Evaluating the Role of Flow Data on Concentration Fluctuations Through the Use of the Comparative Information Yield Curves

Felipe P. J. de Barros[1], Yoram Rubin and Souheil Ezzedine

Civil and Environmental Engineering, University of California – Berkeley

[1] barros@berkeley.edu. Dept. of Civil and Environmental Engineering. Berkeley, 626 Davis Hall, Berkeley, CA, 94720-1710, USA, [2] Atmospheric, Earth and Energy Science Dept. LLNL 7000 East Ave. L-208, Livermore, CA 94550

H43A-0970

Abstract

The significance of quantifying concentration fluctuations due to contaminant transport in heterogeneous flows through natural porous media is addressed. The challenge relies on the fact that the concentration field in the subsurface must be modeled in a probabilistic manner since full characterization of the site is impractical. In this paper we make use of the conditioning methodology introduced by Rubin [1991] to present a rational and concise approach to incorporate hydrogeological data on flow and transport processes in heterogeneous porous media. Most importantly, we wish to investigate the impact of conditioning flow data (for example hydraulic conductivity and head measurements) on the concentration variance as a function of the location of the environmentally sensitive target receptor. It is well documented that concentration variance often presents a halo shape. In the past, developments have been made to investigate the factors that influence this halo shape, however, there is still further needs to investigate this halo feature as a function of both travel distances and conditioning. This issue is of practical relevance since it has direct impact in on evaluating human health risk. We extend the use of existing analytical solutions to accommodate conditioning and parametric uncertainty to address these issues. We show how these analytical solutions can improve their predictive capabilities as hydraulic data and parametric uncertainty are accounted for. Also, we illustrate how the concept of Comparative Information Yield Curves (CIYC) can be used to provide better understanding of assessing characterization needs as a function of different flow and transport conditions.

I. Motivation and Research Goals

Motivation: Risk characterization - It is nowadays widely accepted that risk due to groundwater contamination needs to be accessed in a probabilistic manner. Knowing, in a rational way, where to collect data for better and tighter confidence bounds in human health risk is a challenge;

Objective: Investigate how the utility of a given sampling scheme changes with the physical and geometrical configuration of the problem as well as the prediction goal.

Research Questions:

How do characterization needs vary with travel time and transverse position of the environmentally sensitive target relative to the source?

As shown in the literature (e.g. *Rubin*, 2003), concentration variance is largest at the fringe of the plume: How does the worth of data change if the environmental target is positioned either at the centroid of the plume or at its fringe? How does this worth change for later travel times?

II. Theoretical Formulation

We are interested in evaluating the impact of conditioning measurements (such as hydraulic log-conductivity, denoted by $Y = \ln K$) in on the concentration field. Due to heterogeneity, Y is considered as a Spatial Random Function, SRF (e.g., Rubin 2003). Consider a vector of SRF hydrogeological parameters:

$$\boldsymbol{\theta} = \left\{ m_Y, I_Y, \sigma_Y^2 \ldots \right\}$$

and also, a measurement vector $\boldsymbol{d}^* = \{m_i\}$ where i represents a sampling strategy and consider $\boldsymbol{I_o}$ to denote prior knowledge. For a high Péclet condition, the concentration moments for a tracer migration are given by:

$$\left\langle C\left(\boldsymbol{x},t \middle| \boldsymbol{\theta},\boldsymbol{d}^*,\boldsymbol{I_o}\right)\right\rangle = C_o \prod_{i=1}^{N} \frac{1}{2}\left\{ erf\left[\frac{x_i - U_i t + L_i}{\sqrt{2X_{ii}\left(t\middle|\boldsymbol{\theta},\boldsymbol{d}^*,\boldsymbol{I_o}\right)}}\right] - erf\left[\frac{x_i - U_i t - L_i}{\sqrt{2X_{ii}\left(t\middle|\boldsymbol{\theta},\boldsymbol{d}^*,\boldsymbol{I_o}\right)}}\right]\right\};$$

$$\sigma_C^2\left(\boldsymbol{x},t\middle|\boldsymbol{\theta},\boldsymbol{d}^*,\boldsymbol{I_o}\right) = \left\langle C\left(\boldsymbol{x},t\middle|\boldsymbol{\theta},\boldsymbol{d}^*,\boldsymbol{I_o}\right)\right\rangle^2 \left[\frac{C_o}{\left\langle C\left(\boldsymbol{x},t\middle|\boldsymbol{\theta},\boldsymbol{d}^*,\boldsymbol{I_o}\right)\right\rangle} - 1\right].$$

Where N = 1, 2 and 3 (space dimension). X_{ii} is the particle displacement covariance. The source size is given by L_i and U_i is the mean flow velocity in the ith direction. The initial concentration is C_o. With the above moments, one may obtain the coefficient of variation for concentration (CV_C) conditional on prior information or on hydraulic measurements.

III. Solution Methodology

We evaluate the conditional displacement covariance tensor by integrating the conditional velocity covariance over time (*Rubin*, 1991):

$$X_{ij}^c(t) \equiv X_{ij}^c\left(t\middle|\boldsymbol{\theta},\boldsymbol{d}^*,\boldsymbol{I_o}\right) = \int_t\int_t u_{ij}^c\left[\left\langle \boldsymbol{X}_i^c(t')\right\rangle,\left\langle \boldsymbol{X}_j^c(t'')\right\rangle\right]dt'dt''$$

$$u_{ij}^c(x,x') = u_{lk}(x,x') - \sum_{p=1}^{N}\lambda_p(x)C_{U_kY}(x',x_p)$$

$$\sum_{p=1}^{N}\lambda_p(x)C_Y(x_p,x_t) = C_{U_kY}(x,x_t)$$

The above integrations were evaluated using the algorithm described in *Ezzedine* (1997) and further details in *Rubin* (1991). To investigate parametric uncertainty, we need to infer the PDF parameters. We use Bayesian inference:

$$f_{\theta|D,I_o}\left(\boldsymbol{\theta}\middle|\boldsymbol{d}^*,\boldsymbol{I_o}\right) = \frac{f_{D|\theta,I_o}\left(\boldsymbol{d}^*\middle|\boldsymbol{\theta},\boldsymbol{I_o}\right)g_{prior}\left(\boldsymbol{\theta}\middle|\boldsymbol{I_o}\right)}{\int f_{D|\theta,I_o}\left(\boldsymbol{d}^*\middle|\tilde{\boldsymbol{\theta}},\boldsymbol{I_o}\right)g_{prior}\left(\tilde{\boldsymbol{\theta}}\middle|\boldsymbol{I_o}\right)d\tilde{\boldsymbol{\theta}}};$$

$g_{prior}\left(\boldsymbol{\theta}\middle|\boldsymbol{I_o}\right)$: Prior PDF determined using Maximum Relative Entropy (see details in *Rubin*, 2003)

With the inferred PDF parameters we can calculate its relative entropy:

$$RE_{\theta|D,I_o} = \int_{-\infty}^{+\infty} f_{\theta|D,I_o}\left(\boldsymbol{\theta}\middle|\boldsymbol{d}^*,\boldsymbol{I_o}\right)\ln\left[\frac{f_{\theta|D,I_o}\left(\boldsymbol{\theta}\middle|\boldsymbol{d}^*,\boldsymbol{I_o}\right)}{g_{prior}\left(\boldsymbol{\theta}\middle|\boldsymbol{I_o}\right)}\right]d\boldsymbol{\theta}$$

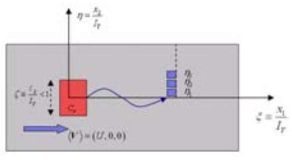
For each level of relative entropy, we can evaluate the gain of information in CV_c by evaluating:

$$\Delta CV_C(\%) = \frac{\left|CV_C^{Unc} - CV_C^{Cond}\right|}{CV_C^{Unc}}\times 100$$

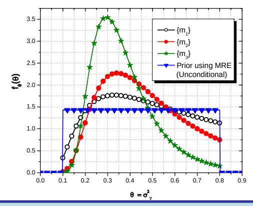
By plotting RE_θ vs ΔCV_C we obtain the Comparative Information Yield Curves See *de Barros and Rubin* (2008) and *de Barros et al*. (submitted) for further details.

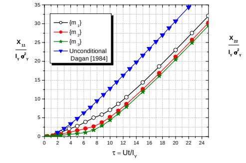
IV. Results for 2D Case

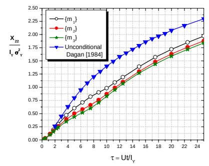
Problem set-up:



Inferred PDF parameters from measurements

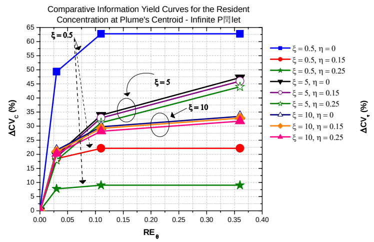


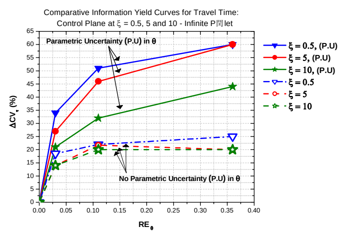


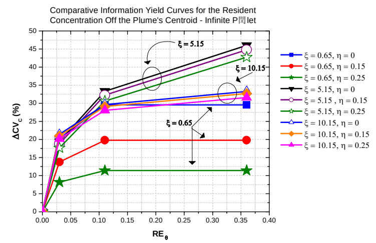


Unconditional and conditional longitudinal, X_{11}, and transversal, X_{22}, particle displacement covariances.

Comparative information yield curves as a function of transversal and longitudinal distance. Plots below obtained at the centroid of the plume (left figure) and off-centroid (right figure):







Notice the change of the slope for each of the curves. With the same level of parameter uncertainty and conditioning measurements, the value of information changes according to the position of the target. Differences between the coefficient of variation for resident concentration and travel time are also observed. Major difference observed at early travel time due to the halo effect of the concentration variance.

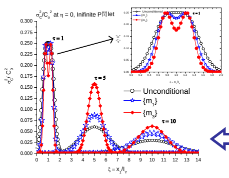
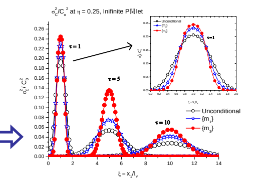


Illustration of how the concentration variance changes for early and late travel times.



V. Concluding Remarks

(I) The use of the *comparative information yield curves* proved useful in investigating uncertainty reduction. **(II)** For the 2D case investigated, the value of information varies with both transversal and longitudinal position of the environmentally sensitive target. Especially if the target is aligned at the fringe of the plume (where the concentration variance is largest – mainly at early travel times). **(III)** Characterization needs vary if uncertainty is addressed in terms of travel times or resident concentration. **(IV)** Parametric uncertainty has a strong role in defining characterization needs.

References

de Barros, F.P.J. and Rubin, Y., A risk-driven approach for site characterization, *Water Resour. Res.*, Vol.44, W01414 , 2008;

de Barros *et al.*, The concept of comparative information yield curves and their application to risk-based site characterization, Submitted to *Water Resour. Res*

Ezzedine, S., Fast computation of head and velocity covariances in three-dimensional statistically axisymmetric heterogeneous porous media, *Water. Resour. Res.*, Vol.33, 1997

Rubin Y, Prediction of Tracer Plume Migration in Disordered Porous Media by the Method of Conditional Probabilities, *Water Resour. Res.*, 27(6), 1991

Rubin, Y. Applied Stochastic Hydrogeology. Oxford Press, 1st Edition, 2003.;