



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

A Bayesian Modeling Approach for Estimation of a Shape-Free Groundwater Age Distribution using Multiple Tracers

A. Massoudieh, A. Visser, S. Sharifi, H. Broers

May 28, 2013

Applied Geochemistry

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

1 A Bayesian modeling approach for 2 estimation of a shape-free groundwater 3 age distribution using multiple tracers

4
5 Authors: Arash Massoudieh^{1*}, Ate Visser², Soroosh Sharifi¹, Hans Peter Broers^{3,4,5}.

6 1- Civil Engineering, The Catholic University of America, Washington, DC, USA 20064.

7 2- Chemical Sciences Division, Lawrence Livermore National Laboratory, 7000 East
8 Avenue Livermore, CA, USA 94550

9 3- Deltares, Unit Soil and Groundwater Systems, Utrecht, The Netherlands

10 4- TNO, Geological Survey of the Netherlands, Utrecht, The Netherlands

11 5- VU University, Critical Zone Hydrology Group, Amsterdam, The Netherlands

12 **Abstract**

13 Due to the mixing of groundwaters with different ages in aquifers, groundwater age is more
14 appropriately represented by a distribution rather than a scalar number. To infer a groundwater
15 age distribution from environmental tracers, a mathematical form is often assumed for the
16 shape of the distribution and the parameters of the mathematical distribution are estimated
17 using deterministic or stochastic inverse methods. The prescription of the mathematical form
18 limits the exploration of the age distribution to the shapes that can be described by the selected
19 distribution. In this paper, the use of freeform histograms as groundwater age distributions is
20 evaluated. A Bayesian Markov Chain Monte Carlo approach is used to estimate the fraction of
21 groundwater in each histogram bin. The method was able to capture the shape of a hypothetical
22 gamma distribution from the concentrations of four age tracers. The number of bins that can be
23 considered in this approach is limited based on the number of tracers available. The histogram

24 method was also tested on groundwater age data sets from Holten (The Netherlands) and the
25 La Selva Biological Station in Costa-Rica, and compared to a number of mathematical forms.
26 According to standard Bayesian measures of model goodness, the best mathematical
27 distribution performs better than the histogram distributions in terms of the ability to capture
28 the observed tracer data relative to their complexity. Among the histogram distributions, the
29 four bin histogram performs better in most of the cases. The Monte Carlo simulations showed
30 strong correlations in the posterior estimates of bin contributions, indicating that these bins
31 cannot be well constrained using the available age tracers. The fact that mathematical forms
32 overall perform better than the freeform histogram does not undermine the benefit of the
33 freeform approach, especially for the cases where a larger amount of observed data is available
34 and when the real groundwater distribution is more complex than can be represented by simple
35 mathematical forms.

36

37 Keywords: Groundwater dating, Bayesian Inference, Shape Free Histogram, Environmental
38 Tracers.

39

40 **1. Introduction**

41 The subsurface travel time of groundwater – here also referred to as groundwater age – is
42 important information for assessing the vulnerability of wells to contamination (Bethke and
43 Johnson, 2008; Broers and G., 2005; Glynn and Plummer, 2005; Kralik and Keimel, 2003;
44 Manning et al., 2005), evaluating the history and fate of contaminants (Bohlke and Denver, 1995;
45 Hinsby et al., 2001) and demonstrating the effectiveness and timescales of groundwater quality
46 management strategies (Hansen et al., 2010; Laier, 2004; Visser et al., 2007; Wassenaar et al.,
47 2006; Zoellmann et al., 2001). The age of pumped groundwater also reflects the sustainability of
48 groundwater resources under climate change (Singleton and Moran, 2010).

49 Initial conceptual models for "groundwater age" assumed a uniform age of groundwater at a
50 specific location in the aquifer (i.e. piston flow model). Recent work demonstrated that even at
51 the smallest scale (e.g. ~200m), due to the diffusive and dispersive mixing of ages in the aquifer
52 (Engesgaard et al., 1996; Gelhar et al., 1992; Weissmann et al., 2002), the age of a groundwater
53 parcel is a distribution (Bethke and Johnson, 2008; Massoudieh et al., 2012). The degree of
54 mixing is still heavily debated (Castro et al., 1998; Cirpka and Attinger, 2003; Solomon et al.,
55 1992; Weissmann et al., 2002), as are the underlying processes and consequent shapes of the
56 probability density function (Engdahl et al., 2012; Engdahl et al., 2013; Weissmann et al., 2002).
57 Production wells and springs undoubtedly collect groundwater with a wide age distribution
58 (Manning et al., 2005).

59 Unfortunately, age is not a quantity that can be measured directly (Massoudieh and Ginn, 2011).
60 Therefore, it needs to be derived from analytical or numerical modeling of groundwater fluxes
61 (Goode, 1996; Troldborg et al., 2008; Visser et al., 2009; Woolfenden and Ginn, 2009) or by
62 converting measured concentrations of a number of tracers to a groundwater age distribution
63 (Corcho Alvarado et al., 2007; Lehmann et al., 2003; Plummer et al., 2001; Solomon et al., 2010;
64 Sültenfuß et al., 2011; Visser et al., 2013). At the time-scale relevant for well vulnerability and
65 response (years to decades), a number of age tracers (e.g. ⁸⁵Krypton (Smethie et al., 1992),
66 tritium-helium (Poreda et al., 1988; Schlosser et al., 1988), chlorofluorocarbons (Busenberg and
67 Plummer, 1992), sulfur-hexafluoride (Busenberg and Plummer, 2000), and ³⁹Argon (Loosli, 1983;
68 Loosli et al., 1989; Oeschger et al., 1974)) are suitable.

69 Since a single sample of a single age tracer is incapable of identifying the entire age distribution,
70 it needs to be inferred from a combination of age tracers or a time series of age tracers
71 (Maloszewski and Zuber, 1982; Morgenstern et al., 2010) or other parameters such as electrical
72 conductivity (Cirpka and Attinger, 2003; Molina-Giraldo et al., 2010). The age distribution is
73 often deconvoluted using mathematical models prescribing its shape (Maloszewski and Zuber,
74 1993; Maloszewski and Zuber, 1998). For instances where such models are too restrictive due to
75 the complexity of true age distribution or the heterogeneity of the aquifer, a shape-free age

76 histogram approach was developed (Liao and Cirpka, 2011; Massoudieh and Ginn, 2011; Visser
77 et al., 2013).

78 Deconvolution is often deterministically performed, using oversimplifying assumptions
79 (Rinaldo et al., 2011) providing a single age distribution. Several studies have included
80 parameter uncertainty (Cirpka et al., 2007; Corcho Alvarado et al., 2007; Sültenfuß et al., 2011)
81 or applied a multitude of age models to estimate the uncertainty of the age distribution
82 (Solomon et al., 2010; Visser et al., 2013). The starting point for improving theories and datasets
83 is a formal framework of hypothesis testing (Beven, 2010). Including an explicit recognition of
84 input uncertainty yields probability density functions of the age model parameters and
85 consequently age distributions. Larocque et al., (2009) presented the first systematic and
86 integrated assessment of bias and uncertainty associated with the estimation of groundwater
87 flow rates using tracers and pumping tests in heterogeneous aquifer systems.

88 Bayesian inference techniques are more commonly applied in catchment hydrology (Vrugt et
89 al., 2008) and unsaturated zone hydrology (Vrugt et al., 2001). Recent studies investigated the
90 propagated uncertainties in excess air models (Jung et al., 2012; Sun et al., 2010) and used
91 Bayesian methods for contaminant source identification (Fox and Papanicolaou, 2008;
92 Massoudieh and Kayhanian, 2013; Zeng et al., 2012). So far, only a single study has applied
93 Bayesian methods to derive age distribution from groundwater age tracers (Massoudieh et al.,
94 2012). This work will extend that study by applying Bayesian inference to a shape free age
95 histogram model. The goal of this study is to evaluate the possibility of using measured
96 concentrations of multiple environmental tracers to infer the groundwater age distribution
97 when the form of the groundwater distribution is not assumed *a priori* but rather is considered
98 to be a histogram with a given bin size but unknown values in each bin. The approach is tested
99 on three increasingly complex cases: a synthetic example, a multi age tracer data set from a
100 production well field in a simple hydrological setting (Visser et al., 2013) and a dataset
101 involving degrading tracers in a complex hydrogeologic system (Solomon et al., 2010). In this
102 study, we investigated the uncertainty of a histogram age distribution estimated from multiple
103 age tracers, at different levels of complexity. The histogram age distribution provides a more

104 free-form presumed age distribution compared to the traditionally used lumped parameter
105 models where the presumed form of the distribution is highly restricted to a prescribed
106 mathematical form.

107 **2. Methods**

108 **2.1 Tracers**

109 Figure 1 shows the expected concentrations (decay-corrected to 2010 if necessary) of common
110 groundwater age tracers in groundwater under piston-flow conditions, on a logarithmic time-
111 scale. Tracer concentrations are averaged over exponentially increasing age bins, and then
112 scaled to the maximum concentration. The logarithmic time scale reflects the exponentially
113 increasing desire for higher resolution when dating younger groundwater. Three patterns
114 emerge from this figure: Type 1: gradually decreasing concentrations with age, resulting from
115 radioactive decay (^{39}Ar , ^{14}C) or increasing anthropogenic releases into the modern atmosphere
116 (SF_6 , CFCs) or both (^{85}Kr). Type 2: increasing concentrations with age (radiogenic ^4He) and Type
117 3: a pulse due to nuclear testing and subsequent decreasing concentrations in younger
118 groundwater (^3H , tritiogenic ^3He , ^{14}C).

119 A combination of Type 1 tracers is ideal for deconvoluting the age distribution using
120 mathematical models, granted that the different tracers exhibit distinctly different histories
121 (Kass and Raftery, 1995). Four distinct Type 1 curves are visible: (1) ^{85}Kr and SF_6 , (2) CFCs, (3)
122 ^{39}Ar and (4) ^{14}C , with “half-lives” of 10.7 years for ^{85}Kr and SF_6 , 25-34 years for CFCs, 269 years
123 for ^{39}Ar and 5730 years for ^{14}C . (“Half-life” referring to the time that half of the maximum
124 concentration is observed.)

125 The linear increasing concentrations of radiogenic helium (Type 2) observed in ideal cases often
126 require calibration against a different age tracer (e.g. Plummer et al., 2012) and provide a
127 distinctly different perspective. The linear ingrowth of radiogenic ^4He results in an increasing
128 sensitivity to the oldest groundwater component in the mixture.

129 The pulsed input caused by nuclear testing (Type 3) is most pronounced in ^3H and ^3He . These
 130 tracers are specifically sensitive to that particular period. The combination of ^3H and ^3He
 131 benefits deconvolution if only a small portion of groundwater dates back to the nuclear test
 132 period because of the overwhelming initial tritium concentrations. The impact of nuclear testing
 133 on ^{14}C is more prolonged due to its longer half-life and the longer lifetime of CO_2 in the
 134 atmosphere, yet less important because of the distance from the age window that is typically
 135 dated with ^{14}C .

136 The example datasets we analyze contain two combinations of tracers: ^{85}Kr , ^3H , ^3He and ^{39}Ar in
 137 Holten, and SF_6 , CFCs, ^3H , ^4He and ^{14}C in La Selva.

138 **2.2 The relationship between tracer concentration and groundwater age** 139 **distribution**

140 Similar to the work by Massoudieh et al. (Massoudieh et al., 2012) assuming that the
 141 contribution from mineral dissolution is devoid of tracers and that the adsorbed tracers
 142 undergo the same rate of decay as the mobile tracers, the general form of the equation for tracer
 143 concentration of decaying and linearly accumulating tracers respectively can be written based
 144 on Maloszewski and Zuber (1982) as:

$$145 \quad c_i(\mathbf{x}, \lambda_i, t) = f \cdot f_m \int_0^\infty c_{i,o}(t - R_i \tau') e^{-\lambda_i \tau} \rho(\mathbf{x}, t, \tau) d\tau + (1 - f) c_{old,i} \quad (1a)$$

$$146 \quad c_i(\mathbf{x}, \lambda_i, t) = f \int_0^\infty \lambda_i \tau \rho(\mathbf{x}, t, \tau) d\tau + c_{i,o}(t - \tau) + (1 - f) c_{old,i} \quad (1b)$$

147 where $c_i(\mathbf{x}, \lambda_i, t)$ is the measured tracer concentration (or isotope's concentration normalized by
 148 the concentration of the stable isotope) of tracer i at location \mathbf{x} and time t , $\rho(\mathbf{x}, t, \tau)$ is the
 149 groundwater age distribution, $c_{i,o}(t - \tau)$ is the concentration (or the isotope ratio) in the
 150 recharge water at time $t - \tau$. In Eq. (1a), λ_i is the decay rate of tracer i for decaying tracers and
 151 in Eq. (1b), it is the accumulation rate of isotope i in case of linearly accumulating isotope

152 tracers. R_i is the retardation factor for tracer i , f_m is the fraction of carbon with atmospheric
 153 source as opposed to mineral dissolution source (only applicable to ^{14}C), f is the fraction of
 154 young groundwater and $c_{old,i}$ is the tracer concentration in the old fraction of groundwater
 155 which is assumed to be fixed.

156 2.3 Bayesian Inference

157 The purpose of the Bayesian inference is to obtain the possibility space of freeform age
 158 distribution given the tracer data. This is done by generating the joint probability distribution of
 159 the model parameters including the parameters in Eq. (1a and b) and the fraction of
 160 groundwater in each bin of the histogram representing the age distribution. The inferred
 161 probability distribution of model parameters and values in each bin based on the observed
 162 tracer concentrations are referred to as the posterior distribution of the model parameter. To
 163 quantify a groundwater age distribution from available age tracer data, a shape-free discrete
 164 groundwater age distribution model was applied. The model describes the groundwater age
 165 distribution by a number of age bins with a uniform age distribution within each bin. Because
 166 most groundwater age tracers (^3H , tritiogenic ^3He , ^{85}Kr , SF_6 , CFCs) relate to the period since
 167 1950, one bin represents groundwater that recharged prior to 1950. The remaining bins cover
 168 the period from present to 1950. For five bins, the first four bins each span 15 years with breaks
 169 at 2010, 1995, 1980, 1965 and 1950. The groundwater age distribution of the old groundwater bin
 170 cannot be further refined, unless multiple old groundwater tracers (^{39}Ar , ^{14}C , radiogenic ^4He) are
 171 available. Ideally, the age distribution of the old groundwater fraction is homogeneous within
 172 the aquifer and has a distinct tracer signature. A discrete groundwater age distribution model
 173 with n bins is defined by $n-1$ parameters. The number of bins is ideally smaller or equal to the
 174 number of tracer measurements, to ensure that the problem is not under defined. The goal here
 175 is to estimate the contribution of groundwater with age a falling within bin i of the histogram:

$$176 \rho_m(X, t, a) = \begin{cases} \phi_i & \text{for } (i-1)\Delta h < a < i\Delta h, \quad i = 1 \cdots n-1 \\ 1 - \Delta h \sum_{i=1}^{n-1} \phi_i & \text{for } a > (n-1)\Delta h \end{cases} \quad (2)$$

177 where $\phi_1, \phi_2, \dots, \phi_n$ are the fraction of groundwater in each bin normalized by size of the bins. .
 178 The goal of Bayesian Inference is to infer the probability density functions of the parameters (i.e.
 179 $\Phi = [\phi_1, \phi_2, \dots, \phi_n]$) and also the fraction of tracers with atmospheric sources, f_m , and the fraction
 180 of young groundwater f while considering the uncertainties in decay rates, λ_i and also the
 181 observed concentrations. If we consider vector $\mathbf{C} = [c_1, c_2, \dots, c_m]$ to represent the random vector
 182 comprising the *true* concentrations (or isotope ratios) of all chemicals, and vector
 183 $\hat{\mathbf{C}} = [\hat{c}_1, \hat{c}_2, \dots, \hat{c}_m]$ to be the observed concentrations, and assuming that the observation errors for
 184 tracer concentrations are log-normally distributed and multiplicative with an equal standard
 185 deviation σ for all the tracers, the Bayes theorem can be written as (Massoudieh et al., 2012):

$$p(\Phi, f, f_m, \lambda_1, \lambda_2, \dots, c_m, c_b, \sigma | \hat{\mathbf{C}}) \propto$$

$$e^{-\sum_{i=1}^m \frac{[\ln(\hat{c}_i) - \ln(c_i)]^2}{2\sigma^2}} \cdot e^{-\frac{(\hat{c}_c - c_c)^2}{2\sigma_c^2}} \cdot U(f_m, 0, 1) \cdot U(f, 0, 1) \cdot \left\{ \frac{1}{\prod \lambda_i} e^{-\sum_{i=1}^m \frac{[\ln(\mu_{\lambda_i}) - \ln(\lambda_i)]^2}{2\sigma_{\lambda_i}^2}} \right\} \cdot e^{-\left\{ \frac{[\mu_{cb} - c_b]^2}{\sigma_{cb}^2} + \frac{[\mu_{cm} - c_m]^2}{\sigma_{cm}^2} \right\}}$$

186 (3)

187 where m is the number of tracers, c_c and \hat{c}_c are the true and observed concentrations of any non-
 188 decaying signature tracer that is used to determine the source of dominant isotopes of the
 189 tracers (e.g. ^{13}C) and is considered to have a Gaussian and additive error structure with a fixed
 190 standard deviation of σ_c . λ_i is the decay rate of tracer i , μ_{λ_i} and σ_{λ_i} are, respectively, the mean
 191 and standard deviation of prior distributions of decay rates for tracer i . μ_{cb} and μ_{cm} are the
 192 means of the prior normal distribution of non-decaying tracer concentrations with biogenic and
 193 mineral sources, respectively. $U(x, 0, 1)$ represent uniform distribution between 0 and 1. σ_{cb} and
 194 σ_{cm} are the standard deviations of the prior normal distribution of non-decaying tracer
 195 concentrations with biogenic and mineral sources, respectively, c_b and c_m are the concentrations
 196 of signature tracers with biogenic and mineral sources, respectively. For example when ^{14}C is
 197 used for dating groundwater it is important to know what fraction of total inorganic carbon in
 198 the aqueous phase is contributed through mineral dissolution and ^{13}C isotope ratio contain
 199 information that can be used for this purpose. In a case where the tracer observed error

200 structure is normally distributed and the observed standard deviation is not the same for all the
201 tracers, the likelihood function (the first term on the right hand-side of Eq. (3)) should be
202 replaced by a joint normal distribution using a variance-covariance matrix instead of a single
203 variance. The assumption about the error structure used in the likelihood function can
204 potentially have a significant effect on the inferred age distribution due to the fact that it
205 determines how much weight to be given to each of the tracers.

206 *Markov-Chain Monte Carlo*

207 In order to find the expected values, i.e. confidence intervals of the posterior joint probability
208 density function (PDF), and to evaluate the correlations between the inferred parameters, Eq. 3
209 has to be integrated over the parameter space. It is clear that since the number of dimensions of
210 the parameter space is large, integration of Eq. 3 is prohibitive. Markov-Chain Monte Carlo
211 (MCMC) methods are relatively simple methods to generate a large number of samples of
212 parameters, based on the posterior distribution (Gamerman and Hedibert, 2006; Kaipio and
213 Somersale, 2004). Algorithms such as Metropolis-Hasting or Gibbs Sampling provide a way to
214 generate samples according to a large-dimensional posterior joint probability density function
215 (JPDF). Here we use the Metropolis-Hasting Algorithm (Metropolis et al., 1953). A C++ code is
216 written to perform the MCMC simulation. The number of Markov Chains can be determined by
217 the user. In the example application presented in the next section, 500,000 samples resulted in
218 convergence of the MCMC method. The first 100,000 samples were left out as “burn-in” period.
219 In order to reduce the burn-in period, a hybrid genetic algorithm (Massoudieh et al., 2008) was
220 used prior to the MCMC sample generation to find the neighborhood of the optimal parameters
221 and the MCMC is initialized from the optimal parameter set.

222 The effectiveness of the free form age distribution with varying number of bins as well as
223 mathematical forms was assessed by the Deviance Information Criteria (DIC) and Bayes Factors
224 (BF).

225 *Deviance Information Criteria (DIC)*

226 DIC (Spiegelhalter et al., 2002) is a measure that is used to compare the goodness of different
227 model structures applied to the same data. It takes into account how well a model structure can
228 reproduce the observed data while explicitly considering the complexity level of the model:

229
$$DIC = 2\overline{D(\Theta)} - D(\bar{\Theta})$$
 (5)

230 where $\overline{D(\Theta)}$ is the mean deviance and $D(\bar{\Theta})$ is the deviance of the mean, respectively defined
231 as:

232
$$\overline{D(\Theta)} = -2E[\ln p(\hat{C} | \Theta)] - 2\ln f(\hat{C})$$
 (6)

233
$$D(\bar{\Theta}) = -\ln p(\hat{C} | \bar{\Theta}) - 2\ln f(\hat{C})$$
 (7)

234 where f is a standardizing term that is only a function of the observed data, Θ represents the
235 parameters that are random-vector distributed based on the posterior distribution, and $\bar{\Theta}$ is the
236 expected values of the posterior distribution of parameters. Because $2\ln f(\hat{C})$ is only a function
237 of the observed data, it is not affected by the model structure, and therefore, becomes irrelevant
238 when comparing two models. The DIC can be easily calculated using the results of the MCMC
239 calculations. A lower DIC value indicates that the modeled tracers deviate less from the
240 measured tracers given the model complexity and that the model performs better.

241 *Bayes Factor (BF)*

242 To compare different age distribution forms in terms of their ability to reproduce the measured
243 data, the Bayes factor method (Jeffreys, 1935; Kass and Raftery, 1995) was used. The Bayes
244 factor for comparing two models M_1 and M_2 assuming an equal prior probability for the models
245 is defined as:

$$B_{12} = \frac{\int_{\Theta_1} p(\hat{C} | \Theta_1, M_1) p(\Theta_1)}{\int_{\Theta_2} p(\hat{C} | \Theta_2, M_2) p(\Theta_2)} \quad (8)$$

247 B_{12} represents the ratio of the chance that model 1 is the true model to the chance that model 2
 248 is the true model, $p(\hat{C} | \Theta, M)$ is the likelihood of observing concentrations \hat{C} given model M ,
 249 and $p(\Theta)$ is the prior density of the parameters. Since we are comparing multiple models here,
 250 only the numerator of Eq. (8) is calculated for each model as:

$$I_k = \int_{\mathbf{x}_s} p(\hat{C} | \Theta_s, M_k) p(\Theta_s) \quad (9)$$

252 The integrals of posterior distribution are estimated using the Monte Carlo method (Carlin and
 253 Chib, 1995):

$$\int_{\Theta_s} p(\hat{C} | \mathbf{X}_s, M_s) p(\Theta_s) \approx \frac{1}{n} \sum p(\hat{C} | \Theta_s^{(k)}, M_s) p(\Theta_s^{(k)}) \quad (10)$$

255
 256 In this paper, the Bayes Factor is expressed as the logarithm of I , whereby a high $\text{Log}(I)$ indicates
 257 that the model performs better and therefore represents a high chance/probability that the
 258 model is correct. Bayes factor can be viewed as measuring the relative success of a model at
 259 predicting the data while implicitly considering the model complexity in evaluating the
 260 evidence in support of a model (Horneman et al., 2008).

261 **2.4 Experiments and study areas**

262 *2.4.1 Synthetic case*

263 To first evaluate the ability of the method to infer histograms that are close to the true age
 264 distribution from groundwater age tracer data, a hypothetical case of a known age distribution
 265 is used. The hypothetical age distribution is a gamma distribution,

266 $gamma(x; k, \theta) = \Gamma(k)^{-1} \theta^{-k} x^{k-1} e^{-x/\theta}$, with a shape parameter α of 4 and a scale parameter θ of
 267 5. This gamma age distribution is used to calculate the average concentration of four tracers
 268 (^3H , ^3He , ^{85}Kr and ^{39}Ar) for the mixture, convoluting the tracer inputs of Figure 1 with the
 269 gamma age distribution. Next, the Bayesian inference scheme is applied to reconstruct the age
 270 distribution using 3, 4 and 10 bin histograms based on the calculated tracer concentrations. The
 271 ability of the Bayesian inference scheme to estimate the histogram contributions is evaluated by
 272 measures of spread and bias of the estimated age distribution with respect to the true values:

$$273 \quad \xi = \int_0^{\infty} E \left\{ [\tilde{\rho}(a) - \rho(a)]^2 \right\} da \quad (12)$$

$$274 \quad b = \int_0^{\infty} \left\{ E[\tilde{\rho}(a)] - \rho(a) \right\}^2 da \quad (13)$$

275 where ξ is a measure of spread of the estimated age distributions $\tilde{\rho}(a)$ around the true age
 276 distribution $\rho(a)$ which can be interpreted as a measure of uncertainty, and b is a measure of
 277 bias representing how much the expected value of age distribution deviates from the true age
 278 distribution.

279 2.4.2 Holten

280 The production well field is located near Holten, in the eastern part of the Netherlands. A
 281 phreatic groundwater system is present in the (partly ice-pushed) fluvial and periglacial
 282 deposits of 90 to 120 m thickness. The Holten well field produces 2.5 million cubic meters
 283 groundwater using 19 wells within 1 km² from a depth of 10 to 70 meters below the surface. The
 284 well field is vulnerable to contamination because of the agricultural land use and residential
 285 areas in the capture area, the sandy aquifer, expected short groundwater residence times and
 286 relatively shallow groundwater levels (4-8 m below surface). Most wells were constructed
 287 between 1960 and 1973 and are screened fairly shallow, between 15 and 30 m below the surface.
 288 In 1985, the three deep wells were drilled and screened from 45 to 70 m below the surface. The
 289 differences in screen depths were expected to result in distinctly different groundwater age

290 distributions (Visser et al. 2013). Samples for ^{85}Kr , $^3\text{H}/^3\text{He}$, and ^{39}Ar analysis were collected on 20
291 and 21 April 2010 from seven wells (four shallow, three deep), capturing 69% of the total
292 drinking water production.

293 For the Holten study area, the concentrations of ^3H in historical precipitation were
294 reconstructed based on five nearby stations and the Vienna station of the Global Network of
295 Isotopes in Precipitation (GNIP, IAEA/WMO, 2010). The time series of the ^{85}Kr activity in the
296 atmosphere collected and measured at Freiburg im Breisgau (Institute of Atmospheric Research
297 (IAR), Freiburg, Germany) was used for Holten without correction, which is likely a lower
298 estimate because of closer the proximity to the reprocessing facilities Sellafield and La Hague
299 compared to Freiburg.

300 All of the wells contain ^3H and ^{85}Kr , indicating a fraction of modern groundwater is present in
301 the mixture sampled by each well. Tritogenic ^3He concentrations range from 6.4 TU in one of
302 the deep wells to 27 TU in one of the shallow wells. The ^{39}Ar activities in the shallow production
303 wells are all over 90 (± 10) percent of the modern atmospheric activities (% modern), while the
304 ^{39}Ar activities of the three deep production wells range from 77 to 51 (± 8) % modern. The deep
305 production well samples also contain a radiogenic helium-4 component of 10-24% of
306 atmospheric equilibrium helium concentrations. Table (1) provides a summary of measured
307 tracer concentrations in the sample wells.

308 *2.4.3 La Selva*

309 La Selva Biological Station is located on the Caribbean coastal plain at the foot of Volcan Barva.
310 Annual precipitation ranges from 4240 mm at the study site, to more than 8000 mm at an
311 elevation of 700 m (Solomon et al., 2010). There is strong evidence that groundwater flow at this
312 site consists of two distinct end-members: (1) high solute bedrock groundwater representing
313 inter-basin groundwater flow and (2) low-solute groundwater derived from recharge that falls
314 within the drainage area of the Sura and Salto streams. The main source of dissolved inorganic
315 carbon is magmatic outgassing. Dissolution of carbonate minerals does not have a significant
316 contribution in Dissolved Inorganic Carbon (DIC) (Genereux et al., 2009). The site characteristics

317 and sampling procedures have been extensively described in (Genereux et al., 2009; Solomon et
318 al., 2010; Webb, 2007). The following tracers were used: ^{14}C , CFC-11, CFC-12, CFC-113, ^3H , SF_6 ,
319 and ^4He . Among these tracers, ^4He is the only one that accumulates in the groundwater as a
320 result of decay of radionuclides ^{238}U , ^{232}Th , and ^{235}U . The historical concentrations of tracers
321 entering the groundwater system have been described in detail before (Solomon et al., 2010).
322 The parameters and tracer concentrations used in the Bayesian modeling are summarized in
323 Table (2). The analysis was performed on samples collected from four wells identified by
324 numbers 7, 11, 16 and 30. The observed concentrations of the tracers at the four sampling
325 locations are listed in Table (2). Different mathematical forms of age distribution has been
326 examined on exactly the same data previously (Massoudieh et al., 2012).

327 **3. Results**

328 *3.1 Synthetic case*

329 Figure 3 shows the hypothetical gamma distribution and the inferred 3, 4 and 10 bin
330 histograms. Calculated concentrations of ^{39}Ar , ^{85}Kr , ^3H and tritogenic ^3He based on the
331 presumed gamma distribution were used to infer the freeform age distribution. In principle, the
332 histogram method is capable of reproducing the mathematical shape of the age distribution.
333 When 4 and 10 bin histograms are used, the inferred histograms fit the younger parts of the age
334 distribution (i.e. more recent years) better than the older parts. The 3 bin histogram does not
335 have the flexibility to represent the early part of the synthetic histogram as well due to the fact
336 that the distribution between 0 and 20 years old water is considered uniform. It should be noted
337 that even though considering a 10-bin histogram results in an over parameterized problem with
338 respect to the values of each bin, due to the constraints between the bin values as a result of
339 applying the tracer data the cumulative distribution is constrained and follow the shape of the
340 true age distribution. It is interesting (although not unexpected) that although in the 10-bin
341 model each bin value can vary within a large range the negative correlation between the
342 adjacent bins is such that their summation does not vary as much. Table (3) shows the measures
343 of model goodness of fit $\text{Log}(I)$ and DIC as well as ξ and b that show how much the inferred

344 distributions deviate from the true gamma distributions when 3, 4 and 10 bin histograms are
345 applied to the hypothetical gamma distribution. $\text{Log}(I)$ is a measure that indicates how good the
346 age distribution represented as a histogram can reproduce the tracer concentrations, so as it is
347 expected, the larger the number of bins results in a higher flexibility of the model in terms of the
348 age distributions it can reproduce which in turns results in a higher the value of $\text{Log}(I)$.
349 However, the relative improvement in the $\text{Log}(I)$ as a result of an increase from four bins to 10
350 bins is smaller than the relative improvement as a result of an increase in the number of bins
351 from 3 to 4. On the other hand DIC is a measure of goodness of models that explicitly accounts
352 for the model complexity. The three bin model has the lowest DIC which means it is the best
353 model. This implies that accounting for the model complexity, a three bin histogram is sufficient
354 for describing the four tracer data obtained based on a gamma age distribution. This result may
355 change if a larger number of tracers are used.

356 ξ and b measure the ability of the model to capture the known true age distribution in terms of
357 spread and bias. It appears that the four bin histogram results in a smaller spread as well as a
358 smaller bias. This means that although the 10 bin histogram results in a better fitting of the
359 tracer concentrations (as measured according to the Bayes factor), $\text{Log}(I)$ it is less capable to
360 reproduce the true age distribution which is an indication of the model being over-
361 parameterized.

362

363 3.4.2 Holten

364 The Bayesian Inference model was used to infer the parameters of seven presumed distribution
365 types: 3, 4 and 10 bin histograms, and exponential, gamma, log-normal and inverse Gaussian
366 distributions. The inverse Gaussian distribution is equivalent to advection dispersion in a one-
367 dimensional system. For all the cases an old fraction is also considered to be present with zero
368 ^{85}Kr , ^3H and tritiogenic ^3He , while the concentration of ^{39}Ar in the old fraction is assumed to be
369 45 percent modern.

370 Figure 4 shows the inferred cumulative age distributions for three of the wells in the Holten site.
371 The wells selected for presentation in this figure cover a wide range of old groundwater
372 fractions. Well 59-05 has a young fraction of 60%-90%, well 72-22 has a young fraction of more
373 than 90% and well 85-35 has 20%-60% young groundwater. In addition to the freeform
374 histogram age distributions, the results of the four mathematical forms of age distributions are
375 also presented. For wells 59-05 and 72-22, all the models infer a very large fraction of water to
376 have an age of less than 20 years while for well 85-35 a more uniform distribution is estimated.
377 The results of 3, 4 and 10 bins are roughly consistent and to a large degree consistent with the
378 mathematical age distribution forms. The variation of the estimated bin contributions is larger
379 for 59-05. The combination of tracer concentrations of 59-05 cannot be well represented by any
380 of the age models, pointing to an internal inconsistency in the tracer data, given the
381 assumptions on the initial concentrations of the tracers in recharging groundwater underlying
382 all the models. This is reflected in relatively low values for $\text{Log}(I)$ and high values for DIC, both
383 indicating poor model performance.

384 The posterior distribution of the old fraction of groundwater is shown in Figure 5 for three of
385 the wells. The posterior old fraction is not significantly affected by the choice of the number of
386 bins. The only exception is for well 72-22 where the 10 bin model results in a much smaller old
387 fraction. This is due to the fact that in this well the fraction of old groundwater is small and for
388 the case of 10 bins the older fraction can be accounted for by the bin closest to the 60 years
389 breakpoint. The same pattern is seen for well 85-35 but to a lesser degree. Figure 6 shows (for
390 well 72-22 as an example) histograms of the posterior distribution of tracer concentrations and
391 observed concentrations (shown as a vertical line). The posterior distribution of ^{39}Ar is lower
392 than the measured value for the 3 bin case and higher for the 10 bin case. In both cases, the
393 posterior distribution is much narrower than the analytical uncertainty of the ^{39}Ar measurement
394 (6-10 percent modern). The performance of 3, 4, and 10 bin models in reproducing the observed
395 tracer concentration of all the wells is summarized in Figure 7. Overall the observed
396 concentrations in all the cases are within the 95% equal-tails Credible Intervals (CIs)(Leray et al.,
397 Accepted) of the posterior distributions. Although in most cases the 10 bin model results in a

398 narrower confidence band for most of the tracers, no definitive conclusion can be made on what
399 number of bins results in an overall better reproduction of the observed tracers or the
400 uncertainty associated with it. In almost all the cases there is a trade-off between reproducing
401 some tracers with better confidence and some with a lesser when a different number of bins is
402 used.

403 Table 4 summarizes the measures of goodness of fit for all the wells and the seven presumed
404 age distributions. According to the DIC measure of goodness of fit, in most cases the log-normal
405 age distribution perform better than the other models, except for well 72-22 where the inverse
406 Gaussian model performs better. This indicates that when the level of complexity is taken into
407 account, mathematical forms can represent important features of the age distribution with a
408 smaller number of parameters. Among the histogram models, in the majority of the wells, the 4
409 bin histogram slightly outperforms both the 3 bin and 10 bin histograms. This means in most
410 cases, 4 bins contain enough complexity that can be determined using the number of tracers
411 used, in this case four. When the Bayes Factor is used to evaluate the performance of the
412 mathematical models, the histograms almost always outperform the mathematical forms.
413 Surprisingly, in most cases the three bin model outperforms a larger number of bins.

414 The MCMC approach also allows the evaluation of the internal correlations of the estimated
415 histogram bin contributions. As an example, scatter plots of accepted a posteriori bin
416 contributions are plotted for well 72-22 for the case when four bins are used (Figure 7). A high
417 correlation between two bins indicates that the method is unable to independently estimate
418 each bin value although it may estimate their sum with good confidence. For example a strong
419 correlation is apparent between bins 1 and 2 and to a smaller degree between bins 3 and 4. In
420 contrast, the correlations between non-adjacent bins and bins 2 and 3 are small. Note that the
421 range of estimated bin contributions over which the correlation is expressed is much larger for
422 bin 1 and bin 2. This is expressed in the cumulative age distribution as a wider range in the
423 estimated contribution of groundwater of 0-15 years, and a smaller range in the estimated
424 contributions of groundwater of 0-30 years (Fig. 4b). Both bin 3 and bin 4 contribute less than
425 0.03 to the total age distribution. This is expressed in the cumulative age distribution as a series

426 of near-horizontal lines between 30 and 60 years. The other combinations of bins do not show a
427 correlation and the Bayesian Inference method is capable of independently estimating these
428 contributions based on the available tracers. When the negative correlation with a slope of
429 around one between two adjacent bins is present, the cumulative distribution may still be
430 expressed with a good confidence although individual bins may not. Correlations between
431 estimated parameters often indicate poorly chosen parameters. Internal correlations of
432 histogram bin estimates could guide the choice of bin-width such that these correlations are
433 minimal.

434 *2.4.3 La Selva*

435 Figure 8 shows the inferred 4 bin age distributions for the four wells considered at the La Selva
436 site. Although the uncertainty associated to the values in each bin is large, some insight can be
437 gained from the results regarding the likely general shape of the distribution. For example, in the
438 case of wells 7, 16 and 30, the Bayesian inference method is incapable of accurately estimating
439 the fraction of groundwater that is younger than 15 years (first age range bin). At the same time,
440 it is better able to estimate the fraction of water that is younger than 30 years (bin 1 and 2
441 combined). This is illustrated by the wide band of possible solutions in the age histogram at 15
442 years, and a narrower band of possible solutions in the cumulative age histogram at 30 years.
443 This effect may indicate that the combination of tracers is not capable of discerning the water
444 from each of the bins. For the case of well 11, the concentration of He and ^{14}C is indicative of a
445 large fraction of groundwater older than 60 years. In this case the inferred age distribution of
446 young fraction (0-60 years old) seems to be close to uniform and the uncertainty regarding the
447 fractions in each bin is such that the method is practically incapable of providing any
448 meaningful information regarding the fractions in each bin. Figure 9 shows a comparison
449 between the observed tracer concentrations and their modeled posterior distribution obtained
450 using the 4 bin model for well 11. In the case of well 11 the model is doing a better job in
451 reproducing observed concentrations of CFC-11, CFC-12, ^3H and ^{14}C . The observed
452 concentration of most other tracers are within the 95% equal-tail credible intervals of the
453 posterior except for He and SF₆. Figure 10 represent the posterior 95% CI brackets obtained

454 using 3, 4 and 10 bin histograms for the four wells considered for the La Selva data. The 3 and 4
455 bin models result in very similar outcomes. The posterior 95% CI brackets obtained using the 10
456 bin model are similar to the ones obtained using 3 and 4 bin models in most cases while in some
457 cases the brackets are narrower. The measures of goodness models $Log(I)$, and DIC for 3, 4 and
458 10 bin models as well as three mathematical forms are summerized in Table 5. As it is expected,
459 the DIC measure tends to attribute a better model goodness index to the simpler models while
460 the $Log(I)$ measure for the more complicated models are better. In general, the best
461 mathematical model outperforms the histogram forms based on both measures of goodness.
462 Among the histogram models 3 bin and 4 bin models alternatively are assigned better DICs
463 while 4 bin model does better according to $Log(I)$ except for the case of well 11 where the 10 bin
464 model outperforms 3 and 4 bin models. When mathematical forms of age distribution are used,
465 the log normal model performs best in terms of DIC for wells 7, 16 and 30 while the simpler
466 exponential model is assigned a lower (better) DIC for the case of well 11.

467

468 2.4.4 Synthesis

469 Because the same tracers were used in the synthetic case and for the Holten dataset, the Bayes
470 Factor and DIC values can be compared directly. All of the 3 bin histograms of the Holten case
471 have a higher DIC than the 3 bin histogram in the synthetic case, showing that the Holten age
472 distributions cannot be approximated by a 3 bin histogram as well as the synthetic case. The
473 three bin models result in comparable values for Bayes Factor for most wells and much better
474 results for some wells (e.g. 72-22, 85-33). In contrast, 4 bin histograms for all Holten wells
475 (except 85-33) have a lower DIC than the synthetic case, showing that the 4 bin histogram is
476 better capable of reproducing the tracers from the Holten wells than from the synthetic gamma
477 distribution. The opposite is the case when analyzing the differences between the Bayes Factor
478 for the Holten and synthetic case. All of the 10 bin histograms in Holten have smaller Bayes
479 Factors compared to the 10 bin results in the hypothetical case, while the comparison between
480 Holten and the synthetic case in terms of DIC yields mixed results.

481 Due to the difference between the number and nature of tracers used at two sites, no direct
482 comparison can be made between the performances of the inferred age distributions at each
483 site. Comparing figures 7 and 11 shows that reproducing the tracer concentrations has been
484 more successful for the Holten data. The uncertainties associated with age distribution are
485 larger for the tracer data collected at La Selva site compared to the Holten site. This can be
486 attributed to the fact that for several of the tracers in La Selva, there are additional uncertainties
487 due to the environmental factors such as the adsorption of CFCs and ^{14}C to the soil matrix, and
488 the decay rates of CFCs.

489 **4. Discussion and conclusions**

490 Since it is not possible to estimate the exact groundwater age distribution from a limited
491 number of age tracer measurements, a mathematical form (e.g. exponential or inverse-Gaussian)
492 is often presumed. Either deterministic or stochastic approaches are then used to infer the
493 parameters defining it. Bayesian inference provides a powerful tool to assess the uncertainties
494 of such estimates and evaluate or compare the appropriateness of the presume age distributions
495 using measures such as DIC and Bayes factor. The DIC measure considers the complexity of the
496 models when assigning the measure while the Bayes Factor does not do that in an explicit way
497 although the model complexity is factored in implicitly. When these measures have been used
498 in the past to determine the best mathematical form representing an age distribution for a single
499 site, no definitive conclusion has been made on the best form of the age distribution function
500 describing the combination of tracer concentration in all of the wells. This is due to the fact that
501 each well may be extracting groundwater from different sources, following different travel
502 paths and undergoing different mixing than the other wells, and also due to the effect of local
503 heterogeneities at the proximity of the wells on the age distribution at each well. It is therefore
504 desirable to allow the inference technique to be able to assign the most appropriate form of the
505 age distribution, rather than it to be pre-determined. In this study we examined a Markov Chain
506 Monte Carlo method to infer the groundwater age distribution from tracer data using a free-
507 form (histogram) age distribution.

508 The method was first applied to a hypothetical case where the tracer concentrations are
509 obtained based on a known gamma age distribution. The combination of ^{85}Kr , ^3H , ^3He and ^{39}Ar
510 was sufficient to reproduce the shape of the mathematical age distribution with a 3 bin or 4 bin
511 age histogram. A 10 bin age histogram showed larger deviations from the original cumulative
512 distribution between 30 and 50 years, as a result of deviations of the estimated fraction of 10-25
513 years. Clearly, the 10 bin age histogram was not sufficiently constrained by the four age tracers.
514 Based on quantitative measures of discrepancy between inferred and true age distribution, the 4
515 bin model provides the closest approximation. While the 10 bin histogram is better capable of
516 reproducing the tracer data, it does not reproduce the true age distribution as well as 3 and 4
517 bin models. This is an indication that the 10 bin model is over-parameterized and the inferred
518 age distributions are non-unique.

519 The method was also applied to the tracer data collected from seven wells at the Holten site in
520 Netherlands where four tracers (^{85}Kr , ^3H , ^3He , and ^{39}Ar) were used, and the La Selva site in
521 Costa Rica where seven tracers including three CFCs, SF_6 , ^{14}C , He and ^3H were used. For each
522 sample, two measures of model effectiveness (DIC and Bayes factor) were used to compare the
523 performance of 3, 4 and 10 bin histogram age distributions as well as mathematical age
524 distributions (exponential, gamma, inverse-Gaussian and log-normal). For most wells at both
525 sites, the log-normal age distribution performed better than the rest of the mathematical
526 distributions as well as the histogram forms. Among the histogram forms, in most cases the four
527 bin and in some cases three bin models performed best when the model complexity was taken
528 into consideration using the DIC measure. It seems that the 10 bin histogram is over-
529 parameterized and the three bin histogram is not adequate for describing the observed tracer
530 concentrations in most cases. When the logarithm of Bayes Factors are used for evaluating the
531 models, models with higher complexity (i.e. 10 bin models) are given the preference since
532 complexity is not explicitly considered

533 Ideally, a multitude of age models is investigated to convert the measured tracer concentrations
534 to a groundwater age distribution. An ensemble of parallel models will represent the model

535 error – no model is capable of capturing the actual age distribution – rather than merely the
536 parameterization error associated with a single model.

537 The age histogram approach provides flexibility to model any shape of age distribution, within
538 the limitations of assigning a uniform distribution within each bin. Due to simplicity of the
539 model, more parameters are required to closely mimic a typical groundwater age distribution,
540 such as the exponential model while the number of parameters (histogram bins) that can be
541 confidently estimated is limited by the number of measurements of distinctly different
542 groundwater age tracer that are available. Age distributions represented by larger number of
543 bins can become preferred if a larger number of tracers or water samples taken at different
544 times are available. The observation that the mathematical forms generally perform better than
545 the histogram form can be attributed to the fact that they can capture smooth features of the
546 true groundwater age distribution more effectively using a smaller number of parameters.
547 However, this does not undermine the value of the freeform histograms since they free us from
548 testing a number of conceivable mathematical forms to find the best ones. Age histograms also
549 do not infer the shape of the pre-modern groundwater age distribution for which often no
550 tracers are available, unlike some of the mathematical models that infer an old-tail into infinity
551 (e.g. the exponential model).

552 Age histograms are also potentially capable of inferring the shape of groundwater age when it
553 does not conform to any conventional mathematical distribution. A larger amount of tracer data
554 can support more complex models and the flexibility of the histogram approach has been
555 demonstrated to be able to reconstruct a highly detailed age distribution based on electrical
556 conductivity time series (Cirpka et al., 2007; Liao and Cirpka, 2011).

557 The uncertainties affecting inference of age distribution can be classified into three categories:
558 The first category is *observation and measurement error* which stems from the uncertainty in the
559 analytical methods as well as the uncertainty due to local and temporal heterogeneities in the
560 concentration of tracers. The second category of uncertainty is referred to as *external forcing*
561 uncertainty. This includes 1) assumptions about the historical (decay corrected) concentrations
562 of tracers in precipitation or the atmosphere, 2) transport of the tracers through the unsaturated

563 zone, leading to delay and dispersion of the signal, 3) non-conservative transport of tracers,
564 caused by sorption, degradation or diffusive exchange. The third category of uncertainty is the
565 *model structural error* which is due to the fact that any presumed form of age distribution is at
566 best an approximation of the real form. In complicated systems with large heterogeneities, at
567 the proximity of the sampling point the true form of age distribution has been shown to be
568 more complex than any of the commonly used mathematical forms (Weissmann et al., 2002).

569 Measurement and analytical uncertainty is often the best constrained source of uncertainty.
570 Some aspects of the second category of uncertainty may be investigated separately. Availability
571 of local or nearby sampling stations of atmospheric tracer mixing ratios (Busenberg and
572 Plummer, 2000) or concentrations in precipitation (Kralik et al., 2003) reduces the uncertainty in
573 the tracer input function. Incorporating the uncertainty of the tracer input functions is
574 challenging because these are time series rather than scalar numbers. In the simplest case, the
575 reconstructed input function is biased. However, local contamination of atmospheric mixing
576 ratios due to nearby industrial activities (Ho and Schlosser, 2000) can be variable in time,
577 affecting only a portion of the tracer history. Although in this work, this source of uncertainty is
578 ignored, developing methods to include this in the analysis is critical.

579 Sampling tracer concentrations in the unsaturated zone (Foster and Smith-Carington, 1980) or in
580 short screened monitoring wells with an assumed narrow age distribution (Visser et al., 2013)
581 can reduce the uncertainty in the dispersion and lag of tracers in the unsaturated zone (Cook
582 and Bohlke, 2000; Corcho Alvarado et al., 2007). A combination of age tracers from such
583 sampling points can demonstrate the non-conservative behavior of some of the tracers (e.g.
584 Bauer et al., 2001). Incorporating the uncertainty of the non-conservative transport behavior in
585 the unsaturated zone and subsurface is also challenging (Visser et al. 2009; 2013). Because of
586 variations in mean water tables and the heterogeneity of subsurface properties in the capture
587 area, effective parameters representing the entire aquifer need to be applied to describe such
588 processes.

589 Without such supplementary information, separating the first and second category from the
590 model structural error is a challenging task. In most stochastic inverse methods applied to

591 parameter estimation of hydrogeological models they are lumped together. Separating these
592 sources of uncertainty can be of enormous value since it provides a way to assess the value of
593 information content of the tracer data and also a more direct way to evaluate various model
594 structures.

595 **Acknowledgements**

596 Part of this work was performed under the auspices of the U.S. Department of Energy by
597 Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. LLNL-JRNL-
598 637575.

599 **5. References**

- 600 Bauer, S., Fulda, C. and Schafer, W., 2001. A multi-tracer study in a shallow aquifer using age dating
601 tracers ^3H , ^{85}Kr , CFC-113 and SF_6 - Indication for retarded transport of CFC-113. *Journal of*
602 *Hydrology*, 248(1-4): 14-34.
- 603 Bethke, C.M. and Johnson, T.M., 2008. Groundwater age and groundwater age dating. *Annual Review of*
604 *Earth and Planetary Sciences*, 36: 121-152.
- 605 Beven, K.J., 2010. Preferential flows and travel time distributions: defining adequate hypothesis tests for
606 hydrological process models. *Hydrological Processes*, 24(12): 1537-1547.
- 607 Bohlke, J.K. and Denver, J.M., 1995. COMBINED USE OF GROUNDWATER DATING, CHEMICAL, AND
608 ISOTOPIC ANALYSES TO RESOLVE THE HISTORY AND FATE OF NITRATE CONTAMINATION IN 2
609 AGRICULTURAL WATERSHEDS, ATLANTIC COASTAL-PLAIN, MARYLAND. *Water Resources*
610 *Research*, 31(9): 2319-2339.
- 611 Broers, H.P. and G., F.C.V., 2005. Monitoring Strategies at Phreatic Wellfields: A 3D Travel Time
612 Approach. *Ground Water*, 43(6): 850-862.
- 613 Busenberg, E. and Plummer, L.N., 1992. Use of Chlorofluorocarbons (CCl_3F and CCl_2F_2) as Hydrologic
614 Tracers and Age-Dating Tools: the Alluvium and Terrace System of Central Oklahoma. *Water*
615 *Resources Research*, 28(9): 2257-2283.
- 616 Busenberg, E. and Plummer, L.N., 2000. Dating young groundwater with sulfur hexafluoride: Natural and
617 anthropogenic sources of sulfur hexafluoride. *Water Resources Research*, 36(10): 3011-3030.
- 618 Carlin, B.P. and Chib, S., 1995. BAYESIAN MODEL CHOICE VIA MARKOV-CHAIN MONTE-CARLO
619 METHODS. *Journal of the Royal Statistical Society Series B-Methodological*, 57(3): 473-484.
- 620 Castro, M.C., Jambon, A., de Marsily, G. and Schlosser, P., 1998. Noble gases as natural tracers of water
621 circulation in the Paris Basin: 1. Measurements and discussion of their origin and mechanisms of
622 vertical transport in the basin. *Water Resour. Res.*, 34(10): 2443-2466.
- 623 Cirpka, O.A. and Attinger, S., 2003. Effective dispersion in heterogeneous media under random transient
624 flow conditions. *Water Resources Research*, 39(9).

625 Cirpka, O.A. et al., 2007. Analyzing Bank Filtration by Deconvoluting Time Series of Electric Conductivity.
626 Ground Water, 45(3): 318-328.

627 Cook, P.G. and Bohlke, J.K., 2000. Determining timescales for groundwater flow and solute transport. In:
628 P.G. Cook and A.L. Herczeg (Editors), Environmental Tracers in Subsurface Hydrology. Kluwer
629 Academic Publishers, Boston, pp. p. 1-30.

630 Corcho Alvarado, J.A. et al., 2007. Constraining the age distribution of highly mixed groundwater using
631 ³⁹Ar: A multiple environmental tracer (³H/³He, ⁸⁵Kr, ³⁹Ar, and ¹⁴C) study in the semiconfined
632 Fontainebleau Sands Aquifer (France). Water Resources Research, 43(3): W03427.

633 Engdahl, N.B., Ginn, T.R. and Fogg, G.E., 2012. Non-Fickian dispersion of groundwater age. Water
634 Resources Research, 48(7): W07508.

635 Engdahl, N.B., Ginn, T.R. and Fogg, G.E., 2013. Using groundwater age distributions to estimate the
636 effective parameters of Fickian and non-Fickian models of solute transport. Advances in Water
637 Resources, 54(0): 11-21.

638 Engesgaard, P., Jensen, K.H., Molson, J., Frind, E.O. and Olsen, H., 1996. Large-scale dispersion in a sandy
639 aquifer: Simulation of subsurface transport of environmental tritium. Water Resources
640 Research, 32(11): 3253-3266.

641 Foster, S.S.D. and Smith-Carington, A., 1980. The interpretation of tritium in the Chalk unsaturated zone.
642 Journal of Hydrology, 46(3&4): 343-364.

643 Fox, J.F. and Papanicolaou, A.N., 2008. An un-mixing model to study watershed erosion processes.
644 Advances in Water Resources, 31(1): 96-108.

645 Gamerman, D. and Hedibert, F.L., 2006. Markov Chain Monte Carlo - Stochastic Simulation for Bayesian
646 Inference. Chapman & Hall/CRC.

647 Gelhar, L.W., Welty, C. and Rehfeldt, K.R., 1992. A critical review of data on field-scale dispersion in
648 aquifers. Water Resources Research, 28(7): 1955-1974.

649 Genereux, D.P., Webb, M. and Solomon, D., 2009. Chemical and isotopic signature of old groundwater
650 and magmatic solutes in a Costa Rican rain forest: Evidence from carbon, helium, and chlorine.
651 Water Resources Research, 45(8): W08413.

652 Glynn, P.D. and Plummer, L.N., 2005. Geochemistry and the understanding of ground-water systems.
653 Hydrogeology Journal, 13: 263-287.

654 Goode, D.J., 1996. Direct simulation of groundwater age. Water Resources Research, 32(2): 289-296.

655 Hansen, B., Thorling, L., Dalgaard, T. and Erlandsen, M., 2010. Trend Reversal of Nitrate in Danish
656 Groundwater - a Reflection of Agricultural Practices and Nitrogen Surpluses since 1950.
657 Environmental Science & Technology, 45(1): 228-234.

658 Hinsby, K. et al., 2001. The modern water interface: recognition, protection and development—advance
659 of modern waters in European aquifer systems. Geological Society, London, Special Publications,
660 189(1): 271.

661 Ho, D.T. and Schlosser, P., 2000. Atmospheric SF₆ near a large urban area. Geophysical Research Letters,
662 27(11): 1679-1682.

663 Horneman, A. et al., 2008. Degradation rates of CFC-11, CFC-12 and CFC-113 in anoxic shallow aquifers
664 of Arahazar, Bangladesh. Journal of Contaminant Hydrology, 97(1-2): 27-41.

665 Hua, Q. and Barbetti, M., 2004. Review of tropospheric bomb (super 14) C data for carbon cycle
666 modeling and age calibration purposes. Radiocarbon, 46(4): 1273-1298.

667 IAEA/WMO, 2010. Global Network of Isotopes in Precipitation. The GNIP Database. Accessible at:
668 <http://isohis.iaea.org>.

669 Jeffreys, H., 1935. Some tests of significance, treated by the theory of probability. Proceedings of the
670 Cambridge Philosophical Society, 31: 203-222.

671 Jung, M., Wieser, M., von Oehsen, A. and Aeschbach-Hertig, W., 2012. Properties of the closed-system
672 equilibration model for dissolved noble gases in groundwater. Chemical Geology(0).

673 Kaipio, J. and Somersale, E., 2004. Statistical and Computational Inverse Problems (Applied
674 Mathematical Sciences) (v. 160). Springer.

675 Kass, R.E. and Raftery, A.E., 1995. BAYES FACTORS. *Journal of the American Statistical Association*,
676 90(430): 773-795.

677 Kralik, M. and Keimel, T., 2003. Time-input, an innovative groundwater-vulnerability assessment
678 scheme: application to an alpine test site. *Environmental Geology*, 44(6): 679-686.

679 Kralik, M., Papesch, W. and Stichler, W., 2003. Austrian Network of Isotopes in Precipitation (ANIP):
680 Quality assurance and climatological phenomenon in one of the oldest and densest networks in
681 the world (IAEA-CN-104/157), *Isotope Hydrology and Integrated Water Resources Management*.
682 International Atomic Energy Agency and the International Association of Hydrogeologists,
683 Vienna, pp. 146-149.

684 Laier, T., 2004. Nitrate monitoring and CFC-age dating of shallow groundwaters - an attempt to check
685 the effect of restricted use of fertilizers. . In: L. Razowska-Jaworek and A. Sadurski (Editors),
686 Nitrate in Groundwaters, IAH Selected papers on hydrogeology. A.A. Balkema Leiden, the
687 Netherlands, pp. 247-258.

688 Larocque, M., Cook, P.G., Haaken, K. and Simmons, C.T., 2009. Estimating Flow Using Tracers and
689 Hydraulics in Synthetic Heterogeneous Aquifers. *Ground Water*, 47(6): 786-796.

690 Lehmann, B.E. et al., 2003. A comparison of groundwater dating with ^{81}Kr , ^{36}Cl and ^4He in four wells of
691 the Great Artesian Basin, Australia. *Earth and Planetary Science Letters*, 211(3-4): 237-250.

692 Leray, S. et al., Accepted. Temporal evolution of age data under transient pumping conditions. *Journal of*
693 *Hydrology*.

694 Liao, Z. and Cirpka, O.A., 2011. Shape-free inference of hyporheic traveltime distributions from synthetic
695 conservative and "smart" tracer tests in streams. *Water Resour. Res.*, 47(7): W07510.

696 Loosli, H.H., 1983. A dating method with ^{39}Ar . *Earth and Planetary Science Letters*, 63(1): 51-62.

697 Loosli, H.H., Lehmann, B.E. and Balderer, W., 1989. Argon-39, argon-37 and krypton-85 isotopes in Stripa
698 groundwaters. *Geochimica Et Cosmochimica Acta*, 53(8): 1825-1829.

699 Maiss, M. and Brenninkmeijer, C.A.M., 1998. Atmospheric SF_6 : Trends, sources, and prospects.
700 *Environmental Science and Technology*, 32(20): 3077-3086.

701 Maloszewski, P. and Zuber, A., 1982. Determining the turnover time of groundwater systems with the
702 aid of environmental tracers : 1. Models and their applicability. *Journal of Hydrology*, 57(3-4):
703 207-231.

704 Maloszewski, P. and Zuber, A., 1993. Principles and practice of calibration and validation of
705 mathematical models for the interpretation of environmental tracer data in aquifers. *Advances*
706 *in Water Resources*, 16(3): 173-190.

707 Maloszewski, P. and Zuber, A., 1998. A general lumped parameter model for the interpretation of tracer
708 data and transit time calculation in hydrologic systems - Comments. *Journal of Hydrology*, 204(1-
709 4): 297-300.

710 Manning, A.H., Solomon, D.K. and Thiros, S.A., 2005. H-3/He-3 age data in assessing the susceptibility of
711 wells to contamination. *Ground Water*, 43(3): 353-367.

712 Massoudieh, A. and Ginn, T.R., 2011. The theoretical relation between unstable solutes and
713 groundwater age. *Water Resour. Res.*, 47(10): W10523.

714 Massoudieh, A. and Kayhanian, M., 2013. Bayesian Chemical Mass Balance Method for Surface Water
715 Contaminant Source Apportionment. *Journal of Environmental Engineering*, 139(2): 250-260.

716 Massoudieh, A., Mathew, A. and Ginn, T.R., 2008. Column and batch reactive transport experiment
717 parameter estimation using a genetic algorithm. *Computers & Geosciences*, 34(1): 24-34.

718 Massoudieh, A., Sharifi, S. and Solomon, D.K., 2012. Bayesian evaluation of groundwater age distribution
719 using radioactive tracers and anthropogenic chemicals. *Water Resources Research*, 48.

720 Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E., 1953. Equations of State
721 Calculations by Fast Computing Machines. *Journal of Chemical Physics*, 21(6): 1087-1092.

722 Molina-Giraldo, N., Bayer, P., Blum, P. and Cirpka, O.A., 2010. Propagation of Seasonal Temperature
723 Signals into an Aquifer upon Bank Infiltration. *Ground Water*: no-no.

724 Morgenstern, U., Stewart, M.K. and Stenger, R., 2010. Dating of streamwater using tritium in a post
725 nuclear bomb pulse world: continuous variation of mean transit time with streamflow. *Hydrol.*
726 *Earth Syst. Sci.*, 14(11): 2289-2301.

727 Oeschger, H. et al., 1974. ³⁹Ar dating of groundwater, *Isotope Techniques in Groundwater Hydrology*.
728 International Atomic Energy Agency, Vienna, pp. 179-190.

729 Plummer, L. et al., 2012. Old groundwater in parts of the upper Patapsco aquifer, Atlantic Coastal Plain,
730 Maryland, USA: evidence from radiocarbon, chlorine-36 and helium-4. *Hydrogeology Journal*,
731 20(7): 1269-1294.

732 Plummer, L.N. and Busenberg, E., 2006. Chlorofluorocarbons in the atmosphere, Use of
733 Chlorofluorocarbons in Hydrology – A Guidebook. IAEA, Vienna., pp. 9-14.

734 Plummer, L.N. et al., 2001. Groundwater residence times in Shenandoah National Park, Blue Ridge
735 Mountains, Virginia, USA: a multi-tracer approach. *Chemical Geology*, 179(1-4): 93-111.

736 Poreda, R.J., Cerling, T.E. and Salomon, D.K., 1988. Tritium and Helium-Isotopes as Hydrologic Tracers in
737 a Shallow Unconfined Aquifer. *Journal of Hydrology*, 103(1-2): 1-9.

738 Rinaldo, A. et al., 2011. Catchment travel time distributions and water flow in soils. *Water Resour. Res.*,
739 47(7): W07537.

740 Schlosser, P., Stute, M., Dorr, H., Sonntag, C. and Munnich, K.O., 1988. Tritium/³He Dating of Shallow
741 Groundwater. *Earth and Planetary Science Letters*, 89(3-4): 353-362.

742 Singleton, M.J. and Moran, J.E., 2010. Dissolved noble gas and isotopic tracers reveal vulnerability of
743 groundwater in a small, high-elevation catchment to predicted climate changes. *Water Resour.*
744 *Res.*, 46: W00F06.

745 Smethie, J.W.M., Solomon, D.K., Schiff, S.L. and Mathieu, G.G., 1992. Tracing groundwater flow in the
746 Borden aquifer using krypton-85. *Journal of Hydrology*, 130(1-4): 279-297.

747 Solomon, D.K., Genereux, D.P., Plummer, L.N. and Busenberg, E., 2010. Testing mixing models of old and
748 young groundwater in a tropical lowland rain forest with environmental tracers. *Water*
749 *Resources Research*, 46.

750 Solomon, D.K., Poreda, R.J., Schiff, S.L. and Cherry, J.A., 1992. Tritium and helium 3 as groundwater age
751 tracers in the Borden aquifer. *Water Resources Research*, 28(3): 741-755.

752 Spiegelhalter, D.J., Best, N.G., Carlin, B.R. and van der Linde, A., 2002. Bayesian measures of model
753 complexity and fit. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 64:
754 583-616.

755 Sültenfuß, J., Purtschert, R. and Fühnböter, J., 2011. Age structure and recharge conditions of a coastal
756 aquifer (northern Germany) investigated with ³⁹Ar, ¹⁴C, ³H, He isotopes and Ne. *Hydrogeology*
757 *Journal*, 19(1): 221-236.

758 Sun, T., Hall, C.M. and Castro, M.C., 2010. Statistical properties of groundwater noble gas paleoclimate
759 models: Are they robust and unbiased estimators? *Geochemistry, Geophysics, Geosystems*,
760 11(2): Q02002.

761 Troldborg, L., Jensen, K.H., Engesgaard, P., Refsgaard, J.C. and Hinsby, K., 2008. Using Environmental
762 Tracers in Modeling Flow in a Complex Shallow Aquifer System. *Journal of Hydrologic*
763 *Engineering*, 13(11): 1037-1048.

764 Visser, A., Broers, H.P., Purtschert, R., Sültenfuß, J. and De Jonge, M., 2013. Groundwater age
765 distributions at a public drinking water supply well field derived from multiple age tracers (⁸⁵Kr,
766 ³H/³He, noble gases and ³⁹Ar). *Water Resources Research*, in review.

- 767 Visser, A., Broers, H.P., Van der Grift, B. and Bierkens, M.F.P., 2007. Demonstrating Trend Reversal of
768 Groundwater Quality in Relation to Time of Recharge determined by $^3\text{H}/^3\text{He}$. *Environmental*
769 *Pollution*, 148 (3): 797-807.
- 770 Visser, A., Heerdink, R., Broers, H.P. and Bierkens, M.F.P., 2009. Travel Time Distributions Derived from
771 Particle Tracking in Models Containing Weak Sinks. *Ground Water*, 47(2): 237-245.
- 772 Vrugt, J.A., Bouten, W. and Weerts, A.H., 2001. Information Content of Data for Identifying Soil Hydraulic
773 Parameters from Outflow Experiments. *Soil Sci. Soc. Am. J.*, 65(1): 19-27.
- 774 Vrugt, J.A., ter Braak, C.J.F., Clark, M.P., Hyman, J.M. and Robinson, B.A., 2008. Treatment of input
775 uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo
776 simulation. *Water Resources Research*, 44(12): W00B09.
- 777 Walker, S.J., Weiss, R.F. and Salameh, P.K., 2000. Reconstructed histories of the annual mean
778 atmospheric mole fractions for the halocarbons CFC-11, CFC-12, CFC-113, and carbon
779 tetrachloride. *Journal of Geophysical Research-Oceans*, 105(C6): 14285-14296.
- 780 Wassenaar, L.I., Hendry, M.J. and Harrington, N., 2006. Decadal geochemical and isotopic trends for
781 nitrate in a transboundary aquifer and implications for agricultural beneficial management
782 practices. *Environmental Science and Technology*, 40(15): 4626-4632.
- 783 Webb, M.D., 2007. Carbon, Chlorine, and Oxygen Isotopes as Tracers of Interbasin Groundwater Flow at
784 La Selva Biological Station, Costa Rica, North Carolina State University, Raleigh, North Carolina.
- 785 Weissmann, G.S., Zhang, Y., LaBolle, E.M. and Fogg, G.E., 2002. Dispersion of groundwater age in an
786 alluvial aquifer system. *Water Resour. Res.*, 38(10): 1198.
- 787 Woolfenden, L.R. and Ginn, T.R., 2009. Modeled Ground Water Age Distributions. *Ground Water*,
788 9999(9999).
- 789 Zeng, L., Shi, L., Zhang, D. and Wu, L., 2012. A sparse grid based Bayesian method for contaminant
790 source identification. *Advances in Water Resources*, 37(0): 1-9.
- 791 Zoellmann, K., Kinzelbach, W. and Fulda, C., 2001. Environmental tracer transport (H-3 and SF6) in the
792 saturated and unsaturated zones and its use in nitrate pollution management. *Journal of*
793 *Hydrology*, 240(3-4): 187-205.

794

795

796

797 Tables

798

799 **Table 1: Tracer concentrations for the samples from seven wells at the Holten site.**

ID	top m below surface	bottom	^{85}Kr		^3H		$^3\text{He}_{\text{trit}}$		^{39}Ar	
			dpm/ccKr	+/-	TU	+/-	TU eq.	+/-	% modern	+/-
59-05	19	30	38	1.2	6.48	0.14	21.1		100	10
67-19	14	23	39.7	1.5	6.09	0.15	6.4	0.5	104	8
72-22	15	34	35.9	1.3	6.75	0.16	14.0	0.5	93	6
73-29	16	31	39.4	1.6	6.38	0.16	8.7	0.5	100	10
85-33	46	68	2.7	0.2	1.23	0.05	9.2	0.5	51	8
85-34	48	74	8.4	0.5	3.98	0.11	21.1	0.5	77	7
85-35	45	75	8.7	0.3	3.82	0.11	20.4	0.5	74	6

800

801

802

Table 2: Observed tracer concentrations and the prior information of transport parameters used in La Selva analysis.

	SF6 (fmol/L)	CFC-11 (pmol/kg)	CFC-12 (pmol/kg)	CFC-113 (pmol/kg)	³ H (TU)	¹⁴ C (pmC)	δ ¹³ C	⁴ He (cm ³ STP/g)
Well 7	1.1	3.02	1.66	0.257	0.61	117.1	-26.00	4.81 × 10 ⁻⁸
Well 11	0.169	0.446	0.138	0.0172	0.36	21.7	-7.85	2.50 × 10 ⁻⁷
Well 16	0.931	3.02	1.62	0.243	0.72	116.9	-24.34	5.06 × 10 ⁻⁸
Well 30	1.18	3.13	1.45	0.187	0.66	83.4	-20.20	7.41 × 10 ⁻⁸
K _{oc} (g/g ⁻¹)	195	97	356	316				0
Log K _{ow}	0.226	2.53	2.16	3.16				0
decay rate (yr ⁻¹)					5.6×10 ⁻²	1.2×10 ⁻⁴		2.77×10 ⁻¹⁰
Prior <i>c₀</i>						70		4.5×10 ⁻⁷
Biogenic δ ¹³ C	-25							
Rock δ ¹³ C	-2.4							

Table 3: Measured of goodness of fit for the hypothetical gamma age distribution represented using 3, 4, and 10 bin histograms. Log (I) and DIC reflect the ability of the model to reproduce the synthetic tracer concentrations, ξ and b are the spread and bias of the histograms around the actual hypothetical gamma age distribution.

	$Log(I)$	DIC	ξ	b
3 bins	21.37	174	0.0514	0.0380
4 bins	25.83	198	0.0259	0.0252
10 bins	26.67	209	0.0884	0.0786

Table 4: Measures of goodness of fit $\text{Log}(I)$ and DIC for different age distributions used to describe the tracer concentrations in the seven wells at the Holten site.

	well 59-05		well 67-19		well 72-22		well 73-29		well 85-33		well 85-34		well 85-35	
	$\text{Log}(I)$	DIC	$\text{Log}(I)$	DIC	$\text{Log}(I)$	DIC								
3 bins	21.25	183.60	21.71	189.79	25.79	430.03	22.66	201.84	24.25	737.15	20.02	180.59	20.43	209.11
4 bins	21.09	179.62	20.09	152.31	24.84	196.01	20.76	180.38	24.80	429.38	21.03	186.66	21.46	190.35
10 bins	21.10	182.46	21.42	183.40	25.27	213.82	23.82	204.42	24.57	395.623	22.55	183.51	24.44	302.80
Exponential	20.83	253.78	20.10	236.09	21.28	404.12	20.60	226.72	22.90	804.62	19.28	146.22	19.68	148.69
Log-normal	19.88	131.80	20.34	123.46	25.16	189.72	22.01	127.00	22.82	219.98	18.78	124.41	19.22	128.45
Gamma	19.21	114.58	21.96	145.1	21.52	119.27	21.67	131.49	18.31	122.55	19.82	146.95	19.41	130.87
Inv Gaussian	20.53	135.01	22.88	165.13	26.36	179.49	22.91	141.81	22.47	244.89	19.81	131.95	20.21	153.15

Table 5: Measures of goodness of fit $Log(I)$ and DIC for 3, 4 and 10 bin age histograms as well as three mathematical models for four wells at the La Selva site.

	well 7		well 11		well 16		well 30	
	$Log(I)$	DIC	$Log(I)$	DIC	$Log(I)$	DIC	$Log(I)$	DIC
3 bins	68.47	629.68	71.43	684.82	71.14	650.95	70.79	651.66
4 bins	68.70	630.84	71.52	675.56	71.44	652.33	71.12	651.07
10 bins	68.22	641.67	71.58	727.08	70.43	655.96	70.57	661.94
Exponential	69.42	617.56	71.63	653.85	71.92	637.81	71.83	638.48
Log-normal	68.52	590.63	71.88	680.51	72.21	612.23	71.12	618.26
Gamma	69.95	594.89	70.89	655.60	72.77	618.46	72.22	617.98

Figures:

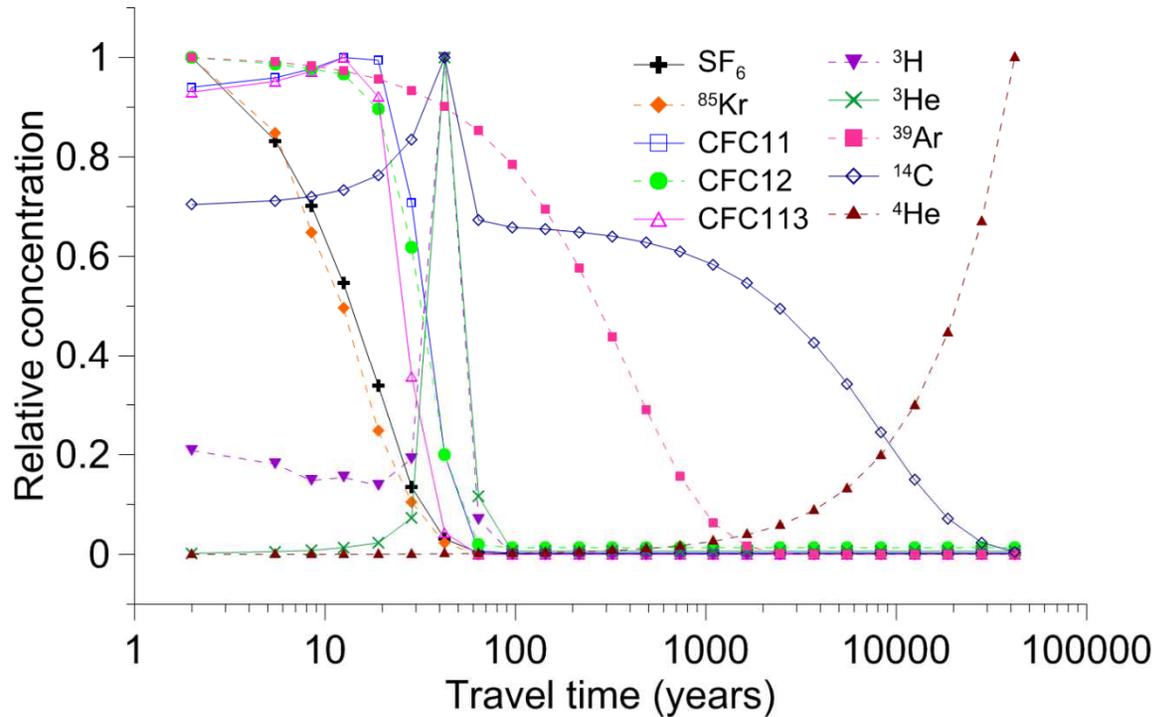


Figure 1: Input functions of common groundwater age tracers on a logarithmic time scale. Tracer concentrations are averaged over exponentially increasing age bins, and then scaled to the maximum historical concentration. Tracer concentrations obtained from (Hua and Barbetti, 2004 (¹⁴C); IAEA/WMO, 2010 (³H); Maiss and Brenninkmeijer, 1998 (SF₆); Plummer and Busenberg, 2006 (CFCs); Walker et al., 2000 (CFCs))and Institute of Atmospheric Research (IAR), Freiburg, Germany (⁸⁵Kr)

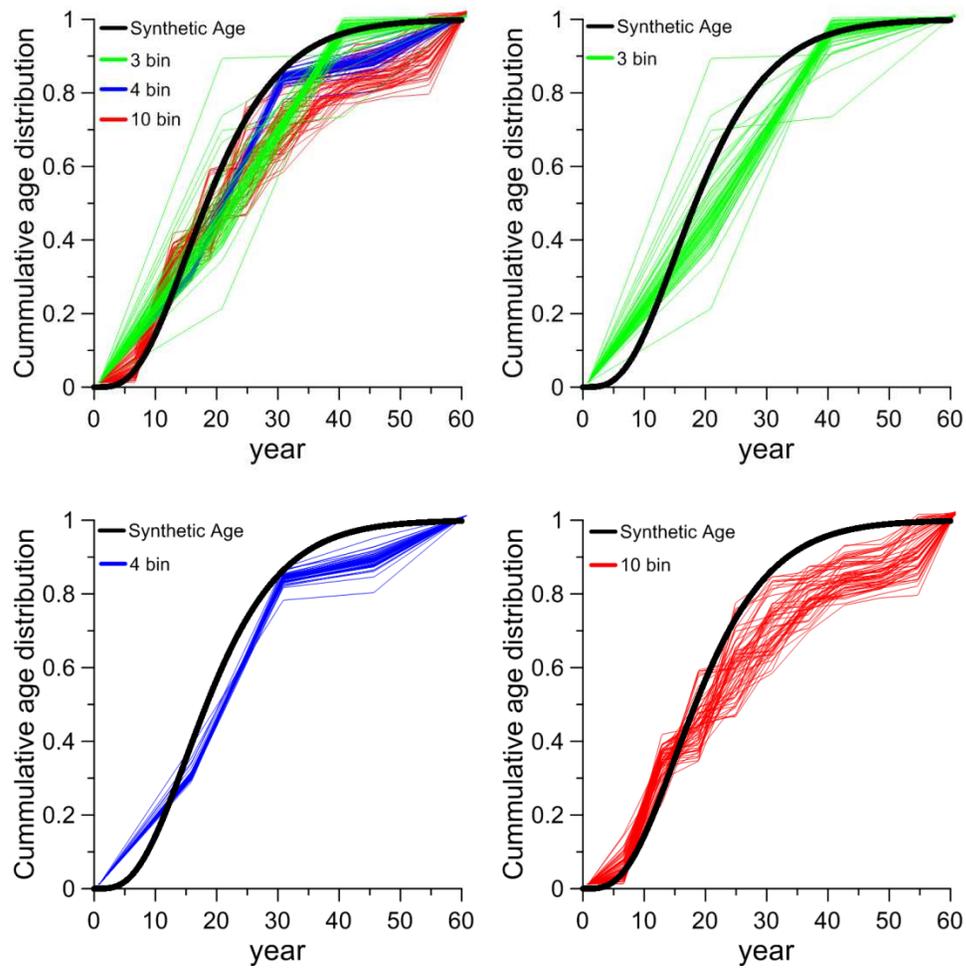


Figure 2: Hypothetical Gamma distribution with parameters $\lambda = 4$ and $k = 5$ and 60 samples from the posterior inferred distributions using b) 3 bins, c) 4 bins and d) 10 bins histograms and a) all three overlaid.

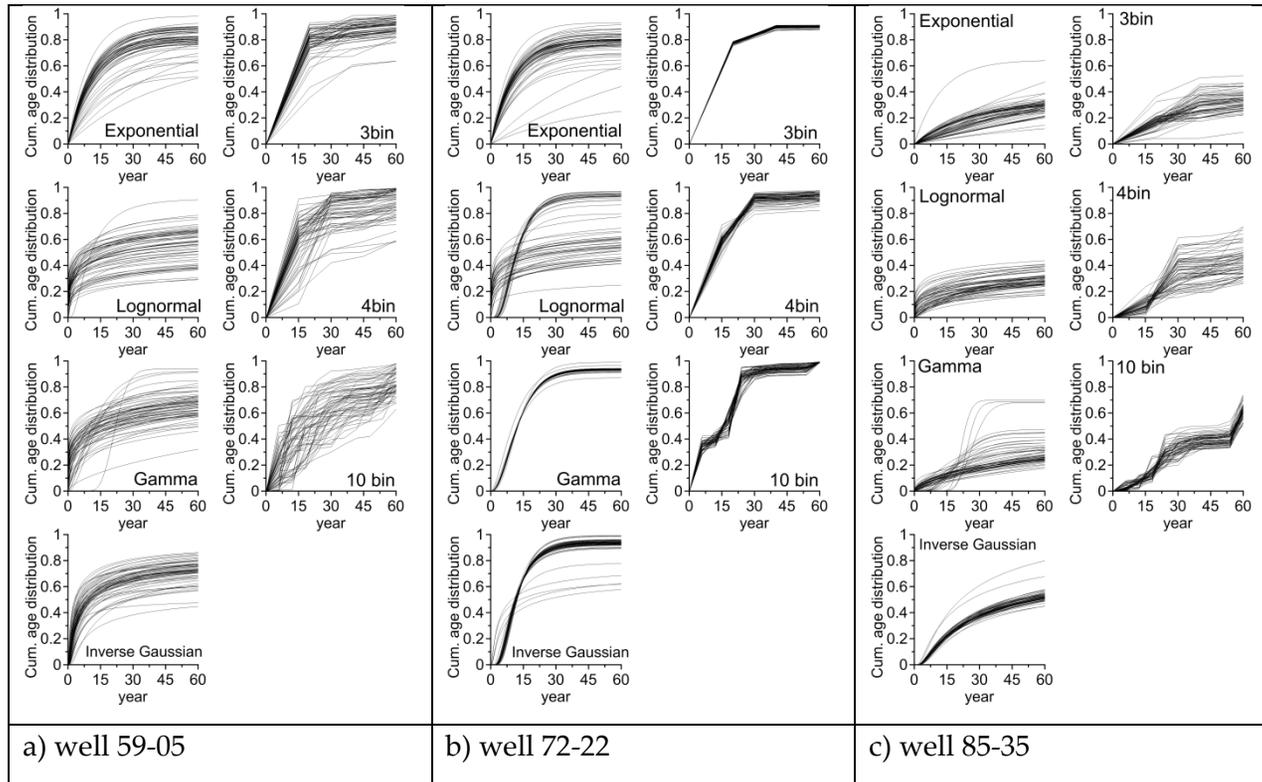


Figure 3: Inferred exponential, lognormal, gamma age distributions as well as 3, 4, and 10 bin histograms for tracer data at wells 59-05, 72-22 and 85-35 at the Holten site.

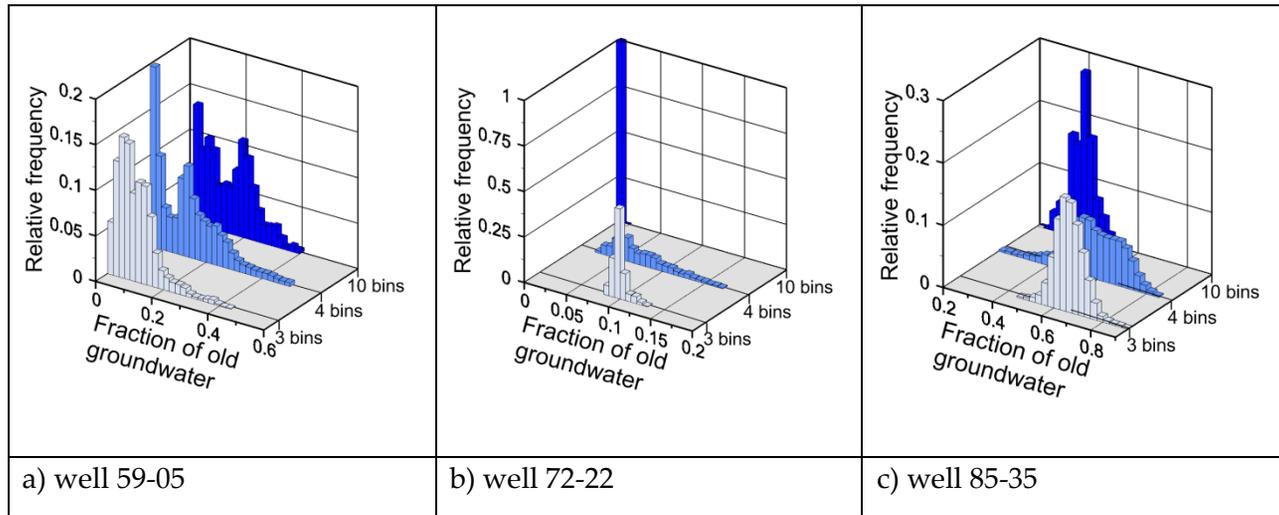


Figure 4: Posterior distribution of old fraction of groundwater for a) well 59-05, b) well 72-22 and c) well 85-33 at the Holten site.

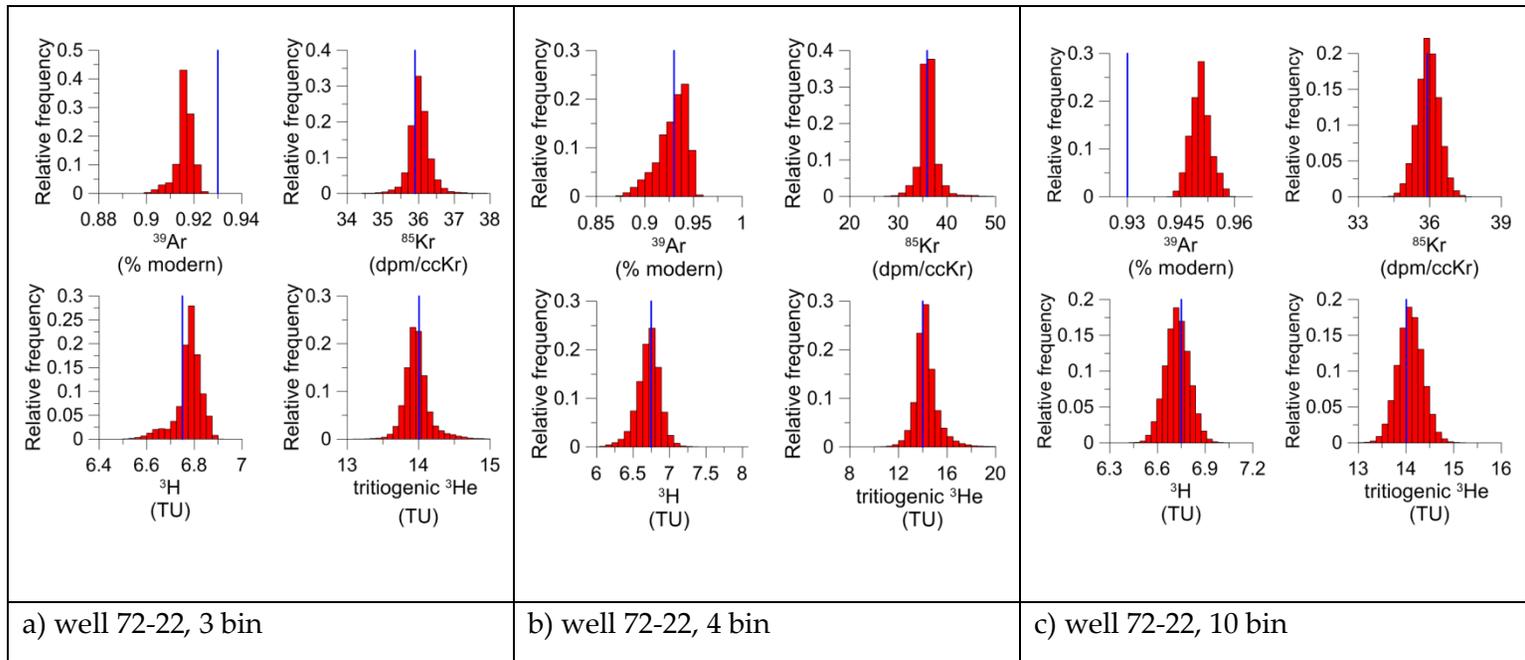


Figure 5: Observed concentrations (blue) and posterior distribution of concentrations (red) of the four tracers used in the Holten site analysis based on a 3 bin (a), 4 bin (b) and 10 bin (c) age distribution, for well 72-22 as an example.

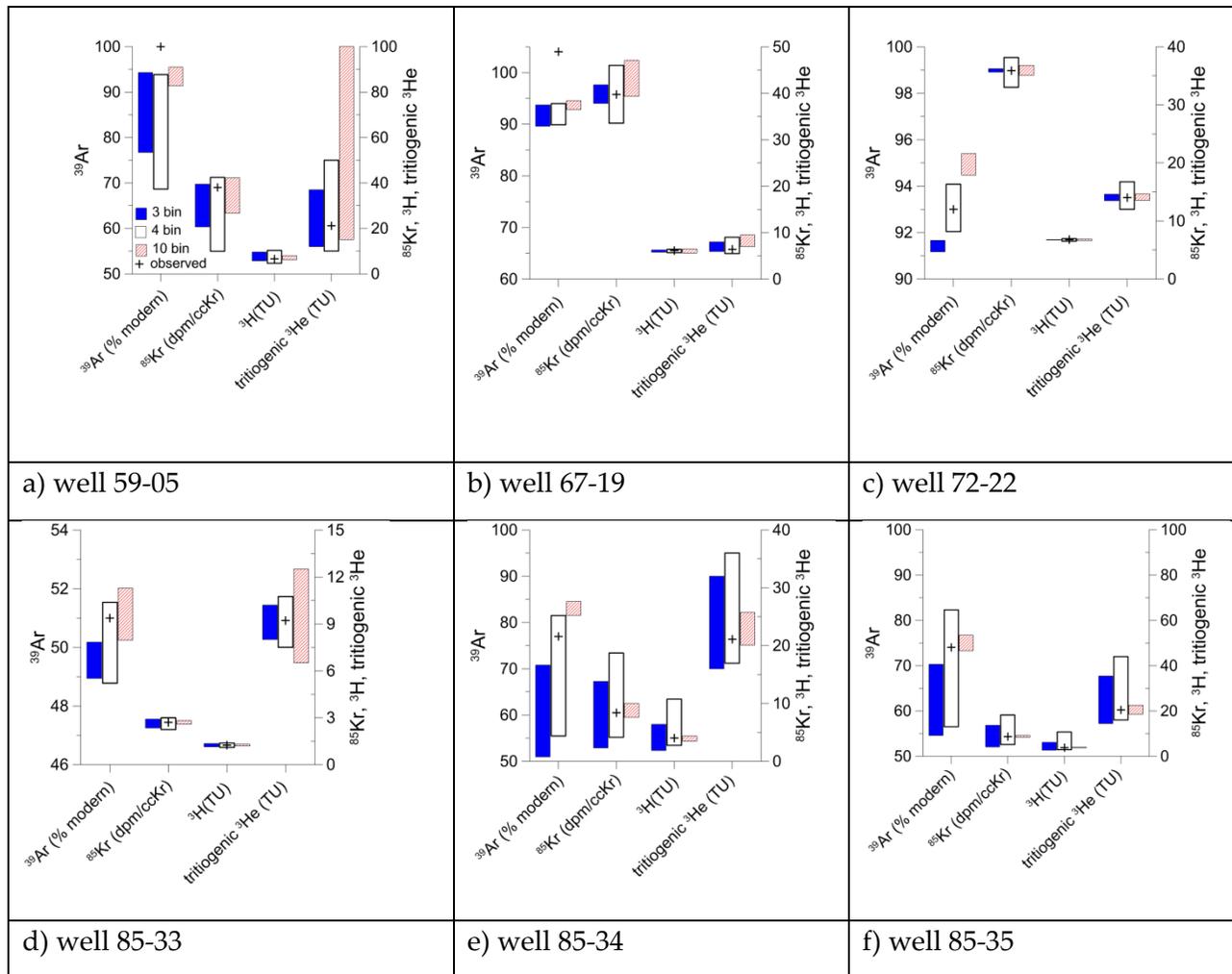


Figure 6: The 95% Credible intervals (CI) for the posterior concentrations of tracers using 3, 4 and 10 bin histogram models for six wells at the Holten site.

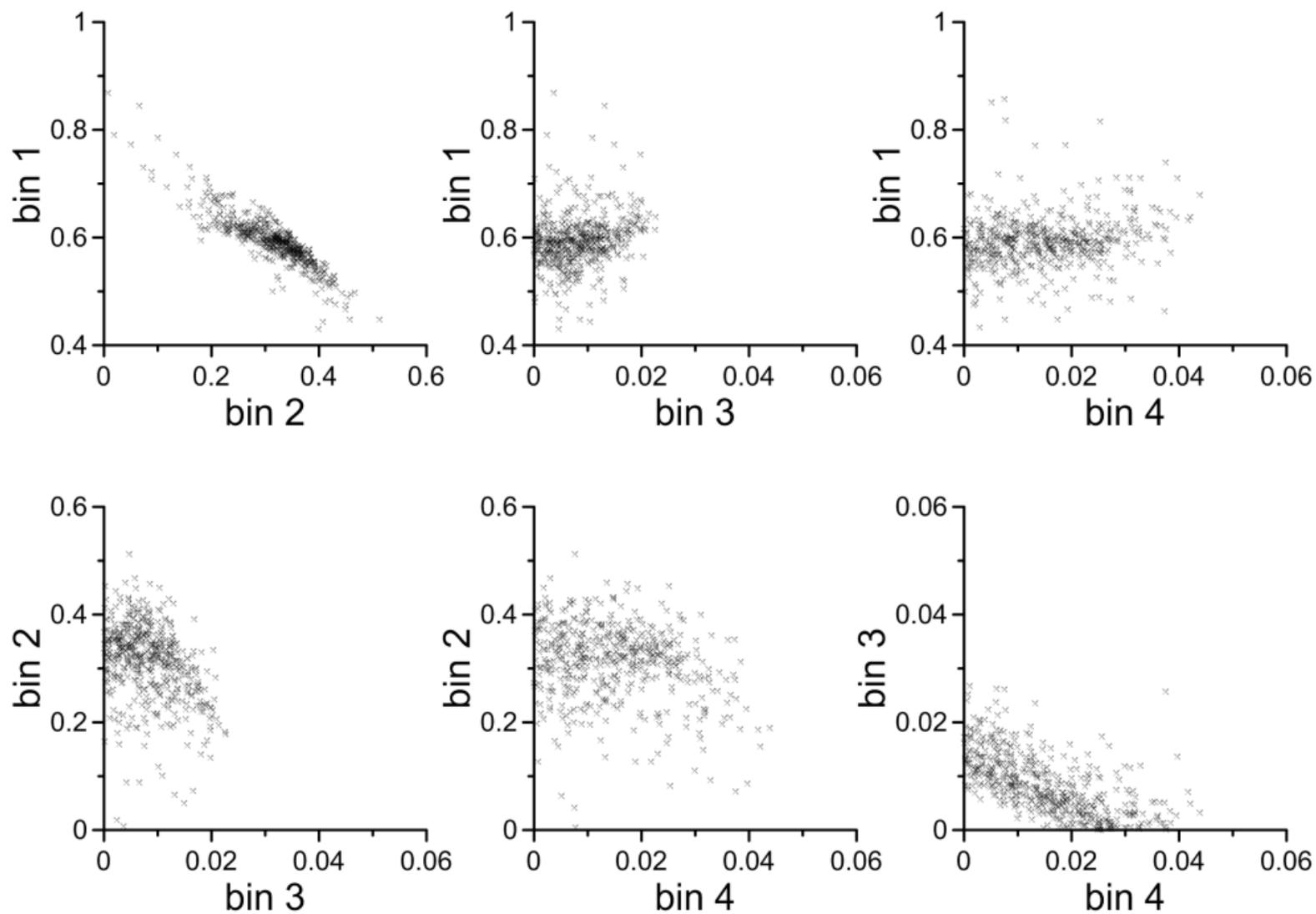


Figure 7: Scatter plots showing the correlations between the inferred fractions of water in each age bin for the 4 bin histogram model applied to well 72-22 at the Holten Site.

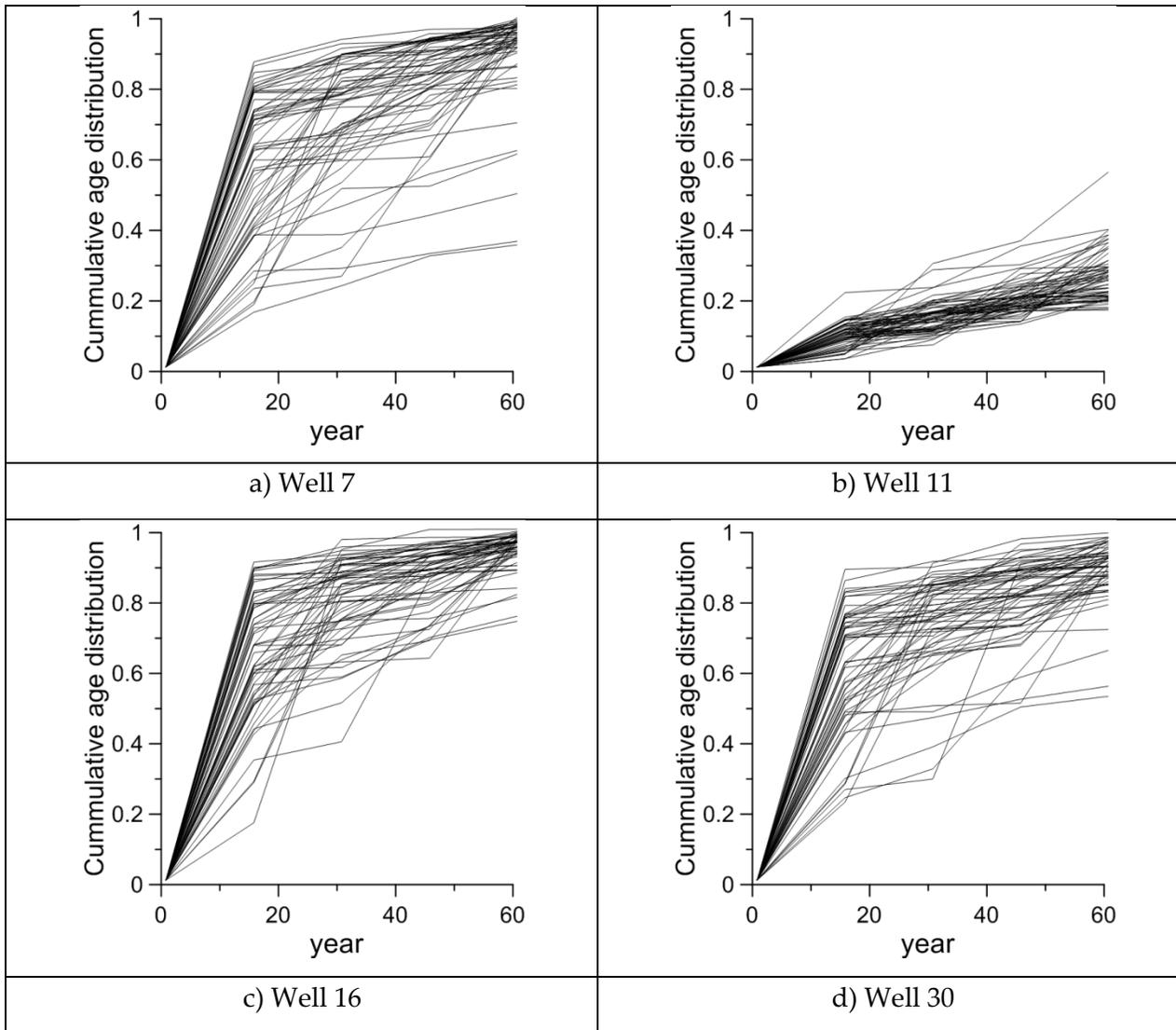


Figure 8: Samples of inferred 4 bin age distributions for a) wells 7, b) well 11, c) well 16 and d) well 30 at the La Selva site

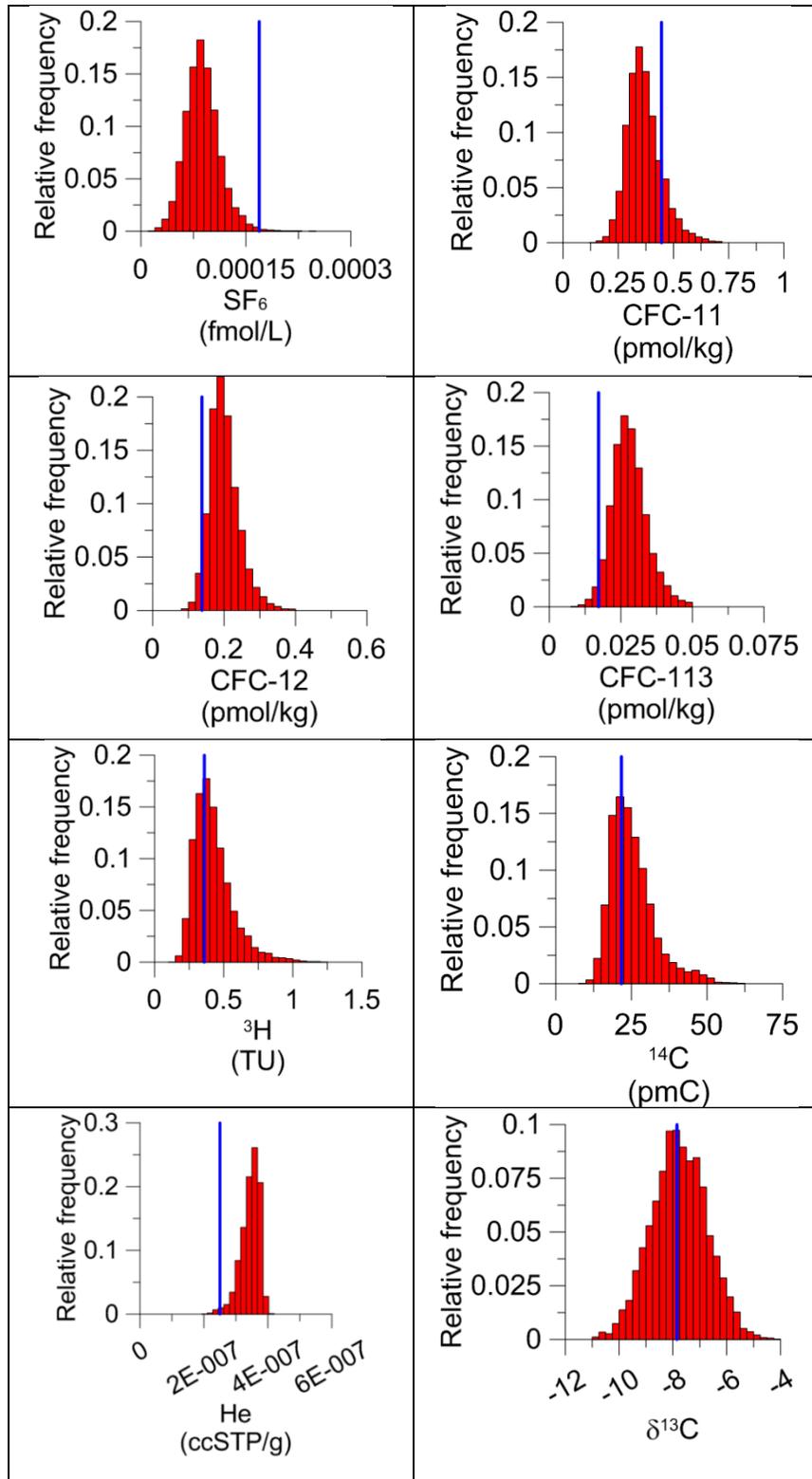


Figure 9: Observed concentration (blue) and posterior distribution of concentrations (red) of the eight tracers used for well 11 based on a 4 bin age distribution.

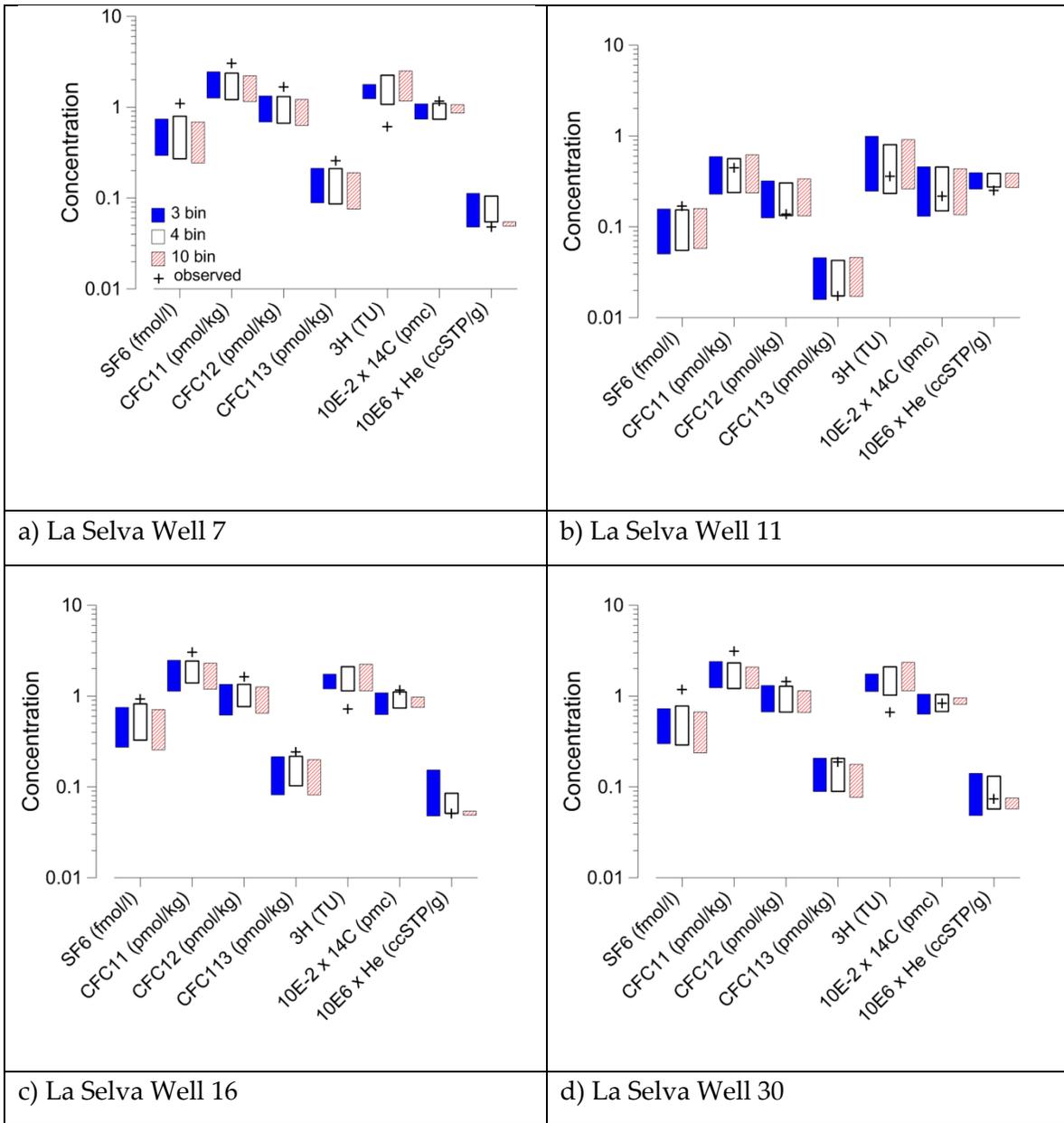


Figure 10: Observed and posterior 95% Credibility Interval (CI) brackets of tracer concentrations for four wells in the La Selva site inferred based on 3, 4 and 10 bin histogram models.