Bayesian evaluation of groundwater age distribution using radioactive tracers and anthropogenic chemicals

Arash Massoudieh,[1] Soroosh Sharifi,[1] and D. Kip Solomon[2]

Received 1 January 2012; revised 9 July 2012; accepted 17 July 2012; published 20 September 2012.

[1] The development of a Bayesian modeling approach for estimation of the age distribution of groundwater using radioactive isotopes and anthropogenic chemicals is described. The model considers the uncertainties associated with the measured tracer concentrations as well as the parameters affecting the concentration of tracers in the groundwater, and it provides the posterior probability densities of the parameters defining the groundwater age distribution using a Markov chain Monte Carlo method. The model also incorporates the effect of dissolution of aquifer minerals on diluting the ^{14}C signature and the uncertainties associated with this process on the inferred age distribution parameters. Two demonstration modeling cases have been performed. First, the method was applied to simulated tracer concentrations at a discharge point of a hypothetical 2-D vertical aquifer with two recharge zones, leading to a mixed groundwater age distribution under different presumed uncertainties. When the error variance of the observed tracer concentrations is considered unknown, the method can estimate the parameters of the fitted exponential-lognormal distribution with a relatively narrow credible interval when five hypothetical samples are assumed to be collected at the discharge point. However, when a single sample is assumed, the credible intervals become wider, and credible estimations of the parameters are not obtained. Second, the method was applied to the data collected at La Selva Biological Station in Costa Rica. In this demonstration application, nine different forms of presumed groundwater age distributions have been considered, including four single forms and five mixed forms, assuming the groundwater consists of distinct young and old fractions. For the medium geometrical standard deviation $\delta_{c,i} = 1.41$, the model estimates a young groundwater age of between 0 and 350 years, with the largest odds being given to a mean age of approximately 100 years, and a fraction of young groundwater of between 15% to roughly 60%, with the largest odds for 30%. However, the method cannot definitively rule out larger fractions of young groundwater. The model provides a much more uncertain estimation of the age of old groundwater, with a credible interval of between 20,000 to 200,000 years.

Citation: Massoudieh, A., S. Sharifi, and D. K. Solomon (2012), Bayesian evaluation of groundwater age distribution using radioactive tracers and anthropogenic chemicals, *Water Resour. Res.*, *48*, W09529, doi:10.1029/2012WR011815.

1. Introduction

[2] Estimating groundwater age has received increasing attention due to its applications in assessing the sustainability of water withdrawal from aquifers, evaluating the vulnerability of groundwater resources due to contamination of near-surface recharge waters [*Bethke and Johnson*, 2008; *Glynn and Plummer*, 2005; *Plummer*, 1999], and estimating the times needed for groundwater quality management strategies to affect the groundwater quality [*Zoellmann et al.*, 2001]. Furthermore, knowing the groundwater age in an aquifer can help in determining some physical characteristics of the aquifer affecting the flow, such as the hydraulic conductivity field and its spatial variability. This leads to a better understanding of the nature of groundwater flow, and results in a more realistic calibration of groundwater flow models [e.g., *Portniaguine and Solomon*, 1998; *Reilly et al.*, 1994; *Szabo et al.*, 1996]. Groundwater age can also be used to evaluate the history and fate of contaminants [*Bohlke and Denver*, 1995] and anthropogenic activities [*Hinsby et al.*, 2001]. Various tracers, including radioactively decaying tracers, such as $^{3}H/^{3}He$, ^{85}Kr, ^{36}Cl, ^{39}Ar, ^{32}Si, ^{14}C, tracers with variable atmospheric concentrations as a result of anthropogenic activities such as CFCs and SF_6 and linearly accumulating ones, such as ^{4}He and ^{40}Ar, have been used for groundwater dating by various researchers [*Bauer et al.*, 2001; *Bethke and Johnson*, 2008; *Bruce et al.*, 2007; *Busenberg and Plummer*, 1992; *Castro et al.*, 1998; *Lehmann et al.*, 2003; *Plummer and Sprinkle*, 2001]. The use of many of these tracers in groundwater dating entails various challenges, including but

[1]Department of Civil Engineering, Catholic University of America, Washington, D.C., USA.
[2]Department of Geology and Geophysics, University of Utah, Salt Lake City, Utah, USA.

Corresponding author: A. Massoudieh, Department of Civil Engineering, Catholic University of America, 620 Michigan Ave., NE, Washington, D.C. 20064, USA. (massoudieh@cua.edu)

0043-1397/12/2012WR011815

not limited to (1) the fact that some of these tracers do not move with the same velocity as the pore water due to their adsorption to the rock or soil matrix [*Cook and Solomon*, 1995, 1997], (2) the uncertainties associated with their measured concentrations due to measurement error as well as the spatial and temporal variations of the tracer concentrations, (3) the fact that some of these isotopes might be produced in the medium or altered due to release of their corresponding stable isotopes through the dissolution of the rock matrix (e.g., ^{14}C), hence diluting the isotope signature in the groundwater [*Geyh*, 1999], and (4) the uncertainties regarding the persistence or degradation rates of some of the tracers, in particular synthetic chemicals such as CFCs. A related challenge is the collection of samples from multiple, but unknown, flow paths that result in arbitrary mixtures of transit times.

[3] Two general classes of models have been used to evaluate groundwater age data [*Cook and Bohlke*, 2000]: (1) models that predict variations in ages within the aquifer; since ages increase with depth in most aquifers, these are referred to as "groundwater stratigraphy" models, and (2) models that predict the integrated age of water discharging from the system, known as "groundwater discharge" models. The choice between these two categories depends on the type of data available and the objectives of the study. Most of the earlier stratigraphy models have implicitly assumed a uniform age for all of the pore water at a specific location of the aquifer (i.e., the piston flow model), and consequently have obtained relationships between the groundwater age and the ratio between the tracer concentration or isotope ratios at that location and those of the recharging waters [*Plummer*, 1999]. These methods provide an age estimate based on a single tracer; however, in some cases, the estimates have been observed to be different when multiple isotopes with different decay or accumulation rates have been used to date the same groundwater [e.g., *Bruce et al.*, 2007; *Lehmann et al.*, 2003; *Massoudieh and Ginn*, 2011; *Troldborg et al.*, 2008]. These inconsistencies have been attributed to the hydrodynamic mixing of waters with various ages and therefore to the presence of a mixture of different ages of a tracer in a sample [*Bethke and Johnson*, 2008; *Busenberg and Plummer*, 1992; *Goode*, 1996; *Walker and Cook*, 1991]. In several studies where forward modeling using particle-tracking or momentum approaches has been used to determine the forms of age distribution, it has been shown that representing groundwater age as a single value can result in misleading results. It has also been shown that single value determination of groundwater age can result in unstable estimates of age in heterogeneous aquifers, meaning small perturbations in space or time can lead to significant changes in the age estimates [*Troldborg et al.*, 2008; *Varni and Carrera*, 1998; *Weissmann et al.*, 2002]. Therefore, it is suggested that it is more appropriate and realistic to consider the groundwater age as a distribution defined by

$$\rho(a) = \frac{dR(a)}{da} \tag{1}$$

where $R(a)$ is the fraction of water having an age smaller than a. No practical way to directly determine the groundwater age distribution experimentally using environmental tracers has yet been demonstrated [*Massoudieh and Ginn*, 2011]. Therefore, in many cases prior assumptions about the form of $\rho(a)$ are made and then least square or maximum likelihood methods are used to determine the parameters defining these distributions [e.g., *Solomon et al.*, 2010]. The piston flow model (PFM; i.e., Dirac delta age distribution), the exponential age distribution model (EM), and their combination (EMP model), suggested by *Maloszewski and Zuber* [1993] and here referred to as exponential-Dirac age distribution, have been widely used.

[4] When using anthropogenic or some radio tracers for groundwater dating, a source of uncertainty is the fact that despite the main assumption often used, some tracers do not move with the same velocity as the groundwater due to adsorption and desorption to aquifer solid materials. It has been shown that CFCs, for example, can undergo adsorption and desorption to a certain degree [*Bauer et al.*, 2001; *Choung and Allen-King*, 2010; *Cook et al.*, 1995; *Horneman et al.*, 2008; *Weissmann et al.*, 2002] and also that ^{14}C in dissolved inorganic carbon (DIC) can interact with the aquifer minerals [*Garnier*, 1985; *Plummer et al.*, 2004; *Sebol et al.*, 2007], resulting in some levels of retardation.

[5] Another source of uncertainty is the fact that the decay rate due to biodegradation for some of the synthetic chemicals used for groundwater dating is not known with certainty. For example, when CFCs are used for groundwater dating, it is often assumed that their concentrations are not impacted by biodegradation. It has been shown that, in reality, CFCs can undergo biodegradation under anaerobic conditions [*Lovley and Woodward*, 1992; *Oster et al.*, 1996; *Scheutz and Kjeldsen*, 2003; *Sonier et al.*, 1994; *Weissmann et al.*, 2002]. In natural aquifers, over long periods of time, patches of anaerobic zones can be present at different scales and some biodegradation can occur [*Weissmann et al.*, 2002]. Due to the unknown distribution and dynamics of anaerobic zones, estimation of the effective biodegradation rates over the path of these environmental tracers involves a great deal of uncertainty.

[6] The goal of the presented study is to provide a method that can assess the uncertainties in inference of the groundwater age distribution as a result of uncertainties in the fate and transport parameters as well as the uncertainties due to possible measurement error and lack of representativeness in observed tracer concentrations. The method uses several tracers for age estimation and also provides a flexible method in terms of incorporating any prior knowledge that may be available about the parameter values controlling the fate and transport of tracers in the aquifer. The impact of mineral dissolution, retardation, biodegradation, and the uncertainties associated with them can be incorporated into the model. The method is also flexible in terms of the forms of the presumed groundwater age distributions.

2. Mathematical Model

[7] Adsorption and desorption, as well as the potential release of some of the environmental tracers as a result of mineral dissolution, make the age distribution of the water differ from the age distribution of the tracers. To accommodate for this difference, the concept of tracer age distribution $\rho_i(\mathbf{x}, t, a)$ is introduced here, which represents the distribution of the time a in the past, a particular tracer sampled at location $\mathbf{x}$, and the time t has entered the ground through recharge or other sources. In general, the particles of a tracer in the

pore water can consist of particles with atmospheric sources (i.e., the molecules that have entered the system through percolation of precipitation water) and molecules with mineral dissolution sources (i.e., the molecules with their latest source being mineral dissolution). Therefore, the tracer age distribution can be expressed as

$$\rho_i(\mathbf{x}, t, a) = f_{m,i}\rho_{m,i}(\mathbf{x}, t, a) + \left(1 - f_{m,i}\right)\rho_{d,i}(\mathbf{x}, t, a) \tag{2}$$

where $f_{m,i}$ is the fraction of the tracer transported in the pore water from the recharge, and $\rho_{m,i}(\mathbf{x}, t, a)$ and $\rho_{d,i}(\mathbf{x}, t, a)$ are the age distributions of the tracer transported and generated through dissolution, respectively. It should be noted that for most of the environmental tracers used for groundwater dating, except for a few such as ^{14}C, the contribution from minerals is effectively zero (i.e., $f_{m,i} = 1$). The relationship between the concentration or isotope ratio of a first-order decaying tracer i and the age can be written as [*Maloszewski and Zuber*, 1982]

$$\begin{aligned} c_i(\mathbf{x}, \lambda, t) &= \int_0^\infty c_{i,o}(t-\tau)e^{-\lambda_i\tau}\rho_i(\mathbf{x}, t, \tau)d\tau \\ &= \int_0^\infty c_{i,o}(t-\tau)e^{-\lambda_i\tau}\left[f_{m,i}\rho_{m,i}(\mathbf{x}, t, \tau)\right. \\ &\quad \left. + \left(1 - f_{m,i}\right)\rho_{d,i}(\mathbf{x}, t, \tau)\right]d\tau \end{aligned} \tag{3a}$$

where $c_i(\mathbf{x}, \lambda_i, t)$ is the measured tracer concentration (or isotope ratio) of tracer i at location $\mathbf{x}$ and time t, and $c_{i,o}(t-\tau)$ is the concentration (or isotope ratio) in the recharge water at time $t - \tau$. In equation (3a), λ_i is the decay rate of tracer i. Also, for linearly accumulating stable nuclides, the following equation can be used:

$$\begin{aligned} c_i(\mathbf{x}, \lambda_i, t) &= \int_0^\infty \left[c_{i,o}(t-\tau) + \lambda_i\tau\right]\rho_i(\mathbf{x}, t, \tau)d\tau \\ &= \int_0^\infty \left[c_{i,o}(t-\tau) + \lambda_i\tau\right]\left[f_m\rho_{m,i}(\mathbf{x}, t, \tau)\right. \\ &\quad \left. + \left(1 - f_{m,i}\right)\rho_{d,i}(\mathbf{x}, t, \tau)\right]d\tau. \end{aligned} \tag{3b}$$

[8] Equation (3b) represents the linear accumulation rate of isotope i in the case of a linearly accumulating isotope where λ_i is the linear accumulation rate of tracer i. The relationship between the water age and tracer age for a reactive tracer with retardation factor of R_i undergoing equilibrium linear sorption can be written as

$$\rho_{m,i}(\mathbf{x}, t, a) = \frac{1}{R_i}\rho_m(\mathbf{x}, t, a/R_i) \tag{4}$$

where $\rho_m(X, t, a)$ is the age distribution of water. The goal here is to estimate $\rho_m(X, t, a)$ using the observed concentrations of multiple tracers. Incorporating equation (4) into equations (3a) and (3b), and imposing a change of variable $\tau' = \tau/R_i$ on the first terms inside the integrals yields

$$\begin{aligned} c_i(\mathbf{x}, \lambda_i, t) &= f_m \int_0^\infty c_{i,o}(t - R_i\tau')e^{-\lambda R_i\tau'}\rho_m(\mathbf{x}, t, \tau')d\tau' \\ &\quad + \left(1 - f_{m,i}\right)\int_0^\infty c_{i,o}(t-\tau, \lambda_i)e^{-\lambda_i\tau}\rho_{d,i}(\mathbf{x}, t, \tau)d\tau \end{aligned} \tag{5a}$$

$$\begin{aligned} c_i(\mathbf{x}, \lambda_i, t) &= f_{m,i}\int_0^\infty \left[c_{i,o}(t - R_i\tau') + \lambda_i\tau' R_i\right]\rho_m(\mathbf{x}, t, \tau')d\tau' \\ &\quad + \left(1 - f_{m,i}\right)\int_0^\infty \left[c_{i,o}(t-\tau) + \lambda_i\tau\right]\rho_{d,i}(\mathbf{x}, t, \tau)d\tau. \end{aligned} \tag{5b}$$

[9] In most practical cases, when the decay time scale of the tracer is small compared to the age of the minerals (e.g., for the case of ^{14}C), the second term in equation (5a) becomes negligible.

[10] In addition to the tracers directly used for age determination, there are some stable isotopes that are indirectly used to characterize some of the processes affecting the age estimation. Particularly, the δ^{13}C value can be used to infer the fraction of carbon originating from recharge or mineral dissolution (i.e., biogenic and mineral sources). This is due to the differential uptake of ^{13}C by plants [*Mazor*, 2003]. The ^{13}C content, assuming two contributing sources of biogenic and mineral, can therefore be expressed as

$$c_i(\mathbf{x}, \lambda_i, t) = f_m c_b + \left(1 - f_{i,m}\right)c_m \tag{6}$$

where c_b and c_m are the normalized ^{13}C concentrations in biogenic and mineral sources. Similarly, ^{36}Cl concentration can often be used more effectively for inferring the fraction of old groundwater than for direct estimation of age distribution because of addition of subsurface ^{36}Cl.

3. Bayesian Inference

[11] The exact form of the age distribution is not known and is typically approximated by a presumed distribution function. Various known or mixed distribution functions can be assumed for the water age, including but not limited to exponential, lognormal, binary, inverse-Gaussian, gamma, and various combinations of them:

$$\rho_m(X, t, a) = \rho_m(a; \phi_1, \phi_2, \ldots, \phi_n) \tag{7}$$

where ϕ_1, ϕ_2, … ϕ_n are the parameters defining the age distribution. Here, the goal of Bayesian modeling is to infer the probability density functions of the parameters of such distribution functions (i.e., ϕ_1, ϕ_2, …, ϕ_n) and also the fraction of tracers with atmospheric sources, f_m, while considering the uncertainties in the parameters affecting the transport and fate of tracers (henceforth referred to as fate and transport parameters), including retardation factors, R_i, decay rates, λ_i, and also the observed concentrations. Therefore, the parameters are considered random variables where some prior information about them might be available and is expressed as probability density functions while the known information

about them is represented by probability distributions referred to as prior distributions. The true concentrations (or isotope ratios) of tracers at the sampling locations will also be considered random variables. It is worth noting that the uncertainties in the observed concentrations can stem from the measurement errors and from the fact that, due to the temporal and spatial heterogeneity of tracer concentrations, the sample does not exactly represent the local average or "true" tracer concentrations in the aquifer.

[12] If we consider vector $\mathbf{C}$ to represent the random vector comprising the true concentrations (or isotope ratios) of all chemicals that can be directly calculated for a given set of parameters using the models introduced the previous section (equations (5a), (5b), and (6)), and vector $\hat{\mathbf{C}}$ to be the observed concentrations, and if we assume a known standard deviation representing the variability of $\mathbf{C}$, then the probability of observing $\hat{\mathbf{C}}$ conditioned to a set of given age distribution parameters and fate and transport parameters is calculated using the Bayes theorem as

$$p(\phi_1,\phi_2,\ldots,f_m,R_1,R_2,\ldots,\lambda_1,\lambda_2,\ldots,\Gamma|\hat{C}) = \frac{p(\hat{C}|\phi_1,\phi_2,\ldots,f_m,R_1,R_2,\ldots,\lambda_1,\lambda_2,\ldots,\Gamma)p(\phi_1,\phi_2,\ldots,f_m,R_1,R_2,\ldots,\lambda_1,\lambda_2,\ldots,\Gamma)}{p(\hat{C})} \quad (8)$$

where Γ is the variance-covariance matrix of the observation error (or a transformation of it), $p(\hat{C}|\phi_1,\phi_2,\ldots,f_m,R_1,R_2,\ldots,\Gamma)$ is the likelihood function, $p(\phi_1, \phi_2, \ldots, f_m, R_1, R_2, \ldots)$ is the prior distribution of the parameters, and $p(\hat{C})$ is a normalizing factor which is equal to the integral of the numerator over the entire parameter space:

$$p(\hat{C}) = \int_{-\infty}^{\infty} p(\hat{C},\phi_1,\phi_2,\ldots,f_m,R_1,R_2,\ldots\lambda_1,\lambda_2,\ldots,\Gamma)\, d\phi_1 d\phi_2 \cdots df_m dR_1 dR_2 \cdots d\lambda_1 d\lambda_2 \cdots d\Gamma. \quad (9)$$

[13] Assuming that the prior information for each of the parameters is independent, the equation for the posterior probability density can be written as

$$p(\phi_1,\phi_2,\ldots,f_m,R_1,R_2,\ldots\lambda_1,\lambda_2,\ldots,\Gamma|\hat{C}) = \frac{p(\hat{C}|\phi_1,\phi_2,\ldots,f_m,R_1,R_2,\ldots,\lambda_1,\lambda_2\ldots,\Gamma)p(\phi_1)p(\phi_2)\cdots p(f_m)p(R_1)p(R_2)\cdots p(\lambda_1)p(\lambda_2)\cdots p(\Gamma)}{\int_{-\infty}^{\infty} p(\hat{C},\phi_1,\phi_2,\ldots,f_m,R_1,R_2,\ldots,\lambda_1,\lambda_2\ldots,\Gamma)d\phi_1 d\phi_2\cdots df_m dR_1 dR_2\cdots d\lambda_1 d\lambda_2\cdots d\Gamma}. \quad (10)$$

[14] For organic synthetic tracers, such as CFCs, adsorption mostly takes place to the soil organic matter, and therefore the retardation factor can be written as

$$R_i = 1 + \frac{\rho_b}{\theta} f_{oc} K_{oc,i} \quad (11)$$

where ρ_b is the soil bulk density, θ is the water content (equal to the porosity under saturated condition), f_{oc} is the soil organic content, and $K_{oc,i}$ is the organic-water partition coefficient for tracer i. If the uncertainties associated with ρ_b, θ, and K_{oc} are small enough to be ignored, then the only stochastic element in retardation factor will be f_{oc}, and thus the posterior equation can be written as

$$p(\phi_1,\phi_2,\ldots,f_m,f_{oc},\lambda_1,\lambda_2,\ldots,\Gamma|\hat{C}) = \frac{p(\hat{C}|\phi_1,\phi_2,\ldots,f_m,f_{oc},\lambda_1,\lambda_2\ldots,\Gamma)p(\phi_1)p(\phi_2)\cdots p(f_m)p(f_{oc})p(\lambda_1)p(\lambda_2)\cdots p(\Gamma)}{\int_{-\infty}^{\infty} p(\hat{C},\phi_1,\phi_2,\ldots,f_m,f_{oc},\lambda_1,\lambda_2\ldots,\Gamma)d\phi_1 d\phi_2\cdots df_m df_{oc} d\lambda_1 d\lambda_2\cdots d\Gamma}. \quad (12)$$

3.1. Likelihood Function

[15] Here $p(\hat{\mathbf{C}}|\phi_1,\phi_2,\ldots,f_m,R_1,R_2,\ldots\lambda_1,\lambda_2,\ldots,\Gamma)$ is the likelihood of observing a set of tracer concentrations/isotope ratios of $\hat{\mathbf{C}}$ conditioned to the parameters. By knowing the form and magnitude of error in the observed tracer concentrations based on the employed analytical method and the possible nonrepresentativeness of samples as a result of temporal and spatial heterogeneities, this probability can be calculated for any set of parameters from:

$$p(\hat{\mathbf{C}}|\phi_1,\phi_2,\ldots,f_m,R_1,R_2,\ldots,\Gamma) = f(\hat{\mathbf{C}}|\mathbf{C}) = \varphi(\boldsymbol{\varepsilon} = \hat{\mathbf{C}} - \mathbf{C};\Gamma) \quad (13)$$

where φ is the distribution of error $\boldsymbol{\varepsilon}$, and $\mathbf{C}$ is the "true" tracer concentration obtained from the model presented in equations (5a), (5b), and (6).

[16] Here we assume that the observation errors for tracer concentrations are lognormally distributed and multiplicative. Assuming a lognormal distribution for measurement error and variability due to spatial and temporal heterogeneities has been shown to be more realistic for natural systems compared to the normal distribution, since spatially heterogeneous and temporally variable environmental quantities representing the properties of natural systems are often nonnegative and skewed and can vary by orders of magnitude [*El-Shaarawi*, 1989; *El-Shaarawi and Lin*, 2007; *Helsel and Cohn*, 1988; *Limpert et al.*, 2001; *Shumway et al.*, 2002]. Inspecting the distribution of residuals after performing the MCMC parameter estimation on the real data (section 5) also shows that the lognormal and multiplicative error structure is more appropriate than normal and additive error structure. The use of lognormal error structure also allows us to assume that the observation error of each tracer is proportional with the magnitude of concentration. Making this assumption and also assuming the observation errors of tracers are independent reduces the variance-covariance matrix to a scalar number representing the scale factor of the lognormal distribution representing the observed error (σ) [*Keats et al.*, 2009]. Therefore the likelihood function can be written as

$$f(\hat{\mathbf{C}}|\mathbf{C}) = \frac{1}{\sqrt{2\pi}^m \prod_{i=1}^{m}\prod_{j=1}^{n} \hat{c}_i \ln(\sigma)} e^{-\sum_{j=1}^{n}\sum_{i=1}^{m}\frac{[\ln(\hat{c}_{i,j})-\ln(c_i)]^2}{2\ln(\sigma)^2}} \quad (14)$$

where $\hat{c}_{i,j}$ is the observed concentration of tracer i in the jth sample collected at the discharge location, n is the number of samples analyzed at the location and σ is the geometric standard deviation of observation error. The error associated with the concentration of stable isotopes (e.g., δ^{13}C) is considered to be normally distributed and additive:

$$f(\hat{c}_c|c_c) = \frac{1}{\sqrt{2\pi\sigma_c^2}} e^{-\sum_{j=1}^{n} \frac{(\hat{c}_{c,j}-c_c)^2}{2\sigma_c^2}} \tag{15}$$

where $\hat{c}_{c,j}$ is the observed isotope content of the stable isotope in sample j (e.g., δ^{13}C) collected from the same well and c_c is the true isotope content. and σ_c is the standard deviation of the measurement error of the stable isotope that is considered deterministic.

3.2. Prior Distributions

[17] The prior distributions for decay rates and retardation factor (or soil organic matter in the case of using equation (12)) were considered to be lognormal. In the demonstration cases presented in this study the tracers are organic and adsorb to soil organic matter so R is assumed to be a direct function of f_{oc} through equation (12)). If no prior information about the contribution from mineral dissolution is present, the prior distribution of f_m will be considered uniform, varying between 0 and 1. Also, the prior densities for c_m and c_b were considered to be Gaussian. Incorporating these prior distributions into equation (13) leads to

$$\begin{aligned} &p\left(\phi_1, \phi_2, \ldots, f_m, f_{oc}, \lambda_1, \lambda_2, \ldots, c_m, c_b, \sigma_{c,i} \middle| \hat{C}\right) \\ &\propto e^{-\sum_{j=1}^{n}\sum_{i=1}^{m} \frac{\left[\ln(\hat{c}_{i,j}) - \ln(c_i)\right]^2}{2\ln(\sigma)^2}} e^{-\sum_{j=1}^{n} \frac{(\hat{c}_{c,j}-c_c)^2}{2\sigma_c^2}} U(f_m, 0, 1) \\ &\cdot \left\{ \frac{1}{f_{oc}} e^{-\frac{\left[\ln\left(\mu_{foc}\right) - \ln(f_{oc})\right]^2}{2\ln\left(\sigma_{foc,i}\right)^2}} \right\} \left\{ \frac{1}{\prod \lambda_i} e^{-\sum_{i=1}^{m} \frac{\left[\ln\left(\mu_{\lambda_i}\right) - \ln(\lambda_i)\right]^2}{2\ln\left(\sigma_{\lambda,i}\right)^2}} \right\} e^{\left\{ \frac{[\mu_{cb}-c_b]^2}{\sigma_{cb}{}^2} + \frac{[\mu_{cm}-c_m]^2}{\sigma_{cm}{}^2} \right\}} \end{aligned} \tag{16}$$

where $U(f_m, 0, 1)$ is the uniform distribution between 0 and 1 and μ_{foc} is the prior geometric mean of the aquifer's organic content.

3.3. Markov Chain Monte Carlo

[18] In order to find the expected values and the credible intervals of the posterior joint PDF and to evaluate the correlations between the inferred parameters, equation (16) has to be integrated over the parameter space. It is clear that since the number of dimensions of the parameter space is large, and since calculating the values of true tracer concentration C given a set of parameters requires evaluating a complex integral, integration of equation (16) is prohibitive. Markov chain Monte Carlo (MCMC) methods are relatively simple methods for generating parameter samples, based on a posterior distribution [*Gamerman and Hedibert*, 2006; *Kaipio and Somersale*, 2004]. Algorithms such as Metropolis-Hasting or Gibbs sampling provide a way to generate samples according to a large-dimensional posterior joint probability density function (JPDF). Here we use the Metropolis-Hasting algorithm [*Metropolis et al.*, 1953]. The algorithm generates a Markov chain in which each new set of parameters depends on the previous parameter set. The algorithm uses proposal densities Q_1, which are the probabilities of moving to a proposed parameter set vector, X', and are conditioned to the current parameter set $X^{(k)} = \{\phi_1^{(k)}, \phi_2^{(k)}, \ldots, f_m^{(k)}, f_{oc}^{(k)}, \lambda_1^{(k)}, \lambda_2^{(k)}, \ldots, c_m^{(k)}, c_b^{(k)}, \sigma^{(k)}\}$. The proposal parameter set X' is accepted as the next parameter set based on the following criteria:

$$X^{(k+1)} = \left\{ \begin{array}{ll} X' & \text{if} \quad u(0,1) < \min\left\{ \frac{P(X')}{P(X^{(k)})} \cdot \frac{Q(X^{(k)};X')}{Q(X';X^{(k)})}, 1 \right\} \\ X^{(k)} & \text{otherwise} \end{array} \right\} \tag{17}$$

where $u(0, 1)$ is a uniformly distributed random number between 0 and 1, and P is the posterior probability density. The following transition functions are used to produce the proposed parameter sets:

$$\phi_i' = \phi_i^{(k)} e^{\varsigma z} \tag{18a}$$

$$\lambda_i' = \lambda_i^{(k)} e^{\varsigma z} \tag{18b}$$

$$f_{oc}' = f_{oc}^{(k)} e^{\varsigma z} \tag{18c}$$

$$\frac{f_m'}{1-f_m'} = \frac{f_m^{(k)}}{1-f_m^{(k)}} e^{\varsigma z} \tag{18d}$$

$$c_b' = c_b^{(k)} + \varsigma z \tag{18e}$$

$$c_m' = c_m^{(k)} + \varsigma z \tag{18f}$$

$$\sigma' = \sigma^{(k)} + \varsigma z \tag{18g}$$

where z is a random number generated from the standard normal distribution and ς is the perturbation factor for the random walk Metropolis Hasting algorithm. The ratio between the proposal probabilities is therefore calculated as

$$\frac{Q(X^{(k)};X')}{Q(X';X^{(k)})} = \frac{\left[\prod \phi_i^{(k)}\right]\left[\prod \lambda_i^{(k)}\right] f_{oc}^{(k)} f_m^{(k)} \left(1-f_m^{(k)}\right)}{[\prod \phi']_i [\prod \lambda_i'] f'_{oc} f'_m (1-f'_m)}. \tag{19}$$

[19] A C++ code is written to perform the MCMC simulation. The number of Markov chains can be determined by the user. In the example application presented in the next section, 8 chains were used and 1,000,000 samples resulted in convergence of the MCMC method. The first 100,000 samples were left out as "burn-in" period.

4. Hypothetical Case Study

[20] The method was applied to a hypothetical aquifer presented in Figure 1, which shows a vertical cross section

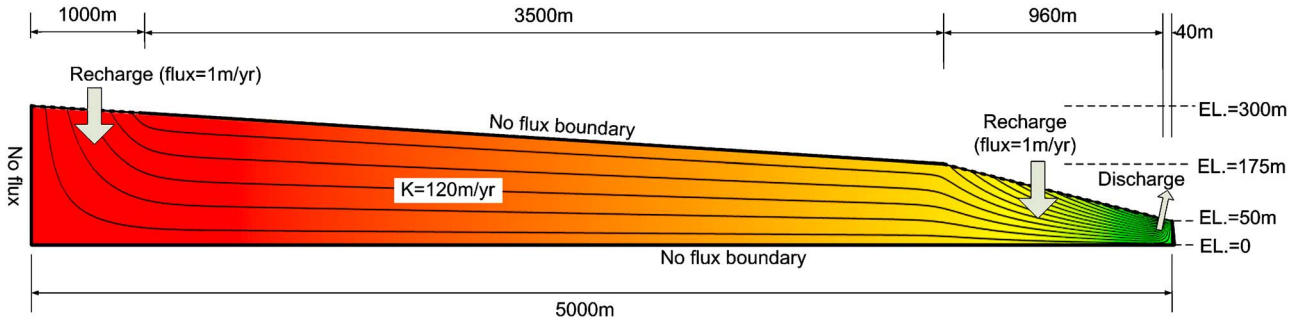

Figure 1. The schematic of the hypothetical 2-D (vertical) aquifer model discussed in section 4.

of an aquifer with a flow field that was assumed to be uniform with respect to the axis normal to the page (2-D vertical representation). The hypothetical case was designed in such a way that it produces a mixed (i.e., old and new) groundwater age distribution at the discharge region indicated on the right side of the figure. There are two recharge zones: one located at a distance between 4000 and 5000 m and the other between 0 and 960 m from the discharge zone. The recharge was assumed to occur at a constant rate of 1 m/yr, and the hydraulic conductivity of the domain was assumed to be 120 m/yr. Seven tracers, SF_6, CFC-11, CFC-12, CFC-113, ^3H, ^{14}C, and ^4He, were considered for this test case. The concentration boundary conditions for the seven tracers were calculated based on Henry's law and atmospheric concentrations obtained from the *International Atomic Energy Agency* (*IAEA*) [2011], *Plummer and Busenberg* [2006], http://isohis.iaea.org, ftp://cdiac.ornl.gov/pub/ndp057, and http://cdiac.ornl.gov/trends/co2/cent-scha.html and are presented in Figure 2. Two hypothetical cases were studied. In the first case it was assumed that five samples at different times have been collected and each have been analyzed for the 8 tracers. The five samples were assumed to represent the variability of the tracer concentration. In the second hypothetical case only one sample was used. In a real world problem due to the uncertainties associated with the transport parameters (λ and f_{oc}), their actual (true) values are different than their expected values perceived by the modeler and used in the model. In addition, the observed concentrations of tracers are also different from the actual locally averaged concentrations due to measurement errors as well as lack of representativeness of the samples. To incorporate these uncertainties into the demonstration modeling presented here, noise based on the prior distributions of the parameters was added to the fate and transport parameters λ and f_{oc} and to the modeled observed tracer concentrations. The geometric standard deviation of observed error σ was considered to be equal to 1.044. Concentration of samples were calculated by adding noise according to the prior densities of fate and transport parameters reflecting the uncertainties associated with them as well as adding noise to the final outcome according to the observation error. The values of five tracer concentrations are presented in Table 1. For the case with only one tracer, only the first observation in the table was used. The fate and transport parameters including the decay rates λ_i and f_{oc} used to generate the synthetic tracers' observed concentrations are listed in Table 2, and the values of $\delta^{13}C$ in biogenic and mineral carbon sources are listed in Table 3. The contribution of dissolved inorganic carbon from mineral dissolution, f_m, was considered to be equal to 0.8 for the forward simulation of ^{14}C content at the discharge zone. The modeled groundwater age distribution was obtained by taking the derivative of the breakthrough curve associated with a step function release of a nonreactive tracer at the recharge boundaries. Several distribution forms were attempted to fit the data, and it was found that the mixed exponential-lognormal distribution results in the best match. Figure 3 shows the modeled groundwater age distribution and the fitted exponential-lognormal distribution. To test the performance of the method, we evaluated how the Bayesian approach estimates the parameters of a presumed exponential-lognormal form for the groundwater age distribution. Also, a presumed single exponential distribution for the groundwater age was tested in order to evaluate the effect of the uncertainties on the capability of the Bayesian method to discriminate between models in terms of capturing the observed tracer concentrations. Figure 4 shows the posterior distribution of the following model parameters: the mean of the exponential (young) part of the distribution, μ_1, the fraction of young groundwater represented by the exponential distribution, f, the location factor of the lognormal distribution representing the old fraction of groundwater, μ_2 and the estimated observation error geometrical standard deviation σ respectively for both cases of one and five samples. As it can be seen for the case of one sample the method is unable to narrow down the credible intervals enough to provide any useful information about the age distribution parameters. Also as it is seen in Figure 4d the method cannot estimate the geometrical standard deviation for observation error with high confidence. This is expected since in order to estimate the observed errors as a result of spatial and temporal heterogeneities in the tracer concentration for each tracer, one sample is not adequate. When five samples are used the method does a much better job in estimating the presumed parameters of age distribution. In Figure 5, the posterior distributions of the tracer concentrations for four of the tracers, CFC-12, ^3H, ^{14}C, and ^4He, are presented. Again for the case of a single sample the posterior distributions of tracer concentrations are spread due to the fact that the standard deviation of the observation error is over estimated. For the case of five samples, in all cases the method is capable of reproducing the observed concentrations successfully and with a narrow credible interval.

[21] To compare the models in terms of their ability to reproduce the measured data, the Bayes factor method

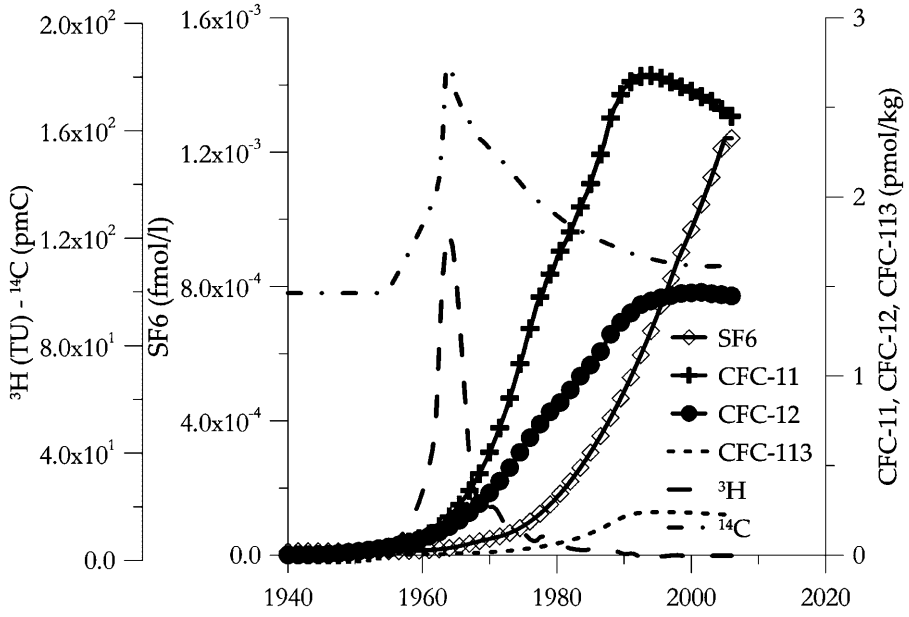

Figure 2. Historical record of the atmospheric concentrations of the tracers used for both the hypothetical aquifer and the La Selva site [*IAEA*, 2011; *Plummer and Busenberg*, 2006] (http://isohis.iaea.org; ftp://cdiac.ornl.gov/pub/ndp057; http://cdiac.ornl.gov/trends/co2/cent-scha.html).

[*Jeffreys*, 1935; *Kass and Raftery*, 1995] was used. The Bayes factor for comparing two models M_1 and M_2 assuming an equal prior probability for the models is defined as

$$B_{12} = \frac{\int_{\Theta_1} p(\hat{C}|\mathbf{X}_1, M_1)p(\mathbf{X}_1)}{\int_{\Theta_2} p(\hat{C}|\mathbf{X}_2, M_2)p(\mathbf{X}_2)}. \quad (20)$$

[22] B_{12} represents the ratio of the odds that model 1 is the true model to the odds that model 2 is the true model, $p(\hat{C}|\mathbf{X}, M)$ is the likelihood of observing concentrations $\hat{C}$ given model M, and $p(\mathbf{X})$ is the prior density of the parameters. For comparing more than two models, a measure of capability of each groundwater age distribution to reproduce the observed data (posterior odds of each model), I_s, is defined as

$$I_s = \frac{\int_{\mathbf{X}_s} p(\hat{C}|\mathbf{X}_s, M_k)p(\mathbf{X}_s)}{\sum_{l=1}^{m} \int_{\Theta_l} p(\hat{C}|\mathbf{X}_l, M_l)p(\mathbf{X}_l)} \quad (21)$$

where m is the total number of age distribution forms considered. The integrals of posterior distribution are estimated using the Monte Carlo method [*Carlin and Chib*, 1995]:

$$\int_{\Theta_s} p(\hat{C}|\mathbf{X}_s, M_s)p(\Theta_s) \approx \frac{1}{n}\sum p(\hat{C}|\mathbf{X}_s^{(k)}, M_s)p\left(\mathbf{X}_s^{(k)}\right) \quad (22)$$

[23] In addition to the capability of an age distribution to reproduce the observed data, another factor in evaluating the goodness of different forms of age distributions in a practical sense is how much information they provide about the age distribution parameters. If the number of parameters of a model is large, it can better match the observed data due to a larger degree of flexibility. However, using such a model may lead to unidentifiable model parameters or to strong posterior internal correlations between them and therefore wide credible intervals for the posterior model parameters (henceforth what is meant by identifiability or lack thereof is the ability of the model to update the prior distributions of the parameters to a practically useful level as a result of the data). Therefore, the level of information each model provides can be considered to be inversely proportional to the coefficient of variation (CV) of model parameters (in this case age distribution parameters). Due to the difference in the number of parameters of different age distributions

Table 1. Modeled Tracer Concentrations Using the Hypothetical Aquifer Model, Without Noise Added To The Parameters and With Random Noise According to the Prior Distributions of Parameters and the Obtained Tracer Concentrations

Sample	SF_6 (fmol/L)	CFC-11 (pmol/kg)	CFC-12 (pmol/kg)	CFC-113 (pmol/kg)	^{3}H (TU)	^{14}C (pm C)	^{4}He (cm^3 STP/g)	δ^{13}C (‰)
1	1.32×10^{-41}	2.39×10^{-1}	1.41×10^{-1}	2.80×10^{-2}	0.853	95.2	1.46×10^{-7}	−1.81
2	1.24×10^{-4}	2.35×10^{-1}	1.47×10^{-1}	2.75×10^{-2}	0.825	82.0	1.35×10^{-7}	0.767
3	1.40×10^{-4}	2.46×10^{-1}	1.50×10^{-1}	2.59×10^{-2}	0.871	85.9	1.46×10^{-7}	−1.42
4	1.34×10^{-4}	2.27×10^{-1}	1.49×10^{-1}	2.60×10^{-2}	0.838	81.4	1.41×10^{-7}	−9.82
5	1.34×10^{-4}	2.32×10^{-1}	1.64×10^{-1}	2.77×10^{-2}	0.828	85.7	1.40×10^{-7}	−2.00

Table 2. Parameter Expected Values and Standard Deviations Used in the Bayesian Dating

	Distribution	SF_6	CFC-11	CFC-12	CFC-113	^3H	^{14}C	^4He
Prior decay rates (yr^{-1})	Lognormal							
Geometric mean μ_λ[a]		0.00	3.21×10^{-1}	5.70×10^{-2}	3.00×10^{-2}	5.60×10^{-2}	1.20×10^{-4}	
Geometric SD σ_λ[b]		1.41	1.41	1.41	1.41	0.0	0.0	
Accumulation rate	Lognormal							
Geometric mean μ_λ								2.77×10^{-10} cm^3 STP/g/yr
Geometric SD σ_λ[b]								1.41
K_{OC}[c] ($g\ g^{-1}$)	Deterministic	195	97	356	316			0
Log K_{OW}[d]		0.226	2.53	2.16	3.16		K_D = 83 mL/g	0
Prior soil organic content f_{oc}	Lognormal							
Geometric mean μ_{foc}						0.001		
Geometric SD σ_{foc}						1.41		
Henry's law constants k_H[d] (mols/kg)	Deterministic	2.2×10^{-4}	1.3×10^{-2}	3.55×10^{-3}	3.08×10^{-3}			
Ambient aqueous concentration	Deterministic							4.1×10^{-8}

[a]Using $\ln(2)/t_{1/2}$, where $t_{1/2}$ is the half-life obtained from *Hinsby et al.* [2007].
[b]Assumed.
[c]Hazardous substance data bank (http://toxnet.nlm.nih.gov).
[d]*Plummer and Busenberg* [2006].

evaluated here, the CVs only for the overall mean age of the groundwater are shown and used for comparing the models. The Bayes factors and the CV values obtained for the cases of presumed exponential-lognormal and single exponential model are presented in Table 4 for the cases of five and one samples. Table 4 summarizes the Bayes factors and the coefficient of variations for the presumed exponential and exponential-lognormal groundwater age distributions for the cases of one and five samples. In both cases of one and five samples the method can identify the right form of the distribution as the odds of the exponential-lognormal distribution are significantly larger than that the one for the exponential distribution. The method however identifies the correct model with a much better confidence when 5 samples are used. The coefficients of variations of the age of the young fraction of groundwater are much larger for the case of one sample indicating that the method is unable to narrow down the age distribution parameters. Also although the exponential model is simpler (it has two less parameters than the exponential-lognormal model), the coefficients of variation obtained for it are larger than the exponential-lognormal model.

5. Example Application

[24] The method was applied to four samples from the data collected by *Solomon et al.* [2010] and *Webb* [2007] at La Selva Biological Station in the Costa Rican rain forest. La Selva Biological Station is located on the Caribbean coastal plain at the foot of Volcan Barva. Annual precipitation ranges from 4240 mm at the study site, to more than 8000 mm at an elevation of 700 m [*Solomon et al.*, 2010]. There is strong evidence that groundwater flow at this site consists of two distinct end-members: (1) high solute bedrock groundwater representing interbasin groundwater flow and (2) low-solute groundwater derived from recharge that falls within the drainage area of the Sura and Salto streams. The main source of dissolved inorganic carbon (DIC) is magmatic outgassing. There is strong evidence that dissolution of carbonate minerals does not have a significant contribution in DIC [*Genereux et al.*, 2009]. The site characteristics and sampling procedures have been extensively described in *Genereux et al.* [2009], *Solomon et al.* [2010], and *Webb*

Table 3. The $\delta^{13}C$ Values and Their Uncertainty for Various Contributing Carbon Sources and the Observed Values

	Average $\delta^{13}C$ (‰)	SD (‰)
Magmatic[a]	−2.4%	±4
Biogenic[b]	−25%[b]	±4

[a]*Webb* [2007].
[b]*Deines* [1980].

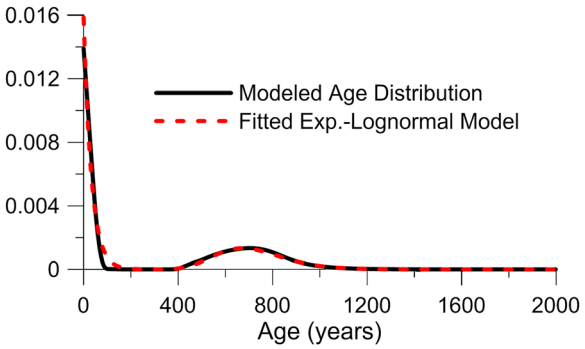

Figure 3. Modeled groundwater age distribution and the fitted exponential-lognormal curve for the hypothetical aquifer: $\rho(a) = f\frac{1}{\mu_1}\exp\left(\frac{a}{\mu_1}\right) + (1-f)\frac{1}{a\sqrt{2\pi\delta^2}}\exp\left(-\frac{\ln(a)-\mu_2}{2\delta^2}\right)$. The parameters were found to be $f = 0.5237$, $\mu_1 = 33.04$ years, $\mu_2 = \ln(702.1$ years), and $\delta = 0.208$.

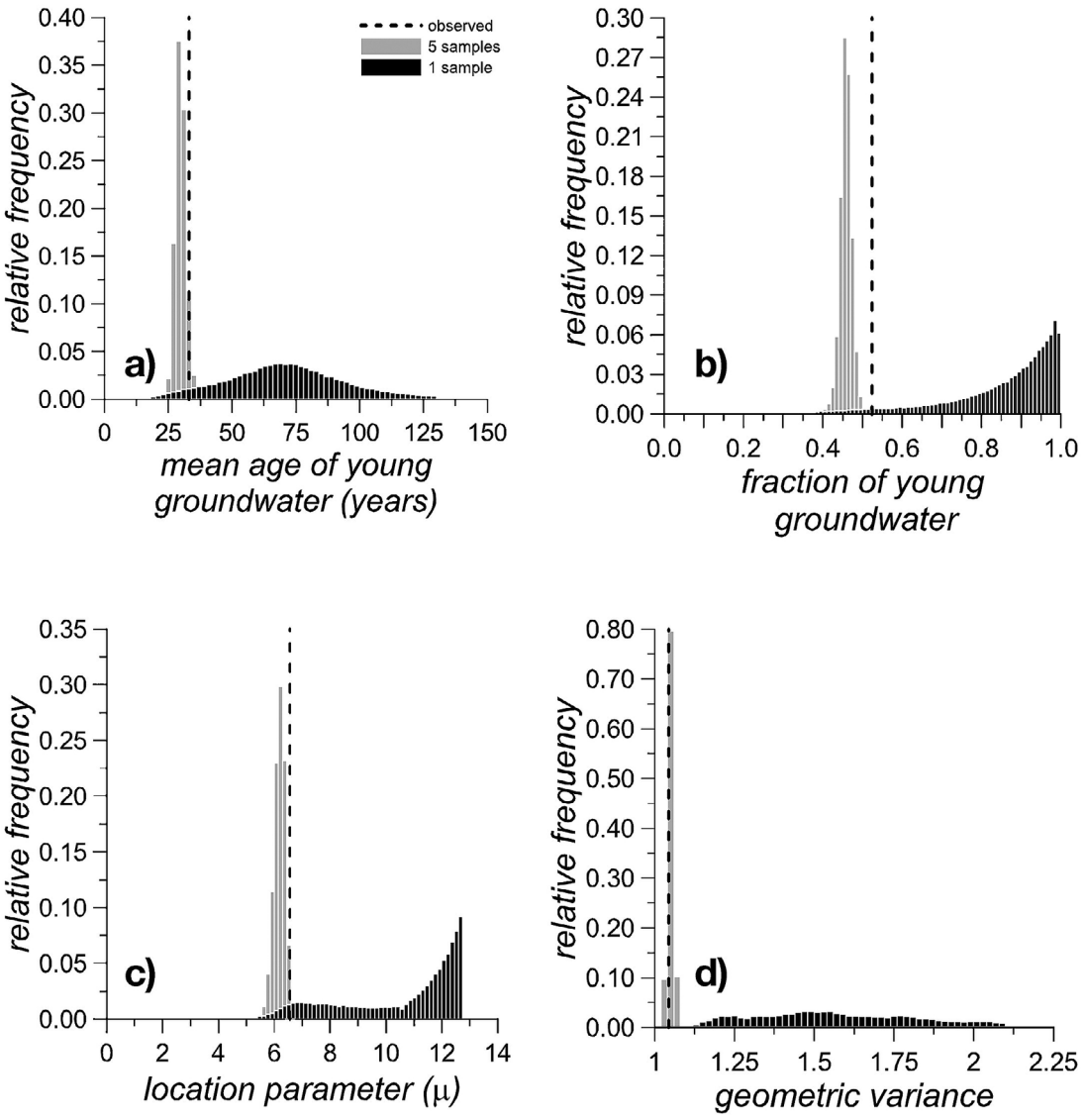

Figure 4. Posterior distributions for (a) mean age of young fraction of groundwater μ_1, (b) fraction of young groundwater f, (c) location parameter for the lognormal distribution μ_2, and (d) geometric variance of observation error for the case of five and one hypothetical samples at the recharge point. The dashed vertical lines represent the values obtained from fitting exponential-lognormal distributions to the data (std: geometrical standard deviation of observation error).

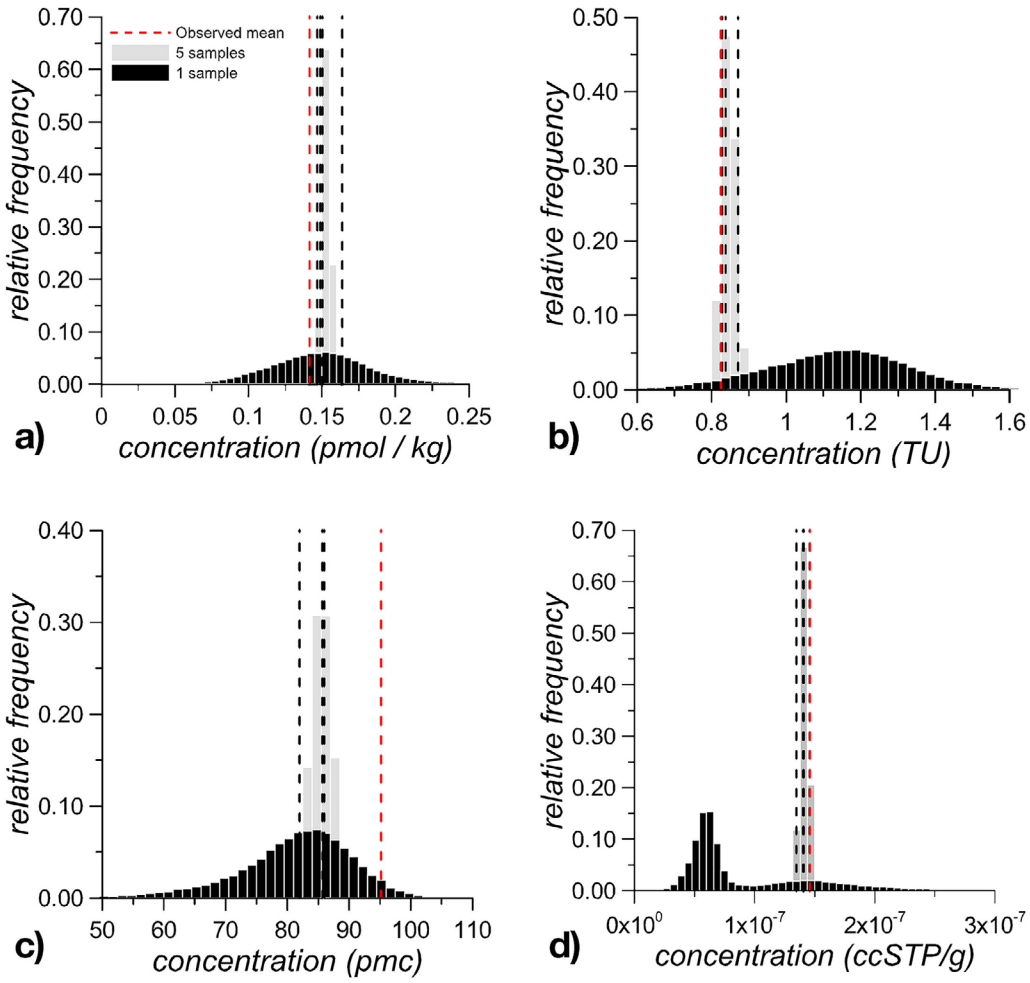

Figure 5. Posterior distributions for concentrations of (a) CFC-12, (b) ^{3}H, (c) ^{14}C, and (d) ^{4}He obtained from the exponential-lognormal distribution for the hypothetical case for a single and five samples at each well. The vertical dashed lines represent the value of tracer concentrations obtained from numerical modeling of the hypothetical aquifer, and the gray vertical bars show the value used in the single sample scenario.

[2007]. The following tracers were used: ^{14}C, CFC-11, CFC-12, CFC-113, ^{3}H, SF_6, and ^{4}He. Among these tracers, ^{4}He is the only one that accumulates in the groundwater as a result of decay of radionuclides including mainly ^{238}U, ^{232}Th, and ^{235}U. The parameters used in the Bayesian modeling are summarized in Table 2. The analysis was performed on four of the samples collected, that is, samples from wells identified by numbers 7, 11, 16, and 30. The observed concentrations of the tracers at the four sampling locations are listed in Table 5. Although multiple samples were collected each well, the variance in measured concentrations represents mostly analytical errors and does not address the representativeness of the samples to a specific location and therefore the standard deviation of observed tracer concentrations due to measurement error and spatial and temporal heterogeneity could not be estimated directly for every well. Therefore, the method was tested with four different assumed geometrical standard deviations (i.e., exponential function of the shape factor) for the lognormal distributions representing the model parameters: the decay rates of CFCs, the accumulation rate of ^{4}He (i.e., $\sigma_{\lambda,i}$), the soil organic content, f_{oc} (i.e., σ_{foc}), and errors in the observed data (σ). It should be noted that the geometric standard deviation of $\hat{\mathbf{C}}$, σ can also be considered a random variable; however, as it was demonstrated in the hypothetical demonstration, without multiple samples the variability of which represents the entire range of uncertainty as a result of heterogeneity and structural error, the method will not be able to narrow down the ranges of the age distribution parameters adequately. Therefore, here we assumed a fixed value for σ

Table 4. Relative Bayes Factors and Coefficient of Variation (CV) for Exponential and Exponential-Lognormal Distributions of One and Five Sample Cases for the Hypothetical Aquifer Shown in Figure 1

	One Sample		Five Samples	
	I_s	CV	I_s	CV
Exponential	0.0221	0.6317	4.90E-27	0.0947
Exponential-lognormal	0.9779	0.3755	1.0000	0.0665

Table 5. Observed Tracer Concentrations and Their Uncertainty for Various Contributing Carbon Sources and the Observed Values at the La Selva Site in Costa Rica

	Observed Concentration Geometrical Mean[a]							
	SF_6 (fmol/L)	CFC-11 (pmol/kg)	CFC-12 (pmol/kg)	CFC-113 (pmol/kg)	^3H (TU)	^{14}C (pm C)	^4He (cm^3 STP/g)	$\delta^{13}C$ (‰)
Well 7	1.100	3.020	1.660	0.257	0.61	117.1	4.81×10^{-8}	−26.00[b]
Well 11	0.169	0.446	0.138	0.0172	0.36	21.7	2.50×10^{-7}	−7.58
Well 16	0.931	3.020	1.620	0.243	0.72	116.9	5.06×10^{-8}	−24.34
Well 30	1.180	3.130	1.450	0.187	0.66	83.4	7.41×10^{-8}	−20.20

[a] *Webb* [2007].
[b] *Genereux et al.* [2009].

based on the known analytical errors of chemical analysis and temporal variations of tracer concentrations. Three values of geometric standard deviation σ, of 2.0, 1.41, and 1.22, representing a 95% probability that the true mean tracer concentration would vary respectively between one fourth and four times, one half and two times, and 1/1.5 and 1.5 times the measured concentration, were tested. This was done to evaluate the sensitivity of the results with respect the uncertainty associated with the observed tracer concentrations. No evidence for biodegradation of SF_6 was found in the literature, and therefore its biodegradation rate coefficient was assumed to be zero. CFCs have been shown to undergo biodegradation under anaerobic conditions [*Lovley and Woodward*, 1992; *Oster et al.*, 1996; *Sebol et al.*, 2007; *Sonier et al.*, 1994]. However, most of the values reported in the literature for the decay of CFCs have been measured under fully anaerobic conditions. We estimated the CFC decay rates based on the values in the literature compiled by *Hinsby et al.* [2007] and the estimated values are presented in Table 2. The rate of accumulation of ^4He was estimated using the observed ^4He, ^{14}C values in a regional spring sample (Gaucimo Spring) that was shown by *Genereux et al.* [2009] to contain only minimal amounts of young water. Due to the large uncertainty about the values a geometrical standard deviation of 1.41 was assumed for both decay rates of CFCs and the accumulation of ^4He, indicating that the actual decay/accumulation rates vary between one half and two times the specified values, with a probability of 95%. No uncertainty was assumed for the decay rates of ^{14}C and ^3H. The retardation factor for CFCs and SF_6 was obtained using equation (11), assuming that the sorption of CFCs mainly takes place to organic carbon. The K_{oc} values were obtained from the hazardous substance data bank (HSDB) (http://toxnet.nlm.nih.gov/) and were considered deterministic, and all the uncertainty associated with the sorption-desorption process was associated to the uncertainties in the soil organic content f_{oc}. The partitioning coefficient K_D for ^{14}C was obtained from *Plummer et al.* [2004] (adopted from *Allard et al.* [1981]). The soil organic content was assumed to be 0.1% with a geometrical standard deviation of 1.41. The Henry's constants were calculated using the formulas suggested by *Warner and Weiss* [1985]. For calculation of Henry's law constants, the temperature was assumed constant and equal to the average temperature in the study site (23.6°C), and the pressure was estimated based on the altitude of the site. Figure 1 presents the aqueous equilibrium concentrations of the tracers used, calculated based on Henry's law and atmospheric concentrations. Table 3 contains the ^{13}C isotopic ratios of biogenic and magmatic sources. A standard deviation of 4% is assumed for the measured as well as the biogenic and magmatic ^{13}C values based on [*Deines*, 1980; *Genereux et al.*, 2009; *Webb*, 2007].

6. Results and Discussion

[25] This section discusses the results of the application of the method to the data collected at the La Selva Biological Station.

[26] Nine forms of age distribution were tested, four single distributions, exponential, Dirac delta, inverse Gaussian [*Chhikara and Folks*, 1989], and gamma, and five mixed distributions (i.e., with PDFs comprising the additions of two functional forms) were tested, double–Dirac delta, exponential-Dirac, double-exponential, gamma-Dirac, and gamma-exponential:

Dirac delta

$$\rho(a) = \delta(a - \beta_1) \tag{23a}$$

Exponential

$$\rho(a) = \frac{1}{\beta_1} e^{-\frac{a}{\beta_1}} \tag{23b}$$

Gamma

$$\rho(a) = a^{\beta_2 - 1} \frac{e^{-a/\beta_1}}{\Gamma(\beta_2)\beta_1^{\beta_2}} \tag{23c}$$

Inverse Gaussian

$$\rho(a) = \left(\frac{\beta_2}{2\pi a^3}\right)^{1/2} e^{\frac{-\beta_2(a-\beta_1)^2}{2\beta_1^2 a}} \tag{23d}$$

Double-Dirac

$$\rho(a) = f\delta(a - \beta_1) + (1 - f)\delta(a - \beta_2) \tag{23e}$$

Double-exponential

$$\rho(a) = f\frac{1}{\beta_1} e^{-\frac{a}{\beta_1}} + (1 - f)\frac{1}{\beta_2} e^{-\frac{a}{\beta_2}} \tag{23f}$$

Exponential-Dirac

$$\rho(a) = f\frac{1}{\beta_1} e^{-\frac{a}{\beta_1}} + (1 - f)\delta(a - \beta_2) \tag{23g}$$

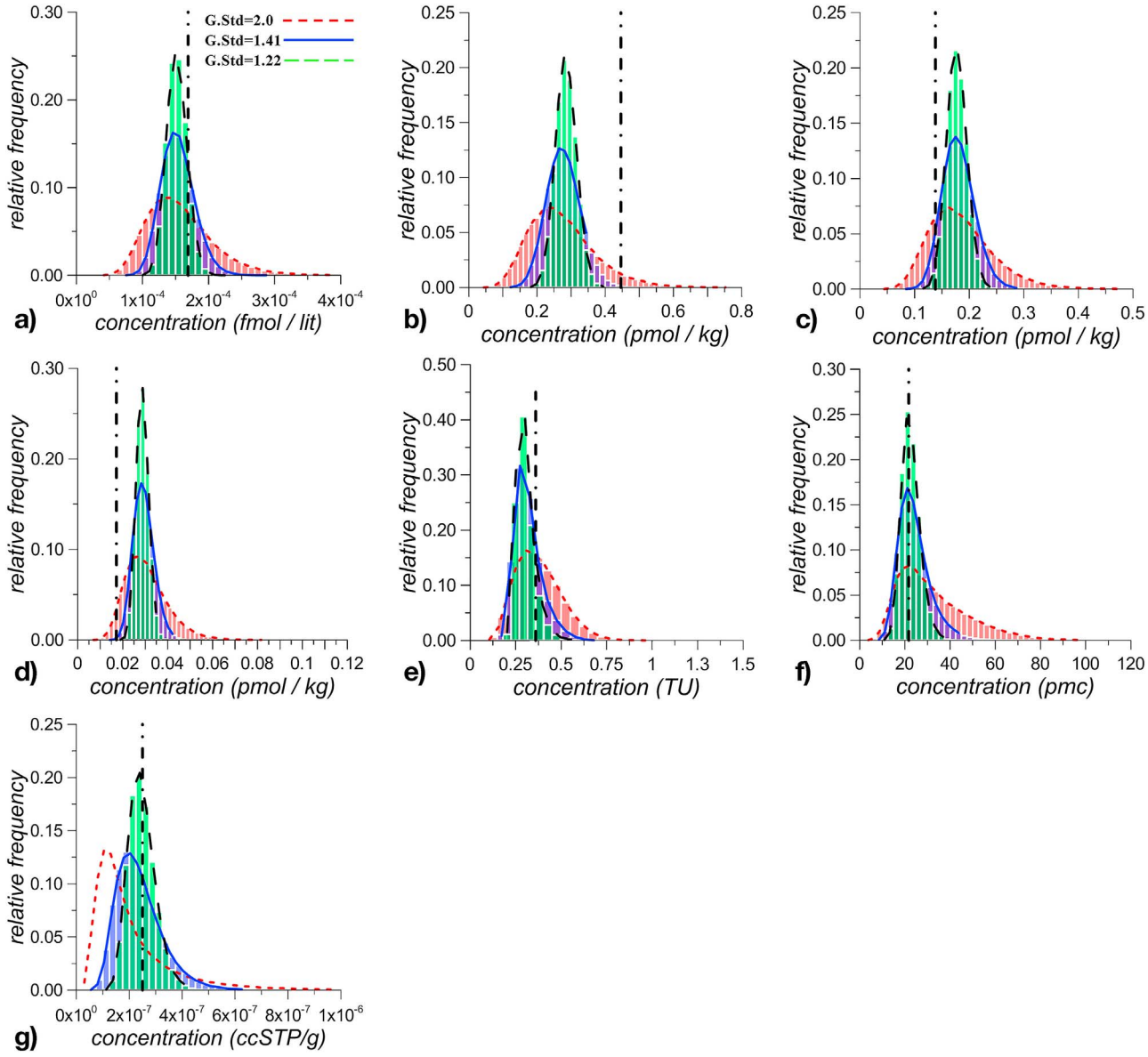

Figure 6. Concentration histograms for well 11 (La Selva site in Costa Rica) and exponential-gamma age distribution for (a) SF_6, (b) CFC-11, (c) CFC-12, (d) CFC-113, (e) ^3H, (f) ^{14}C, and (g) He.

Gamma-Dirac

$$\rho(a) = f a^{\beta_2 - 1} \frac{e^{-a/\beta_1}}{\Gamma(\beta_2)\beta_1^{\beta_2}} + (1-f)\delta(a-\beta_3) \tag{23h}$$

Gamma-exponential

$$\rho(a) = f \frac{1}{\beta_1} e^{-\frac{a}{\beta_1}} + (1-f) a^{\beta_2 - 1} \frac{e^{-a/\beta_3}}{\Gamma(\beta_2)\beta_3^{\beta_2}} \tag{23i}$$

Here β_1 through β_4 are the parameters of the age distribution functions, and f is the ratio of young or old groundwater in the case of mixed age distributions, depending on whether the mean of the first or second term is larger. The inherent assumption when using mixed distributions is that the groundwater consists of distinct old and young fractions.

[27] Figure 6 shows the predicted concentrations of the seven tracers for the case in which the exponential-gamma age distribution was used for well 11 and for the three standard deviations of observed tracer concentrations. As expected, a smaller standard deviation for the observed tracer concentrations results in a smaller spread of the predicted tracer concentration distribution. Generally, the model successfully reproduces the observed tracer concentrations except for CFC-11, which are underestimated by the model. Figure 7 shows the following posterior distributions of the main parameters of the exponential-gamma age distribution: (a) β_1, the mean age of the young portion of groundwater represented by an exponential distribution, (b) f, the fraction of young groundwater, and (c) $\beta_2\beta_3$, the mean age of old groundwater. For the medium geometrical standard deviation $\delta_{c,i}$ = 1.41, the model estimates a young groundwater age of between 0 and 350 years with the largest odds being

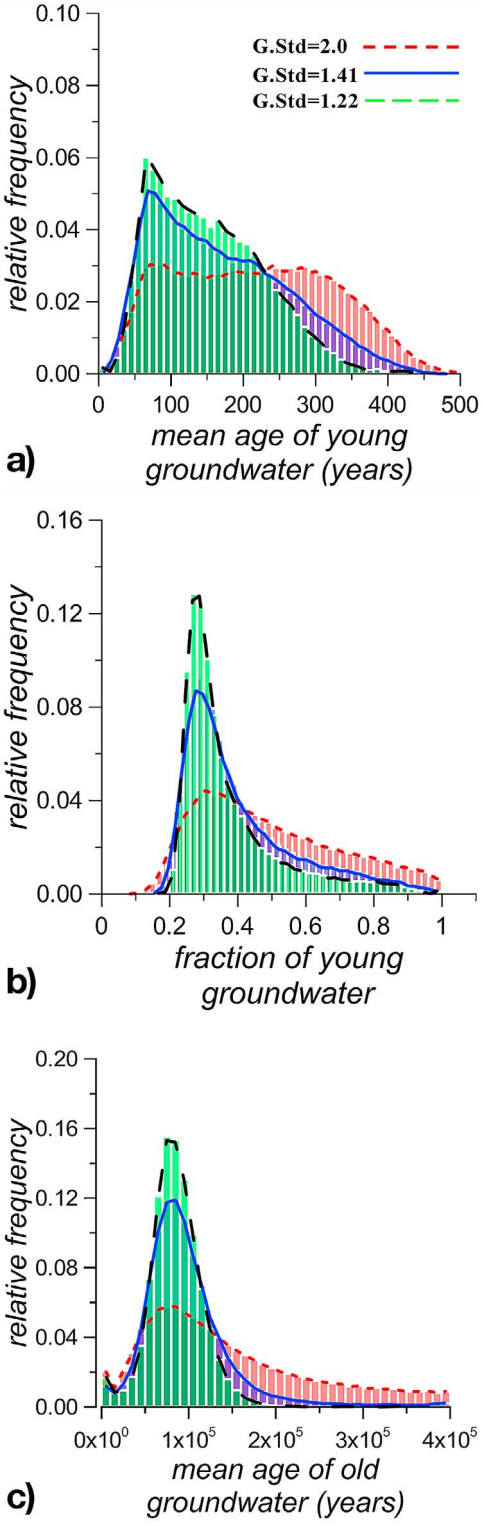

Figure 7. Posterior frequency distributions for well 11 (La Selva site Costa Rica) of the model parameter for exponential-gamma age distribution (a) mean age of young groundwater, (b) fraction of young groundwater f, and (c) mean age of old groundwater.

given to the age of mean groundwater of approximately 100 years and a fraction of young groundwater of between 15% and roughly 60% with the largest odds of 30% (Figure 7b). However, the method cannot definitively rule out larger fractions of young groundwater. The model provides a much more uncertain estimation of the age of old groundwater with a credible interval of between 20,000 and 200,000 years (Figure 7c). This is due to the fact that five out of the seven tracers used in this study were introduced into the environment less than 60 years ago, as well as to the ineffectiveness of ^{14}C in reducing the confidence interval for the age of very old groundwater as a result of the insensitivity of its concentration to age (1) when the ages are several orders of magnitude larger than its half-life and (2) because of the uncertainties associated with the contribution of mineral dissolution (f_m). For well 11, the method is unable to narrow down the acceptable range of ^{14}C contribution from mineral dissolution. This is due to the relatively large standard deviations considered for the ^{13}C concentrations associated with biogenic and magmatic sources of carbon (i.e., 4‰). Therefore, the only tracer that effectively plays a role in determination of the age of the old groundwater was ^{4}He. The overall I_s and CV values for all the wells, and the three observed tracer concentration geometrical standard deviations, are summarized in Table 6. The comparative performance of different models depends on the samples and to some degree on the assumed observed tracer concentration's geometrical standard deviation, $\sigma_{c,i}$.

[28] From the two measures of goodness of models including the I_s and CV, it is clear that for well 11 mixed age distributions perform significantly better than single distributions. This is a strong indicator of the presence of a mixture of young and old groundwaters in the sample taken from this well. This conclusion is confirmed by the significantly higher concentrations of ^{4}He with respect to the ambient atmospheric concentrations and the lower ^{14}C content observed in well 11 (Table 1). For wells 7, 16, and 30, there is not a significant difference between the posterior odds of models for single and mixed age distributions. This indicates that in order to describe the observed tracer concentrations, the presence of old and new fractions of groundwater is unnecessary, and that the method cannot definitively support the fact that the sample collected from these wells contains distinct old and young fractions. The reason for the lack of certainty provided by the method regarding the presence of old fractions of groundwater in the samples from wells 7, 16, and 30 is the uncertainties associated with the accumulation rates of ^{4}He and the lack of any prior knowledge for the age of the old water. However, a better confidence can be gained by using ^{4}He or major ion tracers (e.g., Cl) just as indicators of the fraction of old groundwater by incorporating the known contents of these tracers in old water directly into the analysis. However, doing so will render the method incapable of providing any information about the age of the old fraction of groundwater which may be practically unimportant.

[29] In terms of the coefficients of variation, generally a smaller number of parameters results in smaller CV values. Therefore, the Dirac delta distribution results in the smallest CV for most cases. This is an expected outcome due to the fact that a smaller degrees of freedom results in a smaller chance of internal correlation between the posterior distribution of parameters. Table 7 summarizes geometrical means of the posterior odds of each model separately for well 11 and wells 7, 16, and 30. For well 11, the gamma-exponential model provides the best posterior odds, whereas for wells 7, 16, and 30 collectively the inverse Gaussian

Table 6. I_s Values and Coefficient of Variation (CV) for Four Wells at the La Selva Site in Costa Rica and Three Standard Deviations

	Well 11		Well 16		Well 30		Well 7	
	I_s	CV	I_s	CV	I_s	CV	I_s	CV
Observed Tracer Concentration Geometrical SD = 2.0								
Exponential-Dirac	2.30E-01	0.644	1.09E-01	1.003	1.34E-01	0.979	1.18E-01	1.027
Dirac-Delta	3.01E-03	0.297	1.71E-01	0.747	1.56E-01	0.747	1.78E-01	0.734
Exponential	6.29E-03	0.423	1.57E-01	0.857	1.51E-01	0.846	1.69E-01	0.838
Two-Dirac	4.49E-02	1.754	1.14E-01	0.975	1.16E-01	1.032	1.21E-01	0.999
Inverse Gaussian	3.14E-02	0.338	1.77E-01	0.737	2.07E-01	0.699	1.90E-01	0.769
Two-Exponential	1.42E-01	0.786	1.10E-01	1.004	1.21E-01	0.987	1.12E-01	1.039
Gamma	8.41E-02	0.230	1.95E-02	0.384	3.59E-02	0.359	1.62E-02	0.383
Gamma-Dirac	1.03E-01	0.614	1.21E-01	0.950	5.58E-02	1.067	7.66E-02	1.331
Exponential-gamma	3.55E-01	0.738	2.13E-02	1.218	2.23E-02	1.095	2.02E-02	1.164
Observed Tracer Concentration Geometrical SD = 1.41								
Exponential-Dirac	1.30E-02	1.649	1.06E-01	0.979	8.03E-02	1.050	1.06E-01	1.003
Dirac-Delta	1.76E-07	0.203	1.49E-01	0.746	9.27E-02	0.781	1.59E-01	0.756
Exponential	3.96E-06	0.222	1.47E-01	0.734	9.67E-02	0.752	1.61E-01	0.742
Two-Dirac	8.96E-02	0.588	1.01E-01	0.983	8.02E-02	1.073	1.12E-01	1.052
Inverse Gaussian	5.07E-04	0.324	1.62E-01	0.704	2.09E-01	0.603	1.58E-01	0.743
Two-Exponential	2.14E-01	0.367	1.13E-01	1.009	1.54E-01	0.968	1.11E-01	0.997
Gamma	1.15E-02	0.142	5.03E-03	0.658	1.87E-03	0.446	5.37E-03	0.758
Gamma-Dirac	3.81E-03	1.988	1.25E-01	0.858	2.83E-01	0.729	6.38E-02	1.046
Exponential-gamma	6.67E-01	0.527	9.12E-02	2.680	1.80E-03	1.175	1.24E-01	0.989
Observed Tracer Concentration Geometrical SD = 1.22								
Exponential-Dirac	2.75E-03	0.459	9.82E-02	0.971	1.80E-01	0.836	9.74E-02	1.004
Dirac-Delta	2.19E-18	0.039	1.48E-01	0.629	5.88E-03	0.723	1.46E-01	0.651
Exponential	4.95E-11	0.134	1.33E-01	0.638	6.32E-03	0.659	1.32E-01	0.637
Two-Dirac	4.36E-01	0.451	1.02E-01	0.996	1.33E-01	1.655	1.01E-01	1.002
Inverse Gaussian	6.83E-08	0.321	1.16E-01	0.655	6.64E-02	0.469	9.29E-02	0.699
Two-Exponential	2.77E-02	0.257	1.20E-01	0.944	1.96E-01	0.876	1.08E-01	0.955
Gamma	2.17E-04	0.179	6.52E-02	0.934	1.33E-03	1.193	7.76E-02	0.629
Gamma-Dirac	4.58E-01	0.405	1.51E-01	0.781	3.52E-01	0.548	1.27E-01	0.796
Exponential-gamma	7.49E-02	0.429	6.64E-02	1.107	5.93E-02	1.105	1.18E-01	0.884

model results in the best posterior odds, although the rest of the age distributions are not significantly worse. The superiority of the inverse Gaussian model can be interpreted as the dominance of an advection-dispersion process governing the transport of water molecules in such a way that the transport can be described as a stream tube conceptualization with small lateral mixing. The CV of the mean age associated with the inverse Gaussian model, although relatively small, is still significantly larger than the gamma distribution. It should be noted that the performance of age distributions based on both CV and I_s seems to be case dependent, and in particular depends on the uncertainties associated with the observed concentrations, other model parameters, their values, and the number of tracers used, and no single age distribution can be suggested to perform best in all cases in this study. We further interpret this to mean that some of the wells do not actually sample, in a flow-weighted sense, the entire spectrum of travel times. This underscores the need to develop sample collection methodologies that provide a flow-weighted sampling of all flow paths. Figure 8 summarizes the I_s and CV values for well 11, the three different observed error geometrical standard deviations and all the nine age distributions in a graphical form. Different age distributions forms are sorted from simple (smaller number of parameters) to more complicated on the horizontal axis. By and large the Bayes factor increases with the complexity of age distribution form with some exceptions. For example, the gamma distribution performs worse than both two-Dirac and two-exponential although they have the same number of parameters, which can be interpreted as a confirmation the groundwater consists of distinct old and young fractions. Also the two-Dirac form has a better Bayes Factor compared to two exponential since it better represent young and old fractions of groundwater. As it is expected the CV increases with the complexity of the model particularly when larger values of σ are assumed. Such analyses can be used to find a trade-off between the capability of the model to reproduce the observed tracer concentrations and the level of information it can provide.

Table 7. Geometric Mean of the Posterior Odds of Each Model for All Four Wells at the La Selva Site in Costa Rica

	Geometrical Mean of Posterior Odds	
Model	Well 11	Wells 7, 16, and 30
Exponential-Dirac	0.0201	0.112
Dirac delta	1.05×10^9	0.103
Exponential	1.07×10^6	0.100
Two-Dirac	0.121	0.108
Inverse Gaussian	0.0001	0.144
Two-exponential	0.094	0.125
Gamma	0.0059	0.012
Gamma-Dirac	0.0565	0.126
Gamma-exponential	0.2607	0.036

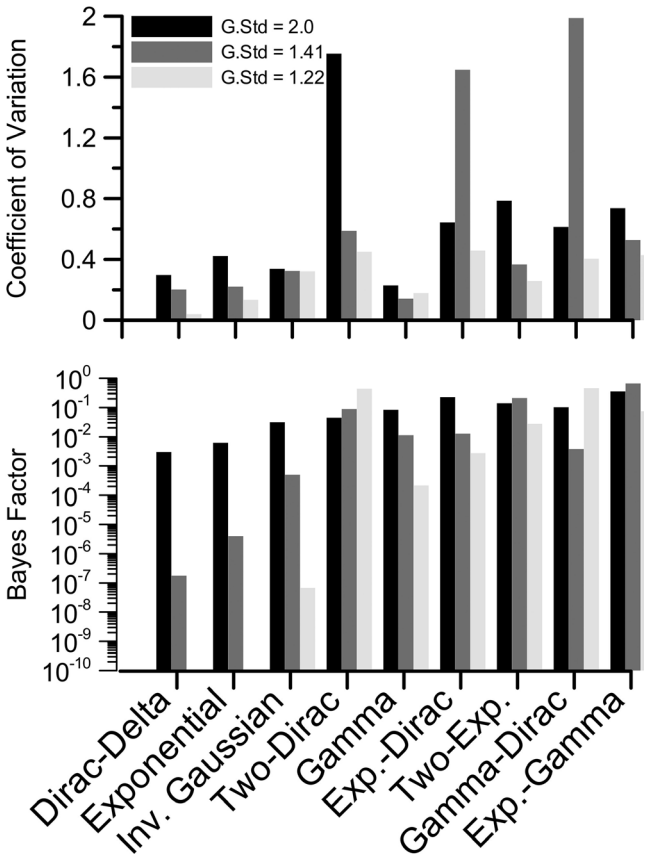

Figure 8. Bayes' factor and coefficient of variation against the assumed levels of uncertainty of the parameters for different distributions.

[30] Figure 9 shows the posterior distributions of mean age of young groundwater for the four wells as a result of four mixed age distribution forms. For well 16, which is deemed to contain a mixture of old and young ages, different models result in somewhat different posterior distributions of the mean age. For other wells, different models result in very close ranges of mean ages of groundwater. Figure 10 shows the contribution of mineral dissolution in carbon for all four wells when the gamma-exponential model is used for well 11 and the exponential age distribution is considered for other wells. The method assigns a significant contribution of dissolved carbon as a result of mineral dissolution. This is supported by the observed ^{13}C values. Figure 11 shows the 95% credible bands (calculated using the equal-tail approach [*Mukhopadhyay*, 2000]) as well as the observed concentrations for all of the tracers and all wells when the gamma-exponential model was used for well 11 and the exponential distribution was used for the rest of the wells. The figure indicates how well each model can explain the observed concentration/isotope ratios of each tracer. For well 11, the gamma-exponential model performs well on all the tracers. For all other wells, the exponential model overestimates the ^{3}H isotope ratio by several folds. We suspect that this might be due to the sharp spike in the record of atmospheric ^{3}H and the fact that this spike was smoothed out significantly due to vadose zone processes (i.e., the input function was measured in the precipitation, but not in water actually recharging the saturated groundwater system).

7. Summary and Conclusions

[31] A Bayesian approach was utilized for estimation of the groundwater age distribution using the measured radioactive isotopes and synthetic chemicals. The method is based on preassumed groundwater age distribution forms. The method then uses the MCMC method to estimate the posterior probability distributions of the parameters defining the assumed distributions based on the uncertainties in measured tracer concentrations and other parameters affecting the transport of tracers in the subsurface. The role of uncertainties in adsorption-desorption, biodegradation of anthropogenic tracers, and the contribution of mineral dissolution in the dilution of ^{14}C signature are incorporated into the model. The method was first tested using a hypothetical case in which the observed concentrations at the discharge region were calculated using forward modeling while incorporating the uncertainties by adding noise to the parameters and the observed concentration. The method was able to infer the parameters of the age distributions and also to select the right model relatively accurately for the cases when five hypothetical samples were assumed to be available at the discharge location representing the variability of tracer concentration as a result of their temporal heterogeneity. However, for the case of a single sample, the ability of the method to infer the age distributions diminishes mainly due to the fact that it is not able to suitably estimate the standard deviation of observed error.

[32] The method was also applied to four samples collected at La Selva Biological Station in Costa Rica. For this case the standard deviation of observation error was considered deterministically. Several functional forms for age distribution were tested, including single distribution functions such as Dirac delta, exponential, gamma, and inverse Gaussian as well as several mixed functional forms consisting of linear combination of two distribution functions representing young and old groundwater. In the case of presumed mixed age distributions, the model in some cases was unable to determine the age of the old fraction with a reasonably small credible interval. This is due to the fact that only one of the tracers used for groundwater dating in this study (i.e., ^{4}He) was appropriate for the estimation of the age of old water and the uncertainties associated with the accumulation rate of this tracer. When there is significant uncertainty in the age of old groundwater (due to either uncertainty in the accumulation rate or mineral dissolution), it may be preferable to use tracers such as ^{4}He or ^{14}C as indicators of the fraction of old groundwater by assigning a probability density function directly to the tracer concentration in the old water.

[33] In order to evaluate the effects of the uncertainties in the measured concentrations of the tracers, four different values of geometrical standard deviations for the observed tracer concentration values were tested. It was found that the credible bands for the groundwater age distribution parameters for young groundwater depends highly on the standard deviations representing the uncertainties in measured tracer concentrations. Due to lack of adequate data to determine some of the parameters affecting the fate and transport of tracers, including the soil organic matter and

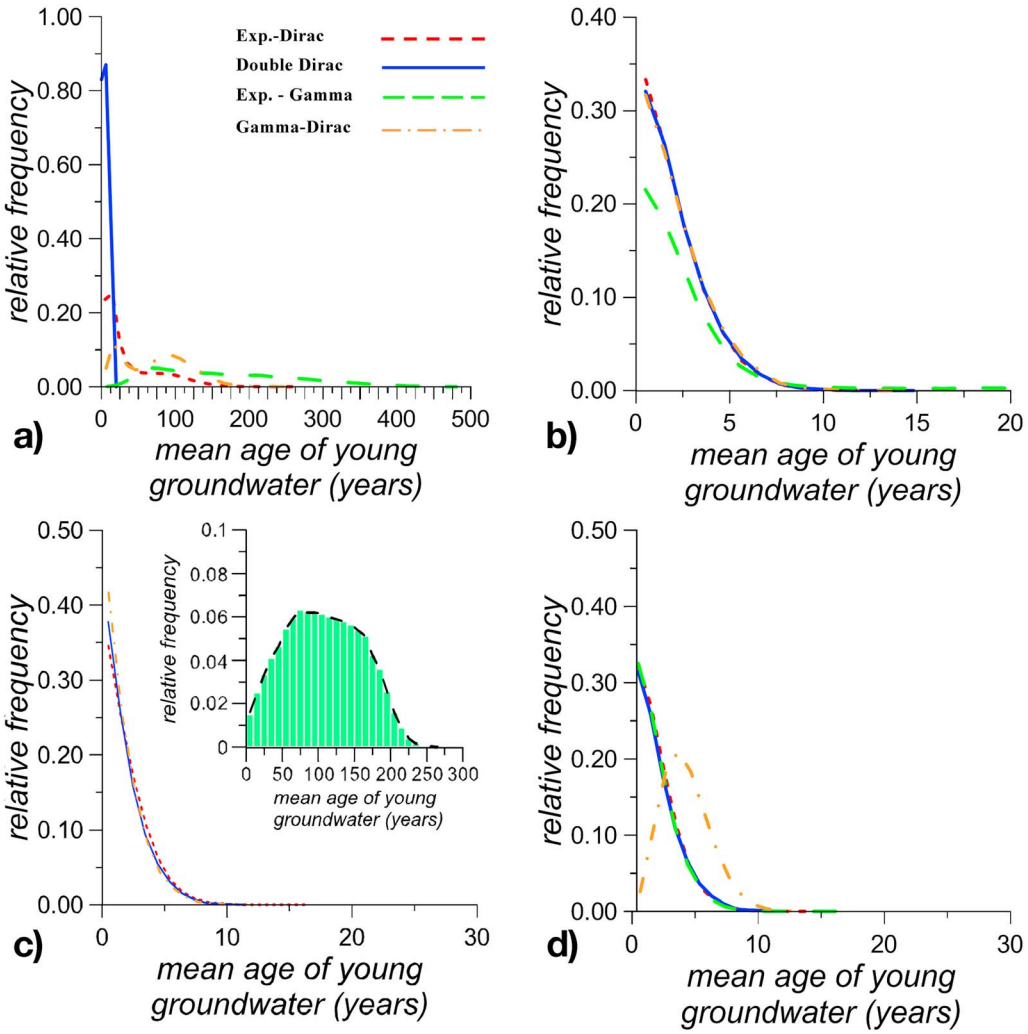

Figure 9. Posterior densities of the mean age of young groundwater for four distributions, including exponential-Dirac, double-Dirac, gamma-Dirac, and gamma-exponential, for (a) well 11, (b) well 16, (c) well 30, and (d) well 7 at the La Selva site Costa Rica.

biodegradation rates of CFCs, large standard deviations were used to represent their uncertainties. This level of uncertainty associated with these parameters resulted in relatively wide confidence interval bands for both young groundwater age parameters and the young fraction of water in cases of mixed age distributions. All of the groundwater age distributions tested here were evaluated in terms of their ability to reproduce the observed tracer concentrations using the Bayes factor method and also the level of confidence at which they estimate the groundwater age distribution as expressed by the coefficient of variations of the posterior distribution of the young fraction of groundwater's mean age.

[34] The performance of groundwater age distributions as measured by these two quantities depends to a large degree on the measured concentration of tracers, and no presumed form of groundwater age distribution can be identified as the one that works best in all cases. This is due in part to the sampling bias, as any given well may not adequately sample a flow-weighted range of all flow paths. However, as is expected, a larger number of the parameters of the presumed age distribution form mostly leads to a wider confidence

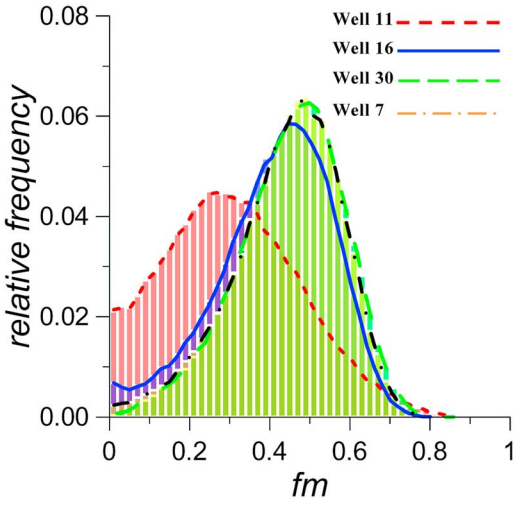

Figure 10. Posterior densities of total dissolved inorganic carbon contributed by mineral dissolution for wells 11, 16, 30, and 7. Exponential-gamma distribution is used for well 11, and exponential distribution is used for wells 16, 30, and 7 at the La Selva site Costa Rica.

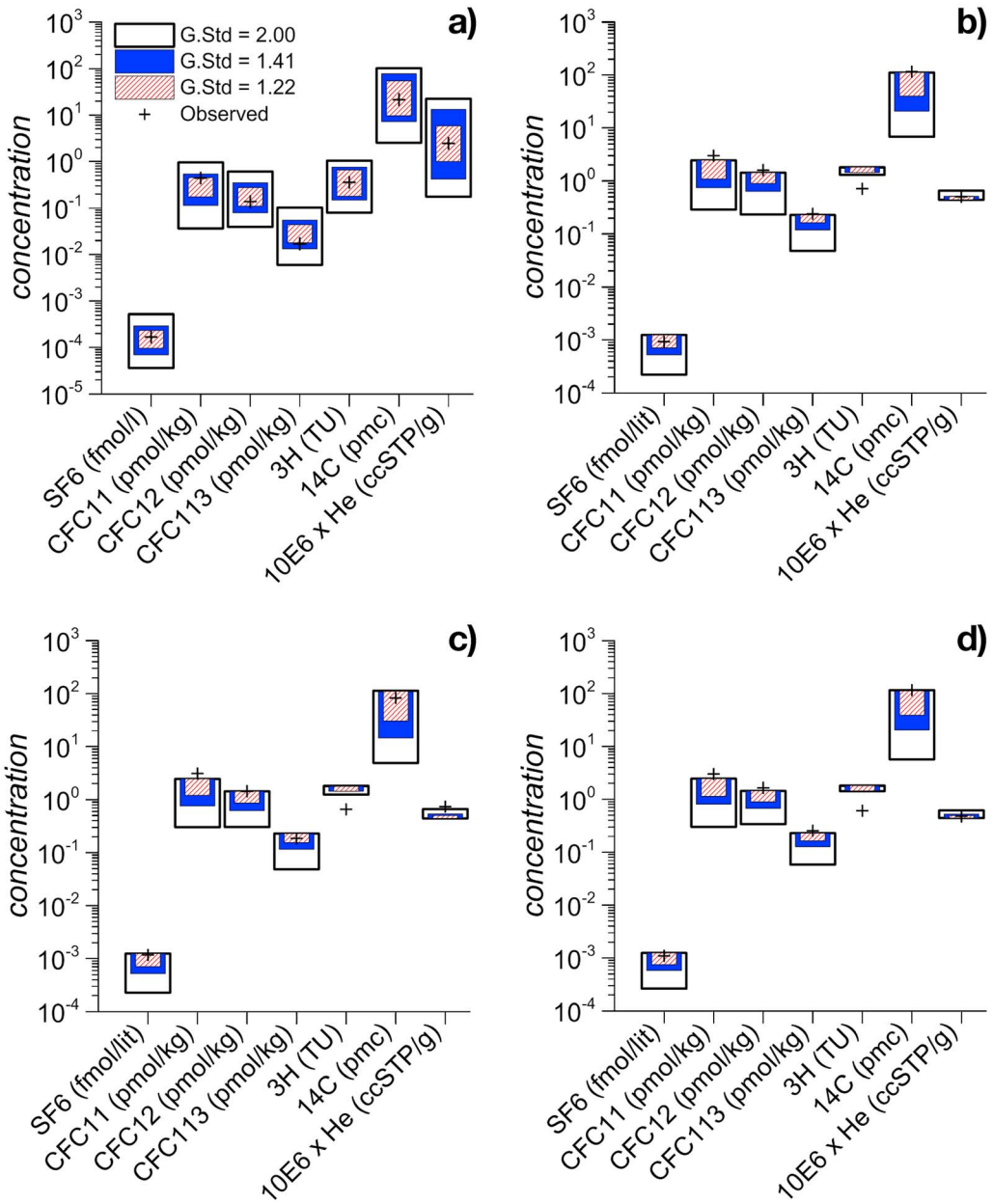

Figure 11. Observed and 95% credible interval brackets of modeled tracer concentrations for (a) well 11 (exponential-gamma), (b) well 16 (exponential), (c) well 30 (exponential), and (d) well 7 (exponential) at the La Selva site Costa Rica.

interval for the parameters defining the age distribution. This is due to the increase in the internal parameter correlations as the number of parameters increases.

[35] **Acknowledgments.** We would like to thank Jennifer McIntosh, Jean Raynald de Dreuzy, and Ming Ye for their constructive comments and suggestions, which led to substantial improvement of the paper.

References

Allard, B., B. Torstenfelt, and K. Andersson (1981), Sorption studies level waste radiological composite analysis of $H^{14}CO_3$ on some geologic media and concrete, in *Scientific Basis for Nuclear Waste Management*, edited by J. G. Moore, E. A. Bryant, and A. L. Harkey, pp. 465–472, Plenum, New York.

Bauer, S., C. Fulda, and W. Schäfer (2001), A multi-tracer study in a shallow aquifer using age dating tracers 3H, ^{85}Kr, CFC-113 and SF_6—Indication for retarded transport of CFC-113, *J. Hydrol.*, *248*(1–4), 14–34.

Bethke, C. M., and T. M. Johnson (2008), Groundwater age and groundwater age dating, *Annu. Rev. Earth Planet. Sci.*, *36*, 121–152.

Bohlke, J. K., and J. M. Denver (1995), Combined use of groundwater dating, chemical, and isotopic analyses to resolve the history and fate of nitrate contamination in two agricultural watersheds, Atlantic coastal plain, Maryland, *Water Resour. Res.*, *31*(9), 2319–2339.

Bruce, D. L., Y. Yechieli, M. Zilberbrand, A. Kaufman, and G. M. Friedman (2007), Delineation of the coastal aquifer of Israel based on repetitive analysis of C-14 and tritium, *J. Hydrol.*, *343*, 56–70.

Busenberg, E., and L. N. Plummer (1992), Use of chlorofluorocarbons (CCl_3F and CCl_2F_2) as hydrologic tracers and age-dating tools: The alluvium and terrace system of central Oklahoma, *Water Resour. Res.*, *28*(9), 2257–2283.

Carlin, B. P., and S. Chib (1995), Bayesian model choice via Markov chain Monte Carlo methods, *J. R. Stat. Soc., B*, *57*(3), 473–484.

Castro, M. C., A. Jambon, G. de Marsily, and P. Schlosser (1998), Noble gases as natural tracers of water circulation in the Paris Basin: 1. Measurements and discussion of their origin and mechanisms of vertical transport in the basin, *Water Resour. Res.*, *34*(10), 2443–2466.

Chhikara, R. S., and L. Folks (1989), *The Inverse Gaussian Distribution: Theory, Methodology, and Applications*, Marcel Dekker, New York.

Choung, S., and R. M. Allen-King (2010), Can chlorofluorocarbon sorption to black carbon (char) affect groundwater age determinations?, *Environ. Sci. Technol.*, *44*(12), 4459–4464.

Cook, P. G., and J. K. Bohlke (2000), Determining timescales for groundwater flow and solute transport, in *Environmental Tracers in Subsurface Hydrology*, edited by P. G. Cook and A. L. Herczeg, pp. 1–30, Kluwer Acad., Boston.

Cook, P. G., and D. K. Solomon (1995), Transport of atmospheric trace gases to the water-table: Implications for groundwater dating with chlorofluorocarbons and krypton-85, *Water Resour. Res.*, *31*(2), 263–270.

Cook, P. G., and D. K. Solomon (1997), Recent advances in dating young groundwater: Chlorofluorocarbons, H-3/He-3 and Kr-85, *J. Hydrol.*, *191*(1–4), 245–265.

Cook, P. G., D. K. Solomon, L. N. Plummer, E. Busenberg, and S. L. Schiff (1995), Chlorofluorocarbons as tracers of groundwater transport processes in a shallow, silty sand aquifer, *Water Resour. Res.*, *31*(3), 425–434.

Deines, P. (1980), The isotopic composition of reduced organic carbon, in *Handbook of Environmental Isotope Geochemistry: The Terrestrial Environment*, edited by P. Fritz and J. C. Fontes, pp. 329–406, Elsevier, Amsterdam.

El-Shaarawi, A. H. (1989), Inferences about the mean from censored water-quality data, *Water Resour. Res.*, *25*(4), 685–690.

El-Shaarawi, A. H., and J. Lin (2007), Interval estimation for log-normal mean with applications to water quality, *Environmetrics*, *18*(1), 1–10.

Gamerman, D., and F. L. Hedibert (2006), *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, 2nd ed., CRC Press, Boca Raton, Fla.

Garnier, J.-M. (1985), Retardation of dissolved radiocarbon through a carbonated matrix, *Geochim. Cosmochim. Acta*, *49*(3), 683–693.

Genereux, D. P., M. Webb, and D. Solomon (2009), Chemical and isotopic signature of old groundwater and magmatic solutes in a Costa Rican rain forest: Evidence from carbon, helium, and chlorine, *Water Resour. Res.*, *45*, W08413, doi:10.1029/2008WR007630.

Geyh, M. A. (1999), Dating of old groundwater— History, potential, limits, and future, in *Isotopes in the Water Cycle: Past, Present and Future of a Developing Science*, edited by P. K. Aggarwal, J. R. Gat, and K. F. O. Froehlich, pp. 221–241, Springer, Dordrecht, Netherlands.

Glynn, P. D., and L. N. Plummer (2005), Geochemistry and the understanding of ground-water systems, *Hydrogeol. J.*, *13*(1), 263–287.

Goode, D. J. (1996), Direct simulation of groundwater age, *Water Resour. Res.*, *32*(2), 289–296.

Helsel, D. R., and T. A. Cohn (1988), Estimation of descriptive statistics for multiply censored water-quality data, *Water Resour. Res.*, *24*(12), 1997–2004.

Hinsby, K., W. Edmunds, H. Loosli, M. Manzano, M. Condesso De Melo, and F. Barbecot (2001), The modern water interface: Recognition, protection and development—Advance of modern waters in European aquifer systems, *Geol. Soc. London Spec. Publ.*, *189*(1), 271–288.

Hinsby, K., A. L. Hojberg, P. Engesgaard, K. H. Jensen, F. Larsen, L. N. Plummer, and E. Busenberg (2007), Transport and degradation of chlorofluorocarbons (CFCs) in the pyritic Rabis Creek aquifer, Denmark, *Water Resour. Res.*, *43*, W10423, doi:10.1029/2006WR005854.

Horneman, A., M. Stute, P. Schlosser, W. Smethie, N. Santella, D. T. Ho, B. Mailloux, E. Gorman, Y. Zheng, and A. van Geen (2008), Degradation rates of CFC-11, CFC-12 and CFC-113 in anoxic shallow aquifers of Araihazar, Bangladesh, *J. Contam. Hydrol.*, *97*(1–2), 27–41.

International Atomic Energy Agency (IAEA) (2011), Global Network of Isotopes in Precipitation (GNIP), http://www-naweb.iaea.org/napc/ih/IHS_resources_gnip.html, Vienna.

Jeffreys, H. (1935), Some tests of significance, treated by the theory of probability, *Proc. Cambridge Philos. Soc.*, *31*, 203–222.

Kaipio, J., and E. Somersale (2004), *Statistical and Computational Inverse Problems*, *Appl. Math. Sci.*, vol. 160, Springer, Berlin.

Kass, R. E., and A. E. Raftery (1995), Bayes factors, *J. Am. Stat. Assoc.*, *90*(430), 773–795.

Keats, A., M. T. Cheng, E. Yee, and F. S. Lien (2009), Bayesian treatment of a chemical mass balance receptor model with multiplicative error structure, *Atmos. Environ.*, *43*(3), 510–519.

Lehmann, B. E., et al. (2003), A comparison of groundwater dating with ^{81}Kr, ^{36}Cl and ^{4}He in four wells of the Great Artesian Basin, Australia, *Earth Planet. Sci. Lett.*, *211*(3–4), 237–250.

Limpert, E., W. A. Stahel, and M. Abbt (2001), Log-normal distributions across the sciences: Keys and clues, *BioScience*, *51*(5), 341–352.

Lovley, D. R., and J. C. Woodward (1992), Consumption of freons CFC-11 and CFC-12 by anaerobic sediments and soils, *Environ. Sci. Technol.*, *26*(5), 925–929.

Maloszewski, P., and A. Zuber (1982), Determining the turnover time of groundwater systems with the aid of environmental tracers: Part 1. Models and their applicability, *J. Hydrol.*, *57*(3–4), 207–231.

Maloszewski, P., and A. Zuber (1993), Principles and practice of calibration and validation of mathematical models for the interpretation of environmental tracer data in aquifers, *Adv. Water Resour.*, *16*(3), 173–190.

Massoudieh, A., and T. R. Ginn (2011), The theoretical relation between unstable solutes and groundwater age, *Water Resour. Res.*, *47*, W10523, doi:10.1029/2010WR010039.

Mazor, E. (2003), *Chemical and Isotopic Groundwater Hydrology*, CRC Press, Boca Raton, Fla.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953), Equations of state calculations by fast computing machines, *J. Chem. Phys.*, *21*(6), 1087–1092.

Mukhopadhyay, N. (2000), *Probability and Statistical Inference*, Marcel Dekker, New York.

Oster, H., C. Sonntag, and K. O. Münnich (1996), Groundwater age dating with chlorofluorocarbons, *Water Resour. Res.*, *32*(10), 2989–3001.

Plummer, L. N. (1999), Dating of young groundwater, in *Isotopes in the Water Cycle: Past, Present and Future of a Developing Science*, edited by P. K. Aggarwal, J. R. Gat, and K. F. O. Froehlich, pp. 193–218, Springer, Dordrecht, Netherlands.

Plummer, L. N., and E. Busenberg (2006), Chlorofluorocarbons in the atmosphere, in *Use of Chlorofluorocarbons in Hydrology: A Guidebook*, pp. 9–14, Int. At. Energy Agency, Vienna.

Plummer, N., and C. Sprinkle (2001), Radiocarbon dating of dissolved inorganic carbon in groundwater from confined parts of the Upper Floridan aquifer, Florida, USA, *Hydrogeol. J.*, *9*(2), 127–150.

Plummer, M. A., L. C. Hull, and D. T. Fox (2004), Transport of carbon-14 in a large unsaturated soil column, *Vadose Zone J.*, *3*(1), 109–121.

Portniaguine, O., and D. K. Solomon (1998), Parameter estimation using groundwater age and head data, Cape Cod, Massachusetts, *Water Resour. Res.*, *34*(4), 637–645.

Reilly, T. E., L. N. Plummer, P. J. Phillips, and E. Busenberg (1994), The use of simulation and multiple environmental tracers to quantify groundwater-flow in a shallow aquifer, *Water Resour. Res.*, *30*(2), 421–433.

Scheutz, C., and P. Kjeldsen (2003), Capacity for biodegradation of CFCs and HCFCs in a methane oxidative counter-gradient laboratory system simulating landfill soil covers, *Environ. Sci. Technol.*, *37*(22), 5143–5149.

Sebol, L. A., W. D. Robertson, E. Busenberg, L. N. Plummer, M. C. Ryan, and S. L. Schiff (2007), Evidence of CFC degradation in groundwater under pyrite-oxidizing conditions, *J. Hydrol.*, *347*(1–2), 1–12.

Shumway, R. H., R. S. Azari, and M. Kayhanian (2002), Statistical approaches to estimating mean water quality concentrations with detection limits, *Environ. Sci. Technol.*, *36*(15), 3345–3353.

Solomon, D. K., D. P. Genereux, L. N. Plummer, and E. Busenberg (2010), Testing mixing models of old and young groundwater in a tropical lowland rain forest with environmental tracers, *Water Resour. Res.*, *46*, W04518, doi:10.1029/2009WR008341.

Sonier, D. N., N. L. Duran, and G. B. Smith (1994), Dechlorination of trichlorofluoromethane (CFC-11) by sulfate-reducing bacteria from an aquifer contaminated with halogenated aliphatic-compounds, *Appl. Environ. Microbiol.*, *60*(12), 4567–4572.

Szabo, Z., D. E. Rice, L. N. Plummer, E. Busenberg, and S. Drenkard (1996), Age dating of shallow groundwater with chlorofluorocarbons, tritium helium 3, and flow path analysis, southern New Jersey coastal plain, *Water Resour. Res.*, *32*(4), 1023–1038.

Troldborg, L., K. H. Jensen, P. Engesgaard, J. C. Refsgaard, and K. Hinsby (2008), Using environmental tracers in modeling flow in a complex shallow aquifer system, *J. Hydrol. Eng.*, *13*(11), 1037–1048.

Varni, M., and J. Carrera (1998), Simulation of groundwater age distributions, *Water Resour. Res.*, *34*(12), 3271–3281.

Walker, G. R., and P. G. Cook (1991), The importance of considering diffusion when using C-14 to estimate groundwater recharge to an unconfined aquifer, *J. Hydrol.*, *128*(1–4), 41–48.

Warner, M. J., and R. F. Weiss (1985), Solubilities of chlorofluorocarbons 11 and 12 in water and seawater, *Deep Sea Res., Part A*, *32*(12), 1485–1497.

Webb, M. D. (2007), Carbon, chlorine, and oxygen isotopes as tracers of interbasin groundwater flow at La Selva Biological Station, Costa Rica, MS thesis, N. C. State Univ., Raleigh.

Weissmann, G. S., Y. Zhang, E. M. LaBolle, and G. E. Fogg (2002), Dispersion of groundwater age in an alluvial aquifer system, *Water Resour. Res.*, *38*(10), 1198, doi:10.1029/2001WR000907.

Zoellmann, K., W. Kinzelbach, and C. Fulda (2001), Environmental tracer transport (H-3 and SF6) in the saturated and unsaturated zones and its use in nitrate pollution management, *J. Hydrol.*, *240*(3–4), 187–205.