# Linear and Nonlinear Uncertainty Analysis

## for the

## NFSEG Groundwater Model

**Watermark Numerical Computing**

**May, 2018**

## **Table of Contents**

# 1. Introduction

During March and April 2018, linear and nonlinear parameter uncertainty analyses were undertaken for the NFSEG groundwater model constructed by St Johns River and Suwannee River Water Management Districts. Linear analysis comprised calculation of parameter identifiabilities and uncertainty reductions accrued through the calibration process. Nonlinear analysis culminated in the generation of 522 parameter fields, all of which can be considered as reasonable, and all of which provide a fit with the calibration dataset which is almost as good as that provided by the parameter field which is deemed to calibrate the model. The uncertainty of any prediction made by the NFSEG model can be assessed by making the prediction with all of these parameter fields.

# 2. Theory

## 2.1 General

Nonlinear model predictive uncertainty analysis is generally implemented using Monte Carlo methods. It requires implementation of a procedure through which the posterior parameter probability distribution is sampled in order to generate a suite of parameter fields. While all of these parameter fields must satisfy calibration constraints (i.e. they must fit the calibration dataset to a level that is commensurate with the amount of measurement noise which accompanies that dataset), the process of sampling the posterior parameter probability distribution is much more difficult than that of obtaining a single parameter field that is deemed to "calibrate" a model. The latter process seeks a unique solution to an ill-posed inverse problem by pursuing a parameter field of minimized error variance. Generally, this is a parameter field that suppresses any heterogeneity that is not supported by the calibration dataset. Model predictions that are made using this parameter field can then also be considered as of minimum error variance. As such, they lie somewhere near the centre of the posterior probability distribution of that prediction.

In contrast to the process required for obtaining a minimum error variance parameter field, nonlinear parameter/predictive uncertainty analysis requires that the posterior parameter probability distribution be sampled in order to seek parameter fields which express heterogeneity rather than suppress it, while still promulgating a good fit with the calibration dataset. However any heterogeneity that is thereby expressed must be realistic from an expert knowledge point of view.

Notionally, the posterior parameter probability distribution can be derived through application of Bayes equation. Let the vector **k** denote parameters employed by a model. Let the operator $Z()$ represent the action of the model on its parameters. Let elements of the vector **h** represent members of the calibration dataset, and let the vector **ε** represent noise associated with measurements that comprise this dataset. Then

$$\mathbf{h} = Z(\mathbf{k}) + \boldsymbol{\varepsilon} \tag{2.1}$$

We use the symbol $P()$ to represent probability. Let $P(\mathbf{k})$ represent the prior probability (i.e. pre-calibration probability) associated with a parameter vector **k**. Bayes equation is used to characterize the posterior (i.e. post-calibration) probability associated with that same parameter vector. This is denoted as $P(\mathbf{k}|\mathbf{h})$, i.e. the probability of **k** conditioned by the calibration dataset **h**. Bayes equation states that:

$$P(\mathbf{k}|\mathbf{h}) \propto P(\mathbf{h}|\mathbf{k})\, P(\mathbf{k}) \tag{2.2}$$

The term $P(\mathbf{h}|\mathbf{k})$ is referred to as the likelihood function. It rises to the extent that a parameter field promulgates a good fit with the calibration dataset.

Direct use of Bayes equation to characterize the posterior parameter probability distribution, or even just to sample it, is extremely difficult (if not impossible) where parameter numbers are high. Hence approximate methods must be sought. The PEST suite provides a number of means to do this. All of them start with a linear approximation to Bayes equation; they then modify parameter fields that are sampled from the thus-approximated posterior parameter probability distribution so that they fit the calibration dataset to an acceptable degree. It is important to note, however, that despite their quasi

linear origin, use of these parameter fields in making a model prediction involves no linearity assumption.

## 2.2 Approximate Implementation of Bayes Equation

Suppose that the following conditions are met:

- The prior probability distribution of parameters is multi Gaussian;
- The probability density function of measurement noise is multi Gaussian;
- The action of the model on its parameters is linear, and hence can be expressed as a matrix.

Then the posterior expected values of parameters (i.e. the mean values of parameters according to their posterior probability distribution) can be obtained through the model calibration process, provided that this process employs regularisation that is in accordance with the prior parameter probability distribution. The covariance matrix that characterizes the posterior parameter probability distribution (we will refer to this as C'(**k**)) can be calculated using either of the following (mathematically equivalent) formulas:

$$C'(\mathbf{k}) = C(\mathbf{k}) - C(\mathbf{k})\mathbf{Z}^t[\mathbf{Z}C(\mathbf{k})\mathbf{Z}^t + C(\boldsymbol{\varepsilon})]^{-1}\mathbf{Z}C(\mathbf{k}) \qquad (2.3a)$$

$$C'(\mathbf{k}) = [\mathbf{Z}^tC^{-1}(\boldsymbol{\varepsilon})\mathbf{Z} + C^{-1}(\mathbf{k})]^{-1} \qquad (2.3b)$$

In these equations C(**k**) is the covariance matrix associated with the prior parameter probability distribution while C(**ε**) is the covariance matrix associated with measurement noise.

C'(**k**) of equations 2.3a and 2.3b can be calculated using the PREDUNC7 program supplied with the PEST suite. This employs the weighted Jacobian matrix calculated using the calibration parameter set in place of **Z**. Samples of the approximate posterior parameter probability distribution can then be generated using the RANDPAR1 utility; these samples are centred on the calibrated parameter field.

Because of model nonlinearity, it is unlikely that random parameter fields which are obtained in this way will promulgate an acceptable fit between model outputs and the calibration dataset. Predictive uncertainties that are calculated using these parameter fields may therefore be too broad. Hence these fields must be adjusted to better respect calibration constraints. Two adjustment options are as follows.

1. Adjust these parameter fields using PEST; the first iteration of the adjustment process can be implemented with very little numerical cost through use of the same Jacobian matrix that was employed by PREDUNC7 to calculate C'(**k**). Thus it does not need to be re-calculated for each random parameter field.
2. Subject these parameter fields to null space projection (see below) to remove those aspects of them that compromise goodness of fit. Then, if model-to-measurement fit is still not good enough, undertake PEST-based parameter adjustment as above. Null space projection can be implemented using the PNULPAR utility. The first iteration of parameter adjustment can employ a pre-calculated Jacobian matrix.

If only one iteration of parameter adjustment is required, both of these options are numerically efficient. However a second iteration of parameter adjustment incurs a high numerical cost, for it requires calculation of a new Jacobian matrix. This requires that one model run be undertaken for

each adjustable parameter (8949 for the NFSEG model). The numerical burden can be somewhat reduced if only "super parameters" are adjusted. These are coefficients applied to the right singular vectors of the weighted Jacobian matrix; only enough of these need be adjusted to span the dimensionality of the calibration solution space – about 1500 for the NFSEG model.

A second iteration of parameter adjustment is unacceptable for the NFSEG model as the numerical burden is too high to countenance. In general, the first of the above two options is preferred. However, the second must be chosen if it eliminates the need for a second iteration of parameter adjustment. The second of these options was indeed adopted for sampling the posterior parameter probability distribution of the NFSEG model. The null space projection methodology that it entails is now described.

Let **J** represent the Jacobian matrix and let **Q** represent the weight matrix used in model calibration (normally diagonal). If subjected to singular value decomposition, the weighted Jacobian matrix $\mathbf{Q}^{\frac{1}{2}}\mathbf{J}$ can be decomposed as:

$$\mathbf{Q}^{\frac{1}{2}}\mathbf{J} = \mathbf{USV}^{t} \tag{2.4}$$

where **U** is an orthonormal matrix whose columns span the range space of $\mathbf{Q}^{\frac{1}{2}}\mathbf{J}$, **V** is an orthonormal matrix whose columns span parameter space, and **S** is a diagonal matrix of singular values. Using the SUPCALC utility provided with the PEST suite, **V** can be partitioned as:

$$\mathbf{V} = [\mathbf{V}_1\ \mathbf{V}_2] \tag{2.5}$$

where the columns of $\mathbf{V}_1$ span the calibration solution space and those of $\mathbf{V}_2$ span the calibration null space (i.e. the space wherein combinations of parameters have minimal effects on model outputs and hence are non-inferable through the calibration process). See Doherty (2015) for further details.

The PNULPAR utility modifies a random parameter vector **k** in the following manner:

1. First it subtracts the parameter vector **k** obtained through calibration of the model to obtain **k** - **k**;
2. It then projects this difference onto the calibration null space through pre-multiplication by $\mathbf{V}_2\mathbf{V}_2^{t}$.
3. It adds **k** to this projected difference.

If the model is run using a PNULPAR-calculated set of parameters, minimal parameter adjustment is normally required for fitting the calibration dataset to an arbitrary level of acceptability. A disadvantage of this method of achieving a good fit with the calibration dataset, however, is that the component of parameter uncertainty that is inherited from measurement noise can be diminished through this process. This is of little consequence if predictions of interest have a high degree of null space dependency. The uncertainties of other predictions may be somewhat under-valued, however. This is rectified to some extent through post-projection parameter adjustment (as was done for the NFSEG model), for it is the solution space component of random parameter fields that must adapt to the requirement for a suitably good level of fit with the calibration dataset; the uncertainty associated with this component is directly inherited from measurement noise associated with that data. Another strategy is to add random noise to members of the calibration dataset prior to adjustment of each parameter field. However, this is problematical in many modelling contexts (including the present one)

as the assumption of measurement-to-measurement statistical independence is not in accordance with the predominately structural origin of model-to-measurement misfit that is typically encountered when calibrating an environmental model. Nor is it possible to provide a suitable stochastic description for this inherently heteroscedastic phenomenon.

## 2.3 Parameter Estimability

Application of the linear theory presented above enables relatively easy calculation of two statistics which provide a measure of the information content of the calibration dataset with respect to parameters employed by the model. Each of these statistics has a value of between 0.0 and 1.0, with the former indicating zero information content, and the latter indicative of sufficient information content to afford unique estimation of its value. Both of these statistics are discussed by Doherty and Hunt (2009).

The first statistic discussed herein is the relative parameter uncertainty variance reduction. The variances of posterior parameter uncertainty comprise the diagonal elements of the $C'(\mathbf{k})$ matrix that is calculated using equation 2.3a or 2.3b. Prior parameter uncertainty variances are available as the corresponding elements of $C(\mathbf{k})$. Let $\sigma_i^{2\prime}$ denote the posterior variance of the *i*'th parameter; let $\sigma_i^2$ denote its prior variance. The relative uncertainty variance reduction of parameter *i* (which we denote as $R_i$) is calculated using the following formula.

$$R_i = \frac{\sigma_i^2 - \sigma_i^{2\prime}}{\sigma_i^2} \qquad\qquad (2.6)$$

The second statistic is the so-called "identifiability" of a parameter. It is defined as the square of the cosine between a vector pointing in the direction of the parameter and the projection of this vector onto the calibration solution space. This is the space defined by the vectors comprising the columns of the $\mathbf{V}_1$ matrix of equation 2.5. This measure of parameter estimability is a little more arbitrary than relative parameter uncertainty variance reduction because it is sensitive to the estimated dimensionality of the solution space. In calibration of an environmental model, the boundary between the solution and null spaces is not sharp. Theoretically, this boundary marks the point at which the potential for error associated with calibration-based estimation of the value of a parameter combination defined by the pertinent column of $\mathbf{V}_1$ is greater than the potential for error based on expert knowledge alone. This is the point at which "over-fitting" begins. Moore and Doherty (2005) show that attempts to estimate values associated with an increasing number of columns of $\mathbf{V}_1$ are accompanied by a growing amplification of measurement noise as singular values associated with these columns diminish in magnitude. The fact that prior uncertainties are themselves uncertain, and that "measurement noise" is dominated by structural noise of unknown statistical properties, makes separation of solution and null spaces based on this premise a somewhat vague undertaking. This vagueness is inherited by the values assigned to the identifiabilities of those parameters that are only mildly informed by the calibration dataset.

# 3. Calculating Calibration-Constrained Parameter Fields

## 3.1 Observations and Parameters

Details of parameters employed by the NFSEG model, and of observations comprising the NFSEG model calibration dataset, are described in chapter 4 of the companion report. In undertaking calibration-constrained random parameter field generation for the NFSEG model, parameters remained unchanged from those employed in the calibration process, with the following exceptions.

- Recharge multiplier parameters which were fixed at a value of 1.0 during the calibration process were decreed as adjustable.
- EVT rate multiplier parameters which were fixed at a value of 1.0 during the calibration process were decreed as adjustable.

Table 3.1 lists parameter groups employed in the random parameter field generation process. With the exception of 8 parameters belonging to the *sc* group (7 of which were fixed and one of which was tied to another parameter), all parameters within all of these groups were log-transformed for the purpose of assigning prior uncertainties, and for the purpose of generating samples from the posterior parameter probability distribution.

**Table 3.1 Parameter groups used in random parameter field generation. A total of 8957 parameters collectively comprise these groups, 8949 of which are adjustable.**

| Parameter group name | Parameterization device | Number of adjustable parameters | Description |
|---|---|---|---|
| k1x | pilot points | 518 | horizontal hydraulic conductivity – layer 1 |
| k3x | pilot points | 1767 | horizontal hydraulic conductivity – layer 3 |
| k5xk3x | pilot points | 201 | horizontal hydraulic conductivity multiplier outside MCU – layer 5 |
| k5x | pilot points | 364 | horizontal hydraulic conductivity – layer 5 |
| k7x | pilot points | 55 | horizontal hydraulic conductivity – layer 7 |
| k2z | pilot points | 556 | vertical hydraulic conductivity – layer 2 |
| k2zk3z | pilot points | 333 | vertical hydraulic conductivity multiplier outside ICU – layer 2 |
| k4zk3z | pilot points | 230 | vertical hydraulic conductivity multiplier outside MCU – layer 4 |
| k4z | pilot points | 139 | vertical hydraulic conductivity – layer 4 |
| k6z | pilot points | 68 | vertical hydraulic conductivity – layer 6 |
| vanis1 | entire layer | 1 | vertical anisotropy – layer 1 |
| vanis2 | zoned according to ICU/non-ICU | 2 | vertical anisotropy – layer 2 |
| vanis3 | pilot points | 154 | vertical anisotropy – layer 3 |
| vanis4 | zoned according to MCU/non-MCU | 2 | vertical anisotropy – layer 4 |
| vanis5 | zoned according to MCU/non-MCU | 2 | vertical anisotropy – layer 5 |
| vanis6 | entire layer | 1 | vertical anisotropy – layer 6 |
| vanis7 | entire layer | 1 | vertical anisotropy – layer 7 |
| lcm | zoned according to lakes | 258 | multiplier applied to lakebed conductance |
| rcm | zoned according to river reaches | 1872 | multiplier applied to river reach conductance |
| sc | zoned according to springs | 371 | GHB conductance at springs |
| rechmul | zones (see fig 3.1) | 904 | multiplier applied to recharge rates |

| evtrmul | zones (see fig 3.1) | 904 | multiplier applied to maximum EVT rates |
| lkzmul | zoned according to lakes | 246 | vertical conductivity multiplier under lakes |

Table 3.2 lists observation groups comprising the calibration dataset.

**Table 3.2 Observation groups used in constraining random parameter fields. A total of 5713 non-zero-weighted observations collectively comprise these groups.**

| Observation group name | Number of observations with non-zero weight | Description |
| --- | --- | --- |
| h2001_lay1 | 228 | Heads in layer 1: 2001 |
| h2001_lay2 | 96 | Heads in layer 2: 2001 |
| h2001_lay3 | 977 | Heads in layer 3: 2001 |
| h2001_lay4 | 13 | Heads in layer 4: 2001 |
| h2001_lay5 | 39 | Heads in layer 5: 2001 |
| h2001_lay7 | 2 | Heads in layer 7: 2001 |
| h2009_lay1 | 581 | Heads in layer 1: 2009 |
| h2009_lay2 | 111 | Heads in layer 2: 2009 |
| h2009_lay3 | 993 | Heads in layer 3: 2009 |
| h2009_lay4 | 10 | Heads in layer 4: 2009 |
| h2009_lay5 | 41 | Heads in layer 5: 2009 |
| h2009_lay7 | 2 | Heads in layer 7: 2009 |
| hd2001_lay3 | 288 | Lateral head gradients in layer 3: 2001 |
| hd2009_lay3 | 274 | Lateral head gradients in layer 3: 2009 |
| td_lay1 | 0 | Temporal head differences: layer 1 |
| td_lay2 | 0 | Temporal head differences: layer 2 |
| td_lay3 | 0 | Temporal head differences: layer 3 |
| td_lay4 | 0 | Temporal head differences: layer 4 |
| td_lay5 | 0 | Temporal head differences: layer 5 |
| td_lay7 | 0 | Temporal head differences: layer 7 |
| wp_dry_2001 | 6 | Minimizes occurrence of dry cells in wetland areas in 2001 |
| wp_wet_2001 | 6 | Minimizes occurrence of flooded cells |
| wp_dry_2009 | 6 | Minimizes occurrence of dry cells in wetland areas in 2009 |
| wp_wet_2009 | 6 | Minimizes occurrence of flooded cells |
| vd_1to3_01 | 114 | Vertical head differences: layer 1 to 3 in 2001 |
| vd_1to3_09 | 141 | Vertical head differences: layer 1 to 3 in 2009 |
| vd_3to5_01 | 17 | Vertical head differences: layer 3 to 5 in 2001 |
| vd_3to5_09 | 19 | Vertical head differences: layer 3 to 5 in 2009 |
| qr01 | 76 | Inflow to river segments between one or more gages: 2001 |
| qr09 | 44 | Inflow to river segments between one or more gages: 2009 |
| qspring01 | 365 | Inflow to springs: 2001 |
| qspring09 | 368 | Inflow to springs: 2009 |
| qs_spring01 | 6 | Inflow to spring groups: 2001 |
| qs_spring09 | 7 | Inflow to spring groups: 2009 |
| qs01 | 11 | Cumulative inflow to river upstream of a gage: 2001 |
| qs09 | 10 | Cumulative inflow to river upstream of a gage: 2009 |
| qlake01 | 255 | Flow to/from lakes: 2001 |
| qlake09 | 258 | Flow to/from lakes: 2009 |

## 3.2 Some Considerations Pertinent to the NFSEG Model

The domain of the NFSEG model is large. The calibration dataset for the NFSEG model is large. It includes both heads and outflows (at rivers and springs). The model is judged by stakeholders on how well all of these elements of the calibration dataset are reproduced by the model. Considerable attention has been given to ensuring that the model is able to match observed heads and outflows

during the 2001 and 2009 calibration periods. In the future, the model will be used to support management which aims to maintain future heads and outflows at desired levels.

This manner in which the NFSEG model was calibrated, and the manner in which it will be deployed has a number of repercussions, some of which are now briefly discussed.

Because of the requirement that the parameter field which is deemed to calibrate the NFSEG model yields a good fit with the calibration dataset, particularly at locations that are pertinent to decision-relevant predictions of future system behaviour, it is equally important that calibration-constrained random parameter fields that are used to explore predictive uncertainty do the same. If not, their characterization of the uncertainties of decision-critical model predictions will be deemed as being too conservative to be useful.

Ideally the fit that is achieved between model outputs and field measurements of head and flow should be commensurate with the noise associated with measurement of these quantities. However, model-to-measurement misfit often exceeds measurement error because of its structural origins, born of model imperfections. For a regional model with a large cell size, calibrated under an assumption of steady state conditions, the magnitude of this structural term is occasionally significant. Furthermore, it differs from location to location in unknown ways. It cannot therefore be characterized as a sequence of random numbers whose magnitude at a given location is a matter of chance, both under calibration and predictive conditions.

Naturally, a prediction of future system behaviour made by the NFSEG model will also be affected by structural error. Ideally, predictive uncertainty that arises from parametric uncertainty (as calculated using calibration constrained random parameter fields) should be supplemented by realisations of structural noise that are added to model predictions. Unfortunately, this is not possible for a number of reasons. These include:

- The location-specific nature of structural error;
- The fact that, for predictions which resemble members of the calibration dataset, some model structural defects will have been "calibrated out" (see Welter and Doherty, 2010; White et al, 2014);
- The likelihood that structural noise as it pertains to a prediction will be different from that which applies during calibration.

The latter is important. While lack of exact adherence to the steady state assumption may incur structural noise under calibration conditions, it will not do so under predictive conditions. Similarly, future predictions of system behaviour (by definition) can be made under the assumption that all system stresses are known, or can be represented with a user-specified stochastic distribution whose properties are known; the same does not apply under historical conditions. Furthermore, in fulfilling its decision-support role, the NFSEG model will often be used to compute differences in heads or flows arising from differences in stresses rather than the actual values of these quantities. It can be assumed that considerable cancellation of structural noise accompanies the making of such differential predictions. (This is further discussed below.)

Because of the presence of structural noise (a feature of all groundwater models), predictive uncertainties that are assessed through use of calibration-constrained stochastic parameter fields

whose calculation is described herein should be viewed as representing lower bounds on the uncertainties of predictions of management interest rather than their true uncertainties. Little can be done about this without recourse to a model from which structural deficiencies are absent.

## 3.3 Prior Parameter Covariance Matrix

As for previous linear analyses conducted on the NFSEG model, the C(**k**) matrix featured in equations 2.3a and 2.3b is considered to be block-diagonal. Submatrices pertaining to many of these blocks are in fact diagonal, this denoting statistical independence of parameters represented by these blocks.

For most parameter types, prior parameter variability assumed for random parameter field generation is somewhat smaller than that employed in previous investigations into parameter and predictive uncertainty using linear analysis. Prior uncertainties employed in previous analyses were used in the initial stages of the present nonlinear investigation (and in linear analysis discussed in the next section of this document). However, difficulties were encountered in attaining a good fit between model outputs and field measurements when deploying the methodologies discussed in chapter 2 of this appendix; a second, numerically expensive, iteration of parameter adjustment was found to be necessary to attain the level of model-to-measurements required for the NFSEG model. This problem was overcome by reducing the prior uncertainties of some parameters.

Parameter groups comprising pilot point parameters were assigned a full covariance matrix based on spatially variable variograms. All variograms are exponential, and hence specified by the equation:

$$\Upsilon(h)=C(0)[1 - \exp (h/a)] \tag{3.1}$$

In equation 3.1 $h$ is distance and C(0) is the overall variance of the hydraulic property in question, this being equal to the sill of the variogram. The range of an exponential variogram is often characterized as 3$a$.

For all pilot point parameters, the variogram "$a$" value was decreed to be pilot-point-specific. This reflects the fact that pilot points are not capable of representing the natural heterogeneity of a complex carbonate aquifer. Instead, their use implies a degree of upscaling, with the extent of upscaling decreasing with increasing spatial density of pilot points. Where spatial density of pilot point emplacement is high, they are capable of representing short range hydraulic property heterogeneity. Alternatively, where it is low, they can only represent long range hydraulic property heterogeneity. To reflect this, the variogram range associated with each pilot point was calculated in the following manner:

1. The average separation between the pilot point to which an "$a$" value must be assigned and its 20 closest neighbours was calculated;
2. The "$a$" value ascribed to the variogram associated with that pilot point was designated as twice this average separation.

This strategy is similar to that applied in previous linear uncertainty analysis.

For all pilot point parameter groups, the variogram sill (applied to the logarithm of hydraulic properties associated with pilot points) was denoted as uniformly 0.16, this implying a standard deviation of parameter variability of 0.4. Variograms assigned to all pilot points are isotropic, except for a number

of pilot points in the k3x and k3xk5x parameter groups to which a horizontal anisotropy of 2.0 with a bearing of 354 degrees was ascribed.

Covariance matrix construction was undertaken using the MKPPSTAT and PPCOV_SVA utilities supplied with the PEST Groundwater Data Utility suite; see Doherty (2014).

For all other parameter types except recharge and EVT multipliers, within-group statistical independence was assumed in filling the pertinent blocks of the C(**k**) matrix. A uniform standard deviation was applied to each member of the group. Details are provided in table 3.3.

**Table 3.3 Standard deviation assigned to each parameter within each respective parameter group. Note that these are actually applied to the log (to base 10) of each parameter. Note also that none of the parameters in this table pertain to pilot points.**

| Parameter group name | Standard deviation ascribed to each parameter within group |
|---|---|
| vanis1 | 0.2 |
| vanis2 | 0.2 |
| vanis4 | 0.2 |
| vanis5 | 0.2 |
| vanis6 | 0.2 |
| vanis7 | 0.2 |
| lcm | 0.5 |
| rcm | 0.4 |
| sc | 0.5 |
| lkzmul | 0.4 |

As has already been discussed, a recharge rate multiplier parameter and an EVT rate multiplier parameter were assigned to each of the 904 watershed-based polygons within the model domain that define zonation of these quantities. These parameters were assumed to be statistically independent of each other between polygons. However, a high degree of negative correlation (-0.94) is assumed to exist between the recharge rate multiplier and the maximum EVT rate multiplier within each polygon. This reflects the fact that they are both calculated by the same HSPF model which is calibrated to reproduce baseflow. Use of this correlation coefficient is equivalent to assuming that the recharge rate multiplier is equal to a random number plus a second random number whose standard deviation is about 25% of that of the first, while the EVT rate multiplier is equal to the same first random number minus a third random number which also has a standard deviation of 25% of that of this first random number. For the present study, the standard deviation of the first random number was chosen such that four standard deviations span a range in log space that is equivalent to multiplying and dividing the "calibrated" value of 1.0 for these recharge and EVT multiplier parameters by a factor of 1.25. A joint covariance matrix for recharge and EVT rate multipliers was constructed accordingly.

### 3.4 Measurement Noise Covariance Matrix

The C($\boldsymbol{\varepsilon}$) matrix which characterizes measurement noise (see equations 2.3a and 2.3b) is calculated slightly differently in nonlinear uncertainty analysis from the way that it is calculated in order to support linear uncertainty analysis. For calculation of C'(**k**), C($\boldsymbol{\varepsilon}$) was assumed to be diagonal, with elements proportional to the inverse of the squared reciprocals of weights used during the calibration process. The proportionality constant applied to all weights was such that the measurement objective

function is approximately equal to the number of observations comprising the calibration dataset minus the dimensionality of the calibration solution space; see Doherty (2015) for details.

Implied in the level of fit that calibration-constrained stochastic parameter fields are required to attain with the calibration dataset is a stochastic characterization of "noise" that is responsible for model-to-measurement misfit. As has already been discussed, much of this "noise" is structural. As such, it is not amenable to statistical characterisation. If it were, indeed, stochastically-characterizable noise, the level of misfit used to constrain random parameter fields could be calculated using standard statistical theory; see, for example, Vecchia and Cooley (1987). For the NFSEG model, however, an objective function 5% greater than that attained by the calibrated parameter field was set as the constraint. The necessity for all random parameter fields used in conjunction with the NFSEG model to achieve a good fit with the model's calibration dataset has already been discussed.

## 3.5 Some Implementation Details

Sampling the posterior probability distribution of NFSEG model parameter fields required that a suite of parameter fields be obtained for which corresponding objective functions are less than 1.05 times that attained by the calibrated model. It was required that the number of these parameter fields be sufficient to provide a reasonable characterization of posterior predictive uncertainties. This, in turn, required that the procedure be numerically efficient. Attainment of a workable level of numerical efficiency required avoidance of a second iteration of random parameter field adjustment. Workflow settings used in this process (including variables used for characterization of prior parameter uncertainties) reflected this necessity. These were the outcome of some trial and error workflow implementations that are not discussed herein.

## 3.6 Sampling a Linear Approximation to the Posterior

The PEST PREDUNC7 utility was employed to calculate a linear approximation to the posterior covariance matrix C'($\mathbf{k}$) using equation 2.3b.

The RANDPAR1 utility was then employed to generate 1000 stochastic parameter field realizations based on the thus-calculated C'($\mathbf{k}$) matrix. Each of these parameter fields was then subjected to null space projection using the PNULPAR utility; the dimensionality of the solution space was assumed to be 1500.

Each of the PNULPAR-generated random parameter fields was then subjected to one iteration of parameter adjustment. As has already been discussed, this was a relatively cheap numerical undertaking as it employed the same Jacobian matrix for adjustment of all parameter fields, namely that which was calculated using the calibrated parameter field. For 522 of these parameter fields the final objective function was less than 1.05 times that associated with the calibration parameter field. These can be considered to comprise samples of the posterior parameter probability distribution. As such, they can be used for predictive uncertainty analysis.

# 4. Outcomes of Nonlinear Analysis

## 4.1 Results

As described previously, a set of 522 calibrated parameter fields were generated using the null-space Monte Carlo approach described in the previous section. An example of one of these fields is shown in figure 4.1. Collectively, these fields can be used to characterize the statistical properties of any model parameter. See, for example, the histograms pertaining to selected parameters depicted in figures 4.2 through 4.4, and the estimated cumulative parameter probability distribution provided in figure 4.5.
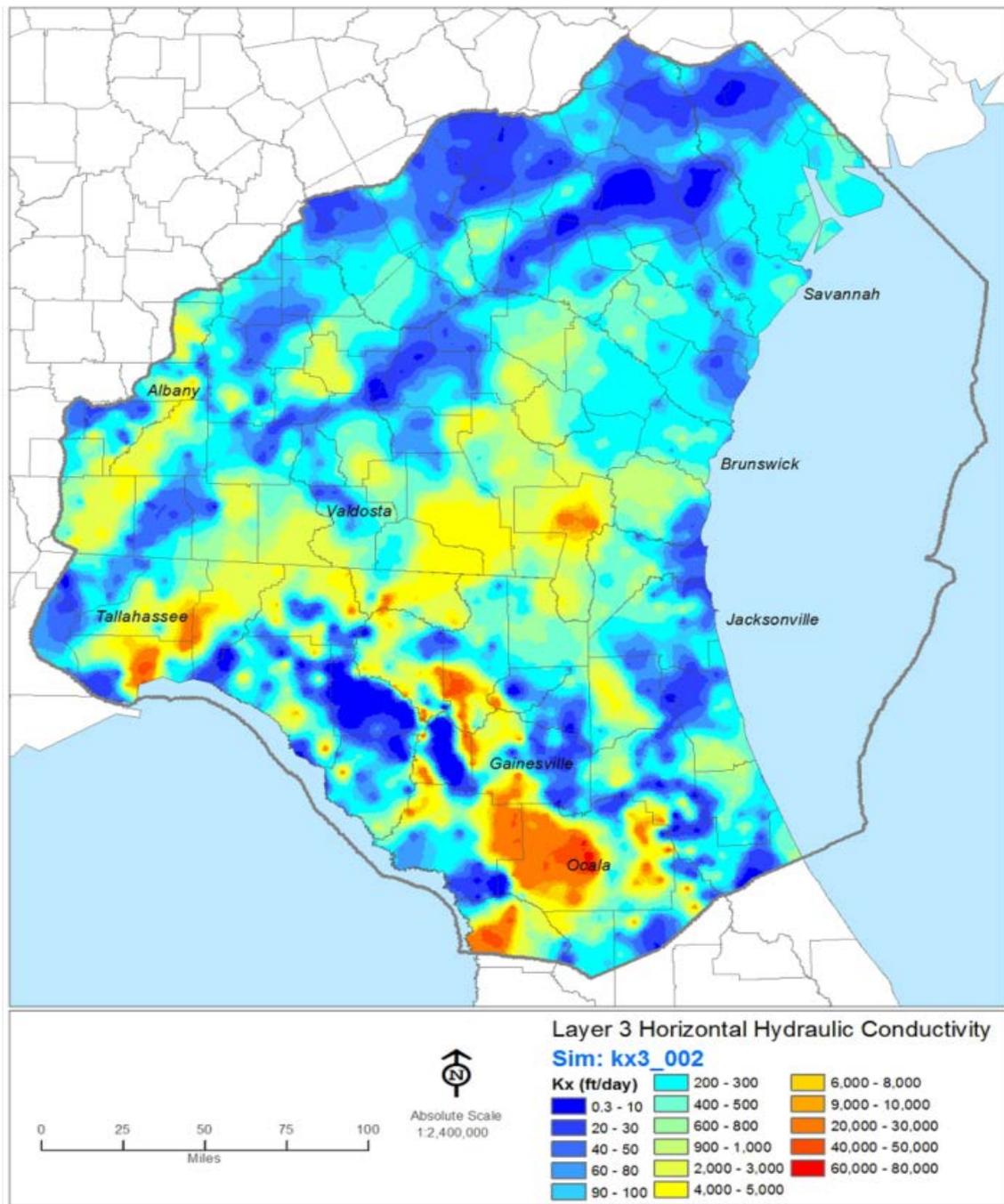
**Figure 4.1. A single post-calibration random parameter field realization of horizontal hydraulic conductivity in layer 3.**
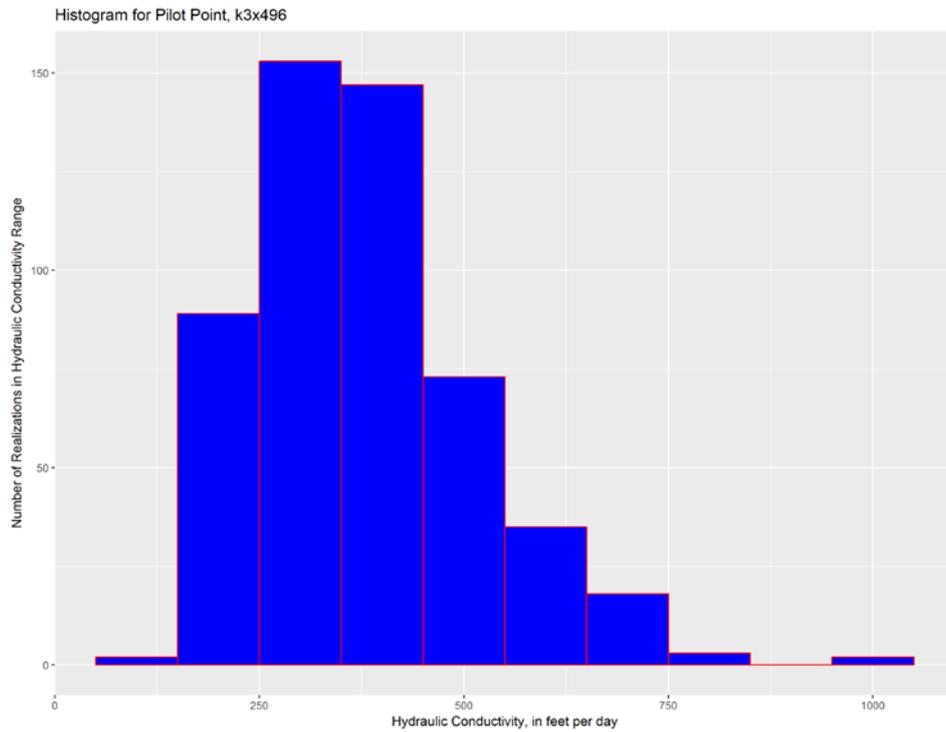
**Figure 4.2. Histogram for model parameter k3x496. This is a horizontal hydraulic conductivity pilot point parameter for layer 3. The histogram was developed from 522 calibration-constrained random parameter values.**
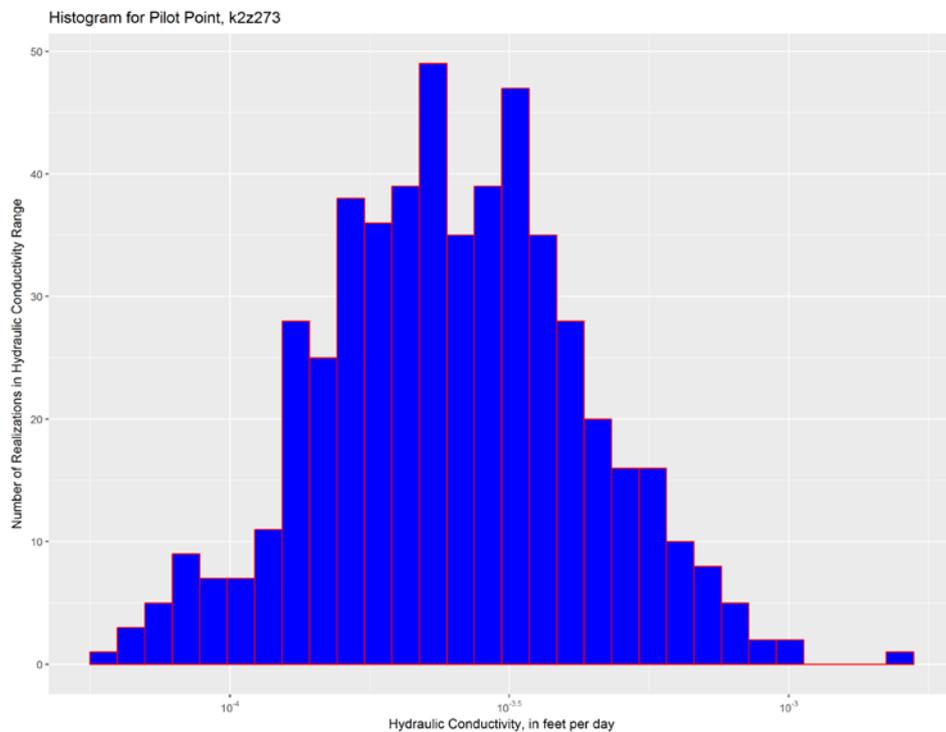


**Figure 4.3. Histogram for model parameter k2z273. This is a vertical hydraulic conductivity pilot point parameter for layer 2. The histogram was developed from 522 calibration-constrained random parameter values.**
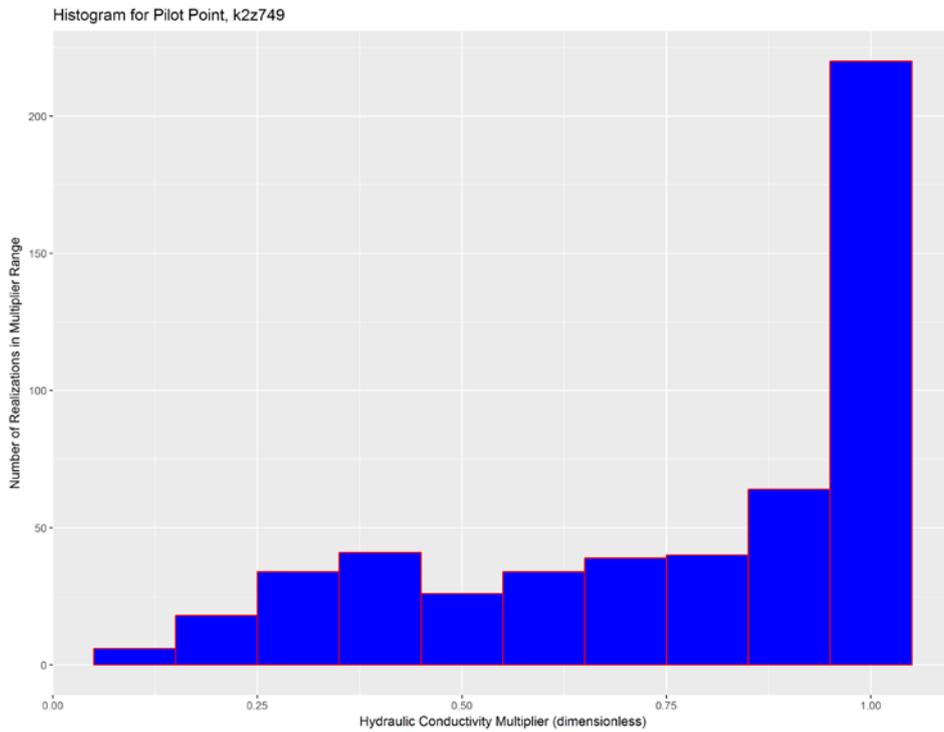
**Figure 4.4. Histogram for model parameter k2z749. This is a vertical hydraulic conductivity multiplier pilot point parameter for layer 2. The histogram was developed from 522 calibration-constrained random parameter values.**
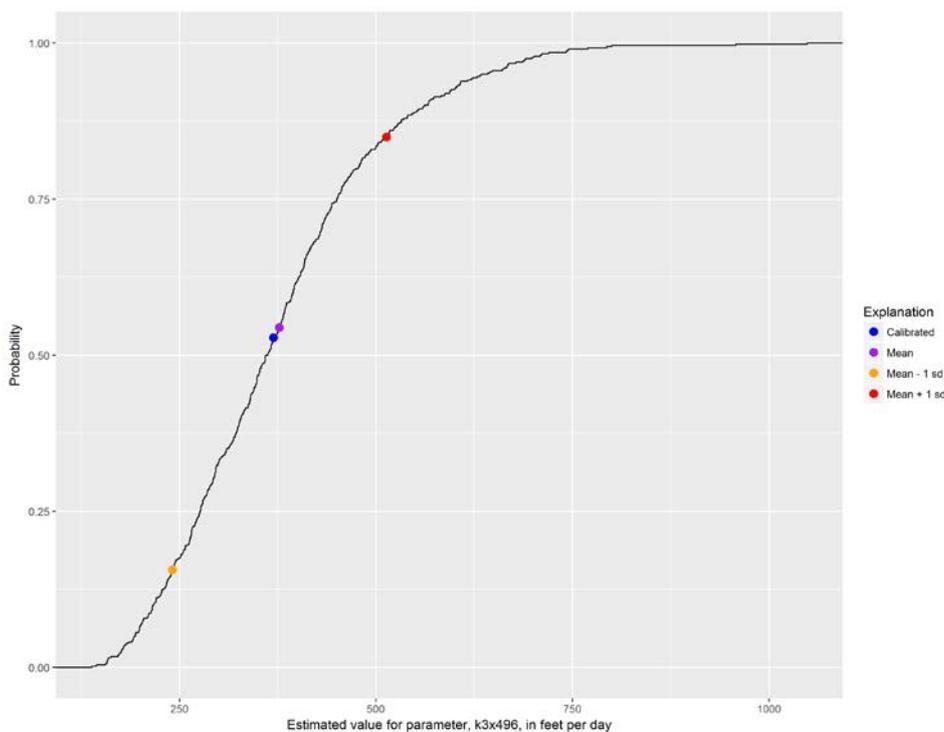


**Figure 4.5 Estimated cumulative probability distribution for model parameter k3496. This is a horizontal hydraulic conductivity pilot point parameter in layer 3. The probability distribution was developed from 522 calibration-constrained random parameter values. Also shown are the mean and calibrated values for this parameter, along with values corresponding to the mean plus and minus one standard deviation.**

Spatial variability of the statistical properties associated with a particular parameter type can be represented by mapping a given statistic to the location of its corresponding model parameter. Examples of these maps are provided in figures 4.6 and 4.7. Part a of these figures shows the geographical distribution of the standard deviation of log (to base 10) layer 3 hydraulic conductivities (figure 4.6), and layer 2 vertical hydraulic conductivities and vertical hydraulic conductivity multipliers (figure 4.7); all of these are pilot point parameters. Part b of these figures shows the spatial distribution of 10 raised to the power of log standard deviation. This is equivalent to the factor by which the mean parameter value at any point must be multiplied and divided to define a range which is roughly equal to its 67% confidence interval.
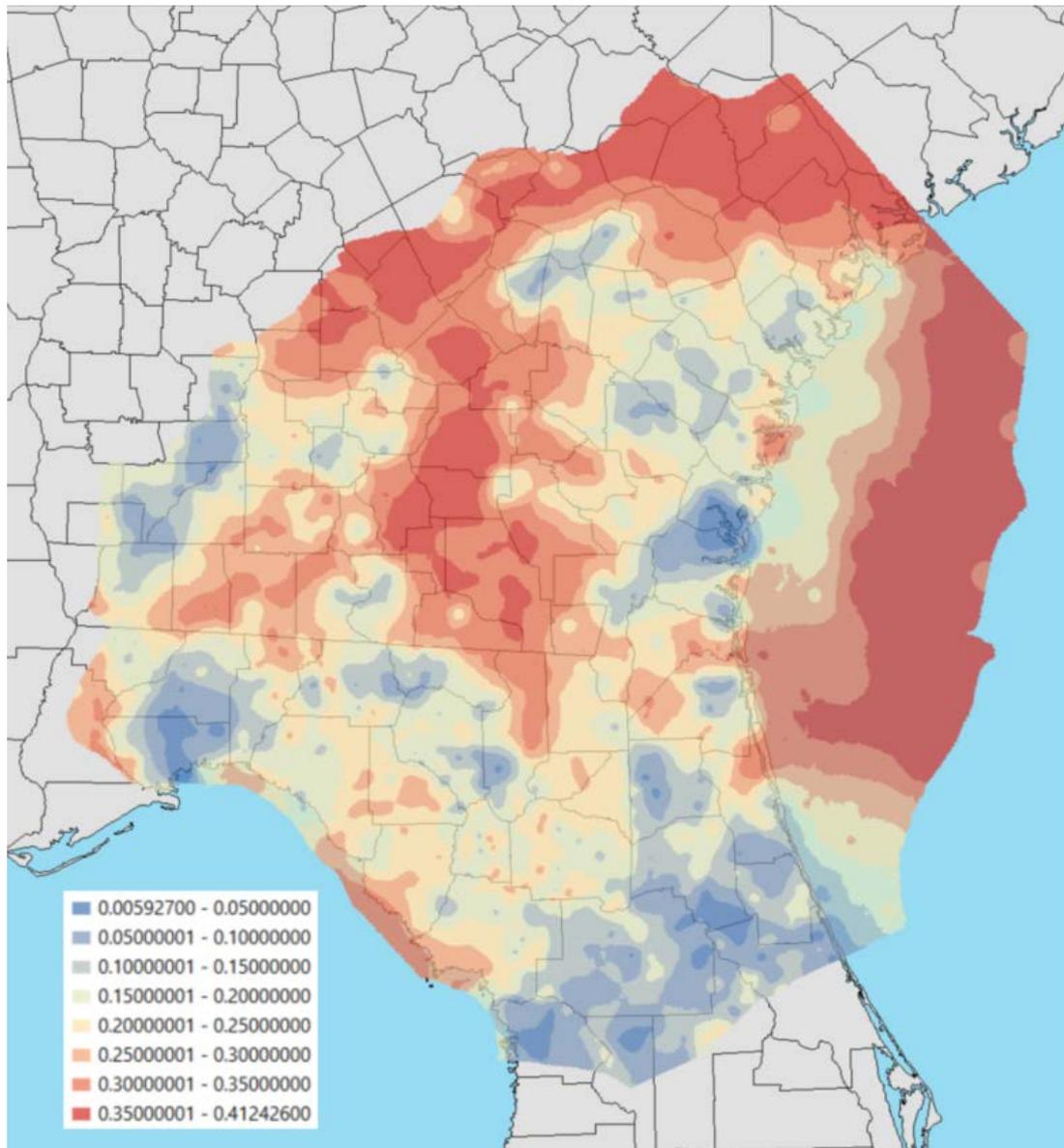


**Figure 4-6a. Standard deviation (in log space) of layer 3 horizontal hydraulic conductivity.**
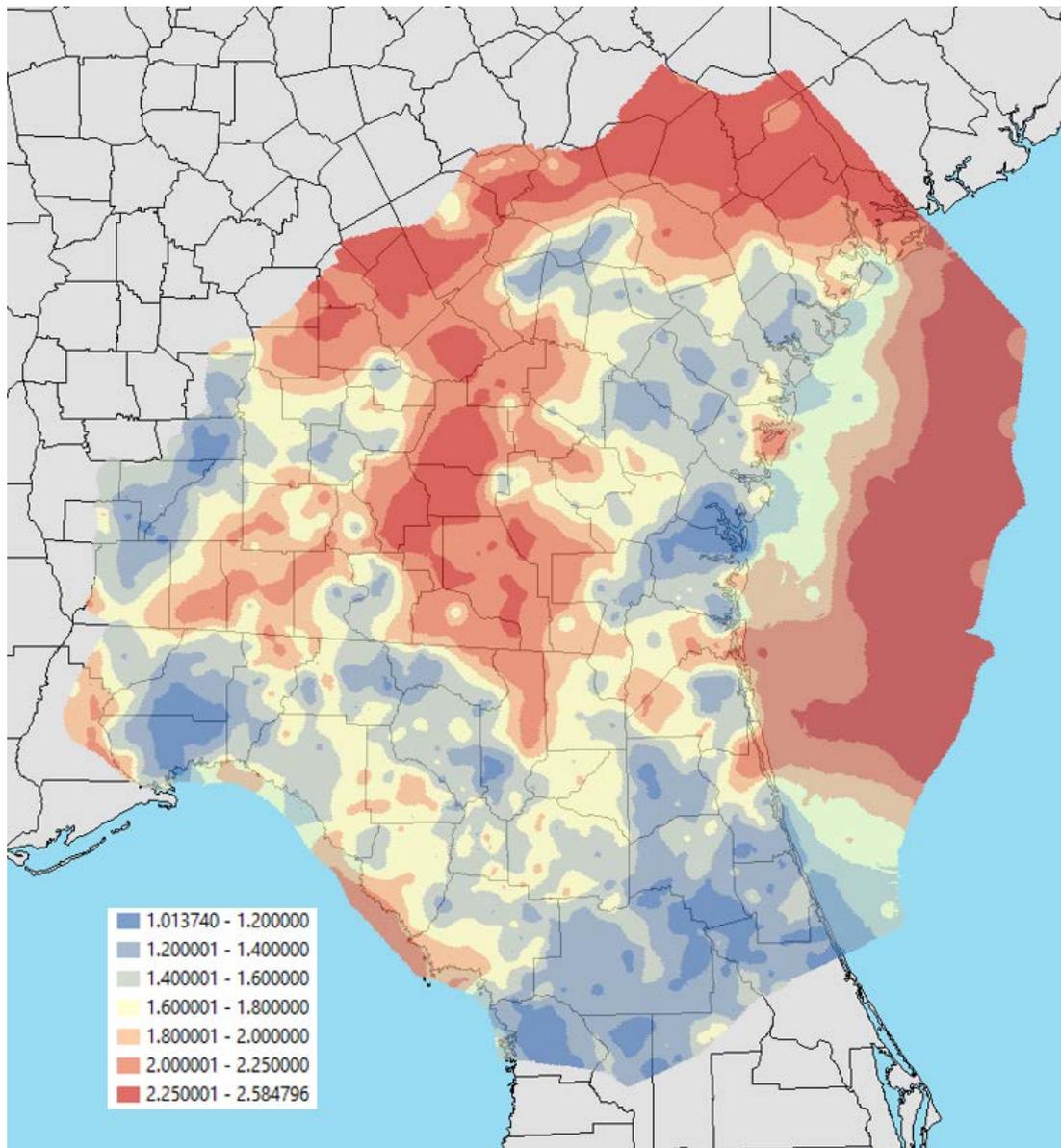
**Figure 4-6b. Factor by which the horizontal hydraulic conductivity in layer 3 must be multiplied and divided to span a range that is roughly equivalent to its 67% confidence interval.**
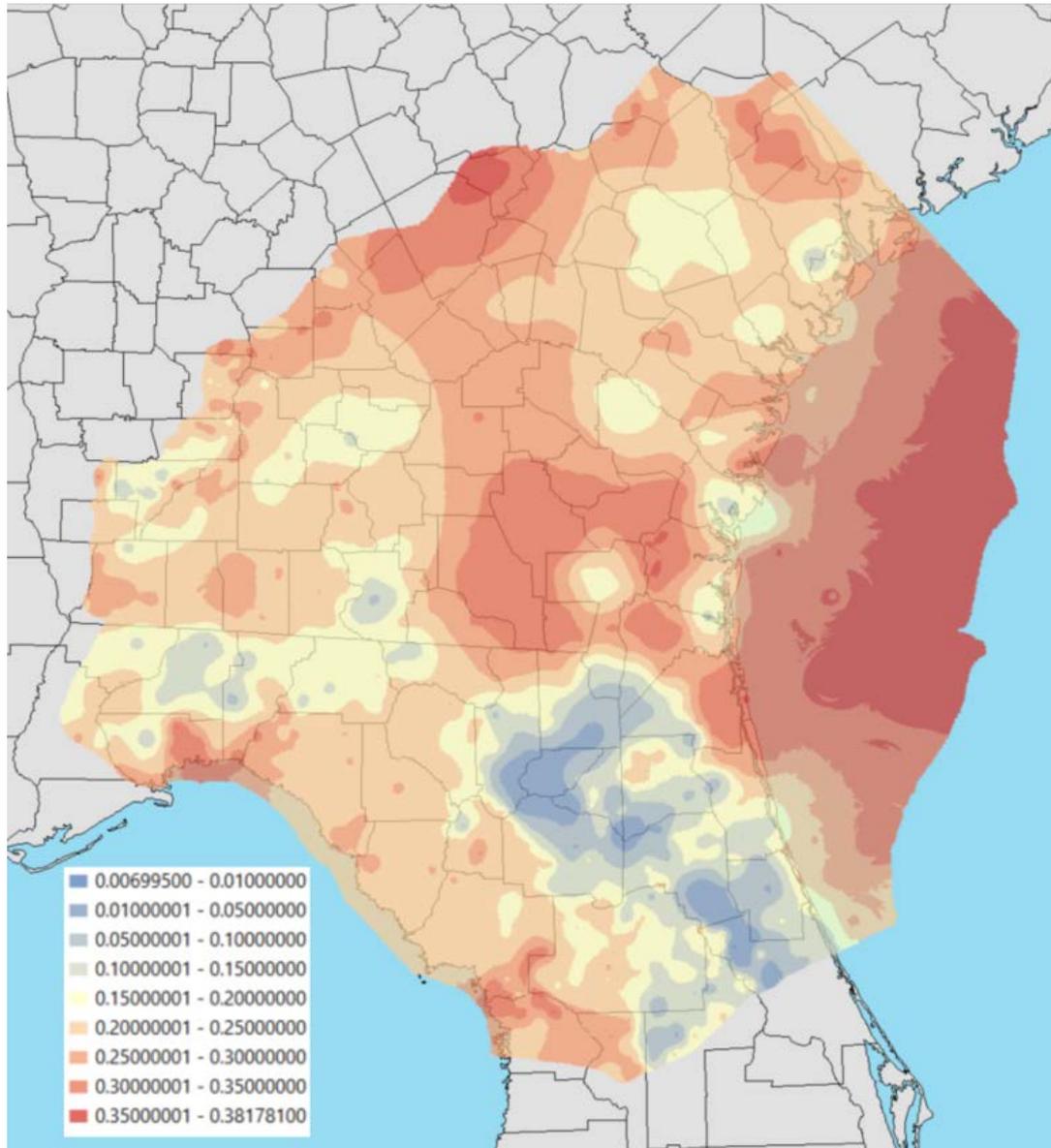
**Figure 4-7a. Standard deviation (in log space) of layer 2 vertical hydraulic conductivity (confined areas) and layer 2 vertical hydraulic conductivity multipliers (unconfined areas).**
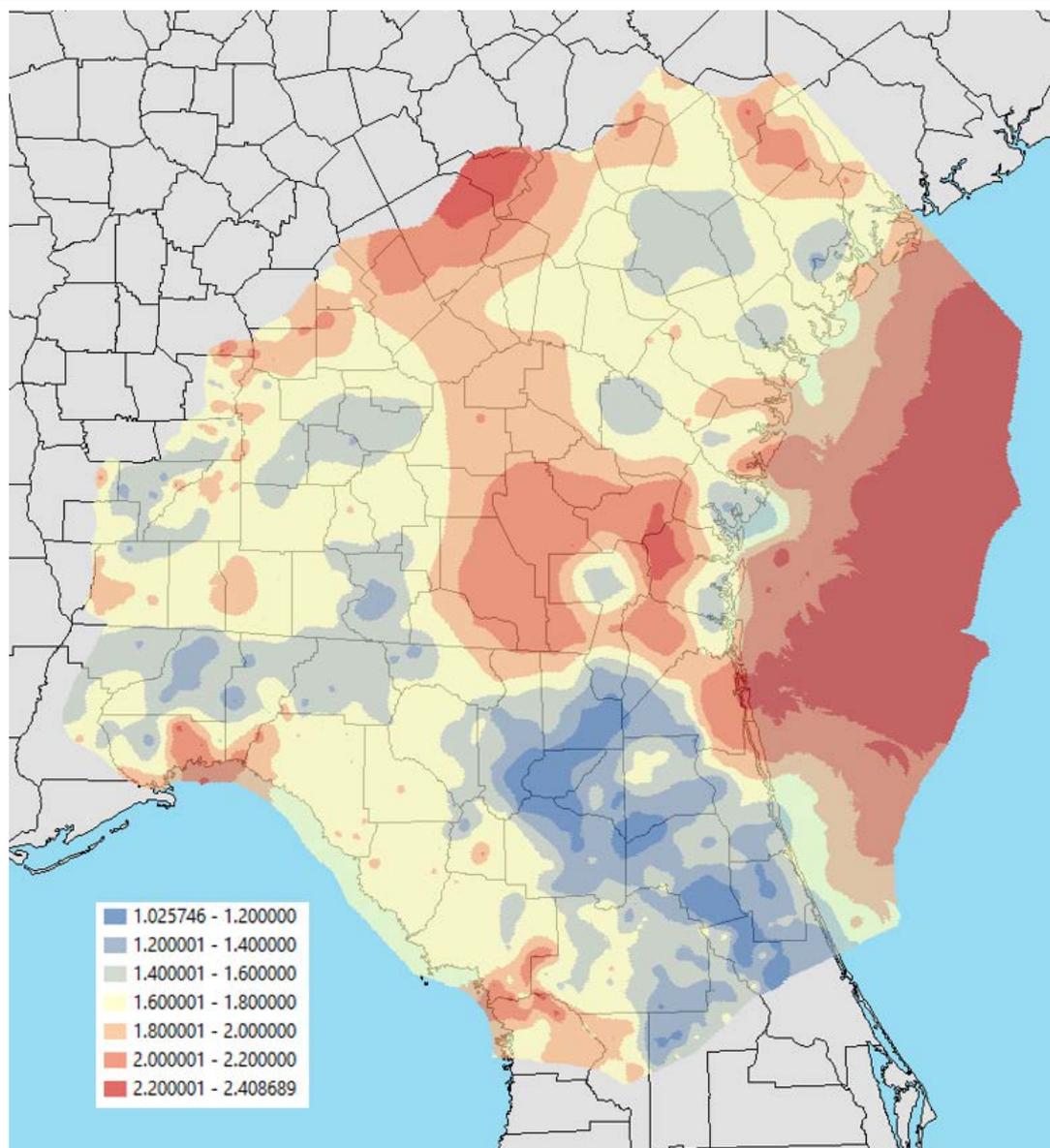
**Figure 4-7b. Factor by which vertical hydraulic conductivity (confined areas) and vertical hydraulic conductivity multipliers (unconfined areas) in layer 2 must be multiplied and divided to span a range that is roughly equivalent to its 67% confidence interval.**

Posterior probability distributions for selected predictions were estimated using the 522 calibration-constrained parameter fields. The locations of these predictions are provided in table 4.1 and figure 4.8. These predictions pertain to a hypothetical year 2035 pumping condition.

**Table 4.1 Names of locations at which predictive uncertainties are evaluated.**

| Prediction location name | Description |
|---|---|
| w00400 | UFA observation well near Lake Lochloosa |
| w00202 | UFA observation well near Lake Brooklyn |
| w00258 | UFA observation well near Lake Geneva |
| w00878 | UFA observation well near Putnam County MFL lakes |
| qs_2315500 | Baseflow to the Suwannee River near White Springs |
| qs_2317620 | Baseflow to the Alapaha River near Jennings |
| qs_2319000 | Baseflow to the Withlacoochee River near Pinetta |

| Prediction location name | Description |
|---|---|
| qr_2319394 | Baseflow pickup in the reach upstream of the Withlacoochee River near Lee |
| qr_2319500 | Baseflow pickup in reach upstream of the Suwannee River near Ellaville |
| qs_2319500 | Baseflow to the Suwannee River at Ellaville |
| qr_2320500a | Baseflow to the reach upstream of the Suwannee River near Branford |
| qs_2320500 | Baseflow to the Suwannee River near Branford |
| qr_2320700 | Baseflow to the Santa Fe River near Graham |
| qr_2321000 | Baseflow to the New River near Lake Butler |
| qs_2321500 | Baseflow to the Santa Fe River near Worthington Springs |
| qs_2322500 | Baseflow to the Santa Fe River near Fort White |
| qr_2322700 | Baseflow to the Ichetucknee River at US Highway 27 near Hildreth |
| qr_2323500 | Baseflow pickup in reach upstream of the Suwannee River near Wilcox |
| qs_2323500 | Baseflow to the Suwannee River near Wilcox |
| qr_2324000 | Baseflow to the Steinhatchee River near Cross City |
| qr_2326000 | Baseflow to the Econfina River near Perry |
| qr_2313700 | Baseflow to the Waccasassa River near Gulf Hammock |
| qr_2326550 | Baseflow pickup in the reach upstream of the Aucilla River near Nutall Rise |
| qspring_s121610002 | Blue Spring near Bronson |
| qspring_s101429027 | Little Fanning Springs near Fanning Spring |
| qspring_s101429001 | Fanning Springs near Wilcox |
| qspring_n011117008 | Madison Blue Spring near Blue Springs |
| qr_lsf_sprgrp | Lower Santa Fe Springs Group |
| qr_iche_sprgrp | Ichetucknee Springs Group |
| qr_wacissa_sprgrp | Wacissa Springs Group |
| qr_silver_sprgrp | Silver Springs Group |
| w3_lakebut | UFA head near Lake Butler |
| w3_hampton | UFA head near Hampton Lake |
| w3_sfelake | UFA head near Santa Fe Lake |
| w3_lkalto | UFA head near Lake Alto |
| w3_paleslk | UFA head near Palestine Lake |
| w3_oceanpo | UFA head near Ocean Pond |
| w3_cherryl | UFA head near Cherry Lake |
| w3_falmout | UFA head near Falmouth |
| w3_whitespr | UFA head near White Springs |
| qspring_s071634012 | Devil's Ear Spring |
| qspring_s081706005 | Poe Spring |
| qspring_s071727012 | Treehouse Spring |
| qspring_s071727010 | Hornsby Spring |
| qspring_s111326002 | Manatee Springs |
| qspring_s041121001 | Lafayette Blue Spring |
| qr_peacock_sprgrp | Peacock and Bonnet Springs |
| qspring_s051334002 | Troy Spring |
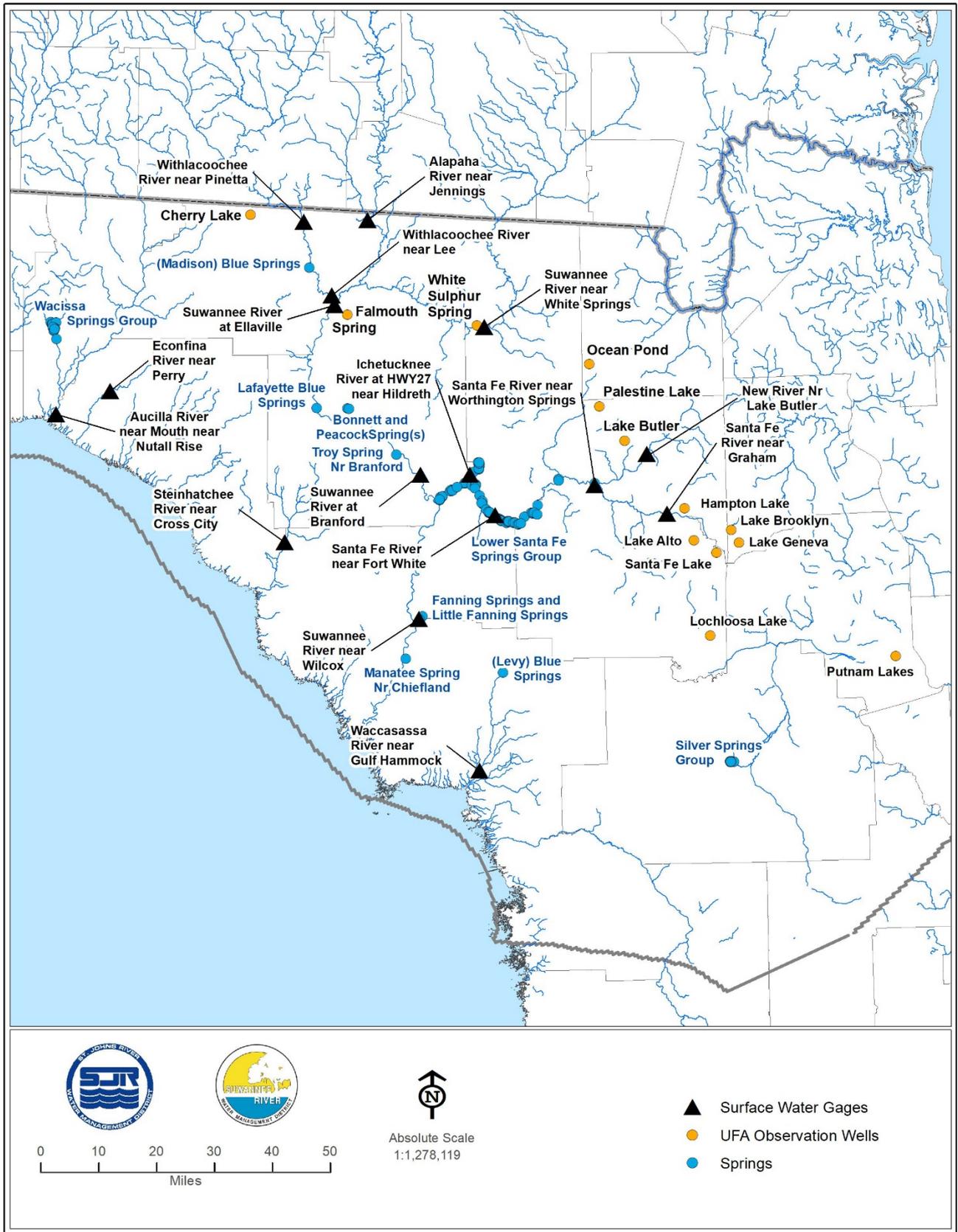
**Figure 4-8. Map of locations at which predictive uncertainty is evaluated. Points shown as orange hexagons are locations of simulated Upper Floridan aquifer groundwater levels. Points shown as green circles are springs. Black triangles represent the downstream limits of simulated river reaches.**

For each prediction, uncertainty statistics (tables 4.2 and 4.3) were calculated for the prediction itself, and for the difference between the prediction value calculated for 2035 and that calculated for 2009 (one of the years used in model calibration). For each of these two types of prediction, uncertainties can be displayed graphically as histograms, and as estimated cumulative probability distributions. Examples of these plots are provided in figures 4.9 through 4.19. The data on which these distributions are based were amassed by running the model 522 times – once for each of the calibration constrained parameter fields whose calculation was discussed in section 3 of this document. Similar plots can be made for any other prediction of interest.

**Table 4.2 Statistical summaries of 2035 predicted values.**

| Prediction location identifier | Prediction location description | Prediction mean | Standard deviation | Coeff. of variation |
|---|---|---|---|---|
| qr_2319394 | Baseflow pickup in the reach upstream of the Withlacoochee River near Lee | -432 | 6.0 | 0.01 |
| qr_2319500 | Baseflow pickup in reach upstream of the Suwannee River near Ellaville | -735 | 10 | 0.01 |
| qr_2320500a | Baseflow to the reach upstream of the Suwannee River near Branford and downstream from Ellaville | -912 | 8.4 | 0.01 |
| qr_2320700 | Baseflow to the Santa Fe River near Graham | -3.5 | 0.39 | 0.11 |
| qr_2321000 | Baseflow to the New River near Lake Butler | -19.4 | 0.79 | 0.04 |
| qr_2322700 | Baseflow to the Ichetucknee River at US Highway 27 near Hildreth | -269 | 4.82 | 0.02 |
| qr_2323500 | Baseflow pickup in reach upstream of the Suwannee River near Wilcox (downstream of Branford and Santa Fe River near Hildreth) | -432 | 6.73 | 0.02 |
| qr_2324000 | Baseflow to the Steinhatchee River near Cross City | -54 | 2.29 | 0.04 |
| qr_2326000 | Baseflow to the Econfina River near Perry | -49 | 1.11 | 0.02 |
| qr_2326550 | Baseflow pickup in the reach upstream of the Aucilla River near Nutall Rise (downstream of Wacissa River near Wacissa and the Aucilla River at Lamont) | -967 | 61.48 | 0.06 |
| qr_iche_sprgrp | Ichetucknee Springs Group | -259 | 3.02 | 0.01 |
| qr_lsf_sprgrp | Lower Santa Fe Springs Group | -794 | 6.55 | 0.01 |
| qr_silver_sprgrp | Silver Springs Group | -466 | 7.88 | 0.02 |
| qr_wacissa_sprgrp | Wacissa Springs Group | -452 | 11.21 | 0.02 |
| qs_2315500 | Baseflow to the Suwannee River near White Springs | -212 | 21.25 | 0.10 |
| qs_2317620 | Baseflow to the Alapaha River near Jennings | -825 | 28.06 | 0.03 |
| qs_2319000 | Baseflow to the Withlacoochee River near Pinetta | -866 | 35.90 | 0.04 |
| qs_2319500 | Baseflow to the Suwannee River at Ellaville | -3070 | 52.84 | 0.02 |
| qs_2320500 | Baseflow to the Suwannee River near Branford | -3981 | 53.75 | 0.01 |
| qs_2321500 | Baseflow to the Santa Fe River near Worthington Springs | -47 | 1.37 | 0.03 |
| qs_2322500 | Baseflow to the Santa Fe River near Fort White | -707 | 6.56 | 0.01 |
| qs_2323500 | Baseflow to the Suwannee River near Wilcox | -5541 | 54.29 | 0.01 |
| qspring_n011117008 | Madison Blue Spring near Blue Springs | -102 | 0.77 | 0.01 |
| qspring_s101429001 | Fanning Springs near Wilcox | -66 | 0.77 | 0.01 |
| qspring_s101429027 | Little Fanning Springs near Fanning Spring | -1.8 | 0.04 | 0.02 |
| qspring_s121610002 | Blue Spring near Bronson | -2.7 | 0.82 | 0.30 |

| w00202 | UFA observation well near Lake Lochloosa | 77.9 | 0.30 | 0.00 |
| w00258 | UFA observation well near Lake Brooklyn | 77.7 | 0.33 | 0.00 |
| w00400 | UFA observation well near Lake Geneva | 63.4 | 0.23 | 0.00 |
| w00878 | UFA observation well near Putnam County MFL lakes | 26.4 | 0.13 | 0.00 |
| W3_CHERRYL | UFA groundwater level near Cherry Lake | 65.5 | 0.66 | 0.01 |
| W3_FALMOUT | UFA groundwater level near Falmouth Spring | 38.6 | 0.10 | 0.00 |
| W3_HAMPTON | UFA groundwater level near Hampton Lake | 65.4 | 0.55 | 0.01 |
| W3_LAKEBUT | UFA groundwater level near Lake Butler | 53.1 | 0.37 | 0.01 |
| W3_LKALTO | UFA groundwater level near Lake Altho | 70.4 | 0.39 | 0.01 |
| W3_OCEANPO | UFA groundwater level near Ocean Pond | 47.8 | 0.25 | 0.01 |
| W3_PALESLK | UFA groundwater level near Palestine Lake | 49.8 | 0.30 | 0.01 |
| W3_SFELAKE | UFA groundwater level near Santa Fe Lake | 77.8 | 0.41 | 0.01 |
| W3_WHITSPR | UFA groundwater level near White Sulphur Springs | 50.6 | 0.22 | 0.00 |
| qr_2313700 | Waccasassa River near Gulf Hammock | -122.3 | 11.74 | -0.10 |
| qr_peackock_sprgrp | Peacock and Bonnet Springs | -44.1 | 0.40 | 0.01 |
| qspring_s041121001 | Lafayette Blue Spring | -54.6 | 3.03 | 0.06 |
| qspring_s051334002 | Troy Spring | -92.7 | 5.36 | 0.06 |
| qspring_s071634012 | Devils Ear Spring | -112.4 | 1.65 | 0.01 |
| qspring_s071727010 | Hornsby Spring | -16.0 | 0.90 | 0.06 |
| qspring_s071727012 | Treehouse Spring | -2.5 | 0.38 | 0.15 |
| qspring_s081706005 | Poe Spring | -41.6 | 0.49 | 0.01 |
| qspring_s111326002 | Manatee Springs | -125.5 | 1.73 | 0.01 |

**Table 4.3 Statistical summaries of predicted changes from 2009 to 2035.**

| Prediction location identifier | Prediction location description | Mean of predicted change | Standard Deviation | Coeff. of Variation |
|---|---|---|---|---|
| qr_2319394 | Baseflow pickup in the reach upstream of the Withlacoochee River near Lee | 16.0 | 1.27 | 0.08 |
| qr_2319500 | Baseflow pickup in reach upstream of the Suwannee River near Ellaville | 42.4 | 2.65 | 0.06 |
| qr_2320500a | Baseflow to the reach upstream of the Suwannee River near Branford and downstream from Ellaville | 17.8 | 0.37 | 0.02 |
| qr_2320700 | Baseflow to the Santa Fe River near Graham | 0.15 | 0.02 | 0.10 |
| qr_2321000 | Baseflow to the New River near Lake Butler | 0.26 | 0.02 | 0.09 |
| qr_2322700 | Baseflow to the Ichetucknee River at US Highway 27 near Hildreth | 7.54 | 0.26 | 0.03 |
| qr_2323500 | Baseflow pickup in reach upstream of the Suwannee River near Wilcox (downstream of Branford and Santa Fe River near Hildreth) | 9.21 | 0.12 | 0.01 |
| qr_2324000 | Baseflow to the Steinhatchee River near Cross City | 0.04 | 0.00 | 0.11 |
| qr_2326000 | Baseflow to the Econfina River near Perry | 0.32 | 0.05 | 0.14 |

| qr_2326550 | Baseflow pickup in the reach upstream of the Aucilla River near Nutall Rise (downstream of Wacissa River near Wacissa and the Aucilla River at Lamont) | 7.94 | 3.13 | 0.39 |
|---|---|---|---|---|
| qr_iche_sprgrp | Ichetucknee Springs Group | 7.42 | 0.25 | 0.03 |
| qr_lsf_sprgrp | Lower Santa Fe Springs Group | 16.6 | 0.73 | 0.04 |
| qr_silver_sprgrp | Silver Springs Group | 21.9 | 1.33 | 0.06 |
| qr_wacissa_sprgrp | Wacissa Springs Group | 2.49 | 2.89 | 1.16 |
| qs_2315500 | Baseflow to the Suwannee River near White Springs | 0.55 | 0.13 | 0.23 |
| qs_2317620 | Baseflow to the Alapaha River near Jennings | -0.34 | 0.15 | 0.44 |
| qs_2319000 | Baseflow to the Withlacoochee River near Pinetta | 8.84 | 1.20 | 0.14 |
| qs_2319500 | Baseflow to the Suwannee River at Ellaville | 67.4 | 4.42 | 0.07 |
| qs_2320500 | Baseflow to the Suwannee River near Branford | 85.2 | 4.56 | 0.05 |
| qs_2321500 | Baseflow to the Santa Fe River near Worthington Springs | 0.75 | 0.05 | 0.06 |
| qs_2322500 | Baseflow to the Santa Fe River near Fort White | 15.4 | 0.77 | 0.05 |
| qs_2323500 | Baseflow to the Suwannee River near Wilcox | 118.97 | 4.66 | 0.04 |
| qspring_n011117008 | Madison Blue Spring near Blue Springs | 2.43 | 0.19 | 0.08 |
| qspring_s101429001 | Fanning Springs near Wilcox | 0.88 | 0.03 | 0.04 |
| qspring_s101429027 | Little Fanning Springs near Fanning Spring | 0.02 | 0.00 | 0.04 |
| qspring_s121610002 | Blue Spring near Bronson | 0.91 | 0.23 | 0.25 |
| w00202 | UFA observation well near Lake Lochloosa | -1.83 | 0.05 | 0.03 |
| w00258 | UFA observation well near Lake Brooklyn | -1.88 | 0.06 | 0.03 |
| w00400 | UFA observation well near Lake Geneva | -0.33 | 0.02 | 0.06 |
| w00878 | UFA observation well near Putnam County MFL lakes | 1.96 | 0.04 | 0.02 |
| W3_CHERRYL | UFA groundwater level near Cherry Lake | -0.41 | 0.05 | 0.11 |
| W3_FALMOUT | UFA groundwater level near Faltmouth Spring | -0.13 | 0.01 | 0.07 |
| W3_HAMPTON | UFA groundwater level near Hampton Lake | -1.25 | 0.05 | 0.04 |
| W3_LAKEBUT | UFA groundwater level near Lake Butler | -1.58 | 0.10 | 0.06 |
| W3_LKALTO | UFA groundwater level near Lake Altho | -1.23 | 0.05 | 0.04 |
| W3_OCEANPO | UFA groundwater level near Ocean Pond | -1.82 | 0.11 | 0.06 |
| W3_PALESLK | UFA groundwater level near Palestine Lake | -1.69 | 0.09 | 0.05 |
| W3_SFELAKE | UFA groundwater level near Santa Fe Lake | -1.60 | 0.06 | 0.03 |
| W3_WHITSPR | UFA groundwater level near White Sulphur Springs | -0.53 | 0.06 | 0.11 |
| qr_2313700 | Waccasassa River near Gulf Hammock | 1.83 | 0.37 | 0.20 |
| qr_peacock_sprgrp | Peacock and Bonnet Springs | 0.56 | 0.02 | 0.04 |
| qspring_s041121001 | Lafayette Blue Spring | 2.09 | 0.16 | 0.07 |
| qspring_s051334002 | Troy Spring | 1.42 | 0.09 | 0.06 |
| qspring_s071634012 | Devils Ear Spring | 1.25 | 0.08 | 0.06 |
| qspring_s071727010 | Hornsby Spring | 0.14 | 0.05 | 0.33 |
| qspring_s071727012 | Treehouse Spring | 0.15 | 0.02 | 0.14 |
| qspring_s081706005 | Poe Spring | 0.45 | 0.04 | 0.08 |
| qspring_s111326002 | Manatee Springs | 0.78 | 0.06 | 0.08 |
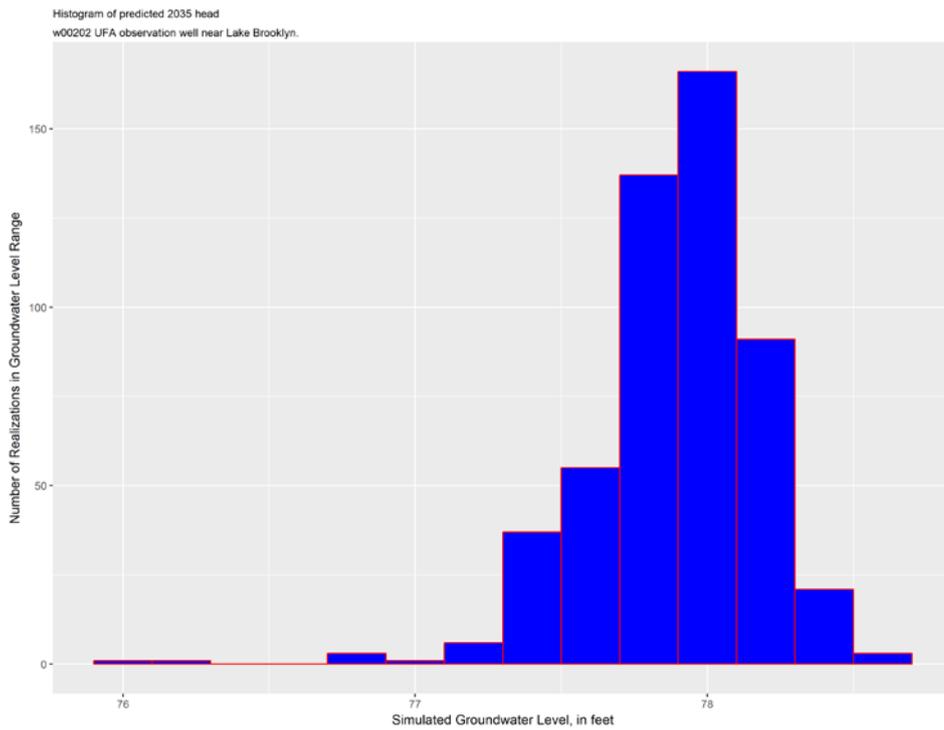
**Figure 4-9. Histogram for the predicted Upper Floridan aquifer groundwater level near Lake Brooklyn for the 2035 hypothetical withdrawal scenario. The histogram was developed from 522 randomly-generated predictions.**
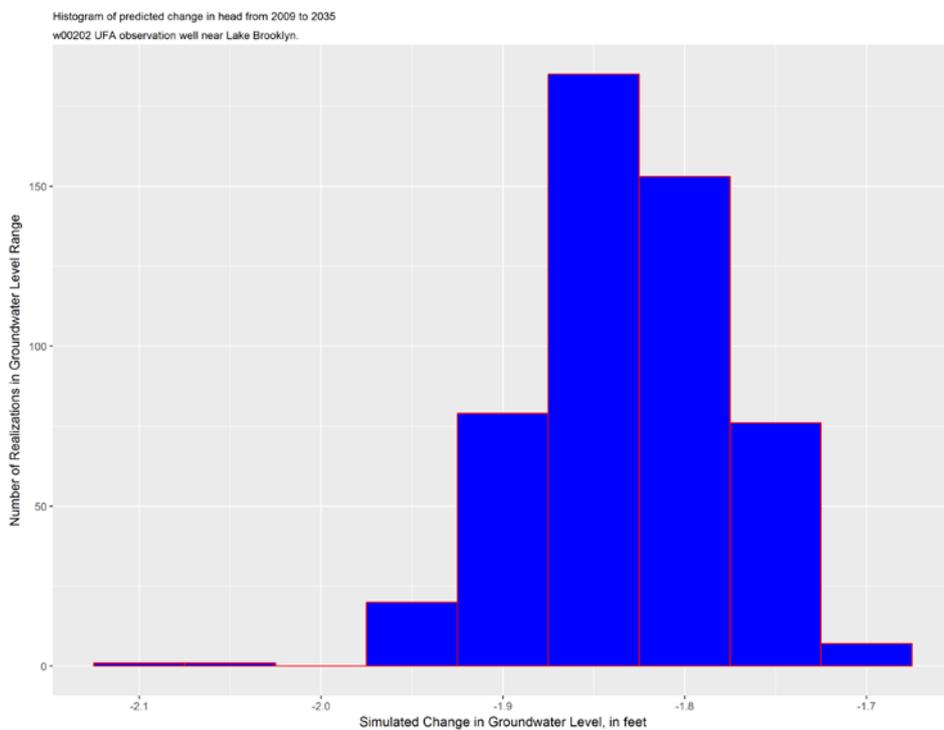


**Figure 4-10 Histogram for the predicted change in the Upper Floridan aquifer groundwater level near Lake Brooklyn from 2009 to the 2035 hypothetical withdrawal scenario. The histogram was developed from 522 randomly-generated predictions.**
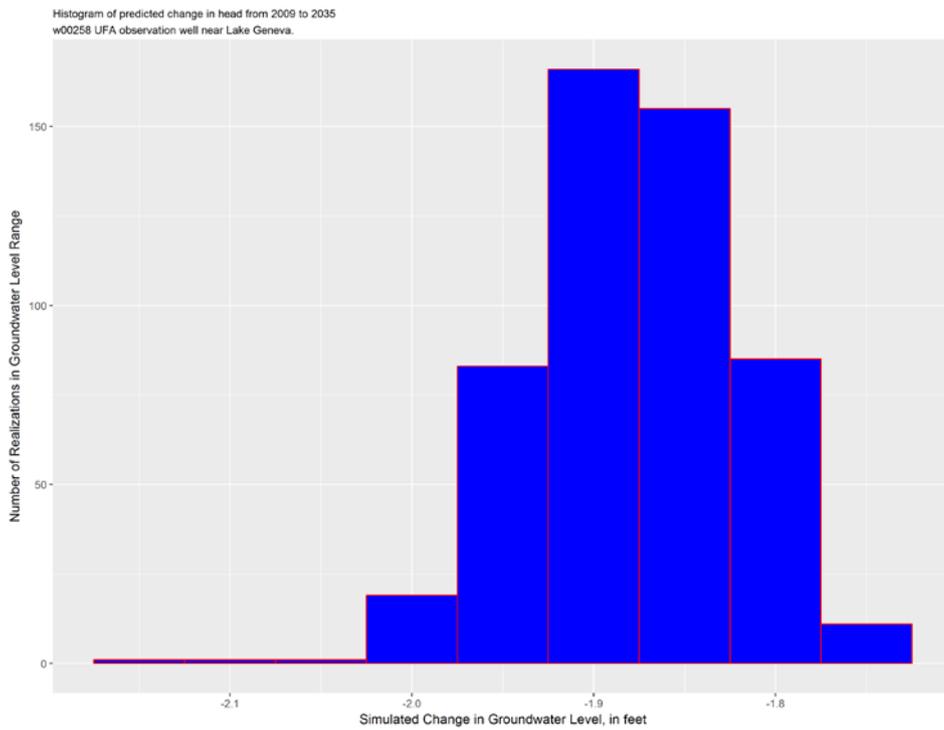
**Figure 4-11. Histogram for the predicted change in the Upper Floridan aquifer groundwater level near Lake Geneva from 2009 to the 2035 hypothetical withdrawal scenario. The histogram was developed from 522 randomly-generated predictions.**
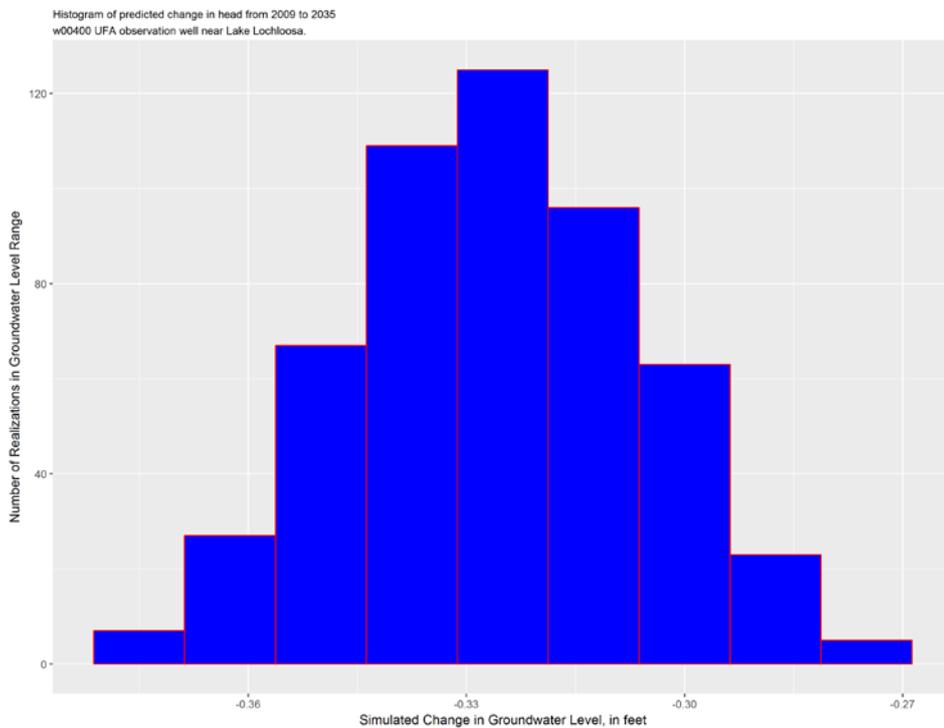


**Figure 4-12. Histogram for the predicted change in the Upper Floridan aquifer groundwater level near Lake Lochloosa from 2009 to the 2035 hypothetical withdrawal scenario. The histogram was developed from 522 randomly-generated predictions.**
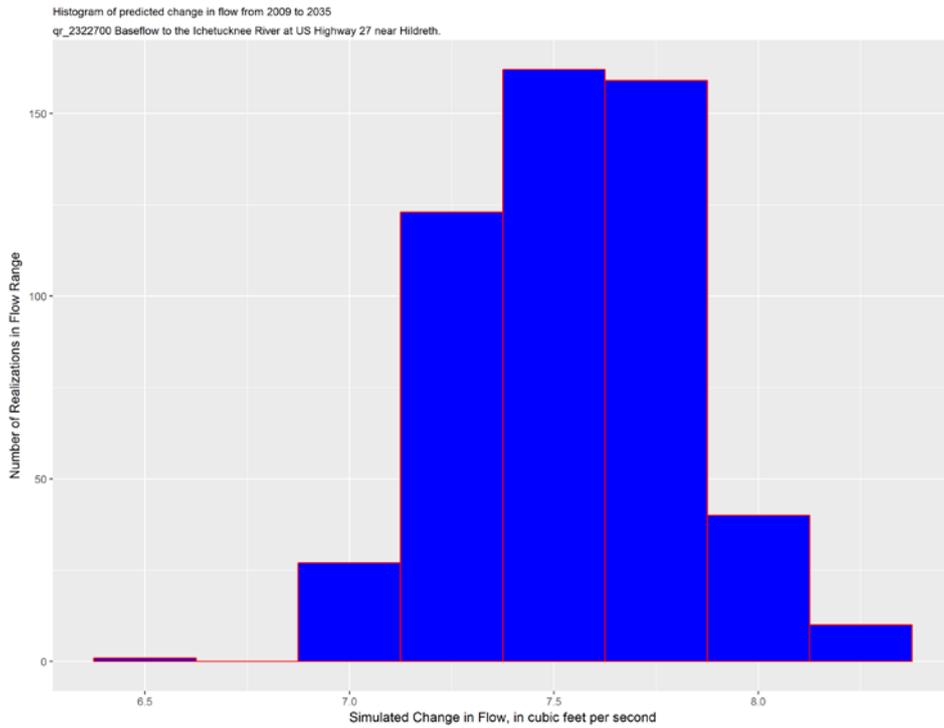
**Figure 4-13. Histogram for the predicted change in flow in the Ichetucknee River near US Highway 27 near Hildreth from 2009 to the 2035 hypothetical withdrawal scenario. The histogram was developed from 522 randomly-generated predictions.**
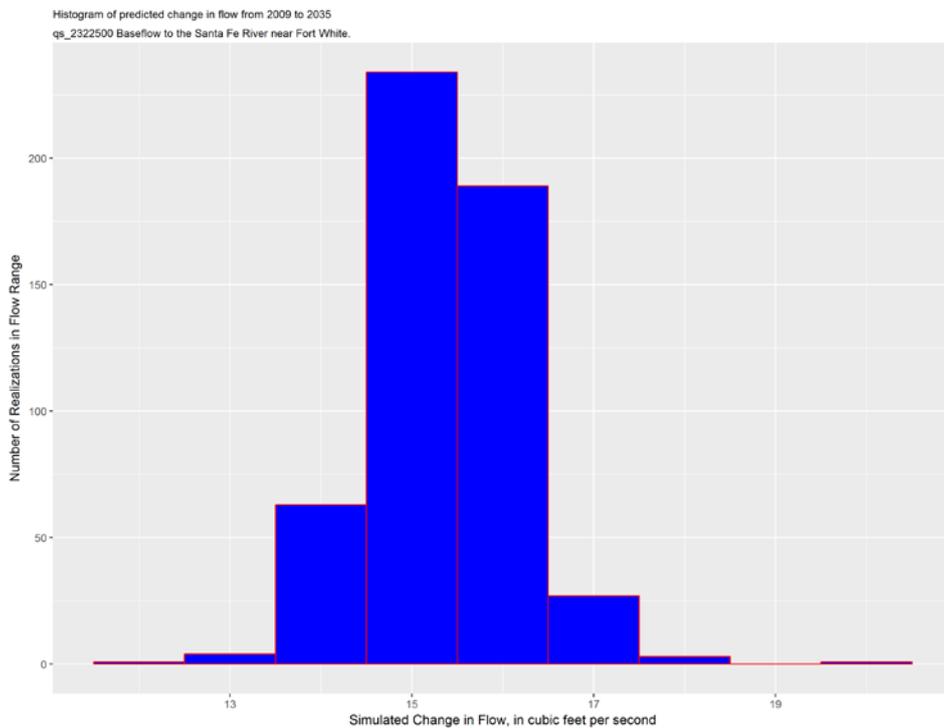


**Figure 4-14. Histogram for the predicted change in flow in the Ichetucknee River near US Highway 27 near Hildreth from 2009 to the 2035 hypothetical withdrawal scenario. The histogram was developed from 522 randomly-generated predictions.**
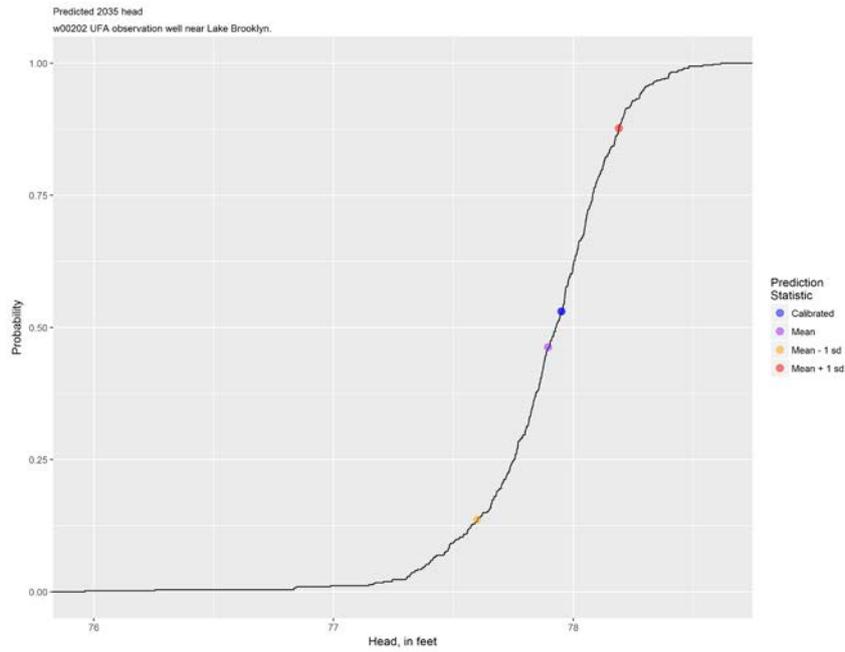
**Figure 4-15. Estimated cumulative probability distribution for the predicted Upper Floridan aquifer groundwater level near Lake Brooklyn for the 2035 hypothetical withdrawal scenario. The probability distribution was developed from 522 randomly-generated predictions. Also shown are the mean predicted value and the prediction based on the calibrated parameter field, along with values corresponding to the mean plus and minus one standard deviation.**
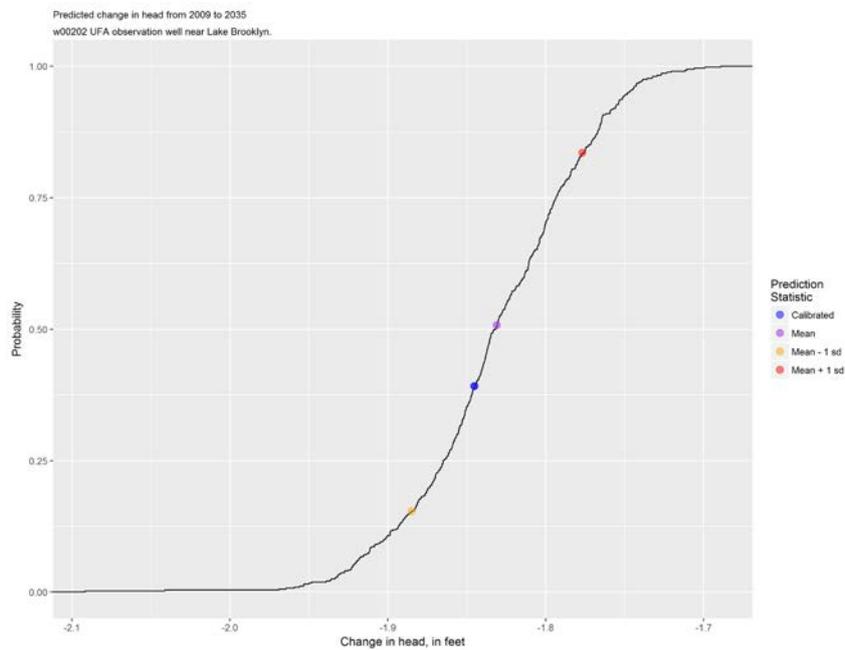


**Figure 4-16. Estimated cumulative probability distribution for the predicted change in the Upper Floridan aquifer groundwater level near Lake Brooklyn from 2009 to the 2035 hypothetical withdrawal scenario. The probability distribution was developed from 522 randomly-generated predictions. Also shown are the mean predicted value and the prediction based on the calibrated parameter field, along with values corresponding to the mean plus and minus one standard deviation.**
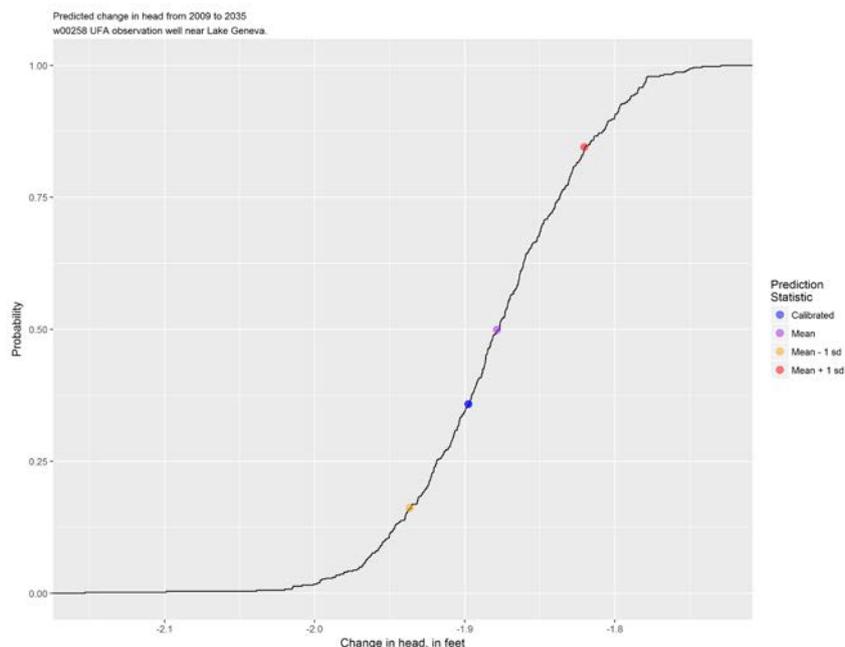
**Figure 4-17. Estimated cumulative probability distribution for the predicted change in the Upper Floridan aquifer groundwater level near Lake Geneva from 2009 to the 2035 hypothetical withdrawal scenario. The probability distribution was developed from 522 randomly-generated predictions. Also shown are the mean predicted value and the prediction based on the calibrated parameter field, along with values corresponding to the mean plus and minus one standard deviation.**
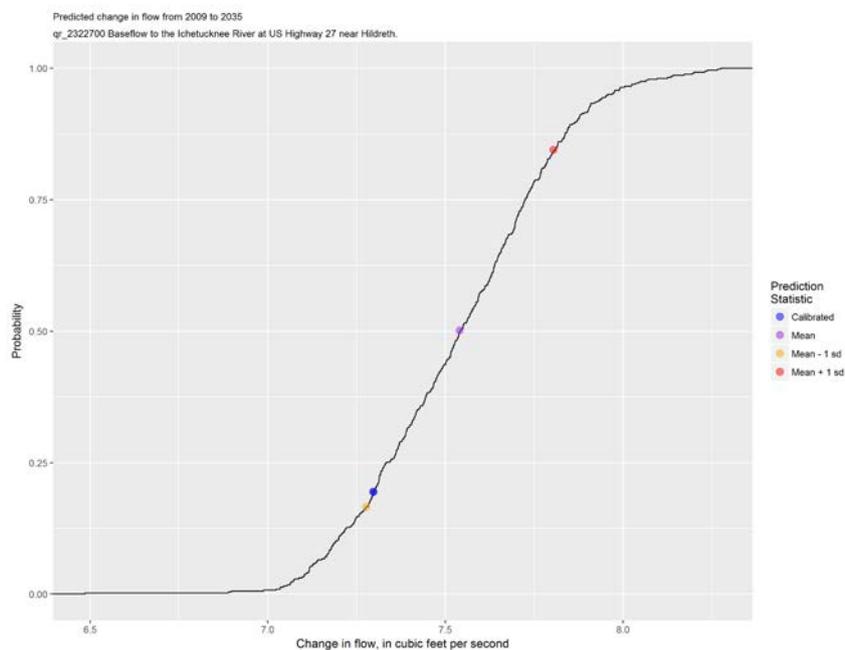


**Figure 4-18. Estimated cumulative probability distribution for the predicted change in the flow of the Ichetucknee River at US Highway 27 near Hildreth from 2009 to the 2035 hypothetical withdrawal scenario. The probability distribution was developed from 522 randomly-generated predictions. Also shown are the mean predicted value and the prediction based on the calibrated parameter field, along with values corresponding to the mean plus and minus one standard deviation.**
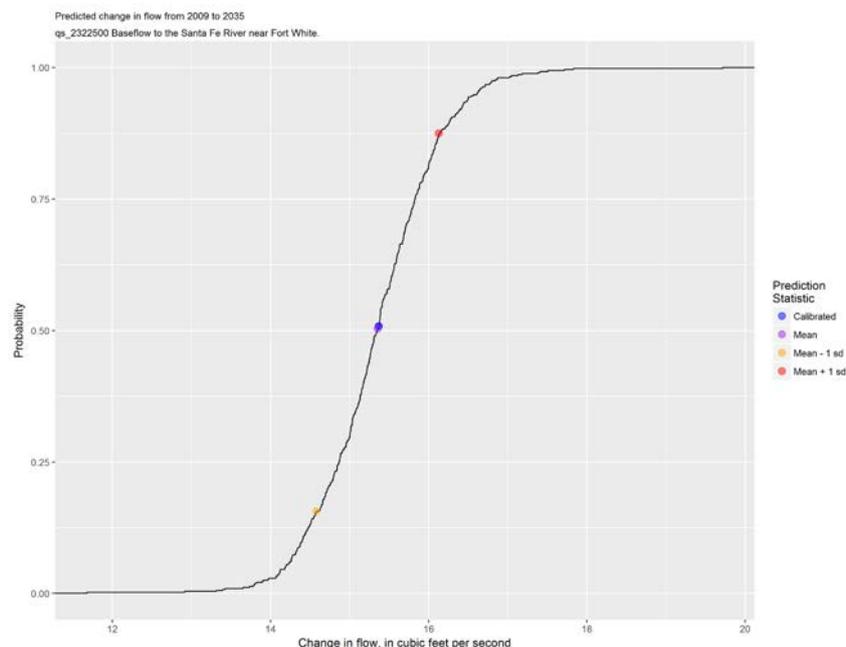
**Figure 4-19. Estimated cumulative probability distribution for the predicted change in the flow of the Santa Fe River near Fort White from 2009 to the 2035 hypothetical withdrawal scenario. The probability distribution was developed from 522 randomly-generated predictions. Also shown are the mean predicted value and the prediction based on the calibrated parameter field, along with values corresponding to the mean plus and minus one standard deviation.**

Intuitively, uncertainties associated with temporal predictive differences are likely to be much smaller than those associated with the individual predictions. This is an outcome of the high degree of temporal correlation that is likely to exist between predictions made at the same location but at different times under system stress regimes that are not too different. To see why this is so, let the uncertainty variance of a prediction made at time $t_1$ be specified as $\sigma^2_1$ and the uncertainty variance of a prediction made at the same location but at time $t_2$ be specified as $\sigma^2_2$. Let us designate the covariance between these predictions as $\sigma^2_{12}$. Using standard relationships for propagation of variance, the variance of uncertainty of the difference between these predictions is calculated as:

$$\sigma^2_{1-2} = \sigma^2_1 + \sigma^2_2 - 2\sigma_{12} \tag{4.2}$$

If the two predictions have about the same uncertainty and experience a correlation coefficient close to 1, it can be easily shown that $\sigma^2_{1-2}$ approaches zero.

While equation 4.2 can partially explain the small uncertainties associated with predictive differences displayed in the above figures, intuition also suggests that the smallness of these uncertainties may nevertheless overstate the accuracy of model-predicted differences in system state arising from differences in system stress. It is suggested that this is probably the case, for the uncertainties computed in the manner discussed in section 3 of this document do not account for model errors. These arise from numerous sources that afflict both the calibration and predictive processes. To the extent that errors can be endowed with a stochastic description, their temporal correlation is likely to be high. Hence equation 4.2 applies to them as well. This, it is hoped, will indeed raise the integrity of predictive differences to a higher level than that of predictive absolutes. However, quantification of the effect of model errors (also referred to herein as "structural noise") is not possible without

recourse to a model from which these errors are absent. Obviously, no such model is available. This matter is further discussed in section 6.

# 5. Outcomes of Linear Analysis

## 5.1 Statistics used in Calculations

As for nonlinear analysis, linear analysis requires a C(**k**) matrix and a C(**ε**) matrix. These are used in equations 2.3a and 2.3b to calculate C'(**k**). They are also used by the PEST SUPCALC utility in estimating the dimensionality of the calibration solution space.

The C(**k**) matrix used for linear analysis is slightly different from that which was employed for nonlinear analysis, in that many parameters are considered to have slightly greater prior uncertainties. These are the same uncertainties as those employed in previous linear analysis conducted on the NFSEG model. (The reasons for use of diminished prior uncertainties in nonlinear analysis were explained in section 3.) However, variograms and correlations used in linear and nonlinear analysis were the same for all parameters. Prior parameter variances used in linear analysis are tabulated by group in table 5.1; this is the counterpart to table 3.3 provided in the previous section.

| Parameter group name | Standard deviation ascribed to each parameter within group |
| --- | --- |
| vanis1 | 0.25 |
| vanis2 | 0.25 |
| vanis4 | 0.25 |
| vanis5 | 0.25 |
| vanis6 | 0.25 |
| vanis7 | 0.25 |
| Lcm | 1.0 |
| Rcm | 0.6 |
| Sc | 1.0 |
| Lkzmul | 0.6 |

**Table 5.1 Standard deviations assigned independently to each parameter within each respective parameter group for use in linear analysis. Note that these are actually applied to the log (to base 10) of each parameter. Note also that none of the parameters appearing in the above table are represented by pilot points in the NFSEG model.**

C(**ε**) is, once again, assumed to be diagonal. For each observation group the PEST PWTADJ2 utility was used to apply a factor to measurement weights used in the calibration process such that the objective function achieved for each observation group is equal to the number of non-zero weighted observations comprising the group. The diagonal elements of C(**ε**) were estimated as the inverse squared reciprocal of thus-calculated weights.

## 5.2 Results

Maps of identifiability and relative uncertainty variance reduction for selected parameter fields are shown in figures 5-1 through 5.7. These maps are broadly similar for each parameter type. However, identifiabilities tend to be more "polarised" than relative uncertainty variance reduction, with parameters tending to adopt values that are either close to one or close to zero. This is an outcome of the somewhat artificial nature of this parameter, and its dependence on the number of dimensions that are assigned to the calibration solution space. It also reflects the "in or out" nature of each vector comprising the columns of the $V_1$ matrix of equation 2.5 as the dimensionality of the solution space is varied.
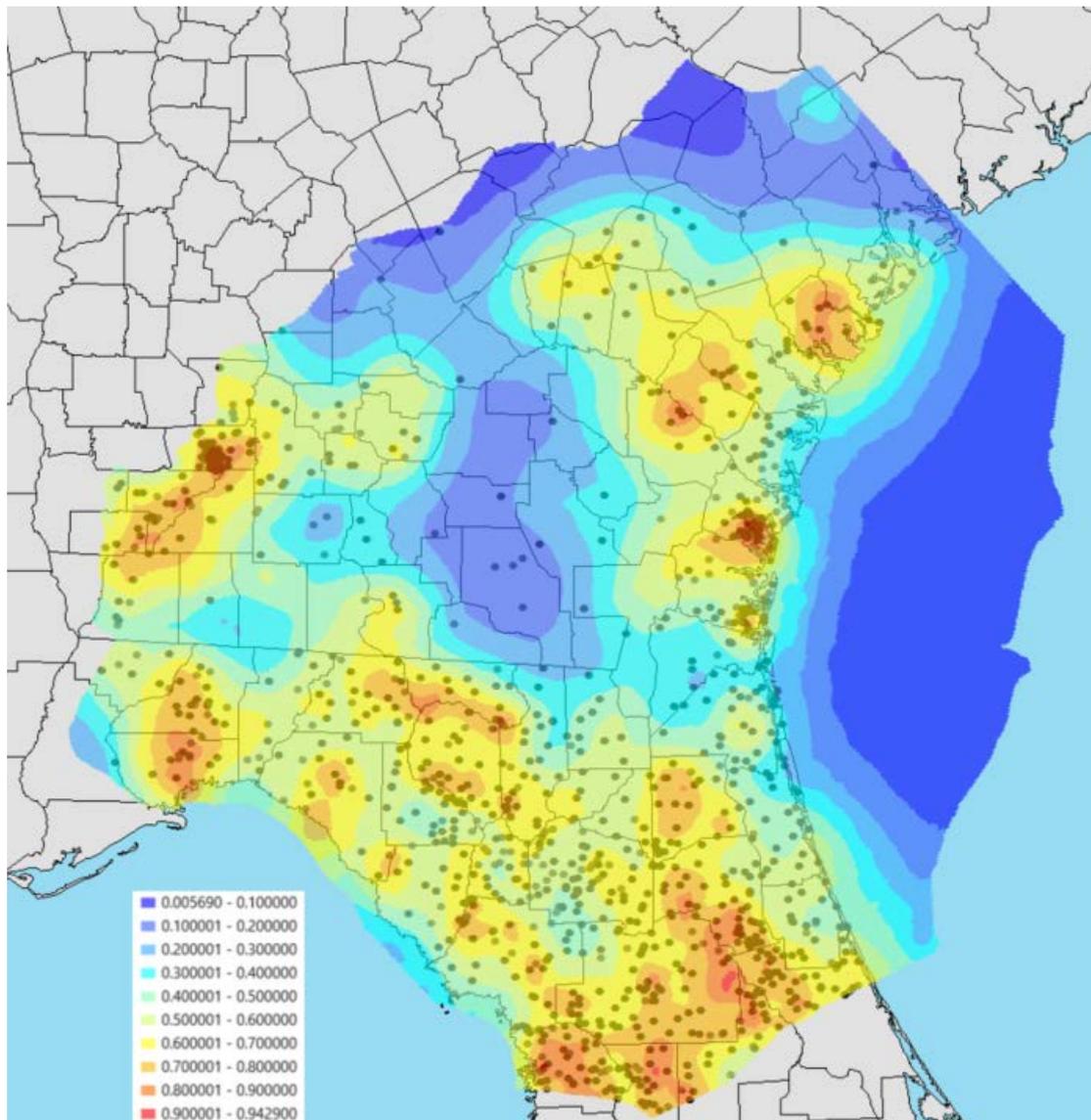
**Figure 5-1. Relative parameter uncertainty variance reduction of layer 3 horizontal hydraulic conductivity pilot points. Locations of layer 3, non-zero weighted, groundwater-level observations in 2001 and/or 2009 are superimposed on this map.**
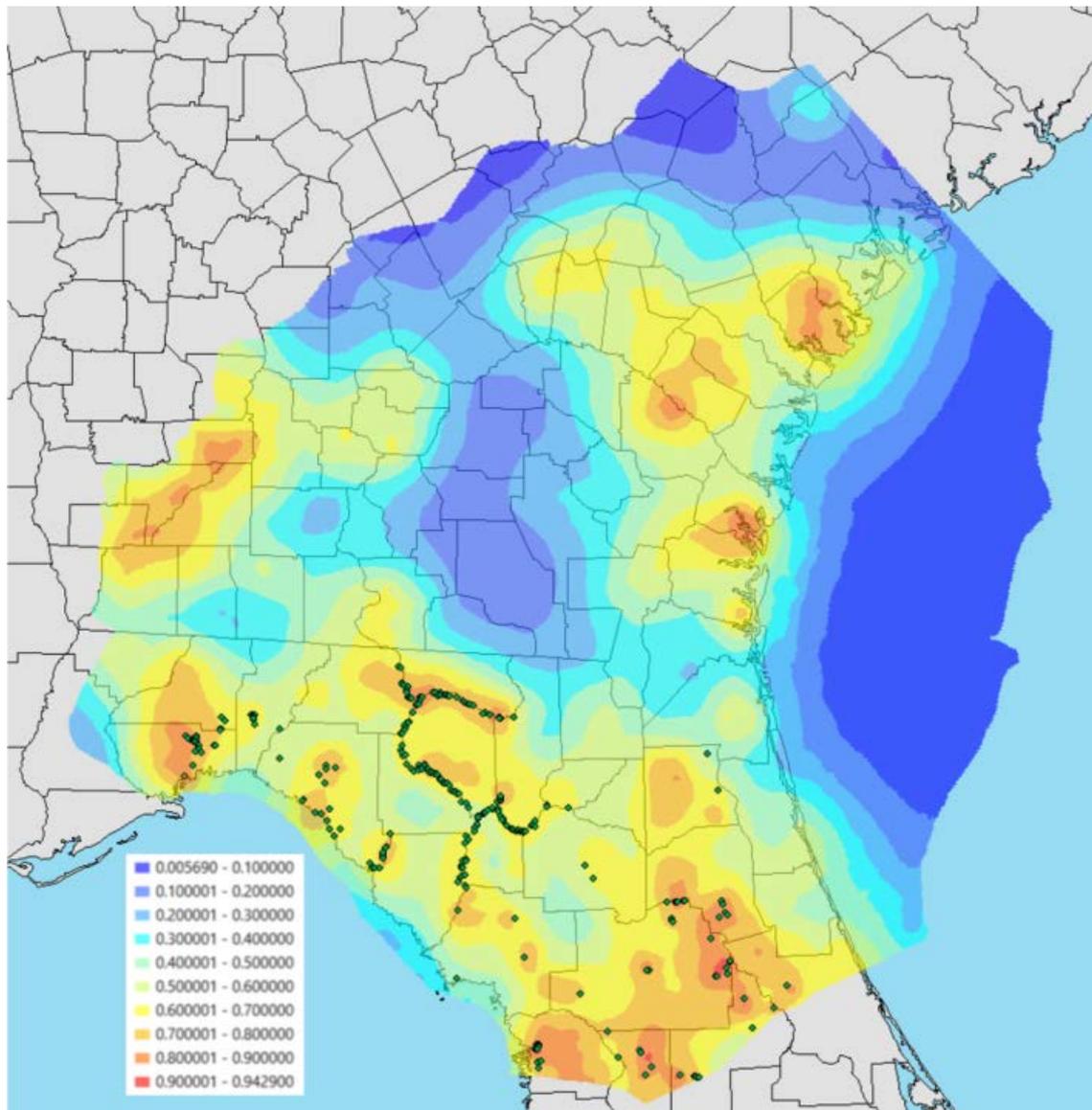
**Figure 5-2. Relative parameter uncertainty variance reduction of layer 3 horizontal hydraulic conductivity pilot points. Locations of non-zero weighted spring flow observations in 2001 and/or 2009 are superimposed on this map.**
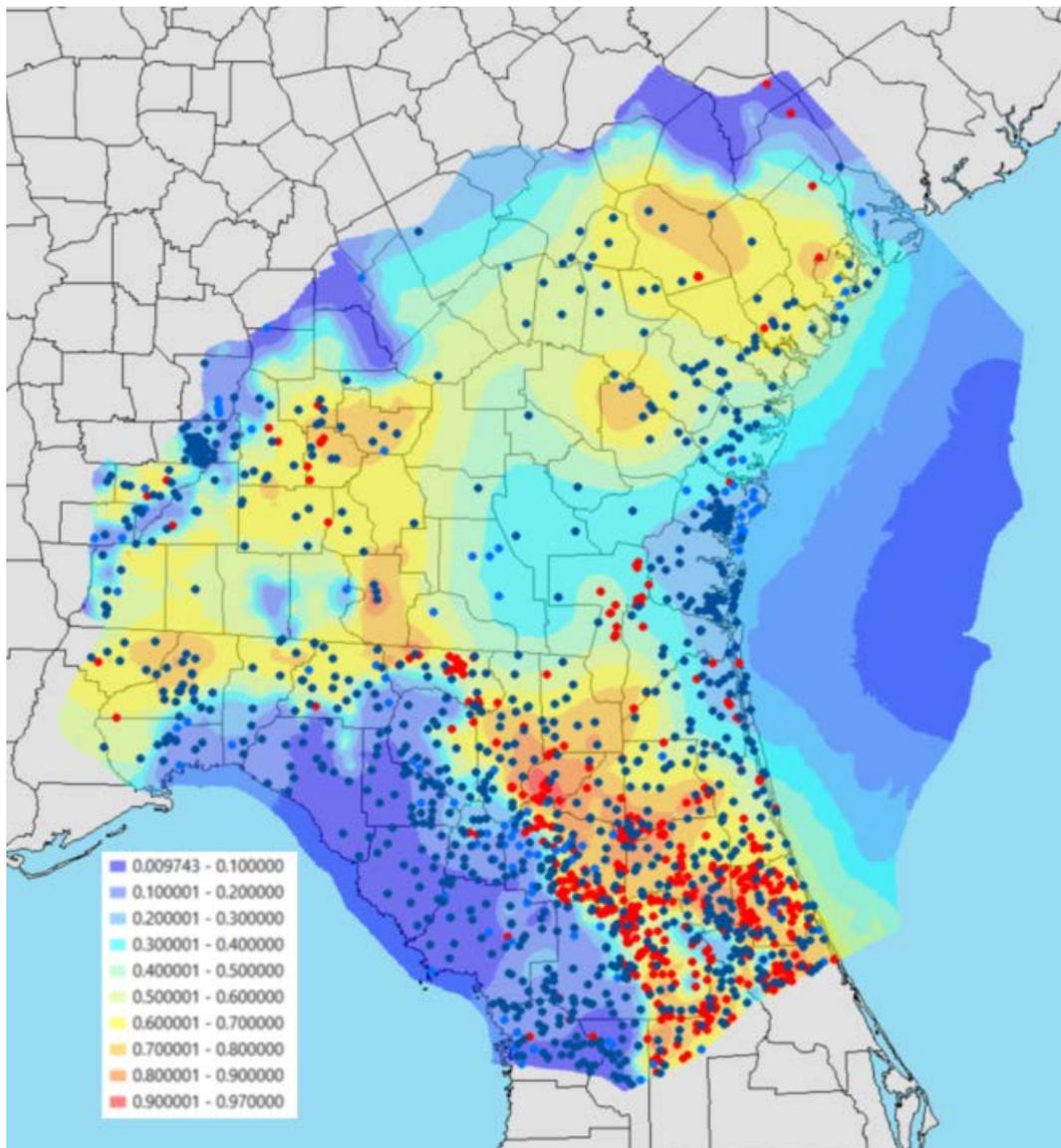
**Figure 5-3. Relative parameter uncertainty variance reduction of layer 2 vertical hydraulic conductivity pilot points. Locations of non-zero weighted, layer 1 (red points) and layer 3 (blue points) observations in 2001 and/or 2009 are superimposed on this map.**
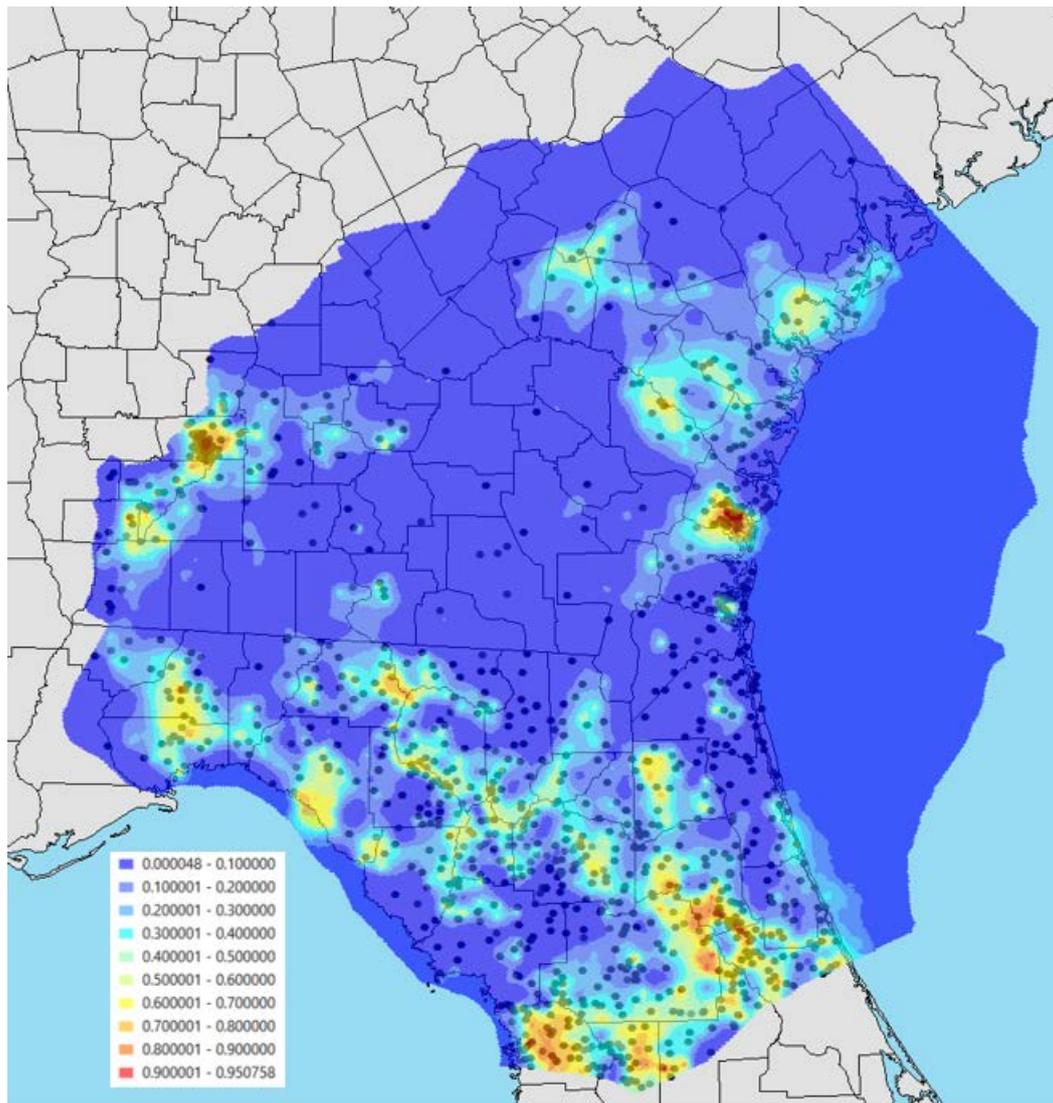
**Figure 5-4. Identifiability of layer 3 horizontal-hydraulic conductivity pilot points. Locations of layer 3, non-zero weighted, groundwater-level observations in 2001 and/or 2009 are superimposed on this map.**
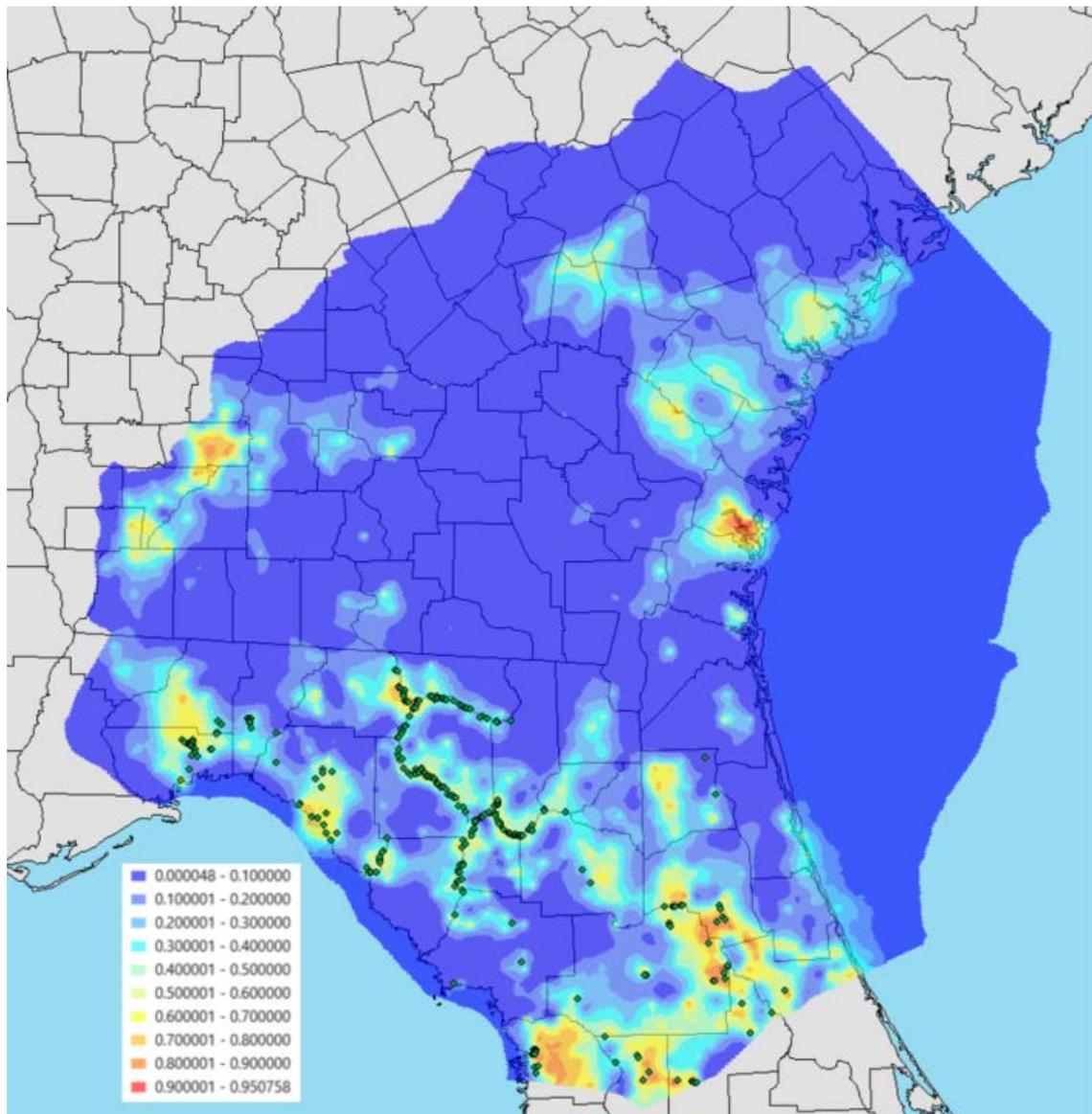
**Figure 5-5. Identifiability of layer 3 horizontal-hydraulic conductivity pilot points. Locations of non-zero weighted spring flow observations in 2001 and/or 2009 are superimposed on this map.**
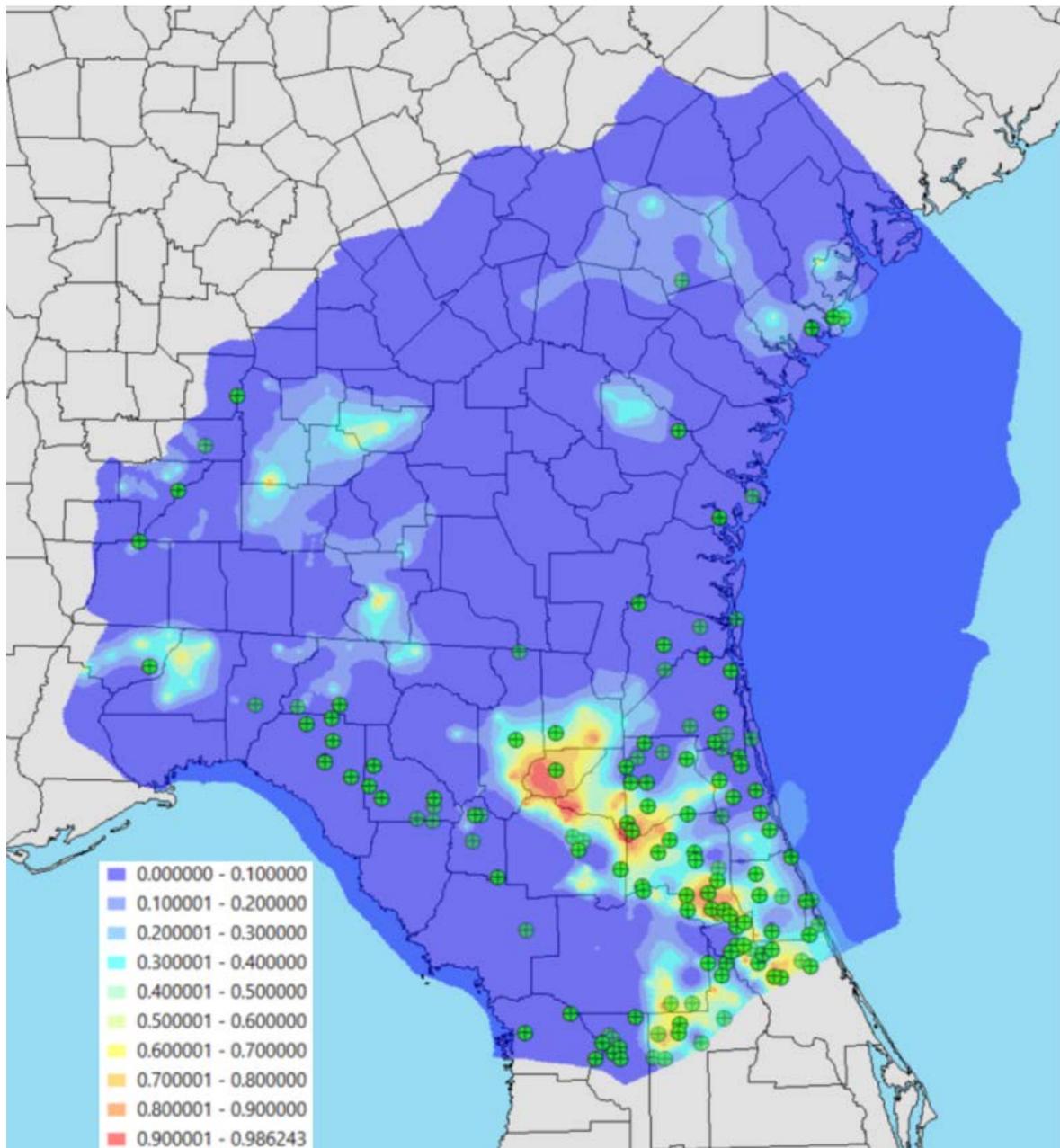
**Figure 5-6. Identifiability of layer 2 vertical-hydraulic conductivity pilot points. Locations of non-zero weighted vertical head difference observations across layer 2 in 2001 and/or 2009 are superimposed on this map.**
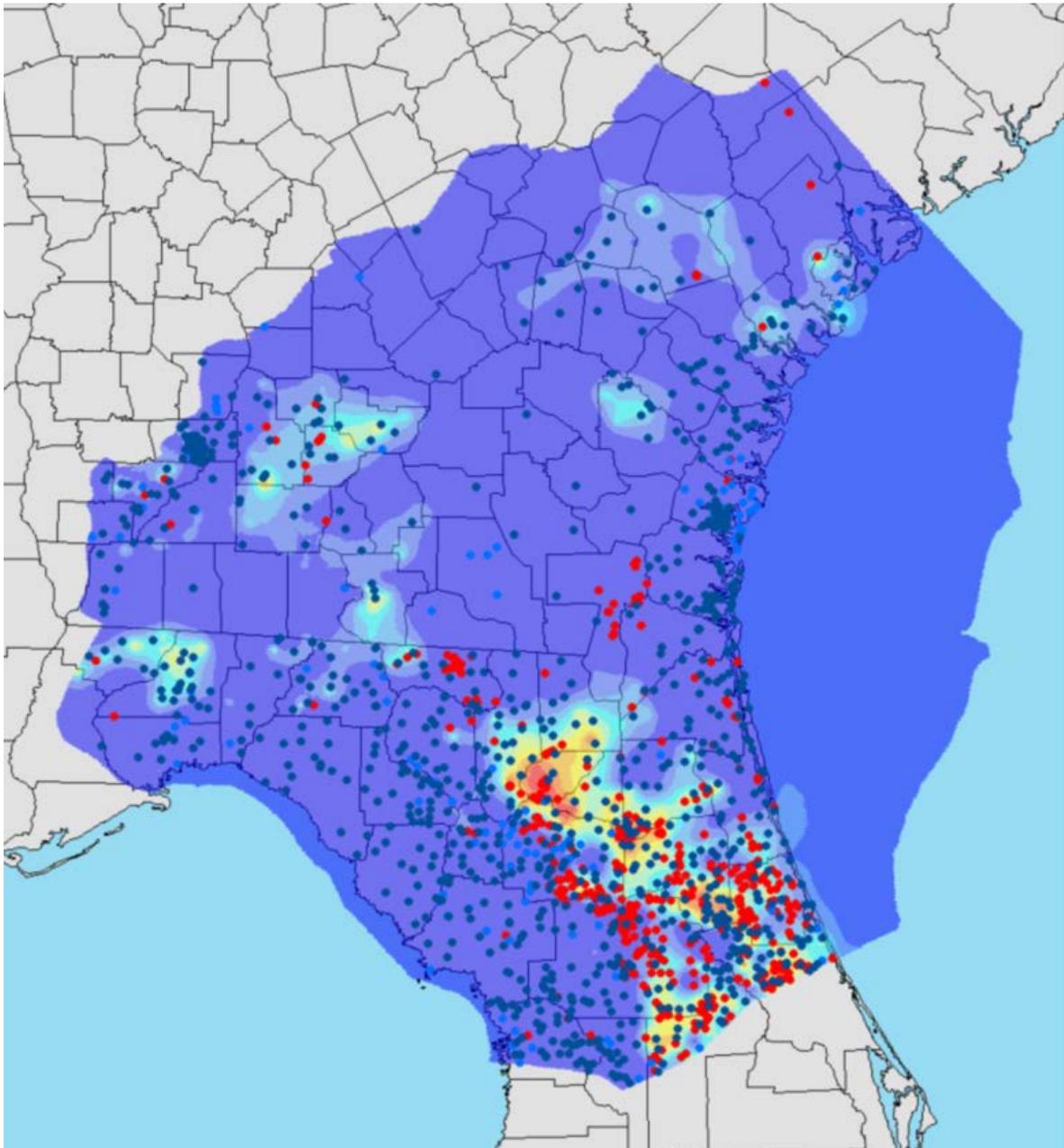
**Figure 5-7. Identifiability of layer 2 vertical-hydraulic conductivity pilot points. Locations of non-zero weighted, 2001 or 2009 groundwater-level observations in layer 1 (red points) and/or and layer 3 (blue points) are superimposed on this map.**

For pilot point parameters, the spatially diffuse nature of relative parameter uncertainty variance reduction compared to corresponding identifiabilities also arises from the fact that prior parameter uncertainties are reflected in posterior parameter uncertainties. Use of a variogram to characterize prior uncertainties results in posterior spatial correlation of these parameters. Hence if a parameter is informed by a local head, spring or baseflow measurement, that information is then passed to neighbouring (spatially correlated) parameters. This two-step passage of information is not reflected in identifiabilities; the latter statistic describes only the first of these two steps.

Notwithstanding differences in detail, the broad scale patterns exhibited by maps of identifiability and relative parameter uncertainty variance reduction can be explained in large part by variations in

spatial density of observations comprising the calibration dataset, and the sensitivities of model-calculated counterparts to these observations to pertinent model parameters. Taking k2z parameters as an example (i.e. vertical hydraulic conductivities in layer 2), the highest values for both identifiability and relative parameter uncertainty variance reduction occur in areas where the Upper Floridan aquifer is confined by the intermediate confining unit, and where high densities of observations of both groundwater level and vertical groundwater-level differences are available. Conversely, both of these statistics have lower values where groundwater level observations are sparse, and in unconfined areas where flow is primarily horizontal and (accordingly) where hydraulic property heterogeneity is reflected in the values assigned to horizontal hydraulic conductivity. Similarly, identifiabilities and relative uncertainty variance reductions are highest for k3x parameters (i.e. horizontal hydraulic conductivities in layer 3) in confined areas, and in unconfined areas where the spatial density of groundwater level observations is highest.

# 6. Discussion and Conclusions

The suite of calibration-constrained, random parameter fields that was calculated using the methodology described in chapter 2 of this appendix can be used to place a lower bound on uncertainties associated with predictions of management interest made by the NFSEG model.

As has been discussed, many predictions required by the NFSEG model are similar in nature to members of the dataset against which the NFSEG model was calibrated. Hence, they are solution space dominated. This has two effects which work in opposite directions and are difficult to quantify.

For a prediction that is solution space dominated, model structural defects can be "calibrated out" to some extent. As Welter and Doherty (2010) and White et al (2014) explain, where a model is defective (as all models are), the link between parameter uncertainty and predictive uncertainty is broken to some extent, as parameters can adopt surrogate roles to compensate for model defects without any deleterious effects being felt by solution space dependent predictions. Moreover, the structural deficiencies which give rise to this type of parameter behaviour are rendered invisible through the very act of parameters adopting these compensatory roles (unless adoption of those roles yields parameter values that are obviously erroneous). For these types of prediction, it is the visible expression of model structural defects that present the most serious imposition to characterization of their uncertainties. As was discussed herein, these visible expressions of model structural defects cannot be subjected to stochastic characterization as they are location-, time- and process-specific. Furthermore, their expression during predictive model deployment is likely to be different from their expression during model calibration.

The situation is different for model predictions that possess a high degree of null space dependence. These predictions pertain to locations that are different from those at which measurements comprising the calibration dataset were made, and/or pertain to a very different regional or local stress regime from that which prevailed during model calibration. For these types of predictions, nonuniqueness of parameter combinations that comprise the calibration null space dominates predictive uncertainty. The methodology described in chapter 2 of this appendix is able to characterize this component of predictive uncertainty reasonably well.

It must not be forgotten, however, that separation of solution and null spaces is somewhat artificial when the upscaled nature of parameters employed by a regional model is taken into account. Local process and hydraulic property details can be represented only in an approximate fashion by such a model. For example, where springs occur within the domain of the NFSEG model, a small number of parameters is used to represent a possibly high level of local heterogeneity which governs flow from the spring. In many cases, these upscaled parameters can be adjusted to provide a good fit with historical spring flow. Where this occurs, the roles that these upscaled parameters play in the future are likely to reflect their roles in the past; hence predictions of future spring flow can be made with relative certainty. However, if upscaled parameters cannot be adjusted in order to allow the model to reproduce measurements of historical spring flow, this may indicate a deficit of representation or salient local parameterization and/or structural details in the model. With the addition of more parameters, a better fit may indeed be achieved with the calibration dataset. At the same time the dimensionality of the null space is likely to be increased. Both of these may lead to better predictions of spring flows at the same location, and to a more reliable characterisation of its uncertainty, even if

some of the added parameters adopt surrogate roles to fit historical spring flow. If this is the case, it suggests that more parameters may need to be added to the NFSEG model on an as-needed basis in accordance with specific predictions that are required of it. (Note that the same considerations apply to baseflows and heads as those that apply to springs.)

These considerations should not impugn the performance of the NFSEG model in achieving its ambitious aims. Though course-gridded in relation to some of the features which may impact local groundwater behaviour, its grid is no coarser than that employed by other District models. At the same time, there can be little doubt that its regional nature has promulgated a reduction in sources of uncertainty that emerge from use of boundary conditions that are not actually groundwater flow boundaries. Numerical models of groundwater flow are always be imperfect, and are always compromised.

# 7. References

Doherty, J., 2014. PEST Groundwater Data Utilities. *Watermark Numerical Computing, Brisbane, Australia*. Downloadable from *www.pesthomepage.org*.

Doherty, J., 2015. Calibration and uncertainty analysis for complex environmental models. *Published by Watermark Numerical Computing, Brisbane, Australia*. 227pp. ISBN: 978-0-9943786-0-6. Downloadable from *www.pesthomepage.org*.

Doherty, J., 2016. PEST: Model-Independent Parameter Estimation. Version 13.7. *Watermark Numerical Computing, Brisbane, Australia*. Downloadable from *www.pesthomepage.org*.

Doherty, J. and Hunt, R.J., 2009. Two statistics for evaluating parameter identifiability and error reduction. *Journal of Hydrology*. 366, 119-127.

Doherty, J. and Welter, D., 2010, A short exploration of structural noise, *Water Resour. Res*., 46, W05525, doi:10.1029/2009WR008377.

Moore, C. and Doherty, J., 2005. The role of the calibration process in reducing model predictive error. *Water Resources Research*. Vol 41, No 5. W05050.

Vecchia, A.V. and Cooley, R.L., 1987. Simultaneous confidence and prediction intervals for nonlinear regression models with application to a groundwater flow model. *Water Resour Res*, 23 (7), 1237-1250.

White, J.T., Doherty, J.E. and Hughes, J.D., 2014. Quantifying the predictive consequences of model error with linear subspace analysis. *Water Resour. Res, 50 (2): 1152-1173.* DOI: 10.1002/2013WR014767