Use of Image Analysis to Develop New Benchmarking Datasets for Variable Density Flow Scenarios

R. R. Goswami[1,2] and *T. P. Clement*[1]
[1]Department of Civil Engineering, Auburn University, Auburn, AL, USA
[2]Geosyntec Consultants, Boca Raton, FL, USA

ABSTRACT
Image analysis (IA) techniques are used to measure properties such as concentration and water content in porous media experiments. Measurements made by IA techniques provide extensive spatial information on a temporal scale. However, a robust analysis of the accuracy of IA techniques has not been presented in previous publications. Therefore, results from experiments using IA techniques are not widely accepted as benchmarking datasets. The conventional method for testing the accuracy of IA techniques is the computation of global mass balance error. We demonstrate the limitations of quantifying IA errors based on this conventional method. We also introduce an alternate statistics-based method for estimating errors. The entire discussion is presented using a theoretical test problem. We also present the results of a physical laboratory experiment processed using the IA technique. The dataset is from a sinking plume experiment simulated by injecting saltwater into a freshwater aquifer. The experimental results are also compared with numerical modeling results generated using SEAWAT.

INTRODUCTION
Experimental studies have been conducted by using both qualitative (Goswami and Clement 2007) and quantitative (MRI, Gamma-radiation) methods to analyze porous media systems. One of the rapidly evolving non-invasive quantitative techniques is image analysis (IA). IA records an optical image property (such as pixel intensity, hue, color, etc.) and relates it to a system property such as concentration or water content. All IA techniques have the following two major parts: (i) acquiring and processing the digital images, and (ii) selecting a calibration-relationship between the intended system property and the image property (e.g., concentration-intensity relationship). A variety of errors may be introduced while acquiring and processing digital images. The most common source of error is non-uniform lighting. There may also be errors in image acquisition and processing due to hardware and software problems (Hansen et al. 2006). Since the calibration-relationship is not known, selection of the best-fit function for correlating the image property with the system property would also involve errors. We propose to categorize all possible errors into the following two categories:

1. Calibration-Relationship Error (CRE): error in the approximate calibration-relationship
2. Experimental Error (EE): error in the image property (e.g., intensity) due to various experimental factors such as non-uniform lighting, hardware problems, etc.

In the published literature, researchers have primarily used comparison of global mass balance errors and dispersion coefficients to test the accuracy of the results. Calculation of mass balance requires use of the entire range of measured values; therefore, they cannot be used to quantify local errors.

The main body of this paper is organized into three major sections. In the first section, we generate a theoretical test problem and use it to demonstrate the limitations of the commonly used mass-balance method to measure errors. In the second section, a statistics-based error estimation method is proposed. In the third section, we provide and compare results from our physical experiment and the application of IA technique and numerical modeling.

THEORETICAL TEST PROBLEM

The theoretical test problem was generated by considering diffusion of a point source in a two-dimensional domain. The following analytical solution was used to simulate the concentration profiles (Fischer et al. 1979):

$$C(x,z) = \frac{M}{4\pi t\sqrt{D_x D_z}} \exp\left(-\frac{x^2}{4D_x t} - \frac{z^2}{4D_z t}\right) \quad (1)$$

where $C(x,z)$ is the concentration at a given point (x,z) in the domain, t is the time from the start of the point-injection, M is the mass injected in the domain, and D_x and D_z are the diffusion coefficients in x and z directions respectively. The predicted plume concentration distribution is given in Figure 1. The analytical concentrations obtained were converted into pixel intensity values using the following concentration-intensity relationship:

$$I_i = A\left(1 - \exp\left(-BC_i\right)\right) \quad (2)$$

where I_i is the pixel intensity corresponding to a given concentration value C_i, The values of parameters A and B were selected based on empirical criterion (Goswami et al. 2008).

Introducing Experimental Error

Experimental error (EE) is associated with randomly distributed errors in the acquired image property which can be simulated by using the MATLAB® function *randn* (Hansen et al. 2006). The simulated noise was added to the intensity profile corresponding to the theoretical concentration data, thereby generating a '*noisy intensity*' profile. We then used the true concentration-intensity relationship (Equation 2) to estimate concentrations that will have some experimental error (EE). The concentrations obtained from this noisy dataset will be identified as $C_{estimate,\,EE}$.

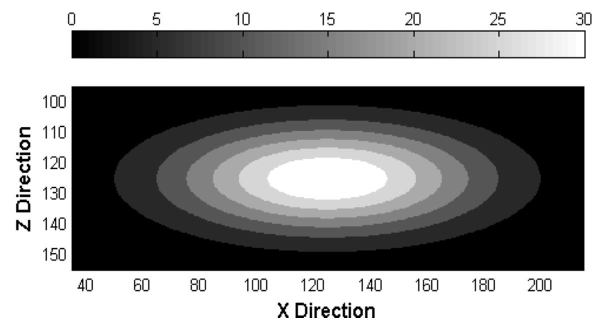

Figure 1. Concentration contours for the true solution (theoretical problem)

Limitations of Using Mass Balance Method

The estimated concentration dataset obtained after introducing experimental errors ($C_{estimate,EE}$) was processed to calculate the values of mass balance error. The total mass, calculated by determining the zeroth spatial moment of the concentration distribution, was found to be very low (<1%). As is commonly done in literature, we assumed the percentage error in mass balance as a quantitative estimate of error in the computed concentration values. Values of two concentration contours [5 mg/l (low) and 20 mg/l (high)] obtained from $C_{estimate,EE}$ were compared with the C_{true} values (obtained using Equation 1). The mass balance error was used to compute the error bounds using the relationship:

$$C_{error\,bound} = C_{true}\left(1 \pm \%\,Mass\ Balance\ Error \times C_{true}\right) \quad (3)$$

Figure 2a presents the 5 mg/L and 20 mg/L concentration contours obtained from the noisy dataset. It can be seen that the $C_{estimate,EE}$ values have speckle-type noise distributed around C_{true} values. The figure also has error bounds estimated using Equation 3. Since the error bounds do not predict the true contours, it can be deduced that mass balance error is not a good indicator of error in $C_{estimate,EE}$. We also employed a trial-and-error approach to generate two contours that will bound the observed spread around the predicted 5 mg/L and 20 mg/L contours. We estimated that a value of ±10% error will bound the 20 mg/l contour (22 mg/l and 18 mg/l) and ±15% error will bound the 5 mg/l contour (4.25 mg/l and 5.75 mg/l). These 'trial and error' bounds are shown in Figure 2b along with $C_{estimate, EE}$ and C_{true}. It can be observed that these trial-and-error contours closely bound C_{true}. Therefore it can be concluded that the error associated with 5 mg/L and 20 mg/L contours are approximately ±15% and ±10%, respectively. These values are considerably higher than the mass balance error, which was less than 1%. These results indicate that the mass-balance method fails to predict the local errors associated with a contour level.

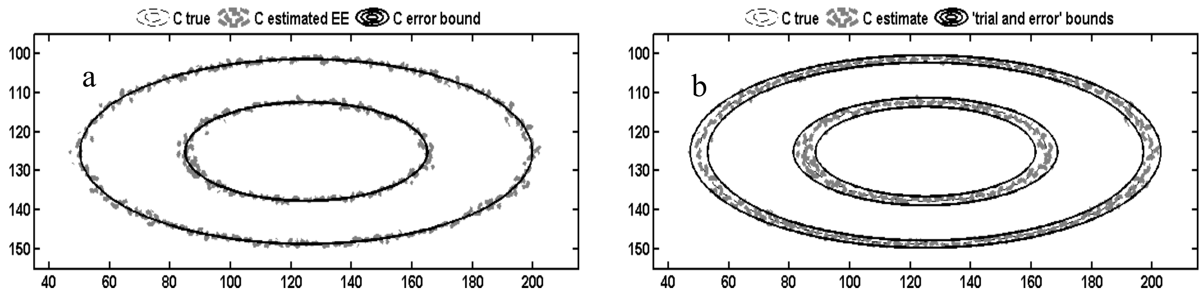

Figure 2. 5 g/l and 20 g/l contours for C_{true} and $C_{estimate, EE}$. 2a shows the $C_{estimate, EE}$ and C_{true} with corresponding mass balance error bound and 2b presents $C_{estimate, EE}$ with 'trial and error' bounds of 10% and 15% for 20 mg/l and 5 mg/l contours, respectively

STATISTICS BASED ERROR ESTIMATION METHOD

Typical IA datasets involve a large number of measurements. Therefore, standard statistical methods can be used to quantify error (Taylor 1997). The two errors listed earlier in the paper can be calculated as follows:

1) *Calibration-Relationship Error (CRE)*

$$\sigma_{c,\text{relationship}} = \sqrt{\frac{\sum_{i=1}^{N}\left(C_{i,measured} - C_{i,estimated}\right)^2}{df}} \tag{4}$$

where $\sigma_{c,relationship}$ is an estimate for CRE computed directly from regression statistics, *N* is the number of measurements in the calibration dataset, *p* is the number of parameters used in the relationship, *df* is the degrees of freedom of the relationship given by *N-p*;

2) *Experimental Error (EE)*

$$\sigma_{c,\text{experiment}} = \left|\frac{dC}{dI}\sigma_I\right| \tag{5}$$

where $\sigma_{c,experiment}$ is the error in concentration due to noise in pixel intensity values (or image property), *dC/dI* is the slope of the concentration-intensity relationship, and σ_i is the randomly distributed error in the recorded pixel intensity field.

BENCHMARKING DATASETS

We conducted a variable-density experiment by injecting a saltwater pulse source into a freshwater aquifer under ambient flow conditions. Detailed information about the methodology and the experiment are available in Goswami et al. (2008). Figure 3a shows the dense-plume pictorial dataset obtained from the experiment. The IA technique we developed was used to extract concentration values from the pictorial dataset and the results are shown in Figure 3b.

Numerical modeling of the experiment was completed using SEAWAT (Langevin and Guo 2006) and the modeled contours are shown in Figure 3c. Comparing Figures 3a, b and c, we can conclude that the IA technique and the numerical results are in good agreement. Using the new statistics-based error estimation technique, we calculated the percentage errors in the concentration values measured by the IA technique. We found approximately 15% error in the concentration values determined by the IA technique.

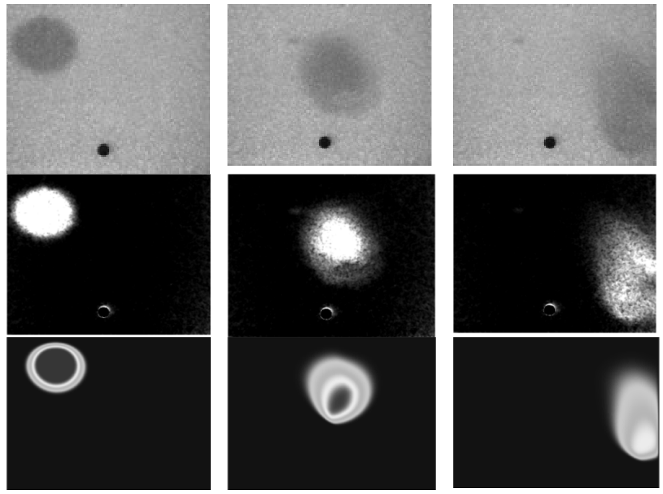

Figure 3. (top) digital images from the experiment, (center) concentrations from the IA technique, (bottom) concentrations from numerical model (SEAWAT)

CONCLUSIONS

This study demonstrates the inability of global mass balance errors to estimate local errors in IA concentration measurements. A new statistics-based approach was proposed to evaluate errors in IA techniques. The new error estimation method was used to quantify IA estimation errors in a sinking plume experiment. The error was found to be approximately 15%.

This study is not intended to represent the state-of-the-art in the application of IA, but to provide a simple method to calculate errors inherent in the application of an IA technique.

REFERENCES

Fischer, H.B., E.J. List, R.C.Y. Koh, J. Imberger, and N.H. Brooks. 1979. Mixing in inland and coastal waters Academic Press, Inc.

Goswami R. R., T. P. Clement. 2007. Laboratory-scale investigation of saltwater intrusion dynamics. Water Resources Research 43, doi:10.1029/2006WR005151.

Goswami R. R., B. Ambale, and T. P. Clement. 2008. Error estimation of concentration measurements obtained from image analysis, under review

Hansen, P.C., J.G. Nagy, and D.P. O'Leary. 2006. Deblurring Images: Matrices, Spectra and Filtering. Society for Industrial and Applied Mathematics.

Langevin, C.D. and W.X. Guo. 2006. MODFLOW/MT3DMS-based simulation of variable-density ground water flow and transport. Ground Water 44: 339-351

Taylor, J.R. 1997. An introduction to error analysis. Second Edition. University Science Books.

Contact Information: Rohit R. Goswami, Geosyntec Consultants, 5901 Broken Sound Pkwy NW, Suite 300, Boca Raton, FL 33487; Phone:561-995-0900, Email:rgoswami@geosyntec.com