Measurement network design including traveltime determinations to minimize model prediction uncertainty

Gijs M. C. M. Janssen,[1,2] Johan R. Valstar,[2] and Sjoerd E. A. T. M. van der Zee[3]

Received 24 August 2006; revised 18 July 2007; accepted 24 September 2007; published 5 February 2008.

[1] Traveltime determinations have found increasing application in the characterization of groundwater systems. No algorithms are available, however, to optimally design sampling strategies including this information type. We propose a first-order methodology to include groundwater age or tracer arrival time determinations in measurement network design and apply the methodology in an illustrative example in which the network design is directed at contaminant breakthrough uncertainty minimization. We calculate linearized covariances between potential measurements and the goal variables of which we want to reduce the uncertainty: the groundwater age at the control plane and the breakthrough locations of the contaminant. We assume the traveltime to be lognormally distributed and therefore logtransform the age determinations in compliance with the adopted Bayesian framework. Accordingly, we derive expressions for the linearized covariances between the transformed age determinations and the parameters and states. In our synthetic numerical example, the derived expressions are shown to provide good first-order predictions of the variance of the natural logarithm of groundwater age if the variance of the natural logarithm of the conductivity is less than 3.0. The calculated covariances can be used to predict the posterior breakthrough variance belonging to a candidate network before samples are taken. A Genetic Algorithm is used to efficiently search, among all candidate networks, for a near-optimal one. We show that, in our numerical example, an age estimation network outperforms (in terms of breakthrough uncertainty reduction) equally sized head measurement networks and conductivity measurement networks even if the age estimations are highly uncertain.

Citation: Janssen, G. M. C. M., J. R. Valstar, and S. E. A. T. M. van der Zee (2008), Measurement network design including traveltime determinations to minimize model prediction uncertainty, *Water Resour. Res.*, *44*, W02405, doi:10.1029/2006WR005462.

1. Introduction

[2] Measurement network design or data worth analysis, whether done manually or using more or less sophisticated mathematical guidelines, is an inherent part of any soil and groundwater investigation, as it naturally evolves from the measurement campaign's purpose: obtaining the necessary information given the limited resources. Accordingly, this field of research receives constant attention in the literature. Nowadays, the literature offers a wide range of design strategies for a large number of applications, which roughly can be divided into five categories: 1) maximizing the likelihood of plume detection [*Massman and Freeze*, 1987a, 1987b; *Meyer and Brill*, 1988; *Meyer et al.*, 1994; *Storck et al.*, 1997]; 2) minimizing the uncertainty in groundwater quality [*Loaiciga*, 1989; *Herrera et al.*, 2000; *Nunes et al.*, 2004; *Wu et al.*, 2005]; 3) model calibration for minimizing the uncertainty in model predictions [*McKinney and Loucks*, 1992; *Wagner*, 1995; *Tiedeman et al.*, 2003, 2004]; 4) optimizing groundwater and remediation management [*James and Gorelick*, 1994; *Wagner*, 1999; *Feyen and Gorelick*, 2005]; 5) model calibration for minimizing model parameter uncertainty [*Hughes and Lettenmaier*, 1981; *Bogardi et al.*, 1985; *Knopman and Voss*, 1989; *Knopman et al.*, 1991; *Sumner et al.*, 1997; *Pardo-Iguzquiza*, 1998; *Chang et al.*, 2005]. The last application hardly ever is a purpose on its own but is usually conducted in order to improve model prediction reliability, which in its own turn can potentially lead to the design of groundwater and remediation management strategies that are more cost-effective [*James and Gorelick*, 1994; *Wagner*, 1999; *Feyen and Gorelick*, 2005].

[3] The design strategies reported in the literature generally seek the optimal placement and/or sampling times for measurements of heads, concentrations, parameters, or a combination of these three. In the past two decades, however, traveltime determinations have found increasing application as another type of information with which flow models can be constrained. By traveltime determinations we mean both groundwater age estimations and tracer arrival time measurements. Groundwater ages can be derived from concentrations of environmental tracers, such as $^{3}H/^{3}He$ [e.g., *Smethie et al.*, 1992], ^{85}Kr [e.g., *Solomon et al.*, 1992], chlorofluorocarbons [e.g., *Dunkle et al.*, 1993], or a com-

[1]Department of Soil Quality, Wageningen University, Wageningen, The Netherlands.
[2]TNO Built Environment and Geosciences – National Geological Survey, Utrecht, The Netherlands.
[3]Department of Soil Physics, Ecohydrology, and Groundwater Management, Wageningen University, Wageningen, The Netherlands.

0043-1397/08/2006WR005462

bination of them [e.g., *Ekwurzel et al.*, 1994; *Reilly et al.*, 1994; *Szabo et al.*, 1996]. By tracer arrival times we mean the advective (mean) arrival times that are obtained in field experiments using conservative tracers [e.g., *Anderman and Hill*, 2001]. When accurate, traveltime determinations can be more informative than head and conductivity measurements, as the sensitivity of heads to parameters is usually limited and the spatial correlation range between traveltime and conductivity is often larger than the correlation range of the conductivity itself [*Harvey and Gorelick*, 1995; *Sheets et al.*, 1998; *Stute and Schlosser*, 2000].

[4] Manual calibration of flow models using tracer derived groundwater ages has been performed by, for example, *Reilly et al.* [1994], *Sheets et al.* [1998] and *Izbicki et al.* [2004]. Systematic, mathematical approaches to parameter inference from age data or data on tracer arrival time are given by *Harvey and Gorelick* [1995], *Portniaguine and Solomon* [1998], *Woodbury and Rubin* [2000], *Cirpka and Kitanidis* [2001], and *Feyen et al.* [2003]. However, traveltime determinations have never been incorporated in the design of optimal measurement strategies. The purpose of the present study is therefore to propose an algorithm that optimally configures measurement networks including age or tracer arrival time determinations.

[5] We will focus our network design on minimizing contaminant breakthrough prediction uncertainty, thus making a contribution to the 3rd network design application category mentioned earlier. Aiming a measurement campaign at reducing prediction uncertainty makes sense, as "it is rarely feasible to improve the representation of all aspects of a simulated system. Thus it is of interest to identify the particular attributes of a flow system that are most important to the relevant predictions, and to focus field characterization on these attributes" [*Tiedeman et al.*, 2003]. Yet, this category has received little attention in the literature. If we narrow down our search to those studies that take correlations between measurements into account, then the most important contributions are given by *Wagner* [1995] and *Tiedeman et al.* [2003, 2004]. These authors all used the first-order linear statistical inference method [*Dettinger and Wilson*, 1981] to infer the prediction covariance matrix from sensitivity matrices and the posterior parameter covariance matrix. They recalculated the posterior parameter covariance matrix in every network evaluation that is performed, which in their approaches requires repeated inversion of a squared estimation Jacobian, i.e., the matrix of state measurement sensitivities. This can become computationally demanding if the number of unknown parameters is large, for example in case of a highly discretized hydraulic conductivity field. Here, we will adopt an adjusted version of the representer-based inverse method [*Valstar*, 2001; *Valstar et al.*, 2004], a Bayesian algorithm which can efficiently calculate the linearized covariances between measurements and predictions, and between the measurements themselves. On the basis of these covariances, the posterior covariances of the states and predictions can be calculated for every possible measurement set, without having to recalculate the posterior parameter covariance matrix.

[6] The purpose of the proposed method is to provide insight in where in the flow field traveltime information would be most valuable for the reduction of prediction uncertainty. A presupposition of the proposed method is that the traveltimes can be derived with a reasonable and quantifiable reliability from field measurements. As will be discussed in section 5, for groundwater age this is a very challenging issue in itself, but how to do that is out of the scope of this paper.

[7] In the next section, we will discuss the concepts and the mathematical development of our first-order design method. The method will be illustrated with a numerical, synthetic example, the details and results of which are given in section 3 and 4, respectively. Conclusions and a discussion on some aspects of the method follow in section 5.

2. Theory

[8] As we feel that the method is better explained by focusing it directly on a tangible illustrative example, first a hypothetical problem statement is introduced. Subsequently, an outline of the first-order design method is given, followed by a presentation on how we arrive at the necessary prior and posterior covariances in the applied Bayesian framework.

2.1. Problem Statement

[9] The illustrative example with which the design method will be explained is schematized in Figure 1. It represents a two-dimensional cross-section of a confining layer (y_2-y_3) that protects Aquifer 2 (y_3-y_4) from a conservative contaminant released out of a source zone (x_7-x_8) at the top of Aquifer 1 (y_1-y_2). The control plane at which breakthrough is evaluated is located at y_3 along the bottom of the area of interest (x_6-x_9, y_2-y_3).

[10] All parameters and dimensions in the example problem are assumed constant and known, except for the hydraulic conductivity (K) in the shaded center part (x_5-x_{10}, y_2-y_3) of the confining layer, which is assumed to be Gaussian distributed with known mean and variogram functions. A full description of the system's quantitative properties will follow in section 3 where the system visualized in Figure 1 is used in the example calculations.

[11] A steady state head distribution is obtained by assigning recharge on top of Aquifer 1 and by imposing constant head boundaries in Aquifer 2. Except for the upper boundary and the left and right boundaries of Aquifer 2, all boundaries are closed. For illustration, Figure 1 also shows the contaminant plume that spreads from the contaminant source zone when the confining layer is modeled as a homogeneous medium with K set to its geometric mean (see section 3) and the boundary conditions are imposed as in the computational examples (section 3).

[12] The question underlying the theoretical development given in the remainder of section 2 is how to optimally choose between age, head and conductivity measurements and how to optimally distribute them over the potential sampling locations, indicated by the stars in Figure 1.

2.2. Outline of the First-Order Design Method

[13] If the control plane over which the breakthrough is predicted receives a significant inflow that does not originate from the contaminant source zone, as is the case in Figure 1, the contaminant breakthrough time probability

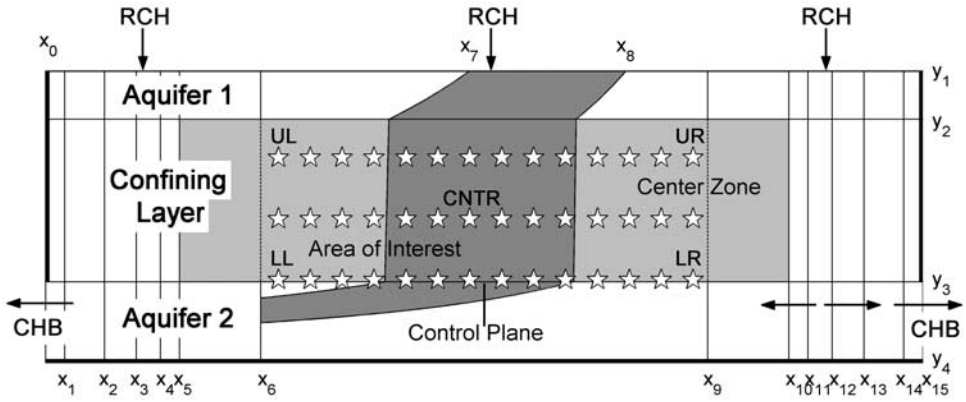

Figure 1. Schematic overview of the numerical grid used in the calculations, showing the prior contaminant flow (in dark grey). RCH = recharge. CHB = constant head boundary. $x_0 = 0$ m, $x_1 = 2.0$ m, $x_2 = 2502.0$ m, $x_3 = 2752.0$ m, $x_4 = 2777.0$ m, $x_5 = 2779.5$ m, $x_6 = 2879.5$ m, $x_7 = 3009.5$ m, $x_8 = 3109.5$ m, $x_9 = 3159.5$ m, $x_{10} = 3259.5$ m, $x_{11} = 3262.0$ m, $x_{12} = 3287.0$ m, $x_{13} = 3537.0$ m, $x_{14} = 6037.0$ m, $x_{15} = 6039.0$ m. $y_1 = 0$ m, $y_2 = -2.0$ m, $y_3 = -22.0$ m, $y_4 = -32.0$ m. Bold lines indicate no flow boundaries. Potential measurement locations labeled UR, UL, LR, LL and CNTR are the locations where the head variances are evaluated in Figure 4. Arrows indicate flow directions.

distribution $p(\tau_{BT})$ is a function of the traveltime probability of the entire inflow and the contaminant breakthrough location probability:

$$p(\tau_{BT}) = p(t_0 + \tau) = \int_{x^{CP}=A}^{x^{CP}=B} p(\tau; x^{CP}) p(x^{CP} \in \chi) dx^{CP} \quad (1)$$

or in discretized form:

$$p(\tau_{BT}) \approx \sum_{x^{CP}=A}^{x^{CP}=B} \frac{1}{2} \left[p(\tau; x^{CP}) p(x^{CP} \in \chi) + p(\tau; x^{CP} + \Delta x^{CP}) p(x^{CP} + \Delta x^{CP} \in \chi) \right] \Delta x^{CP} \quad (2)$$

[14] In equations (1)–(2), τ_{BT} is the contaminant breakthrough time evaluated at the control plane, t_0 is the starting time of the contamination, τ is arrival time (or groundwater age at the control plane), $p(\tau; x^{CP})$ is the marginal arrival time probability evaluated at x^{CP}, x^{CP} is a location at the control plane, $p(x^{CP} \in \chi)$ is the probability that x^{CP} belongs to the contaminant breakthrough zone χ (χ is the collection of all breakthrough locations χ), and A and B are the spatial x limits of the control plane (x_6 and x_9 in Figure 1). Δx^{CP} is the discretization interval along the control plane.

[15] Assuming that τ and χ are Gaussian distributed, we have

$$p(\tau; x^{CP}) \approx N\left(\tau; \sigma^2_{\tau(x^{CP})}, \mu_{\tau(x^{CP})}\right) \quad (3)$$

$$p(x^{CP} \in \chi) \approx \sum_{x^{SZ}=C}^{x^{SZ}=D} \frac{1}{2} \left[N\left(x^{CP}; \sigma^2_{\chi(x^{sz})}, \mu_{\chi(x^{sz})}\right) + N\left(x^{CP}; \sigma^2_{\chi(x^{sz}+\Delta x^{sz})}, \mu_{\chi(x^{sz}+\Delta x^{sz})}\right) \right] \Delta x^{SZ}, \quad (4)$$

where $N(p_1; p_2, p_3)$ represents the probability density of p_1 according to the normal distribution, parameterized with variance p_2 and mean p_3, $\tau(x^{CP})$ is the arrival time of a particle arriving at the control plane at location x^{CP}, $\sigma^2_{\tau(x^{CP})}$ is the arrival time (or age) variance at x^{CP}, and $\mu_{\tau(x^{CP})}$ is the mean arrival time (or age) at x^{CP}. $\chi(x^{SZ})$ is the breakthrough location of a particle that originated from x^{SZ}, where x^{SZ} is a location in the contaminant source zone. $\sigma^2_{\chi(x^{sz})}$ and $\mu_{\chi(x^{sz})}$ are the breakthrough location variance and mean (along the x axis) of a particle originating from x^{SZ}, respectively, Δx^{SZ} is the discretization interval along the contaminant source zone, and C and D are the spatial x limits of the contaminant source zone (x_7 and x_8 in Figure 1).

[16] Finally, the variance of the breakthrough time $\sigma^2_{\tau_{BT}}$ can be calculated as:

$$\sigma^2_{\tau_{BT}} = \frac{\int_{\tau_{BT}=0}^{\infty} (\tau_{BT} - E(\tau_{BT}))^2 p(\tau_{BT}) dt}{\int_{\tau_{BT}=0}^{\infty} p(\tau_{BT}) dt}$$

$$\text{with} \quad E(\tau_{BT}) = \frac{\int_{\tau_{BT}=0}^{\infty} \tau_{BT} p(\tau_{BT}) dt}{\int_{\tau_{BT}=0}^{\infty} p(\tau_{BT}) dt}, \quad (5)$$

where $E(\tau_{BT})$ is the expected contaminant breakthrough time. Equation (5) is the objective function that is to be minimized in the search for an optimal network design.

[17] From equations (1)–(5) it follows that to evaluate the performance of different measurement network designs, the influences of the observable variables at their potential sampling locations on the prediction of the breakthrough time $\tau(x^{CP})$ and the breakthrough locations $\chi(x^{SZ})$ have to be known. This makes $\tau(x^{CP})$ and $\chi(x^{SZ})$ our goal variables and they will be called as such throughout the remainder of this paper.

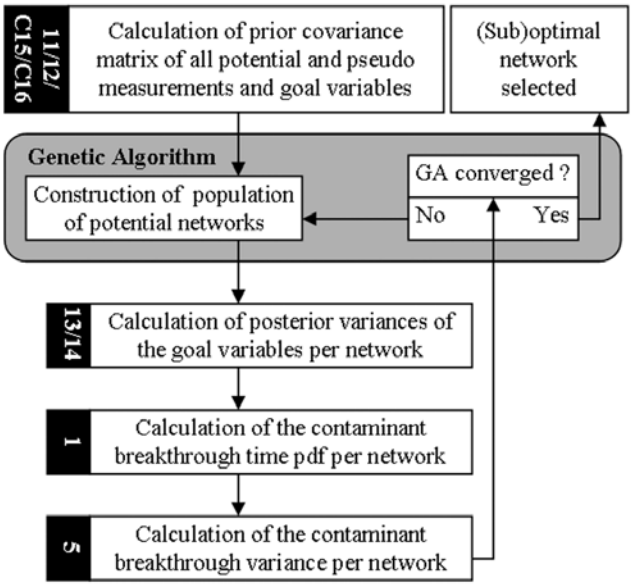

Figure 2. Flowchart of the proposed network design algorithm. The numbers in the black boxes refer to the equations using which the actions described in the text boxes attached to them are performed.

[18] The recently proposed representer-based inverse method [*Valstar et al.*, 2004] provides an efficient way to calculate these influences. The representers calculated in this method are equivalent to the linearized covariances between the observable variable at the potential sampling location and the variables for which the representers are defined. As such, they provide a first-order estimation of the prior covariances of the observable variables and the goal variables $\tau(x^{CP})$ and $\chi(x^{SZ})$. These covariances can subsequently be used to approximate the posterior covariances of the goal variables, and therefore the posterior breakthrough time uncertainty, before the measurements are actually taken.

[19] Thus the presented algorithm can efficiently calculate, at first order, the expected posterior breakthrough time variance $\sigma^2_{\tau_{BT}}$ (equation (5)) of the contaminant for every candidate measurement network design. A Genetic Algorithm was used to efficiently search for a near-optimal design that minimizes $\sigma^2_{\tau_{BT}}$. For clarification, a flowchart of the design method is given in Figure 2.

2.3. Bayesian Framework: Derivation of Ln(Breakthrough Time) and Breakthrough Location Representers

[20] The covariances between the breakthrough location $\chi(x^{SZ})$ and any observation are given at first order by breakthrough location representers. These breakthrough representers have been derived before (N. P. A. J. van de Wiel, manuscript in preparation), though more generally as particle location representers.

[21] The covariances between the breakthrough times $\tau(x^{CP})$ and any observation, as well as the covariance between groundwater ages and any other measurement, are given at first order by traveltime representers, which also have been derived before (N. P. A. J. van de Wiel, manuscript in preparation). In that work, the derivation is carried out for Gaussian distributed traveltimes. However, unless the traveltime is evaluated in the large displacement regime (i.e., after many correlation distances), where by virtue of the Central Limit Theorem [e.g., *Zhang*, 2002, p. 61] the Gaussian assumption might be valid, the traveltime probability density function will, in general, show a significant skewness. In the literature, therefore, skewed distributions are often adopted for the traveltime, such as the lognormal [e.g., *Simmons*, 1982; *Cvetkovic et al.*, 1992; *Kovar et al.*, 2005] or the inverse-Gaussian [e.g., *Cirpka and Kitanidis*, 2000] distribution. Here, because of the analytical simplicity of logarithmic transformations, we will treat the traveltime as a lognormally distributed variable, transform traveltime estimations to the Gaussian distribution by taking their natural logarithm to make them comply with the Bayesian framework of the representer-based inverse algorithm, and derive Ln(traveltime) representers accordingly.

[22] Consider the flow, the particle traveltime, and the particle location equations:

$$\phi_1 = A_{fg}(\alpha)h_g - q_f = 0 \tag{6}$$

$$\phi^s_{2_i} = 0 = t^s_i - t^{s-1}_i - \Delta t^s_i\left(\alpha, h, x^s_i, x^{s-1}_i, y^s_i, y^{s-1}_i\right) \tag{7}$$

$$\phi^s_{3_i} = 0 = \begin{cases} x^s_i - x^{s-1}_i - \Delta x^s_i\left(\alpha, h, y^s_i, y^{s-1}_i\right) & \text{for } x^s_i \neq x_{cell\ boundary} \\ x^s_i - x_{cell\ boundary} & \text{for } x^s_i = x_{cell\ boundary} \end{cases} \tag{8}$$

$$\phi^s_{4_i} = 0 = \begin{cases} y^s_i - y^{s-1}_i - \Delta y^s_i\left(\alpha, h, x^s_i, x^{s-1}_i\right) & \text{for } y^s_i \neq y_{cell\ boundary} \\ y^s_i - y_{cell\ boundary} & \text{for } y^s_i = y_{cell\ boundary} \end{cases} \tag{9}$$

In these and following equations, subscripted indices show vector and matrix dimensions and how, where applicable, matrix-vector multiplications should be performed. Products of terms containing the same index twice should be summed over that index. The ranges of all indices used are given in Table 1. Furthermore, in equations (6)–(9), h is the vector of nodal heads, q is the vector of driving forces, and $A(\alpha)$ is the system matrix depending on the unknown parameters α (=Ln(K)), t_i is the traveltime of particle i, Δt^s_i is the duration of travel step s, x_i and y_i are the locations along respectively the x- and y-axis of particle i, and Δx^s_i and Δy^s_i are the distances traveled along respectively the x- and y-axis by particle i during travel step s. For this study we applied a numerical particle tracking scheme in which travel step sizes are limited by reaching cell boundaries rather than by reaching set time step sizes. In this case, Δt^s_i and Δy^s_i depend on x^s_i and x^{s-1}_i if travel step s is limited by reaching an x-boundary of a cell. Furthermore, Δt^s_i and Δx^s_i depend on y^s_i and y^{s-1}_i if travel step s is limited by reaching a y-boundary of a cell.

[23] The representer-based inverse method searches for the maximum a posteriori estimates of the parameters, given the measurements. In this paper, the measurements that are taken into account are head (h), conductivity (K), groundwater age (t), and particle breakthrough

Table 1. Ranges of Indicators and Indices

Index	Ranges From 1 to the Number of:
i	Tracked particles, which equals the number of potential age determinations plus the number of pseudo breakthrough time and pseudo breakthrough location measurements
f, g	head state variables
k, l	uncertain parameters
n, p, q	measurements

location (χ) measurements. As will be explained in section 2.4, χ measurements will be incorporated in the algorithm as pseudo observations only, to enable the calculation of cross covariances between the goal variable $\chi(x^{SZ})$ and the other measurement types.

[24] For a steady state system, if all parameters α and measurement errors v are assumed to be multivariate Gaussian distributed with known covariances and they are not cross-correlated, the maximum a posteriori estimates of the parameters can be found by minimizing the following objective function:

$$J = [\alpha - \bar{\alpha}]^T \left[P_\alpha^{-1}\right][\alpha - \bar{\alpha}] + [z - M(h, \alpha, \mathrm{Ln}(t), \chi)]^T \cdot \left[P_v^{-1}\right][z - M(h, \alpha, \mathrm{Ln}(t), \chi)], \quad (10)$$

where J is the objective function value, z is the vector of measurement values, $M(\ )$ is a linear function that interpolates the vector of model predictions at the nodal points to the locations of the measurements, $\overline{\alpha}$ is the prior mean of the parameters, P_v is the covariance matrix of the measurement errors ν, and P_α is the prior covariance matrix of the parameters. By the Lagrange Method, we expand the objective function as:

$$J^* = J + 2\lambda_h[\phi_1 - q_f] + \sum_i^{N_i}\sum_s^{N_{s_i}}\left[2\lambda_{t_i}^s\left(\phi_{2_i}^s\right)\right] + \sum_i^{N_i}\sum_s^{N_{s_i}}\left[2\lambda_{x_i}^s\left(\phi_{3_i}^s\right)\right] + \sum_i^{N_i}\sum_s^{N_{s_i}}\left[2\lambda_{y_i}^s\left(\phi_{4_i}^s\right)\right], \quad (11)$$

where N_i is the number of tracked particles, N_{s_i} is the total number of steps traveled by particle i, λ_h is the head adjoint vector, λ_t is the traveltime adjoint vector, and λ_x and λ_y are the x and y location adjoint vector, respectively. In the minimum of objective function J^*, the variation of J^* is zero for any variation of α, h, x, y, t, λ_h, λ_x, λ_y and λ_t. Forcing this condition on equation (11) yields a system of 9 coupled Euler-Lagrange equations (see Appendix A). The solution of the system of Euler-Lagrange equations gives the set of parameters at which the derivatives of the extended objective function (equation (11)) with respect to the parameters are zero. Assuming that equation (11) has only 1 global minimum and no local minima, this is the set of parameters that optimally obeys the observations given the prior information.

[25] The nonlinear system of Euler-Lagrange equations is solved by first linearizing its individual equations. Then, the unknowns are expanded into a finite number of representer terms, each of which represents the influence of one particular measurement unknown, see Appendix B. Subsequently, the representer expressions and the unknowns are updated iteratively. The representer expansions enable the decoupling of the Euler-Lagrange equations into a set of expressions for the representers and their coefficients that can be sequentially solved (see Appendix C). In our first-order network design approach, the inverse algorithm is terminated in the first iteration after the Ln(τ) representers and the breakthrough location representers are obtained. Updating the unknowns in preparation for the second iteration and further requires measurement information, which is assumed to be unavailable at this stage of the network design.

[26] From Appendices C9 and C7 it follows that after the first iteration the Ln(τ) representer (Θ^*) for a particle i is given by:

$$\Theta_i^{*N_{s_i}} = \frac{1}{t_i^{N_{s_i}}}\left[\sum_{s=N_{s_i}}^{1}\frac{\partial\Delta t_i^s}{\partial\alpha_l}\Psi_l + \sum_{s=N_{s_i}}^{1}\frac{\partial\Delta t_i^s}{\partial h_g}\Xi_g + \sum_{s=N_{s_i}}^{1}\frac{\partial\Delta t_i^s}{\partial x_i^s}\mathrm{X}_i^s + \sum_{s=N_{s_i}}^{1}\frac{\partial\Delta t_i^s}{\partial x_i^{s+1}}\mathrm{X}_i^{s+1} + \sum_{s=N_{s_i}}^{1}\frac{\partial\Delta t_i^s}{\partial y_i^s}\mathrm{Y}_i^s + \sum_{s=N_{s_i}}^{1}\frac{\partial\Delta t_i^s}{\partial y_i^{s+1}}\mathrm{Y}_i^{s+1}\right], \quad (12)$$

with terminal conditions $\mathrm{X}_i^{N_s+1} = 0$ and $\mathrm{Y}_i^{N_s+1} = 0$, and the breakthrough location representer is given by:

$$\mathrm{X}_i^{N_{s_i}} = \sum_{s=1}^{N_{s_i}}\frac{\partial\Delta x_i^s}{\partial\alpha_l}\Psi_l + \sum_{s=1}^{N_{s_i}}\frac{\partial\Delta x_i^s}{\partial h_g}\Xi_g + \sum_{s=1}^{N_{s_i}}\frac{\partial\Delta x_i^s}{\partial y_i^s}\mathrm{Y}_i^s + \sum_{s=1}^{N_{s_i}}\frac{\partial\Delta x_i^s}{\partial y_i^{s-1}}\mathrm{Y}_i^{s-1}, \quad (13)$$

with boundary conditions $\mathrm{Y}_i^0 = 0$. In equations (12) and (13), Ψ is the parameter representer (see Appendices B and C5), Ξ is the head representer (see Appendix B and C6), and $t_i^{N_{si}}$ is the breakthrough time of particle i. Equation (12) is in backward notation, because particles defined to calculate breakthrough time representers are tracked backward from the locations at the control plane where breakthrough time uncertainty information is desired.

2.4. Computation of Posterior Breakthrough Time Probability

[27] The first-order posterior variances of the goal variables can now be calculated as:

$$P_{\mathrm{Ln}(\tau_i)}^{posterior} = P_{\mathrm{Ln}(\tau_i)}^{prior} - \Theta_{ip}^{*N_{s_i}}\left((M(\Xi, \Psi, \Theta^*, \mathrm{X}) + P_v)^{-1}\right)_{pq}\Theta_{qi}^{*N_{s_i}} \quad (14)$$

$$P_{\chi_i}^{posterior} = P_{\chi_i}^{prior} - \mathrm{X}_{ip}^{N_{s_i}}\left((M(\Xi, \Psi, \Theta^*, \mathrm{X}) + P_v)^{-1}\right)_{pq}\mathrm{X}_{qi}^{N_{s_i}}, \quad (15)$$

where $M(\Xi, \Psi, \Theta^*, \mathrm{X})_{pq}$ is a $p \times q$ representer matrix that contains the prior cross covariances between all observations, $P_{\mathrm{Ln}(\tau_i)}^{prior}$ is calculated by defining a pseudo measurement of τ_i and calculating the Ln(τ) representer for this measurement, and $P_{\chi_i}^{prior}$ is calculated by defining a pseudo measurement for χ_i and calculating the breakthrough location representer for this measurement.

[28] $P^{posterior}_{\mathrm{Ln}(\tau_i)}$ and $P^{posterior}_{\chi_i}$ can subsequently be filled in for $\sigma^2_{\tau(x^{CP})}$ and $\sigma^2_{\chi(x^{sz})}$ in equations (3) and (4), respectively (note that τ in equations (1)–(3) is replaced by its natural logarithm). $\mu_{\tau(x^{CP})}$ and $\mu_{\chi(x^{SZ})}$ in equations (3) and (4) are approximated by their first guess estimates, $\mathrm{Ln}(\tau_F(x^{CP}))$ and $\chi_F(x^{SZ})$ (see also Appendix B), obtained by running a simulation with $\alpha = \overline{\alpha}$. Now all necessary information is available to compute $p(\tau_{BT})$ according to equation (2) and subsequently the variance of the breakthrough time $\sigma^2_{\tau_{BT}}$ according to equation (5).

[29] Note that although many network evaluations are performed in the Genetic Algorithm, this repeated evaluation requires calculating the representer matrix $M(\Xi, \Psi, \Theta^*, X)$ (for all potential measurements and pseudo measurements) only once. With $M(\Xi, \Psi, \Theta^*, X)$ known, $P^{posterior}_{\mathrm{Ln}(\tau_i)}$ and $P^{posterior}_{\chi_i}$ can be calculated for every candidate network by selecting the relevant variances and covariances from this matrix and performing the necessary operations with them according to equations (14)–(15).

3. Numerical Application to a Synthetic Test Case

[30] We will now demonstrate the design methodology by numerically filling in the problem that was introduced in section 2.1 and Figure 1 and performing example calculations with the thus created test case.

[31] Additional to the discretization of the flow model as shown in Figure 1, the area between x_5 and x_{10} is discretized into 240 equally sized ($\Delta x = 2.0$ m) columns, and additionally the confining layer (y_2–y_3) is discretized into 40 equally sized rows ($\Delta y = 0.5$ m). The purpose of all areas outside the area of interest is solely to reduce the impact of boundary conditions on the flow in the area of interest.

[32] The porosity is 0.4 everywhere, and Aquifer 1 and Aquifer 2 have a hydraulic conductivity (K) of 3.0 and 0.6 m/d, respectively. The Gaussian distribution of Ln(K) (= Y) in the center part of the confining layer (9,600 cells) has a geometric mean $\mathrm{Ln}(K_G)$ of −3.0 Ln(m/d), a variance σ^2_Y of 2.0 unless stated otherwise, and an exponential variogram model with horizontal and vertical correlation ranges of 75 m and 25 m, respectively. Outside the shaded center part (x_0–x_5 and x_{10}–x_{13}) the confining layer is modeled as a homogeneous deposit with K equal to the K_G assigned to the center part. A steady state head distribution is obtained by assigning recharge (250 mm/a) to the top of all cells of Aquifer 1 and by imposing a constant head of 0.0 m and 15.0 m in the utmost left and utmost right cell of Aquifer 2, respectively. This results in a flow divide between x_{13} and x_{14}, as indicated by the arrows in Figure 1.

[33] The stars in the area of interest indicate 42 potential measurement locations. All measurements suffer from measurement errors, which are assumed to be Gaussian distributed. The measurement error variances are taken to be 0.0001 m^2 and 0.001 $(m/d)^2$ for h and K measurements, respectively. For t determinations, the estimation error standard deviation is taken as a percentage of the expected (untransformed) age (t_F) at the sampled location in the prior realization of the confining layer, in which all stochastic parameters are set at their prior means ($\alpha = \mathrm{Ln}(K) = \overline{\alpha} =$ −3.0 Ln(m/d)). This percentage varies between different examples.

[34] At the control plane, in every column one pseudo age determination was defined to enable the calculation of the covariances between the goal variable $\mathrm{Ln}(\tau(x^{CP}))$ and the measurements in a design. The sensitivity terms in equation (12) (the Ln(τ) representer equation) were computed by backward tracking of particles traveling from the pseudo measurement locations at the control plane to the inflow zone ($y = y_1$) and calculating the time step sensitivities of every travel step using the ADV2 package of MODFLOW 2000 [*Anderman and Hill*, 2001].

[35] Likewise, in the contaminant source zone, one pseudo location measurement was defined in every column, to enable the calculation of the covariances between the goal variable $\chi(x^{SZ})$ and the measurements in a design. The terms in the breakthrough location (χ) representer equation (13) were computed by forward tracking of particles traveling from the pseudo measurement location in the contaminant source zone to the control plane and evaluating the displacement sensitivities of every travel step, again using the ADV2 package of MODFLOW 2000.

[36] In the example calculations, the desired number of measurements in the resulting designs was always fixed, although during the optimization it was allowed to vary. This was achieved by adding an extra term to the objective function (equation (5)) in the Genetic Algorithm. This term consisted of a multiplication of the absolute value of the squared difference between the required number of measurements and the actual number of measurements in the design under consideration, and a multiplication coefficient. This multiplication coefficient was chosen differently in every design optimization, as it affects the convergence of the Genetic Algorithm.

4. Results

[37] In this section, we will first investigate the validity of the first-order Bayesian methodology by checking the normality of Ln(t), and by comparing linearized variance predictions as given by the representer method with Monte Carlo results for increasing parameter variability. The Monte Carlo simulations were performed using MODFLOW [*McDonald and Harbaugh*, 1984] in combination with the particle-tracking software MODPATH [*Pollock*, 1994]. The number of Monte Carlo runs was 10000 for every calculation.

[38] Subsequently, we will take a look at the correlations between the observable variables and the goal variables. Knowing and understanding these correlations is useful to explain the observation networks as designed by the first-order method. The details of the procedure followed for the calculation of the correlations will be given in section 4.2. In section 4.3, three examples of near-optimal network designs will be given and explained on the basis of the calculated correlations.

[39] Finally, in section 4.4, we will compare the performance of networks consisting solely of age determinations with the performance of networks consisting solely of head and solely of conductivity measurements.

4.1. Validity of the First-Order Design Method

[40] As our inverse method is defined in a Bayesian framework, it is important that all unknown parameters and dependent variables are (at least approximately) Gaussian distributed. It is well known that if Y is Gaussian

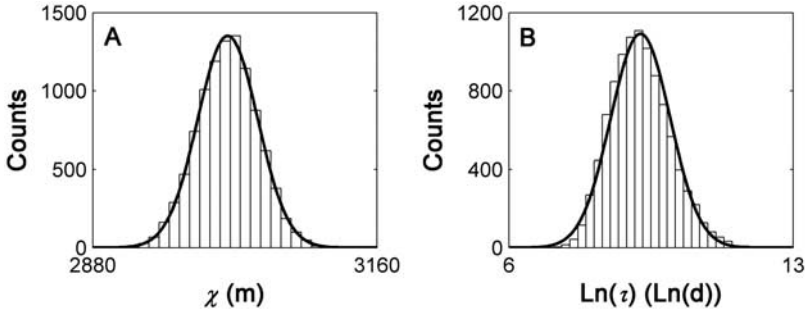

Figure 3. Histograms of the distribution of breakthrough locations (a) and breakthrough times (b) for the Monte Carlo simulation of a particle starting from the center of the contaminant source zone, compared with the theoretical normal probability density function using the Monte Carlo derived means and variances (black lines). $\sigma_Y^2 = 2.0$.

distributed, at first order so are the hydraulic head and the particle displacement [*Dagan*, 1989]. Particle traveltimes, however, can be significantly skewed if the traveled number of correlation scales is limited (see section 2.3). Fortunately, the natural logarithm of arrival time is approximately Gaussian if σ_Y^2 is not too large. This is shown in Figure 3, which gives the histogram of breakthrough times for a particle starting from the middle of the contaminant source zone. Also indicated in this figure is the Gaussian probability density function (pdf) parameterized with the mean and variance of breakthrough times calculated from the Monte Carlo results. The good correspondence of the Monte Carlo-derived Ln(τ) histogram with the Gaussian distribution will deteriorate as σ_Y^2 gets larger and the untransformed traveltimes become more skewed. However, for the range of σ_Y^2 values for which the linear theory applied here can be assumed valid (see below) the natural logarithm of the traveltime can be considered as a sufficiently Gaussian distributed variable.

[41] Figure 4 addresses the question regarding the applicability of the linear theory for larger σ_Y^2. Figure 4 compares the variances of the state variables as calculated by the representer approach with the variances of these variables computed with a Monte Carlo approach. Figure 4a shows the results for Ln(τ) of a particle originating from the center of the contaminant source zone. It should be noted here that the observed underestimation of the variances by the first-order method is caused by two distinct factors: increasing non-linearity of the flow equation and increasing non-normality of Ln(τ) as σ_Y^2 increases. Still, up to $\sigma_Y^2 = 3.0$ the difference between linearization and Monte Carlo results is less than 10%.

[42] Figure 4b compares Monte Carlo variances with linearized variances of the particle breakthrough location χ.

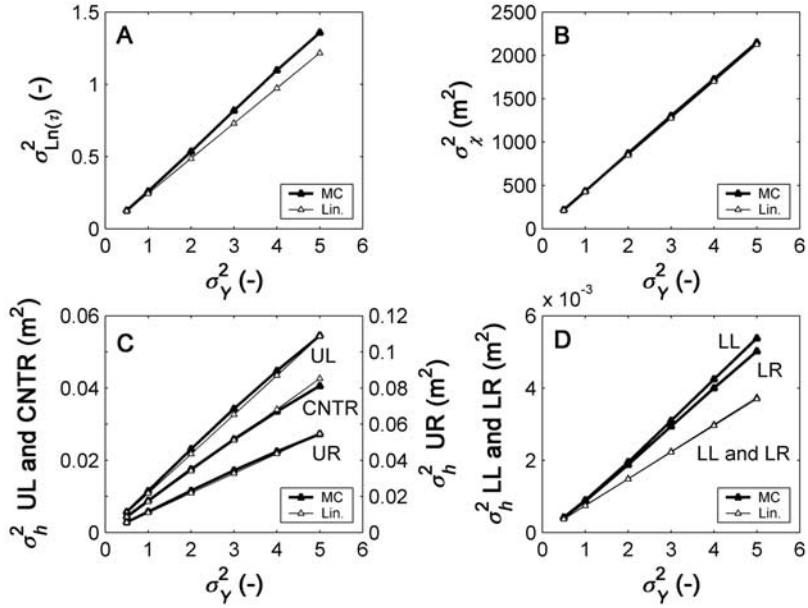

Figure 4. Linearized Ln(t) (a), χ (b) and h (c and d) variance predictions as given by the representer approach compared with variances obtained from a Monte Carlo series, as function of σ_Y^2. In Figures 4c and 4d, the labels UL, LL, UR, LR and CNTR correspond with the labels in Figure 1 and identify the locations in the in at which the variances were evaluated.

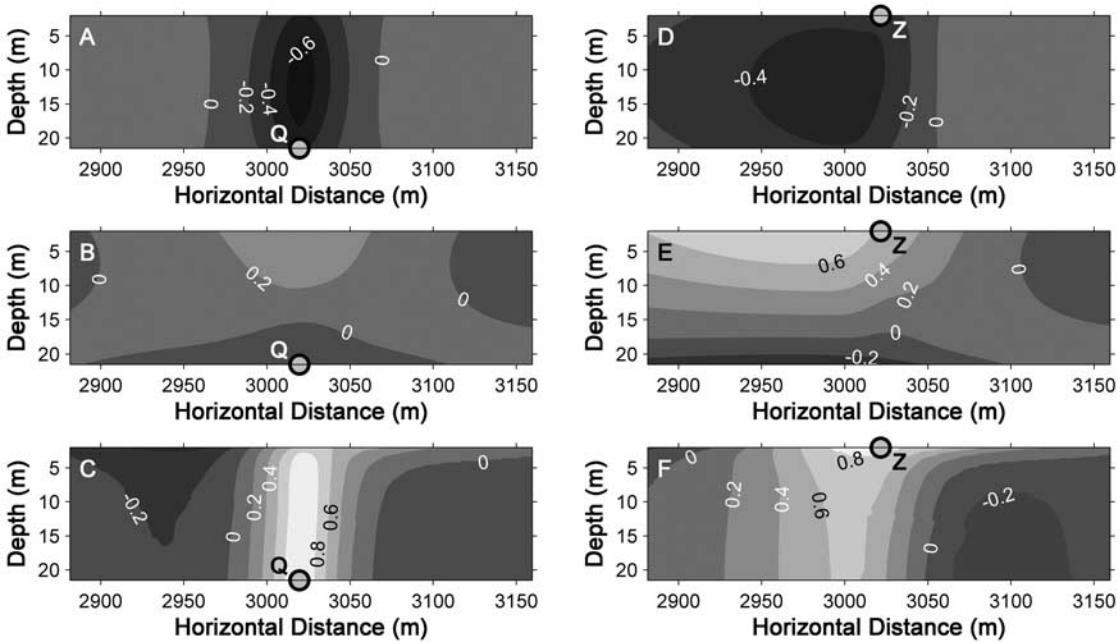

Figure 5. The correlation between the groundwater age at or the traveltime to Q and the log conductivity (a), the head (b) and the groundwater age (c) in every grid block of the center area. The correlation between the breakthrough location of the particle a priori expected to break through at Q (this particle enters the confining layer at Z) and the natural logarithm of the conductivity (d), the head (e) and the groundwater age (f) in every grid block of the center area.

The correspondence is excellent for the entire range investigated. The same holds for the head variance in the upper part of the domain (Figure 4c). Only for the head variance in the lower part, the correspondence between the first-order results and the Monte Carlo results deteriorates rather quickly as σ_Y^2 increases. It is conceivable that the larger influence of the constant head boundaries in this part of the domain increases the nonlinearity of the flow equation. Nevertheless, considering the small absolute value of the underestimation of the head variance (note the factor 10 difference between the y-axes of Figures 4c and 4d), the error made is not expected to greatly influence contaminant breakthrough time uncertainty predictions.

[43] Based on the result given in Figure 4, for our synthetic example the outcome of our first-order design strategy as outlined in section 2 is considered reliable up to a variance of Y of 3.0.

4.2. Correlations Between Observed Variables and Goal Variables

[44] Figure 5 shows plots of the spatial distribution of the correlations between the observable variables (h, Ln(K) and Ln(t)) and respectively the natural logarithm of the groundwater age at Q (Ln(t_Q), see Figures 5a–5c, in which Q is the center of the a priori expected contaminant breakthrough zone) and the breakthrough location χ_Q of the particle a priori expected to break through at Q (Figures 5d–5f). The particle that a priori is expected to break through at Q enters the confining layer at Z (see Figures 5d–5f).

[45] Figure 5a gives the correlation between the conductivity everywhere in the domain and the groundwater age at Q. These correlations were computed by defining a pseudo age determination at Q calculating the parameter representer of this measurement throughout the domain. The representer values were subsequently divided by the square root of the product of the prior parameter variance and the age variance at Q, to yield the correlations. The age variance at Q was calculated by computing the value of the Ln(τ) representer at location Q.

[46] Figures 5b, 5d and 5e were produced in a similar manner. For Figures 5b and 5e, however, the prior σ_h^2 field, which in contrast to the parameter variance is not known beforehand, was inter- and extrapolated from prior head variances calculated at the potential measurement locations. Inter- and extrapolation of these values is justified, as the head variance field is likely to be smooth.

[47] Figure 5c was constructed by defining in every grid cell a pseudo traveltime measurement and calculating their Ln(t) representers for Ln(t) at the location of the pseudo measurement itself (yielding the prior $\sigma^2_{\mathrm{Ln}(t)}$ field) and for Ln(t_Q) (yielding the cross covariances). Figure 5f was produced in a similar manner, with the distinction that now for every pseudo traveltime measurement the location representer was calculated for the particle that a priori is expected to break through in Q.

[48] From Figure 5 it appears that Ln(t_Q) is particularly strongly correlated (correlation coefficient $\rho > 0.6$), in a narrow zone above Q, with Ln(K) and Ln(t). Whereas the Ln(t) – Ln(t_Q) correlation (Figure 5c) obviously is the strongest at Q (here $\rho = 1$), the Ln(K) – Ln(t_Q) (Figure 5a) correlation is strongest around the center of the domain. This is because here Ln(K) maximally affects (via its covariance function) the conductivity of the area in which the particle moving toward Q is expected to travel, and therefore maximally affects Ln(t_Q). A similar correlation

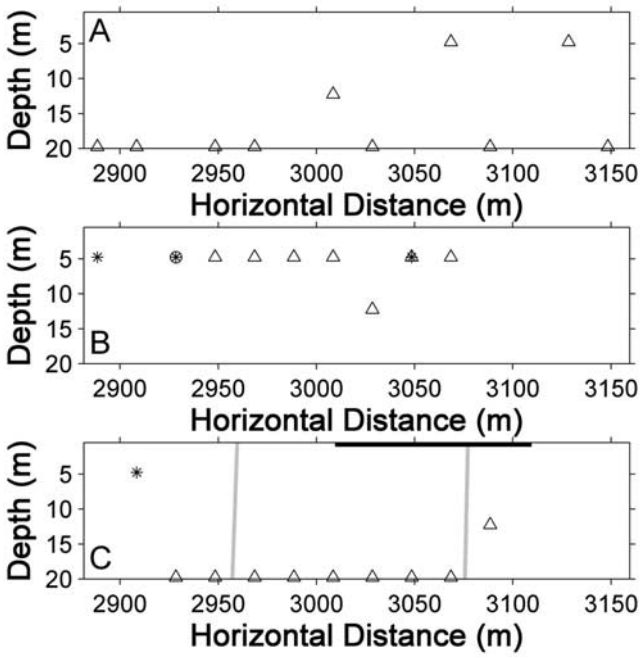

Figure 6. Near-optimal designs for minimizing the groundwater age uncertainty at the control plane (a), for minimizing the contaminant arrival location uncertainty (b), and for minimizing the contaminant breakthrough time uncertainty (c). The domains shown in this figure represent the confining unit in the area of interest (see Figure 1). The black bold line in Figure 6c represents a projection of the contaminant source zone on the top of the confining layer. Grey bold lines delineate the (a priori) expected contaminant flow through the confining layer. Δ = t measurement, * = h measurement, ○ = K measurement.

pattern was found by *Harvey and Gorelick* [1995] (their Figure 6b).

[49] Figure 5b shows that the head is only very weakly correlated with Ln(t). This is due to the small sensitivity of heads to the parameters.

[50] The curvature of the positive Ln(t) − Ln(t_Q) correlation zone in Figure 5c is caused by the fact that the particle arriving in Q has traveled a certain horizontal distance downgradient in Aquifer 1 before entering the confining layer (see also Figure 1). Traveltimes in the upper right part of the domain are correlated with the residence times of the corresponding particles in Aquifer 1, which are on their turn correlated with the residence time in Aquifer 1 of the particle arriving in Q.

[51] Whereas Ln(t_Q) particularly shows strong correlations with Ln(K) and Ln(t), χ_Q is especially strongly correlated with h and Ln(t) (Figures 5e and 5f, respectively). The large $h - \chi_Q$ correlations found in the upper part of the domain (Figure 5e) are related to the particle transport through Aquifer 1: because of the relatively large expected horizontal distance traveled in Aquifer 1 (see also Figure 1), χ_Q is determined for an important part by the hydraulic gradient in Aquifer 1, which on its own turn is correlated with the heads in the upper part of the confining layer.

[52] χ_Q − Ln(K) correlations (Figure 5d) are relatively weak, due to the limited vertical correlation of Ln(K) values. Figure 5f reveals that χ_Q shows strong correlations with Ln(t), particularly around Z. This makes sense, again because the horizontal distance traveled within the confining layer is expected to be very small and χ_Q is determined strongly by the distance travelled in Aquifer 1, which in turn is correlated with the vertical flow velocity in Aquifer 1. This vertical flow velocity is strongly correlated with the groundwater age near Z.

4.3. Examples of Network Designs

[53] Figures 6 shows examples of near-optimal 10-measurement network designs, in the search of which the Genetic Algorithm was allowed to choose among h, K and t measurements. The age estimation error standard deviation was set at 10% of the untransformed a priori expected value.

[54] In Figure 6a, the purpose of the observation network design was to minimize the groundwater age uncertainty at the control plane. The optimal design solely consists of age determinations, which is not surprising as they are strongly correlated with the groundwater age at the control plane. The horizontal correlation of traveltime is not strong enough for all determinations to be placed next to each other at the control plane, which would be a logical configuration as then they directly sample the goal variables. Instead, the upper right part of the domain is also sampled, for reasons explained in the discussion of Figure 5c.

[55] For the design shown Figure 6b, the sampling objective was to minimize the contaminant arrival location variance. K, h and t measurements are located in those areas where, based on Figures 5d–5f, strong correlations with the breakthrough location collection χ are expected.

[56] For Figure 6c, the objective was to minimize $\sigma^2_{\tau_{BT}}$ which is a combination of the objectives used in Figures 6a and 6b (see equation 1). The head measurement in the upper left-hand part of the confining layer primarily serves to constrain arrival location (see Figures 5e and 5b). The fact that the age determinations are placed at the control plane (as in Figure 6a), rather than in the upper part of the confining layer (as in Figure 6b), indicates that they are primarily important to obtain information on $t(x^{CP})$, rather than on $\chi(x^{SZ})$. Apart from one head measurement, no measurements are chosen that are specifically aimed at reducing $p(x^{CP} \in \chi)$. The age determinations sample a narrower area than in Figure 6a, because their placement is now focused at obtaining information on the traveltime toward the expected contaminant breakthrough zone, rather than toward the entire control plane. The focus of the design in Figure 6c on the reduction of arrival time uncertainty is probably due to the relatively wide contaminant source zone, which limits the reduction of the contaminant breakthrough location variance that can be achieved.

4.4. Performance of Age Estimation Networks Compared With *K* and *h* Networks

[57] Figure 7 shows the posterior breakthrough variance of (near-)optimal designs as a function of network size, for observation networks containing only one measurement type. The figure shows that, if the traveltime estimations are of a reasonable reliability, the uncertainty reduction that can be realized using traveltime determinations is larger

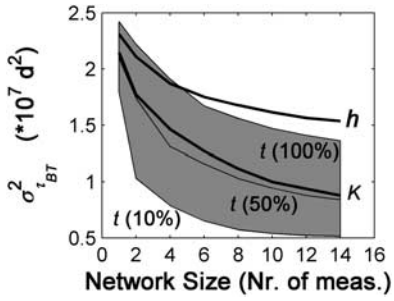

Figure 7. First-order estimation of the posterior breakthrough time uncertainty for networks consisting of only one observation type (h = pressure head, K = conductivity, t = groundwater age), as a function of network size. The given percentages refer to different levels of the age estimation error standard deviation. The prior breakthrough time uncertainty (no measurements) for this example was $4.85*10^7$ d^2.

than can be achieved with the other measurement types (with reasonable network sizes). After having studied Figure 5 this is not surprising anymore: heads are only very weakly correlated with traveltime, and conductivity is weakly correlated with arrival location. Groundwater age is the only variable that is strongly correlated with both arrival time and arrival location. Furthermore, groundwater age is more strongly correlated with arrival time than K is, and also more strongly correlated with arrival location than the head is. According to Figure 7, in our synthetic example an age determination network outperforms a K and an h network if the age estimation error standard deviation is less than 50% of the expected age, even though the error variances of the K and h measurements were chosen unrealistically small. So, if traveltime estimations of a reasonable accuracy are available, more information can be obtained with fewer measurements than with the other two measurement types.

5. Conclusions and Discussion

[58] A methodology was proposed that incorporates traveltime determinations into measurement network design. The methodology was focused on minimizing model prediction uncertainty (specifically, contaminant breakthrough time uncertainty) and is one of very few that take correlations between observations into account when doing so. Moreover, by directly calculating the covariances between observations and predictions, instead of evaluating the influence of the observations on the predictions via the posterior parameter covariance matrix, a major computational advantage is accomplished compared to previously reported network design algorithms aimed at prediction uncertainty minimization.

[59] As the Gaussian assumption often is not valid for traveltimes, they were transformed to comply with the Bayesian framework applied here. The natural logarithm of traveltime was shown to be approximately Gaussian distributed for systems of low to medium heterogeneity.

[60] Based on this result, we derived expressions for the linearized covariances between the measurable variables and the natural logarithm of traveltime (equation (C24)) and breakthrough time (equation (12)). The latter, together with the cross covariances between the observable variables and particle breakthrough locations, given by equation (13), are necessary for the calculation of posterior breakthrough time uncertainty given a certain set of measurements. In a synthetic example of contaminant breakthrough in a confining layer, the discrepancy between the linearized approximation of the prior variance of the natural logarithm of traveltime and Monte Carlo results was less then 10% for variances of the natural logarithm of the conductivity up to 3.0.

[61] Age estimations, if of a reasonable quality, were shown to be more valuable for the reduction of breakthrough time uncertainty than head and conductivity measurements. In our numerical example, even if the age estimation error standard deviation was taken as large as 50% of the expected value and head and conductivity measurement error variances were assumed to be very small, (near-)optimal age estimation networks still outperformed equally sized (near-)optimal networks of head or conductivity measurements.

[62] For sake of conciseness, we chose not to involve other considerations, like for example cost minimization, into the network design, other than restricting the designs to a certain number of measurements. However, such considerations can be accommodated in the Genetic Algorithm by expanding the objective function with additional terms. An example of a Genetic Algorithm application involving both prediction uncertainty minimization and budget constraints is given by *Wagner* [1995].

[63] In our calculations we assumed Gaussian distributed measurement errors for the log-transformed groundwater age estimations. It should be noted, however, that there are many factors contributing in different ways to the uncertainties involved in the translation of tracer concentrations to groundwater age, and the resultant of these factors might not always warrant the use of a Gaussian or even symmetric uncertainty structure. The most important sources of uncertainty are (note that not all uncertainty sources apply to all age dating techniques): pore-scale dispersion and macrodispersion resulting in mixed-age samples [*Maloszewski and Zuber*, 1982; *Bethke and Johnson*, 2002; *Weissmann et al.*, 2002; *Castro and Goblet*, 2005; *Manning et al.*, 2005], analytical error [*Solomon et al.*, 1992; *Ekwurzel et al.*, 1994], influence of uncertain recharge temperature affecting tracer concentrations at time of recharge [*Dunkle et al.*, 1993], nonconservative behavior of the tracer [*Dunkle et al.*, 1993], contamination [*Dunkle et al.*, 1993], water table fluctuations, and incomplete confinement of reactive decay products (e.g., ^{3}He [*Solomon et al.*, 1992]). It is yet unclear how reliable uncertainty estimates for age determinations can be obtained and whether corrections and transformations to the determinations can be formulated that can effectively make the uncertainty structure Gaussian. For arrival times in a conservative tracer test, on the other hand, these issues are far less problematic, considering the fact that the injection and detection times are relatively easy to determine and that the structure of the uncertainties involved in extracting the advective arrival time is likely to be either near-Gaussian or transformable to near-Gaussian.

[64] In the present study, reactive-dispersive behavior of the tracers, and the influence of instationarity on this

behavior, are accounted for only indirectly by treating them as sources of uncertainty to the traveltime estimations that are themselves used in a strictly advective, conservative and stationary inverse computational framework. As such, the approach is suitable for situations in which reaction, dispersion and instationarity are thought to be of secondary importance for the tracer. The assumption of stationarity will often be reasonable, as traveltimes determined for site characterization usually represent a long-term averaging over velocity fluctuations and are therefore likely to be relatively stationary.

[65] To keep the presentation simple, transport of the contaminant in the numerical example was assumed to be strictly advective and conservative as well. Although we recognize that reactive-dispersive behavior of the contaminant might be even a bigger source of uncertainty than the residence time of the groundwater within the transport volume, the considered measurement types, as treated in this paper, can only provide information on the latter factor. Therefore reactive-dispersive behavior of the contaminant is beyond the scope of the presented methodology.

[66] The method is applicable to real world cases as long as using the Bayesian framework and the linear theory is warranted, which respectively places demands on data availability (e.g., parameter statistics and uncertainties, measurement values and uncertainties, etc.) and puts restrictions on parameter variability. With regard to the computational demand of the method we can mention that the computation of the prior covariance matrix $M(\Xi, \Psi, \Theta^*, X)$ constitutes most of the computational burden. For the examples presented in section 4.3, this took about 45 min on a Dell personal computer with a Pentium 4 2.6 GHz processor and 2.5 GB RAM. This computer time increases with the number of potential and pseudo measurements: for every potential or pseudo head measurement a groundwater model run is performed twice and a particle tracking run is performed, for every potential and pseudo conductivity measurement a groundwater model and a particle tracking run are performed once, and for every potential and pseudo traveltime and location measurement two groundwater model runs and two particle-tracking runs are performed. Finally, for every potential and pseudo measurement a convolution with the parameter covariance matrix is required, so the number of unknown parameters (9,600 in this paper) also greatly influences computation times.

Appendix A: Derivation of the Euler-Lagrange Equations

[67] Forcing the conditions for a minimum in equation (11) yields a system of 9 coupled Euler-Lagrange equations. Besides the flow equation (6), the particle traveltime equation (7) and the particle location equations (8) and (9), this system consists of a traveltime adjoint equation, two particle location adjoint equations, a head adjoint equation, and a parameter equation, all of which are derived below.

A1. *X* Location Adjoint Equation

[68] The conditions for a minimum in the extended objective function (equation (11)) prescribe the variation of (11) with respect to the x location of tracked particles to be zero. This yields:

$$\frac{\partial J}{\partial x_i^s} + \frac{2\lambda_h A(\alpha) h}{\partial x_i^s} + \frac{\partial \sum_i^{N_i} \sum_s^{N_{s_i}} 2\lambda_{t_i}^s \left(\phi_{2_i}^s\right)}{\partial x_i^s} + \frac{\partial \sum_i^{N_i} \sum_s^{N_{s_i}} 2\lambda_{x_i}^s \left(\phi_{3_i}^s\right)}{\partial x_i^s} + \frac{\partial \sum_i^{N_i} \sum_s^{N_{s_i}} 2\lambda_{y_i}^s \left(\phi_{4_i}^s\right)}{\partial x_i^s} = 0 \tag{A1}$$

The derivatives in the third, fourth and the fifth term should only be taken for those travel steps s where x^s is really variable (i.e., not fixed by reaching an x-boundary of a cell in that travel step). Taking this into account, filling in equations (6)–(9) into (A1), working out the derivatives and dividing by 2 yields:

$$\begin{aligned} &- \frac{\partial M_p(h, \alpha, \mathrm{Ln}(t), \chi)}{\partial x_i^s} \left[P_v^{-1}\right]_{pn} [z_n - M_n(h, \alpha, \mathrm{Ln}(t), \chi)] \\ &- \lambda_{t_i}^s \frac{\partial \Delta t_i^s \left(\alpha, h, x_i^s, x_i^{s-1}, y_i^s, y_i^{s-1}\right)}{\partial x_i^s} \\ &- \lambda_{t_i}^{s+1} \frac{\partial \mathrm{D}t_i^{s+1} \left(\alpha, h_i^{s+1}, x_i^s, y_i^{s+1}, y_i^s\right)}{\partial x_i^s} \\ &+ \left(\lambda_{x_i}^s - \lambda_{x_i}^{s+1}\right) - \lambda_{y_i}^s \frac{\partial \mathrm{D}y_i^s \left(\alpha, h_i^s, x_i^{s-1}\right)}{\partial x_i^s} \\ &- \lambda_{y_i}^{s+1} \frac{\partial \mathrm{D}y_i^{s+1} \left(\alpha, h, x_i^{s+1}, x_i^s\right)}{\partial x_i^s} = 0 \qquad \text{for } x_i^s \neq x_{cell\ boundary} \end{aligned} \tag{A2}$$

and

$$\begin{aligned} &- \frac{\partial M_p(h, \alpha, \mathrm{Ln}(t), \chi)}{\partial x_i^s} \left[P_v^{-1}\right]_{pn} [z_n - M_n(h, \alpha, \mathrm{Ln}(t), \chi)] \\ &+ \lambda_{x_i}^s = 0 \qquad \text{for } x_i^s = x_{cell\ boundary} \end{aligned} \tag{A3}$$

Note that the second and the fifth term of equation (A2) are always zero because Δt_i^s and Δy_i^s only depend on the x coordinates of travel step s if the step size is limited by reaching the x-boundary of a cell, in which case equation (A3) applies. Taking this into account and rearranging yields the x location adjoint equation:

$$\begin{aligned} \lambda_{x_i}^s = \lambda_{x_i}^{s+1} &+ \frac{\partial M_p(h, \alpha, \mathrm{Ln}(t), \chi)}{\partial x_i^s} \left[P_v^{-1}\right]_{pn} \\ &\cdot [z_n - M_n(h, \alpha, \mathrm{Ln}(t), \chi)] \\ &+ \lambda_{t_i}^{s+1} \frac{\partial \Delta t_i^{s+1} \left(\alpha_i^{s+1}, x_i^s, y_i^{s+1}, y_i^s\right)}{\partial x_i^s} \\ &+ \lambda_{y_i}^{s+1} \frac{\partial \Delta y_i^{s+1} \left(\alpha, h, x_i^{s+1}, x_i^s\right)}{\partial x_i^s} \qquad \text{for } x_i^s \neq x_{cell\ boundary} \end{aligned} \tag{A4}$$

$$\lambda_{x_i}^s = \frac{\partial M_p(h,\alpha,\mathrm{Ln}(t),\chi)}{\partial x_i^s}\left[P_v^{-1}\right]_{pn}[z_n - M_n(h,\alpha,\mathrm{Ln}(t),\chi)] = 0 \quad \text{for } x_i^s = x_{cell\ boundary} \tag{A5}$$

with boundary condition $\lambda_{x_i}^{N_{s_i+1}} = 0$.

A2. *Y* Location Adjoint Equation

[69] Following the same reasoning as in the derivation of the x location adjoint equation, the y location adjoint equation is given by:

$$\begin{aligned}\lambda_{y_i}^s = {} & \lambda_{y_i}^{s+1} + \frac{\partial M_p(h,\alpha,\mathrm{Ln}(t),\chi)}{\partial y_i^s}\left[P_v^{-1}\right]_{pn} \\ & \cdot [z_n - M_n(h,\alpha,\mathrm{Ln}(t),\chi)] \\ & + \lambda_{t_i}^{s+1}\frac{\partial \Delta t_i^{s+1}(\alpha,h,x_i^{s+1},x_i^s,y_i^{s+1},y_i^s)}{\partial y_i^s} \\ & + \lambda_{x_i}^{s+1}\frac{\partial \Delta x_i^{s+1}(\alpha,h,y_i^{s+1},y_i^s)}{\partial y_i^s} \quad \text{for } y_i^s \neq y_{cell\ boundary}\end{aligned} \tag{A6}$$

$$\lambda_{y_i}^s = \frac{\partial M_p(h,\alpha,\mathrm{Ln}(t),\chi)}{\partial y_i^s}\left[P_v^{-1}\right]_{pn} \cdot [z_n - M_n(h,\alpha,\mathrm{Ln}(t),\chi)] = 0 \quad \text{for } y_i^s = y_{cell\ boundary} \tag{A7}$$

with boundary condition $\lambda_{t_i}^{N_{s_i+1}} = 0$.

A3. Traveltime Adjoint Equation

[70] The conditions for a minimum in the extended objective function (equation (11)) prescribe the variation of (11) with respect to the traveltime to be zero. This yields:

$$\begin{aligned}\frac{\partial J}{\partial t_i^s} + \frac{2\lambda_h A(\alpha)h}{\partial t_i^s} + \frac{\partial \sum_i^{N_i}\sum_s^{N_{s_i}} 2\lambda_{t_i}^s\left(\phi_{2_i}^s\right)}{\partial t_i^s} \\ + \frac{\partial \sum_i^{N_i}\sum_s^{N_{s_i}} 2\lambda_{x_i}^s\left(\phi_{3_i}^s\right)}{\partial t_i^s} + \frac{\partial \sum_i^{N_i}\sum_s^{N_{s_i}} 2\lambda_{y_i}^s\left(\phi_{4_i}^s\right)}{\partial t_i^s} = 0\end{aligned} \tag{A8}$$

[71] The second, fourth and fifth term of equation (A8) are zero. Working out the other terms and dividing by 2 directly yields the traveltime adjoint equation:

$$\lambda_{t_i}^s = \lambda_{t_i}^{s+1} + \frac{\partial M_p(h,\alpha,\mathrm{Ln}(t),\chi)}{\partial t_i^s}\left[P_v^{-1}\right]_{pn}[z_n - M_n(h,\alpha,\mathrm{Ln}(t),\chi)] \tag{A9}$$

with boundary condition $\lambda_{t_i}^{N_{s_i+1}} = 0$.

A4. Head Adjoint Equation

[72] The conditions for a minimum in the extended objective function (equation (11)) prescribe the variation of (11) with respect to the heads to be zero. This yields:

$$\begin{aligned}\frac{\partial J}{\partial h} + \frac{2\lambda_h A(\alpha)h}{\partial h} + \frac{\partial \sum_i^{N_i}\sum_s^{N_{s_i}} 2\lambda_{t_i}^s\left(\phi_{2_i}^s\right)}{\partial h} \\ + \frac{\partial \sum_i^{N_i}\sum_s^{N_{s_i}} 2\lambda_{x_i}^s\left(\phi_{3_i}^s\right)}{\partial h} + \frac{\partial \sum_i^{N_i}\sum_s^{N_{s_i}} 2\lambda_{y_i}^s\left(\phi_{4_i}^s\right)}{\partial h} = 0\end{aligned} \tag{A10}$$

Working out equation (A10), dividing by 2 and rearranging yields the head adjoint equation:

$$\begin{aligned}A_{gf}(\alpha)\lambda_{h_f} = {} & \frac{\partial M_p(h,\alpha,\mathrm{Ln}(t),\chi)}{\partial h_g}\left[P_v^{-1}\right]_{pn} \\ & \cdot [z_n - M_n(h,\alpha,\mathrm{Ln}(t),\chi)] \\ & + \sum_i^{N_i}\sum_s^{N_{s_i}} \lambda_{t_i}^s \frac{\partial \Delta t_i^s(\alpha,h,x_i^s,x_i^{s-1},y_i^s,y_i^{s-1})}{\partial h_g} \\ & + \sum_i^{N_i}\sum_s^{N_{s_i}} \lambda_{x_i}^s \frac{\partial \Delta x_i^s(\alpha,h,y_i^s,y_i^{s-1})}{\partial h_g} \\ & + \sum_i^{N_i}\sum_s^{N_{s_i}} \lambda_{y_i}^s \frac{\partial \Delta y_i^s(\alpha,h,x_i^s,x_i^{s-1})}{\partial h_g}\end{aligned} \tag{A11}$$

A5. Parameter Equation

[73] The conditions for a minimum in the extended objective function (equation (11)) prescribe the variation of (11) with respect to the parameters to be zero. This yields:

$$\begin{aligned}\frac{\partial J}{\partial \alpha} + \frac{2\lambda_h A(\alpha)h}{\partial \alpha} + \frac{\partial \sum_i^{N_i}\sum_s^{N_{s_i}} 2\lambda_{t_i}^s\left(\phi_{2_i}^s\right)}{\partial \alpha} \\ + \frac{\partial \sum_i^{N_i}\sum_s^{N_{s_i}} 2\lambda_{x_i}^s\left(\phi_{3_i}^s\right)}{\partial \alpha} + \frac{\partial \sum_i^{N_i}\sum_s^{N_{s_i}} 2\lambda_{y_i}^s\left(\phi_{4_i}^s\right)}{\partial \alpha} = 0\end{aligned} \tag{A12}$$

[74] Working out equation (A12), dividing by 2 and rearranging yields the parameter equation:

$$\begin{aligned}\alpha_l = \overline{\alpha}_l - P_\alpha \Bigg[& h_g \frac{\partial A(\alpha)}{\partial \alpha_k}\lambda_{h_f} \\ & - \sum_i^{N_i}\sum_s^{N_{s_i}} \lambda_{t_i}^s \frac{\partial \Delta t_i^s(\alpha,h,x_i^s,x_i^{s-1},y_i^s,y_i^{s-1})}{\partial \alpha_k} \\ & - \sum_i^{N_i}\sum_s^{N_{s_i}} \lambda_{x_i}^s \frac{\partial \Delta x_i^s(\alpha,h,y_i^s,y_i^{s-1})}{\partial \alpha_k} \\ & - \sum_i^{N_i}\sum_s^{N_{s_i}} \lambda_{y_i}^s \frac{\partial \Delta y_i^s(\alpha,h,x_i^s,x_i^{s-1})}{\partial \alpha_k} \Bigg]\end{aligned} \tag{A13}$$

Appendix B: Representer Expansions

[75] In the representer-based inverse algorithm [*Valstar et al.*, 2004], the unknowns in the nonlinear system of Euler-Lagrange equations are expanded in finite series, allowing the equations in the system to be decoupled and iteratively solved. Every measurement adds a term to this finite series, consisting of (1) a basis function or representer, quantifying the influence of the measurement on the estimate of the variable for which the representer is defined, and (2) its coefficient, quantifying the weight given to the representer, which depends on the misfit between measurement value and measurement prediction. The definitions of the representer functions are:

$$\lambda_{t_i} = \sum_{p=1}^{N_p} \Lambda_{ip} b_p \tag{B1}$$

$$\lambda_{x_i} = \sum_{p=1}^{N_p} \Pi_{ip} b_p \tag{B2}$$

$$\lambda_{y_i} = \sum_{p=1}^{N_p} \mathrm{H}_{ip} b_p \tag{B3}$$

$$\lambda_{h_f} = \sum_{p=1}^{N_p} \Gamma_{fp} b_p \tag{B4}$$

$$\alpha_l = \overline{\alpha}_l + \sum_{p=1}^{N_p} \Psi_{lp} b_p \tag{B5}$$

$$h_g = h_{F_g} + h_{corr_g} + \sum_{p=1}^{N_p} \Xi_{gp} b_p \tag{B6}$$

$$x_i = x_{F_i} + x_{corr_i} + \sum_{p=1}^{N_p} \mathrm{X}_{ip} b_p \tag{B7}$$

$$y_i = y_{F_i} + y_{corr_i} + \sum_{p=1}^{N_p} \mathrm{Y}_{ip} b_p \tag{B8}$$

$$t_i^{\eta} = t_{F_i} + t_{corr_i} + \sum_{p=1}^{N_p} t_i^{\eta-1} \Theta_{ip}^{*} b_p \tag{B9}$$

where b is the vector of representer coefficients, N_p is the number of measurements, Λ_{ip} is the traveltime adjoint representer of measurement p, calculated for particle i, Π_{ip} is the x location adjoint representer of measurement p, calculated for particle i, H_{ip} is the y location adjoint representer of measurement p, calculated for particle i, Γ_{fp} is the head adjoint representer of measurement p, calculated for the head state variable f, Ψ_{lp} is the parameter representer of measurement p, calculated for parameter l, Ξ_{gp} is the head representer of measurement p, calculated for head state variable g, X_{ip} and Y_{ip} are respectively the x and y location representer of measurement p, calculated for particle i, η is the iteration number and Θ_{ip}^{*} is the $\mathrm{Ln}(t)$ representer of measurement p, calculated for particle i. In order to be able to expand the untransformed traveltime t_i, Θ_{pi}^{*} is multiplied by the derivative $dt_i/d\ln t_i = t_i$, which is estimated using its value in the previous iteration ($t_i^{\eta-1}$). h_F, x_F, y_F, and t_F are the solutions obtained by solving equations (6)–(9) with $\alpha = \overline{\alpha}$. h_{corr}, x_{corr}, y_{corr}, and t_{corr} are correction terms. In our first-order design method, the algorithm is terminated after the first iteration, the unknown variables are not actually updated (this would require actual sampling) and therefore the representer coefficients b and the correction terms h_{corr}, x_{corr}, y_{corr}, and t_{corr} do not have to be calculated. The algorithm is initiated using $\alpha = \overline{\alpha}$, $h = h_F$, $t_i = t_{i_f}$, $x_i = x_{i_f}$, $y_i = y_{i_f}$, and $\lambda_h = \lambda_{t_i} = \lambda_{x_i} = \lambda_{y_i} = 0$.

Appendix C: Representer Derivations

[76] For the sake of readability, the explicit statement of the dependencies of A, Δt, Δx and Δy are omitted from now on.

C1. *X* Location Adjoint Representers

[77] Inserting the representer expansions (B1)–(B9) into the x location adjoint equations (A4)–(A5) yields:

$$\Pi_{ip}^{s} b_p = \Pi_{ip}^{s+1} b_p + \frac{\partial M_p(h, \alpha, \mathrm{Ln}(t), \chi)}{\partial x_i^s} \left[P_v^{-1}\right]_{pn} (z_n - M_n(\,)) + \Lambda_{ip}^{s+1} b_p \frac{\partial \Delta t_i^{s+1}}{\partial x_i^s} + \mathrm{H}_{ip}^{s+1} b_p \frac{\partial \Delta y_i^{s+1}}{\partial x_i^s} \quad \text{for} \quad x_i^s \neq x_{cell\ boundary} \tag{C1}$$

$$\Pi_{ip}^{s} b_p = \frac{\partial M_p(h, \alpha, \mathrm{Ln}(t), \chi)}{\partial x_i^s} \left[P_v^{-1}\right]_{pn} (z_n - M_n(\,)) \quad \text{for} \quad x_i^s = x_{cell\ boundary} \tag{C2}$$

where $M_n(\,) = M_n(h_F + h_{corr} + \Xi b, \overline{\alpha}_l + \Psi b, \mathrm{Ln}(t_F + t_{corr} + t_F \Theta^{*} b), x_F + x_{corr} + \mathrm{X} b)$. The coefficients b for all p measurements are now chosen as:

$$b_p = \left[P_v^{-1}\right]_{pn} (z_n - M_n(\,)) \tag{C3}$$

Equations (C1)–(C2) can only be fulfilled for nonzero b if:

$$\Pi_{ip}^{s} = \Pi_{ip}^{s+1} + \frac{\partial M_p(h, \alpha, \mathrm{Ln}(t), \chi)}{\partial x_i^s} + \Lambda_{ip}^{s+1} \frac{\partial \Delta t_i^{s+1}}{\partial x_i^s} + \mathrm{H}_{ip}^{s+1} \frac{\partial \Delta y_i^{s+1}}{\partial x_i^s} \quad \text{for} \quad x_i^s \neq x_{cell\ boundary} \tag{C4}$$

$$\Pi_{ip}^{s} = \frac{\partial M_p(h, \alpha, \mathrm{Ln}(t), \chi)}{\partial x_i^s} = 0 \quad \text{for} \quad x_i^s = x_{cell\ boundary} \tag{C5}$$

with boundary condition $\Pi_{ip}^{N_{si}+1} = 0$. The second term on the right hand side of (C4) and the term on the right hand side of (C5) are nonzero only in case of (pseudo) measurements of χ_i (by definition at $s = N_{s_i}$). In that case $\frac{\partial M_p(\chi_i)}{\partial x_i^{N_{s_i}}} = 1$.

C2. *Y* Location Adjoint Representers

[78] Following the same reasoning as above, it follows for the y location adjoint representers:

$$\mathrm{H}_{ip}^{s} = \mathrm{H}_{ip}^{s+1} + \frac{\partial M_p(h,\alpha,\mathrm{Ln}(t),\chi)}{\partial y_i^s} + \Lambda_{ip}^{s+1}\frac{\partial \Delta t_i^{s+1}}{\partial y_i^s} + \Pi_{ip}^{s+1}\frac{\partial \Delta x_i^{s+1}}{\partial y_i^s} \quad \text{for } y_i^s \neq y_{cell\ boundary} \tag{C6}$$

$$\mathrm{H}_{ip}^{s} = \frac{\partial M_p(h,\alpha,\mathrm{Ln}(t),\chi)}{\partial y_i^s} = 0 \quad \text{for } y_i^s = y_{cell\ boundary} \tag{C7}$$

with boundary condition $\mathrm{H}_{ip}^{N_{si}+1} = 0$. The second term on the right hand side of (C6) and the term on the right hand side of (C7) are always zero. Therefore (C6)–(C7) can be simplified into:

$$\mathrm{H}_{ip}^{s} = \mathrm{H}_{ip}^{s+1} - \Lambda_{ip}^{s+1}\frac{\partial \Delta t_i^{s+1}}{\partial y_i^s} - \Pi_{ip}^{s+1}\frac{\partial \Delta x_i^{s+1}}{\partial y_i^s} \quad \text{for } y_i^s \neq y_{cell\ boundary} \tag{C8}$$

$$\mathrm{H}_{ip}^{s} = 0 \quad \text{for } y_i^s = y_{cell\ boundary} \tag{C9}$$

with boundary condition $\mathrm{H}_{ip}^{N_{si}+1} = 0$.

C3. Traveltime Adjoint Representers

[79] Inserting the representer expansions (B1)–(B9) into the traveltime adjoint equation (A9) yields:

$$\Lambda_{ip}^{s} b_p = \Lambda_{ip}^{s+1} b_p + \frac{\partial M_p(h,\alpha,\mathrm{Ln}(t),\chi)}{\partial t_i^s}\left[P_v^{-1}\right]_{pn}(z_n - M_n(\ )) \tag{C10}$$

After inserting (C3) into (C10), the resulting equation can only be fulfilled for non-zero b if:

$$\Lambda_{ip}^{s} = \Lambda_{ip}^{s+1} + \frac{\partial M_p(h,\alpha,\mathrm{Ln}(t),\chi)}{\partial t_i^s} \tag{C11}$$

with boundary condition $\Lambda_{ip}^{N_s+1} = 0$. The second term of equation (C11) is nonzero only in case of measurements of $\mathrm{Ln}(t_i)$. In that case $\frac{\partial M_p(\mathrm{Ln}(t_i))}{\partial t_i^s} = \frac{1}{t_i^s}$, with t_i^s approximated by its most recent estimate.

C4. Head Adjoint Representers

[80] Inserting the representer definitions (B1)–(B9) into the head adjoint equation (A11) yields:

$$A_{gf}\Gamma_{fp}b_p = \frac{\partial M_p(h,\alpha,\mathrm{Ln}(t),\chi)}{\partial h_g}\left[P_v^{-1}\right]_{pn}(z_n - M_n(\ )) + \sum_i^{N_i}\sum_s^{N_{s_i}}\Lambda_{ip}^{s}b_p\frac{\partial \Delta t_i^s}{\partial h_g} + \sum_i^{N_i}\sum_s^{N_{s_i}}\Pi_{ip}^{s}b_p\frac{\partial \Delta x_i^s}{\partial h_g} + \sum_i^{N_i}\sum_s^{N_{s_i}}\mathrm{H}_{ip}^{s}b_p\frac{\partial \Delta y_i^s}{\partial h_g} \tag{C12}$$

[81] Equation (C12) can only be fulfilled for nonzero b if

$$A_{gf}\Gamma_{fp} = \frac{\partial M_p(h,\alpha,\mathrm{Ln}(t),\chi)}{\partial h_g} + \sum_i^{N_i}\sum_s^{N_{s_i}}\Lambda_{ip}^{s}\frac{\partial \Delta t_i^s}{\partial h_g} + \sum_i^{N_i}\sum_s^{N_{s_i}}\Pi_{ip}^{s}\frac{\partial \Delta x_i^s}{\partial h_g} + \sum_i^{N_i}\sum_s^{N_{s_i}}\mathrm{H}_{ip}^{s}\frac{\partial \Delta y_i^s}{\partial h_g} \tag{C13}$$

C5. Parameter Representers

[82] Inserting the representer definitions into the parameter equation (A13) yields:

$$\bar{\alpha}_l + \Psi_{lp}b_p = \bar{\alpha}_l - P_{\alpha_{lk}}\left[h_g\frac{\partial A_{gf}}{\partial \alpha_k}\Gamma_{fp}b_p - \sum_i^{N_i}\sum_s^{N_{s_i}}\Lambda_{ip}b_p\frac{\partial \Delta t_i^s}{\partial \alpha_k} - \sum_i^{N_i}\sum_s^{N_{s_i}}\Pi_{ip}b_p\frac{\partial \Delta x_i^s}{\partial \alpha_k} - \sum_i^{N_i}\sum_s^{N_{s_i}}\mathrm{H}_{ip}b_p\frac{\partial \Delta y_i^s}{\partial \alpha_k}\right] \tag{C14}$$

Equation (C14) can only be fulfilled for nonzero b if:

$$\Psi_{lp} = -P_{\alpha_{lk}}\left[h_g\frac{\partial A_{gf}}{\partial \alpha_k}\Gamma_{fp} - \sum_i^{N_i}\sum_s^{N_{s_i}}\Lambda_{ip}\frac{\partial \Delta t_i^s}{\partial \alpha_k} - \sum_i^{N_i}\sum_s^{N_{s_i}}\Pi_{ip}\frac{\partial \Delta x_i^s}{\partial \alpha_k} - \sum_i^{N_i}\sum_s^{N_{s_i}}\mathrm{H}_{ip}\frac{\partial \Delta y_i^s}{\partial \alpha_k}\right] \tag{C15}$$

C6. Head Representers

[83] The head representers should be the exact linearizations of the heads around the head estimates of the previous iteration. For a derivation see *Valstar* [2001] and *Valstar et al.* [2004]. The result is:

$$A_{fg}\Xi_{gp} = -\frac{\partial A_{fg}}{\partial \alpha_k}\Psi_{kp}h_g \tag{C16}$$

C7. *X* Location Representers

[84] For the calculation of the x location representer, an exact linearization of the x location equation (8) is performed around the most recent estimate of x_i. Since the variation in the parameters induced by the measurements p is given by $\Psi_{lp}\partial b_p$ (B5), the variation of the heads by $\Xi_{gp}\partial b_p$ (B6), the variation of the y location of particle i by $\mathrm{Y}_{ip}\partial b_p$ (B8), and the variation of the traveltime of particle i by $t_i\Theta_{ip}^{*}\partial b_p$ (B9), a linear relationship between the x

location of particle i on one hand and the parameters, heads, y location of particle i and traveltime of particle i on the other implies that the variation in x locations of particles, given by $X_{ip}\partial b_p$ (see equation (B7)), is

$$X^s_{ip}\partial b_p = \sum_{s'=1}^{s}\frac{\partial \Delta x_i^{s'}}{\partial \alpha_l}\Psi_{lp}\partial b_p + \sum_{s'=1}^{s}\frac{\partial \Delta x_i^{s'}}{\partial h_g}\Xi_{gp}\partial b_p + \sum_{s'=1}^{s}\frac{\partial \Delta x_i^{s'}}{\partial y_i^{s'}}Y^{s'}_{ip}\partial b_p + \sum_{s'=1}^{s}\frac{\partial \Delta x_i^{s'}}{\partial y_i^{s'-1}}Y^{s'-1}_{ip}\partial b_p, \quad \text{for } x_i^s \neq x_{cell\ boundary} \tag{C17}$$

$$X^s_{ip}\partial b_p = \sum_{s'=1}^{s-1}\frac{\partial \Delta x_i^{s'}}{\partial \alpha_l}\Psi_{lp}\partial b_p + \sum_{s'=1}^{s-1}\frac{\partial \Delta x_i^{s'}}{\partial h_g}\Xi_{gp}\partial b_p + \sum_{s'=1}^{s-1}\frac{\partial \Delta x_i^{s'}}{\partial y_i^{s'}}Y^{s'}_{ip}\partial b_p + \sum_{s'=1}^{s-1}\frac{\partial \Delta x_i^{s'}}{\partial y_i^{s'-1}}Y^{s'-1}_{ip}\partial b_p, \quad \text{for } x_i^s = x_{cell\ boundary} \tag{C18}$$

with boundary condition $Y^0_{ip} = 0$. The function of s' is the same as that of s. Dividing (C17)–(C18) by ∂b_p and putting it in sequential notation yields the x location representer:

$$X^s_{ip} = X^{s-1}_{ip} + \frac{\partial \Delta x_i^s}{\partial \alpha_l}\Psi_{lp} + \frac{\partial \Delta x_i^s}{\partial h_g}\Xi_{gp} + \frac{\partial \Delta x_i^s}{\partial y_i^s}Y^s_{ip} + \frac{\partial \Delta x_i^s}{\partial y_i^{s-1}}Y^{s-1}_{ip}, \quad \text{for } x_i^s \neq x_{cell\ boundary} \tag{C19}$$

$$X^s_{ip} = X^{s-1}_{ip} \quad \text{for } x_i^s = x_{cell\ boundary} \tag{C20}$$

with boundary conditions $X^0_{ip} = Y^0_{ip} = 0$.

C8. *Y* Location Representers

[85] The same reasoning holds for the y location representers:

$$Y^s_{ip} = Y^{s-1}_{ip} + \frac{\partial \Delta x_i^s}{\partial \alpha_l}\Psi_{lp} + \frac{\partial \Delta y_i^s}{\partial h_g}\Xi_{gp} + \frac{\partial \Delta y_i^s}{\partial x_i^s}X^s_{ip} + \frac{\partial \Delta y_i^s}{\partial x_i^{s-1}}X^{s-1}_{ip}, \quad \text{for } y_i^s \neq y_{cell\ boundary} \tag{C21}$$

$$Y^s_{ip} = Y^{s-1}_{ip} \quad \text{for } y_i^s = y_{cell\ boundary} \tag{C22}$$

with boundary conditions $Y^0_{ip} = X^0_{ip} = 0$.

C9. Ln(Traveltime) Representers

[86] For the calculation of the Ln(t) representer, an exact linearization of the traveltime equation (7) is performed around the most recent estimate of the traveltime. Following similar considerations as above, the variation in Ln(t_i), given by $t_i\Theta^*_{ip}\partial b_p$ (see equation (B9)), is

$$t_i^s\Theta^{*s}_{ip}\partial b_p = \sum_{s'=1}^{s}\frac{\partial \Delta t_i^{s'}}{\partial \alpha_l}\Psi_{lp}\partial b_p + \sum_{s'=1}^{s}\frac{\partial \Delta t_i^{s'}}{\partial h_g}\Xi_{gp}\partial b_p + \sum_{s'=1}^{s}\frac{\partial \Delta t_i^{s'}}{\partial x_i^{s'}}X^{s'}_{ip}\partial b_p + \sum_{s'=1}^{s}\frac{\partial \Delta t_i^{s'}}{\partial x_i^{s'-1}}X^{s'-1}_{ip}\partial b_p + \sum_{s'=1}^{s}\frac{\partial \Delta t_i^{s'}}{\partial y_i^{s'}}Y^{s'}_{ip}\partial b_p + \sum_{s'=1}^{s}\frac{\partial \Delta t_i^{s'}}{\partial y_i^{s'-1}}Y^{s'-1}_{ip}\partial b_p \tag{C23}$$

[87] Dividing (C23) by ∂b_p, rearranging, and putting it in sequential notation yields the Ln(t) representer:

$$\Theta^{*s}_{ip} = \frac{1}{t_i^s}\left[\Theta^{*s-1}_{ip} + \frac{\partial \Delta t_i^s}{\partial \alpha_l}\Psi_{lp} + \frac{\partial \Delta t_i^s}{\partial h_g}\Xi_{gp} + \frac{\partial \Delta t_i^s}{\partial x_i^s}X^s_{ip} + \frac{\partial \Delta t_i^s}{\partial x_i^{s-1}}X^{s-1}_{ip} + \frac{\partial \Delta t_i^s}{\partial y_i^s}Y^s_{ip} + \frac{\partial \Delta t_i^s}{\partial y_i^{s-1}}Y^{s-1}_{ip}\right] \tag{C24}$$

with boundary conditions $\Theta^{*0}_{ip} = X^0_{ip} = Y^0_{ip} = 0$.

C10. Correction Terms

[88] The correction terms h_{corr}, t_{corr}, x_{corr}, and y_{corr} are chosen in such a way that the forward flow (6), the particle traveltime (7), the particle x-location (8), and the particle y-location equation (9) equation are fulfilled, respectively. For h_{corr}, first the unknown variables in (6) are replaced with the representer expansions (B1)–(B9), and subsequently the head representer equation (C16), multiplied by b_p and summed over all measurements, is inserted. This yields [*Valstar*, 2001; *Valstar et al.*, 2004]:

$$A_{fg}h_{corr_g} = q_f + \frac{\partial A_{fg}}{\partial \alpha_k}(\alpha - \bar{\alpha})h_g - A_{fg}h_{F_g} \tag{C25}$$

[89] For t_{corr}, first the unknown variables in (7) are replaced with the representer expansions (B1)–(B9), and subsequently the Ln(traveltime) representer equation (C24), multiplied by b_p and summed over all measurements, is inserted. This yields:

$$t^s_{corr_i} = t_i^{s-1} + \Delta t_i^s - t^s_{F_i} - \Theta^{*s-1}_{ip}b_p - \frac{\partial \Delta t_i^s}{\partial \alpha_l}\Psi_{lp}b_p - \frac{\partial \Delta t_i^s}{\partial h_g}\Xi_{gp}b_p - \frac{\partial \Delta t_i^s}{\partial x_i^s}X^s_{ip}b_p \tag{C26}$$

[90] Since $t_i^{s-1} - \Theta^{*s-1}_{ip}b_p = t^{s-1}_{corr_i} + t^{s-1}_{F_i}$ (see equation (B9)), we have:

$$t^s_{corr_i} = t^{s-1}_{corr_i} + t^{s-1}_{F_i} - t^s_{F_i} + \Delta t_i^s - \frac{\partial \Delta t_i^s}{\partial \alpha_l}\Psi_{lp}b_p - \frac{\partial \Delta t_i^s}{\partial h_g}\Xi_{gp}b_p - \frac{\partial \Delta t_i^s}{\partial x_i^s}X^s_{ip}b_p - \frac{\partial \Delta t_i^s}{\partial x_i^{s-1}}X^{s-1}_{ip} - \frac{\partial \Delta t_i^s}{\partial y_i^s}Y^s_{ip} - \frac{\partial \Delta t_i^s}{\partial y_i^{s-1}}Y^{s-1}_{ip} \tag{C27}$$

[91] Following similar reasoning, we have for x_{corr}:

$$x^s_{corr_i} = x^{s-1}_{corr_i} + x^{s-1}_{F_i} - x^s_{F_i} + \Delta x^s_i - \frac{\partial \Delta x^s_i}{\partial \alpha_l}\Psi_{lp}b_p - \frac{\partial \Delta x^s_i}{\partial h_g}\Xi_{gp}b_p - \frac{\partial \Delta x^s_i}{\partial y^s_i}\mathrm{Y}^s_{ip}b_p - \frac{\partial \Delta x^s_i}{\partial y^{s-1}_i}\mathrm{Y}^{s-1}_{ip}b_p \quad \text{for } x^s_i \neq x_{cell\ boundary} \tag{C28}$$

$$x^s_{corr_i} = -x^s_{F_i} + x_{cell\ boundary} \quad \text{for } x^s_i \neq x_{cell\ boundary} \tag{C29}$$

And for y_{corr}:

$$y^s_{corr_i} = y^{s-1}_{corr_i} + y^{s-1}_{F_i} - y^s_{F_i} + \Delta y^s_i - \frac{\partial \Delta y^s_i}{\partial \alpha_l}\Psi_{lp}b_p - \frac{\partial \Delta y^s_i}{\partial h_g}\Xi_{gp}b_p - \frac{\partial \Delta y^s_i}{\partial x^s_i}\mathrm{X}^s_{ip}b_p - \frac{\partial \Delta y^s_i}{\partial x^{s-1}_i}\mathrm{X}^{s-1}_{ip}b_p \quad \text{for } y^s_i \neq y_{cell\ boundary} \tag{C30}$$

$$y^s_{corr_i} = -y^s_{F_i} + y_{cell\ boundary} \quad \text{for } y^s_i \neq y_{cell\ boundary} \tag{C31}$$

[92] During iteration η, $\Psi_{lp}b_p$, $\Xi_{gp}b_p$, $\mathrm{X}^s_{ip}b_p$, $\mathrm{X}^{s-1}_{ip}b_p$, $\mathrm{Y}^s_{ip}b_p$ and $\mathrm{Y}^{s-1}_{ip}b_p$ in equations (C27)–(C31) are not known and replaced by the estimations from the previous iteration: $(\alpha_l^{\eta-1} - \overline{\alpha}_l)$, $(h_g^{\eta-1} - h_{F_g}^{\eta-1} - h_{corr_g}^{\eta-1})$, $(x_i^{s^{\eta-1}} - x_{F_i}^{s^{\eta-1}} - x_{corr_i}^{s^{\eta-1}})$, $(x_i^{s-1^{\eta-1}} - x_{F_i}^{s-1^{\eta-1}} - x_{corr_i}^{s-1^{\eta-1}})$, $(y_i^{s^{\eta-1}} - y_{F_i}^{s^{\eta-1}} - y_{corr_i}^{s^{\eta-1}})$, and $(y_i^{s-1^{\eta-1}} - y_{F_i}^{s-1^{\eta-1}} - y_{corr_i}^{s-1^{\eta-1}})$, respectively.

References

Anderman, E. R., and M. C. Hill (2001), MODFLOW-2000, the U.S. Geological Survey modular ground-water model - documentation of the advective-transport observation (ADV2) package, version 2, open-file report 01-54, U.S. Geological Survey, Denver, CO.

Bethke, C. M., and T. M. Johnson (2002), Paradox of groundwater age: Correction, *Geology*, *30*, 385–388.

Bogardi, I., A. Bardossy, and L. Duckstein (1985), Multicriterion network design using geostatistics, *Water Resour. Res.*, *21*, 199–208.

Castro, M. C., and P. Goblet (2005), Calculation of ground water ages - A comparative analysis, *Ground Water*, *43*, 368–380.

Chang, L.-F., N.-Z. Sun, and W. W.-G. Yeh (2005), Optimal observation network design for parameter structure identification in groundwater modeling, *Water Resour. Res.*, *41*, W03002, doi:10.1029/2004WR003514.

Cirpka, O. A., and P. K. Kitanidis (2000), An advective-dispersive stream tube approach for the transfer of conservative-tracer data to reactive transport, *Water Resour. Res.*, *36*, 1209–1220.

Cirpka, O. A., and P. K. Kitanidis (2001), Sensitivity of temporal moments calculated by the adjoint-state method and joint inversing of head and tracer data, *Adv. Water Res.*, *24*, 89–103.

Cvetkovic, V. D., A. M. Shapiro, and G. Dagan (1992), A solute flux approach to transport in heterogeneous formations 2. Uncertainty analysis, *Water Resour. Res.*, *28*, 1377–1388.

Dagan, G. (1989), *Flow and Transport in Porous Formations*, Springer-Verlag, Berlin.

Dettinger, M. D., and J. L. Wilson (1981), First order analysis of uncertainty in numerical methods of groundwater flow 1, Mathematical development, *Water Resour. Res.*, *17*, 149–161.

Dunkle, S. A., L. N. Plummer, E. Busenberg, P. J. Phillips, J. M. Denver, A. Hamilton, R. L. Michel, and T. B. Coplen (1993), Chlorofluorocarbons (CCl3F and CCl2F2) as dating tools and hydrologic tracers in shallow groundwater of the Delmarva Peninsula, Atlantic Coastal plane, *Water Resour. Res.*, *29*, 3837–3860.

Ekwurzel, B., P. Schlosser, W. M. Smethie Jr., L. N. Plummer, E. Busenberg, R. L. Michel, R. Weppernig, and M. Stute (1994), Dating of shallow groundwater: Comparison of the transient tracers 3H/3He, chlorofluorocarbons, and 85Kr, *Water Reso* , *30*, 1693–1708.

Feyen, L., and S. M. Gorelick (2005), Framework to evaluate the worth of hydraulic conductivity data for optimal groundwater resources management in ecologically sensitive areas, *Water Resour. Res.*, *41*, W03019, doi:10.1029/2003WR002901.

Feyen, L., J. J. Gomez-Hernandez, P. J. Ribeiro Jr., K. J. Beven, and F. De Smedt (2003), A Bayesian approach to stochastic capture zone delineation incorporating arrival times, conductivity measurements, and hydraulic head observations, *Water Resour. Res.*, *39*(5), 1126, doi:10.1029/2002WR001544.

Harvey, C. F., and S. M. Gorelick (1995), Mapping hydraulic conductivity: Sequential conditioning with measurements of solute arrival time, hydraulic head, and local conductivity, *Water Resour. Res.*, *31*, 1615–1626.

Herrera, G. S., J. Guarnaccia, and G. F. Pinder (2000), A methodology for the design of space-time groundwater quality sampling networks, in *Proceedings of Computational Methods in Water Resources XIII*, Vol. *1*, edited by L. R. Bentley, J. F. Sykes, C. A. Brebbia, W. G. Gray, and G. F. Pinder, pp. 579–585, A. A. Balkema, Rotterdam.

Hughes, J. P., and D. P. Lettenmaier (1981), Data requirements for Kriging: Estimation and network design, *Water Resour. Res.*, *17*, 1641–1650.

Izbicki, J. A., C. L. Stamos, and P. Martin (2004), Comparison of groundwater flow model particle tracking results and isotopic data in the Mojave River ground-water basin, southern California, USA, *J. Hydrol.*, *292*, 30–47.

James, B. R., and S. M. Gorelick (1994), When enough is enough: The worth of monitoring data in remediation design, *Water Resour. Res.*, *30*, 3499–3513.

Knopman, D. S., and C. I. Voss (1989), Multiobjective sampling design for parameter estimation and model discrimination in groundwater solute transport, *Water Resour. Res.*, *25*, 2245–2258.

Knopman, D. S., C. I. Voss, and S. P. Garabedian (1991), Sampling design for groundwater solute transport: Tests of methods and analysis of Cape Cod tracer test data, *Water Resour. Res.*, *27*, 925–949.

Kovar, K., A. Leijnse, G. J. M. Uffink, M. J. H. Pastoors, J. H. C. Mülschlegel, and W. J. Zaadnoordijk (2005), Reliability of travel times to groundwater abstraction wells: Application of the Netherlands Groundwater Model, LGM, Report 703717013/2005, RIVM, Bilthoven.

Loaiciga, H. O. (1989), An optimization approach for groundwater quality monitoring network design, *Water Resour. Res.*, *25*, 1771–1782.

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

Manning, A. H., D. K. Solomon, and S. A. Thiros (2005), $^3H/^3He$ Age data in assessing the susceptibility of wells to contamination, *Ground Water*, *43*, 353–367.

Massman, J., and R. A. Freeze (1987a), Groundwater contamination from waste management sites: The interaction between risk-based engineering design and regulatory policy. 1. Methodology, *Water Resour. Res.*, *23*, 351–367.

Massman, J., and R. A. Freeze (1987b), Groundwater contamination from waste management sites: The interaction between risk-based engineering design and regulatory policy. 2. Results, *Water Resour. Res.*, *23*, 368–380.

McDonald, M. G., and A. W. Harbaugh (1984), *A modular three-dimensional finite-difference ground-water flow model*, USGS, Reston.

McKinney, D. C., and D. P. Loucks (1992), Network design for predicting groundwater contamination, *Water Resour. Res.*, *28*, 133–147.

Meyer, P. D., and E. D. Brill (1988), A method for locating wells in a groundwater monitoring network under conditions of uncertainty, *Water Resour. Res.*, *24*, 1277–1282.

Meyer, P. D., A. J. Valocchi, and J. W. Eheart (1994), Monitoring network design to provide initial detection of groundwater contamination, *Water Resour. Res.*, *30*, 2647–2659.

Nunes, L. M., E. Paralta, M. C. Cunha, and L. Ribeiro (2004), Groundwater nitrate monitoring network optimization with missing data, *Water Resour. Res.*, *40*, W02406, doi:10.1029/2003WR002469.

Pardo-Iguzquiza, E. (1998), Optimal selection of number and location of rainfall gauges for areal rainfall estimation using geostatistics and simulated annealing, *J. Hydrol.*, *210*, 206–220.

Pollock, D. W. (1994), User's Guide for MODPATH/MODPATH-PLOT, Version 3: A particle-tracking post-processing package for MODFLOW, the U.S. Geological Survey finite-difference ground-water flow model, Open-File Report, 94-464, U.S. Geological Survey, Reston.

Portniaguine, O., and D. K. Solomon (1998), Parameter estimation using groundwater age and head data, Cape Cod, Massachusetts, *Water Resour. Res.*, *34*, 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*, 421–433.

Sheets, R. A., E. S. Bair, and G. L. Rowe (1998), Use of $^{3}H/^{3}He$ ages to evaluate and improve groundwater flow models in a complex buried-valley aquifer, *Water Resour. Res.*, *34*, 1077–1089.

Simmons, C. S. (1982), A stochastic transport representation of dispersion in one-dimensional porous media systems, *Water Resour. Res.*, *18*, 1193–1214.

Smethie, W. M., Jr., D. K. Solomon, S. L. Schiff, and G. G. Mathieu (1992), Tracing groundwater flow in the Borden aquifer using krypton-85, *J. Hydrol.*, *130*, 279–297.

Solomon, D. K., R. J. Poreda, S. L. Schiff, and J. A. Cherry (1992), Tritium and Helium3 as groundwater age tracers in the Borden aquifer, *Water Resour. Res.*, *28*, 741–755.

Storck, P., J. W. Eheart, and A. J. Valocchi (1997), A method for the optimal location of monitoring wells for detection of groundwater contamination in three-dimensional heterogeneous aquifers, *Water Resour. Res.*, *33*, 2081–2088.

Stute, M., and P. Schlosser (2000), Tritium/^{3}Helium measurements as calibration tools in groundwater transport modelling, in *Tracers and Modelling in Hydrogeology*, edited by A. Dassargues, pp. 33–38, IAHS, Oxford.

Sumner, N. R., P. M. Fleming, and B. C. Bates (1997), Calibration of a modified SFB model for twenty-five Australian catchments using simulated annealing, *J. Hydrol.*, *197*, 166–188.

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 plane, *Water Resour. Res.*, *32*, 1023–1038.

Tiedeman, C. R., M. C. Hill, F. A. D'Agnese, and C. C. Faunt (2003), Methods for using groundwater model predictions to guide hydrogeologic data collection, with application to the Death Valley regional groundwater flow system, *Water Resour. Res.*, *39*(1), 1010, doi:10.1029/2001WR001255.

Tiedeman, C. R., D. M. Ely, M. C. Hill, and G. M. O'Brien (2004), A method for evaluating the importance of system state observations to model predictions, with application to the Death Valley regional groundwater flow system, *Water Resour. Res.*, *40*, W12411, doi:10.1029/2004WR003313.

Valstar, J. R. (2001), Inverse modeling of groundwater flow and transport, Ph.D.Thesis, Delft Univ. of Technology, Delft.

Valstar, J. R., D. B. McLaughlin, C. B. M. te Stroet, and F. C. van Geer (2004), A representer-based inverse method for groundwater flow and transport applications, *Water Resour. Res.*, *40*, W05116, doi:10.1029/2003WR002922.

Wagner, B. J. (1995), Sampling design methods for groundwater modeling under uncertainty, *Water Resour. Res.*, *31*, 2581–2591.

Wagner, B. J. (1999), Evaluating data worth for ground-water management under uncertainty, *J. Water Resour. Plann. Manage.*, *125*, 281–288.

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

Woodbury, A. D., and Y. Rubin (2000), A full-Bayesian approach to parameter inference from tracer travel moments and investigation of scale effects at the Cape Cod experimental site, *Water Resour. Res.*, *36*, 159–171.

Wu, J., C. Zheng, and C. C. Chien (2005), Cost-effective sampling network design for contaminant plume monitoring under general hydrogeological conditions, *J. Contam. Hydrol.*, *77*, 41–65.

Zhang, D. (2002), *Stochastic methods for flow in porous media: coping with uncertainties*, Academic Press, San Diego.

G. M. C. M. Janssen, Department of Soil Quality, Wageningen University, Wageningen, The Netherlands. (gijs.janssen@tno.nl)

J. R. Valstar, TNO Built Environment and Geosciences – National Geological Survey, Utrecht, The Netherlands.

S. E. A. T. M. van der Zee, Department of Soil Physics, Ecohydrology, and Groundwater Management, Wageningen University, Wageningen, The Netherlands.