ERDC TR-06-1

Engineer Research and Development Center

US Army Corps of Engineers®
Engineer Research and Development Center

FASST Soil Moisture Prediction Error at Yuma, Arizona

Brian E. Skahill and Susan Frankenstein

April 2006

ERDC TR-06-1
April 2006

FASST Soil Moisture Prediction Error at Yuma, Arizona

Brian E. Skahill

Coastal and Hydraulics Laboratory
U.S. Army Engineer Research and Development Center
72 Lyme Road
Hanover, New Hampshire 03755-1290

Susan Frankenstein

Cold Regions Research and Engineering Laboratory
U.S. Army Engineer Research and Development Center
72 Lyme Road
Hanover, New Hampshire 03755-1290

Final report

Prepared for U.S. Army Corps of Engineers
Washington, DC 20314-1000

ABSTRACT

A linear predictive error variance analysis was employed using the one-dimensional state-of-the-ground model FASST (Fast All-season Soil STrength) and calibration data for Yuma, Arizona (Frankenstein and Koenig 2004). The analysis was performed with the intent to quantify soil moisture predictive error and to examine ways in which it could be reduced, in particular, to demonstrate a methodology wherein one can examine the contribution from each individual model parameter to overall soil moisture predictive error variance.

CONTENTS

ILLUSTRATIONS

TABLES

PREFACE

This report was prepared by Brian E. Skahill, Watershed Systems Group, U.S. Army Engineer Research and Development Center, Coastal and Hydraulics Laboratory (ERDC-CHL), Vicksburg, Mississippi, and by Susan Frankenstein, Geophysical Sciences Branch, U.S. Army Engineer Research and Development Center, Cold Regions Research and Engineering Laboratory (ERDC-CRREL), Hanover, New Hampshire.

Funding for this work was provided by the Department of Defense Basic Research (6.1) Program.

This report was prepared under the general supervision of Janet P. Hardy, Acting Chief, Geophysical Sciences Branch, CRREL; Dr. Lance D. Hansen, Deputy Director, CRREL; and James L. Wuebben, Acting Director, CRREL.

The Commander and Executive Director of the Engineer Research and Development Center is Colonel James R. Rowan. The Director is Dr. James R. Houston.

FASST Soil Moisture Prediction Error at Yuma, Arizona

BRIAN E. SKAHILL AND SUSAN FRANKENSTEIN

1 INTRODUCTION

The effort described herein is based on the linear analysis presented by Moore and Doherty (2005) wherein predictive error variance is computed subsequent to regularized inversion as a means of model calibration. The objectives of this document, using the linear analysis of Moore and Doherty (2005) and the FASST (Fast All-season Soil STrength) model for Yuma, Arizona (Frankenstein and Koenig 2004), are to

1. Quantify and compare notional predictive error variance for two specified soil moisture predictions (one immediately following a storm event and the other following a period with no precipitation);
2. Examine the role of the calibration process in reducing notional predictive error variance for the two specified soil moisture predictions;
3. Measure the worth of additional data in terms of how much it would reduce soil moisture prediction error;
4. Examine the contribution from each individual model parameter to overall soil moisture predictive error variance.

The analysis was performed with the intent to quantify soil moisture predictive error and to examine ways in which it could be reduced. The analysis is presented with the understanding that model parameters are related to underlying hydrologic processes that are represented in the FASST simulator. FASST is a one-dimensional state-of-the-ground dynamic model. As additional hydrologic processes are represented in FASST (e.g., lateral flow) and future versions of FASST are deployed, analyses similar to that contained herein could guide the identification of the level of model complexity that is required, depending on the availability of calibration data, to predict soil moisture with minimal error variance.

2 BACKGROUND

Moore and Doherty (2005) applied linear analysis to derive an equation to determine predictive error variance consequent to regularized inversion as a means of model calibration:

$$\sigma^2_{s-\hat{s}} = \mathbf{y}^t(\mathbf{I} - \mathbf{R})C(\mathbf{p})(\mathbf{I} - \mathbf{R})^t\mathbf{y} + \mathbf{y}^t\mathbf{G}C(\boldsymbol{\varepsilon})\mathbf{G}^t\mathbf{y}, \tag{1a}$$

$$\sigma^2_{s-\hat{s}} = \mathbf{y}^t\mathbf{V}_2\mathbf{V}^t_2C(\mathbf{p})\mathbf{V}_2\mathbf{V}^t_2\mathbf{y} + \sigma^2_h\mathbf{y}^t\mathbf{V}_1\mathbf{E}_1^{-1}\mathbf{V}^t_1\mathbf{y}, \tag{1b}$$

$$\sigma^2_{s-\hat{s}} = \sigma^2_p\mathbf{y}^t\mathbf{V}_2\mathbf{V}^t_2\mathbf{y} + \sigma^2_h\mathbf{y}^t\mathbf{V}_1\mathbf{E}_1^{-1}\mathbf{V}^t_1\mathbf{y} \tag{1c}$$

where

$\sigma^2_{s-\hat{s}}$	=	predictive error variance
$\mathbf{p}$	=	m true model parameters
$C(\mathbf{p})$	=	covariance matrix of $\mathbf{p}$
$\mathbf{y}$	=	sensitivities of the prediction to model parameters
$\mathbf{X}$	=	model sensitivity matrix
$\mathbf{Q}$	=	cofactor matrix whose diagonal elements contain the squares of the observation weights
$\boldsymbol{\varepsilon}$	=	measurement and structural noise
$C(\boldsymbol{\varepsilon})$	=	covariance of the measurement and structural noise (assumed to be diagonal)
$\mathbf{V} = [\mathbf{V}_1\ \mathbf{V}_2]$	=	matrix of eigenvectors of the normal matrix $\mathbf{X}^t\mathbf{QX}$
$\mathbf{V}_1$	=	eigenvectors associated with the k largest eigenvalues of $\mathbf{X}^t\mathbf{QX}$
$\mathbf{V}_2$	=	eigenvectors associated with the remaining eigenvalues of $\mathbf{X}^t\mathbf{QX}$
$\mathbf{E}_1$	=	diagonal matrix containing the k largest eigenvalues of $\mathbf{X}^t\mathbf{QX}$
$\mathbf{I}$	=	identity matrix
$\mathbf{R} = \mathbf{V}_1\mathbf{V}^t_1$	=	model resolution matrix

σ^2_h = the reference variance

σ^2_p = assumed constant variance for the model parameters.

$\mathbf{G} = \mathbf{V_1}\mathbf{E_1}^{-1}\mathbf{V}^t_1\mathbf{X}^t\mathbf{Q}$ and $\hat{\mathbf{p}} = \mathbf{Gh} = \mathbf{G}(\mathbf{Xp} + \boldsymbol{\varepsilon})$, where $\hat{\mathbf{p}}$ represents the m estimated model parameters (that is, the set of parameters corresponding to **p** calculated through the calibration process) and **h** is the vector of *n* field measurements.

Based on an inspection of equations (1a), (1b), and (1c), the following observations are summarized (Moore and Doherty 2005).

A. The orthonormal basis vectors, **V**, that span the range of model (parameter) space can be subdivided into two separate subspaces, $\mathbf{V_1}$ and $\mathbf{V_2}$. The vectors that comprise $\mathbf{V_2}$ span the "calibration null space" (Tonkin and Doherty 2005), which consists of the model null space and the orthonormal basis vectors of **V** with corresponding near-zero eigenvalues. Near-zero singular values amplify data noise and lead to potential instability during the inverse process (i.e., calibration). The calibration null space often includes system detail represented in the model that is simply beyond the reach of the calibration dataset to infer. The remaining orthonormal basis vectors of **V** define $\mathbf{V_1}$, and they span the "calibration solution space" (Tonkin and Doherty 2005). The generalized inverse solution is expressed as a linear combination of vectors occupying the $\mathbf{V_1}$ subspace. Decomposition of **V** into $\mathbf{V_1}$ and $\mathbf{V_2}$ is based on the specification of a regularization parameter that defines truncation. This can be based on, among other things, an examination of the singular value spectrum, or in the case described herein, specified in a manner to minimize the error variance for a specific model prediction.

B. Predictive error variance is the sum of two explicit terms:

a. The degree to which measurement and structural noise contribute to model error, and

b. The contribution from the calibration null space to predictive error variance. If a prediction is sensitive to linear combinations of parameters that cannot be inferred through the calibration process (and this lies within the calibration null space), then the error variance of that prediction is unchanged from what it would have been if the model had not been calibrated at all.

C. If regularization is not required in solution of the inverse problem because the system is simple enough and the calibration dataset is informative enough for the null space to be reduced to zero dimensions, predictive error variance is solely the contribution resulting from model calibration (the second term of equations [1a], [1b], and [1c]), because the model resolution matrix, **R** =

$\mathbf{VV}^t = \mathbf{I}$, is equal to the identity matrix for overdetermined systems. Model predictive error variance is thus solely a function of measurement noise.

D. If a given prediction is strongly influenced by uncapturable system detail (because it is sensitive to that detail), then the first term represented in equations (1a), (1b), and (1c), for the predictive error variance, may be quite large in relation to the second.

E. For an uncalibrated model, the second term in equations (1a), (1b), and (1c) is zero, the model resolution matrix is the null matrix, and total predictive error variance is equal to $\mathbf{y}^t C(\mathbf{p})\mathbf{y}$, as expected.

F. Where a model is a perfect replicator of reality (and hence there is no structural component of measurement noise), and where the observational component of measurement noise is zero, a perfect fit between model outputs and field measurement could be obtained. In this case, the second term of the above equations is zero, but the first may still be substantial.

G. In the extreme case that the prediction sensitivity is parallel to a truncated eigenvector (in $\mathbf{V_2}$), then the calibration process does nothing to reduce the uncertainty of that prediction, for the second term is zero regardless of the amount of measurement noise.

H. On the other hand, if the prediction sensitivity is parallel to a retained eigenvector (in $\mathbf{V_1}$), then the first term in equations (1a), (1b), and (1c) is zero, the second term is nonzero, and its magnitude depends upon the amount of measurement /structural noise associated with the calibration dataset (which the calibration process can provide information on) and also on the relative magnitude of the eigenvalue, within the singular value spectrum, that corresponds to the eigenvector to which $\mathbf{y}$ is parallel.

I. In general, as fewer eigenvectors are truncated, the first term in equations (1a), (1b), and (1c) falls approximately in a monotonic manner, whereas the second term rises in an approximately monotonic manner. Moreover, assuming the observation data set contains sufficient information for the calibration process to reduce predictive error variance, a minimum for the predictive error variance is typically achieved at a specific truncation level.

J. Among others, the equations for predictive error variance can be used prior to the calibration process to determine the optimal singular value truncation level (which is roughly related to model complexity) for a given prediction. At this optimum level, the predictive error variance is reduced to as low as it can be, given the data available for calibration, and the noise associated with this data.

K. Equations (1a) through (1c) do not involve the use of actual field data, only the sensitivities of the model outputs (corresponding to field measurements) with respect to the input parameters. Hence "notional calibration exercises" can be undertaken with different posited measurements included in the calibration process. The worth of different data acquisition strategies can thus be compared in terms of their ability to reduce model predictive error variance.

L. By repeatedly recalculating model predictive error variance under the assumption that certain parameters, or parameter types, are perfectly known (and hence need not be estimated through the calibration process), the reduction in predictive uncertainty accrued through this process can be used as a measure of that parameter's or parameter type's contribution to model predictive error.

Moore and Doherty (2005) also discuss the consequences of parameter lumping, or zonation, on the computation of predictive error variance. See Moore and Doherty (2005) for details of the derivation of equations (1a), (1b), and (1c), a lengthier discussion of their meaning, and an example application. Equations (1a), (1b), and (1c) can be computed using utility programs that are now part of the public domain software PEST (Parameter ESTimation) (Doherty 2005). We used PEST (Doherty 2004), and the PEST surface water utilities (Doherty 2003) for the analysis presented in the following section.

As we stated previously, this analysis is based on a linearity assumption and may thus be somewhat in error as far as predictions of actual error variance are concerned. However, while the exact results will be in error, the analysis will be informative in a relative sense and will allow useful conclusions to be drawn. One could also apply the nonlinear extension of the theory presented in Moore and Doherty (2005). However, unlike linear analysis, which, as discussed above, relies only on characterizations of model predictive noise rather than on the exact values of the measurements themselves, nonlinear analysis can proceed only in the context of a specific calibration dataset.

3 ANALYSIS

The objectives of the analysis presented herein, using the linear analysis of Moore and Doherty (2005), the FASST model, and calibration data for Yuma, Arizona (Frankenstein and Koenig, 2004), are to

1. Compute the notional predictive error variance for two soil moisture predictions (one immediately following a storm event and the other following a period with no precipitation);
2. For each prediction, determine the minimum notional predictive error variance that is achievable with the available calibration data;
3. Conduct a notional data acquisition exercise to examine the reduction that could be accrued in the error variance for a given soil moisture prediction through the acquiring of such extra data;
4. Compute the contribution to overall soil moisture predictive error variance from each individual adjustable model parameter.

With respect to items 1 and 2, the notional predictive error variance analysis for the two soil moisture predictions involved calculating $\sigma^2_{s-\hat{s}}$ at all possible truncation levels. The notional analysis described in item 3 required building a new time series of model weather input, assigning new observation weights, and recalculating $\sigma^2_{s-\hat{s}}$. The contribution to overall soil moisture predictive error variance from each individual adjustable model parameter was determined by

I. Recording the minimum of the predictive error variance curve (i.e., the curve of predictive error variance versus the number of singular values before truncation) for each prediction. At this stage all parameters are involved in the parameter estimation process.

II. Removing one parameter from the parameter estimation process and repeating the exercise, in each case finding the minimum of the error variance curve.

III. Comparing minima obtained in items I and II to establish the contribution to the predictive error variance from the selected parameter.

With the assumption that the limited observation dataset that is available for the Yuma site would be used in its entirety for model calibration, predictions were based on a dual model run (Doherty 2004). The only difference between the weather dataset employed for the predictions and that used for the calibration effort is that the magnitude of all precipitation data was doubled (Fig. 1). For convenience, the date and times associated with the weather data for the pre-

diction run was set equal to the corresponding date and time for the observed weather data. Two surface soil moisture predictions were specified, one immediately following the two storm events that occurred on 26 March 1993, and the other after a period with no precipitation. Table 1 summarizes the specifics for the two predictions.

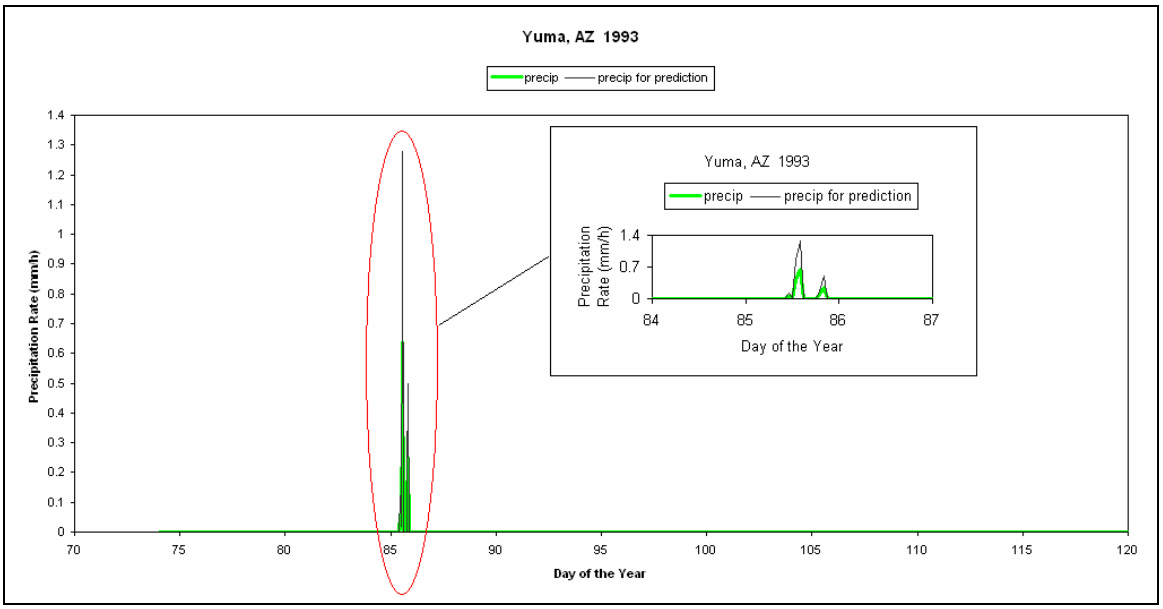

Figure 1. Precipitation data used for calibration and prediction.

Table 1. Surface soil moisture predictions	
Prediction	**Description**
y_1	Maximum surface soil moisture for 27 March 1993 (86th day of the year)
y_2	Mean surface soil moisture for 26 April 1993 (116th day of the year)

3.1 Compute Notional Predictive Error Variance; Determine Minimums

We used Equation 1 to compute notional predictive error variance for the two specified predictions using truncated singular valued decomposition (TSVD). This involved

1. Interfacing PEST with FASST for the Yuma site

   a. Specifying adjustable model parameters and subsequently preparing PEST template files. The names and meanings of the sixteen adjustable model parameters are listed in Table 2. Given that iddm = bddm/(1.0 – por), vfs = 1.0 – por, void = por/(1.0 – por), and mwc = por, there are in fact only twelve adjustable model parameters. It can also be argued that tcbdm depends on bddm (Farouki 1981), but for this study, we consider them to be independent of one another. In order to better accommodate scaling issues resulting from the use of different units for different parameters, and in an attempt to decrease the degree of nonlinearity of the parameter estimation problem, we estimated the logs of these parameters instead of their native values. Past experience has demonstrated that greater efficiency and stability of the parameter estimation process can often be achieved through this means (Doherty and Skahill in press).

   b. Writing software to extract and process simulated FASST surface soil moisture data and surface temperature data into the "site sample file" format (Doherty 2003).

   c. Specifying weights for the observed surface soil moisture data (see Fig. 2) and the observed surface temperature data (see Fig. 3) to be equal to the inverse of the assumed standard deviation of the observation noise (Aster et al. 2005).

   d. Using PEST surface water utilities to process observed and simulated surface soil moisture data and surface temperature data, and the specified predictions $\mathbf{y_1}$ and $\mathbf{y_2}$ (which were both assigned a weight of zero), to support the preparation of a PEST control file that then would guide subsequent calibration and predictive error analyses.

Table 2. Name and meaning of FASST input parameters.

Name	Meaning
bddm	bulk density of dry material (g/cm^3)
iddm	intrinsic density of dry material (g/cm^3)
vfs	volume fraction of solids (0.0 – 1.0)
por	porosity (0.0 – 1.0)
void	void ratio (0.0 – 1.0)
sani	Surface albedo – normal incidence
lse	Longwave surface emissivity
qc	quartz content (0.0 – 1.0)
of	organic fraction (0.0 – 1.0)
tcbdm	thermal conductivity of the bulk dry material (W/m*K)
shdm	specific heat of dry material (J/kg*K)
shc	saturated hydraulic conductivity (cm/s)
rwc	residual water content (vol/vol) (0.0 – 1.0)
mwc	maximum water content (vol/vol) (0.0 – 1.0)
vGBph	van Genuchten Bubbling pressure head (cm)
vGe	van Genuchten exponent (n)

2. Computing derivatives for all model outputs (including the two predictions [see Table 1]) with respect to all adjustable model parameters.

3. Defining $C(\mathbf{p})$, a pre-calibration assessment of the inherent variability of each parameter, based on expert knowledge of the system and system properties (Guymon et al. 1993, Koenig 1994, Sullivan et al. 1997, Schaap 1999, Peck 2002, Frankenstein and Koenig 2004), to be diagonal and each diagonal term to be equal to the square of the standard deviation associated with each adjustable model parameter. We approximated the standard deviation in each case as follows: $\sigma_i = 0.25^* [\log_{10}(UB) - \log_{10}(LB)]$, where UB is the parameter's upper bound and LB is the parameter's lower bound.

4. Defining an observation reference variance, σ^2_r, where it is assumed that $C(\boldsymbol{\varepsilon}) = \sigma^2_r\mathbf{Q}$. We estimated this on the basis of model to measurement misfit achieved through a calibration process.

5. Computing the predictive error variance at all possible truncation levels.

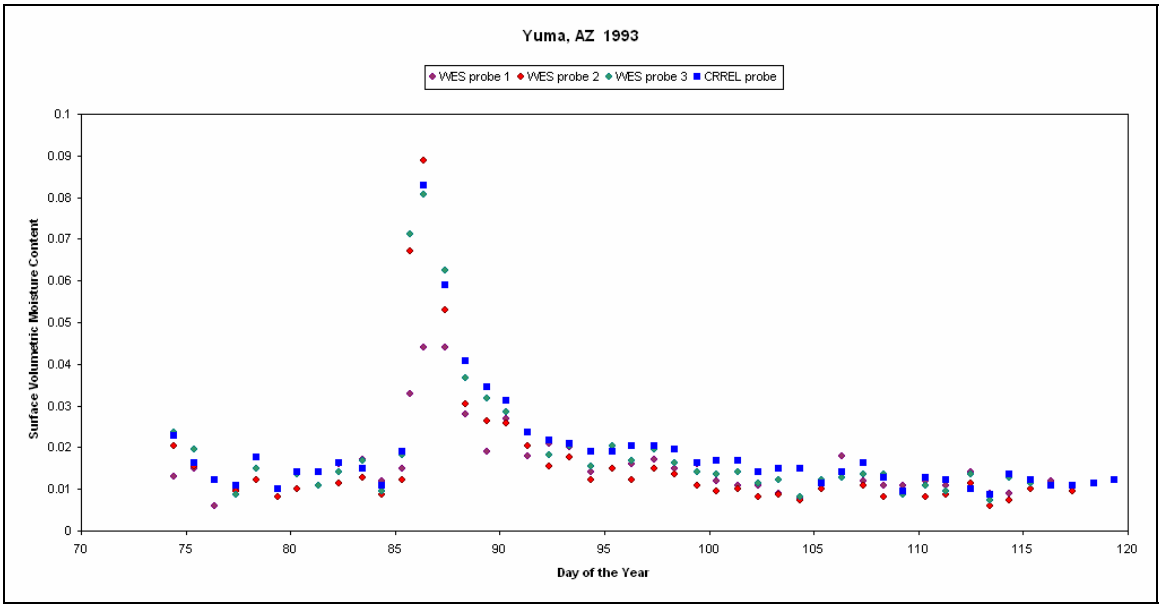

Figure 2. Observed surface soil moisture data. (See Frankenstein and Koenig [2004] for additional details.)

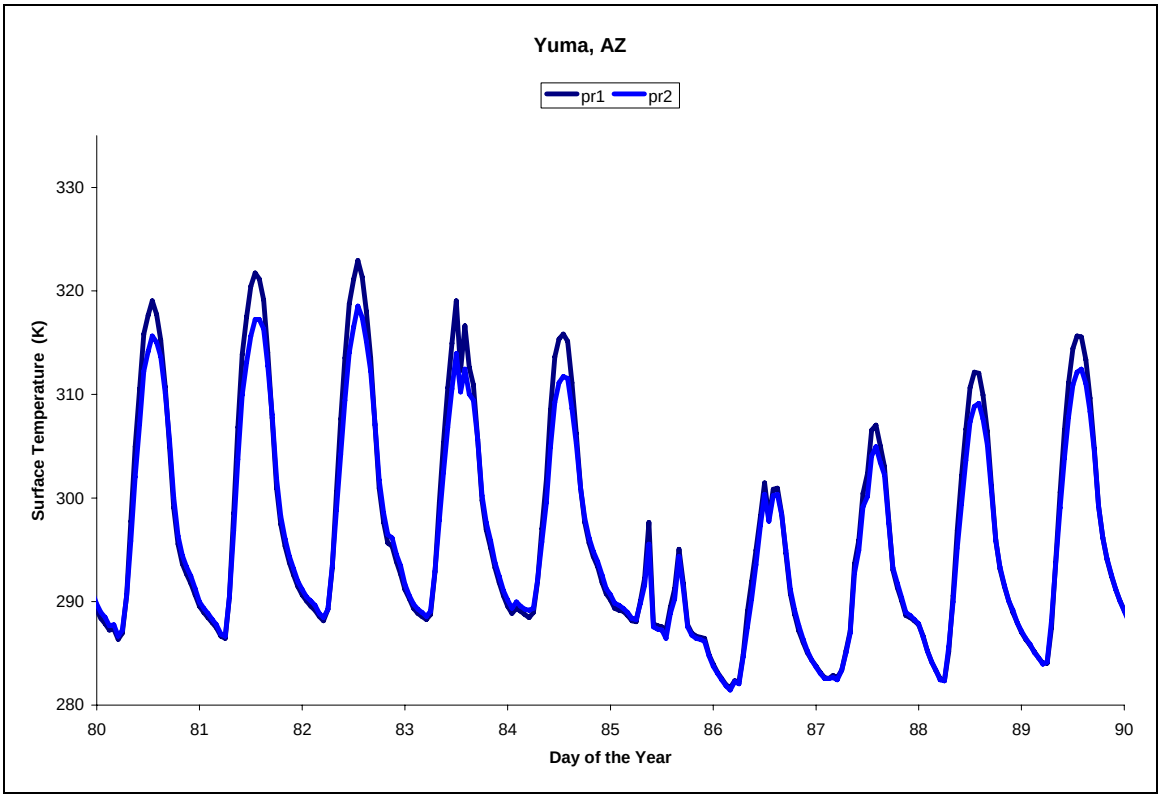

Figure 3. Observed surface temperature data. (See Frankenstein and Koenig [2004] for additional details.)

Results from the notional predictive error analysis are presented in Table 3 and Figure 4 for the first prediction, $\mathbf{y_1}$, and Table 4 and Figure 5 for the second prediction, $\mathbf{y_2}$.

Table 3. Notional predictive error variance for y_1.

Singular values	1st term	2nd term	Total variance	Standard deviation
0	1.33E-02	0	1.33E-02	0.1154064
1	1.31E-02	1.04E-08	1.31E-02	0.114401
2	1.31E-02	1.08E-08	1.31E-02	0.1144537
3	1.29E-02	2.82E-08	1.29E-02	1.14E-01
4	1.21E-02	1.69E-07	1.21E-02	1.10E-01
5	1.13E-02	1.04E-06	1.13E-02	1.06E-01
6	1.00E-02	8.38E-06	1.01E-02	0.1002882
7	1.05E-02	8.76E-06	1.05E-02	0.1024523
8	3.78E-03	4.74E-05	3.83E-03	6.19E-02
9	1.21E-04	1.96E-04	3.18E-04	1.78E-02
10	1.28E-04	1.97E-04	3.24E-04	1.80E-02
11	2.35E-08	4.68E-04	4.68E-04	2.16E-02
12	0	4.69E-04	4.69E-04	2.17E-02

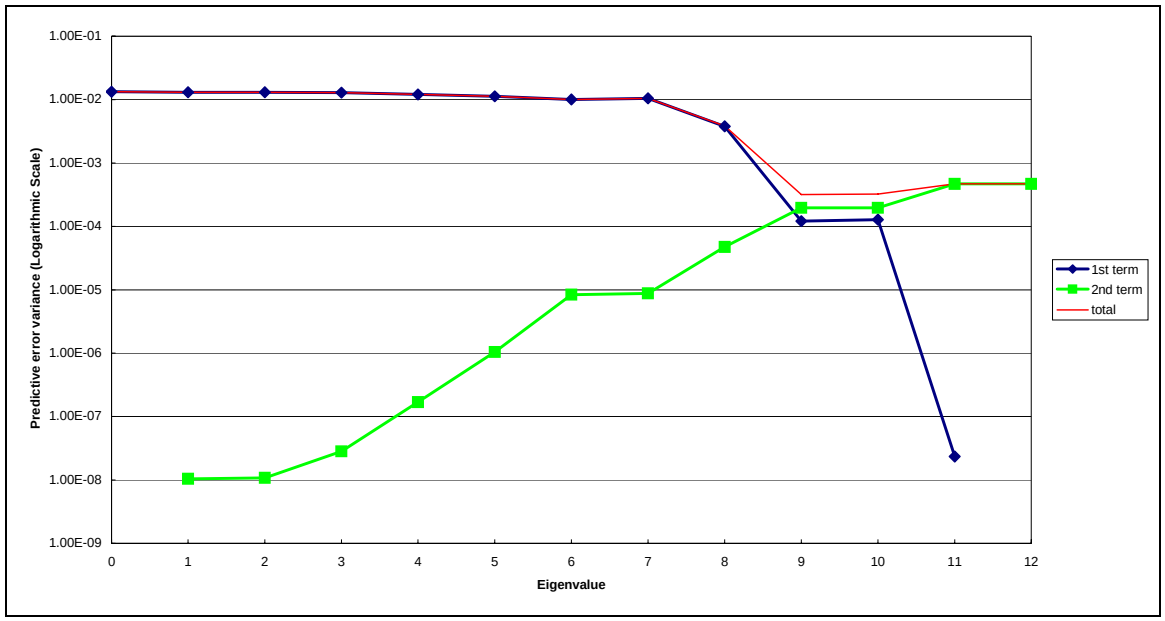

Figure 4. Notional predictive error variance for y_1.

Table 4. Notional predictive error variance for y_2.				
Singular values	**1st term**	**2nd term**	**Total variance**	**Standard deviation**
0	3.72E-03	0	3.72E-03	6.10E-02
1	1.46E-04	7.27E-08	1.46E-04	1.21E-02
2	4.21E-05	9.52E-08	4.22E-05	6.49E-03
3	2.28E-05	1.16E-07	2.29E-05	4.78E-03
4	1.20E-06	1.51E-07	1.35E-06	1.16E-03
5	7.20E-07	1.55E-07	8.75E-07	9.35E-04
6	9.71E-07	1.60E-07	1.13E-06	1.06E-03
7	1.04E-06	1.61E-07	1.20E-06	1.10E-03
8	3.44E-08	1.83E-07	2.17E-07	4.66E-04
9	6.46E-08	1.84E-07	2.49E-07	4.99E-04
10	3.44E-08	1.89E-07	2.23E-07	4.72E-04
11	6.22E-13	2.62E-07	2.62E-07	5.12E-04
12	0	2.62E-07	2.62E-07	5.12E-04

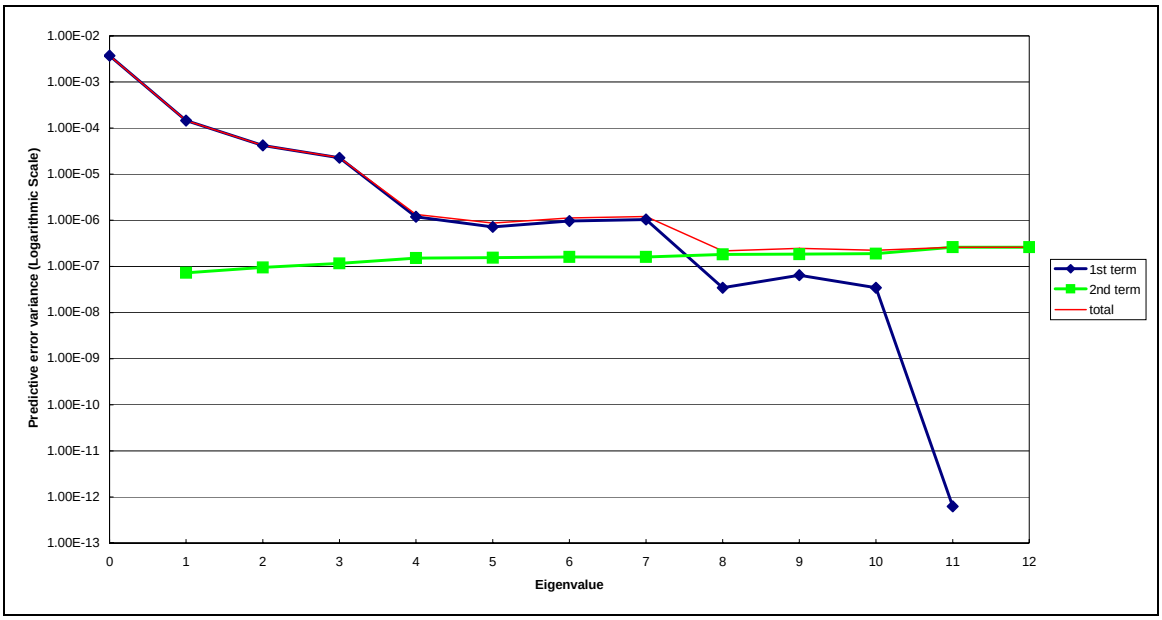

Figure 5. Notional predictive error variance for y_2.

Examining the results for the notional predictive error analysis for the two specified predictions, in each case we see

1. The second and first term rising and falling, respectively, in an approximate monotonic manner as fewer eigenvectors are trunctated.
2. The minimum predictive error variance occurring at a specific truncation level.

We calculated the pre-calibration predictive error variance using zero singular values. In this case, the second term in equations (1a), (1b), and (1c) is zero, as noted above (also see Tables 3 and 4). The difference between the pre-calibration predictive error variance and the minimum predictive error variance is a measure of the worth of the calibration process in reducing predictive error variance. Hence, for both $\mathbf{y_1}$ and $\mathbf{y_2}$, calibration would reduce predictive error variance. However, a greater reduction in predictive error variance would be accrued through calibration for the second prediction, by a factor of approximately 5.8E-05, relative to the first prediction, where the reduction is approximately 2.4E-02. Results from the notional predictive error analysis also indicate that there would be more predictive uncertainty associated with the prediction immediately following the two storm events than the prediction following a period with no precipitation. We also can observe that the minimum predictive error variance for $\mathbf{y_1}$ required using nine singular values, whereas the minimum predictive error variance for $\mathbf{y_2}$ required using eight singular values. In summary, the notional predictive error variance analysis indicated that

- The calibration process would be of value in reducing predictive uncertainty for both predictions ($\mathbf{y_1}$ and $\mathbf{y_2}$), but that a much greater reduction would be achieved, by model calibration, for the second prediction relative to the first prediction.
- There would be greater predictive uncertainty associated with soil moisture prediction using FASST for the Yuma site for periods immediately following storm events in comparison to periods with no precipitation.

These findings are likely a function of the observation data set that we used for model calibration, in particular, to the variance associated with the soil moisture observations (see Figure 2) immediately following the two storm events (both occurring on 26 March 1993 [Fig. 1], and also to the fact that just two storm events occurred during the entire calibration period (15 March 1993–30 April 1993).

3.2 Data Acquisition and Reduction in Error Variance

We performed a single notional data acquisition exercise to examine the reduction that could be achieved in reducing the error variance for the soil moisture prediction immediately following the two manufactured storm events. Rather than construct an altogether new weather input data file, we recalculated the weights for the observed soil moisture data immediately following the storms to be commensurate with the weights assigned for the remaining observed soil moisture data. Hence, this notional data acquisition exercise examined the benefits that could be gained in terms of reducing predictive error variance for $\mathbf{y_1}$ based on the assumption that soil moisture observations following storm events were collected with greater precision. The minimum error variance for $\mathbf{y_1}$ was reduced by a factor of approximately 38 percent (see Table 5).

Table 5. Notional predictive error variance for y_1 based on assumed data acquisition strategy.

Singular values	1st term	2nd term	Total variance	Standard deviation
0	1.33E-02	0	1.33E-02	0.1154064
1	1.31E-02	1.03E-08	1.31E-02	0.1143415
2	1.31E-02	1.06E-08	1.31E-02	0.1143799
3	1.29E-02	2.76E-08	1.29E-02	1.14E-01
4	1.20E-02	1.63E-07	1.20E-02	1.10E-01
5	1.11E-02	1.09E-06	1.11E-02	1.06E-01
6	9.40E-03	8.58E-06	9.41E-03	9.70E-02
7	9.91E-03	9.24E-06	9.91E-03	9.96E-02
8	2.54E-03	4.92E-05	2.59E-03	5.09E-02
9	3.82E-05	1.65E-04	2.03E-04	1.42E-02
10	2.99E-05	1.68E-04	1.98E-04	1.41E-02
11	1.27E-08	2.15E-04	2.15E-04	1.47E-02
12	0	2.16E-04	2.16E-04	1.47E-02

3.3 Contribution to Predictive Error Variance from Individual Parameters

We computed the contribution to overall soil moisture predictive error variance from each individual adjustable model parameter as described above. This allows one to assess the relative importance of each model parameter to a specific prediction. The results of the analysis are presented in Tables 6 and 7 for the first and second prediction, respectively.

Table 6. Percent reduction accrued in minimum predictive error variance for y_1 assuming each parameter, in turn, is perfectly known.

Adjustable parameter	% reduction in minimum predictive error variance	
	Before calibration	After calibration
bddm	0.64	10.43
por	0.09	–2.36
sani	0.00	0.52
lse	0.00	–0.35
qc	0.63	3.35
of	0.00	0.01
tcbdm	0.38	21.07
shdm	0.00	–3.19
shc	94.24	71.60
rwc	0.06	38.58
vgbph	0.09	17.56
vge	3.86	–8.08

Table 7. Percent reduction accrued in minimum predictive error variance for y_2 assuming each parameter, in turn, is perfectly known.

Adjustable parameter	% reduction in minimum predictive error variance	
	Before calibration	After calibration
bddm	0.00	0.01
por	0.14	–1.47
sani	0.00	5.40
lse	0.00	8.74
qc	0.00	2.27
of	0.00	0.00
tcbdm	0.00	–0.57
shdm	0.00	–4.65
shc	0.00	0.45
rwc	0.21	17.03
vgbph	0.14	11.55
vge	99.51	–1.58

Examining the contents of Table 6 and Table 7, the following observations can be made:

1. For prediction $\mathbf{y_1}$
    a. Before calibration,
        i. Obtaining prior information about the saturated hydraulic conductivity would, by far, go the farthest in reducing the predictive error variance.
        ii. Obtaining prior information concerning the van Genuchten exponent would be the second, although distant, most important parameter in terms of reducing the predictive error variance.
        iii. Obtaining prior information concerning the bulk density of dry material, quartz content, thermal conductivity of the bulk dry material, porosity, van Genuchten bubbling pressure head, and the residual water content would result only in a possible slight reduction in predictive error variance.
        iv. Obtaining prior information concerning surface albedo-normal incidences, longwave surface emissivity, organic fraction, and specific heat of dry material would be of no value in reducing predictive error variance.
    b. After calibration,
        i. The parameter estimate for the saturated hydraulic conductivity contributes the most to predictive error variance.
        ii. In decreasing order, the parameter estimates obtained for the residual water content, the thermal conductivity of the bulk dry material, the van Genuchten bubbling pressure head, and the bulk density of dry material are also important to predictive error variance.
        iii. Knowledge obtained through the calibration process for the following parameters is of little to no importance to the predictive error variance.
            1. por
            2. sani
            3. qc
            4. of
            5. shdm
            6. vge

2. For prediction $\mathbf{y_2}$
   a. Before calibration,
      i. Obtaining prior information concerning the van Genuchten exponent would explain almost all of the predictive error variance.
      ii. Obtaining prior information concerning the porosity, van Genuchten bubbling pressure head and the residual water content would possibly result in a slight reduction in predictive error variance.
      iii. Obtaining prior information concerning the remaining parameters would serve no value in reducing the predictive error variance.
   b. After calibration,
      i. In decreasing order, the parameter estimates for the residual water content, the van Genuchten bubbling pressure head, longwave surface emissivity, surface albedo-normal incidences, and quartz content are of most importance to predictive error variance.
      ii. The remaining parameters are of little to no value to predictive error variance.

For periods immediately following storm events, parameter estimates obtained through the calibration of soil properties related to water movement and retention, especially the saturated hydraulic conductivity (shc), would be most important to the soil moisture predictive error variance. For soil moisture prediction following periods with no precipitation, parameter estimates obtained through the calibration of soil properties related to both the soil hydraulic and thermal properties both would be of value to minimizing the predictive error variance.

For both predictions, the organic fraction parameter (of) was determined to be of no importance to predictive error variance before or after calibration because that component of the prediction sensitivity vector was zero for the second prediction and almost zero for the first (see Appendix A and Doherty [2005]).

3.4 Calibration

After rescaling the weights for the two observation groups such that they each made an equal contribution to the objective function at the beginning of the calibration process, we employed truncated singular valued decomposition to support regularized inversion of FASST for the Yuma site. Truncation was specified to occur when the ratio of all remaining eigenvalues to the maximum system eigenvalue was less than 10^{-10}. The Marquardt lambda was not employed to further stabilize the parameter estimation process. Table 8 lists the estimated parameter values obtained using TSVD. Examining the diagonal elements of the model resolution matrix listed in Appendix B, it is apparent that the parameters bddm, por, lse, tcbdm, shdm, and vge were estimated well whereas the parameters sani, qc, and of were estimated poorly.

Table 8. Estimated model parameter values obtained for the FASST model for Yuma, Arizona, using TSVD.

Parameter	Value
bddm	1.140000
por	0.5045182
sani	0.3500000
lse	0.9900000
qc	0.1500000
of	1.9995444E-06
tcbdm	0.3486975
shdm	850.6000
shc	6.2560347E-03
rwc	4.2122053E-03
vgbph	26.55427
vge	1.366374

Figures 6 and 7 are plots of the calibrated model output along with the measured data.

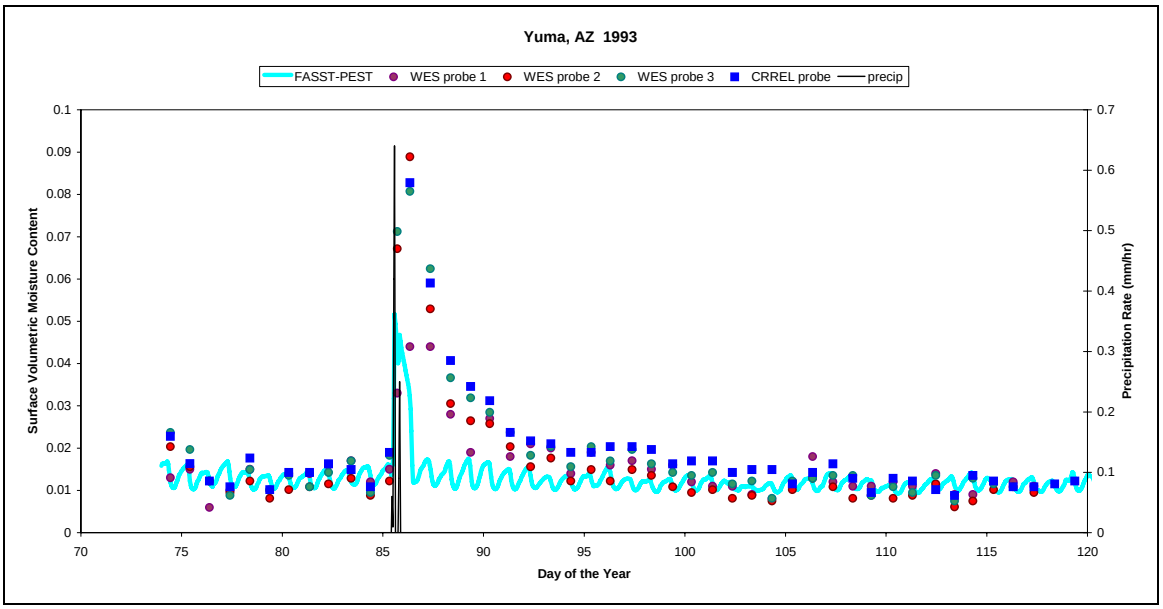

Figure 6. Plot of simulated surface soil moisture and observed surface soil moisture.

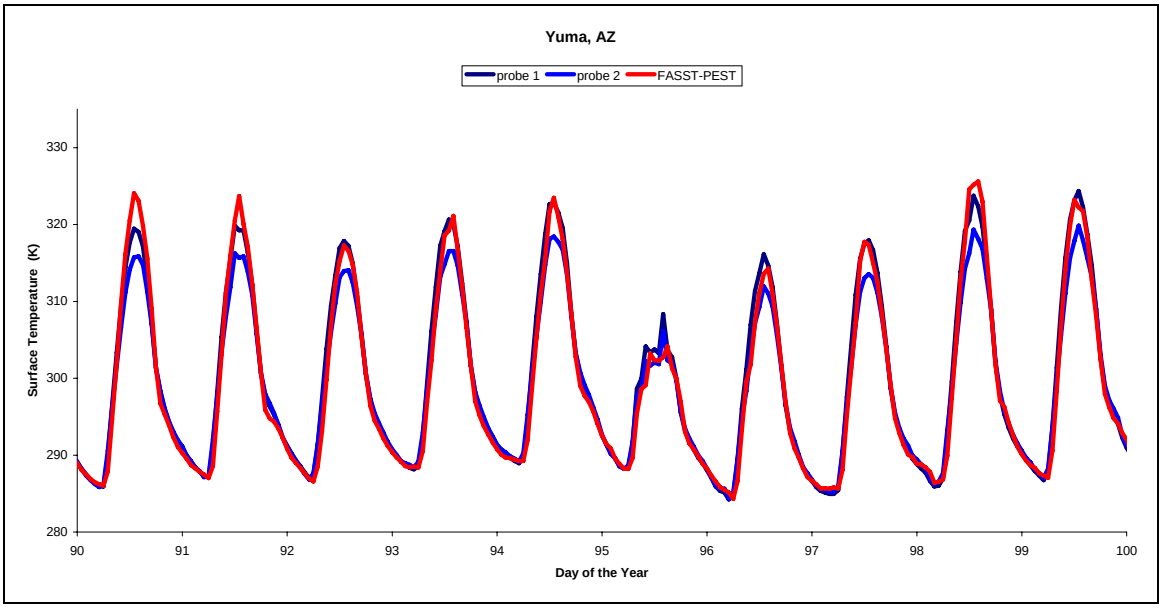

Figure 7. Plot of simulated surface soil temperature and observed surface soil temperature.

4 CONCLUSIONS

We employed the linear analysis of Moore and Doherty (2005) using the FASST model with data collected at Yuma, Arizona (Frankenstein and Koenig 2004) to

1. Quantify and compare notional predictive error variance for two specified soil moisture predictions (one immediately following a storm event and the other following a period with no precipitation),
2. Examine the role of the calibration process in reducing notional predictive error variance for the two specified soil moisture predictions,
3. Measure the worth of additional data in terms of how much it would reduce soil moisture prediction error,
4. Examine the contribution from each individual model parameter to overall soil moisture predictive error variance.

The underlying intent of the notional predictive error variance analysis presented herein for the one-dimensional state-of-the-ground dynamic model FASST, using observed data for the Yuma site, was to demonstrate a method that could later be used to assess the relative worth of vertical processes versus horizontal processes for soil moisture prediction. In the future, either a multi-dimensional version of FASST or an alternate model will be used to address these issues.

A “simpler,” effectively vertical, one-dimensional model could be built from a more complex multi-dimensional model by fixing parameter values in space. Hence, the complex model could be used to compute the predictive error variance for the “simpler” model built from it. The complex model could also be used to calculate the predictive error variance when full heterogeneity is allowed. If the predictive error variance for the “simpler,” effectively vertical, one-dimensional model is not much greater than that of the complex model, then the “simpler” model is just as good, because the observation dataset used to support model calibration does not have the information content required to support the complex model anyway. Also, one could determine the contribution to the error variance of key model predictions made by different parameter types before and after calibration. If there are parameter types that make a large contribution before calibration but a reduced contribution after calibration, and if these parameters pertain to processes not represented, or effectively eliminated, in the “simpler” model, then this says that, whatever simplification takes place, these parameters and processes must be included because (a) the prediction depends on them and

(b) the calibration dataset contains information about them and hence the complex model is required for a complete calibration. Hence, this would say something about the cost of simplification—or at least certain (inappropriate) types of simplification. Moreover, if a prediction is dependent on some processes/parameters that are in the complex model, and if the contribution to uncertainty of those parameters is large and cannot be reduced through calibration, then the complex model can potentially give wrong answers. Error variance analysis could be used to determine whether the prediction would be any more wrong if these processes were ignored or vastly simplified.

Lastly, as was noted above for an individual parameter associated with the one-dimensional FASST model, if the prediction is insensitive to certain processes in the complex model, then those processes could be omitted.

REFERENCES

Aster, R.C., B. Borchers, and C.H. Thurber (2005) *Parameter Estimation and Inverse Problems*, Elsevier, Inc., 301 p.

Doherty, J. (2003) *PEST: Surface Water Utilities*. Brisbane, Australia, and University of Idaho, Idaho Falls, Idaho: Watermark Numerical Computing.

Doherty, J. (2004) *PEST: Model-Independent Parameter Estimation, User Manual, Version 5*. Brisbane, Australia: Watermark Numerical Computing.

Doherty, J. (2005) *Addendum to the PEST Manual*. Brisbane, Australia: Watermark Numerical Computing.

Doherty, J., B. Skahill (in press) An advanced regularization methodology for use in watershed model calibration. Accepted for publication in *Journal of Hydrology*.

Farouki, O.T. (1981) Thermal Properties of Soils. U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire, Monograph 81-1.

Frankenstein, S., and G. Koenig (2004) Fast All-season Soil STrength (FASST). U.S. Army Engineer Research and Development Center, Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire, Special Report SR-04-1.

Moore, C., and J. Doherty (2005) Role of the calibration process in reducing model predictive error. *Water Resources Research*, **41**, W05020, doi:10.1029/2004WR003501.

Guymon, G.L., R.L. Berg, and T.V. Hromadka (1993) Mathematical Model of Frost Heave and Thaw Settlement in Pavements. U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire, CRREL Report 93-2.

Koenig, G.G. (1994) Smart Weapons Operability Enhancement (SWOE) Joint Test and Evaluation (JT & E) Program Final Report. Dr. James P. Welch, Joint Test Director, SWOE JT & E, SWOE Report 94-10, Annex D.

Moore, C., and J. Doherty (2005) Role of the calibration process in reducing model predictive error. *Water Resources Research*, **41**, W05020, doi:10.1029/2004WR003501.

Peck, L. (2002) Battlespace Terrain Reasoning and Awareness Soil Properties Database Population. U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire.

Schaap, M.C. (1999) ROSETTA Model, Version 1.0. http://www.ars.usda.gov/Services/docs.htm?docid=8953.

Sullivan, P.M., C.D. Bullock, N.A. Renfroe, M.R. Albert, G.G. Koenig, L. Peck, and K. O'Neill (1997) Soil Moisture Strength Prediction Model Version II (SMSP II). U.S. Army Corps of Engineers Waterways Experiment Station, Vicksburg, Mississippi, Technical Report GL-97-15.

Tonkin, M.J., and J. Doherty (2005) A hybrid regularized inversion methodology for highly parameterized environmental models. *Water Resources Research*, **41**, W10412, doi:10.1029/2005WR003995.

APPENDIX A. PREDICTION SENSITIVITIES FOR y_1 AND y_2

Prediction sensitivities for y_1

12 1 2

8.9044822E-02

3.7992827E-02

7.8319554E-03

1.1813355E-02

–2.9261424E-02

–3.2721702E-19

–6.2365748E-02

6.6851879E-03

9.8464122E-02

1.0933653E-02

1.0997549E-02

–0.1438406

* row names

bddm

por

sani

lse

qc

of

tcbdm

shdm

shc

rwc

vgbph

vge

* column names

pred1_max

Prediction sensitivities for y_2

12 1 2

3.1413017E-04

2.4589594E-02

5.2337759E-04

2.5389455E-03

9.7859866E-05

0.000000

–3.2236191E-06

–2.8450404E-04

–1.1299936E-04

1.1236558E-02

7.3991269E-03

–0.3860382

* row names

bddm

por

sani

lse

qc

of

tcbdm

shdm

shc

rwc

vgbph

vge

* column names

pred2_mean

APPENDIX B. MODEL RESOLUTION MATRIX ASSOCIATED WITH CALIBRATED MODEL

12 12 1

9.98E-01	-3.85E-03	-8.95E-15	-1.29E-03	2.38E-03	9.96E-09	-9.35E-03	-1.35E-03	-4.01E-02	4.44E-03	1.29E-02	2.87E-04
-3.85E-03	9.92E-01	-2.51E-15	-2.42E-03	-2.16E-03	-1.57E-07	-1.94E-02	-2.73E-03	-7.72E-02	-1.46E-03	4.19E-02	5.99E-04
-8.95E-15	-2.51E-15	1.25E-28	-1.09E-15	2.65E-17	3.86E-22	4.84E-15	1.31E-15	4.61E-16	2.88E-15	1.91E-15	2.40E-17
-1.29E-03	-2.42E-03	-1.09E-15	9.99E-01	9.10E-04	3.62E-10	-6.05E-03	-9.01E-04	-2.90E-02	7.94E-03	1.23E-03	1.85E-04
2.38E-03	-2.16E-03	2.65E-17	9.10E-04	7.61E-05	3.07E-09	3.47E-03	1.22E-03	1.09E-03	6.00E-03	3.76E-03	1.41E-04
9.96E-09	-1.57E-07	3.86E-22	3.62E-10	3.07E-09	5.17E-13	-3.77E-07	-4.24E-08	2.35E-07	4.83E-07	2.44E-07	-1.40E-08
-9.35E-03	-1.94E-02	4.84E-15	-6.05E-03	3.47E-03	-3.77E-07	9.53E-01	-6.62E-03	-1.91E-01	7.21E-03	8.51E-02	1.43E-03
-1.35E-03	-2.73E-03	1.31E-15	-9.01E-04	1.22E-03	-4.24E-08	-6.62E-03	9.99E-01	-2.82E-02	2.78E-03	9.61E-03	2.03E-04
-4.01E-02	-7.72E-02	4.61E-16	-2.90E-02	1.09E-03	2.35E-07	-1.91E-01	-2.82E-02	1.11E-01	2.07E-01	1.01E-01	5.87E-03
4.44E-03	-1.46E-03	2.88E-15	7.94E-03	6.00E-03	4.83E-07	7.21E-03	2.78E-03	2.07E-01	6.97E-01	4.10E-01	-1.94E-04
1.29E-02	4.19E-02	1.91E-15	1.23E-03	3.76E-03	2.44E-07	8.51E-02	9.61E-03	1.01E-01	4.10E-01	2.50E-01	-2.66E-03
2.87E-04	5.99E-04	2.40E-17	1.85E-04	1.41E-04	-1.40E-08	1.43E-03	2.03E-04	5.87E-03	-1.94E-04	-2.66E-03	1.00E+00

* row and column names

bddm

por

sani

lse

qc

of

tcbdm

shdm

shc

rwc

vgbph

vge

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
April 2006

2. REPORT TYPE
Technical Report

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
FASST Soil Moisture Prediction Error at Yuma, Arizona

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Brian E. Skahill and Susan Frankenstein

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army ERDC
CRREL
72 Lyme Road
Hanover NH 03755-1290

U.S. Army ERDC
CHL
3909 Halls Ferry Road
Vicksburg MS 39180-6199

8. PERFORMING ORGANIZATION REPORT NUMBER
ERDC TR-06-1

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Corps of Engineers
Washington, DC 20314-1000

10. SPONSOR/MONITOR'S ACRONYM(S)
USACE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

Available from NTIS, Springfield, Virginia 22161.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

A linear predictive error variance analysis was employed using the one-dimensional state-of-the-ground dynamic model FASST (Fast All-season Soil STrength) and calibration data for Yuma, Arizona (Frankenstein and Koenig 2004). The analysis was performed with the intent to quantify soil moisture predictive error and to examine ways in which it could be reduced, in particular, to demonstrate a methodology wherein one can examine the contribution from each individual model parameter to overall soil moisture predictive error variance.

15. SUBJECT TERMS
FASST
Model calibration
Predictive error variance
Regularization
Soil moisture

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	U	37	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. 239.18