbol{f}$ with respect to $\boldsymbol{\theta}$,

$$\boldsymbol{Df} = \left[\frac{\partial f_i}{\partial \theta_j}\right] \tag{5}$$

theoretically evaluated at the drift set of parameters $\overline{\boldsymbol{\theta}} = E(\boldsymbol{\theta}_*)$, and

$$\left.\begin{aligned} \boldsymbol{R}_m &= \boldsymbol{\omega}_m^{1/2}\boldsymbol{Df}_m\left(\boldsymbol{Df}'\boldsymbol{\omega}\boldsymbol{Df}\right)^{-1}\boldsymbol{Df}_m'\boldsymbol{\omega}_m^{1/2} \\ \boldsymbol{R}_{md} &= \boldsymbol{\omega}_m^{1/2}\boldsymbol{Df}_m\left(\boldsymbol{Df}'\boldsymbol{\omega}\boldsymbol{Df}\right)^{-1}\boldsymbol{Df}_d'\boldsymbol{\omega}_d^{1/2} \\ \boldsymbol{R}_{dm} &= \boldsymbol{\omega}_d^{1/2}\boldsymbol{Df}_d\left(\boldsymbol{Df}'\boldsymbol{\omega}\boldsymbol{Df}\right)^{-1}\boldsymbol{Df}_m'\boldsymbol{\omega}_m^{1/2} \\ \boldsymbol{R}_d &= \boldsymbol{\omega}_d^{1/2}\boldsymbol{Df}_d\left(\boldsymbol{Df}'\boldsymbol{\omega}\boldsymbol{Df}\right)^{-1}\boldsymbol{Df}_d'\boldsymbol{\omega}_d^{1/2} \end{aligned}\right\}, \tag{6}$$

where $\boldsymbol{f}_m$ and $\boldsymbol{f}_d$ are model-function and direct-information partitions of $\boldsymbol{f}$, and $\boldsymbol{D}(\ldots)$ is the derivative operator applied as in equation 5 to yield the $n_m \times p$ and the $n_d \times p$ matrices $\boldsymbol{Df}_m$ and $\boldsymbol{Df}_d$. Similar partitioning is used for the n-vector $\boldsymbol{Q}$, defined as

$$\boldsymbol{Q} = \boldsymbol{\omega}^{1/2}\boldsymbol{Df}\left(\boldsymbol{Df}'\boldsymbol{\omega}\boldsymbol{Df}\right)^{-1}\boldsymbol{Dg}' = \begin{bmatrix} \boldsymbol{\omega}_m^{1/2}\boldsymbol{Df}_m\left(\boldsymbol{Df}'\boldsymbol{\omega}\boldsymbol{Df}\right)^{-1}\boldsymbol{Dg}' \\ \boldsymbol{\omega}_d^{1/2}\boldsymbol{Df}_d\left(\boldsymbol{Df}'\boldsymbol{\omega}\boldsymbol{Df}\right)^{-1}\boldsymbol{Dg}' \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}_m \\ \boldsymbol{Q}_d \end{bmatrix}, \tag{7}$$

where $\boldsymbol{Dg}' = \left[\partial g / \partial \theta_i\right]$ is the column p-vector of sensitivities of the prediction $g(\boldsymbol{\gamma\theta})$.

The UNC Process described in 'The UNC Process' section can be used to compute confidence or prediction intervals for parameters ($\hat{\boldsymbol{\theta}}$) of the Parameter-Estimation Process, and for most types of predictions ($g(\boldsymbol{\gamma\theta}_*)$ or $g(\boldsymbol{\beta})+\varepsilon$ where ε is a random observation error) that can be computed by a MODFLOW-2000 model calibrated by the Parameter-Estimation Process. The programs RESAN2-2k, BEALE2-2k, and CORFAC-2k that also are described in this report are valuable in the evaluation of results from the Parameter-Estimation Process and in the preparation of input values for UNC.

The UNC Process

The purpose of UNC is to compute confidence and prediction intervals for variables of the form $g(\boldsymbol{\gamma\theta}_*)$. Both the model functions $f(\boldsymbol{\gamma\theta}_*)$ and predictions $g(\boldsymbol{\gamma\theta}_*)$ can be nonlinear, and the intervals can be individual, Scheffé type, or, with correction factors not given in this report, other simultaneous types (Hill, 1994, p. 26-33). Calculation procedures are the same for all types of confidence intervals and for all types of prediction intervals; only the critical values described after equation 8 are different. Extensive information on the theory, calculation procedures, and use of confidence and prediction intervals can be found in Graybill (1976, chapters 6-10), Seber and Wild (1989, chapter 5), Hill (1994, p. 26-37), Cooley and Vecchia (1987), Vecchia and Cooley (1987), Cooley (1997), Christensen and Cooley (1999a, b), Cooley and Naff (1990, chapter 5), and Cooley (2004). Cooley (2004) summarizes much of the information found in the other references.

Confidence Intervals

A confidence interval for $g(\boldsymbol{\gamma\theta}_*)$ is defined by two confidence limits, the maximum and minimum values of $g(\boldsymbol{\gamma\theta})$ over a likelihood region. The definition of the region depends on the type of confidence interval being computed. For example, for a Scheffé interval the likelihood

region is the standard confidence region given in terms of $S(\hat{\boldsymbol{\theta}})$, and for an individual confidence interval the likelihood region is of the same form but has a smaller diameter (Cooley, 2004, p. 53). If we assume that the maximum and minimum values of $g(\boldsymbol{\gamma\theta})$ occur on the boundary of the likelihood region as they do for linear problems, then these extreme values can be found using the method of Lagrange multipliers (Vecchia and Cooley, 1987). Cooley (1999, p. 118) states that this assumption is almost non-restrictive; the assumption has never been violated in any of the calculations made by the authors of this report. Therefore, confidence and prediction intervals computed by the UNC Process are based on the method of Lagrange multipliers.

A $(1-\alpha)\times 100$ percent confidence interval is computed by finding extreme values (the confidence limits) of the Lagrange function

$$L(\boldsymbol{\theta},\lambda) = g(\boldsymbol{\gamma\theta}) + \frac{1}{2\lambda}\left(d_\alpha^2 - S(\boldsymbol{\theta}) + S(\hat{\boldsymbol{\theta}})\right), \tag{8}$$

where λ^{-1} is the Lagrange multiplier and d_α^2 is the critical value that depends on the type of confidence interval being computed. For an individual confidence interval $d_\alpha^2 = c_c t_{\alpha/2}^2\, S(\hat{\boldsymbol{\theta}})/(n-p)$ and for a Scheffé interval $d_\alpha^2 = c_r F_\alpha(p, n-p) S(\hat{\boldsymbol{\theta}})/(n-p)$. Here c_c and c_r are correction factors defined in 'The RESAN2-2k Program' section, $t_{\alpha/2}$ is the $(1-\alpha/2)\times 100$ percentile of the cumulative *t* distribution, $F_\alpha(p, n-p)$ is the upper α point of the *F* distribution with *p* and $n-p$ degrees of freedom, *p* is the number of model parameters, and *n* is the total number of observations used to estimate $\hat{\boldsymbol{\theta}}$. Theoretically, the equations to compute extreme values are obtained simply by taking the derivatives of $L(\boldsymbol{\theta},\lambda)$ with respect to $\boldsymbol{\theta}$ and λ^{-1}, then setting the results to zero. However, this yields a set of nonlinear equations to be solved whenever $g(\boldsymbol{\gamma\theta})$ or $\boldsymbol{f}(\boldsymbol{\gamma\theta})$ (or both) are nonlinear. Another method was developed by Vecchia and Cooley (1987) and involves linearizing the nonlinear functions $g(\boldsymbol{\gamma\theta})$ and (or) $\boldsymbol{f}(\boldsymbol{\gamma\theta})$ first, then taking the derivatives of $L(\boldsymbol{\theta},\lambda)$ and setting the results to zero. The resulting linear system of equations then can be written to yield an iterative solution to the nonlinear problem. The iterative solution is derived in appendix A; the solution $\boldsymbol{\theta}_{r+1}$ at iteration *r*+1 is given in terms of quantities computed at iteration *r* as

$$\lambda_{r+1} = \pm\left(\frac{d_\alpha^2 - S(\boldsymbol{\theta}_r) + S(\hat{\boldsymbol{\theta}}) + (\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\theta}_r))'\boldsymbol{\omega}^{1/2}\boldsymbol{R}_r\boldsymbol{\omega}^{1/2}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\theta}_r))}{\boldsymbol{Q}_r'\boldsymbol{Q}_r}\right)^{1/2} \tag{9}$$

and

$$\boldsymbol{\theta}_{r+1} = \boldsymbol{\theta}_r + \lambda_{r+1}(\boldsymbol{Df}_r'\boldsymbol{\omega}\boldsymbol{Df}_r)^{-1}\boldsymbol{D}g_r' + (\boldsymbol{Df}_r'\boldsymbol{\omega}\boldsymbol{Df}_r)^{-1}\boldsymbol{Df}_r'\boldsymbol{\omega}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\theta}_r)), \tag{10}$$

where subscript *r* indicates evaluation using parameter set $\boldsymbol{\theta}_r$ and $\boldsymbol{\theta}_0$ is the initial, user supplied, set. Use of the positive sign in equation 9 yields the maximum and use of the negative sign yields

the minimum. At convergence of the process $\boldsymbol{\theta}_r \approx \boldsymbol{\theta}_{r+1} \approx \widetilde{\boldsymbol{\theta}}$, the set of parameters defining either limit $g(\boldsymbol{\gamma}\widetilde{\boldsymbol{\theta}})$.

To control round-off errors resulting from large variations in magnitude of the elements of $\boldsymbol{Df}_r'\boldsymbol{\omega}\boldsymbol{Df}_r$ and to put equation 10 into a form for the Marquardt-like solution scheme developed by Vecchia and Cooley (1987), $\boldsymbol{Df}_r$, $\boldsymbol{Dg}_r$, $\boldsymbol{\theta}_r$, and $\boldsymbol{\theta}_{r+1}$ need to be scaled. Let $\boldsymbol{A}_r$ be a diagonal matrix composed of the square roots of the inverses of the diagonal elements of $\boldsymbol{Df}_r'\boldsymbol{\omega}\boldsymbol{Df}_r$. Then let

$$\boldsymbol{S}_r = \boldsymbol{Df}_r\boldsymbol{A}_r, \tag{11}$$

$$\boldsymbol{Z}_r = \boldsymbol{A}_r\boldsymbol{Dg}_r', \tag{12}$$

and

$$\boldsymbol{d}_{r+1} = \boldsymbol{A}_r^{-1}(\boldsymbol{\theta}_{r+1} - \boldsymbol{\theta}_r). \tag{13}$$

With the above scaled matrix and vectors, equation 10 transforms to

$$\boldsymbol{d}_{r+1} = \lambda_{r+1}(\boldsymbol{S}_r'\boldsymbol{\omega}\boldsymbol{S}_r)^{-1}\boldsymbol{Z}_r + (\boldsymbol{S}_r'\boldsymbol{\omega}\boldsymbol{S}_r)^{-1}\boldsymbol{S}_r'\boldsymbol{\omega}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta}_r)). \tag{14}$$

Note that $\boldsymbol{Q}_r$ and $\boldsymbol{R}_r$ in equation 9 are invariant under the scaling used for equations 11 and 12, so can be written in terms of $\boldsymbol{S}_r$ and $\boldsymbol{Z}_r$.

Vecchia and Cooley (1987) found that a Marquardt-like modification (Seber and Wild, 1989, p. 624) of equation 14 often was effective for solving ill-conditioned problems. We have used numerous tests to confirm this finding. The modification is simply to add $\mu_{r+1}\boldsymbol{I}$ to $\boldsymbol{S}_r'\boldsymbol{\omega}\boldsymbol{S}_r$ in equation 14 to obtain

$$\boldsymbol{d}_{r+1} = \lambda_{r+1}(\boldsymbol{S}_r'\boldsymbol{\omega}\boldsymbol{S}_r + \mu_{r+1}\boldsymbol{I})^{-1}\boldsymbol{Z}_r + (\boldsymbol{S}_r'\boldsymbol{\omega}\boldsymbol{S}_r + \mu_{r+1}\boldsymbol{I})^{-1}\boldsymbol{S}_r'\boldsymbol{\omega}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta}_r)), \tag{15}$$

where μ_{r+1} is a parameter obtained using the algorithm given by Cooley and Naff (1990, p. 71-72).

Another modification involves damping of the parameter displacement vector $\boldsymbol{d}_{r+1}$. This often is necessary to achieve convergence because $\boldsymbol{d}_{r+1}$ is frequently too large (Cooley and Vecchia, 1987). Thus, parameter set $\boldsymbol{\theta}_{r+1}$ is computed using a damped form of equation 13:

$$\boldsymbol{\theta}_{r+1} = \rho_{r+1}\boldsymbol{A}_r\boldsymbol{d}_{r+1} + \boldsymbol{\theta}_r. \tag{16}$$

Damping parameter ρ_{r+1} is computed using the algorithm given by Cooley (1993c, p. 2-3) except that t_{r+1} used by that algorithm is defined in the present report as the value of $\delta_j^{r+1}/\theta_{sj}$ for which $|t_{r+1}| = \max_j |\delta_j^{r+1}/\theta_{sj}|$, where δ_j^{r+1} is an element of $\boldsymbol{A}_r\boldsymbol{d}_{r+1}$, θ_{sj} is an element of $\boldsymbol{\theta}_s$, and $\boldsymbol{\theta}_s$ is a

user-supplied scaling vector. The scaling vector adjusts for differing sizes of elements of $\boldsymbol{\theta}$, so should reflect the user's best knowledge of $\boldsymbol{\theta}_*$, except that no element of $\boldsymbol{\theta}_s$ can be zero.

We consider the iteration scheme to have converged if either $|t_{r+1}|$ or $|S(\boldsymbol{\theta}_{r+1}) - S(\boldsymbol{\theta}_r)|/S(\boldsymbol{\theta}_r) + |S(\boldsymbol{\theta}_r) - S(\boldsymbol{\theta}_{r-1})|/S(\boldsymbol{\theta}_{r-1})$ is small enough. The convergence criteria for $|t_{r+1}|$ and $|S(\boldsymbol{\theta}_{r+1}) - S(\boldsymbol{\theta}_r)|/S(\boldsymbol{\theta}_r) + |S(\boldsymbol{\theta}_r) - S(\boldsymbol{\theta}_{r-1})|/S(\boldsymbol{\theta}_{r-1})$ are user supplied. Another test that overrides the above two also is used, if specified by the user. Sometimes $g(\boldsymbol{\gamma}\boldsymbol{\theta}_{r+1})$ can stabilize even though one or more elements of $\boldsymbol{\theta}$ have not converged. To address this case the user can specify that the process has converged whenever $|2 \times (g(\boldsymbol{\gamma}\boldsymbol{\theta}_{r+1}) - g(\boldsymbol{\gamma}\boldsymbol{\theta}_r))/(g(\boldsymbol{\gamma}\boldsymbol{\theta}_{r+1}) + g(\boldsymbol{\gamma}\boldsymbol{\theta}_r))|$ is smaller than another user-supplied criterion. This criterion is not used when $|(g(\boldsymbol{\gamma}\boldsymbol{\theta}_{r+1}) + g(\boldsymbol{\gamma}\boldsymbol{\theta}_r))| < 10^{-12}$.

Weighted Residuals

Cooley (2004, p. 56-57) showed that weighted residuals from a regression constrained so that $g(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}}) = g(\boldsymbol{\gamma}\boldsymbol{\theta}_*)$ (where $\tilde{\boldsymbol{\theta}}$ is the constrained regression estimate) can be defined as $(\boldsymbol{I} - \boldsymbol{Q}\boldsymbol{Q}'/\boldsymbol{Q}'\boldsymbol{Q})\boldsymbol{\omega}^{1/2}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}}))$ and are equal to weighted residuals $\boldsymbol{\omega}^{1/2}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}))$ from the unconstrained regression if model intrinsic nonlinearity and model combined intrinsic nonlinearity as defined by Cooley (2004, p. 35-36) both are small. Parameter set $\tilde{\boldsymbol{\theta}}$ can be the set computed for either confidence limit because minimization of $S(\boldsymbol{\theta})$, subject to the constraint $g(\boldsymbol{\gamma}\boldsymbol{\theta}) = g(\boldsymbol{\gamma}\boldsymbol{\theta}_*)$, is the same as finding an extreme value of $g(\boldsymbol{\gamma}\boldsymbol{\theta})$, subject to the constraint $S(\boldsymbol{\theta}) - S(\hat{\boldsymbol{\theta}}) = d_\alpha^2$, if $g(\boldsymbol{\gamma}\boldsymbol{\theta}_*)$ assumed for the constrained regression is the confidence limit for $g(\boldsymbol{\gamma}\boldsymbol{\theta}_*)$ computed using d_α^2. Thus, the constrained weighted residuals for both confidence limits for $g(\boldsymbol{\gamma}\boldsymbol{\theta}_*)$ are computed and printed if specified by the user. Vector $\boldsymbol{Q}$ is computed using the initial set of parameters $\boldsymbol{\theta}_0$ because if $\tilde{\boldsymbol{\theta}}$ were used, the term in $(\boldsymbol{I} - \boldsymbol{Q}\boldsymbol{Q}'/\boldsymbol{Q}'\boldsymbol{Q})\boldsymbol{\omega}^{1/2}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}}))$ expressing model combined intrinsic nonlinearity, that is $(\boldsymbol{R} - \boldsymbol{Q}\boldsymbol{Q}'/\boldsymbol{Q}'\boldsymbol{Q})\boldsymbol{\omega}^{1/2}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}}))$, always would be zero, as shown in appendix B.

Prediction Intervals

As for a confidence interval, a prediction interval for a predicted observation of $g(\boldsymbol{ß})$, Y_p, is defined by two limits, the maximum and minimum values of $g(\boldsymbol{\gamma}\boldsymbol{\theta}) + v$ over a likelihood region, where v is the predicted error (Cooley, 2004, p. 59-60). The form of the likelihood region is developed in Cooley (2004, p. 61) and the prediction limits are assumed to lie on the edge of the likelihood region so that the method of Lagrange multipliers can again be used to find the extreme values. To obtain the Lagrange function we assume that the weight matrix incorporating the prediction, $\boldsymbol{\omega}_a$ of Cooley (2004, p. 63), is block diagonal of the form

$$\boldsymbol{\omega}_a = \begin{bmatrix} \boldsymbol{\omega} & \boldsymbol{0} \\ \boldsymbol{0} & \omega_p \end{bmatrix} \tag{17}$$

as described by Cooley (2004, p. 63). In equation 17, ω_p is the weight for the prediction that is analogous to diagonal elements of $\boldsymbol{\omega}$. The form given by equation 17 not only simplifies calculations, but also does not require the user to estimate second moments between observations and the prediction, which often would be unknown. However, if such second moments actually exist so that the second-moment matrix that includes the prediction is of the form (Cooley, 2004, equation 5-87)

$$\boldsymbol{\Omega}_a = \begin{bmatrix} \boldsymbol{\Omega} & \boldsymbol{C} \\ \boldsymbol{C}' & b\hat{\omega}_p^{-1} \end{bmatrix}, \tag{18}$$

then a correction factor c_p, described in 'The RESAN2-2k Program' section, must be used. In equation 18, $\boldsymbol{C} = E\left(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\theta}_*)\right)\left(Y_p - g(\boldsymbol{\gamma\theta}_*)\right)/\sigma_\varepsilon^2$, $\hat{\omega}_p^{-1} = E\left(Y_p - g(\boldsymbol{\gamma\theta}_*)\right)^2/\sigma_\varepsilon^2$, and $b = tr\left(\boldsymbol{\omega}^{1/2}\boldsymbol{\Omega}\boldsymbol{\omega}^{1/2}\right)/n$, where $tr(\ldots)$ signifies matrix trace.

On the basis of the assumed structure of $\boldsymbol{\omega}_a$, Cooley (2004, p. 65) stated the Lagrange function as

$$L(\boldsymbol{\theta}, v, \lambda) = g(\boldsymbol{\gamma\theta}) + v + \frac{1}{2\lambda}\left(d_\alpha^2 - S(\boldsymbol{\theta}) - \omega_p v^2 + S\left(\hat{\boldsymbol{\theta}}\right)\right), \tag{19}$$

where critical value d_α^2 is equal to $c_p t_{\alpha/2}^2 (n-p) S\left(\hat{\boldsymbol{\theta}}\right)/(n-p)$ for an individual prediction interval. As explained by Cooley (2004, p. 65), sometimes the second moment of $Y_p - g(\boldsymbol{\gamma\theta}_*)$, $E\left(Y_p - g(\boldsymbol{\gamma\theta}_*)\right)^2$, can be estimated more easily than $\hat{\omega}_p$ for use in equation 19. In this case, the term $E\left(Y_p - g(\boldsymbol{\gamma\theta}_*)\right)^{-2} s^2$ can replace the general weight ω_p in equation 19.

Iterative solution of the extreme value problem based on equation 19 is derived in appendix A using the same method as used for equation 8. The solution is given as

$$\lambda_{r+1} = \pm\left(\frac{d_\alpha^2 - S(\boldsymbol{\theta}_r) + S\left(\hat{\boldsymbol{\theta}}\right) + \left(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\theta}_r)\right)' \boldsymbol{\omega}^{1/2} \boldsymbol{R}_r \boldsymbol{\omega}^{1/2} \left(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\theta}_r)\right)}{\boldsymbol{Q}_r' \boldsymbol{Q}_r + \omega_p^{-1}}\right)^{1/2}, \tag{20}$$

$$v_{r+1} = \omega_p^{-1}\lambda_{r+1}, \tag{21}$$

and equation 10. All of the developments following equation 10 for a confidence interval are used for a prediction interval.

The RESAN2-2k Program

The purpose of analyzing residuals is to test whether or not the assumptions made for nonlinear regression and uncertainty analysis seem to be violated (Cooley and Naff, 1990, p. 167). RESAN2-2k focuses on detection of model and system types of intrinsic nonlinearity as defined by Cooley (2004, p. 35-36), as well as the traditional graphical examination of residuals

plots for lack of model fit and testing of residuals for indications of non-normality. More on the purposes, theory, and methods of residuals analysis is given in Draper and Smith (1998, chapters 2, 7, and 8), Cooley and Naff (1990, chapter 5), Hill (1992, p. 56-66, 88-89), Hill (1994), Cooley (2004), and references cited in these works.

A Test for Model Intrinsic Nonlinearity

Model intrinsic nonlinearity causes confidence regions, confidence intervals, and prediction intervals to be too small unless they are corrected with a correction factor. Cooley (2004, section 5) showed that (1) the size of a confidence region is proportional to $c_r F_\alpha(p, n-p)$, where c_r is the correction factor; (2) the size of a Scheffé interval is proportional to $\left(c_r F_\alpha(p, n-p)\right)^{1/2}$; (3) the size of an individual confidence interval is proportional to $\left(c_c t^2_{\alpha/2}(n-p)\right)^{1/2}$, where c_c is the correction factor; and (4) the size of an individual prediction interval is proportional to $\left(c_p t^2_{\alpha/2}(n-p)\right)^{1/2}$, where c_p is the correction factor. The correction factors are defined by

$$c_r = \frac{\sigma_\varepsilon^2 + \left(\gamma_{wr}\sigma_\beta^2 + \gamma_{Ir}\sigma_\varepsilon^4\right)/p}{\sigma_\varepsilon^2 + \left(\hat{\gamma}_w\sigma_\beta^2 + \hat{\gamma}_I\sigma_\varepsilon^4\right)/(n-p)}, \tag{22}$$

$$c_c = \frac{\sigma_\varepsilon^2 + \gamma_{wc}\sigma_\beta^2 + \gamma_{Ic}\sigma_\varepsilon^4}{\sigma_\varepsilon^2 + \left(\hat{\gamma}_w\sigma_\beta^2 + \hat{\gamma}_I\sigma_\varepsilon^4\right)/(n-p)}, \tag{23}$$

and, for the form of prediction interval computed by UNC,

$$c_p = \frac{\sigma_\varepsilon^2 + \gamma_{wa}\sigma_\beta^2 + \gamma_{Ia}\sigma_\varepsilon^4}{\sigma_\varepsilon^2 + \left(\hat{\gamma}_w\sigma_\beta^2 + \hat{\gamma}_I\sigma_\varepsilon^4\right)/(n-p)}. \tag{24}$$

Variables contained in equations 22-24 are defined as follows: σ_β^2 is a scalar multiplier for the variance of the vector of system characteristics, $\boldsymbol{\beta}$, defined by $Var(\boldsymbol{\beta}) = \boldsymbol{V}_\beta\sigma_\beta^2$; σ_ε^2 is a scalar multiplier for the variance of the observation-error vector, $\boldsymbol{\varepsilon}$, of order *n*, defined by $Var(\boldsymbol{\varepsilon}) = \boldsymbol{V}_\varepsilon\sigma_\varepsilon^2$; and $\hat{\gamma}_w\sigma_\beta^2, \hat{\gamma}_I\sigma_\varepsilon^4, \gamma_{wr}\sigma_\beta^2, \gamma_{Ir}\sigma_\varepsilon^4, \gamma_{wc}\sigma_\beta^2, \gamma_{Ic}\sigma_\varepsilon^4, \gamma_{wa}\sigma_\beta^2,$ and $\gamma_{Ia}\sigma_\varepsilon^4$ are component correction factors defined by Cooley (2004, section 5). (The subscripts *r* and *c* were added in the present report.)

The factors examined by RESAN2-2k, $\hat{\gamma}_I\sigma_\varepsilon^4$ and $\gamma_{Ir}\sigma_\varepsilon^4$, measure the importance of model intrinsic nonlinearity. Cooley (2004, p. 50) showed that $\gamma_{Ir}\sigma_\varepsilon^4 = -\hat{\gamma}_I\sigma_\varepsilon^4$. Thus, only $\hat{\gamma}_I\sigma_\varepsilon^4$ needs to be computed, and this factor is expressed by Cooley (2004, equation 6-3) as

$$\begin{aligned}
&\hat{\gamma}_I \sigma_\varepsilon^4 = 2E\left(\boldsymbol{f}(\boldsymbol{\gamma\theta}_*) - \boldsymbol{f}_0(\boldsymbol{\gamma\theta}_*) - \boldsymbol{f}\left(\boldsymbol{\gamma\hat{\theta}}\right) + \boldsymbol{f}_0\left(\boldsymbol{\gamma\hat{\theta}}\right)\right)' \boldsymbol{\omega}^{1/2}(\boldsymbol{I} - \boldsymbol{R})\boldsymbol{\omega}^{1/2}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\theta}_*)) \\
&+ E\left(\boldsymbol{f}(\boldsymbol{\gamma\theta}_*) - \boldsymbol{f}_0(\boldsymbol{\gamma\theta}_*) - \boldsymbol{f}\left(\boldsymbol{\gamma\hat{\theta}}\right) + \boldsymbol{f}_0\left(\boldsymbol{\gamma\hat{\theta}}\right)\right)' \boldsymbol{\omega}^{1/2}(\boldsymbol{I} - \boldsymbol{R})\boldsymbol{\omega}^{1/2} \cdot \\
&\left(\boldsymbol{f}(\boldsymbol{\gamma\theta}_*) - \boldsymbol{f}_0(\boldsymbol{\gamma\theta}_*) - \boldsymbol{f}\left(\boldsymbol{\gamma\hat{\theta}}\right) + \boldsymbol{f}_0\left(\boldsymbol{\gamma\hat{\theta}}\right)\right) + E\left(\boldsymbol{Y} - \boldsymbol{f}\left(\boldsymbol{\gamma\hat{\theta}}\right)\right)' \boldsymbol{\omega}^{1/2}\boldsymbol{R}\boldsymbol{\omega}^{1/2}\left(\boldsymbol{Y} - \boldsymbol{f}\left(\boldsymbol{\gamma\hat{\theta}}\right)\right) \\
&= 2E\left(\boldsymbol{f}_m(\boldsymbol{\gamma\theta}_*) - \boldsymbol{f}_{0m}(\boldsymbol{\gamma\theta}_*) - \boldsymbol{f}_m\left(\boldsymbol{\gamma\hat{\theta}}\right) + \boldsymbol{f}_{0m}\left(\boldsymbol{\gamma\hat{\theta}}\right)\right)' \boldsymbol{\omega}_m^{1/2} \cdot \\
&\left((\boldsymbol{I}_m - \boldsymbol{R}_m)\boldsymbol{\omega}_m^{1/2}(\boldsymbol{Y}_m - \boldsymbol{f}_m(\boldsymbol{\gamma\theta}_*)) - \boldsymbol{R}_{md}\boldsymbol{\omega}_d^{1/2}(\boldsymbol{Y}_d - \boldsymbol{f}_d(\boldsymbol{\gamma\theta}_*))\right) \\
&+ E\left(\boldsymbol{f}_m(\boldsymbol{\gamma\theta}_*) - \boldsymbol{f}_{0m}(\boldsymbol{\gamma\theta}_*) - \boldsymbol{f}_m\left(\boldsymbol{\gamma\hat{\theta}}\right) + \boldsymbol{f}_{0m}\left(\boldsymbol{\gamma\hat{\theta}}\right)\right)' \boldsymbol{\omega}_m^{1/2}(\boldsymbol{I}_m - \boldsymbol{R}_m)\boldsymbol{\omega}_m^{1/2} \cdot \\
&\left(\boldsymbol{f}_m(\boldsymbol{\gamma\theta}_*) - \boldsymbol{f}_{0m}(\boldsymbol{\gamma\theta}_*) - \boldsymbol{f}_m\left(\boldsymbol{\gamma\hat{\theta}}\right) + \boldsymbol{f}_{0m}\left(\boldsymbol{\gamma\hat{\theta}}\right)\right) + E\left(\boldsymbol{Y} - \boldsymbol{f}\left(\boldsymbol{\gamma\hat{\theta}}\right)\right)' \boldsymbol{\omega}^{1/2}\boldsymbol{R}\boldsymbol{\omega}^{1/2}\left(\boldsymbol{Y} - \boldsymbol{f}\left(\boldsymbol{\gamma\hat{\theta}}\right)\right),
\end{aligned} \tag{25}$$

where $\boldsymbol{f}_0(\boldsymbol{\gamma\theta}_*)$ is the linear-model approximation of $\boldsymbol{f}(\boldsymbol{\gamma\theta}_*)$ defined in Cooley (2004, equation 5-104 and text preceding equation 6-2), $\boldsymbol{f}_0\left(\boldsymbol{\gamma\hat{\theta}}\right)$ is the linear-model approximation of $\boldsymbol{f}\left(\boldsymbol{\gamma\hat{\theta}}\right)$ defined in Cooley (2004, equation 5-104 and text preceding equation 6-2), and $\boldsymbol{I}$ is the identity matrix.

The last term of equation 25 is the expected value of the product of row vector $(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\hat{\theta}}))'\boldsymbol{\omega}^{1/2}\boldsymbol{R}$ and its transpose. If $\boldsymbol{Df}$ in $\boldsymbol{R}$ is evaluated using $\hat{\boldsymbol{\theta}}$, then $\boldsymbol{R}\boldsymbol{\omega}^{1/2}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\hat{\theta}}))$ is always $\boldsymbol{0}$ (Cooley, 2004, p. 39). Also, if model intrinsic nonlinearity is absent, then $\boldsymbol{R}$ is constant (that is, the same for any set $\boldsymbol{\theta}$) (Cooley, 2004, p. 39). Thus, a test for the importance of intrinsic nonlinearity in the denominators of the correction factors is to compare the estimate of the last term of equation 25, $(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\hat{\theta}}))'\boldsymbol{\omega}^{1/2}\boldsymbol{R}\boldsymbol{\omega}^{1/2}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma\hat{\theta}}))$, where $\boldsymbol{R}$ is not computed using $\hat{\boldsymbol{\theta}}$, with another appropriate term in equations 22-24. This term is obtained from equation 5-16 of Cooley (2004, p. 48), which shows that model intrinsic nonlinearity is negligible in the denominator of the correction factors if $\left|\hat{\gamma}_I \sigma_\varepsilon^4\right| << (n-p)\sigma_\varepsilon^2 + \hat{\gamma}_w \sigma_\beta^2 = b(n-ap)\sigma_\varepsilon^2$, where a and b are factors that equal 1 if $b\boldsymbol{\omega}^{-1} = \boldsymbol{\Omega}$, the inverse of the correct weight matrix for the Gauss-Markov method of weighted regression to find $\hat{\boldsymbol{\theta}}$ (Cooley, 2004, section 4). Specifically,

$$a = tr\left(\boldsymbol{R}(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{\Omega}(\boldsymbol{\omega}/b)^{1/2}\right)/p = \left(tr\left(\boldsymbol{R}_m(\boldsymbol{\omega}_m/b)^{1/2}\boldsymbol{\Omega}_m(\boldsymbol{\omega}_m/b)^{1/2}\right) + tr(\boldsymbol{R}_d)/b\right)/p \tag{26}$$

and

$$b = tr\left(\boldsymbol{\omega}^{1/2}\boldsymbol{\Omega}\boldsymbol{\omega}^{1/2}\right)/n = \left(tr\left(\boldsymbol{\omega}_m^{1/2}\boldsymbol{\Omega}_m\boldsymbol{\omega}_m^{1/2}\right) + n_d\right)/n\,. \tag{27}$$

Because $b\sigma_\varepsilon^2$ is estimated by s^2 defined by (Cooley, 2004, p. 50) as

$$s^2 = \frac{S\left(\hat{\boldsymbol{\theta}}\right)}{n - ap}, \tag{28}$$

$b(n-ap)\sigma_\varepsilon^2$ is estimated by $S\left(\hat{\boldsymbol{\theta}}\right)$, where

$$S(\hat{\boldsymbol{\theta}}) = \left(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}})\right)' \boldsymbol{\omega} \left(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}})\right). \tag{29}$$

Thus, the test is that $(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}))'\boldsymbol{\omega}^{1/2}\boldsymbol{R}\boldsymbol{\omega}^{1/2}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}))$ should be much smaller than $S\left(\hat{\boldsymbol{\theta}}\right)$ if $\boldsymbol{R}$ is computed using some reasonable set of parameters $\boldsymbol{\theta}$ not equal to $\hat{\boldsymbol{\theta}}$. A set of direct observations or initial estimates for $\boldsymbol{\theta}_*$ should work. If desired, vector $\boldsymbol{R}\boldsymbol{\omega}^{1/2}\left(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}})\right)$ can be examined to determine where in observation space the model intrinsic nonlinearity is large.

Because $\gamma_{Ir}\sigma_\varepsilon^4 = -\hat{\gamma}_I\sigma_\varepsilon^4$, the test also can be used to assess the importance of model intrinsic nonlinearity in the numerator of c_r, except that the test criterion should be changed. From Cooley (2004, p. 48) model intrinsic nonlinearity is negligible in the numerator of c_r when $\left|\gamma_{Ir}\sigma_\varepsilon^4\right| << p\sigma_\varepsilon^2 + \gamma_w\sigma_\beta^2 = bap\sigma_\varepsilon^2$. The term $bap\sigma_\varepsilon^2$ is estimated by $aps^2 = apS\left(\hat{\boldsymbol{\theta}}\right)/(n - ap)$, so the test is that $(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}))'\boldsymbol{\omega}^{1/2}\boldsymbol{R}\boldsymbol{\omega}^{1/2}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}))$ should be much smaller than $apS\left(\hat{\boldsymbol{\theta}}\right)/(n - ap)$. When a is unknown, the more conservative criterion $pS\left(\hat{\boldsymbol{\theta}}\right)/(n - p)$ should be used.

A Test for Model and System Types of Intrinsic Nonlinearity

Cooley (2004, p. 38-39) showed that both the mean weighted residual $\sum_i \boldsymbol{\omega}_i^{1/2}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}))/n$ (where $\boldsymbol{\omega}_i^{1/2}$ is row i of $\boldsymbol{\omega}^{1/2}$) and the slope of the plot of weighted residuals $\boldsymbol{\omega}_i^{1/2}(\boldsymbol{Y} - \boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}))$ in relation to weighted model functions $\boldsymbol{\omega}_i^{1/2}\boldsymbol{f}\left(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}\right)$ should be small if the model $\boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}})$ is adequate and both model and system types of intrinsic nonlinearity are small. Formal t tests could be used to determine the significance of these measures, but graphical examination as performed by Cooley (2004, section 7) on test problems and comparisons with synthetic residual sets as described next should be adequate to detect significant model and system types of intrinsic nonlinearity as well as significant model inadequacy as described in the references cited at the beginning of this section.

A Test for Normality of the Weighted Residuals

Examination of the probability plot of weighted residuals $\boldsymbol{\omega}_i^{1/2}\left(\boldsymbol{Y} - \boldsymbol{f}\left(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}\right)\right)$ will sometimes reveal obvious departures from normality. However, as discussed by Cooley and Naff (1990, p. 168-170), departures from a standard normal distribution are expected for weighted residuals because they always are correlated and heteroscedastic. Under ideal conditions of no model or system intrinsic nonlinearity and zero-mean normal distributions of $\boldsymbol{\beta}$ and $\boldsymbol{\varepsilon}$, Cooley (2004, p. 38) showed that the weighted residuals have the normal distribution

$$\boldsymbol{\omega}^{1/2}\left(\boldsymbol{Y} - \boldsymbol{f}\left(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}\right)\right) \sim N\left(\boldsymbol{0}, (\boldsymbol{I} - \boldsymbol{R})(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{\Omega}(\boldsymbol{\omega}/b)^{1/2}(\boldsymbol{I} - \boldsymbol{R})b\sigma_\varepsilon^2\right). \tag{30}$$

If Gauss-Markov estimation is used so that $b\boldsymbol{\omega}^{-1} = \boldsymbol{\Omega}$, then equation 30 becomes

$$\boldsymbol{\omega}^{1/2}\left(\boldsymbol{Y} - \boldsymbol{f}\left(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}\right)\right) \sim N\left(\boldsymbol{0}, (\boldsymbol{I} - \boldsymbol{R})b\sigma_\varepsilon^2\right), \tag{31}$$

which is of the form of the distribution of standard linear-regression weighted residuals.

The weighted residual distribution can be graphically compared with synthetic weighted residual distributions generated according to equation 30 (or 31) as described by Cooley and Naff (1990, p. 168-171) to determine whether or not the weighted residuals appear to have the specified normal distribution. However, a more precise test also can be used. To make the test, first synthetic weighted residuals are generated from equation 30 or 31 (with the estimate s^2 replacing $b\sigma_\varepsilon^2$) as follows.

1. Standard normal deviates are generated from $\boldsymbol{u} \sim N(\boldsymbol{0}, \boldsymbol{I}s^2)$. Vector $\boldsymbol{u}_T$ is initialized as $\boldsymbol{u}_T = \boldsymbol{u}$.
2. If specified in the input, $\boldsymbol{u}$ is modified to be the synthetic, weighted, true error vector $\boldsymbol{u}_T = \boldsymbol{L}\boldsymbol{u}$, where $\boldsymbol{L}$ is the lower triangular Cholesky factor of $(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{\Omega}(\boldsymbol{\omega}/b)^{1/2}$. (See Kitanidis (1997, p. 237-238).)
3. Vector $\boldsymbol{u}_T$ is modified according to equation 30 (or 31) to give $\boldsymbol{d}^k = (\boldsymbol{I} - \boldsymbol{R})\boldsymbol{u}_T$, the set of synthetic, weighted residuals for realization *k*.
4. Steps 1-3 are repeated for a large number of realizations, *M*.

Note that skipping step 2 yields synthetic weighted residuals distributed according to equation 31 (with the estimate s^2 replacing $b\sigma_\varepsilon^2$). This step would be skipped if Gauss-Markov estimation were used or if $\boldsymbol{\Omega}$ were unknown, in which case $b\sigma_\varepsilon^2$ would have to be estimated by $S(\hat{\boldsymbol{\theta}})/(n-p)$ or *a* in the estimate s^2 would have to be approximated using equation 76 in 'The CORFAC-2k Program' section of this report.

Next, the following quantities are computed:

$$\bar{\boldsymbol{d}} = \sum_k \boldsymbol{d}^k / M, \tag{32}$$

the mean *n*-vector over all *M* realizations,

$$\boldsymbol{s} = \left[\sum_k d_i^k d_i^k \right], \tag{33}$$

the vector of sum of squared values of d_i^k over all realizations, and

$$\boldsymbol{v} = \left[\sqrt{(s_i - M\bar{d}_i\bar{d}_i)/(M-1)} \right], \tag{34}$$

the sample standard deviation *n*-vector.

A normal probability plot of vectors $\bar{\boldsymbol{d}} \pm 2\boldsymbol{v}$ gives a band within which the weighted residuals usually might be expected to lie (Cooley, 2004, section 7). Even though $\bar{d}_i \pm 2v_i$ defines an approximate 95 percent individual confidence interval, interpretation of the plot is

limited because each interval is an individual one so that all intervals in the band cannot be interpreted simultaneously. A better test for normality is obtained as follows as a generalization of a test given by Shapiro and Francia (1972) and used by Hill (1992, p. 63). First, the grand mean is computed:

$$\overline{\overline{d}} = \sum_i \overline{d}_i / n , \tag{35}$$

Next the square of the correlation between synthetic weighted residuals for each realization k and the means over all realizations is obtained:

$$c^k = \frac{\left(\sum_i \left(d_i^k - \sum_j d_j^k / n \right) \left(\overline{d}_i - \overline{\overline{d}} \right) \right)^2}{\sum_i \left(d_i^k - \sum_j d_j^k / n \right)^2 \sum_i \left(\overline{d}_i - \overline{\overline{d}} \right)^2} . \tag{36}$$

Then, all values of c^k are ordered from smallest to largest.

By definition, the probability that c^k has some specified value c_s or a smaller value is

$$P = \sum_k I\left(c_s \le c^k\right) / M , \tag{37}$$

where $I\left(c_s \le c^k\right)$ is the indicator function $I = 0$ when $c_s > c^k$ and $I = 1$ when $c_s \le c^k$. Therefore, a 95 percentile is given by $c_s = c^{k_1}$, where $k_1 = \text{int}(0.95M)$ and $\text{int}(\cdots)$ indicates truncation to the nearest integer. Similar 90 and 99 percentiles also can be computed. These percentiles should be compared to c_d, the square of the correlation between weighted residuals $\hat{\boldsymbol{e}}_i = \boldsymbol{\omega}_i^{1/2} \left(\boldsymbol{Y} - \boldsymbol{f}\left(\boldsymbol{\gamma} \hat{\boldsymbol{\theta}} \right) \right)$ and the means $\overline{d}_i$, which is

$$c_d = \frac{\left(\sum_i \left(\hat{e}_i - \sum_j \hat{e}_j / n \right) \left(\overline{d}_i - \overline{\overline{d}} \right) \right)^2}{\sum_i \left(\hat{e}_i - \sum_j \hat{e}_j / n \right)^2 \sum_i \left(\overline{d}_i - \overline{\overline{d}} \right)^2} . \tag{38}$$

The comparison will tell if the weighted residuals are significantly different from the specified normal distribution at some predetermined level of significance equal to 100 minus the selected percentile. Also, by letting $c_s = c_d$, the probability P of c_d or a smaller value can be obtained, which is a more direct test.

The BEALE2-2k Program

The purpose of BEALE2-2k is to compute measures of total model nonlinearity, model intrinsic nonlinearity, and model combined intrinsic nonlinearity in the vicinity of $\hat{\boldsymbol{\theta}}$. These measures relate to component correction factors $\hat{\gamma}_I\sigma_\varepsilon^4, \gamma_{Ir}\sigma_\varepsilon^4, \gamma_{Ic}\sigma_\varepsilon^4$, and $\gamma_{Ia}\sigma_\varepsilon^4$ that compose correction factors c_r, c_c, and c_p.

A Measure of Average Total Model Nonlinearity

Total model nonlinearity is the sum of model intrinsic nonlinearity and parameter effects nonlinearity, the latter of which is nonlinearity that can be eliminated by some generally unknown transformation of parameters $\boldsymbol{\varphi}(\boldsymbol{\theta})$ (Draper and Smith, 1998, p. 528-529). In effect, intrinsic nonlinearity is the smallest possible total model nonlinearity for any transformation of parameters. The measure of total nonlinearity was developed to determine when linear theory can be used to compute confidence regions (Beale, 1960; Guttman and Meeter, 1965; Cooley and Naff, 1990, p. 187-188).

Cooley and Naff (1990, p. 187-188) defined a measure of average total nonlinearity as a combination of measures developed by Beale (1960) and Linssen (1975). For reasons discussed by Cooley (2004, p. 85) the measure used in this report was altered slightly from the one defined by Cooley and Naff (1990). It is defined as the average weighted sum of squared discrepancies between nonlinear and linear model values on the edge of the linear probability region having diameter $apb\sigma_\varepsilon^2$ and centered on the drift set of parameters $\overline{\boldsymbol{\theta}}$. Therefore, the measure of average total model nonlinearity is calculated using estimates as (Cooley, 2004, p. 87)

$$\begin{aligned}\hat{N} &= \frac{1}{aps^2}\sum_{l=1}^{2p}\left(\boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta}_l)-\boldsymbol{f}_0(\boldsymbol{\gamma}\boldsymbol{\theta}_l)\right)'\boldsymbol{\omega}\left(\boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta}_l)-\boldsymbol{f}_0(\boldsymbol{\gamma}\boldsymbol{\theta}_l)\right)\Big/(2p)\\ &= \frac{1}{aps^2}\sum_{l=1}^{2p}\left(\boldsymbol{f}_m(\boldsymbol{\gamma}\boldsymbol{\theta}_l)-\boldsymbol{f}_{0m}(\boldsymbol{\gamma}\boldsymbol{\theta}_l)\right)'\boldsymbol{\omega}_m\left(\boldsymbol{f}_m(\boldsymbol{\gamma}\boldsymbol{\theta}_l)-\boldsymbol{f}_{0m}(\boldsymbol{\gamma}\boldsymbol{\theta}_l)\right)\Big/(2p),\end{aligned} \tag{39}$$

where $\boldsymbol{f}_0(\cdots)$ and $\boldsymbol{f}_{0m}(\cdots)$ indicate linear model values, and the sets $\boldsymbol{\theta}_l, l=1,2,\cdots,2p,$ are calculated as given by Cooley and Naff (1990, p. 189):

$$\boldsymbol{\theta}_l = \hat{\boldsymbol{\theta}} + \lambda\left(\boldsymbol{D}\hat{\boldsymbol{f}}'\boldsymbol{\omega}\boldsymbol{D}\hat{\boldsymbol{f}}\right)^{-1}\boldsymbol{l}', \tag{40}$$

in which

$$\lambda = \pm\left(\frac{aps^2}{\boldsymbol{l}\left(\boldsymbol{D}\hat{\boldsymbol{f}}'\boldsymbol{\omega}\boldsymbol{D}\hat{\boldsymbol{f}}\right)^{-1}\boldsymbol{l}'}\right)^{1/2} \tag{41}$$

and $\boldsymbol{l}$ is the row p-vector

$$\boldsymbol{l} = [0 \;\; 0 \;\cdots\; 1 \;\cdots\; 0], \tag{42}$$

with the 1 being in column l. Sensitivity matrix $\boldsymbol{D\hat{f}}$ is evaluated at $\hat{\boldsymbol{\theta}}$. Note that the diameter of the linear confidence region aps^2 can be written in terms of c_r as evaluated in CORFAC-2k (which assumes no model intrinsic nonlinearity) as

$$aps^2 = \frac{apS(\hat{\boldsymbol{\theta}})}{n-ap} = \frac{(n-p)a}{n-ap}\, p\, \frac{S(\hat{\boldsymbol{\theta}})}{n-p} = c_r p \frac{S(\hat{\boldsymbol{\theta}})}{n-p}. \tag{43}$$

If $\boldsymbol{\Omega}$ is unknown so that a cannot be calculated using equation 26, then a may be estimated using equation 76 derived in 'The CORFAC-2k Program' section.

By analogy with the standard criteria for ranking nonlinearity first given by Beale (1960, p. 60) and later augmented by Cooley and Naff (1990, p. 189), we consider the model to be highly nonlinear if $\hat{N} > 1$ so that the numerator of $\hat{N}$ is greater than the diameter aps^2 of the confidence region; nonlinear if $1 \geq \hat{N} > 0.09$; moderately nonlinear if $0.09 \geq \hat{N} > 0.01$; and essentially linear if $\hat{N} \leq 0.01$. Linear theory for computing confidence regions seems to produce good approximate results when $\hat{N} \leq 0.09$ (Cooley and Naff, 1990, p. 189).

A Measure of Average Model Intrinsic Nonlinearity

Average model intrinsic nonlinearity is assessed by the measure $\hat{N}_{\min}$ modified by Cooley (2004, p. 85) from forms given earlier by Beale (1960) and Linssen (1975). As for the measure of total nonlinearity, this measure is defined in terms of the linear probability region centered on $\overline{\boldsymbol{\theta}}$ and having diameter $apb\sigma_\varepsilon^2$. It is calculated using the same estimates and linear confidence region as used for $\hat{N}$. Therefore, the model intrinsic nonlinearity measure is calculated as (Cooley, 2004, p. 85)

$$\begin{aligned}
\hat{N}_{\min} &= \frac{1}{aps^2} \sum_{l=1}^{2p} \left(\boldsymbol{f}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_0(\boldsymbol{\gamma\theta}_l) - \boldsymbol{D\hat{f}\psi}_l\right)' \boldsymbol{\omega} \left(\boldsymbol{f}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_0(\boldsymbol{\gamma\theta}_l) - \boldsymbol{D\hat{f}\psi}_l\right) \Big/ (2p) \\
&= \frac{1}{aps^2} \sum_{i=1}^{2p} \Big(\left(\boldsymbol{f}_m(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_{0m}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{D\hat{f}}_m \boldsymbol{\psi}_l\right)' \boldsymbol{\omega}_m \left(\boldsymbol{f}_m(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_{0m}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{D\hat{f}}_m \boldsymbol{\psi}_l\right) \\
&\quad + \left(\boldsymbol{D\hat{f}}_d \boldsymbol{\psi}_l\right)' \boldsymbol{\omega}_d \; \boldsymbol{D\hat{f}}_d \boldsymbol{\psi}_l \Big) \Big/ (2p),
\end{aligned} \tag{44}$$

where $\boldsymbol{\theta}_l$ is computed using equations 40-42, and

$$\begin{aligned}
\boldsymbol{\psi}_l &= \left(\boldsymbol{D\hat{f}}' \boldsymbol{\omega} \boldsymbol{D\hat{f}}\right)^{-1} \boldsymbol{D\hat{f}}' \boldsymbol{\omega} \left(\boldsymbol{f}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_0(\boldsymbol{\gamma\theta}_l)\right) \\
&= \left(\boldsymbol{D\hat{f}}' \boldsymbol{\omega} \boldsymbol{D\hat{f}}\right)^{-1} \boldsymbol{D\hat{f}}_m' \boldsymbol{\omega}_m \left(\boldsymbol{f}_m(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_{0m}(\boldsymbol{\gamma\theta}_l)\right).
\end{aligned} \tag{45}$$

The measure $aps^2\hat{N}_{\min}$ is similar to the second expected value in $\hat{\gamma}_I\sigma_\varepsilon^4$ given by equation 25. Hence, the measure provides another test of the importance of model intrinsic nonlinearity in c_r, c_c, and c_p in addition to the test using the term $(\boldsymbol{Y}-\boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}))'\boldsymbol{\omega}^{1/2}\boldsymbol{R}\boldsymbol{\omega}^{1/2}(\boldsymbol{Y}-\boldsymbol{f}(\boldsymbol{\gamma}\hat{\boldsymbol{\theta}}))$ given after equation 29. By analogy with the former test, $aps^2\hat{N}_{\min}$ should be much smaller than $S(\hat{\boldsymbol{\theta}})$ when applied as a test of the size of $\hat{\gamma}_I\sigma_\varepsilon^4$ and should be much smaller than aps^2 when applied as a test of the size of $\gamma_{Ir}\sigma_\varepsilon^4$. Therefore, the test criterion for $\hat{\gamma}_I\sigma_\varepsilon^4$ is that $\hat{N}_{\min}$ should be much smaller than $(n-ap)/(ap)$, and the test criterion for $\gamma_{Ir}\sigma_\varepsilon^4$ is that $\hat{N}_{\min}$ should be much smaller than 1. The ranking used to classify nonlinearity for $\hat{N}$ also can be used for $\hat{N}_{\min}$ because $\hat{N}_{\min}$ is just the smallest possible value for $\hat{N}$.

A Measure of Model Combined Intrinsic Nonlinearity for Confidence Intervals

Model combined intrinsic nonlinearity affects c_c because $\gamma_{Ic}\sigma_\varepsilon^4$ is a component of this correction factor. Cooley (2004, equation 6-20) expressed the component correction factor $\gamma_{Ic}\sigma_\varepsilon^4$ as

$$\begin{aligned}
\gamma_{Ic}\sigma_\varepsilon^4 &= 2E\Bigg(\Big(\boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta}_*)-\boldsymbol{f}_0(\boldsymbol{\gamma}\boldsymbol{\theta}_*)-\boldsymbol{f}(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})+\boldsymbol{f}_0(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})\Big)'\boldsymbol{\omega}^{1/2}\\
&-\Big(g(\boldsymbol{\gamma}\boldsymbol{\theta}_*)-g_0(\boldsymbol{\gamma}\boldsymbol{\theta}_*)-g(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})+g_0(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})\Big)\frac{\boldsymbol{Q}'}{\boldsymbol{Q}'\boldsymbol{Q}}\Bigg)\Bigg(\boldsymbol{I}-\boldsymbol{R}+\frac{\boldsymbol{Q}\boldsymbol{Q}'}{\boldsymbol{Q}'\boldsymbol{Q}}\Bigg)\boldsymbol{\omega}^{1/2}\big(\boldsymbol{Y}-\boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta}_*)\big)\\
&+E\Bigg(\Big(\boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta}_*)-\boldsymbol{f}_0(\boldsymbol{\gamma}\boldsymbol{\theta}_*)-\boldsymbol{f}(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})+\boldsymbol{f}_0(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})\Big)'\boldsymbol{\omega}^{1/2}\\
&-\Big(g(\boldsymbol{\gamma}\boldsymbol{\theta}_*)-g_0(\boldsymbol{\gamma}\boldsymbol{\theta}_*)-g(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})+g_0(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})\Big)\frac{\boldsymbol{Q}'}{\boldsymbol{Q}'\boldsymbol{Q}}\Bigg)\Bigg(\boldsymbol{I}-\boldsymbol{R}+\frac{\boldsymbol{Q}\boldsymbol{Q}'}{\boldsymbol{Q}'\boldsymbol{Q}}\Bigg)\cdot\\
&\Bigg(\boldsymbol{\omega}^{1/2}\Big(\boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta}_*)-\boldsymbol{f}_0(\boldsymbol{\gamma}\boldsymbol{\theta}_*)-\boldsymbol{f}(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})+\boldsymbol{f}_0(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})\Big)-\frac{\boldsymbol{Q}}{\boldsymbol{Q}'\boldsymbol{Q}}\Big(g(\boldsymbol{\gamma}\boldsymbol{\theta}_*)-g_0(\boldsymbol{\gamma}\boldsymbol{\theta}_*)-g(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})+g_0(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})\Big)\Bigg)\\
&+E\Bigg(\Big(\boldsymbol{Y}-\boldsymbol{f}(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})\Big)'\boldsymbol{\omega}^{1/2}\Bigg(\boldsymbol{R}-\frac{\boldsymbol{Q}\boldsymbol{Q}'}{\boldsymbol{Q}'\boldsymbol{Q}}\Bigg)\boldsymbol{\omega}^{1/2}\Big(\boldsymbol{Y}-\boldsymbol{f}(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})\Big)\Bigg)-\hat{\gamma}_I\sigma_\varepsilon^4,
\end{aligned} \tag{46}$$

where $g(\boldsymbol{\gamma}\boldsymbol{\theta})$, with $\boldsymbol{\theta}=\boldsymbol{\theta}_*$ or $\tilde{\boldsymbol{\theta}}$, is a model prediction for which a confidence interval is to be computed, $g_0(\boldsymbol{\gamma}\boldsymbol{\theta})$ is the linearized approximation of it analogous to $\boldsymbol{f}_0(\boldsymbol{\gamma}\boldsymbol{\theta})$, and $\tilde{\boldsymbol{\theta}}$ is the weighted regression estimate of $\boldsymbol{\theta}_*$ that is constrained so that $g(\boldsymbol{\gamma}\boldsymbol{\theta}_*)=g(\boldsymbol{\gamma}\tilde{\boldsymbol{\theta}})$.

Cooley (2004, p. 86-87) developed a measure of model combined intrinsic nonlinearity, $\hat{M}_{\min}$, similar to the second expected value in equation 46. As for $\hat{N}_{\min}$, the measure is defined in terms of a likelihood region centered on $\bar{\boldsymbol{\theta}}$, but this time having diameter $\xi b\sigma_\varepsilon^2$, where

$$\xi=\boldsymbol{Q}'(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{\Omega}(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{Q}/\boldsymbol{Q}'\boldsymbol{Q}=\big(\boldsymbol{Q}_m'(\boldsymbol{\omega}_m/b)^{1/2}\boldsymbol{\Omega}_m(\boldsymbol{\omega}_m/b)^{1/2}\boldsymbol{Q}_m+\boldsymbol{Q}_d'\boldsymbol{Q}_d/b\big)/\boldsymbol{Q}'\boldsymbol{Q}. \tag{47}$$

Thus, it is calculated using estimates as

$$\begin{aligned}\hat{M}_{\min} &= \frac{1}{\xi s^2}\sum_{l=1}^{2}\left(\boldsymbol{f}(\gamma\boldsymbol{\theta}_l)-\boldsymbol{f}_0(\gamma\boldsymbol{\theta}_l)-\boldsymbol{D}\hat{\boldsymbol{f}}\boldsymbol{\psi}_l\right)'\boldsymbol{\omega}\left(\boldsymbol{f}(\gamma\boldsymbol{\theta}_l)-\boldsymbol{f}_0(\gamma\boldsymbol{\theta}_l)-\boldsymbol{D}\hat{\boldsymbol{f}}\boldsymbol{\psi}_l\right)\Big/2 \\ &= \frac{1}{\xi s^2}\sum_{l=1}^{2}\Big(\left(\boldsymbol{f}_m(\gamma\boldsymbol{\theta}_l)-\boldsymbol{f}_{0m}(\gamma\boldsymbol{\theta}_l)-\boldsymbol{D}\hat{\boldsymbol{f}}_m\boldsymbol{\psi}_l\right)'\boldsymbol{\omega}_m\left(\boldsymbol{f}_m(\gamma\boldsymbol{\theta}_l)-\boldsymbol{f}_{0m}(\gamma\boldsymbol{\theta}_l)-\boldsymbol{D}\hat{\boldsymbol{f}}_m\boldsymbol{\psi}_l\right) \\ &+\left(\boldsymbol{D}\hat{\boldsymbol{f}}_d\boldsymbol{\psi}_l\right)'\boldsymbol{\omega}_d\,\boldsymbol{D}\hat{\boldsymbol{f}}_d\boldsymbol{\psi}_l\Big)\Big/2,\end{aligned} \tag{48}$$

where $\boldsymbol{\theta}_l$ is given by Cooley (2004, p. 87) as

$$\boldsymbol{\theta}_l = \hat{\boldsymbol{\theta}} \pm \left(\frac{\xi s^2}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}}}\right)^{1/2}\left(\boldsymbol{D}\hat{\boldsymbol{f}}'\boldsymbol{\omega}\boldsymbol{D}\hat{\boldsymbol{f}}\right)^{-1}\boldsymbol{D}\hat{g}', \tag{49}$$

in which the carets over $\boldsymbol{f}$, g, and $\boldsymbol{Q}$ indicate evaluation using $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$,

$$\boldsymbol{\psi}_l = \boldsymbol{\psi}_l^0 + \frac{\left(\boldsymbol{D}\hat{\boldsymbol{f}}'\boldsymbol{\omega}\boldsymbol{D}\hat{\boldsymbol{f}}\right)^{-1}\boldsymbol{D}\hat{g}'}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}}}\left(g(\gamma\boldsymbol{\theta}_l)-g_0(\gamma\boldsymbol{\theta}_l)\right) \tag{50}$$

and

$$\begin{aligned}\boldsymbol{\psi}_l^0 &= -\left(\frac{\left(\boldsymbol{D}\hat{\boldsymbol{f}}'\boldsymbol{\omega}\boldsymbol{D}\hat{\boldsymbol{f}}\right)^{-1}\boldsymbol{D}\hat{g}'\hat{\boldsymbol{Q}}'}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}}}-\left(\boldsymbol{D}\hat{\boldsymbol{f}}'\boldsymbol{\omega}\boldsymbol{D}\hat{\boldsymbol{f}}\right)^{-1}\boldsymbol{D}\hat{\boldsymbol{f}}'\boldsymbol{\omega}^{1/2}\right)\boldsymbol{\omega}^{1/2}\left(\boldsymbol{f}(\gamma\boldsymbol{\theta}_l)-\boldsymbol{f}_0(\gamma\boldsymbol{\theta}_l)\right) \\ &= -\left(\frac{\left(\boldsymbol{D}\hat{\boldsymbol{f}}'\boldsymbol{\omega}\boldsymbol{D}\hat{\boldsymbol{f}}\right)^{-1}\boldsymbol{D}\hat{g}'\hat{\boldsymbol{Q}}'_m}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}}}-\left(\boldsymbol{D}\hat{\boldsymbol{f}}'\boldsymbol{\omega}\boldsymbol{D}\hat{\boldsymbol{f}}\right)^{-1}\boldsymbol{D}\hat{\boldsymbol{f}}'_m\boldsymbol{\omega}_m^{1/2}\right)\boldsymbol{\omega}_m^{1/2}\left(\boldsymbol{f}_m(\gamma\boldsymbol{\theta}_l)-\boldsymbol{f}_{0m}(\gamma\boldsymbol{\theta}_l)\right).\end{aligned} \tag{51}$$

Note that the diameter of the linear likelihood region ξs^2 can be written in terms of c_c as evaluated in CORFAC-2k (which assumes no model combined intrinsic nonlinearity) as

$$\xi s^2 = \frac{\boldsymbol{Q}'(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{\Omega}(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{Q}}{\boldsymbol{Q}'\boldsymbol{Q}}\frac{S(\hat{\boldsymbol{\theta}})}{n-ap} = \frac{(n-p)\boldsymbol{Q}'(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{\Omega}(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{Q}}{\boldsymbol{Q}'\boldsymbol{Q}(n-ap)}\frac{S(\hat{\boldsymbol{\theta}})}{n-p} = c_c\frac{S(\hat{\boldsymbol{\theta}})}{n-p}. \tag{52}$$

If $\boldsymbol{\Omega}$ is unknown so that c_c cannot be exactly computed, then the approximate bound for c_c given by equation 83 derived in the sub-section titled 'The CORFAC-2k Program' can be used. In all cases, $\boldsymbol{Q}$ is replaced with $\hat{\boldsymbol{Q}}$ to make the calculations.

Note that $\gamma_{Ic}\sigma_\varepsilon^4$ as defined by equation 46 contains $\hat{\gamma}_I\sigma_\varepsilon^4$. A measure similar to the sum $(\gamma_{Ic}+\hat{\gamma}_I)\sigma_\varepsilon^4$ uses two components in addition to $\hat{M}_{\min}$. These are (Cooley, 2004, p. 86)

$$\hat{B}_U = \frac{1}{\xi s^2}\sum_{l=1}^{2}\left(g(\gamma\boldsymbol{\theta}_l) - g_0(\gamma\boldsymbol{\theta}_l)\right)^2 / \left(2\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}}\right) \tag{53}$$

and

$$\begin{aligned}\hat{B}_L &= \frac{1}{\xi s^2}\sum_{l=1}^{2}\left(\boldsymbol{f}(\gamma\boldsymbol{\theta}_l) - \boldsymbol{f}_0(\gamma\boldsymbol{\theta}_l) - \boldsymbol{D}\hat{\boldsymbol{f}}\boldsymbol{\psi}_l^0\right)'\boldsymbol{\omega}\left(\boldsymbol{f}(\gamma\boldsymbol{\theta}_l) - \boldsymbol{f}_0(\gamma\boldsymbol{\theta}_l) - \boldsymbol{D}\hat{\boldsymbol{f}}\boldsymbol{\psi}_l^0\right)\Big/2 \\ &= \frac{1}{\xi s^2}\sum_{l=1}^{2}\Big(\left(\boldsymbol{f}_m(\gamma\boldsymbol{\theta}_l) - \boldsymbol{f}_{0m}(\gamma\boldsymbol{\theta}_l) - \boldsymbol{D}\hat{\boldsymbol{f}}_m\boldsymbol{\psi}_l^0\right)'\boldsymbol{\omega}_m\left(\boldsymbol{f}_m(\gamma\boldsymbol{\theta}_l) - \boldsymbol{f}_{0m}(\gamma\boldsymbol{\theta}_l) - \boldsymbol{D}\hat{\boldsymbol{f}}_m\boldsymbol{\psi}_l^0\right) \\ &\quad + \left(\boldsymbol{D}\hat{\boldsymbol{f}}_d\boldsymbol{\psi}_l^0\right)'\boldsymbol{\omega}_d\boldsymbol{D}\hat{\boldsymbol{f}}_d\boldsymbol{\psi}_l^0\Big)\Big/2.\end{aligned} \tag{54}$$

The measure of model combined intrinsic nonlinearity related to the sum is given by the maximum in absolute value of $\hat{M}_{\min} + 2\hat{B}_U$ and $\hat{M}_{\min} - 2\hat{B}_L$ (Cooley, 2004, p. 87).

From Cooley (2004, equation 5-49), model combined intrinsic nonlinearity is negligible in the correction factor c_c if $\left|\gamma_{Ic}\sigma_\varepsilon^4\right| << \sigma_\varepsilon^2 + \gamma_{wc}\sigma_\beta^2 = \boldsymbol{Q}'(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{\Omega}(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{Q}b\sigma_\varepsilon^2 / \boldsymbol{Q}'\boldsymbol{Q} = \xi b\sigma_\varepsilon^2$. Thus, $(\gamma_{Ic} + \hat{\gamma}_I)\sigma_\varepsilon^4 / (\xi b\sigma_\varepsilon^2)$ should be much greater than $(\xi b\sigma_\varepsilon^2 + apb\sigma_\varepsilon^2)/(\xi b\sigma_\varepsilon^2) = (\xi + ap)/\xi$, which has a value greater than 1, so that, as a conservative test criterion, $\hat{M}_{\min} + 2\hat{B}_U$ and $\left|\hat{M}_{\min} - 2\hat{B}_L\right|$ should both be much less than 1. It is likely that the same ranking used to classify nonlinearity for $\hat{N}$ also can be used to classify model intrinsic nonlinearity for its measure, although this ranking might be conservative in the present instance.

Cooley (2004, p. 66) showed that when model intrinsic nonlinearity and model combined intrinsic nonlinearity both are small, standard linear confidence intervals are accurate approximations of nonlinear intervals. Thus, because $\hat{M}_{\min}$ measures the magnitude of these two sources of nonlinearity, $\hat{M}_{\min}$ can be used to indicate when standard linear confidence intervals can be used. On the basis of preliminary results given by Cooley (2004, section 7), linear confidence intervals should be good approximations when $\hat{M}_{\min} \le 0.01$.

A Measure of Model Combined Intrinsic Nonlinearity for Prediction Intervals

The development and measure of model combined intrinsic nonlinearity for prediction intervals is analogous to the development and measure for confidence intervals. The measure for prediction intervals was not explicitly derived by Cooley (2004) but follows from the equations derived for confidence intervals with some added terms obtained from equations 6-50 - 6-55 of Cooley (2004, p. 84-85). For the prediction intervals calculated by UNC, the diameter of the likelihood region analogous to ξs^2 is $\xi_a s^2$, where ξ_a is defined using equation 7-12 from Cooley (2004, p. 113) as

$$\begin{aligned}\xi_a &= \frac{\boldsymbol{Q}'(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{\Omega}(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{Q} - 2\omega_p^{-1/2}(\omega_p/b)^{1/2}\boldsymbol{C}'(\boldsymbol{\omega}/b)^{1/2}\boldsymbol{Q} + \hat{\omega}_p^{-1}}{\boldsymbol{Q}'\boldsymbol{Q} + \omega_p^{-1}} \\ &= \frac{\boldsymbol{Q}_m'(\boldsymbol{\omega}_m/b)^{1/2}\boldsymbol{\Omega}_m(\boldsymbol{\omega}_m/b)^{1/2}\boldsymbol{Q}_m + \boldsymbol{Q}_d'\boldsymbol{Q}_d/b - 2\omega_p^{-1/2}(\omega_p/b)^{1/2}\boldsymbol{C}_m'(\boldsymbol{\omega}_m/b)^{1/2}\boldsymbol{Q}_m + \hat{\omega}_p^{-1}}{\boldsymbol{Q}'\boldsymbol{Q} + \omega_p^{-1}}.\end{aligned} \tag{55}$$

We assume in this report that errors $Y_p - g(\boldsymbol{\gamma\theta}_*)$ in predictions are not correlated with errors $\boldsymbol{Y}_d - \boldsymbol{f}_d(\boldsymbol{\gamma\theta}_*)$ in direct information so that $\boldsymbol{C}_d = \boldsymbol{\theta}$. The three measures analogous to equations 48, 53, and 54 are

$$\begin{aligned}\hat{M}^{\,a}_{\min} &= \frac{1}{\xi_a s^2}\sum_{l=1}^{2}\Big(\big(\boldsymbol{f}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_0(\boldsymbol{\gamma\theta}_l) - \boldsymbol{D}\hat{\boldsymbol{f}}\boldsymbol{\psi}_{al}\big)'\boldsymbol{\omega}\big(\boldsymbol{f}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_0(\boldsymbol{\gamma\theta}_l) - \boldsymbol{D}\hat{\boldsymbol{f}}\boldsymbol{\psi}_{al}\big) \\ &+ \psi_{pl}\omega_p\psi_{pl}\Big)\Big/2 \\ &= \frac{1}{\xi_a s^2}\sum_{l=1}^{2}\Big(\big(\boldsymbol{f}_m(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_{0m}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{D}\hat{\boldsymbol{f}}_m\boldsymbol{\psi}_{al}\big)'\boldsymbol{\omega}_m\big(\boldsymbol{f}_m(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_{0m}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{D}\hat{\boldsymbol{f}}_m\boldsymbol{\psi}_{al}\big) \\ &+ \big(\boldsymbol{D}\hat{\boldsymbol{f}}_d\boldsymbol{\psi}_{al}\big)'\boldsymbol{\omega}_d\boldsymbol{D}\hat{\boldsymbol{f}}_d\boldsymbol{\psi}_{al} + \psi_{pl}\omega_p\psi_{pl}\Big)\Big/2,\end{aligned} \tag{56}$$

$$\hat{B}^a_U = \frac{1}{\xi_a s^2}\sum_{l=1}^{2}\big(g(\boldsymbol{\gamma\theta}_l) - g_0(\boldsymbol{\gamma\theta}_l)\big)^2\Big/\Big(2\big(\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}} + \omega_p^{-1}\big)\Big), \tag{57}$$

and

$$\begin{aligned}\hat{B}^a_L &= \frac{1}{\xi_a s^2}\sum_{l=1}^{2}\Big(\big(\boldsymbol{f}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_0(\boldsymbol{\gamma\theta}_l) - \boldsymbol{D}\hat{\boldsymbol{f}}\boldsymbol{\psi}^0_{al}\big)'\boldsymbol{\omega}\big(\boldsymbol{f}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_0(\boldsymbol{\gamma\theta}_l) - \boldsymbol{D}\hat{\boldsymbol{f}}\boldsymbol{\psi}^0_{al}\big) + \psi^0_{pl}\omega_p\psi^0_{pl}\Big)\Big/2 \\ &= \frac{1}{\xi_a s^2}\sum_{l=1}^{2}\Big(\big(\boldsymbol{f}_m(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_{0m}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{D}\hat{\boldsymbol{f}}_m\boldsymbol{\psi}^0_{al}\big)'\boldsymbol{\omega}_m\big(\boldsymbol{f}_m(\boldsymbol{\gamma\theta}_l) - \boldsymbol{f}_{0m}(\boldsymbol{\gamma\theta}_l) - \boldsymbol{D}\hat{\boldsymbol{f}}_m\boldsymbol{\psi}^0_{al}\big) \\ &+ \big(\boldsymbol{D}\hat{\boldsymbol{f}}_d\boldsymbol{\psi}^0_{al}\big)'\boldsymbol{\omega}_d\boldsymbol{D}\hat{\boldsymbol{f}}_d\boldsymbol{\psi}^0_{al} + \psi^0_{pl}\omega_p\psi^0_{pl}\Big)\Big/2,\end{aligned} \tag{58}$$

where $\boldsymbol{\theta}_l$ is obtained from equation H-12 in Cooley (2004, p. 205) as

$$\boldsymbol{\theta}_l = \hat{\boldsymbol{\theta}} \pm \left(\frac{\xi_a s^2}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}} + \omega_p^{-1}}\right)^{1/2}\big(\boldsymbol{D}\hat{\boldsymbol{f}}'\boldsymbol{\omega}\boldsymbol{D}\hat{\boldsymbol{f}}\big)^{-1}\boldsymbol{D}\hat{g}', \tag{59}$$

and remaining variables are obtained straightforwardly from equations 6-51 - 6-55 in Cooley (2004, p. 84-85) as

$$\boldsymbol{\psi}_{al} = \boldsymbol{\psi}_{al}^{0} + \frac{\left(\hat{\boldsymbol{D}}\boldsymbol{f}'\boldsymbol{\omega}\hat{\boldsymbol{D}}\boldsymbol{f}\right)^{-1}\hat{\boldsymbol{D}}\boldsymbol{g}'}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}} + \omega_p^{-1}}\left(g(\boldsymbol{\gamma}\boldsymbol{\theta}_l) - g_0(\boldsymbol{\gamma}\boldsymbol{\theta}_l)\right), \tag{60}$$

$$\begin{aligned}\boldsymbol{\psi}_{al}^{0} &= -\left(\frac{\left(\hat{\boldsymbol{D}}\boldsymbol{f}'\boldsymbol{\omega}\hat{\boldsymbol{D}}\boldsymbol{f}\right)^{-1}\hat{\boldsymbol{D}}\boldsymbol{g}'\hat{\boldsymbol{Q}}'}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}} + \omega_p^{-1}} - \left(\hat{\boldsymbol{D}}\boldsymbol{f}'\boldsymbol{\omega}\hat{\boldsymbol{D}}\boldsymbol{f}\right)^{-1}\hat{\boldsymbol{D}}\boldsymbol{f}'\boldsymbol{\omega}^{1/2}\right)\boldsymbol{\omega}^{1/2}\left(\boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta}_l) - \boldsymbol{f}_0(\boldsymbol{\gamma}\boldsymbol{\theta}_l)\right) \\ &= -\left(\frac{\left(\hat{\boldsymbol{D}}\boldsymbol{f}'\boldsymbol{\omega}\hat{\boldsymbol{D}}\boldsymbol{f}\right)^{-1}\hat{\boldsymbol{D}}\boldsymbol{g}'\hat{\boldsymbol{Q}}_m'}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}} + \omega_p^{-1}} - \left(\hat{\boldsymbol{D}}\boldsymbol{f}'\boldsymbol{\omega}\hat{\boldsymbol{D}}\boldsymbol{f}\right)^{-1}\hat{\boldsymbol{D}}\boldsymbol{f}_m'\boldsymbol{\omega}_m^{1/2}\right)\boldsymbol{\omega}_m^{1/2}\left(\boldsymbol{f}_m(\boldsymbol{\gamma}\boldsymbol{\theta}_l) - \boldsymbol{f}_{0m}(\boldsymbol{\gamma}\boldsymbol{\theta}_l)\right),\end{aligned} \tag{61}$$

$$\psi_{pl} = \psi_{pl}^{0} - \frac{\omega_p^{-1}}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}} + \omega_p^{-1}}\left(g(\boldsymbol{\gamma}\boldsymbol{\theta}_l) - g_0(\boldsymbol{\gamma}\boldsymbol{\theta}_l)\right), \tag{62}$$

$$\begin{aligned}\psi_{pl}^{0} &= \frac{\omega_p^{-1}}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}} + \omega_p^{-1}}\hat{\boldsymbol{Q}}'\boldsymbol{\omega}^{1/2}\left(\boldsymbol{f}(\boldsymbol{\gamma}\boldsymbol{\theta}_l) - \boldsymbol{f}_0(\boldsymbol{\gamma}\boldsymbol{\theta}_l)\right) \\ &= \frac{\omega_p^{-1}}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}} + \omega_p^{-1}}\hat{\boldsymbol{Q}}_m'\boldsymbol{\omega}_m^{1/2}\left(\boldsymbol{f}_m(\boldsymbol{\gamma}\boldsymbol{\theta}_l) - \boldsymbol{f}_{0m}(\boldsymbol{\gamma}\boldsymbol{\theta}_l)\right).\end{aligned} \tag{63}$$

The criteria for assessing the importance of model combined intrinsic nonlinearity in c_p are completely analogous to the criteria pertaining to c_c; equations 5-93 and 7-12 of Cooley (2004) replace equation 5-49 of Cooley (2004) in the development. Thus, $\hat{M}_{\min}^{a} + 2\hat{B}_U^{a}$ and $\left|\hat{M}_{\min}^{a} - 2\hat{B}_L^{a}\right|$ should be much less than 1, and the ranking used to classify nonlinearity for $\hat{N}$ can again be used to classify model combined nonlinearity for its measures.

The measure $\hat{M}_{\min}^{a}$ can be used in the same way as $\hat{M}_{\min}$ was used for confidence intervals to indicate when standard linear prediction intervals should be accurate approximations of nonlinear ones. Thus, linear prediction intervals should be good approximations of nonlinear ones when $\hat{M}_{\min}^{a} \le 0.01$.

The CORFAC-2k Program

The purpose of CORFAC-2k is to compute correction factors c_r, c_c, and c_p, together with components of them, assuming that model intrinsic nonlinearity and model combined intrinsic nonlinearity are negligible. The component correction factors $\hat{\gamma}_w\sigma_\beta^2$, $\gamma_{wr}\sigma_\beta^2$, $\gamma_{wc}\sigma_\beta^2$, and $\gamma_{wa}\sigma_\beta^2$ are equal to zero when weight matrix $\boldsymbol{\omega}$ is equal to the inverse of the matrix $\boldsymbol{\Omega}$. When $\boldsymbol{\Omega}$ is unknown, approximations and bounds are developed for the correction factors.

Calculation of the Correction Factors When $\boldsymbol{\Omega}$ Is Known

Component correction factor $\hat{\gamma}_w\sigma_\beta^2$ is defined using equations 5-10 and F-133 of Cooley (2004) as

$$\hat{\gamma}_w \sigma_\beta^2 = \big((b-1)n - (ba-1)p\big)\sigma_\varepsilon^2, \tag{64}$$

and $\gamma_{wr}\sigma_\beta^2$ is defined using equations 5-12 and F-135 of Cooley (2004) as

$$\gamma_{wr}\sigma_\beta^2 = (ba-1)p\sigma_\varepsilon^2, \tag{65}$$

where *a* and *b* are given by equations 26 and 27 and $a \geq 1$ (Cooley, 2004, p. 52). Note that if $b \geq 1$, then $\hat{\gamma}_w \sigma_\beta^2 \leq 0$, and if $b = 1$, then $\hat{\gamma}_w \sigma_\beta^2 = -\gamma_{wr}\sigma_\beta^2$. Correction factor c_r is obtained by combining equations 64 and 65 according to equation 22, assuming $\hat{\gamma}_I \sigma_\varepsilon^4$ and $\gamma_{Ir}\sigma_\varepsilon^4$ are both zero, to get

$$c_r = \frac{(n-p)a}{n-ap}. \tag{66}$$

Note from equations 26 and 27 that if $\boldsymbol{\omega}$ is proportional to, but not equal to, $\boldsymbol{\Omega}^{-1}$, then $a = 1$, but $b \neq 1$ and component correction factors from equations 64 and 65 are both nonzero. However, *b* cancels when obtaining equation 66 so that $\boldsymbol{\omega}$ only has to be proportional to $\boldsymbol{\Omega}^{-1}$ for c_r to be equal to 1. Also note that at least when model intrinsic nonlinearity is negligible, $c_r \geq 1$, so that uncorrected confidence regions and uncorrected Scheffé intervals would be too small unless $\boldsymbol{\omega} \propto \boldsymbol{\Omega}^{-1}$. To compute c_r, *a* and thus $\boldsymbol{\Omega}$ have to be known.

Component correction factor $\gamma_{wc}\sigma_\beta^2$ is defined using equation F-146 of Cooley (2004, p. 187) and equation 47 as

$$\gamma_{wc}\sigma_\beta^2 = (b\xi - 1)\sigma_\varepsilon^2. \tag{67}$$

Variable ξ given by equation 52 is not necessarily greater than 1, so $\gamma_{wc}\sigma_\beta^2$ is not necessarily greater than zero. Correction factor c_c is obtained by combining equations 64 and 67 according to equation 23, assuming $\hat{\gamma}_I \sigma_\varepsilon^4$ and $\gamma_{Ic}\sigma_\varepsilon^4$ are both zero, to get

$$c_c = \frac{(n-p)\xi}{n-ap}, \tag{68}$$

where again $\boldsymbol{\omega}$ only has to be proportional to $\boldsymbol{\Omega}$ for c_c to equal 1. Matrix $\boldsymbol{\Omega}$ has to be known in order to compute c_c.

Finally, component correction factor $\gamma_{wa}\sigma_\beta^2$ is defined using equation 5-93 of Cooley (2004, p. 63) and equation 55 as

$$\gamma_{wa}\sigma_\beta^2 = \big(b\xi_a - 1\big)\sigma_\varepsilon^2. \tag{69}$$

As for equation 67, ξ_a is not necessarily greater than 1, so $\gamma_{wa}\sigma_\beta^2$ is not necessarily greater than zero. Correction factor c_p is obtained by combining equations 64 and 69 according to equation 24, assuming $\hat{\gamma}_I\sigma_\varepsilon^4$ and $\gamma_{Ia}\sigma_\varepsilon^4$ are both zero, to get

$$c_p = \frac{(n-p)\xi_a}{n-ap}, \tag{70}$$

which behaves like c_c does. However, c_p can differ from 1 even when $\boldsymbol{\omega}$ is equal to $\boldsymbol{\Omega}^{-1}$ because of the term involving $\boldsymbol{C}$ in equation 55.

Approximate Calculation of the Correction Factors When $\boldsymbol{\Omega}$ Is Unknown

When $\boldsymbol{\Omega}_m$ is unknown, a, ξ, and ξ_a must be approximated. To obtain the approximations, we set $\boldsymbol{\omega}_m$ equal to the diagonal weight matrix $\hat{\boldsymbol{\omega}}_m$ composed of the inverses of the diagonals of $\boldsymbol{\Omega}_m$, we let $\boldsymbol{\omega}_d = \hat{\boldsymbol{\omega}}_d = \boldsymbol{\Omega}_d^{-1}$, and we set ω_p equal to $\hat{\omega}_p$. (Usually, $\hat{\boldsymbol{\omega}}_m$ will only be an approximation of the inverse of the diagonals of $\boldsymbol{\Omega}_m$. This is discussed briefly later in this section.) Then we make use of approximations from equations 5-20, 5-56, and 5-117 of Cooley (2004). To approximate a we use equation 5-20 in equation 5-10 of Cooley (2004), noting that b=1 and $\boldsymbol{\omega}_d = \boldsymbol{\Omega}_d^{-1}$ so that equation 5-20 need only be applied to the model function partition. Therefore,

$$tr\left(\left(\boldsymbol{I}_m - \hat{\boldsymbol{R}}_m\right)\left((1-c)\boldsymbol{I}_m + c\boldsymbol{I}_m\right)\right) + tr\left(\boldsymbol{I}_d - \hat{\boldsymbol{R}}_d\right) = (1-c)t_m + cs_m + t_d = n - ap, \tag{71}$$

where c is an effective correlation, $\boldsymbol{I}_m$ is an $n_m \times n_m$ matrix of 1's, partitions of $\hat{\boldsymbol{R}}$ are partitions of $\boldsymbol{R}$ computed using $\hat{\boldsymbol{\omega}}$, and the traces are evaluated as

$$t_m = tr\left(\left(\boldsymbol{I}_m - \hat{\boldsymbol{R}}_m\right)\boldsymbol{I}_m\right) = \sum_{i=1}^{n_m}\left(1-\hat{r}_{ii}\right) = n_m - \sum_{i=1}^{n_m}\hat{r}_{ii}, \tag{72}$$

$$s_m = tr\left(\left(\boldsymbol{I}_m - \hat{\boldsymbol{R}}_m\right)\boldsymbol{I}_m\right) = \sum_{i=1}^{n_m}\sum_{j=1}^{n_m}\left(\delta_{ij} - \hat{r}_{ij}\right) = n_m - \sum_{i=1}^{n_m}\sum_{j=1}^{n_m}\hat{r}_{ij}, \tag{73}$$

$$t_d = tr\left(\boldsymbol{I}_d - \hat{\boldsymbol{R}}_d\right) = \sum_{i=n_m+1}^{n}\left(1-\hat{r}_{ii}\right) = n_d - \sum_{i=n_m+1}^{n}\hat{r}_{ii}. \tag{74}$$

In equations 72-74, $\hat{r}_{ij}$ is an element of $\hat{\boldsymbol{R}}$ and δ_{ij} is the Kronecker delta ($\delta_{ij} = 1$ when $i = j$ and $\delta_{ij} = 0$ when $i \neq j$).

Equation 71 can be solved for c by making use of the fact that $t_m + t_d = n - p$. The result is

$$c = \frac{(a-1)p}{t_m - s_m}, \tag{75}$$

which can serve as a definition of *c*. Note that if $t_m > s_m, c \geq 0$. The condition $t_m > s_m$ has always been found to hold. If *c* is presumed known, then equation 75 can be solved for *a* to get

$$a = \frac{c}{p}(t_m - s_m) + 1, \tag{76}$$

which has a lower bound of *a*=1 when *c*=0 and an upper bound of $1 + (t_m - s_m)/p$ when *c*=1. By assuming a reasonable upper limit for *c*, c_e, an approximate upper limit for *a* can be computed using equation 76.

The variable ξ can be bounded using equation 5-56 from Cooley (2004, p. 56). However, a more refined approximation can be obtained as follows. From equation 5-56 from Cooley (2004) and the fact that the approximation need not be applied for the direct information,

$$\xi = \left(\hat{\boldsymbol{Q}}_m' \hat{\boldsymbol{\omega}}_m^{1/2} \boldsymbol{\Omega}_m \hat{\boldsymbol{\omega}}_m^{1/2} \hat{\boldsymbol{Q}}_m + \hat{\boldsymbol{Q}}_d' \hat{\boldsymbol{Q}}_d\right) / \hat{\boldsymbol{Q}}' \hat{\boldsymbol{Q}} \leq \left(V_{mx} + \hat{\boldsymbol{Q}}_d' \hat{\boldsymbol{Q}}_d\right) / \hat{\boldsymbol{Q}}' \hat{\boldsymbol{Q}}, \tag{77}$$

where

$$V_{mx} = \max_s \left(\sum_{i(s)} Q_{mi} \right)^2 \tag{78}$$

and the notation indicates the maximum of either the squared sum of positive values of the Q_{mi}'s or the squared sum of negative values of the Q_{mi}'s (Cooley, 2004, p. 55-56). Also, if $\boldsymbol{\Omega}_m$ is diagonal so that $\hat{\boldsymbol{\omega}}_m^{1/2} \boldsymbol{\Omega}_m \hat{\boldsymbol{\omega}}_m^{1/2} = \boldsymbol{I}_m$, then $\hat{\boldsymbol{Q}}_m' \hat{\boldsymbol{\omega}}_m^{1/2} \boldsymbol{\Omega}_m \hat{\boldsymbol{\omega}}_m^{1/2} \hat{\boldsymbol{Q}}_m = \hat{\boldsymbol{Q}}_m' \hat{\boldsymbol{Q}}_m$ so that $\xi = 1$. Therefore, a value of $c_e \leq 1$ always exists so that

$$\xi \leq \left((1 - c_e) \hat{\boldsymbol{Q}}_m' \hat{\boldsymbol{Q}}_m + c_e V_{mx} + \hat{\boldsymbol{Q}}_d' \hat{\boldsymbol{Q}}_d\right) / \hat{\boldsymbol{Q}}' \hat{\boldsymbol{Q}}. \tag{79}$$

It is almost certain that equation 79 with c_e<1 would produce a closer bound than equation 77.

Finally, ξ_a is bounded using equation 5-117 of Cooley (2004, p. 68) and a development completely analogous to the one used to obtain equation 79 to obtain

$$\xi_a \leq \left((1 - c_e)\left(\hat{\boldsymbol{Q}}_m' \hat{\boldsymbol{Q}}_m + \hat{\omega}_p^{-1}\right) + c_e V_{mxa} + \hat{\boldsymbol{Q}}_d' \hat{\boldsymbol{Q}}_d\right) / \left(\hat{\boldsymbol{Q}}' \hat{\boldsymbol{Q}} + \hat{\omega}_p^{-1}\right), \tag{80}$$

where V_{mxa} is evaluated using equation 78 as

$$\left.\begin{aligned} V_{mxa} &= V_{mx} \ \text{ for } \sum_{i(s)} Q_{mi} \text{ as a sum of positive values} \\ V_{mxa} &= V_{mx} + 2\hat{\omega}_p^{-1/2} V_{mx}^{1/2} + \hat{\omega}_p^{-1} \ \text{ for } \sum_{i(s)} Q_{mi} \text{ as a sum of negative values} \end{aligned}\right\}. \tag{81}$$

Correction factors c_r, c_c, and c_p are approximated by letting $c = c_e$ and using equations 71, 76, 79, and 80 to get

$$c_r \approx \frac{n-p}{p} \frac{c_e(t_m - s_m) + p}{(1-c_e)t_m + c_e s_m + t_d}, \tag{82}$$

$$c_c \approx \frac{n-p}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}}} \frac{(1-c_e)\hat{\boldsymbol{Q}}_m'\hat{\boldsymbol{Q}}_m + c_e V_{mx} + \hat{\boldsymbol{Q}}_d'\hat{\boldsymbol{Q}}_d}{(1-c_e)t_m + c_e s_m + t_d}, \tag{83}$$

and

$$c_p \approx \frac{n-p}{\hat{\boldsymbol{Q}}'\hat{\boldsymbol{Q}} + \hat{\omega}_p^{-1}} \frac{(1-c_e)\left(\hat{\boldsymbol{Q}}_m'\hat{\boldsymbol{Q}}_m + \hat{\omega}_p^{-1}\right) + c_e V_{mxa} + \hat{\boldsymbol{Q}}_d'\hat{\boldsymbol{Q}}_d}{(1-c_e)t_m + c_e s_m + t_d}. \tag{84}$$

Note that if c_e=0 is correct, then c_r=1, c_c=1, and c_p=1, which also are correct. If c_e=1, then c_r, c_c, and c_p, attain their maximum values, which could be very large because the denominators could be very small. For the correction factors examined thus far in tests by the authors of this report, c_e=0.8 has yielded bounds for all values of c_c and c_p. This value of c_e also has yielded bounds for most values of c_r and has yielded good approximations of c_r for the remainder.

For field problems $\hat{\boldsymbol{\omega}}_m$ will only be an approximation of the matrix of inverses of the diagonal elements of $\boldsymbol{\Omega}_m$. The results of example 2 of Cooley (2004, section 7) suggest that a reasonable estimate of $\hat{\boldsymbol{\omega}}_m$ yields very little error in the correction factors. Further work by the senior author suggests that use of the approximation for $\hat{\boldsymbol{\omega}}_m$ obtained for the Tude aa case studied by Christensen and Cooley (1999b) also yields very little error. It appears that careful hydrogeologic work can yield a satisfactory estimate for $\hat{\boldsymbol{\omega}}_m$.

Approximate Calculation of the Correction Factors When $\boldsymbol{\Omega}$ Is Partly Known

In order to make the approximations in equations 71, 76, 79, and 80, we assumed the matrix $\hat{\boldsymbol{\omega}}_m^{1/2}\boldsymbol{\Omega}_m\hat{\boldsymbol{\omega}}_m^{1/2}$ to be similar to a correlation matrix so that all elements have magnitudes ≤1 (Cooley, 2004, p. 48-49). This implies that $\hat{\boldsymbol{\omega}}_m$ is diagonal with elements equal to the inverses of the diagonal elements of $\boldsymbol{\Omega}_m$. Matrix $\boldsymbol{\Omega}_m$ has the form $\boldsymbol{V}_{\varepsilon m} + \boldsymbol{V}_m$, where $\boldsymbol{V}_{\varepsilon m}$ is the diagonal block in $\boldsymbol{V}_\varepsilon$ for model functions and $\boldsymbol{V}_m = E(\boldsymbol{f}_m(\boldsymbol{\beta}) - \boldsymbol{f}_m(\boldsymbol{\gamma}\boldsymbol{\theta}_*))(\boldsymbol{f}_m(\boldsymbol{\beta}) - \boldsymbol{f}_m(\boldsymbol{\gamma}\boldsymbol{\theta}_*))'/\sigma_\varepsilon^2$, which is the model-error second-moment matrix for model functions divided by σ_ε^2 (Cooley, 2004, p. 18-19). Observations such as streamflow gains and losses may have measurement errors that have

known correlation so that $\boldsymbol{V}_{\varepsilon m}$ is not diagonal. In this case, assuming $\hat{\boldsymbol{\omega}}_m$ to be diagonal would result in a loss of information. If $\boldsymbol{V}_{\varepsilon m}$ is known, then to preserve this information $\hat{\boldsymbol{\omega}}_m$ should be written in the form

$$\hat{\boldsymbol{\omega}}_m = (\boldsymbol{V}_{\varepsilon m} + \boldsymbol{D}_m)^{-1}, \tag{85}$$

which gives $\hat{\boldsymbol{\omega}}_m^{1/2}\boldsymbol{\Omega}_m\hat{\boldsymbol{\omega}}_m^{1/2}$ the form $\boldsymbol{C}_m = (\boldsymbol{V}_{\varepsilon m} + \boldsymbol{D}_m)^{-1/2}(\boldsymbol{V}_{\varepsilon m} + \boldsymbol{V}_m)(\boldsymbol{V}_{\varepsilon m} + \boldsymbol{D}_m)^{-1/2}$, where $\boldsymbol{D}_m$ is the diagonal matrix of diagonal elements of $\boldsymbol{V}_m$. Unless $\boldsymbol{V}_{\varepsilon m}$ is diagonal, $\boldsymbol{C}_m$ does not have a correlation-like form. However, as $\boldsymbol{V}_{\varepsilon m}$ approaches $\boldsymbol{0}$, $\boldsymbol{C}_m$ approaches the matrix $\boldsymbol{D}_m^{-1/2}\boldsymbol{V}_m\boldsymbol{D}_m^{-1/2}$, which is similar to a correlation matrix; as $\boldsymbol{V}_m$ approaches $\boldsymbol{0}$, $\boldsymbol{C}_m$ approaches $\boldsymbol{I}$, the identity matrix; and if $\boldsymbol{V}_{\varepsilon m}$ is diagonal, $\boldsymbol{C}_m$ has the standard form. Thus, $\boldsymbol{C}_m$ has the form assumed for the approximations in the limits. We assume that elements of $\boldsymbol{C}_m$ have magnitudes ≤1 (at least in final effect) for all matrices $\boldsymbol{V}_{\varepsilon m}$, so that equations 82, 83, and 84 can be used to compute correction factors when $\hat{\boldsymbol{\omega}}_m$ is defined by equation 85. (This assumption held in tests made during development of the computer code for this report.) Note, however, that as in the standard case where $\hat{\boldsymbol{\omega}}_m$ is diagonal, $\boldsymbol{D}_m$ is approximate because $\boldsymbol{V}_m$ is unknown.