Hydrostratigraphy Modeling of the Southern Hills Aquifer System and Faults

Basic Information

<table>
<thead>
<tr>
<th>Title:</th>
<th>Hydrostratigraphy Modeling of the Southern Hills Aquifer System and Faults</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Number:</td>
<td>2012LA85B</td>
</tr>
<tr>
<td>Start Date:</td>
<td>3/1/2012</td>
</tr>
<tr>
<td>End Date:</td>
<td>2/28/2012</td>
</tr>
<tr>
<td>Funding Source:</td>
<td>104B</td>
</tr>
<tr>
<td>Congressional District:</td>
<td>6th</td>
</tr>
<tr>
<td>Research Category:</td>
<td>Ground-water Flow and Transport</td>
</tr>
<tr>
<td>Focus Category:</td>
<td>Groundwater, Solute Transport, Methods</td>
</tr>
<tr>
<td>Descriptors:</td>
<td></td>
</tr>
<tr>
<td>Principal Investigators:</td>
<td>Frank Tsai</td>
</tr>
</tbody>
</table>

Publications

3. Frank Tsai and Jeffrey Hanor, 2012, Hierarchical Multimodel Saltwater Intrusion Remediation and Sampling Designs: A BMA Tree Approach, Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Louisiana, 10 pages. (USGS 104G)
4. Frank Tsai, 2012, Feasibility Study of Scavenging Approach to Stop Saltwater Toward Water Wells, Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Louisiana, 10 pages. (USGS 104B)
Hydrostratigraphy Modeling of the Southern Hills Aquifer System and Faults

December 2012
Problem and Research Objectives

The Southern Hills aquifer system shown in Figure 1(a) was designated to be the sole source aquifer for southeastern Louisiana and southwestern Mississippi (Buono 1983; USEPA Region 4 & 6). The Southern Hills aquifer system also provides an essential amount of high quality groundwater for industries in Louisiana. In 2000, 290 million gallons per day of groundwater were withdraw from the Southern Hills aquifer system, of which 49% were used for public supply and 39% were used for industries (Sargent 2002). Figure 1(b) shows the 2009 groundwater withdrawals by 33 public suppliers and industries in East Baton Rouge Parish. Due to excessive groundwater withdrawal, many freshwater sands in the Southern Hills aquifer system north of the Baton Rouge fault are being contaminated by saltwater intrusion from brackish aquifers south of the fault (Tomaszewski 1996; Lovelace 2009).

Fig. 1: (a) The Southern Hills aquifer system (map source: USEPA Region 6). The yellow box is the current study area under saltwater intrusion, shown. (b) 2009 groundwater withdrawals by 33 public suppliers and industries from the aquifers in East Baton Rouge Parish (data source: CAGWCC).

Saltwater intrusion, recent drought (based on the U.S. Drought Monitoring), climate change and concerns of groundwater shortage/contamination by natural gas shale fracking prompts the development of a “Statewide Groundwater Management Plan” by the Louisiana Department of Natural Resources (DNR) starting in 2010 (Ecology and Environment, Inc. 2010). This Plan is especially intriguing to citizens in the Southern Hills aquifer system after the oil and gas industries expressed a strong interest in developing the Tuscaloosa Marine Shale (Welsh 2011), which covers almost the entire Florida Parishes shown in Figure 2. Recently, the DNR issued “Recommendations for a Statewide Ground Water Management Plan” (Ecology and Environment, Inc. 2011) that “requests the legislature to develop a program to fund the development of aquifer-wide groundwater availability models...” The Recommendation recommended Louisiana GWMTF (Ground Water Management Advisory Task Force) to study and identify (1) type and frequency of modeling suggested per area; (2) initial and annual maintenance cost to implement suggested modeling projects per area; and (3) sustainable funding sources for each project.
Developing reliable *groundwater availability models* needs to understand the complexity of an aquifer system. The proposal aims to develop a geostatistical hydrostratigraphy technique to understand the complexity of the Southern Hills aquifer system. The technique can also be applied to other aquifer systems.

**Fig. 2:** Extent of Tuscaloosa Marine Shale (John et al. 1997)

The goal of the project is to construct detailed hydrostratigraphy of the Southern Hills aquifer system. The scope of the project includes collection of geophysical logs, driller logs and well schedules from the DNR, USGS, LGS and CAGWCC, interpretation of well logs into sand-clay binary hydrostratigraphy, calibration of three-dimensional hydrostratigraphy, and identification of faults in the aquifer system.

**Objectives**
To achieve the project goal, we propose the following specific objectives:

**Objective 1** Delineate aquifer structure
The objective is to better understand hydraulic connection from outcrops to the south boundary of the Southern Hills aquifer system.

**Objective 2** Delineate fault structure
The objective is to calculate displacement of aquifers across faults and identify flow pathways that saltwater crosses the Baton Rouge fault.

**Methodology (from Elshall et al., 2013 Hydrogeology Journal)**

1. **Indicator generalization parameterization for hydrostratigraphy modeling**
   In this study, parameterization is conducted in the two-dimensional planar direction along the dip for every one-foot vertical interval. Three-dimensional aquifer-fault architecture is reconstructed by assembling all two-dimensional slices.
   This study utilizes a generalized parameterization (GP) method (Tsai and Yeh 2004; Tsai 2006), which combines the indicator kriging (IK) and indicator zonation (IZ) through a set of data weighting coefficients to obtain nonsmooth conditional estimates. The indicator function \( \{ I(x, v): x \in \text{study area} \} \) is a random function with the indicator random variable \( v \) describing the spatial extent of sand or clay facies. For a given sand-clay cutoff \( \alpha \), the random function of the indicator random variable \( v \) for sand facies is defined as

   \[
   I(x, v) = \begin{cases} 
   1 & v \in \text{Sand}, \quad v(x) \geq \alpha \\
   0 & v \notin \text{Sand}, \quad v(x) < \alpha 
   \end{cases}
   \]

   From equation (1) the indicator outcome (one or zero) indicates the presence of sand facies or clay facies, respectively. The indicator variogram has the same definition as the normal
variogram except that the real random function is replaced by the indicator random function $I(x, v)$. To calculate the expected value $v^*(x_0)$ at location $x_0$, the GP is

$$v^*(x_0) = I(x_0) + \sum_{i=1}^{N} \lambda_i [I(x_i) - I(x_0)] \beta_i$$

(2)

where $N$ is the number of electric well logs, $I(x_i)$ is the indicator data, $\lambda_i$ is the indicator kriging weight, and $\beta_i$ is the data weighting coefficient for a data point of a well log at location $x_i$. $I(x_k)$ is indicator data for a zone defined by well log $k$. Equation (2) shows that GP estimate at unknown location is similar to IK estimate $v^*(x_0) = \sum_{i=1}^{N} \lambda_i I(x_i)$ or the IZ estimate $v^*(x_0) = I(x_k)$ except for the introduction of $\beta_i$ such that $\forall \beta_i = 1$ gives the IK estimate, $\forall \beta_i = 0$ gives the IZ estimate and $0 < \beta_i < 1$ gives the in-between GP estimate.

For zonal delineation, this study uses two-dimensional Voronoi tessellation (Sibson 1980). This is a simple mathematical technique for dividing a space into a number of Voronoi zones, given a set of coplanar points, which are electric well logs data. A Voronoi zone, which is drawn based on bi-sectors for each data point, is a boundary enclosing all the intermediate space lying nearest to that data point than to other data points in the plane. The Voronoi tessellation is considered a neutral and unbiased approach to define the neighborhood of a data point (Tsai and Yeh 2004; Tsai 2006).

2. **Unknown model parameters**

The first unknown model parameter is the formation dip, which establishes data correlation. The Baton Rouge aquifer system gently dips south. Prior geological studies did not quantify the formation dip. The dip was calculated to be $0.30^\circ \pm 0.06^\circ$ from the USGS cross-sectional maps in the area (Griffith 2003) as prior information. To constrain the search space, the dip is set within the range $0.06^\circ \leq \phi \leq 0.60^\circ$. The vertical tolerance of the dip is not reported in any study, and no vertical tolerance is considered.

The second unknown model parameter is the sand-clay cutoff value $\alpha$. The estimated $v$ values, which cutoff value rounds to produce an indicator, could be viewed as the conditional probability with respect to the binary variables (Chiles and Delfiner 1999). The limits of the cutoff value $\alpha$ are constrained to a realistic but flexible range of $0.3 \leq \alpha \leq 0.7$. Other unknown model parameters are the data weighting coefficients $\beta$ of the well logs.

For reconstructing aquifer-fault architecture in the Baton Rouge area shown in Figure 3, the model used 288 geophysical well logs in which the south, middle and north domains have 61, 129 and 98 well logs, respectively, to reconstruct the hydrostratigraphy. When the spacing of well logs is dense, the estimates by the GP method in these areas become insensitive to the data weighting coefficients $\beta$, since the indicator kriging estimates and indicator zonation estimates are similar. To reduce the computational cost of the inverse problem, insensitive values of $\beta$ are identified through sensitivity analysis. Performing the sensitivity analysis starts by calibrating the hydrostratigraphy model only with respect to the dip and cutoff for given $\forall \beta = 1$ to all well logs,
and then using the calibrated model as a reference for the fitting error. Then taking one well log at a time, its data weighting coefficient is evaluated from 0 to 1 incrementally by 0.1 to calculate new fitting errors. Any well log that results in an error difference less than ±0.05% from the calibrated model is considered to have a fixed data weighting coefficient $\beta = 1$. The sensitivity analysis shows that 48 well logs have sensitive $\beta$ coefficients with their number in the south, middle and north domains being 6, 34 and 8 well logs, respectively.

**Fig. 3.** Map of the study area in the Universal Transverse Mercator (UTM) coordinate system. Black dots represent the location of electrical well logs. The bold solid lines are fault lines identified by the surface expression (McCullough and Heinrich 2012). The bold dashed lines are the approximate surface locations of the faults (Griffith 2003). The yellow areas are urban areas, the grey lines are parish borders, the red lines are interstate freeways, the green lines are US highways, and the blue areas and lines are water bodies.

### 3. Inverse problem

The data weighting coefficients $\beta$ along with the dip $\phi$ and sand-clay cutoff $\alpha$ are the unknown model parameters to be estimated using an inversion scheme. The inversion scheme for the IZ, IK and GP is the same expect for the size of the unknown parameters. The IZ inversion has only one unknown parameter that is the dip. The unknown parameters of the IK inversion are the dip and the cutoff. The unknown parameters of the GP inversion are the dip, the cutoff and the data weighting coefficients. The inverse problem is formulated by minimizing the mean squared error between the estimated and observed facies as follows:

$$
\min_{\phi, \alpha, \beta} \frac{1}{2} \left[ \frac{1}{M_{\text{sand}}} \sum_{i=1}^{M_{\text{sand}}} \left[ f_{\text{est}}^{i, \text{obs}}(x) - f_{\text{est}}^{i, \text{obs}}(x) \right]^2 + \frac{1}{M_{\text{clay}}} \sum_{i=1}^{M_{\text{clay}}} \left[ f_{\text{est}}^{i, \text{obs}}(x) - f_{\text{est}}^{i, \text{obs}}(x) \right]^2 \right]
$$

where $M_{\text{sand}}$ and $M_{\text{clay}}$ are the number of data points of the sand facies and clay facies, respectively. The $f_{\text{est}}^{i, \text{obs}}(x)$, $f_{\text{est}}^{i, \text{obs}}(x)$ and $f_{\text{est}}^{i, \text{obs}}(x)$ are the indicator estimate, the observed sand facies indicator and the observed clay facies indicator at a location $x$, respectively. The mean squared error is separated into two error terms with one for each facies to avoid calibration bias toward favoring the fitting of clay over sand since the well logs indicate a clay proportion of about two-thirds by volume within the study area. The proportion of sand calculated from the electric logs is 0.338 and is 0.339 from the drillers’ logs. This separation underlines that reducing the sand error is equally important as reducing the clay error.
To solve the inverse problem, the study adopts the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) (Hansen et al. 2003) as a local derivative-free optimization method for two reasons. First, unlike the derivative-based methods, using the CMA-ES allows for flexible optimization without prior assumptions or restrictions about the model structure. Second, the enhanced search properties of the CMA-ES allow for reaching near global solution. Similar to other generation-based optimization algorithms, the CMA-ES proposes several candidate solutions per search iteration. Each candidate solution is a vector of unknown model parameters, which the model uses to solve for state variables. Then the objective functions of all solutions are calculated and ranked. The CMA-ES adapts a covariance matrix representing the pair-wise dependency between unknown model parameters, which approximates the inverse of the Hessian matrix up to a certain factor. The covariance matrix adaptation uses information from the ranking of the current solutions and from the previous search path. Then the solutions are updated with the covariance matrix and an adaptable step size, which are adapted through two conjugates that implements heuristic control terms. These enhanced search properties allow the CMA-ES to handle ill-conditioned, nonsmooth, discontinuous, nonconvex and multimodal functions. Reviewing the CMA-ES algorithm is beyond the scope of this work, and reader is referred to Hansen (2006) and the references therein.

The inversion scheme steps are as follows. First, the CMA-ES generates candidate solutions, which are sets of unknown model parameters. Second, for each proposed solution the experimental variograms and a theoretical variogram are calculated based on the proposed dip. With respect to experimental variograms it is important to clarify one precaution with respect to location dependence of data correlation is accounted for. The correlation between the data across the faults is prevented, but all the experimental variograms of each domain are grouped together to calculate one theoretical variogram. The theoretical variograms is fitted to the experimental variograms automatically through using the pattern search method. It performs direct directional search for the correlation parameters, which are the nugget, sill and effective range, to minimize the squared root error between the experimental and the theoretical variograms. Third, interpolation function in equation (2) is used to estimate facies distribution at unknown locations. For the inversion purpose the unknown locations are the drillers’ logs locations. For the IZ inversion all the $\beta$ values are set to zero, and thus the cutoff is not needed. Contrariwise, for the IK inversion the $\beta$ values are set to one, thus the estimated facies is rounded to the indicator value by the cutoff. For the GP inversion $\beta$ values are used by the interpolation function in equation (2) to estimate facies distribution at the unknown locations and the cutoff is used to round the indicator. Fourth, the estimated facies are compared to the observed facies to calculate the mean squared error for individual solutions. Then, step 1 is repeated until the mean squared error is minimized.

The outcome of the inversion is the best unknown model parameters set that fits the observed facies. This parameters set (dip, cutoff and $\beta$ values) can be used to plot any 2-dimensional or 3-dimensional diagrams according to the desired grid size. For example, in this study all the cross sections of the faults have a grid size of 50 m along the fault lines. The 3-dimensional diagrams of the aquifer system have a grid size of 200 m in the X and Y directions. The discretization in Z direction is 1-foot (0.34-m) interval.

**Principal Findings and Significance (from Elshall et al., 2013 Hydrogeology Journal)**
1. **Calibration results**

This inversion scheme is used to obtain the optimal data weighting coefficients, dip, and cutoff for the hydrostratigraphy model. The calibration results are shown in Table 1. The variogram structure and cutoff are similar for the indicator zonation (IK), generalized parameterization (GP), and indicator kriging (IZ) methods. The three methods also show the same dip around 0.29° and the same sand proportion around 0.35. The GP shows less fitting error than the IK and IZ methods due to the flexibility of the method.

<table>
<thead>
<tr>
<th>Method</th>
<th>Nugget</th>
<th>Sill</th>
<th>Range [m]</th>
<th>Dip [Deg.]</th>
<th>Cutoff [-]</th>
<th>Sand Proportion</th>
<th>Sand Error[%]</th>
<th>Clay Error[%]</th>
<th>Total Error[%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>IZ</td>
<td>0.062</td>
<td>0.161</td>
<td>8400</td>
<td>0.276</td>
<td>-</td>
<td>0.340</td>
<td>13.02</td>
<td>12.79</td>
<td>12.91</td>
</tr>
<tr>
<td>GP</td>
<td>0.083</td>
<td>0.139</td>
<td>8400</td>
<td>0.289</td>
<td>0.404</td>
<td>0.347</td>
<td>11.96</td>
<td>12.90</td>
<td>12.43</td>
</tr>
<tr>
<td>IK</td>
<td>0.084</td>
<td>0.139</td>
<td>8600</td>
<td>0.286</td>
<td>0.404</td>
<td>0.347</td>
<td>12.04</td>
<td>12.96</td>
<td>12.50</td>
</tr>
</tbody>
</table>

2. **Fault architecture**

To show the differences between the three methods, the architecture of the Denham Springs-Scotlandville fault and the Baton Rouge fault are used as examples. Figure 4 show the juxtaposition at the fault cross sections using the generalized parameterization method. Black areas are clay units north of the fault. Gray areas are clay units south of the fault. Sand units are transparent to show potential hydraulic connections through the fault. It is noted that the faults are three-dimensional zones of deformation, not two-dimensional planes. Determination of permeability of the fault zone is suggested by Bense and Person (2006) and Hanor et al. (2011). Nevertheless, the detailed architecture of the fault zone is not the scope of this study.
Fig. 4. (a) the architecture of the Denham Springs-Scotlandville fault in the modeling area using generalized parameterization, and (b) the architecture of the Baton Rouge fault in the modeling area using generalized parameterization. Black areas are clay units north of the fault. Gray areas are clay units south of the fault. Sand units are transparent to show potential hydraulic connections through the fault. The fault cross sections are based on 3D estimates that follow the UTM_X and UTM_Y coordinates of the fault line in Figure 3. Elevation is feet above NGVD29.

Fig. 5. Hydrostratigraphic architecture of the “1,200-foot” sand to the “2,000-foot” sand in the middle domain of the modeling area in Figure 3. The wells in the figure are public supply and industrial wells.

3. Aquifer architecture
Figure 5 shows the simulated aquifer architecture from the “1,200-foot” sand to the “2,000-foot” sand in the middle domain based on the GP method. The “1,200-foot” sand connects vertically to the “1,500-foot” sand. The “2,000-foot” sand is clearly separated from the “1,700-foot” sand by a confining layer. There are four sand units between the “1,200-foot” sand and the “2,000-foot” sand, which are generally classified as the “1,500-foot” sand and the “1,700-foot” sand (Griffith 2003). Unlike the distinguishable “1,200-foot” sand and the “2,000-foot” sand, the separation of the “1,500-foot” sand from the “1,700-foot” sand is not well-defined in the published cross sections (Rollo 1969; Griffith 2003). The findings of this study also show that they are not clearly separable. Therefore, in this study the “1,500-foot” sand and the “1,700-foot” sand are together treated as a single unit.

4. Conclusion
The study finds strong hydraulic connection between the “1,200-foot” sand and the “1,500-foot” sand. Merger of the sand units indicates groundwater recharge from the “1,200-foot” sand to the “1,500-foot” sand. However, there is a distinct clay confining layer to separate the “2,000-foot” sand from the “1,700-foot” sand. The hydrostratigraphy also reveals four sand deposits that compose the “1,500-foot” sand and the “1,700-foot” sand. In general, sand deposition is not uniform, due to spatial and temporal variations in fluvial processes (Chamberlain 2012). The study shows that there is large amount of missing sand in “1,500-foot” sand in the industrial district and in West Baton Rouge Parish, which is possibly due to the presence of an erosional unconformity (Chamberlain 2012).

The sand unit displacement on the Baton Rouge fault and the Denham Springs-Scotlandville fault is significant. The Baton Rouge fault has higher sand displacement than the Denham Springs-Scotlandville fault. Displacement increases over depth. Due to non-uniform fault throw and sand deposition, the study reveals non-uniform flow pathways that connect different sand units at the fault planes. In particular, the identified flow pathways through the Baton Rouge
fault provide important information for understanding patterns of salinization of freshwater aquifers in the East Baton Rouge Parish.

References


6. Student Support
- Ehsan Beigi, doctoral student (starting spring 2011)
- Nima Chitsazan, doctoral student (starting fall 2010)
- Ahmed Elshall, doctoral student (starting fall 2010)