SYNOPSIS

**Title:** Saltwater Intrusion Management with Conjunctive Use of Surface Water and Ground Water

# Problem and Research Objectives

Saltwater intrusion in coastal aquifers is one of the major issues in coastal water resources management. The encroachment of saltwater from the sea floor is triggered by natural hydrologic processes and human-built environments. Seawater always intrudes geological formations due to the fact that seawater has slightly higher density and much higher dissolved salt concentration than freshwater. However, severe saltwater intrusion is mainly caused by the combination of droughts and excessive groundwater withdrawals. Once saltwater has invaded an aquifer, it could take significant time and cost to regain the virgin aquifer. Effective coastal saltwater intrusion management plans need the better understanding of saltwater intrusion mechanism and development of flow and transport simulation models as a decision-making tool.

In this study, we focus on two research objectives for the saltwater intrusion problem. The first research objective is to simulate saltwater intrusion in coastal aquifers using a lattice Boltzmann method. The saltwater intrusion phenomenon is described by the density-viscosity-dependent groundwater flow and mass transport equations. Our focus is on the understanding of the similarities between the lattice Boltzmann model (LBM) and the macroscopic saltwater intrusion model such that the macroscopic aquifer parameters, e.g., dispersion coefficient and hydraulic conductivity, can be properly represented by the LBM parameters. One of the challenges of using LBM is to cope with the spatial-temporal heterogeneity when particle distribution functions stream to neighboring lattice nodes. We will use the Henry problem to demonstrate the capability of our LBM to solve the saltwater intrusion in the heterogeneous aquifer.

Another challenge in real-world saltwater intrusion problems is the parameter heterogeneity estimation problem when the parameterization method is non-unique and inflexible. Therefore, the second research objective is to develop the maximum weighted log-likelihood estimation (MWLLE) and Bayesian model averaging (BMA) along with the generalized parameterization (GP) method (Tsai and Yeh, 2004; Tsai 2006) to cope with this problem in hydraulic conductive estimation. We will apply the MWLLE and BMA to a real-world case study to estimate the hydraulic conductivity in the Alamitos Gap area, California, where the Alamitos Barrier Project (ABP) has been operated for more than forty years to protect freshwater aquifers from saltwater intrusion.

# Methodology
## 1. Density-Viscosity-Dependent Saltwater Intrusion Model
The groundwater flow equation with changes in water density and viscosity due to the presence of the dissolved salt has been formulated in terms of the freshwater pressure head (Huyakorn et al., 1987; Boufadel et al., 1999; Simpson and Clement, 2003). Using fresh groundwater head in the groundwater flow equation was suggested to improve the numerical efficiency for the case that large static pressures dominate the dynamic pressure differences (Frind, 1982). In this study, we have derived the density-viscosity-dependent groundwater flow equation in terms of fresh groundwater head:

$$\phi S_{s_f} \frac{\partial h_f}{\partial t} + n \frac{\partial \phi}{\partial C} \frac{\partial C}{\partial t} = \left[ \frac{\partial}{\partial x} \left( \frac{\phi}{\lambda} K_f \frac{\partial h_f}{\partial x} \right) + \frac{\partial}{\partial z} \left( \frac{\phi}{\lambda} K_f \frac{\partial h_f}{\partial z} \right) + \frac{\partial}{\partial z} \left( \frac{\phi(\phi-1)}{\lambda} K_f \right) \right] + \frac{\rho_{ss}}{\rho_f} Q_{ss} \qquad (1)$$

where $S_{s_f}$ is the freshwater specific storage, $h_f$ is the fresh groundwater head, $n$ is the porosity, $K_f$ is the freshwater hydraulic conductivity, $\rho_{ss}$ is the water density at the sinks/sources; and $Q_{ss}$ is the flow rate per unit aquifer volume at the sinks/sources. $\phi = \rho/\rho_f$ is the ratio of fluid density to freshwater density. $\lambda = \mu/\mu_f$ is the fluid dynamic viscosity to the freshwater dynamic viscosity.

In general, the dissolved salt is considered as a conservative solute, which usually has very small to zero sorption and chemical reaction in the formation environment. Therefore, the salt transport is described by the advection-dispersion equation (ADE).

$$\frac{\partial nC}{\partial t} + \nabla \cdot (n\mathbf{u}C) = \nabla \cdot (nD\nabla C) + C_{ss}Q_{ss} \qquad (2)$$

where $\mathbf{u}$ is the average pore velocity vector, $D$ is the dispersion coefficient; and $C_{ss}$ is the salinity at the sinks/sources. We recognize that the dispersion coefficient depends on both anisotropy and flow velocity (Scheidegger, 1961; Bear, 1972). However, our current focus is on the scalar dispersion coefficient in LBM. Using a constant dispersion coefficient to study the saltwater intrusion will not destroy essential features of the problem (Henry, 1964; Pinder and Cooper, 1970; Lee and Cheng, 1974).

## 2. Lattice Boltzmann Model (LBM)

The discrete lattice Boltzmann model with the Bhatnagar-Gross-Krook (BGK) collision model has been introduced by Bhatnagar et al. (1954):

$$f_i' - f_i = -\frac{1}{\tau}(f_i - f_i^{eq}) + \frac{F}{N}\Delta t \qquad (3)$$

where $f_i' = f(\mathbf{x} + \mathbf{c}_i \Delta t, t + \Delta t)$, $i = 1, 2, \cdots, N$ are the particle distribution functions after the collision step; $i$ represents the discretized direction; $N$ is the number of lattice directions; $f_i(\mathbf{x}, t)$ are the particle distribution functions after the streaming step; $f_i^{eq}$ are the equilibrium distribution functions (EDFs); $\tau$ is the relaxation parameter; $F$ is the forcing term that represents the sinks/sources in the macroscopic equation, which is invariant of lattice directions; and $\Delta t$ is the lattice time step. The lattice speed is defined as $c = \Delta x / \Delta t$, where $\Delta x$ is the lattice spacing. To solve the ADE (Eq.(2)), D2Q9 EDFs are used (Chen and Doolen, 1998). Using LBM to solve the density-viscosity-dependent groundwater flow equation (Eq.(1)) has lesser numerical instability than in the ADE because the groundwater flow equation principally is a diffusion equation. Therefore, D2Q5 EDFs are sufficient for solving the groundwater flow equation with less computation demand. To cope with the density-viscosity variation in space and time and hydraulic conductivity heterogeneity, in each lattice time step we need to modify the speed of sound in the EDFs in order to take into consideration the heterogeneity effect when the particle distribution functions stream to their neighboring lattice nodes. We have found the equivalent squared speed of sound along each lattice direction to cope with the heterogeneity problem. The new EDFs for each lattice direction can be obtained.

## 3. The Maximum Weighted Log-Likelihood Estimation (MWLLE)

To increase flexibility of a conditional parameterization method in hydraulic conductivity estimation, Tsai and Yeh (2004) and Tsai (2006) have developed a generalized parameterization (GP) method, which is able to conditionally estimate a non-smooth random field. However, due to limited data, there may be many zonation structures and interpolation methods that are equally important according to the measured data. Combinations of these zonation and interpolation methods will result in many possible GP methods, which should be taken into consideration simultaneously in the aquifer parameter estimation and groundwater modeling. To estimate the data weighting coefficients, $\boldsymbol{\beta}$ (Tsai 2006), among the multiple GP methods, this study proposes the weighted log-likelihood (WLL), which combines the log-likelihood functions through the weight of each GP method:

$$\ln L_w\left(\boldsymbol{\beta} \mid \mathbf{u}^{obs}\right) = \sum_{i=1}^{M} W_i \ln L\left(\boldsymbol{\beta} \mid \mathbf{u}^{obs}, \theta^{(i)}\right) \tag{4}$$

where $\ln L_w\left(\boldsymbol{\beta} \mid \mathbf{u}^{obs}\right)$ is the weighted log-likelihood function of the data weighting coefficients given groundwater head observations $\mathbf{u}^{obs}$; $\ln L\left(\boldsymbol{\beta} \mid \mathbf{u}^{obs}, \theta^{(i)}\right)$ is the log-likelihood function of the data weighted coefficients given groundwater head observations and a GP method $\theta^{(i)}$; $W_i$ is the GP method weight, which relates to the selected GP methods and data; and $M$ is the number of the selected GP methods. The sum of the weights is $\sum_{i=1}^{M} W = 1$.

The parsimony principle for the GP weight leads us to consider the posterior probability of a GP method conditioned on the observed groundwater head data, i.e., $W_i = \Pr\left(\theta^{(i)} \mid \mathbf{u}^{obs}\right)$, which can be calculated in terms of the Akaike information criterion, (AIC), Bayesian information criterion (BIC), Kashyap information criterion (KIC), etc. We consider the BIC in this study. The traditional Bayesian weights, especially in the real-world case study, tends to single out the best GP method and overkill other good GP methods because the GP method weights exponentially decease with $\frac{1}{2}\Delta\text{BIC}^{(i)}$, where $\Delta\text{BIC}^{(i)} = \text{BIC}^{(i)} - \text{BIC}_{\min}$ and $\text{BIC}_{\min}$ is the minimum BIC value among the GP methods, the traditional. A straightforward way to overcome this problem is to consider a scaled likelihood function for $\Pr\left(\mathbf{u}^{obs} \mid \boldsymbol{\beta}, \theta^{(i)}\right)$ such that a scaled Bayesian information criterion (SBIC) is resulted

$$\text{SBIC}^{(i)} = \alpha\text{BIC}^{(i)} \tag{5}$$

where $\alpha$ is a scaling factor. We choose $\alpha = 3/\sqrt{L}$, where $L$ is the number of head observations. Therefore, the GP weights are determined by the following

$$W_i = \exp\left(-\tfrac{3}{2\sqrt{L}}\Delta\text{BIC}^{(i)}\right) \Big/ \sum_{i=1}^{M} \exp\left(-\tfrac{3}{2\sqrt{L}}\Delta\text{BIC}^{(i)}\right) \tag{6}$$

Substituting Eq.(6) into Eq.(4), the MWLLE becomes

$$\min_{0 \leq \boldsymbol{\beta} \leq 1} \ -\ln L_w\left(\boldsymbol{\beta} \mid \mathbf{u}^{obs}\right) = -\sum_{i=1}^{M} \exp\left(-\tfrac{3}{2\sqrt{L}}\Delta\text{BIC}^{(i)}\right) \ln L\left(\boldsymbol{\beta} \mid \mathbf{u}^{obs}, \theta^{(i)}\right) \Big/ \sum_{i=1}^{M} \exp\left(-\tfrac{3}{2\sqrt{L}}\Delta\text{BIC}^{(i)}\right) \tag{7}$$

Once the optimal data weighting coefficients are obtained, the GP weights are also determined. Through the Bayesian model averaging (BMA) (Draper, 1995; Hoeting et al., 1999), the conditional mean and conditional covariance of the estimated hydraulic conductivity using multiple GP methods can be obtained via the BMA approach:

$$\text{E}\left[\boldsymbol{\pi} \mid \boldsymbol{\pi}^{data}\right] = \overline{\boldsymbol{\pi}}_{GP} = \sum_{i=1}^{M} W_i \boldsymbol{\pi}_{GP}^{(i)} \tag{8}$$

$$\text{Cov}\left[\boldsymbol{\pi} \mid \boldsymbol{\pi}^{data}\right] = \sum_{i=1}^{M} W_i \left[\text{Cov}_{GP}^{(i)} + \left(\boldsymbol{\pi}_{GP}^{(i)} - \overline{\boldsymbol{\pi}}_{GP}\right)\left(\boldsymbol{\pi}_{GP}^{(i)} - \overline{\boldsymbol{\pi}}_{GP}\right)^{T}\right] \tag{9}$$

where $\pi$ is the log hydraulic conductivity value. The first term in the right side of Eq.(9) is the within-GP covariance and the second term represents the between-GP covariance. The conditional estimation $\boldsymbol{\pi}_{GP}^{(i)}$ and conditional covariance $\text{Cov}_{GP}^{(i)}$ for each GP method have been derived by Tsai (2006).

## Principal Findings and Significance
### 1. Saltwater Intrusion Modeling Using Lattice Boltzmann Method
### 1.1 The Henry Problem
The Henry problem (Henry, 1964) is one of the benchmark problems for validating the density-dependent groundwater flow and mass transport models, especially for the saltwater intrusion problem in coastal aquifers. The parameter values for the Henry problem are listed in Table 1.

**Table 1:** Parameter Values for the Henry Problem.

| Parameters | Value |
|---|---|
| $D$: dispersion coefficient, [m$^2$/sec] | $1.886 \times 10^{-5}$ |
| $K_f$: freshwater hydraulic conductivity, [m/sec] | 0.01 |
| $Q_{in}$: inflow flux, [m$^3$/sec-m] | $6.6 \times 10^{-5}$ |
| $n$: porosity, [-] | 0.35 |
| $\rho_f$: freshwater density, [kg/m$^3$] | 1000 |
| $\rho_s$: seawater density, [kg/m$^3$] | 1025 |
| $C_s$: seawater concentration, [kg/m$^3$] | 35 |

Considering the constant concentration of salt at the seaside, Figure 1(a) shows the LBM results of the Henry problem against the Henry analytical solution revisited by Segol (1994). The 50% isochlor is almost exactly on the analytical solution. Although not shown here, the 25%, 50% and 75% isochlors agree with the Henry analytical solution revisited by Simpson and Clement (2004).

In Figure 1b, the flow field and the fresh groundwater head distribution demonstrate the seawater circulation from the sea floor (Cooper, 1964). The saltwater circulation is characterized by the interface of zero horizontal velocity (solid line) in Figure 1b. The area below the interface represents the landward flow zone, where the water is coming into the aquifer from the seaside. The area above the interface represents the seaward flow zone, where the water flows out of the aquifer. The outflow region at the seaside boundary is $1 \leq z \leq 0.43$. It is noted in Figure 1b that the fresh groundwater equipotential lines are not perpendicular to the top and bottom no-flow boundaries because of the density variation. The salt groundwater equipotential lines are orthogonal to the impermeable boundaries.
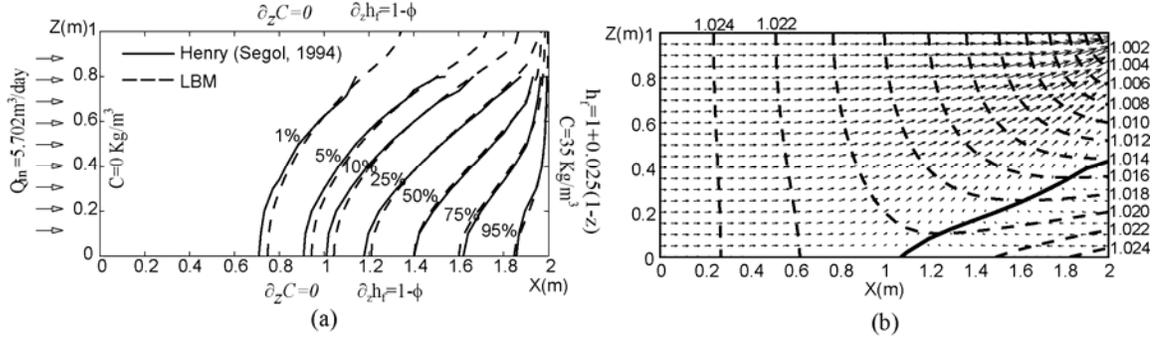
**Figure 1:** (a) The isochlor distribution. (b) The fresh groundwater head distribution and flow field.

## 1.2 Saltwater Intrusion in Heterogeneous Aquifer

Based on our literature review, we don't find any studies using the lattice Boltzmann method to simulate saltwater intrusion in the heterogeneous hydraulic conductivity (K) field. To demonstrate our LBM capability of handling the heterogeneity problem, we consider one correlated K field and one uncorrelated K fields as shown in Figure 2 to test the LBM. The mean of $\log_{10}K$ is -2. The unconditional standard deviation is 0.5m/s. The integral scale along x direction is 0.5 m and along z direction is 0.1m for correlated K. The uncorrelated K field has zero integral scale.
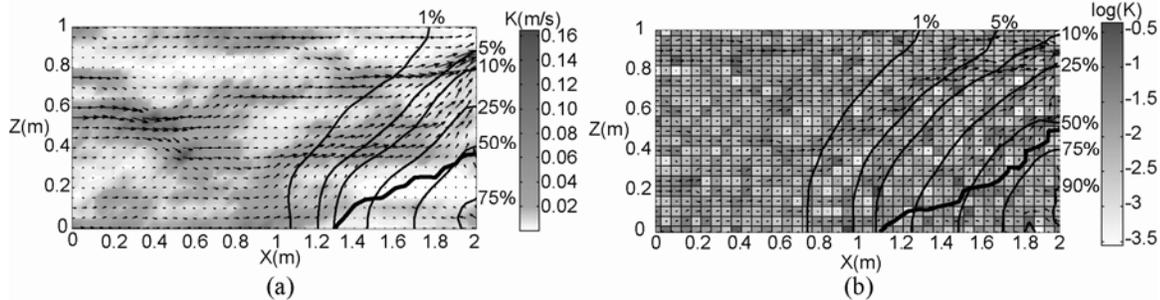


**Figure 2:** The isochlor distribution and flow field with (a) the correlated heterogeneous K field and (b) the uncorrelated heterogeneous K field.

The parameter values in Table 1 are also used for the heterogeneous K case. We consider the mixed Neumann-Cauchy boundary condition at the seaside. Less saltwater intrusion is observed in Figure 2a in comparison with the homogeneous case. High flow velocities are also observed at the high K areas. The isochlors in Figure 2b are very close to those for the homogeneous aquifer (not shown here). This indicates that completely random heterogeneity does not significantly change the scale of the saltwater intrusion from that predicted using the mean K value. However, the correlated K field has a significant impact on the saltwater intrusion result, which is quite different from that obtained by the mean K field.

## 2. Case Study: Alamitos Barrier Project, Southern California

Long-term overproduction of groundwater from the Coastal Plain aquifer in Southern California has significantly lowered the groundwater surface below sea level in extensive areas. The landward gradient from the ocean to these human-built pumping depressions has developed a condition wherein seawater has intruded into the aquifer system which is in hydraulic continuity with Pacific Ocean (Callison et al., 1991). One of the saltwater intrusion remediation actions has

been taken to protect aquifers from saltwater intrusion is the development of regional-scale freshwater barriers, which create local hydraulic ridges along the coastal line via injecting freshwater into aquifers through a series of freshwater injection wells. The Alamitos Barrier Project (ABP) is one of three major freshwater barriers in Southern California, which was constructed in 1964 and has being operated since 1966 to protect the groundwater supplies of the central basin of Los Angeles County and southwest portion of the Coastal Plain area in Orange County from the intrusion of seawater through the Alamitos gap area (Callison et al., 1991).

Groundwater flow simulation is important in order to improve the performance of the barrier operations and better the groundwater management in the Alamitos Gap area, which has 5 major aquifers, R, C, B, A, and I zones overlaying each other in this order. In collaboration with Los Angeles County Department of Public Works (LACDPW), the groundwater model is developed using the 566 groundwater head observation data from 56 head observation wells in the Alamitos Gap area and injection record for 37 injection wells from 1992 to 2002. Location of the groundwater head boreholes and injection wells are shown in Figure 3a. Figure 3a also shows the complexity of I zone. Several places in I zone are missing or merging with other aquifers. The missing and merging areas are interpreted from the log data (Callison et al., 1991). The 148 logs shown in Figure 3b determine the top and bottom elevations of I zone and hydraulic conductivity values at the log sites (Callison et al., 1991). The Seal Beach Fault forms a substantial barrier to the movement of groundwater into or out of the Central Basin for I zone. However, groundwater in I zone does flow in and out of the Central Basin through the erosion gaps in the Recent aquifer (Callison et al., 1991).
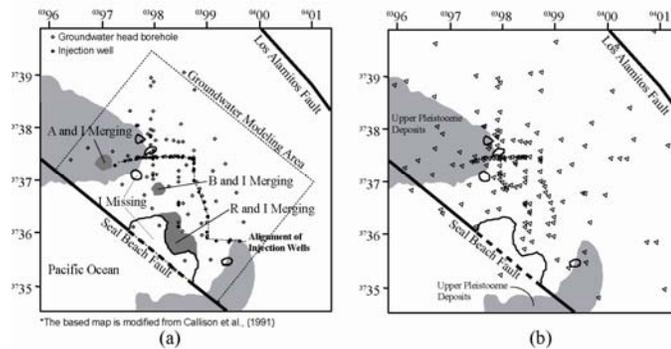


**Figure 3:** (a) The study area (I zone) in the Alamitos Gap area. (b) The log sites where top and bottom elevations and hydraulic conductivity values are available in I zone.

## 2.1 Groundwater Modeling and Parameterization

We adopt MODFLOW-2000 (Harbaugh et al., 2000) for groundwater flow simulation from July 1992 to July 2002 in I zone. Currently, we don't consider saltwater transport in this study. The hydraulic conductivity is considered to be log-normally distributed. The time-varied constant-head boundary conditions are given to the boundaries of the study area as well as the aquifer mergent areas. The 148 hydraulic conductivity values (Figure 3b) show a secondary-order stationary K field. An exponential semivariogram model $\gamma(d) = 0.3257(1 - \text{Exp}(-d/649.7134))$ is obtained, where $d$ is the distance lag. We choose Voronoi tessellation (VT) as the zonation method (Tsai et al., 2003) and choose three interpolation methods, the natural neighbor interpolation (NN) method (Sibson, 1981; Tsai et al., 2005), inverse distance squared interpolation (ID) method (Watson and Philip, 1985; Gotway et al., 1996, and ordinary kriging

(OK) methods (Olea, 1999). Combination of the zonation and interpolation methods results in three GP methods, NN-VT, ID-VT, and OK-VT.

## 2.2 Data Weighting Coefficient Identification in MWLLE

The three GP methods (NN-VT, ID-VT, and OK-VT) are considered in MWLLE. The individual zonation and interpolation methods are not considered in MWLLE because they are a subset of the GP methods. The groundwater head variances are estimated as the mean of the groundwater head variances from the zonation and interpolation methods. Three GP methods use the same data weighting coefficients in this study. We use the combination of a gradient-based method and a local search method to identify the optimal data weighting coefficients. We adopt a BFGS solver (Byrd et al., 1994) to solve MWLLE. The local search method is use to improve the BFGS solution on one data weighting coefficient at a time. Moreover, the adjoint-state method was used to calculate the gradients and tremendously reduces the computation time. In each optimization step, we only need to run three times the groundwater flow equation and three times the adjoint-state equation due to three GP methods.

The optimal data weighting coefficient values and their locations over the study area for the unscaled case ($\alpha = 1$) and scaled case $\alpha = 3/\sqrt{566}$ show no distinct pattern for the distribution of the weighting coefficient values in both cases. However, some areas do show clustering weighting coefficients with value close to 1 or close to zero. The GP method shows its advantage to produce a non-smooth distribution of hydraulic conductivity. Almost one third of the $\beta$ values are 0 in both cases, which will make the estimated hydraulic conductivity distribution preserve the feature of zonation distribution.

For the unscaled case in Table 2, the small difference between the maximum BIC=3764.75 at ID-VT and the minimum BIC=3755.39 at NN-VT results in the dominant GP weight $W_{NN-VT} = 89.0\%$ for NN-VT. Even though the BIC of the OK-VT method is very close to that of the NN-VT method, the GP weight for OK-VT is only 1.2%. Again, using unscaled BIC may overkill good parameterization methods. For example, both ID-VT and OK-VT have small conditional uncertainty and misfit values with respect to NN-VT, but their weights to the hydraulic conductivity estimation are extremely small, which is not logically reasonable.

The scaled case in Table 2 has similar fitting residuals, where NN-VT has the minimum BIC=3756.95 and ID-VT has the maximum BIC=3759.73. The misfit values and the conditional uncertainty in both cases are close to each other. With the scaling factor value $\alpha = 3/\sqrt{566}$, the reasonable GP weights are obtained around one-third for each GP method. The misfit values for MWLLE are obtained using the weighted hydraulic conductivity distribution, which in both cases are very close to NN-VT, the best GP method in this study. The conditional uncertainty for MWLLE in the scaled case is smaller than that in the unscaled case because the ID-VT and OK-VT have smaller conditional uncertainty and similar weights to the NN-VT. Nevertheless, the NN-VT has the height weight 35.70%.

**Table 2:** Identification Results of the Unscaled and Scaled Cases.

| | NN-VT | ID-VT | OK-VT | MWLLE |
|---|---|---|---|---|
| $\alpha = 1$ (unscaled) | | | | |

| | | | | |
|---|---|---|---|---|
| BIC | 3755.39 | 3764.75 | 3759.72 | |
| ΔBIC | 0 | 9.36 | 4.33 | |
| $W_i$ | 89.0% | 0.8% | 1.2% | |
| tr(Cov), uncertainty | 2013 | 1794 | 1823 | 1997 |
| $\alpha = 3/\sqrt{566}$ (scaled) | | | | |
| SBIC | 3756.95 | 3759.73 | 3757.67 | |
| ΔSBIC | 0 | 2.78 | 0.71 | |
| $W_i$ | 35.70% | 30.06% | 34.24% | |
| tr(Cov), uncertainty | 2016 | 1806 | 1818 | 1903 |

## 2.3 Hydraulic Conductivity Estimation and Uncertainty

The weighted conditional hydraulic conductivity estimations obtained by MWLLE are shown in Figure 4. Although the reasonable GP weights are calculated in the scaled case, the difference between the hydraulic conductivity distributions obtained by the unscaled and scaled cases is visually insignificant. This is expected because the similar hydraulic conductivity distributions obtained by individual GP methods will gives similar hydraulic conductivity distributions under different GP weights. However, the significance of the GP weights will be revealed on the conditional covariances and estimation uncertainty in MWLLE, which will distinguish the conditional simulation (CS) results in the groundwater modeling when different GP weights are considered.



(a)Unscaled Case    (b)Scaled Case
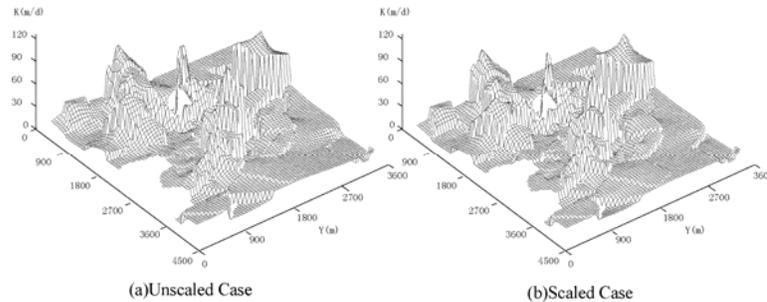
**Figure 4:** The estimated hydraulic conductivity distributions by MWLLE for (a) the unscaled case and (b) the scaled case.

Using the Bayesian model averaging (BMA) for the scaled case, the within-GP variance (Figure 5a), between-GP variance (Figure 5b), and the total variance (Figure 5c) are obtained for conditional simulation on hydraulic conductivity. The between-GP variance is much smaller than the within-GP variance in this case because the similar hydraulic conduction distributions are obtained by the three GP methods.
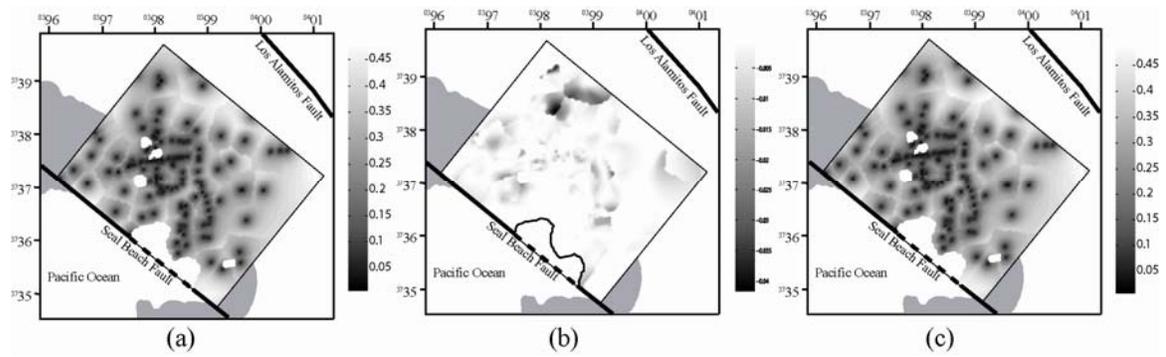
**Figure 5:** Conditional variance distributions (a) within-GP variance, (b) between-GP variance, and (c) total variance.