



Economic valuation of hydrogeological information when managing groundwater drawdown

Downloaded from: <https://research.chalmers.se>, 2019-11-13 18:44 UTC

Citation for the original published paper (version of record):

Sundell, J., Norberg, T., Haaf, E. et al (2019)

Economic valuation of hydrogeological information when managing groundwater drawdown

Hydrogeology Journal, 27(4): 1111-1130

<http://dx.doi.org/10.1007/s10040-018-1906-z>

N.B. When citing this work, cite the original published paper.



Economic valuation of hydrogeological information when managing groundwater drawdown

Jonas Sundell^{1,2,3} · Tommy Norberg⁴ · Ezra Haaf^{2,5} · Lars Rosén¹

Received: 1 June 2018 / Accepted: 26 November 2018
© The Author(s) 2019

Abstract

A procedure is presented for valuation of information analysis (VOIA) to determine the need for additional information when assessing the effect of several design alternatives to manage future disturbances in hydrogeological systems. When planning for groundwater extraction and drawdown in areas where risks—such as land subsidence, wells running dry and drainage of streams and wetlands—are present, the need for risk-reducing safety measures must be carefully evaluated and managed. The heterogeneity of the subsurface calls for an assessment of trade-offs between the benefits of additional information to reduce the risk of erroneous decisions and the cost of collecting this information. A method is suggested that combines existing procedures for inverse probabilistic groundwater modelling with a novel method for VOIA. The method results in (1) a prior analysis where uncertainties regarding the efficiency of safety measures are estimated, and (2) a pre-posterior analysis, where the benefits of expected uncertainty reduction deriving from additional information are compared with the costs for obtaining this information. In comparison with existing approaches for VOIA, the method can assess multiple design alternatives, use hydrogeological parameters as proxies for failure, and produce spatially distributed VOIA maps. The method is demonstrated for a case study of a planned tunnel in Stockholm, Sweden, where additional investigations produce a low number of benefits as a result of low failure rates for the studied alternatives and a cause-effect chain where the resulting failure probability is more dependent on interactions within the whole system rather than on specific features.

Keywords Subsidence · Urban groundwater · Value of information analysis · Inverse probabilistic modelling · Cost-benefit analysis

Introduction

When planning subsurface constructions below the water table, risks associated with groundwater drainage must be

considered. A particularly important risk in this context is land subsidence induced by groundwater drawdown in areas with compressible clay deposits (Sundell 2016). Many examples have been documented where groundwater drawdown-induced land subsidence has led to severe consequences, including Shanghai (Xue et al. 2005) and Beijing (Zhu et al. 2015) in China, Mexico City (Ortega-Guerrero et al. 1999), Bangkok in Thailand (Phien-wej et al. 2006), Las Vegas (Burbey 2002) and Los Angeles (Bryan et al. 2018) in the USA, Stockholm and Gothenburg in Sweden, and Oslo in Norway (Karlsrud 1999; Olofsson 1994).

In infrastructure projects, damage can be avoided by implementing risk-reduction measures such as improved sealing to avoid leakage and artificial infiltration of water into the aquifer to maintain groundwater levels. To evaluate the effect of a planned measure properly, the relevant properties of the hydrogeological system need to be sufficiently understood. A hydrogeological system is often characterized by heterogeneous and anisotropic materials as well as high

✉ Jonas Sundell
jonas.sundell@trafikverket.se; jonas.sundell@chalmers.se

¹ Department of Civil and Environmental Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden

² COWI AB, Box 12076, SE-402 41 Gothenburg, Sweden

³ Present address: Swedish Transport Administration, SE-405 33 Gothenburg, Sweden

⁴ Department of Mathematical Sciences, Chalmers University of Technology and the University of Gothenburg, SE-412 96 Gothenburg, Sweden

⁵ Department of Earth Sciences, University of Gothenburg, Box 100, SE-405 30 Gothenburg, Sweden

temporal and spatial variability in water balance conditions. Because of these characteristics, and since field investigations are both costly and time-consuming, the system cannot be exhaustively investigated, which means that decisions regarding the need for risk-reduction measures must be taken under uncertainty. To decide what is “sufficient”, a trade-off between the benefits of increased knowledge to reduce the risk of inappropriate decisions and the cost of new information can be made in accordance with the principles of Value of Information Analysis (VOIA). VOIA is a cost-benefit analysis (CBA) where the cost of collecting new information is compared with the expected benefits of a reduced risk of making an erroneous decision relative to a reference alternative. The result of the VOIA, from an economic perspective, is a selection of the most appropriate information collection alternatives.

VOIA as a decision support method in hydrogeological systems was introduced into a framework by Freeze et al. (1990) and further described by Freeze et al. (1992). Within this framework, the costs and benefits of alternative designs are compared using hydrogeological simulation models that account for uncertainties. If a hydrogeological system cannot be investigated in all its aspects, the problem is ill-posed, meaning many plausible models can be sufficiently consistent with available observations (Beven 2006; Carrera and Neuman 1986). Inverse probabilistic calibration to identify plausible models can be used to handle such a model structure with more independent rather than dependent parameters (Burrows and Doherty 2015; Carrera et al. 2005; Doherty 2003; Li and Zhang 2018; Siade et al. 2017; Sun 1999; Tonkin and Doherty 2009). The need for inverse calibration combined with VOIA was identified early on by Freeze et al. (1992) but no such method was present at that time. Recently, Kitanidis (2015) found it surprising that the issue had not received more attention given its importance, but points out that the topic is both conceptually difficult and computationally challenging.

A common tool for inverse probabilistic calibration of groundwater models in the numerical code MODFLOW (Harbaugh 2005) is PEST (Parameter ESTimation code). PEST allows inverse calibration of many plausible model parameterizations to be carried out based on a user-defined range. Using these calibrations, model parameter uncertainties can be estimated. Although these parameter uncertainties can be useful, it is the effects (e.g. changed groundwater heads, stream flows and water balance conditions) of a planned disturbance (e.g. groundwater extraction or infiltration) that are of primary interest. Relative parameter values that do not change over time, the future effect of a disturbance event cannot be measured at present; however, the future effect is dependent on the properties of the system described by the model and its parameters in the present state. Therefore, the hypothesis proposed here is that model parameters can serve

as proxies for such effects; furthermore, whether investigations to reduce uncertainties in model parameters also reduce uncertainty regarding future effects is studied. Finally, whether additional investigations of these parameters can change which design alternative to recommend is investigated with VOIA.

The main objective of this paper is to suggest a procedure for VOIA in order to assess the need for additional information when assessing the effect of several design alternatives with regard to future disturbances in hydrogeological systems. The procedure is shown for a specific situation, i.e. planning risk-reduction measures for groundwater drawdown in infrastructure projects in subsidence-sensitive areas where the economic cost of failure is associated with exceeding acceptable groundwater drawdown magnitudes. The general procedure recommended by Freeze et al. (1992) is updated using a VOIA method capable of assessing several design alternatives, and where hydrogeological parameters are used as proxies for failure (in this case critical lowering of groundwater heads). The uncertainty estimation is based on an inverse probabilistic calibration using PEST and MODFLOW. The result introduces spatially distributed VOIA maps as a decision support tool for planning additional investigations. The procedure is demonstrated with the aid of a case study of a planned tunnel in central Stockholm, Sweden, which will be built in crystalline bedrock below soil layers with coarse-grained materials and postglacial clay. First, the general strategy is presented together with the groundwater modelling and the VOIA procedures. Then the Stockholm case study is described.

Method

General strategy

The overall framework (Fig. 1) of the proposed method covers the two steps in VOIA: (1) prior analysis and (2) pre-posterior analysis (dashed box, Fig. 1). Within the scope of decision support, there are two additional steps, which are followed by steps (3) decision on whether to proceed with the design alternative or additional investigations and (4) posterior analysis (Freeze et al. 1992). VOIA evaluates whether additional investigations are worthwhile as a basis for choosing between design alternatives. This evaluation is initiated using a prior analysis based on groundwater modelling results and using the current stage of available information (step 1). The prior analysis is a cost-benefit analysis (CBA) where the investment cost (c_i) of design alternatives (A_i , $i = 1, \dots, m$ represents the numbering of the alternatives) is compared with the expected benefits of reduced risks relative to a reference alternative A_0 . All alternatives, including the reference alternative, involve changed drainage conditions, which are expected to disturb

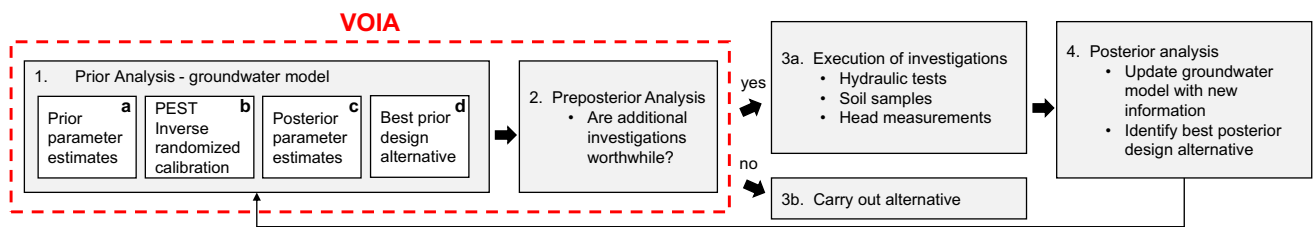


Fig. 1 Principle of the valuation of information analysis (VOIA) for groundwater models as decision support for alternative designs. Modified from Zetterlund et al. (2015)

the current water balance and groundwater head situation. The prior analysis is initiated by assigning prior estimates to groundwater modelling parameters (1a in Fig. 1). With the prior parameter estimates as initial values, a randomized inverse groundwater model calibration in PEST results in several calibrated model solutions representing plausible conditions (equally acceptable parameter settings of the ill-posed problem) for the current undisturbed situation (1b in Fig. 1). The repeated randomizations in the PEST calibrations result in updated posterior parameter estimates (c in Fig. 1). In each of the calibrated solutions, different design alternatives (including A_0) that disrupt the current situation are modelled. From these models, the effect on the water balance and groundwater heads is evaluated. The probability of failure, $P(F)$, for each alternative is calculated by comparing the difference in head relative to the current situation, with a failure criterion defined using risk areas of acceptable groundwater drawdown magnitudes (Sundell et al. 2017). Within the risk areas, the drawdown is not permitted to exceed a certain magnitude. Exceeding this failure criterion results in an economic cost of failure k_F . By multiplying k_F by $P(F)$, the risk cost (R), i.e. the expected failure cost, is set for each alternative i . The reduction in risk cost relative to the reference alternative is a benefit which is compared to the cost of obtaining the new information and the best prior alternative is identified as the alternative with the highest expected net benefit (step 1d in Fig. 1; e.g. Zetterlund et al. 2015). In the pre-posterior analysis (step 2 in Fig. 1), the expected information gain from a planned investigation is calculated by comparing the monetary benefit of the expected information with the cost of conducting the investigation. Based on the result of the VOIA, investigations (step 3a in Fig. 1) or the best prior alternative (step 3b in Fig. 1) are carried out. In the final posterior analysis, the model is updated with the new information and the decision alternatives are re-evaluated.

Prior analysis using a groundwater model

The groundwater modelling process follows general principles (e.g. Freeze et al. 1990; LeGrand and Rosén 2000; Reilly 2001) including: definition of the project goal, data collection, development of a conceptual model, development of a numerical model, model parameterization, calibration,

assessment of a problem using a simulation model and, in the final stage, a prior design suggestion based on the model results. In addition to these steps, the whole process can be repeated when new information is available. In the case study, the numerical model is constructed in MODFLOW (Harbaugh 2005) using the NWT solver (Niswonger et al. 2011) together with the PEST sub-space Monte Carlo (SSMC) (Tonkin and Doherty 2009) technique in the GMS graphic user interface (Aquaveo 2017).

In Fig. 2, the groundwater modelling in the prior analysis is initiated with a definition of a conceptual model with three soil stratification layers and fractured bedrock at the bottom. Based on these materials, the numerical model can be discretized with different layers in a three-dimensional (3D) grid. The model is then parameterized with properties such as recharge (RCH) and hydraulic conductivity (K) for the different materials. Since significant heterogeneity and anisotropy is expected within different materials, fields of material properties are modelled with the aid of pilot points (Doherty 2003) in the different layers. Pilot points (PPs) are a two-dimensional (2D) scatter-point set representing different locations within a material. As recommended in Doherty (2003), PPs should be placed with a high spatial density throughout the model domain to encapsulate heterogeneity and avoid numerical instability. From the PP, a spatially distributed parameter field is modelled using kriging (Matheron 1963) along with a variogram that approximates the heterogeneity of the parameter.

Prior parameter estimates

Based on expected variability, prior estimates of material properties are assigned to the PPs (Fig. 2). If significant material heterogeneity is expected and few or no investigations of material properties are made, the set variability span should be quite large. For locations with hydrogeological investigations such as slug tests, pumping tests or screening curves of soil fractions, a smaller variability span can be set.

Inverse probabilistic calibration with PEST

From observations such as groundwater heads, the model is calibrated using PEST, resulting in posterior estimates at the individual PP. The theoretical considerations of the tools

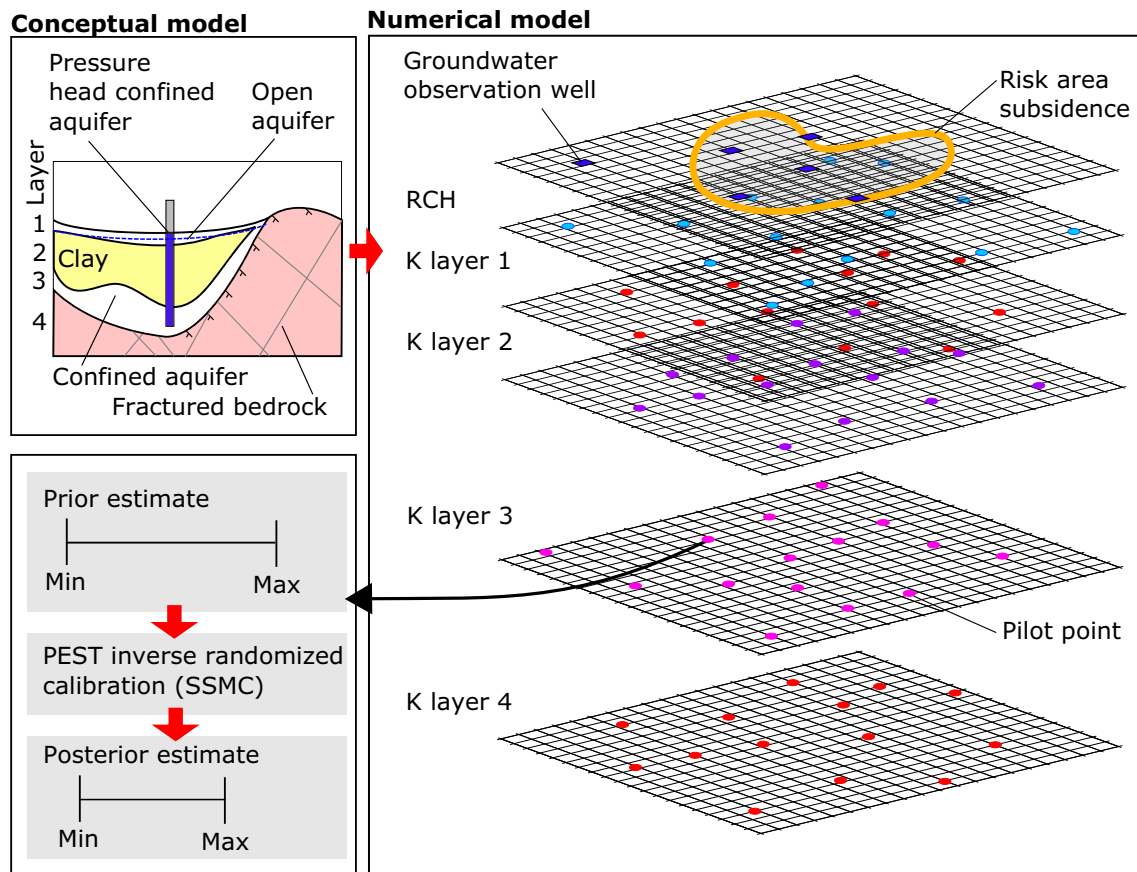


Fig. 2 Procedure for model conceptualization, definition of numerical model with grid layers, and randomized inverse calibration using PEST

included in the PEST software suite (PEST 2018) are documented over a wide range of literature (Burrows and Doherty 2015; Doherty 2003, 2011; Doherty and Hunt 2010; Fioren et al. 2013, 2009; Moore and Doherty 2005; Rossi et al. 2014; Tonkin and Doherty 2005, 2009; Woodward et al. 2016). The goal of an inverse calibration using PEST SSMC is to find the parameter combinations that meet the calibration criterion. Although the process is conditioned on the calibration criterion, it is possible that some randomizations do not fulfil this criterion or result in an unreasonable water balance. In these cases, it is important that the modeller reviews these solutions based on expert knowledge, see discussion on “hydrosense” in Hunt and Zheng (2012). Two specific criteria are used here: solutions where the difference between simulated and measured head for any observation well is greater than 1.5 m, and solutions where the difference in water balance between the inflow and outflow of water is greater than 10%, are ignored.

Posterior parameter estimates

An example with three randomized calibrated solutions is shown in Fig. 3a. Each solution consists of posterior-calibrated parameter combinations of PPs for RCH and K in

the different layers. The groundwater heads are calibrated for each observation well between the three solutions although the values of the PPs between the solutions vary. After the process is repeated for n calibrations, posterior parameter ranges can be calculated. In the example for one PP in Fig. 2, the SSMC process in PEST results in a narrower span for the posterior parameter range compared with its prior estimate.

Best prior design alternative

The effect of changed drainage conditions is modelled using the calibrated randomized solutions. These changes include groundwater leakage into a planned tunnel, as well as safety measures to reduce the effect of the leakage on the water balance and groundwater heads. The changes are represented by different design alternatives (A_i , with index $i = 1, \dots, m$), where A_0 is the reference alternative and m is the number of alternatives. The effect of an alternative simulated in each calibrated solution is illustrated in Fig. 3b. For a model with n calibrated randomized parameter solutions, the difference between these simulations results in a range of possible effects on the water balance and groundwater heads for each design alternative. From these, the likelihood of a certain groundwater level at a specific location can be calculated.

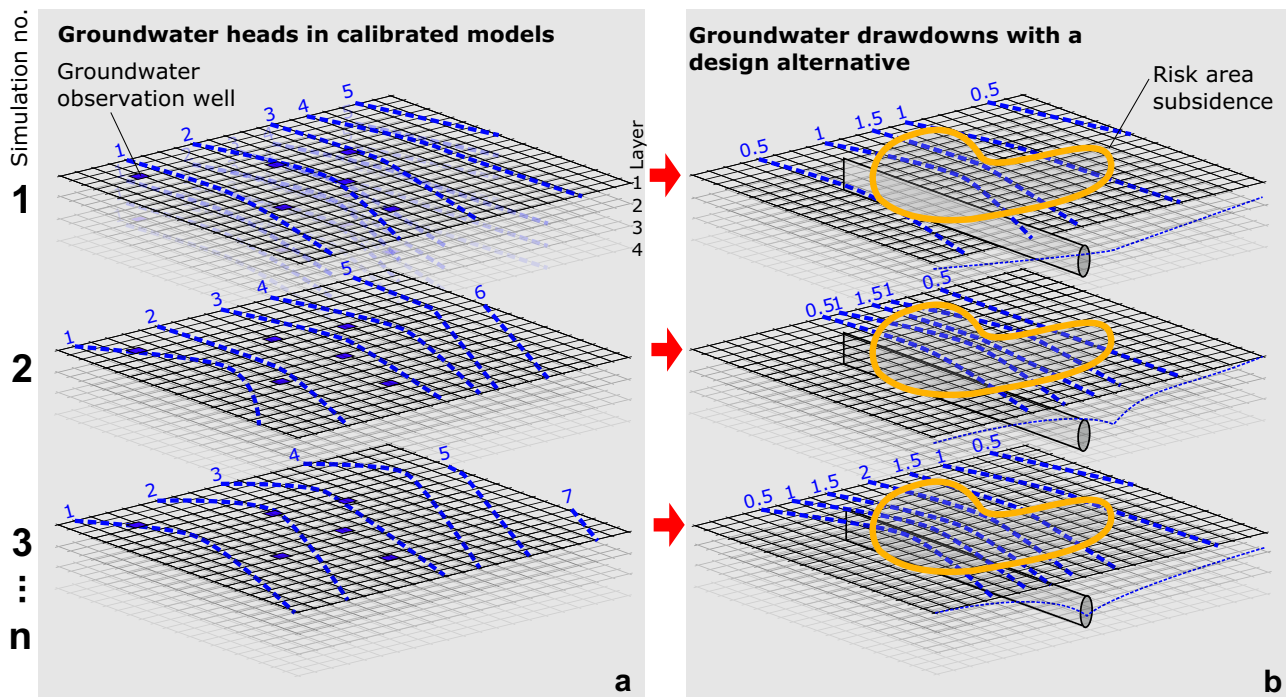


Fig. 3 **a** Three calibrated randomized solutions with hydraulic conductivity and groundwater heads. **b** Modelled effect of a design alternative on groundwater heads in each calibrated solution

To investigate whether a design alternative is acceptable, the simulations are compared with a failure criterion. Commonly, this criterion is defined by a regulation authority. In Sweden, limits for groundwater drainage for major subsurface and water supply projects are decided in the environmental court. In this process, the decision is based on the consequences of the drawdown. In the case study, land subsidence is the main consequence of drawdown. Risk areas for groundwater drawdown-induced land subsidence are defined from locations where the 95th percentile of simulations of subsidence exceed 2 cm for groundwater drawdown magnitudes of 0.5, 1 and 2 m in the confined soil aquifer. These simulations are based on a probabilistic method, where a geostatistics-based soil stratification model (Sundell et al. 2016) is combined with a one-dimensional (1D) elasto-plastic model for the calculation of consolidation settlements (Larsson and Sällfors 1986). See Sundell et al. (2017) for a complete presentation of the method and results. In order to not exceed 2 cm of subsidence the groundwater drawdown within the risk area ($g_{w_{accept}}$) should not exceed 1 m.

The calculation of a design alternative aimed at a failure criterion defined by the risk area for $g_{w_{accept}}$ is illustrated by the upper “Prior analysis” part of Fig. 4 (the risk area is also illustrated in Figs. 3 and 4). To facilitate comparison between the risk area and the modelled changes in groundwater heads, and to enable field evaluations, the comparison is realized for groundwater observation wells within and close to the risk area. Although measurements of groundwater heads exist,

the comparison is made between the simulated value in the calibrated model and the simulated value in the corresponding modified model, since every calibrated solution is a plausible representation of the reality as discussed earlier. In Fig. 4, the difference is calculated between each calibrated solution and the corresponding simulation using a design alternative. If the difference is less than $g_{w_{accept}}$ the simulation is deemed not to have failed (F^c), otherwise it has failed (F). This process is repeated for all n simulations for each alternative design, and the ratio of failed simulations corresponds to $P(F)$. Figure 4a illustrates the maximum groundwater drawdown at any observation well in three design alternatives (A_0, A_1, A_2) compared with $g_{w_{accept}}$.

Exceeding the limits of the failure criterion means that the associated consequences can take effect. If a regulation authority conditions the limit, exceedance can result in fines and delay costs in addition to costs for possible damage. The sum of these costs is the failure cost (k_F).

The risk (R) for each alternative (including A_0) is given by:

$$R_i = P(F_i) \times k_F \quad (1)$$

The benefit of an alternative (B_i) is given by:

$$B_i = R_0 - R_i \quad (2)$$

Each alternative (i) has an investment cost (c_i). The net benefit of an alternative is given by:

$$\Phi_i = B_i - c_i \quad (3)$$

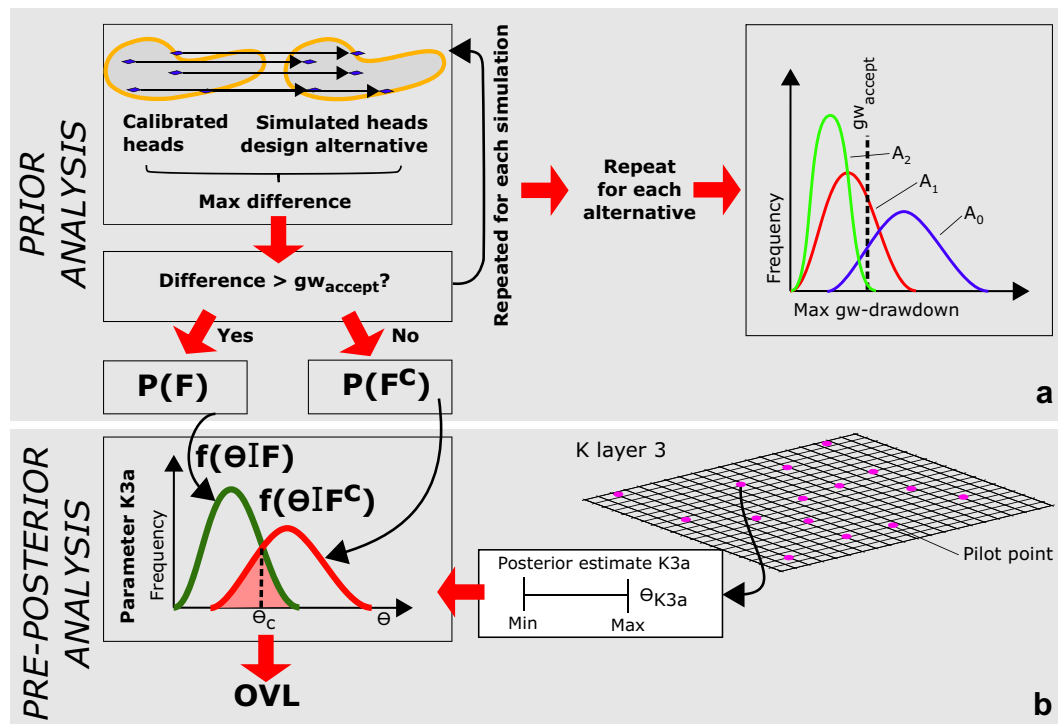


Fig. 4 **a** Prior analysis with a comparison of differences in groundwater heads at observation wells, representing the risk area between a calibrated solution and the corresponding simulation for a design alternative. The ratio of unacceptable solutions (difference in any observation well $> gw_{accept}$) results in the probability of failure $P(F)$. The graph illustrates

the maximum groundwater drawdown at any observation well for three design alternatives (A_0, A_1, A_2). **b** The initial step of the pre-posterior analysis. For each parameter, the posterior parameter estimate is grouped into frequency functions for failed and nonfailed simulations. The overlap of these results in the overlap coefficient (OVL)

Subsequently, the best prior alternative (ϕ_i) is the alternative with the greatest value:

$$\Phi_{prior} = \max(\Phi_i) \tag{4}$$

Pre-posterior analysis

If additional information changes the recommended design alternative, it is evaluated in the subsequent pre-posterior analysis. This process is initiated by separating parameter values (θ) of failed simulations from nonfailed simulations for each design alternative and parameter (“Pre-posterior analysis” part of Fig. 4 illustrated for a parameter represented by a PP named “K3a”). The division results in the relative frequency functions $f(\theta|F)$ and $f(\theta|F^c)$, where the integral of each of these functions equals 1. The overlap coefficient (OVL, Fig. 4) is used to measure the agreement between these two distribution functions (Inman and Bradley 1989). If OVL = 1 the parameters are identical, meaning that additional investigations to find the actual parameter value will not help to determine whether or not a design alternative will meet the tolerability criterion. OVL = 1 results in a probability of detecting failure, given failure $P(D|F) = 0$, and a probability of not detecting failure given nonfailure $P(D^c|F^c) = 0$. On the contrary, if OVL = 0 an additional investigation that

can detect the true parameter value with a high degree of accuracy will determine whether the alternative will meet the tolerability criterion, resulting in $P(D|F) = 1$ and $P(D^c|F^c) = 1$.

Since the parameters are represented by spatially distributed pilot points, OVL can be mapped for each parameter group (RCH and K for the different layers) and design alternative. Locations with low OVL values indicate that additional investigations can detect whether the design alternative will fail or not.

For cases with an OVL between 0 and 1, and in the parameter value range where the functions overlap, it is uncertain whether a sampled value belongs to $f(\theta|F)$ or $f(\theta|F^c)$. In this range, a critical parameter value, θ_c , is selected. For the example in Fig. 4, it is assumed that all values below θ_c belong to $f(\theta|F)$ and above θ_c it is assumed that all values belong to $f(\theta|F^c)$, meaning that $D = [\theta < \theta_c]$ and $D^c = [\theta \geq \theta_c]$. The opposite condition takes affect if it is $f(\theta|F)$ instead of $f(\theta|F^c)$ that contains the highest values. Detection of failure can either result in a true detection, $P(D|F)$ or a false detection (type I error), $P(D|F^c)$. Similarly, D^c can either result in $P(D^c|F^c)$ or the type II error $P(D^c|F)$. If θ_c is chosen as the intersection point of the two functions, as illustrated in Fig. 4, the sum of the type I and type II errors is minimized. With this choice, $OVL = P(D^c|F) + P(D|F^c)$ (Fig. 5).

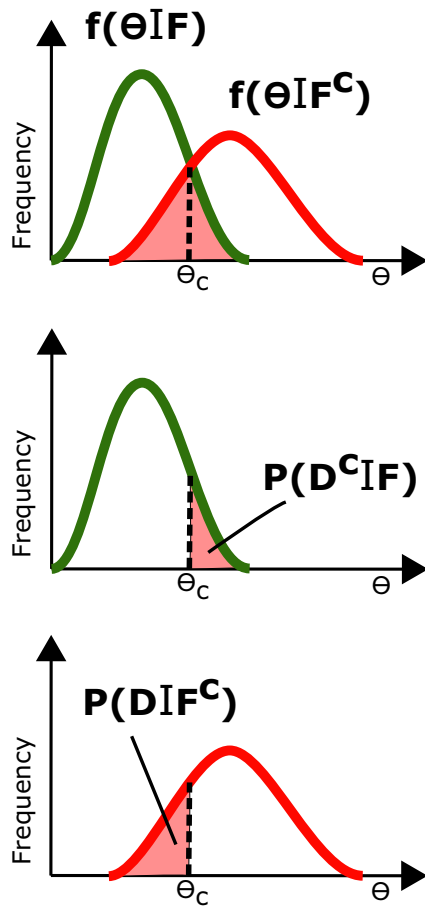


Fig. 5 Choice of critical parameter value (θ_c) based on OVL together with the type II $P(D^c|F)$ and type I errors $P(D|F^c)$

For each parameter and alternative, OVL and the θ_c are calculated by testing each position along the parameter value axis (θ). In each test, $P(D^c|F)$ is calculated for $f(\theta|F)$ and $P(D|F^c)$ is calculated for $f(\theta|F^c)$. The position that minimizes $P(D^c|F) + P(D|F^c)$ gives OVL and θ_c .

In the subsequent step, $P(D|F)$ is compared between the alternatives. This comparison is complicated by the different positions of θ_c between the alternatives, resulting in a different detection event (D_i) for each alternative, as illustrated in Fig. 6. For the case with three alternatives, there are $2^3 = 8$ different detection possibilities:

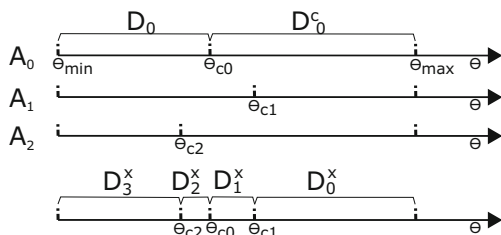


Fig. 6 Example of the location of θ_c for three alternatives (A_0, A_1, A_2) together with the detection intervals (D_j^x)

- $D_0 \cap D_1 \cap D_2$
- $D_0 \cap D_1 \cap D_2^c$
- $D_0 \cap D_1^c \cap D_2$
- $D_0^c \cap D_1 \cap D_2$
- $D_0 \cap D_1^c \cap D_2^c$
- $D_0^c \cap D_1 \cap D_2^c$
- $D_0^c \cap D_1^c \cap D_2$
- $D_0^c \cap D_1^c \cap D_2^c$

To make the comparison between the different detection events for each alternative plausible, detection intervals D_j^x are introduced based on the different positions of θ_{ci} :

$$D_3^x = D_0 \cap D_1 \cap D_2 \tag{5}$$

where failure is identified in all three alternatives,

$$D_2^x = (D_0 \cap D_1 \cap D_2^c) \cup (D_0 \cap D_1^c \cap D_2) \cup (D_0^c \cap D_1 \cap D_2) \tag{6}$$

where failure is identified in two alternatives,

$$D_1^x = (D_0 \cap D_1^c \cap D_2^c) \cup (D_0^c \cap D_1 \cap D_2^c) \cup (D_0^c \cap D_1^c \cap D_2) \tag{7}$$

where failure is identified in one alternative,

$$D_0^x = D_0^c \cap D_1^c \cap D_2^c \tag{8}$$

where failure is not identified for any alternative.

For the example in Fig. 6:

$$D_3^x = D_0 \cap D_1 \cap D_2 \tag{9}$$

$$D_2^x = D_0 \cap D_1 \cap D_2^c \tag{10}$$

$$D_1^x = D_0^c \cap D_1 \cap D_2^c \tag{11}$$

$$D_0^x = D_0^c \cap D_1^c \cap D_2^c \tag{12}$$

The other detection events for failure of three and two alternatives remain empty for the example.

Based on Eqs. (5)–(8), $P(F_i|D_j^x)$ and $P(D_j^x)$ are calculated for each alternative, A_i and detection event, D_j^x from the locations of θ_{ci} and the previous grouping of failed and nonfailed parameter values. From these calculations, the posterior net present value is calculated for each alternative and detection event by making a comparison with the reference alternative:

$$\Phi_i(D_j^x) = B_i(D_j^x)^{-c_i} = k_F [P(F_0|D_j^x) - P(F_i|D_j^x)]^{-c_i} \tag{13}$$

The best net posterior present value for a detection interval is:

$$\Phi_{\text{posterior}}(D_j^x) = \max [\Phi_i(D_j^x)] \tag{14}$$

The pre-posterior present value is the sum of the best posterior present values for all detection events multiplied by their probability:

$$\Phi_{\text{pre-posterior}}(D_j^x) = \sum_j \Phi_{\text{posterior}}(D_j^x) \times P(D_j^x) \tag{15}$$

Finally, the expected net value of information (EVI) is calculated as the difference between the pre-posterior and the prior net present values:

$$EVI = \Phi_{\text{pre-posterior}} - \Phi_{\text{prior}} \quad (16)$$

In these calculations, information is only of value if the investigation has the potential to change the decision on which alternative to recommend. EVI does not take into account the data collection cost (c_p). To do so, the expected net value

(ENV) is calculated:

$$ENV = EVI - c_p \quad (17)$$

Commonly, several investigations are suggested as part of an investigation programme. With the procedure pre-

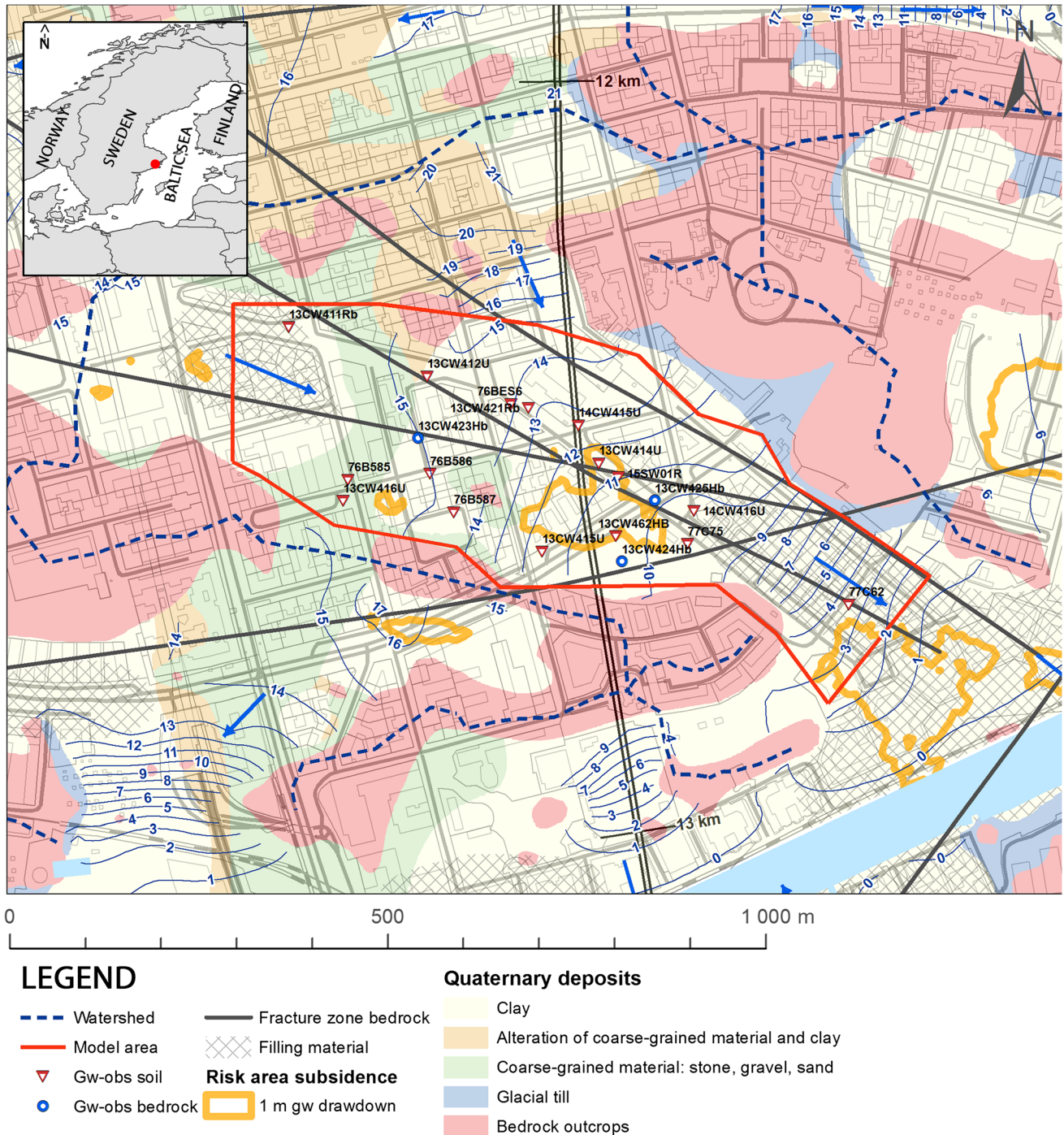


Fig. 7 Map of the model area, including observation wells in the confined aquifer and observation wells in bedrock. Groundwater pressure levels in the confined aquifer are illustrated by iso-lines interpolated from observation data (Sundkvist 2015). The broken blue lines show a groundwater watershed and the blue arrows show the flow direction in the confined

aquifer. The background map illustrates major fracture zones in bedrock, bedrock outcrops, Quaternary soil deposits, and later-age filling material (Stockholms stad 2014). The double lines in a N-S direction indicate the location of the planned tunnel

sented, combinations of several investigations are not considered. Nevertheless, it is reasonable to recommend several locations with a high $\Phi_{\text{pre-posterior}}$ as potential sampling locations unless parameters within the same group are in close proximity.

Posterior analysis

If the pre-posterior result suggests carrying out the investigation programme, the programme is realized, and new information is obtained, the model is updated in a posterior analysis.

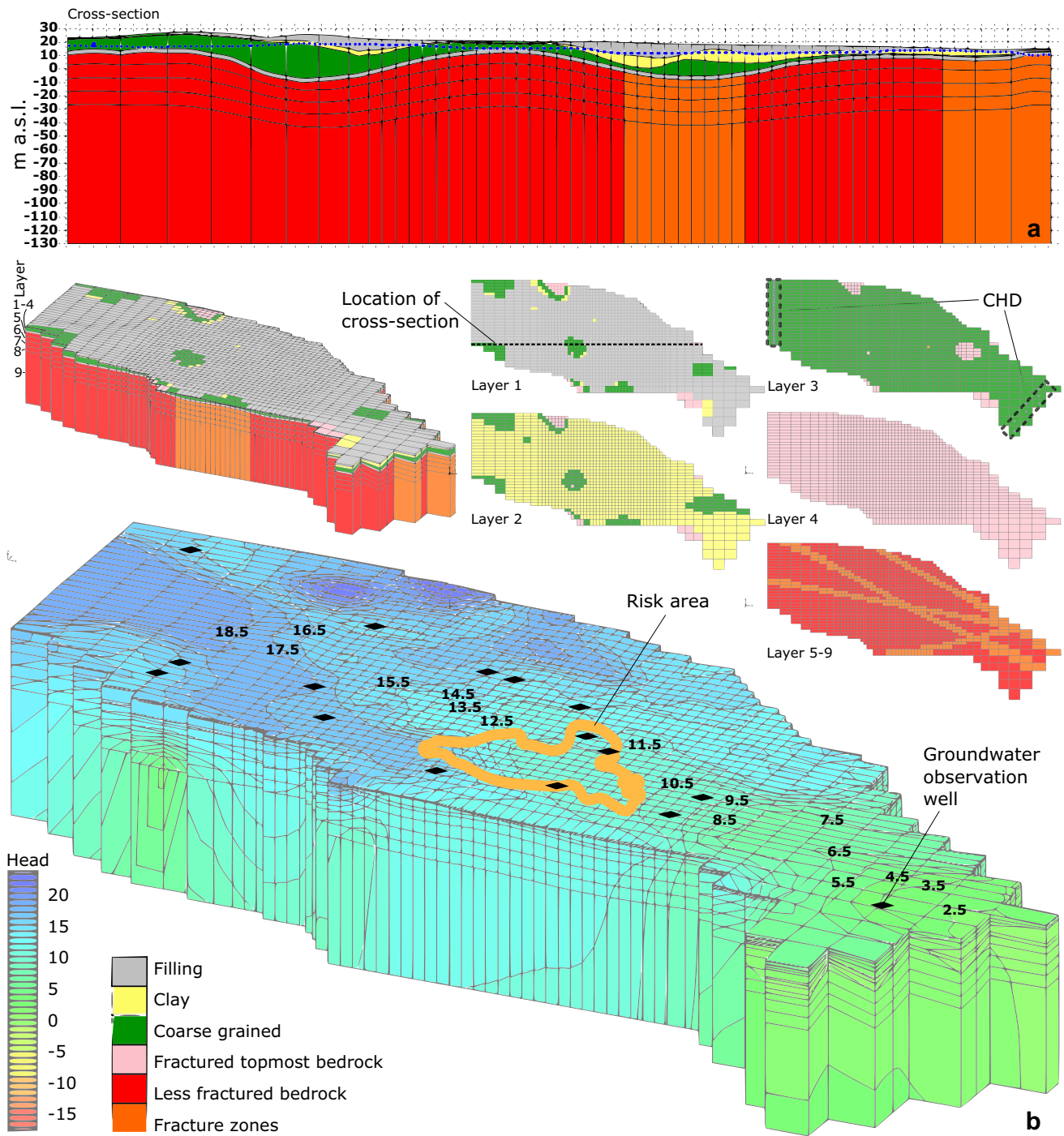


Fig. 8 a Three-dimensional model grid and materials. b Groundwater heads of the initial calibration and locations of the observation wells in layer 3 (black dots)

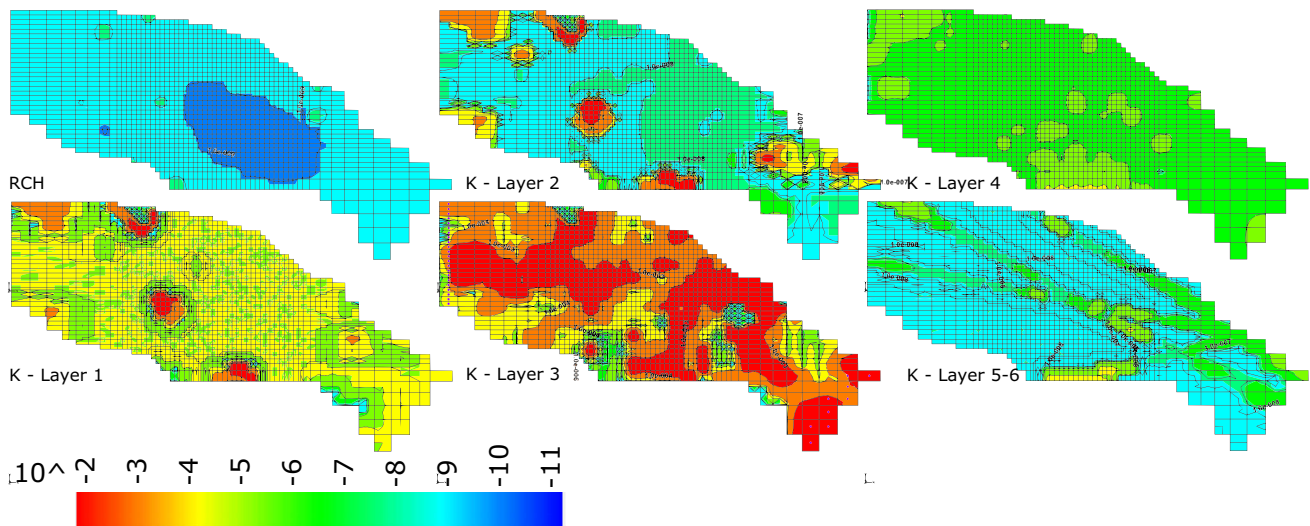


Fig. 9 Recharge (RCH) and hydraulic conductivity (*K*) fields in each layer for the initial calibration

This updating will follow the same procedure as the calibration in the prior analysis. After this step, the process loop continues according to Fig. 1 until a design alternative has been executed.

Case study

The method is applied in an area in central Stockholm, where a utility tunnel for power lines, 5 m in diameter, is modelled. The soil layers in the area are characterized by the Stockholm Esker with coarse-grained material as an open aquifer in the

western part and confined by clay in the eastern part (Fig. 7). The pre-Cambrian crystalline bedrock has dominant fracture zones in an E–W, WNW and NW direction (Persson 1998). Although important data in the area is available from the investigations of the planned tunnel such as a model for soil and bedrock stratification, groundwater head measurements and hydraulic tests, there are significant uncertainties. These uncertainties include information on groundwater recharge (from precipitation and leakage from fresh water and sewage water distribution), drainage into existing tunnels, and *K* values of materials and locations not tested or not tested with sufficient accuracy. Based on an inverse calibrated steady-state SSMC

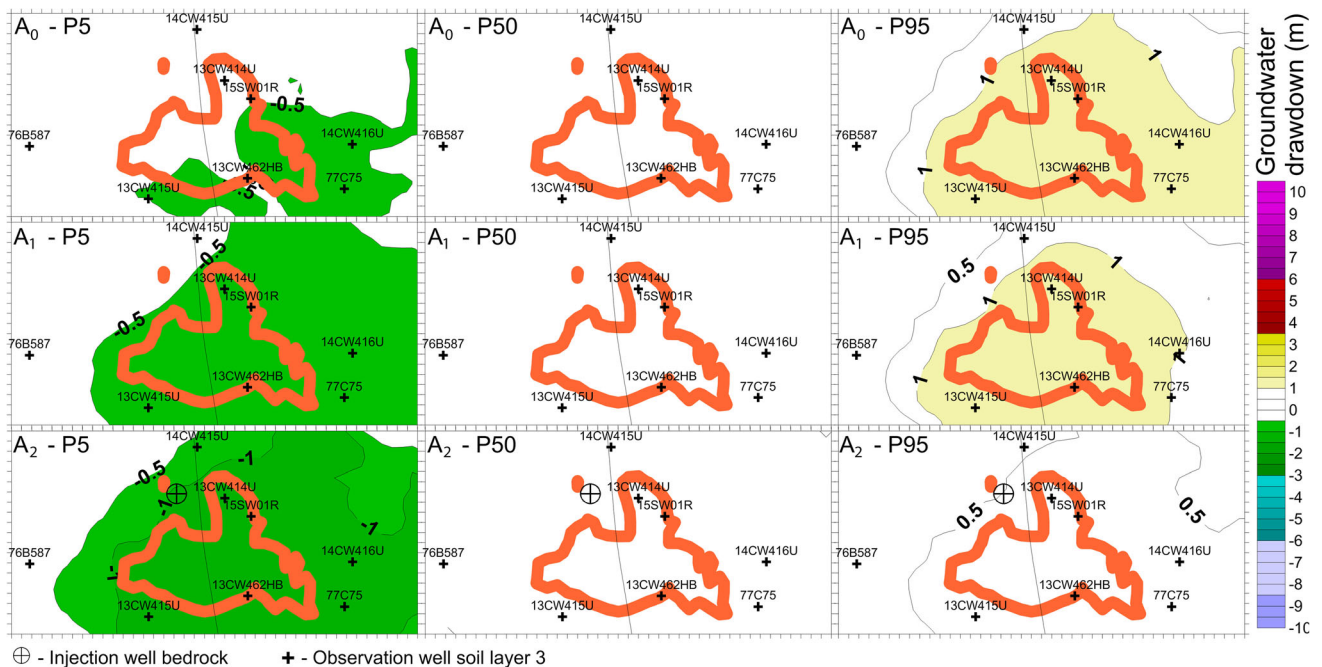


Fig. 10 5th, 50th and 95th percentiles for simulated groundwater drawdowns in A_0 , A_1 and A_2 . A negative value indicates a higher level relative to the calibrated model. The tick marks are located at 10-m intervals

Table 1 Cost-benefit analysis (CBA) for the different alternatives, A_0 , A_1 and A_2 , with 1 m and 0.5 m groundwater drawdown as the failure criterion

Alternative	$P(F)$	k_f	R_i	B_i	c_i	Φ_i
$A_0 - 1$ m	4.7%	20	0.934	0	0	0
$A_1 - 1$ m	3.2%	20	0.639	0.295	0.300	-0.005
$A_2 - 1$ m	1.5%	20	0.295	0.639	8.777	-8.139
$A_0 - 0.5$ m	19.4%	20	3.89	0	0	0
$A_1 - 0.5$ m	11.5%	20	2.31	1.57	0.300	1.272
$A_3 - 0.5$ m	1.7%	20	0.34	3.54	8.777	-5.239

Cost and benefits in MSEK. $P(F)$ probability of failure, k_f cost of failure, R_i risk for each alternative, B_i benefit relative to the reference alternative (A_0), c_i investment cost of each alternative, Φ_i net benefit of an alternative relative to the reference alternative.

model, three design alternatives for the planned tunnel are modelled. A_0 (reference alternative) includes a tunnel without sealing in fracture zones; A_1 , a tunnel with sealing in fracture zones to $K = 10^{-8}$ m/s; and A_2 , same as A_1 but with three injection wells in the bedrock. Details of the case study are referred to as supporting information.

Groundwater model

The groundwater model is constructed in nine different layers with five different materials in bedrock and soil. Based on major expected differences in K , the soil materials are divided into three categories (Fig. 8): filling material, clay and coarse-grained soil (esker material and glacial till). The bedrock is divided into three categories (Fig. 8): a more fractured top

layer, vertical fracture zones, and less fractured bedrock. Each of these materials is assigned prior parameter distributions (see supporting information), and the materials with a dominant response to the effect of the design alternatives are modelled using PPs. The model is calibrated against the observation wells presented in Fig. 7 and is further described in supporting information.

Initial calibration

In a first step, the model is calibrated by manually adjusting the parameter sets to close proximity between the observed and modelled heads. Based on this model, a model parameter field with minimum error variance is calibrated using PEST. In these two steps a conductance value for all tunnels is set at a fixed value, 5×10^{-8} m²/s/m, to represent a reasonable inflow for tunnels in bedrock with cement injection. The K value for the less fractured bedrock is set at a fixed value of 10^{-8} m/s. Since the K of the other materials is several orders of magnitude higher, particularly in the fracture zones that dominate the inflow into the tunnels in bedrock, and it determines the connectivity with overburden layers, variability of K in the less fractured bedrock is of lesser importance for the result. The K value in the filling material is set at a fixed value of 1×10^{-5} m²/s. The resulting heads can be observed in Fig. 8 and fields of RCH and K in Fig. 9. For five wells in soil, the residuals between observed and modelled heads are between 0.2–0.9 m. In all the other wells, the residual is less than 0.2 m. One observation well in bedrock,

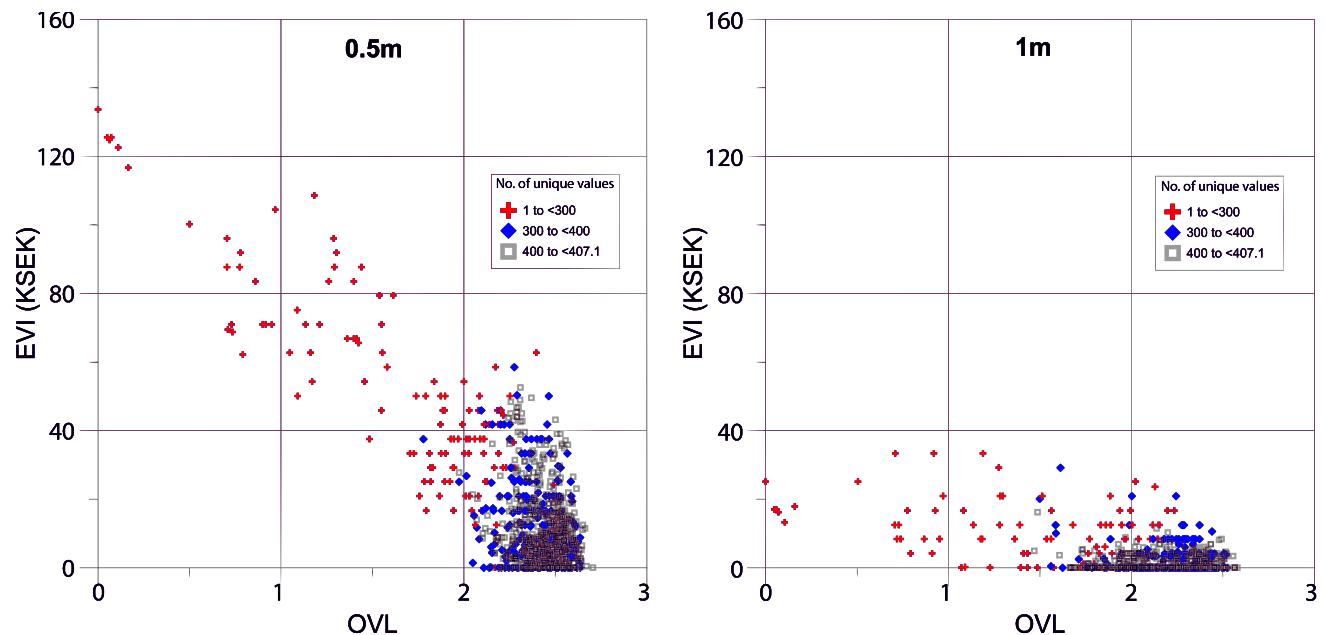


Fig. 11 Correlation between EVI (in thousands SEK: KSEK) and OVL for three classes of unique values (1–300, 300–400 and 400–407 respectively) with **a** 0.5 m and **b** 1 m as a failure criterion

13CW424Hb, is excluded from calibration as it is not sectioned for different fracture zones, which means that it is not possible to determine what the measured head represents. When calibrating the model, the residuals in 13CW424Hb were always greater than 1.5 m, irrespective of the adjustments made.

Inverse probabilistic calibration

Using the result from the PEST calibration, variograms are modelled for the logarithms of each PP. In the SSMC step, the PP is randomized based on log-normal distributions, where the result of the initial PEST calibration represents

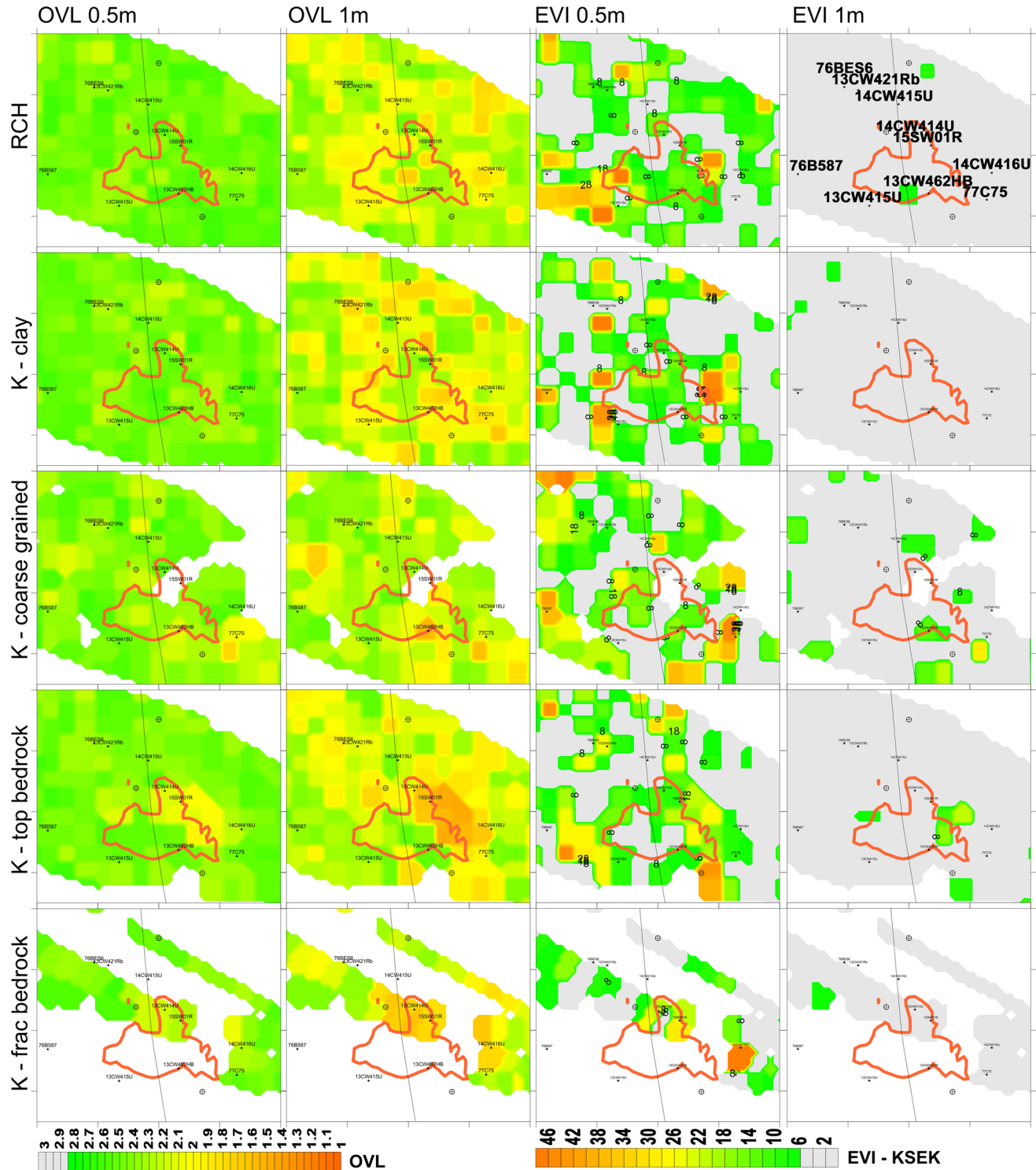


Fig. 12 OVL and EVI (in thousands SEK: KSEK) for $g_{w,accept}$ of 0.5 m and 1 m in each of the PP sets (RCH and K for the different materials)

mean values. The prior uncertainties are defined by the bounds of minimum and maximum values, see Appendix section ‘[Material properties](#)’. Because of this definition, a high value for the standard deviations of each parameter set ($SD = 1.95$) is selected, resulting in prior distributions truncated by their respective bounds. From these definitions of prior distributions, 1,000 SSMC runs are made, which was found to be the practical limit for the power station used for computation in this study. Out of the 1,000 runs, 747 were remaining after nonconverged solutions were removed, 670 were remaining after removing differences in simulated and observed heads greater than 1.5 m, and 563 were remaining after removing solutions with a water balance discrepancy between inflow and outflow greater than 10%. The accepted solutions demonstrate homoscedastic errors without any systematic under— or overestimation between observed and modelled heads. The resulting water balance for the accepted solutions is presented in Appendix section ‘[Water balance: accepted SSMC solution](#)’.

Prior analysis

From the calibrated SSMC solutions, changes in groundwater head and water balance are simulated for the three different

design alternatives. The planned tunnel is located in layer 8. In A_0 (reference alternative), sealing is represented by a conductance value of $5 \times 10^{-8} \text{ m}^2/\text{s}/\text{m}$, as assumed for the other tunnels in the area. In A_1 , the cells representing adjacent fracture zones are sealed to $K = 10^{-8} \text{ m/s}$. A_2 is the same as A_1 but with three injection wells in the fracture zones in the bedrock, layer 5 (Fig. 10). In A_2 , flow into the wells is conditioned to a stop criterion that equals the surface level.

Out of the 563 calibrated solutions, simulations that did not converge in any of the alternatives were removed from further analysis, resulting in 407 remaining solutions. Water balances for the different alternatives are presented in Appendix section ‘[Water balance: design alternatives](#)’.

The difference in groundwater level between the calibrated models and each alternative is presented in Fig. 10. In A_0 , A_1 and A_2 , the median groundwater change is less than 0.5 m for each alternative (P50 in Fig. 10). The 5th percentiles show increased levels in all the alternatives as the injected tunnel can create a barrier in some cases. In the 95th percentiles, all alternatives have reduced levels. The failure criterion is defined for seven observation wells within or in close vicinity to the risk area: 14CW415U, 13CW414U, 15SW01R, 14CW416U, 77C75, 13CW462HB and 13CW415U. These wells are tested for a $g_{w\text{accept}}$ of 0.5 and 1 m respectively. If

Table 2 Groundwater observations for calibration and calculated K values in test wells in soil and bedrock (last three wells)

Obs. well name	gw level masl	Estimated K (m/s)	Test for K estimation	Tested material	Reference
13CW411Rb	15.19	1e-3	Constant rate infiltration test	1 m glacial till (silt and gravel), 1.4 m bedrock	Björgúlfsson (2014)
13CW414U	11.36	—	—	—	—
13CW415U	11.32	—	—	—	—
13CW416U	15.14	—	—	—	—
13CW421Rb	13.20	3e-3	Constant rate pump test	3 m layers with clay, silt and gravel, 3 m bedrock	Björgúlfsson (2014)
14CW415U	12.60	—	—	—	—
14CW416U	9.40	—	—	—	—
76B585	15.18	—	—	—	—
76B586	14.99	—	—	—	—
76B587	14.81	—	—	—	—
76BES6	14.32	—	—	—	—
77C75	9.47	—	—	—	—
77C62	3.37	—	—	—	—
13CW412U	15.98	—	—	—	—
15SW01R	10.73	3 - 7e-4	Constant rate pump test	Soil layer above bedrock	Berzell and Dehkordi (2017)
13CW462HB	10.29	—	—	—	—
13CW425Hb	10.40	1.4e-5 2e-8	Pulse test with packer at 5 m in bedrock	Bedrock, top 5 m Bedrock 5–100 m	Björgúlfsson (2014)
13CW424Hb	10.13	4.3e-7	Constant rate pump test	Bedrock section 38–84 m	Björgúlfsson (2014)
13CW423Hb	13.00	1.4e-5 1.2e-8	Pulse test with packer at 7 m in bedrock	Bedrock, top 7 m Bedrock, 7–81 m	Björgúlfsson (2014)

Table 3 Prior estimates of minimum and maximum values of K for the different materials

Material	Min K (m/s)	Max K (m/s)
Filling material	1e-8	1e-3
Clay – PP: HK_11	1e-11	5e-7
Coarse-grained soil – PP: HK_12	5e-7	1e-2
Fractured uppermost bedrock – PP: HK_13	1e-7	1e-5
Less-fractured bedrock	1e-11	5e-8
Fracture zones – PP: HK_15	1e-8	1e-5

the groundwater drawdown in any of the wells is higher than $g_{w_{accept}}$, the simulation is regarded as a failure.

Although a detailed estimation of k_F is beyond the scope of this paper, an overall principle with reasonable estimates is given here. Failure costs can be both direct and indirect. Direct costs refer to costs for repairing the damage, whereas indirect costs include project delays, a lower market value of the damaged buildings, or inconvenience for the tenants. The buildings within the risk area are founded on friction piles in concrete or steel although the basement floors have shallow foundations. The main concern regarding subsidence is not an

additional pile load but damage to the basement floor (case M 2772–15, Land and Environment Court at the District Court in Nacka, 2016-11-30). Damage to the basement floor can be aesthetic or functional but is not assumed to result in structural damage affecting the stability of the buildings. With an assumed cost of SEK 4,000/m² and a building area of 500 m², k_F is estimated at SEK 2 million (SEK 10 = approximately €1). If instead k_F is assumed to result in delay costs because of not meeting the acceptance criterion, a delay of 1 month is estimated to cost SEK 20 million (case M 2772–15). During this month, the contractor is supposed to improve fracture sealing or infiltration and then continue with the project. This assumption implies that the modelled groundwater head changes in the steady-state model occur and can be observed within a short period of time after leakage into the bedrock tunnel begins.

The failure ratio together with a prior CBA analysis of the different alternatives is presented in Table 1 for 1 and 0.5 m as the failure criterion. The cost of sealing a fracture zone is estimated at SEK 100,000. With three zones crossed in A₁, c_i is calculated. In A₂, the cost of sealing in A₁ added to the installation cost of a permanent infiltration well is estimated at SEK 0.5 million. For the three wells in A₂, a cost for

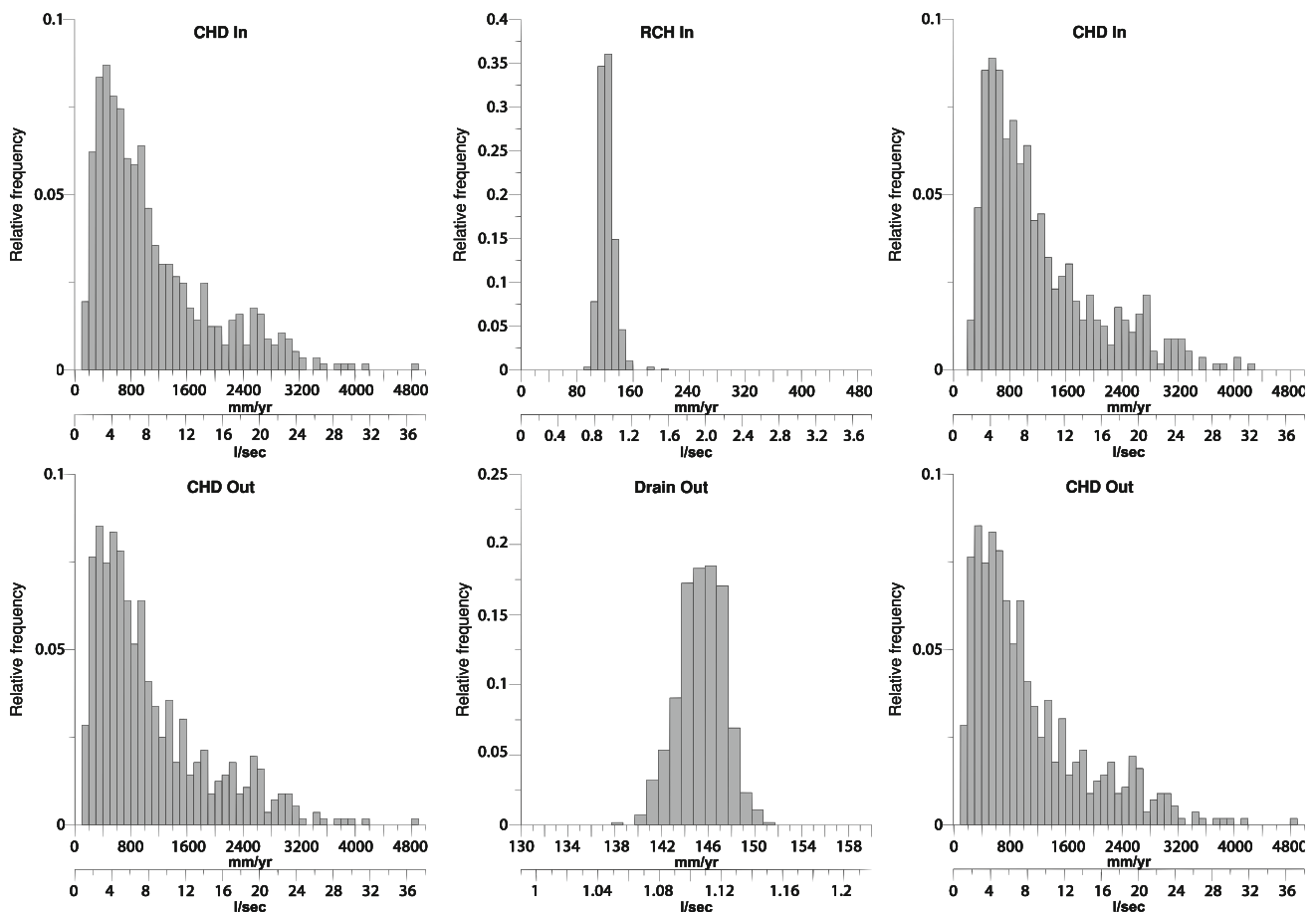


Fig. 13 Histograms with the water balance for the 563 accepted solutions expressed as l/sec and mm/yr/m² over the model area. Note the different scales between the graphs

infiltrating 0.4 l/sec of freshwater with a unit price of SEK 20/ m³ for a 100-year period with discount rate of 3.5%, c_1 is calculated for this alternative. With 1 m as the failure criterion, A_0 is identified as the best prior alternative although the difference is small in comparison with A_1 . If k_F is reduced to SEK 2 million, A_0 is still the best prior alternative. With 0.5 m as the failure criterion, A_1 is identified as the best prior alternative. If k_F is reduced to SEK 2 million, A_0 is the best alternative.

Pre-posterior analysis

Following the description in section ‘Prior parameter estimates’, OVL and EVI are calculated in each of the PP sets (Figs. 11 and 12). OVL is presented as the sum of OVL for all three alternatives, resulting in a theoretical variation of OVL between 0 and 3. The result of the SSMC calibration shows that some PP sets have several values (θ) equal to the bounds of the prior parameter estimates, resulting in fewer unique values (θ) within the entire distribution (red dots in Fig. 11). Since θ and the

bounds have the same value, they cannot be ranked and differences between high and low θ cannot be observed for $(\theta|F^c)$ and $(\theta|F)$. Nevertheless, these θ are ranked in the calculations, which often result in cases with a low OVL and a high EVI, which is an artificial result. To eliminate such cases, only PPs with more than 300 unique θ 's are presented in Fig. 12. This elimination affects K to a large extent for coarse-grained material and fracture zones, indicated by the white areas at the location of eliminated PPs in Fig. 12 (compared with Fig. 9). OVL is in general lower, with 1 m compared to 0.5 m as a failure criterion. Low $P(F)$ values at 1 m result in fewer values in the failure category. Fewer values in one of the categories increases the possibility of both type I and II errors if these values are clustered by chance. Type I errors (false confirmation of a large OVL) occur for clusters close to the bounds of the value ranges for a PP, which can be observed by the scattered presence of low OVL values for 1 m in RCH and K for clay in Fig. 12. For K , uppermost bedrock and fractured bedrock spatial clusters of lower OVL values are observed both for 0.5 and 1 m. These clusters are correlated with locations

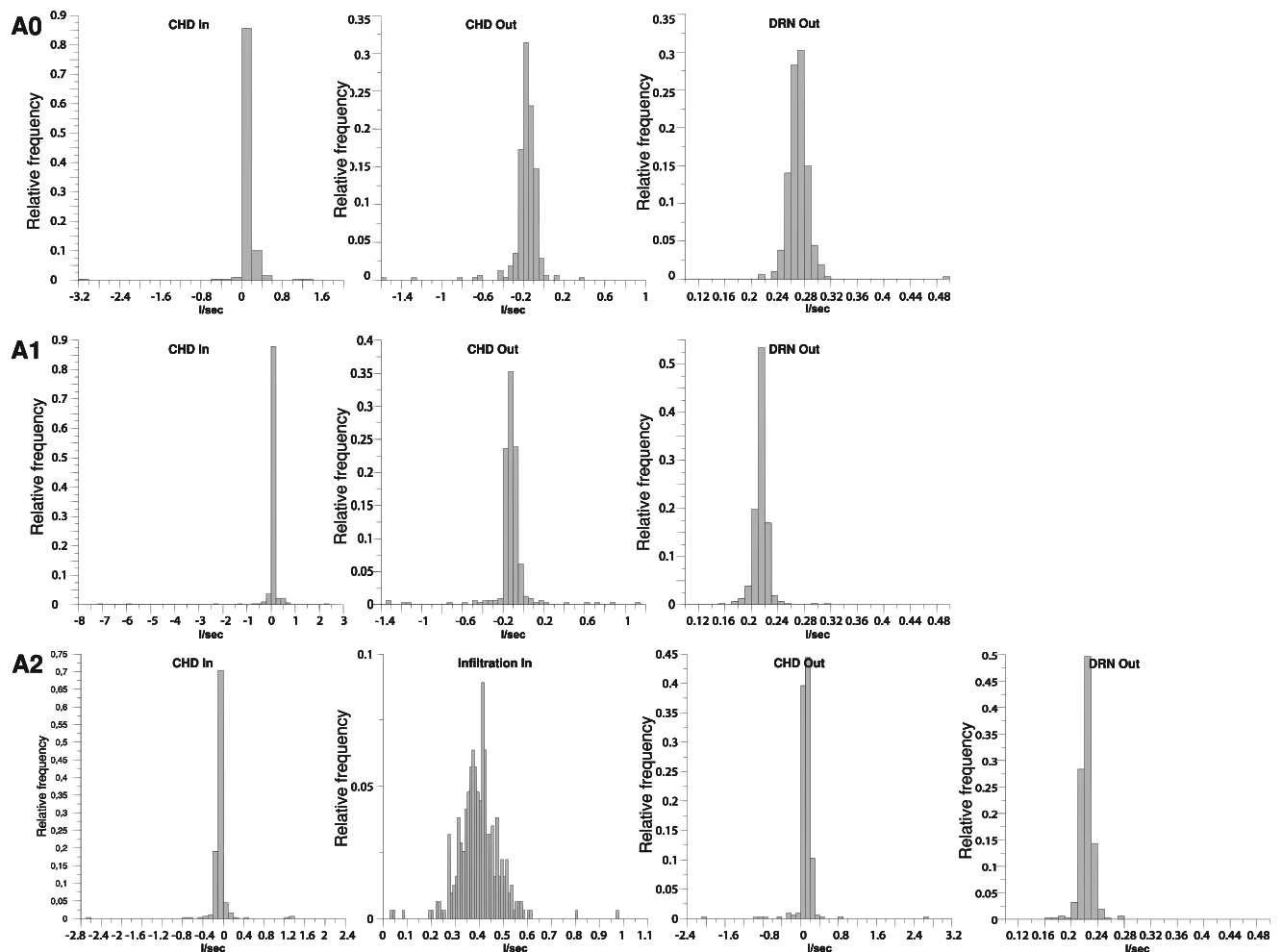


Fig. 14 Change in the water balance between calibrated simulations and alternatives. Note the different scales between the graphs

close to settings that have a direct influence on failure (risk area, tunnel and infiltration wells). With these consistencies, it is reasonable to expect that future observations at the locations of these clusters can improve the possibility of determining failure or no failure (F or F^c) in an alternative.

Although a low OVL indicates that F can be detected, it is possible that sampling at this location has no or low EVI if the sampling does not change the recommended alternative from the prior analysis. The stability of the result is tested using a bootstrap analysis, where the null hypothesis, H_0 , states that EVI is independent of the sample size. Sample sizes of 100, 150, 200, 250, 300, 350 and 407 are tested. In sample sizes <300 , EVI calculations yield unstable results since $P(F)$ is low, which results in no observations of $(\theta|F)$. With 0.5 m as a failure criterion, sample sizes >300 typically result in \pm SEK 25000 for an 80% confidence interval of future observations. This result can be compared with the scatter plot in Fig. 11, which shows a deviation in this range for PPs with the same OVL value. With 1 m as a failure criterion, the difference between A_0 and A_1 is very small (A_2 is never the preferred pre-posterior alternative), which in general results in a small EVI. Since zero is often within the 80% confidence interval of future observations, the result indicates that additional sampling is not beneficial with 0.5 m as a failure criterion.

Even if the number of realizations for 0.5 m as the failure criterion is sufficient to detect a positive EVI, the values are low and no specific PP is a major contribution to the pre-posterior analysis. This is a result of a cause-effect chain (initiated by groundwater leakage and infiltration, groundwater drawdown in different layers in bedrock, and failure determined by drawdown in the soil layer) that is more dependent on the interactions in the whole system rather than specific features. For other problems, where both the cause and the effect are closely related, higher EVI values are expected at PPs connected to the studied phenomena. This situation is likely to occur in water supply studies where failure is related to extracted groundwater quantities, which are in return related to K in the pumped layer. Even if the EVIs are reliable in the presented case study, the highest values (with more than 300 unique θ 's) are about SEK 50000. This means that c_p must be higher than this value to create a positive ENV. In addition, the investigation method needs to be sufficiently accurate to detect the actual value of the PPs. The previous investigations in the case study are both more expensive and quite inaccurate (see the [Appendix](#)), meaning that none of these conditions are assumed to be met. Low ENV values do not mean that the model is unreliable, more that additional information is not expected to be worthwhile. Furthermore, it implies that the recommendation in the prior analysis is sufficient. Since the EVI is calculated for one investigation at a time (a limitation of the method presented), it is possible that combinations of several investigations would produce another result.

Despite large variations in parameter values and water balance, the different realizations indicate low failure rates. The main reason for this is that the confined aquifer is very conductive and has a large groundwater flow relative to the bedrock layers and inflow into the tunnel. Nevertheless, additional investigations do not need to be worthless if they are able to change the conceptual understanding of the system and not only parameter uncertainties that are addressed in the presented method.

Conclusion

This article presents a novel method for VOIA to assess the need for additional information when estimating the effect of multiple design alternatives on future disturbances of hydrogeological systems. The method presented here facilitates a spatial VOIA of hydrogeological information for multiple design alternatives, which to the knowledge of the authors has previously not been possible. With this method, the economic benefit of additional investigation is presented in maps, which can be an important decision support tool with regard to additional investigations. The case study results indicate low expected value of information because of low failure rates for the studied alternatives, and a complex cause-effect chain where the resulting failure probability is more dependent on the interactions within the whole system rather than on specific features. This result means that no additional investigation can be recommended at any specific location, and that the recommendation from the prior analysis is sufficient. It should be emphasized that this conclusion is site-specific and that the value of hydrogeological information in projects relating to groundwater drawdown-induced land subsidence is expected to exhibit a large degree of variation between different projects and different hydrogeological settings.

Future research on implementing the method in less complex cause-effect chains, where both the cause and the effect are more closely related than the presented case study, is recommended for calculation of the relationship between the ability of hydrogeological information to represent the failure criterion and the value of additional hydrogeological information. Furthermore, this ability is expected to improve with additional inverse calibrations. Finally, it is recommended to examine the possibility of calculating the economic benefits of combined investigation programmes and not only one investigation programme at a time.

Acknowledgements Data for this study were supplied by the planned City Link Tunnel project in Stockholm. We would also like to thank the four anonymous reviewers and the editor, whose comments and insights helped to improve the paper.

Funding information We gratefully acknowledge the funding of this work by Formas (contract 2012-1933), BeFo (contract BeFo 333), and the COWI Fund (grant HHT/A-119.23/jat).

Appendix: supporting case study information

Geological and hydrogeological conditions

The model area is illustrated in Fig. 7 together with the risk areas for 1 m groundwater drawdown taken from Sundell et al. (2017). The bedrock in the area is characterized by Precambrian crystalline bedrock of the Svecofennian Orogeny with radiometric ages of 1,750–2,000 years (Stålhös 1969). Dominant fracture zones are located E–W, WNW and NW (Persson 1998). During the last glaciation (Weichsel), glacial till was deposited directly on the bedrock. During deglaciation, substantial glaciofluvial sedimentation took place. The green area in Fig. 7 shows parts of a glaciofluvial esker deposit, the Stockholm Esker. The central parts of the esker consist of coarse-grained sandy, gravelly and stone fractions. At more distal parts of the esker, coarse-grained and fine-grained material is often layered. Previously, the Stockholm Esker was a local topographic high point, but significant excavation and backfilling has evened out its topography. The whole study area is below the highest shoreline (150 m above current mean sea level, masl). As a result, glacial and post-glacial clay has been deposited in the valleys and previously deposited material at higher altitude points has largely been washed out due to abrasion during land uplift. Consequently, postglacial sands together with anthropogenic filling materials on top of the clay deposits make up surficial aquifers.

The clay deposits make up an aquitard between the confined aquifer in the esker, and between the glacial till and the surficial aquifer. At locations where clay layers are nonexistent, hydrostatic conditions do exist and the confined conditions change to open. Figure 7 displays interpolated groundwater levels in the confined aquifer throughout the map area from about 50 observation wells (only the wells used for calibrations are presented in Fig. 7). As can be seen, the whole model area is within the same watershed. Initial groundwater models covered the whole watershed but due to convergence problems because of thin cells at the edges of the soil-covered valley, the model area was reduced. The blue arrows indicate the groundwater flow direction in the confined aquifer. Several subsurface constructions, including the metro, a road tunnel and various facility tunnels, drain the area. Since the location of a few of these are confidential, their locations are not illustrated. Drainage volumes into these constructions are unknown.

Available data

The soil stratigraphy and geotechnical properties of the area have previously been investigated for different infrastructure and building projects. During the design phase of the City Link Tunnel, a large amount of this data was compiled, and additional geotechnical drillings were carried out (Sundell et al. 2016; Sundkvist 2015).

Groundwater head observations from 16 wells in the confined aquifer and three from wells in bedrock are used as calibration data (Fig. 7). During a 4-year period, 2013–2017 most of these wells were measured monthly. During this period, observation wells in the western area close to the Stockholm Esker (heads >13 masl in Fig. 7) indicated small fluctuations, less than 0.5 m. Historical measurements, starting from 1974 in some observation wells, demonstrate larger fluctuations (about 1.5 m) for long periods. In the eastern part, the fluctuations were about 1.5 m in most observation wells during the 4-year period. In the model, all the observation wells in soil are located at layer 3 (see Appendix section ‘Geometry’). Groundwater levels are generally about 5 m below the surface as a result of permeable material in the esker together with existing drainage from tunnels and constructions.

Groundwater observations in the bedrock also demonstrate fluctuations of about 1.5 m. Selected observation measurements for calibration (Table 2), approximate to the average values during the 4-year period. Since the piezometric head in bedrock varies with depth and the observation wells are open boreholes, the observations are assumed to represent a section in each well where a conductive fracture zone is present.

Different hydrogeological tests (Table 2) including slug tests, short-period (few hours) and longer (few days) pump and infiltration tests were performed in three wells in soil (13CW411Rb, 13CW421Rb and 15SW01R) and three wells in bedrock (13CW425Hb, 13CW424Hb and 13CW423Hb). The tests are described in Erlandsson (2014) and the results of analytical evaluations are presented in Björgúlfsson (2014); Sundkvist (2015) and Berzell and Dehkordi (2017). Two tests, 13CW411Rb and 13CW421Rb indicate higher values of K than what could be expected in the indicated fine-grained material. Several of the long-term tests were problematic to evaluate due to a partly uncontrolled test environment in the urban location. Although the SSMC procedure offers fixed-value PPs, this opportunity has not been used since the reliability of individual tests is questionable. Instead, the results are used for prior parameter estimates (see Appendix section ‘Material properties’) of the different materials together with the calculations in the hydrogeological reports for City Link (Sundkvist 2015) and the Metro extension (Berzell and Dehkordi 2017).

Recharge and drainage conditions

The annual average precipitation in the area for the period 1981–2010 is 599 mm, with evapotranspiration of 276 mm, resulting in a runoff of 323 mm (SMHI 2018). For greenfield areas, the average annual recharge to coarse-grained material in the region is estimated at about 300 mm (Rodhe et al. 2006). Most of the natural recharge is expected to take place from the higher altitude areas with glacial till and esker material. Since the model area (0.2 km²) is less than one-third of the watershed (0.71 km²), it is reasonable to assume a significant flow from adjacent areas into the model area, in particular from the esker. An urban setting can locally imply reduced natural recharge as a result of impervious surfaces and stormwater drainage. Recharge in urban areas is often significantly increased by leakage of imported water in fresh water and sewage water distribution systems. Lerner (2002) argues that the complexity and lack of data for these systems make it difficult to make accurate estimations of urban recharge, which is also the case for the model area. Nevertheless, leakage from a distribution system in Sweden is in general about 25% (Statistics Sweden 2017) and with a total number of inhabitants in the area of approximately 50,000 with an average daily consumption of 288 l/sec (Stockholms stad 2017), the leakage in the area approximates to 40 l/sec. Since K in the esker material is expected to be high and the groundwater levels are low, the recharge potential is high. Discharge from the area occurs to the south-east and finds its way into the sea (Fig. 7). Information on leakage rates into the specific tunnels is not available. Flow rates into tunnels in bedrock in Stockholm are in the range 8–10 L/min per 100-m tunnel (case M 2772–15, Land and Environment Court at the District Court in Nacka, 30-11-2016) and vary depending on the sealing, depth, and presence of conductive fractures. Since the groundwater levels are low in the area, small volumes of groundwater evapotranspiration are expected.

Model and assumptions

Geometry

The stratigraphy is modified from the model presented in Sundell et al. (2016) with three layers in soil and one in bedrock. These layers are transferred to a finite-difference grid with nine layers. In this grid, the smallest cells are 10 × 10 m in the central part of the model area and up to 40 m wide in the distal western and eastern parts (Fig. 9). Since both the critical risk area and the planned tunnels are located in the central part, this area has a finer resolution. In the four uppermost layers, both bedrock and soil can be present (Fig. 8). The uppermost layer consists mostly of filling material forming an open aquifer, the second layer of clay forms an aquitard, and the third layer of coarse-grained material forms a confined aquifer. In

layer 4, the uppermost bedrock is assumed to be more fractured than the less fractured bedrock at deeper levels (layers 5–9), whereas in layers 5–9, fracture zones are included. These are also present in cells where materials with “uppermost bedrock” are indicated but are not included in the model since no significant difference in K is assumed between these materials. The sealed zone is not included in the calibration process for the model, only in the models for alternatives.

Boundary conditions

In the western part, cells with constant heads (CHD) in layer 3 represent a flow into the model from the Stockholm esker (Fig. 9). Due to the highly conductive esker with significant storage, these levels are not expected to be affected by future tunnels at the central part of the model. The inflow from the CHD boundary depends on the K of the materials in the esker. CHD in the eastern part represents the outflow from the model. The groundwater levels in the eastern part are low and are controlled by the nearby lake, and the assumption is that they will not be affected significantly by a future tunnel. The CHD boundaries are assigned from an initial simulation with good calibration results from a model covering a larger area. Existing tunnels are modelled as three drains in bedrock layers 8–9 with a total length of 1,020 m. The prior conductance value before the initial calibration for the tunnels is assumed to range between 10^{-11} and 10^{-6} m²/s/m.

Groundwater recharge

As a result of the different recharge sources (precipitation, flow from adjacent areas and leakage), the recharge is expected to vary significantly at different locations within the model. Pilot points every 30 m model this spatial variability. The limits for recharge in the prior model are set at between 10^{-10} and 10^{-7} m/s (32–3,154 mm/yr). A range higher than both the precipitation and the runoff is motivated by possible leakage from water/sewage distribution systems at PPs. Evapotranspiration is simulated using the EVT package with an extinction depth of 0.5 m in the top layer and with a maximum rate of 1.3×10^{-8} m/s (410 mm/yr).

Material properties

Maximum and minimum values of K for the different materials in the prior model are presented in Table 3. The test results and general knowledge of the expected variability within the different materials limit the values. The rather high limit for clay is justified as it is assumed that the layer will be penetrated with piles and old investigations together with a significant interspersed of more coarse-grained materials. Pilot points every 30 m are used for all materials where

variation in properties is of significance to the simulation of effects of different design alternatives. Only the materials that are a dominant response from the system are assigned PPs. Since the leakage into the tunnel is dominated by the fracture zones, the less fractured bedrock is not modelled using PPs. Because the filling material is often dry and of less importance to critical drawdown in layer 3, this material is not modelled using PPs. An anisotropy factor, $K_h/K_v = 3$ is assumed between horizontal and vertical K for soil layers and the uppermost bedrock.

Water balance: accepted SSMC solutions

The resulting water balance for the accepted solutions is presented in the histograms below (Fig. 13). If an individual simulation has high K values at the CHD boundary, representing the esker through a connective field of high K values to the outflow CHD boundary, the resulting inflow from the esker will be high. If high recharge rates from leakages of imported water into the esker cannot be excluded, resulting simulations with a high CHD inflow cannot be excluded. Pilot points representing RCH are in the range of natural recharge to the model and modelled drainage into tunnels falls within the range of expected flow into tunnels in crystalline bedrock. Since groundwater heads in general are several metres below the ground surface, evapotranspiration rates are negligible and are not demonstrated in the graphs.

Water balance: design alternatives

The difference in water balance between the calibrated models and each alternative is presented in the histograms below (Fig. 14). As can be seen for A_0 and A_1 , the CHD inflow is increased whereas the CHD outflow is reduced because of increased drainage from the planned tunnel. In Appendix section 'Available data', the opposite occurs because of the increase in water flowing into the model from infiltration.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Aquaveo (2017) Groundwater modeling system version 10.3 (version 10.3). Aquaveo, Provo, UT. <http://www.aquaveo.com/software/gms-groundwater-modeling-system-introduction>. Accessed November 2018
- Berzell A, Dehkordi E (2017) Miljöprovning för tunnelbana från Kungsträdgården till Nacka och söderort - Bilaga C - PM hydrogeologi (2320-M25-22-00001) [Environmental testing for the subway from Kungsträdgården to Nacka and South: Annex C—PM hydrogeology]. http://www.nyatunnelbanan.sll.se/sites/tunnelbanan/files/PM%20Hydrogeologi_1.pdf. Accessed September 2017
- Beven K (2006) A manifesto for the equifinality thesis. *J Hydrol* 320(1–2):18–36. <https://doi.org/10.1016/j.jhydrol.2005.07.007>
- Björgulfsson P (2014) City link stage 2: hydrogeological investigations evaluation report, PM_G0014: appendix B, S. Kraftnät, Sundbyberg, Sweden
- Bryan R, Mark S, Daniel P, Piyush A, Romain J (2018) Quantifying ground deformation in the Los Angeles and Santa Ana coastal basins due to groundwater withdrawal. *Water Resour Res* 54(5):3557–3582. <https://doi.org/10.1029/2017WR021978>
- Burbey TJ (2002) The influence of faults in basin-fill deposits on land subsidence, Las Vegas Valley, Nevada, USA. *Hydrogeol J* 10(5): 525–538. <https://doi.org/10.1007/s10040-002-0215-7>
- Burrows W, Doherty J (2015) Efficient calibration/uncertainty analysis using paired complex/surrogate models. *Groundwater* 53(4):531–541. <https://doi.org/10.1111/gwat.12257>
- Carrera J, Neuman SP (1986) Estimation of aquifer parameters under transient and steady state conditions: 2. uniqueness, stability, and solution algorithms. *Water Resour Res* 22(2):211–227. <https://doi.org/10.1029/WR022i002p00211>
- Carrera J, Alcolea A, Medina A, Hidalgo J, Slooten LJ (2005) Inverse problem in hydrogeology. *Hydrogeol J* 13(1):206–222. <https://doi.org/10.1007/s10040-004-0404-7>
- Doherty J (2003) Ground water model calibration using pilot points and regularization. *Ground Water* 41(2):170–177. <https://doi.org/10.1111/j.1745-6584.2003.tb02580.x>
- Doherty J (2011) Modeling: picture perfect or abstract art? *Ground Water* 49(4):455–455. <https://doi.org/10.1111/j.1745-6584.2011.00812.x>
- Doherty J, Hunt RJ (2010) Approaches to highly parameterized inversion: a guide to using PEST for groundwater-model calibration. US Geological Survey, Reston, VA
- Erlandsson M (2014) RA_G0012: ground investigation report (RA_G0012). S. Kraftnät, Sundbyberg, Sweden
- Fienen MN, Muffels CT, Hunt RJ (2009) On constraining pilot point calibration with regularization in PEST. *Ground Water* 47(6):835–844. <https://doi.org/10.1111/j.1745-6584.2009.00579.x>
- Fienen MN, D’Oria M, Doherty JE, Hunt RJ (2013) Approaches in highly parameterized inversion: bgaPEST, a Bayesian geostatistical approach implementation with PEST— documentation and instructions (7-C9). US Geological Survey. <http://pubs.er.usgs.gov/publication/tm7C9>. Accessed October 2017
- Freeze RA, Massmann J, Smith L, Sperling T, James B (1990) Hydrogeological decision analysis: 1. a framework. *Ground Water* 28(5):738–766. <https://doi.org/10.1111/j.1745-6584.1990.tb01989.x>
- Freeze RA, James B, Massmann J, Sperling T, Smith L (1992) Hydrogeological decision analysis: 4. the concept of data worth and its use in the development of site investigation strategies. *Groundwater* 30(4):574–588. <https://doi.org/10.1111/j.1745-6584.1992.tb01534.x>
- Harbaugh AW (2005) MODFLOW-2005, the US Geological Survey modular ground-water model: the ground-water flow process. US Geological Survey Reston, VA
- Hunt RJ, Zheng C (2012) The current state of modeling. *Ground Water* 50(3):330–333. <https://doi.org/10.1111/j.1745-6584.2012.00936.x>
- Inman HF, Bradley EL (1989) The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. *Commun Stat Theory Methods* 18(10):3851–3874. <https://doi.org/10.1080/03610928908830127>

- Karlsruh K (1999) General aspects of transportation infrastructure. In: Proceedings of the 12th European Conference on Soil Mechanics and Geotechnical Engineering, Amsterdam, June 1999
- Kitanidis PK (2015) Persistent questions of heterogeneity, uncertainty, and scale in subsurface flow and transport. *Water Resour Res* 51(8):5888–5904. <https://doi.org/10.1002/2015WR017639>
- Larsson R, Sällfors G (1986) Automatic continuous consolidation testing in Sweden. In: Yong RN, Townsend FC (eds) Consolidation of soils: testing and evaluation, ASTM STP 892. American Society for Testing and Materials, West Conshohocken, PA, pp 299–328
- LeGrand HE, Rosén L (2000) Systematic makings of early stage hydrogeologic conceptual models. *Ground Water* 38(6):887–893. <https://doi.org/10.1111/j.1745-6584.2000.tb00688.x>
- Lerner DN (2002) Identifying and quantifying urban recharge: a review. *Hydrogeol J* 10(1):143–152. <https://doi.org/10.1007/s10040-001-0177-1>
- Li L, Zhang M (2018) Inverse modeling of interbed parameters and transmissivity using land subsidence and drawdown data. *Stoch Env Res Risk A* 32(4):921–930. <https://doi.org/10.1007/s00477-017-1396-x>
- Matheron G (1963) Principles of geostatistics. *Econ Geol* 58(8):1246–1266
- Moore C, Doherty J (2005) Role of the calibration process in reducing model predictive error. *Water Resour Res* 41(5). <https://doi.org/10.1029/2004WR003501>
- Niswonger RG, Panday S, Ibaraki M (2011) MODFLOW-NWT: a Newton formulation for MODFLOW-2005 (6-A37). <http://pubs.er.usgs.gov/publication/tm6A37>. Accessed March 2018
- Olofsson B (1994) Flow of groundwater from soil to crystalline rock. *Appl Hydrogeol* 2(3):71–83. <https://doi.org/10.1007/s100400050052>
- Ortega-Guerrero A, Rudolph DL, Cherry JA (1999) Analysis of long-term land subsidence near Mexico City: field investigations and predictive modeling. *Water Resour Res* 35(11):3327–3341. <https://doi.org/10.1029/1999WR900148>
- Persson L (1998) Engineering geology of Stockholm, Sweden. *Bull Eng Geol Environ* 57(1):79–90. <https://doi.org/10.1007/s100640050024>
- PEST (2018) PEST homepage. www.pesthomepage.org. Accessed November 2018
- Phien-vej N, Giao PH, Nutalaya P (2006) Land subsidence in Bangkok, Thailand. *Eng Geol* 82(4):187–201. <https://doi.org/10.1016/j.enggeo.2005.10.004>
- Reilly T (2001) System and boundary conceptualization in ground-water flow simulation. https://pubs.usgs.gov/twri/twri-3_B8/pdf/twri_3b8.pdf. Accessed 2017-12-05
- Rodhe A, Lindström G, Rosberg J, Pers C (2006) Grundvattenbildning i svenska typjordar [Groundwater formation in Swedish soil type]. <ftp://www.w-program.nu/Allan/Grundvattenbildning%20i%20svenska%20typjordar%20Okt%2006.pdf>. Accessed May 2017
- Rossi PM, Ala-aho P, Doherty J, Kløve B (2014) Impact of peatland drainage and restoration on esker groundwater resources: modeling future scenarios for management. *Hydrogeol J* 22(5):1131–1145. <https://doi.org/10.1007/s10040-014-1127-z>
- Siade AJ, Hall J, Karelse RN (2017) A practical, robust methodology for acquiring new observation data using computationally expensive groundwater models. *Water Resour Res* 53(11):9860–9882. <https://doi.org/10.1002/2017WR020814>
- SMHI (2018) Vattenwebb [Water web]. <https://www.smhi.se/klimatdata/ladda-ner-data/vattenwebb>. Accessed 2018-03-27
- Statistics Sweden (2017) Environmental accounts MIR 2017: 1, water use in Sweden 2015. Statistics Sweden. https://www.scb.se/contentassets/bcb304eb5e154bdf9aad3fbcd063a0d3/mi0902_2015a01_br_miftr1701.pdf. Accessed March 2018-03
- Stockholms stad (2014) Building geological map. <http://www.stockholm.se/ByggBo/Kartor-och-lantmaterier/Bestall-kartor/Geoarkivet/>. Accessed November 2014-11
- Stockholms stad (2017) Water consumption, etc. <http://statistik.stockholm.se/images/stories/excel/b048.htm>. Accessed March 2018
- Stålhös G (1969) Stockholms omgivningar: geologi [Stockholm's surroundings: geology]. Norstedt, Stockholm
- Sun N-Z (1999) The stochastic method for solving inverse problems in groundwater modeling. Springer, Dordrecht, The Netherlands, pp 141–193
- Sundell J (2016) Risk estimation of groundwater drawdown in subsidence sensitive areas. Engineering degree, Chalmers University of Technology, Gothenburg, Sweden. <http://publications.lib.chalmers.se/publication/236827-risk-estimation-of-groundwater-drawdown-in-subsidence-sensitive-areas>. Accessed December 2018
- Sundell J, Haaf E, Norberg T, Alén C, Karlsson M, Rosén L (2017) Risk mapping of groundwater-drawdown-induced land subsidence in heterogeneous soils on large areas. *Risk Analysis*. <https://doi.org/10.1111/risa.12890>
- Sundell J, Rosén L, Norberg T, Haaf E (2016) A probabilistic approach to soil layer and bedrock-level modelling for risk assessment of groundwater drawdown induced land subsidence. *Eng Geol* 203(Spec. issue):126–139. <https://doi.org/10.1016/j.enggeo.2015.11.006>
- Sundkvist U (2015) City Link etapp 2: PM Hydrogeologi—Underlag för tillståndsprövning enligt miljöbalken [City link stage 2: PM hydrogeology—submission for permission testing according to the environmental code]. <http://wwwsvkse/siteassets/naturutveckling/utbyggnadsprojekt/stockholmsstrom/dokument/anneberg-skanstull/ansokan-till-mark%2D%2Doch-miljodomstolen/bilaga-5-pm-hydrogeologi.pdf>. Accessed 2017-09-20
- Tonkin M, Doherty J (2005) A hybrid regularized inversion methodology for highly parameterized environmental models. *Water Resour Res* 41(10). <https://doi.org/10.1029/2005WR003995>
- Tonkin M, Doherty J (2009) Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques. *Water Resour Res* 45(12). <https://doi.org/10.1029/2007WR006678>
- Woodward SJR, Wöhling T, Stenger R (2016) Uncertainty in the modelling of spatial and temporal patterns of shallow groundwater flow paths: the role of geological and hydrological site information. *J Hydrol* 534:680–694. <https://doi.org/10.1016/j.jhydrol.2016.01.045>
- Xue Y-Q, Zhang Y, Ye S-J, Wu J-C, Li Q-F (2005) Land subsidence in China. *Environ Geol* 48(6):713–720. <https://doi.org/10.1007/s00254-005-0010-6>
- Zetterlund M, Norberg T, Ericsson LO, Norrman J, Rosén L (2015) Value of information analysis in rock engineering: a case study of a tunnel project in Äspö Hard Rock Laboratory. *Georisk: Assess Manag Risk Eng Syst Geohazards* 9(1):9–24. <https://doi.org/10.1080/17499518.2014.1001401>
- Zhu L, Gong H, Li X, Wang R, Chen B, Dai Z, Teatini P (2015) Land subsidence due to groundwater withdrawal in the northern Beijing Plain, China. *Eng Geol* 193:243–255. <https://doi.org/10.1016/j.enggeo.2015.04.020>