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 Zee3

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

Traveltine 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.


1. Introduction

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].

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, traveltine determinations have found increasing application as another type of information with which flow models can be constrained. By traveltine determinations we mean both groundwater age estimations and tracer arrival time measurements. Groundwater ages can be derived from 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 3H/3He [e.g., Smethie et al., 1992], 36Kr [e.g., Solomon et al., 1992], chlorofluorocarbons [e.g., Dunkle et al., 1993], or a com-
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; State and Schlösser, 2000].

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], Porninaguine 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.

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.

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.

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

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

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 (y2–y3) that protects Aquifer 2 (y3–y4) from a conservative contaminant released out of a source zone (x7–x8) at the top of Aquifer 1 (y1–y2). The control plane at which breakthrough is evaluated is located at y3 along the bottom of the area of interest (x6–x8, y2–y3).

All parameters and dimensions in the example problem are assumed constant and known, except for the hydraulic conductivity (K) in the shaded center part (x5–x10, y2–y3) 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.

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 right 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).

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

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

$$p(\tau_{BT}) = p(t_0 + \tau) = \int_{x^i=0}^{x^f=B} p(\tau; x^{CP})p(x^{CP} \in \chi)dx^{CP}$$  \hspace{1cm} (1)$$

or in discretized form:

$$p(\tau_{BT}) \approx \frac{1}{2} \left[ p(\tau; x^{CP})p(x^{CP} \in \chi) + p(\tau; x^{CP} + \Delta x^{CP})p(x^{CP} + \Delta x^{CP} \in \chi) \right]$$  \hspace{1cm} (2)$$

[14] In equations (1)–(2), $\tau_{BT}$ is the contaminant breakthrough time evaluated at the control plane, $t_0$ is the starting time of the contaminant's arrival, $\tau$ 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 $\chi$ ($\chi$ is the collection of all breakthrough locations $\chi$, and $A$ and $B$ are the spatial $x$ limits of the control plane ($x_6$ and $x_5$ in Figure 1). $\Delta x^{CP}$ is the discretization interval along the control plane.

[15] Assuming that $\tau$ and $\chi$ are Gaussian distributed, we have

$$p(\tau; x^{CP}) \approx N\left(\tau; \sigma_{\chi(cw)}, \mu_{\chi(cw)}\right)$$  \hspace{1cm} (3)$$

$$p(x^{CP} \in \chi) \approx \sum_{i=0}^{N} \left[N\left(x^{CP}; \sigma_{\chi(w+i\Delta w)}, \mu_{\chi(w+i\Delta w)}\right) + N\left(x^{CP}; \sigma_{\chi(w+i\Delta w)}, \mu_{\chi(w+i\Delta w)}\right)\right] \Delta x^{SZ},$$  \hspace{1cm} (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 $x^{CP}$, $\sigma_{\chi(cw)}^2$ is the arrival time (or age) variance at $x^{CP}$, and $\mu_{\chi(cw)}$ is the mean arrival time (or age) at $x^{CP}$. $\chi(x^{SZ})$ is the breakthrough location of a particle that originated from $x^{SZ}$, respectively, $\Delta x^{SZ}$ 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 $\sigma_{\tau_{BT}}^2$ can be calculated as:

$$\sigma_{\tau_{BT}}^2 = \int_{\tau_{BT}=0}^{\infty} (\tau_{BT} - E(\tau_{BT}))^2 p(\tau_{BT})d\tau$$

$$\sigma_{\tau_{BT}}^2 = \int_{\tau_{BT}=0}^{\infty} \int_{\tau_{BT}=0}^{\infty} p(\tau_{BT})d\tau dt$$

$$= \int_{\tau_{BT}=0}^{\infty} \int_{\tau_{BT}=0}^{\infty} \tau_{BT} p(\tau_{BT})d\tau dt$$

$$= \int_{\tau_{BT}=0}^{\infty} \int_{\tau_{BT}=0}^{\infty} p(\tau_{BT})d\tau dt,$$  \hspace{1cm} (5)$$

with $E(\tau_{BT}) = \int_{\tau_{BT}=0}^{\infty} \tau_{BT} p(\tau_{BT})d\tau dt.$

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.
(Sub)optimal network selected

[57x753]W02405

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 representor-based inverse method [Valstar et al., 2004] provides an efficient way to calculate these influences. The representors 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 representors 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 \( \chi(x^{W}) \). 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 \( \sigma_{\tau_{\alpha}}^{2} \) (equation (5)) of the contaminant for every \( y_{i} \) during travel step \( a \). A Genetic Algorithm was used to efficiently search for a near-optimal design that minimizes \( \sigma_{\tau_{\alpha}}^{2} \). 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 \( \chi(x^{W}) \) 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 representor-based inverse algorithm, and derive \( \text{Ln(traveltime)} \) representers accordingly.

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

\[
\phi_{1} = A_{bg}(\alpha)h_{g} - q_{f} = 0
\]  

\[
\phi_{2}^{c} = 0 = t_{i}^{c} - t_{i}^{c-1} - \Delta t_{i}^{c}(\alpha, h, x_{i}^{c}, y_{i}^{c}, y_{i}^{c-1})
\]  

\[
\phi_{5}^{c} = 0 = \left\{ \begin{array}{ll}
\Delta x_{i}^{c}(\alpha, h, x_{i}^{c}, y_{i}^{c-1}) & \text{for } x_{i}^{c} \neq x_{\text{cell boundary}} \\
\Delta y_{i}^{c}(\alpha, h, x_{i}^{c}, y_{i}^{c-1}) & \text{for } y_{i}^{c} = y_{\text{cell boundary}}
\end{array} \right.
\]  

\[
\phi_{6}^{c} = 0 = \left\{ \begin{array}{ll}
\Delta y_{i}^{c}(\alpha, h, x_{i}^{c}, y_{i}^{c-1}) & \text{for } y_{i}^{c} \neq y_{\text{cell boundary}} \\
\Delta x_{i}^{c}(\alpha, h, x_{i}^{c}, y_{i}^{c-1}) & \text{for } x_{i}^{c} = x_{\text{cell boundary}}
\end{array} \right.
\]

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 \( \alpha (=\text{Ln}(K)) \), \( t_{i} \) is the traveltime of particle \( i \), \( \Delta t_{i}^{c} \) 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 \( \Delta x_{i}^{c} \) and \( \Delta y_{i}^{c} \) 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, \( \Delta t_{i}^{c} \) and \( \Delta y_{i}^{c} \) depend on \( x_{i}^{c} \) and \( y_{i}^{c-1} \) if travel step \( s \) is limited by reaching an \( x \)-boundary of a cell. Furthermore, \( \Delta x_{i}^{c} \) and \( \Delta y_{i}^{c} \) depend on \( y_{i}^{c} \) and \( y_{i}^{c-1} \) 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

<table>
<thead>
<tr>
<th>Index</th>
<th>Ranges From 1 to the Number of</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>Tracked particles, which equals the number of potential age determinations plus the number of pseudo breakthrough time and pseudo breakthrough location measurements</td>
</tr>
<tr>
<td>f, g</td>
<td>head state variables</td>
</tr>
<tr>
<td>k, l</td>
<td>uncertain parameters</td>
</tr>
<tr>
<td>q</td>
<td>measurements</td>
</tr>
</tbody>
</table>

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 \( \text{Ln}(\tau) \) 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.

From Appendices C9 and C7 it follows that after the first iteration the \( \text{Ln}(\tau) \) representer (\( \Theta^* \)) for a particle \( i \) is given by:

\[
\Theta^*_i = \frac{1}{\psi_{i\alpha}^N} \Bigg[ \sum_{s=j=1}^{N_i} \frac{\partial \Delta \Theta^*}{\partial \alpha_j} \psi_{ij} + \sum_{s=j=1}^{N_i} \frac{\partial \Delta \Theta^*}{\partial q_{ij}} \frac{\partial q_{ij}}{\partial s} + \sum_{s=j=1}^{N_i} \frac{\partial \Delta \Theta^*}{\partial \Psi_{ij}} \frac{\partial \Psi_{ij}}{\partial s} + \sum_{s=j=1}^{N_i} \frac{\partial \Delta \Theta^*}{\partial x_{ij}} \frac{\partial x_{ij}}{\partial s} + \sum_{s=j=1}^{N_i} \frac{\partial \Delta \Theta^*}{\partial y_{ij}} \frac{\partial y_{ij}}{\partial s} \Bigg],
\]

(12)

where \( \text{with terminal conditions } X_{N+1} = 0 \) and \( Y_{N+1} = 0 \), and the breakthrough location representer is given by:

\[
X_{iN} = \sum_{s=j=1}^{N_i} \frac{\partial \Delta x_i}{\partial \alpha_j} \psi_{ij} + \sum_{s=j=1}^{N_i} \frac{\partial \Delta x_i}{\partial q_{ij}} \frac{\partial q_{ij}}{\partial s} + \sum_{s=j=1}^{N_i} \frac{\partial \Delta x_i}{\partial \Psi_{ij}} \frac{\partial \Psi_{ij}}{\partial s} + \sum_{s=j=1}^{N_i} \frac{\partial \Delta x_i}{\partial x_{ij}} \frac{\partial x_{ij}}{\partial s} + \sum_{s=j=1}^{N_i} \frac{\partial \Delta x_i}{\partial y_{ij}} \frac{\partial y_{ij}}{\partial s},
\]

(13)

with boundary conditions \( Y_0^0 = 0 \). In equations (12) and (13), \( \Psi \) is the parameter representer (see Appendices B and C5), \( \Xi \) is the head representer (see Appendix B and C6), and \( Y_{N+1}^i \) 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

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

\[
P_{\text{posterior}}^{\text{Ln}(\tau_i)} = P_{\text{prior}}^{\text{Ln}(\tau_i)} - \Theta_{\Psi}^{*N_i} \left( (M(\Xi, \Psi, \Theta^*, X) + P_v)^{-1} \right)_{pq} \Theta_{\Psi}^{*N_i}
\]

(14)

\[
P_{\text{posterior}}^{\chi_i} = P_{\text{prior}}^{\chi_i} - X_{\chi_i}^{N_i} \left( (M(\Xi, \Psi, \Theta^*, X) + P_v)^{-1} \right)_{pq} X_{\chi_i}^{N_i},
\]

(15)

where \( M(\Xi, \Psi, \Theta^*, X)_{pq} \) is a \( p \times q \) representer matrix that contains the prior cross covariances between all observations, \( P_{\text{prior}}^{\text{Ln}(\tau)} \) is calculated by defining a pseudo measurement of \( \tau_i \) and calculating the \( \text{Ln}(\tau) \) representer for this measurement, and \( P_{\text{prior}}^{\chi_i} \) is calculated by defining a pseudo measurement for \( \chi_i \) and calculating the breakthrough location representer for this measurement.
and determinations, the estimation error \( \sigma^2_{\tau(\gamma)} \) and \( \sigma^2_{\chi(\gamma)} \) in equations (3) and (4), respectively (note that \( \tau \) in equations (1)–(3) is replaced by its natural logarithm). \( \mu_{\tau(\gamma)} \) and \( \mu_{\chi(\gamma)} \) in equations (3) and (4) are approximated by their first guess estimates, \( \ln(\tau(p(x_{CP})) \) and \( \chi(x_{SZ}) \) (see also Appendix B), obtained by running a simulation with \( \alpha = \pi \). Now all necessary information is available to compute \( p(\tau_{gP}) \) according to equation (2) and subsequently the variance of the breakthrough time \( \sigma^2_{\tau_{br}} \) 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, \( P_{\text{Laplace}} \) and \( P_{\text{Posterior}} \) 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 \( x_3 \) 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 \( \ln(K_C) \) of \(-3.0\) Ln(m/d), a variance \( \sigma^2_T \) 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 \( K_C \) 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 \( x_{12} \) 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 m² 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 (\( \tau(t) \)) at the sampled location in the prior realization of the confining layer, in which all stochastic parameters are set at their prior means (\( \alpha = \ln(K) = \pi = -3.0 \) Ln(m/d)). This percentage varies between different examples.

At the control plane, in every column one pseudo age determination was defined to enable the calculation of the covariances between the goal variable \( \ln(t(x_{CP})) \) and the measurements in a design. The sensitivity terms in equation (12) (the \( \ln(\tau) \) 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].

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 (\( \chi \)) 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.

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.

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.

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

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
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 \( \sigma_Y^2 \) 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 \( \ln(t) \) histogram with the Gaussian distribution will deteriorate as \( \sigma_Y^2 \) gets larger and the untransformed traveltimes become more skewed. However, for the range of \( \sigma_Y^2 \) 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 \( \sigma_Y^2 \). 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 \( \ln(t) \) 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 \( \ln(t) \) as \( \sigma_Y^2 \) 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 \( \chi \).

**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). \( \sigma_Y^2 = 2.0 \).

**Figure 4.** Linearized \( \ln(t) \) (a), \( \chi \) (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 \( \sigma_Y^2 \). 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 which the variances were evaluated.
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 \( Y^2 \) 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.

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

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 \( \chi_Q \) 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).

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 representor of this measurement throughout the domain. The representor 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 \( \ln(t) \) representor at location \( Q \).

Figures 5b, 5d and 5e were produced in a similar manner. For Figures 5b and 5e, however, the prior \( s_h^2 \) 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.

Figure 5c was constructed by defining in every grid cell a pseudo traveltime measurement and calculating their \( \ln(t) \) representors for \( \ln(t) \) at the location of the pseudo measurement itself (yielding the prior \( s_{\ln(t)}^2 \) 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 representor was calculated for the particle that a priori is expected to break through in \( Q \).

From Figure 5 it appears that \( \ln(t_Q) \) is particularly strongly correlated (correlation coefficient \( r > 0.6 \)) 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 \( r = 1 \)), 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...
pattern was found by Harvey and Gorelick [1995] (their Figure 6b).

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.

The curvature of the positive \( \ln(t) - \ln(t_0) \) 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 time 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 \( Q \).

Whereas \( \ln(t_0) \) particularly shows strong correlations with \( \ln(K) \) and \( \ln(t) \), \( \chi_Q \) 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), \( \chi_Q \) 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.

\[ \chi_Q - \ln(K) \] correlations (Figure 5d) are relatively weak, due to the limited vertical correlation of \( \ln(K) \) values. Figure 5f reveals that \( \chi_Q \) 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 \( \chi_Q \) 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

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.

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 5e.

For 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 \( \chi \) are expected.

For Figure 6c, the objective was to minimize \( \sigma_w^2 \), 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 \( \chi(x_{CP}) \). 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

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.
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

A methodology was proposed that incorporates traveltime determinations into measurement network design. The methodology was focused on minimizing 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.

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.

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.

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.

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].

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 dating techniques): pore-scale dispersion and macro-dispersion 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; Ekwarzel 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., 3He [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.
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 instationary 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:

\[
\frac{\partial J}{\partial x^0_i} + 2\lambda_i M_i h + \sum_{j=1}^{N_p} \sum_{s=1}^{N_s} 2\lambda^s_i \left( \phi^s_i \right) \frac{\partial}{\partial x^0_i} + \sum_{j=1}^{N_p} \sum_{s=1}^{N_s} 2\lambda^s_i \left( \phi^s_i \right) \frac{\partial}{\partial x^0_i} = 0 \quad \text{(A1)}
\]

The derivatives in the third, fourth and the fifth term should only be taken for those travel steps \( s \) where \( x^0_i \) 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:

\[
- \frac{\partial M_i(h, \alpha, \ln(t), \chi)}{\partial x^0_i} \left[ p_{v \to m}^{s-1} [z_a - M_i(h, \alpha, \ln(t), \chi)] \right] - \lambda^0_{i} \Delta x^0_i (\alpha, h, x^0_i, x^{s-1}_i, y^0_i, y^{s-1}_i) \\
- \lambda^{x+1}_{i} \frac{\partial M^{x+1}_i(\alpha, h, x^{x+1}_i, y^{x+1}_i)}{\partial x^0_i} \\
+ \left( \phi^0_{i} - \phi^{x+1}_i \right) - \lambda^{x+1}_{i} \frac{\partial \phi^{x+1}_i(\alpha, h, x^{x+1}_i, y^{x+1}_i)}{\partial x^0_i} \\
- \lambda^{x+1}_{i} \frac{\partial \phi^{x+1}_i(\alpha, h, x^{x+1}_i, y^{x+1}_i)}{\partial x^0_i} \quad \text{for } x^0_i \neq x_{\text{cell boundary}} \quad \text{(A2)}
\]

and

\[
- \frac{\partial M_i(h, \alpha, \ln(t), \chi)}{\partial x^0_i} \left[ p_{v \to m}^{s-1} [z_a - M_i(h, \alpha, \ln(t), \chi)] \right] + \lambda^0_{i} = 0 \quad \text{for } x^0_i = x_{\text{cell boundary}} \quad \text{(A3)}
\]

Note that the second and the fifth term of equation (A2) are always zero because \( \Delta x^0_i \) and \( \Delta y^0_i \) 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:

\[
\phi^0_{i} = \phi^{x+1}_i + \frac{\partial M_i(h, \alpha, \ln(t), \chi)}{\partial x^0_i} \left[ p_{v \to m}^{s-1} [z_a - M_i(h, \alpha, \ln(t), \chi)] \right] \\
+ \lambda^{x+1}_{i} \frac{\partial \Delta x^0_i (\alpha, h, x^{x+1}_i, y^{x+1}_i)}{\partial x^0_i} \\
+ \lambda^{x+1}_{i} \frac{\partial \Delta y^0_i (\alpha, h, x^{x+1}_i, y^{x+1}_i)}{\partial x^0_i} \quad \text{for } x^0_i \neq x_{\text{cell boundary}} \quad \text{(A4)}
\]
with boundary condition $\lambda^N_{y^{-1}} = 0$.

### A2. Y Location Adjoint Equation

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

$$\chi_i^k = \lambda^{N_y}_{y^{-1}} + \frac{\partial M_y(h, \alpha, \text{Ln}(t), \chi)}{\partial y_i} \left[ P_{y^{-1}} \right]_{pm} \cdot [z_n - M_y(h, \alpha, \text{Ln}(t), \chi)]$$

$$+ \lambda^{N_y}_{y^{-1}} \frac{\partial \Delta y_i^{t+1}}{\partial y_i} \left( \alpha, h, x_i^{t+1}, y_i^{t+1}, y_i^{t+1} \right)$$

$$+ \lambda^{N_y}_{y^{-1}} \frac{\partial \Delta y_i^{t+1}}{\partial y_i} \left( \alpha, h, x_i^{t+1}, y_i^{t+1}, y_i^{t+1} \right) \text{ for } y_i \neq y_{cell boundary}$$

$$\chi_i^k = \frac{\partial M_y(h, \alpha, \text{Ln}(t), \chi)}{\partial y_i} \left[ P_{y^{-1}} \right]_{pm} \cdot [z_n - M_y(h, \alpha, \text{Ln}(t), \chi)] = 0 \text{ for } y_i = y_{cell boundary}$$

with boundary condition $\lambda^N_{y^{-1}} = 0$.

### A3. Traveltime Adjoint Equation

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:

$$\frac{\partial J}{\partial h} + \frac{2\lambda_A(h, \alpha, \text{Ln}(t), \chi)}{\partial \phi_h}\left[ P_{x^{-1}} \right]_{p_m} \cdot [z_n - M_y(h, \alpha, \text{Ln}(t), \chi)]$$

$$+ \lambda^{N_y}_{y^{-1}} \frac{\partial \Delta x_i^{t+1}}{\partial \phi_h} \left( \alpha, h, x_i^{t+1}, y_i^{t+1}, y_i^{t+1} \right) = 0$$

(A8)

[72] 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:

$$\frac{\partial J}{\partial \phi_h} + \frac{2\lambda_A(h, \alpha, \text{Ln}(t), \chi)}{\partial \phi_h}\left[ P_{x^{-1}} \right]_{p_m}$$

$$+ \sum_{i \in x} \frac{\partial \Delta x_i^{t+1}}{\partial \phi_h} \left( \alpha, h, x_i^{t+1}, y_i^{t+1}, y_i^{t+1} \right) = 0$$

(A9)

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:

$$\chi_i^k = \chi_i^{k+1} + \frac{\partial M_y(h, \alpha, \text{Ln}(t), \chi)}{\partial \phi_h} \left[ P_{x^{-1}} \right]_{p_m} [z_n - M_y(h, \alpha, \text{Ln}(t), \chi)]$$

(A10)

with boundary condition $\chi_i^{N_y} = 0$.

### A4. Head Adjoint Equation

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:

$$\frac{\partial J}{\partial h} + \frac{2\lambda_A(h, \alpha, \text{Ln}(t), \chi)}{\partial \phi_h}\left[ P_{x^{-1}} \right]_{p_m}$$

$$+ \sum_{i \in x} \frac{\partial \Delta x_i^{t+1}}{\partial \phi_h} \left( \alpha, h, x_i^{t+1}, y_i^{t+1}, y_i^{t+1} \right) = 0$$

(A11)

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

$$A_{\phi_h}(\alpha) \lambda_h = \frac{\partial M_y(h, \alpha, \text{Ln}(t), \chi)}{\partial \phi_h}\left[ P_{x^{-1}} \right]_{p_m}$$

$$+ \sum_{i \in x} \lambda_i^k \frac{\partial \Delta x_i^{t+1}}{\partial \phi_h} \left( \alpha, h, x_i^{t+1}, y_i^{t+1}, y_i^{t+1} \right)$$

(A12)

### A5. Parameter Equation

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:

$$\frac{\partial J}{\partial \alpha} + \frac{2\lambda_A(h, \alpha, \text{Ln}(t), \chi)}{\partial \alpha}\left[ P_{x^{-1}} \right]_{p_m}$$

$$+ \sum_{i \in x} \frac{\partial \Delta x_i^{t+1}}{\partial \alpha} \left( \alpha, h, x_i^{t+1}, y_i^{t+1}, y_i^{t+1} \right) = 0$$

(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:

\[ \lambda_i = \sum_{p=1}^{N_p} \lambda_{ip} b_p \]  
(B1)

\[ \lambda_j = \sum_{p=1}^{N_p} \Pi_{jp} b_p \]  
(B2)

\[ \lambda_k = \sum_{p=1}^{N_p} H_{kp} b_p \]  
(B3)

\[ \lambda_l = \sum_{p=1}^{N_p} \Gamma_{lp} b_p \]  
(B4)

representers of measurement \( p_i \), calculated for particle \( i \), \( \Gamma_{ip} \) is the head adjoint representer of measurement \( p_i \), calculated for the head state variable \( f \), \( \Psi_{ip} \) is the parameter representer of measurement \( p_i \), calculated for parameter \( l \), \( \Xi_{ip} \) is the head representer of measurement \( p_i \), calculated for head state variable \( g \), \( X_{ip} \) and \( Y_{ip} \) are respectively the \( x \) and \( y \) location representer of measurement \( p_i \), calculated for particle \( i \). \( \eta \) is the iteration number and \( \Theta_{ip}^{\eta} \) is the \( \eta \)th iteration of measurement \( p_i \), calculated for particle \( i \). In order to be able to expand the untransformed traveltime \( t_i \), \( \Theta_{ip}^{\eta} \) is multiplied by the derivative \( dt_i/dn \). In the previous iteration \( t_i^{\eta-1} \). \( x_i \), \( y_i \), \( x_{ip} \), \( y_{ip} \) and \( t_{ip} \) are the solutions obtained by solving equations (6)–(9) with \( \alpha = \pi \). \( h_{ip} \), \( x_{ip} \), \( x_{ip} \) and \( t_{ip} \) 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 \( h_{ip} \), \( x_{ip} \), \( x_{ip} \) and \( t_{ip} \) do not have to be calculated. The algorithm is initiated using \( \alpha = \pi \), \( h = h_i \), \( t = t_i \), \( x_i = x_{ip} \), \( y_i = y_{ip} \) and \( \lambda_{ip} = \lambda_{ip} = \lambda_{ip} = 0 \).

Appendix C: Representer Derivations

[76] For the sake of readability, the explicit statement of the dependencies of \( A \), \( \Delta t \), \( \Delta x \) and \( \Delta y \) are omitted from now on.

C1. X Location Adjacent Representers

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

\[ \Pi_{ip} b_p = \Pi_{ip}^{\eta+1} b_p + \frac{\partial M_p(h, \alpha, \Delta t_{ip}(t), \chi)}{\partial x_i} [P_{ip}^{-1}]_{pa} (z_a - M_a( )) \]
 \[ + \lambda_{ip}^{\eta+1} b_p \frac{\partial \Delta t_{ip}^{\eta+1}}{\partial x_i} + H_{ip}^{\eta+1} b_p \frac{\partial \Delta y_{ip}^{\eta+1}}{\partial x_i} \quad \text{for} \quad x'_i \neq \text{cell boundary} \]
(C1)

\[ \Pi_{ip} b_p = \frac{\partial M_p(h, \alpha, \Delta t_{ip}(t), \chi)}{\partial x_i} [P_{ip}^{-1}]_{pa} (z_a - M_a( )) \quad \text{for} \quad x'_i = \text{cell boundary} \]
(C2)

where \( M_a( ) = M_a(h_F + h_{ip} + \Xi h, \chi_i + \Psi b, \Delta t_{ip} + t_{ip} + \Theta^* b, x_F + \chi_{ip} + \Theta b) \). The coefficients \( b \) for all \( p \) measurements are now chosen as:

\[ b_p = [P_{ip}^{-1}]_{pa} (z_a - M_a( )) \]
(C3)

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

\[ \Pi_{ip}^{\eta+1} b_p + \frac{\partial M_p(h, \alpha, \Delta t_{ip}(t), \chi)}{\partial x_i} \]
 \[ + \lambda_{ip}^{\eta+1} \frac{\partial \Delta t_{ip}^{\eta+1}}{\partial x_i} + H_{ip}^{\eta+1} \frac{\partial \Delta y_{ip}^{\eta+1}}{\partial x_i} \quad \text{for} \quad x'_i \neq \text{cell boundary} \]
(C4)

\[ \Pi_{ip} b_p = \frac{\partial M_p(h, \alpha, \Delta t_{ip}(t), \chi)}{\partial x_i} = 0 \quad \text{for} \quad x'_i = \text{cell boundary} \]
(C5)
with boundary condition $\Pi_{\alpha}^{N+1} = 0$. 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 $\chi_i$ (by definition at $s = N_\alpha$). In that case $\frac{\partial M_{p}(\chi)}{\partial t_i} = 1$.

### C2. $Y$ Location Adjoint Representers

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

$$
H_y^{(i)} = \frac{\partial M_p(h(\alpha, \ln(t)), \chi)}{\partial y_i^{(i)}}
$$

$$
\quad + \Pi_{\alpha}^{(i)} \frac{\partial \Delta y_i^{(i+1)}}{\partial y_i^{(i)}} + \Pi_{\alpha}^{(i)} \frac{\partial \Delta y_i^{(i+1)}}{\partial y_i^{(i)}}
$$

for $y_i^{(i)} \neq y_{\text{cell boundary}}$

(C6)

$$
H_y^{(i)} = \frac{\partial M_p(h(\alpha, \ln(t)), \chi)}{\partial y_i^{(i)}} = 0
$$

for $y_i^{(i)} = y_{\text{cell boundary}}$

(C7)

with boundary condition $H_y^{N+1} = 0$. 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:

$$
H_y^{(i)} = H_y^{(i+1)} - \Pi_{\alpha}^{(i+1)} \frac{\partial \Delta y_i^{(i+1)}}{\partial y_i^{(i)}} - \Pi_{\alpha}^{(i)} \frac{\partial \Delta y_i^{(i+1)}}{\partial y_i^{(i)}}
$$

for $y_i^{(i)} \neq y_{\text{cell boundary}}$

(C8)

$$
H_y^{(i)} = 0
$$

for $y_i^{(i)} = y_{\text{cell boundary}}$

(C9)

with boundary condition $H_y^{N+1} = 0$.

### C3. Traveltime Adjoint Representers

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

$$
\Lambda_y^{(i)} = \Lambda_y^{(i+1)} + \frac{\partial M_p(h(\alpha, \ln(t)), \chi)}{\partial t_i^{(i)}}[P_{\alpha}^{-1}]_{pp}(z_{\alpha} - M_{\alpha}( ))
$$

(C10)

After inserting (C3) into (C10), the resulting equation can only be fulfilled for nonzero $b$ if:

$$
\Lambda_y^{(i)} = \Lambda_y^{(i+1)} + \frac{\partial M_p(h(\alpha, \ln(t)), \chi)}{\partial t_i^{(i)}}
$$

with boundary condition $\Lambda_y^{N+1} = 0$. The second term of equation (C11) is nonzero only in case of measurements of $\ln(t_i)$. In that case $\frac{\partial M_p(\ln(t_i))}{\partial t_i^{(i)}} = \frac{1}{t_i^{(i)}}$ with $t_i^{(i)}$ approximated by its most recent estimate.

### C4. Head Adjoint Representers

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

$$
A_{\alpha} \Gamma_{\alpha} b_p = \frac{\partial M_p(h(\alpha, \ln(t)), \chi)}{\partial t_i^{(i)}}[P_{\alpha}^{-1}]_{pp}(z_{\alpha} - M_{\alpha}( )
$$

$$
+ \sum_{i} \sum_{s} \Lambda_y^{(i)} \frac{\partial \Delta y_i^{(i)}}{\partial t_i^{(i)}} + \sum_{i} \sum_{s} \Pi_{\alpha}^{(i)} \frac{\partial \Delta y_i^{(i)}}{\partial t_i^{(i)}}
$$

(C12)

$$
A_{\alpha} \Gamma_{\alpha} b_p = \frac{\partial M_p(h(\alpha, \ln(t)), \chi)}{\partial t_i^{(i)}}[P_{\alpha}^{-1}]_{pp}(z_{\alpha} - M_{\alpha}( )
$$

$$
+ \sum_{i} \sum_{s} \Lambda_y^{(i)} \frac{\partial \Delta y_i^{(i)}}{\partial t_i^{(i)}} + \sum_{i} \sum_{s} \Pi_{\alpha}^{(i)} \frac{\partial \Delta y_i^{(i)}}{\partial t_i^{(i)}}
$$

(C13)

### C5. Parameter Representers

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

$$
\bar{\alpha}_t + \Psi_y b_p = \bar{\alpha}_t - P_{\alpha} \left[ \frac{\partial A_{\alpha}}{\partial \alpha_k} \Gamma_{\alpha} b_p - \sum_{i} \sum_{s} \Lambda_y^{(i)} \frac{\partial \Delta y_i^{(i)}}{\partial \alpha_k} + \sum_{i} \sum_{s} \Pi_{\alpha}^{(i)} \frac{\partial \Delta y_i^{(i)}}{\partial \alpha_k} \right]
$$

(C14)

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

$$
\Psi_y = -P_{\alpha} \left[ \frac{\partial A_{\alpha}}{\partial \alpha_k} \Gamma_{\alpha} b_p - \sum_{i} \sum_{s} \Lambda_y^{(i)} \frac{\partial \Delta y_i^{(i)}}{\partial \alpha_k} + \sum_{i} \sum_{s} \Pi_{\alpha}^{(i)} \frac{\partial \Delta y_i^{(i)}}{\partial \alpha_k} \right]
$$

(C15)

### C6. Head Representers

[82] 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:

$$
A_{\alpha} \Xi_y = - \frac{\partial A_{\alpha}}{\partial \alpha_k} \Psi_y b_p
$$

(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 $\Psi_y \partial b_p$ (B5), the variation of the heads by $\Xi_y \partial b_p$ (B6), the variation of the $y$ location of particle $i$ by $Y_y \partial b_p$ (B8), and the variation of the traveltime of particle $i$ by $t_i \partial t_i^{(i)} \partial b_p$ (B9), a linear relationship between the $x$
location of particle $i$ on one hand and the parameters, heads, heads, $h_i$, 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

$$X_{ip}^t \partial b_p = \sum_{j=1}^{s} \frac{\partial \Delta x_j}{\partial \xi_i} Y_{ip} \partial b_p + \sum_{j=1}^{s} \frac{\partial \Delta y_j}{\partial \eta_i} x_{ip} \partial b_p + \sum_{j=1}^{s} \frac{\partial \Delta x_j}{\partial \eta_i} Y_{ip} \partial b_p$$

$$+ \sum_{j=1}^{s} \frac{\partial \Delta y_j}{\partial \eta_i} x_{ip} \partial b_p,$$

for $x_i^t \neq x_{cell \ boundary}$ (C17)

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

$$X_{ip}^t = X_{ip}^{t-1} + \frac{\partial \Delta x_j}{\partial \xi_i} Y_{ip} \partial b_p + \frac{\partial \Delta y_j}{\partial \eta_i} x_{ip} \partial b_p$$

$$+ \frac{\partial \Delta x_j}{\partial \eta_i} Y_{ip} \partial b_p + \frac{\partial \Delta y_j}{\partial \eta_i} x_{ip} \partial b_p,$$

for $x_i^t = x_{cell \ boundary}$ (C19)

$$X_{ip}^t = X_{ip}^{t-1} \quad \text{for } x_i^t = x_{cell \ boundary}$$

with boundary conditions $X_{ip}^0 = Y_{ip}^t = 0.$

**C8. Y Location Representers**

The same reasoning holds for the $y$ location representers:

$$Y_{ip}^t = Y_{ip}^{t-1} + \frac{\partial \Delta x_j}{\partial \xi_i} Y_{ip} \partial b_p + \frac{\partial \Delta y_j}{\partial \eta_i} x_{ip} \partial b_p$$

$$+ \frac{\partial \Delta x_j}{\partial \eta_i} Y_{ip} \partial b_p + \frac{\partial \Delta y_j}{\partial \eta_i} x_{ip} \partial b_p,$$

for $y_i^t \neq y_{cell \ boundary}$ (C21)

$$Y_{ip}^t = Y_{ip}^{t-1} \quad \text{for } y_i^t = y_{cell \ boundary}$$

with boundary conditions $Y_{ip}^0 = Y_{ip}^0 = 0.$

**C9. Ln(Traveltime) Representers**

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 $\text{Ln}(t)$, given by $t_{ip}^t \Theta^* \partial b_p$ (see equation (B9)), is

$$t_{ip}^t \Theta^* \partial b_p = \sum_{j=1}^{s} \frac{\partial \Delta t_j}{\partial \xi_i} Y_{ip} \partial b_p + \sum_{j=1}^{s} \frac{\partial \Delta t_j}{\partial \eta_i} X_{ip} \partial b_p$$

$$+ \sum_{j=1}^{s} \frac{\partial \Delta t_j}{\partial \eta_i} Y_{ip} \partial b_p + \sum_{j=1}^{s} \frac{\partial \Delta t_j}{\partial \eta_i} Y_{ip} \partial b_p$$

(C23)

Dividing (C23) by $\partial b_p$, rearranging, and putting it in sequential notation yields the Ln(t) representer:

$$\Theta^*_{ip} = \frac{1}{t_{ip}^t \partial b_p} \left[ \Theta^*_{ip} \partial b_p + \frac{\partial \Delta t_j}{\partial \xi_i} Y_{ip} \partial b_p + \frac{\partial \Delta t_j}{\partial \eta_i} Y_{ip} \partial b_p + \frac{\partial \Delta t_j}{\partial \eta_i} Y_{ip} \partial b_p \right]$$

(C24)

with boundary conditions $\Theta^*_{ip} = Y_{ip}^t = Y_{ip}^0 = 0.$

**C10. Correction Terms**

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]:

$$A_{ib} h_{corr} = q_f + \frac{\partial A_{ib}}{\partial \alpha} (\alpha - \bar{\alpha}) h_g - A_{ib} b_p$$

(C25)

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:

$$t_{corr}^t = t_{corr}^t \Delta t_i + \Theta_{ip}^t b_p - \frac{\partial \Delta t_j}{\partial \xi_i} Y_{ip} \partial b_p$$

(C26)

Since $t_{corr}^t = \Theta_{ip}^t b_p = t_{corr}^t + t_{corr}^t$ (see equation (B9)), we have:

$$t_{corr}^t = t_{corr}^t \Delta t_i + \Theta_{ip}^t b_p - \frac{\partial \Delta t_j}{\partial \eta_i} Y_{ip} \partial b_p$$

(C27)
Following similar reasoning, we have for $x_{corr}$:

$$x'_{corr} = x'_{corr} + x'_{ti} - x'_{pi} + \Delta x'_{i} - \frac{\partial \Delta x'_{i}}{\partial \Psi_{ip}} \Psi_{ip} - \frac{\partial \Delta x'_{ip}}{\partial x'_{i}} x'_{ip} + \frac{\partial \Delta x'_{ip}}{\partial x'_{i}} x'_{ip} p_{ip} - \frac{\partial \Delta x'_{ip}}{\partial x'_{ip}} x'_{ip}^{-1} p_{ip}$$

for $x' \neq x_{cell} \ boundary$ (C28)

And for $y_{corr}$:

$$y'_{corr} = y'_{corr} + y'_{ti} - y'_{pi} + \Delta y'_{i} - \frac{\partial \Delta y'_{i}}{\partial \Psi_{ip}} \Psi_{ip} - \frac{\partial \Delta y'_{ip}}{\partial x'_{i}} X_{ip} - \frac{\partial \Delta y'_{ip}}{\partial x'_{ip}} X_{ip}^{-1} p_{ip} - \frac{\partial \Delta y'_{ip}}{\partial x'_{ip}} X_{ip}^{-1} p_{ip}$$

for $y' \neq y_{cell} \ boundary$ (C29)

During iteration $\eta$, $\Psi_{ip} p_{ip} \equiv_{\Delta} x'_{corr} X_{ip}^{0} X_{ip} X_{ip}^{-1} p_{ip} Y_{ip}^{0} p_{ip}$ in equations (C27)–(C31) are not replaced by the estimations from the previous iteration:

\[
(\alpha_{i}^{0} - \alpha_{i}^{c}), (h_{ip}^{0} - h_{ip}^{c}), (x_{i}^{0} - x_{i}^{c}), (x_{i}^{-1} - x_{i}^{0} - x_{i}^{1} - x_{i}^{c} - x_{i}^{corr}), (y_{i}^{0} - y_{i}^{c} - y_{i}^{corr}), (y_{i}^{-1} - y_{i}^{0} - y_{i}^{1} - y_{i}^{corr}), \]

respectively.

References


---

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


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