

Optimization for Early-Warning Monitoring Networks in Well Catchments Should Be Multi-objective, Risk-Prioritized and Robust Against Uncertainty

Felix Bode¹ · Wolfgang Nowak¹ · Matthias Loschko²

Received: 4 April 2015 / Accepted: 1 October 2015 / Published online: 23 November 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Groundwater abstraction wells are commonly protected by zones of restricted land use. Such well protection areas typically cannot cover the entire well catchment, and numerous risk sources remain. Each risk source could release contaminants at any time, affect the well earlier or later, and thus put the quality of supplied water at risk. In this context, it seems fortunate that most well catchments are equipped with monitoring networks. Such networks, however, often grew historically while following diverse purposes that changed with time. Thus, they are often inadequate (or at least suboptimal) as reliable risk control mechanism. We propose to optimize existing or new monitoring networks in a multi-objective setting. The different objectives are minimal costs, maximal reliability in detecting recent or future contaminant spills, and early detection. In a synthetic application scenario, we show that (1) these goals are in fact competing, and a multi-objective analysis is suitable, (2) the optimization should be made robust against predictive uncertainty through scenario-based or Monte Carlo uncertainty analysis, (3) classifying the risk sources (e.g., as severe, medium, almost tolerable) is useful to prioritize the monitoring needs and thus to obtain better compromise solutions under budgetary constraints, and (4) one can defend the well against risk sources at unknown locations through an adequate model for the residual risk. Overall, the concept brings insight into the costs of reliability, the costs of early warning, the costs of uncertainty, and into the trade-off between covering only severe risks versus the luxury situation of controlling almost tolerable risks as well.

✉ Felix Bode
felix.bode@iws.uni-stuttgart.de

Wolfgang Nowak
wolfgang.nowak@iws.uni-stuttgart.de

Matthias Loschko
matthias.loschko@uni-tuebingen.de

¹ Modelling Hydraulic and Environmental Systems, University of Stuttgart, Pfaffenwaldring 5a, 70569 Stuttgart, Germany

² Centre for Applied Geosciences, University of Tuebingen, Hoelderlinstr. 12, 72074 Tuebingen, Germany

Keywords Optimal monitoring · Well protection · Detection probability · Early-warning time · Multi-objective optimization · Uncertainty · Risk

1 Introduction: Motivation and State of the Art in Monitoring Well Catchments

Ninety-seven percent of the world's usable freshwater is stored as groundwater. Groundwater abstraction wells produce a major share of the worldwide daily water demand (Sampat 2000), but they are often threatened by a large number of hazardous land use activities (called *risk sources* in the following). Each risk source may or may not release contaminants into the aquifer, and thus put the produced water quality at risk. The World Health Organization proposed Water Safety Plans to control this risk (Davison et al. 2005) by (1) knowing the entire inventory of risks within the well catchment, (2) identifying measures to control these risk sources, and (3) ensuring that they are in fact controlled.

For risk reduction, the most common measure is to restrict the land use within well protection zones (USEPA 1993). Yet, it is typically not possible to remove all risks through declaring protection zones due to two reasons: (1) Restricting the land use everywhere in a large well catchment is often impossible, especially in urban regions, and (2) there is uncertainty in the actual outline of the well catchment. Thus, there is always an inventory of risk sources in the catchment that remains to be assessed and controlled. Accordingly, there is ample literature on capture zone delineation and its uncertainty (e.g., Varljen and Shafer 1991; Jacobson et al. 2002; Moutsopoulos et al. 2008; Stauffer et al. 2005), on aquifer vulnerability (e.g., Aller et al. 1987; Overheu et al. 2014; Zwahlen 2003), on (probabilistic) well vulnerability (e.g., Enzenhoefer et al. 2012; Frind et al. 2006; Enzenhoefer et al. 2014), and on risk analysis (e.g., Cushman et al. 2001; Tartakovsky 2013). While such works help to *evaluate* the risk sources to which the production well is exposed, it is not yet helpful in *controlling* them.

In order to track the quality of groundwater prior to pumping, most well catchments are equipped with monitoring networks (at least in countries like Germany and Denmark). Such networks often grew historically while following diverse purposes that changed with time (e.g., monitoring groundwater levels or risks that ceased to exist). Therefore, unfortunately, they are inadequate (or at least suboptimal) for rigorous and cost-efficient risk control. A prioritization according to the severity of the perceived risk is sometimes done implicitly (through adding monitoring wells to monitor the risk sources that are currently perceived as the most severe ones), but seldom in a coordinated fashion. In situations like this, a re-optimization of the existing monitoring networks can be advisable. Opposed to that, there are well catchments in other countries without any monitoring networks for controlling risk sources. Absence of monitoring networks often coincides with the absence of corresponding budgets or regulations. However, it is in strong contrast with the recommendation of the World Health Organization to use a risk control structure (see above). In order to initialize cost-minimal risk control in such cases, again, optimization of monitoring networks is advisable.

Thus, in any case, one could achieve or increase reliability and cost efficiency of monitoring-based risk control through formal, risk-prioritized optimization. When doing so, it is apparent that the goals of monitoring are manifold and often competing (e.g., Reed and Minsker 2004). In the current context, they should include at least the following three objectives: (1) a maximal detection probability of contaminant spills, because otherwise the risks cannot be controlled at a high reliability level, (2) a maximal early-warning time

(i.e., earliest possible detection of new contaminant spills after they have occurred), because countermeasures such as installing additional water treatment steps require time for implementation, (3) minimal costs for installation and operation of the monitoring network. In the following, these three objectives will be called *detection probability*, *early-warning time*, and *costs*. It is intuitive that these objectives are partly competing (a cheap monitoring network cannot offer a high detection probability and a large early-warning time). Hence, concepts of multi-objective optimization (e.g., [Marler and Arora 2004](#)) seem appropriate.

Along the lines of long-term groundwater quality monitoring (see review by [Loaiciga et al. 1992](#)), there is a large body of literature on multi-objective groundwater quality monitoring (e.g., [Reed and Minsker 2004](#)). The focus of these studies is mostly on long-term water quality monitoring or on monitoring the evolution of a single, known contaminant plume, but not on early warning. Typical objectives include the estimation of plume shapes (e.g., [MacFarlane et al. 1983](#)), estimation of total contaminant fluxes (e.g., [Schwede and Cirpka 2010](#); [Trolborg et al. 2008](#)), and contaminant source identification ([Saenton and Illangasekare 2004](#); [Michalak and Kitanidis 2004](#)).

Monitoring networks as early-warning systems are often associated with disaster management for natural hazards such as tsunami or earthquake warning systems, e.g., [Allen and Kanamori \(2003\)](#). Alternatively, there is research on detection sensor networks in the signal processing literature (e.g., [Chamberland and Veeravalli 2003](#)), but without existing applications to well catchments. In the area of hydrogeology, there are two relevant studies that investigate the monitoring of a single landfill (e.g., [Yenigül et al. 2006](#); [Meyer et al. 1994](#)). Both studies use the objectives of high detection probability and early warning. The first study, however, works with a cost formulation for increased groundwater remediation costs after late detection, while the second study minimizes the area covered by the contaminant plume before the plume is detected. [Meyer et al. \(1994\)](#) also minimized the installation costs of a monitoring network counted through the number of required monitoring wells, which is the same approach as we use in our current study. The necessity to consider robustness of their monitoring networks against uncertainty in hydrogeology and spill location is done in both previous studies through Monte Carlo simulation over hydraulic conductivity and potential leak locations. In our current approach, we also consider uncertain ambient flow conditions and water production rates. Additionally, we propose an approach for treating contamination spills at entirely unknown locations. The major difference between our current approach and these two studies, however, is that these studies do not consider the protection of a groundwater well against an inventory of risk sources. Instead, they focus on monitoring of a single landfill. Thus, to the best knowledge of the authors, there is no study to date that approaches the problem of robust optimal monitoring for multiple risk sources as an early-warning system in drinking water well catchments. In [Nowak et al. \(2015\)](#), the authors already described the general concept behind our current approach, yet without any implementation, without application to a synthetic test case, and without the concept for entirely unknown risk sources. Also, we refine now the definition of our objective functions.

2 Goals and Approach

Our overall goal is to provide risk control in well catchments through optimal early-warning monitoring networks. We wish to detect with maximum probability all possible future contaminations that might affect the well's water quality. Additionally, we wish to provide maximal

early-warning time for installing counteractions. Finally, we seek minimal costs for installing and operating the monitoring network.

As already described in Nowak et al. (2015), the considered three objectives are competing. Multi-objective optimization will reveal the competition and trade-offs between the goal attainment levels in these three objectives, and will provide a transparent decision basis. For the optimization procedure, the first step is to formulate a corresponding multi-objective optimization problem (see Sect. 3.1).

In a second step, scenario analyses or Monte Carlo simulation should be used to represent predictive uncertainty. Uncertainty in predicting the yet nonexistent contaminant transport from the risk sources to the drinking water well is substantial. Appropriate treatment of these uncertainties and a corresponding robust formulation of the multi-objective optimization problem will make the optimization results robust against the considered uncertainties. The affected objective functions need to be formulated according to well-known rules of robust optimization (see Sect. 3.2).

In a third step, qualitative risk categorization is used. Sufficient monitoring is not practically affordable, if the risk inventory of the featured well catchment is large. A categorization of the risk sources according to their severity and a corresponding re-formulation of the optimization problem will help to prioritize and better focus the monitoring network. Therefore, separate optimization objectives for each risk category need to be introduced (see Sect. 3.3).

The last step is to introduce an additional risk category for unknown risks together with corresponding additional optimization objectives. This is necessary because knowledge of the risk inventory is unlikely to be complete, especially in densely populated areas. The early-warning system can be made robust against the residual risk only through an appropriate representation in the optimization (see Sect. 3.4).

After presenting this approach in Sect. 3, we present in Sect. 4 a synthetic application scenario and provide implementation details. Sect. 5 uses variants of the application scenario in order to demonstrate stepwise the benefits of our proposed approach (see above), followed by a conclusion and outlook in Sect. 6.

3 The Proposed Concept and Methods in Detail

Our approach assumes that a sufficiently well-calibrated simulator for flow and transport in the well catchment is available. For the treatment of uncertainty and unresolved heterogeneity, please see Sects. 3.2 and 5.2. In practical applications, such a simulator will typically solve the groundwater flow equation at steady state, and the advection–dispersion equation. Steady-state groundwater flow can be described by

$$-\nabla \cdot (K \nabla \phi) = q_s \quad \text{in } \Omega, \tag{1}$$

with hydraulic conductivity $K(\mathbf{x})$, hydraulic head ϕ , and source and sink term q_s in the domain Ω . Eq. 1 is subject to the general boundary conditions:

$$-(K \nabla \phi) \cdot \mathbf{n} = \hat{q} \quad \text{on } \Gamma_1 \text{ and} \tag{2}$$

$$\phi = \hat{\phi} \quad \text{on } \Gamma \setminus \Gamma_1, \tag{3}$$

using the prescribed fluxes \hat{q} and heads $\hat{\phi}$ on the Neumann boundary Γ_1 and on the Dirichlet boundary $\Gamma \setminus \Gamma_1$. The normal vector \mathbf{n} points outward on the domain. Advective–dispersive transport for conservative tracers is described by:

$$\frac{\partial c}{\partial t} + \nabla \cdot (\mathbf{vc} - \mathbf{D}\nabla c) = 0 \quad \text{in } \Omega, \tag{4}$$

with concentration c , time t , velocity $\mathbf{v} = \mathbf{q}/n_e$, Darcy velocity \mathbf{q} , effective porosity n_e and hydromechanic dispersion tensor \mathbf{D} (Scheidegger 1954):

$$\mathbf{D} = (\alpha_t \|\mathbf{v}\| + D_m) \mathbf{I} + (\alpha_\ell - \alpha_t) \frac{\mathbf{v}\mathbf{v}^T}{\|\mathbf{v}\|}. \tag{5}$$

Here, α_ℓ and α_t are longitudinal and transversal dispersivities, D_m is the molecular diffusion coefficient, and \mathbf{I} is the identity matrix. The boundary conditions for Eq. 4 are given by:

$$-\mathbf{n} \cdot \mathbf{vc} + \mathbf{n} \cdot (\mathbf{D}\nabla c) = \hat{J} \quad \text{on } \Gamma_2 \text{ and} \tag{6}$$

$$c = \hat{c} \quad \text{on } \Gamma \setminus \Gamma_2, \tag{7}$$

with \hat{J} as a prescribed normal flux density and \hat{c} as prescribed concentrations.

We also assume that there is a list of known locations where contaminant spills could occur. These risk sources will typically include agriculture, housing areas with oil tanks, sewer systems and wastewater treatment plants, and industrial sites.

First, using the available simulator, we solve a reverse transport problem (Neupauer and Wilson 2002) that starts at the production well. The resulting backward plume identifies the transport-relevant catchment of the well that needs to be monitored. Then, we discretize the identified catchment with a fine spatial mesh of potential monitoring well locations M_j , $j = 1, \dots, n_M$. For each potential location, the drilling costs should be known. All risk sources R_i , $i = 1, \dots, n_R$ within the identified catchment are relevant in the following. We assume that conservative estimates of contaminant mass m_i in case of spill events are available for each relevant risk source R_i , and that the corresponding chemical detection limit c_i^{det} is known.

Second, starting from all relevant risk sources, we solve instantaneous-release forward transport problems to simulate the contaminant plumes that would emerge in case a risk source actually released contaminants into the subsurface. Following the discussion of risk estimation in well catchments in Enzenhofer et al. (2015), we distinguish here the different spatial extent of risk sources (e.g., point sources, line sources and areal sources versus point sources with uncertain position along a line or within an area) through corresponding source geometries. From all these transport simulations, we store the following data:

$$\delta_{ij} = \begin{cases} 1 & \text{a contamination released at risk source } R_i \text{ exceeds the} \\ & \text{detection limit } c_i^{\text{det}} \text{ at monitoring candidate location } M_j \\ 0 & \text{no detectable concentration at } M_j. \end{cases} \tag{8}$$

If $\delta_{ij} = 1$, we can also obtain the travel time τ_{ij}^{det} between contaminant release at risk source R_i and first exceedance of the detection limit c_i^{det} at monitoring candidate M_j . We also extract a travel time $\tau_{i,\text{well}}$ for first arrival of a critical concentration c_i^{crit} at the production well. The difference yields the respective early-warning time t_{ij} :

$$t_{ij} = \begin{cases} \max(\tau_{i,\text{well}} - \tau_{ij}^{\text{det}}, 0) & \text{early-warning time (if } \delta_{ij} = 1) \\ 0 & \text{no early warning } (\delta_{ij} = 0). \end{cases} \tag{9}$$

Further, we extract the duration $\Delta t_{ij}^{\text{vis}}$ for which the contaminant plume from risk source R_i is visible at a monitoring candidate M_j , i.e., the duration for which the simulated plume from R_i exceeds the detection limit c_i^{det} at the position of M_j .

3.1 Multi-objective Optimization Formulation

We formulate the multi-objective optimization problem at hand as follows:

$$\mathbf{d}_{\text{opt}} = \arg \min_{\mathbf{d} \in \mathbf{D}} [f_{\text{det}}, f_{\text{warn}}, f_{\text{cost}}]. \tag{10}$$

Here, \mathbf{d}_{opt} is the optimal set of decision variables \mathbf{d} . The decision variables \mathbf{d} characterize the planned monitoring system (e.g., number and positions of monitoring wells, filtering depth and window, frequency of sampling), \mathbf{D} is the space of allowable designs (e.g., restricted to accessible positions within the catchment, maximum admissible installation and operation costs), and f_{det} , f_{warn} and f_{cost} are the objective functions that assign goal attainment levels to each design $\mathbf{d} \in \mathbf{D}$. These three objective functions are explained in the following.

Detection probability The values δ_{ij} from Eq. 8 express, which contaminant plumes from the risk sources R_i can in principle be detected by the monitoring wells M_j . However, the monitoring wells are only sampled in time intervals of $\Delta t_j^{\text{sample}}$, while the plumes exceed the detectable concentration c_i^{det} only for a duration $\Delta t_{ij}^{\text{vis}}$. This fact causes a probability that a plume may pass unnoticed through a monitoring well position due to unfortunate timing. To account for this effect, we correct from δ_{ij} to a probability of detection P_{ij}^{det} as follows:

$$P_{ij}^{\text{det}}(\mathbf{d}) = \min \left(\frac{\Delta t_{ij}^{\text{vis}}}{\Delta t_j^{\text{sample}}(\mathbf{d})}, 1 \right) \cdot \delta_{ij}. \tag{11}$$

The aggregation of the obtained probabilities P_{ij}^{det} over all risk sources $R_i, i = 1, \dots, n_R$ and over all monitoring wells suggested from the candidates $M_j, j = 1, \dots, n_M$ by any given design \mathbf{d} is rather complex, and hence outsourced to the Appendix (see Eqs. 15, 16).

Early-warning time The achievable early-warning time values for different risk sources may vary between none and several decades, such that a simple linear relation between individual early-warning time values and the overall early-warning performance of the monitoring network is not suitable. Therefore, we use a nonlinear utility function (see Fig. 1), defined by a desirable minimal early-warning time \hat{t}^{min} , a corresponding early-warning utility \hat{u}^{min} , and the maximal useful early-warning time \hat{t}^{max} . For an early-warning time t for any given risk source that falls between $\hat{t} = 0$ and \hat{t}^{min} , the growth of the utility is steep. Beyond \hat{t}^{min} , the utility grows slower, until it reaches its maximum value of one at \hat{t}^{max} . Finally, for longer early-warning time than \hat{t}^{max} , there is no additional benefit. Since \hat{u}^{min} defines the steepness of the utility curve between 0 and \hat{t}^{min} , it can be seen as a catchment-specific weighting factor for early or late detection and needs to be chosen carefully. Changes in all three values (\hat{u}^{min} , \hat{t}^{min} and \hat{t}^{max}) might influence the final results more or less distinctively.

While the objective behind f_{warn} is straightforward, the mathematical formulation of f_{warn} is surprisingly complex—especially when a risk source is visible to several monitoring wells. Thus, the definition of f_{warn} is postponed to the ‘‘Appendix.’’

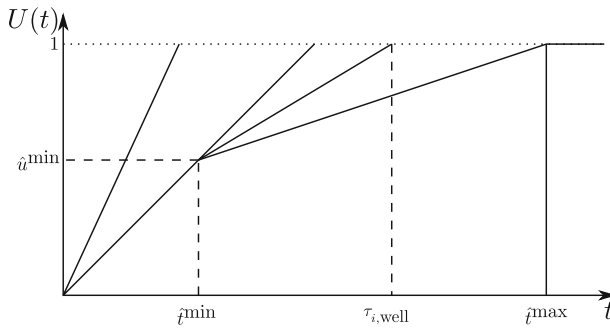


Fig. 1 Case-specific utility function $U_i(t)$ defined in Eq. 18 (see “Appendix”)

Table 1 Parameter values chosen for the objective functions and their required components defined in Eqs. 8, 9, 11 and 15–18

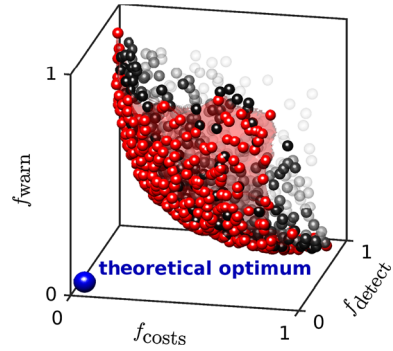
Name	Symbol	Units	Value
Sampling interval	$\Delta t_j^{\text{sample}}$	(years)	1
Non-dimensional spill mass	m_0	(–)	1
Detection limit	c_i^{det}	($1/\text{m}^3$)	$1 \cdot 10^{-12}$
Critical concentration	c_i^{crit}	($1/\text{m}^3$)	$1 \cdot 10^{-12}$
Desirable minimal early-warning time	\hat{t}^{min}	(years)	10
Maximal useful early-warning time	\hat{t}^{max}	(50)	50
Minimal desirable early-warning utility	\hat{u}^{min}	(–)	0.7

Costs The objective function for costs, f_{cost} , could be a summation of depth-specific drilling costs that are spatially variable, plus operation costs that are composed of costs per sampling round within a given budget period. For the sake of simplicity, we use only drilling costs, and the drilling depth is an identical value across the entire domain. As a slight complexity, we use spatially variable drilling costs that are linked to the zonation pattern of hydraulic conductivity. The spatial variations, however, are rather small ($\pm 10\%$) and have almost no effect on the optimization results. Therefore, in the remaining study, we mostly use the number of installed monitoring wells to express costs. Table 1 provides the relevant parameters in the objective functions chosen in our study.

Each objective is formulated as minimization problem, i.e., f_{det} is the probability of not detecting any emitted plume, f_{warn} evaluates the time lost between contaminant spill and detection, and f_{cost} specifies the costs of installation and operation. Furthermore, all objective functions are normalized to the interval $[0, 1]$, such that the theoretical optimal value in all individual objectives is zero. While f_{det} falls between zero and one by definition, we express f_{warn} through a utility function that we chose to be normalized to that interval, and the costs function f_{cost} is normalized through division by a maximal admissible cost value.

The solution of a multi-objective optimization problem (e.g., Marler and Arora 2004) is not a unique optimum, because multiple competing objectives do not infer a unique ranking. Instead, solutions are ranked according to their domination. A solution is said to be dominant over a second solution, if it is better in at least one objective, and equal in all others. The multi-objective solution is the set of non-dominated solutions, called Pareto front. The final

Fig. 2 Illustration of multi-objective optimization results for monitoring networks



decision is found by discussion and then choosing a case-specific compromise solution from the Pareto front.

Figure 2 is a graphical illustration of how the optimization results may look like. The figure shows the performances in all three objective functions for the set of possible monitoring networks considered during an optimization. Each sphere represents one possible network design and is located in the diagram according to its goal attainment levels. The theoretical optimum (blue) is idealistic: Fully certain and early detection cannot come at zero costs. The red spheres mark the Pareto front: None of them is dominated by any other sphere. All dominated solutions/designs are shown in black and drawn transparent according to their domination rank (from black to invisible for increasing rank).

3.2 Extending the Multi-objective Optimization for Model Uncertainty

The objective functions f_{det} and f_{warn} use the simulation-based data specified in Sect. 3.1. Hence, they rely on model-based predictions about flow and transport in the catchment, which are subject to uncertainty. The uncertainty arises from many sources, such as from model errors, from numerical errors, from the post-calibration uncertainty of the flow and transport simulator, from unresolved aquifer inhomogeneities, from uncertainty in the hydrogeological boundary conditions, and from different scenarios in water extraction rates and/or hydrological conditions (e.g., wet, medium, or dry seasonal conditions; high, medium, or low pumping rates at the well) (Carrera 1993). When not accounting for this uncertainty, the optimized monitoring network will most likely achieve a lower performance (when applied in practice) than wrongly predicted during the optimization. Vice versa, accounting for this uncertainty at least to some extent will make the optimized monitoring networks robust against the considered uncertainties.

Thus, our concept proposes to perform Monte Carlo simulations (e.g., using calibration-constrained Monte Carlo simulations (Tarantola 2005) or Bayesian model averaging (Hoeting et al. 1999)) or at least scenario analyses. Each scenario or Monte Carlo realization k , $k = 1, \dots, n_{MC}$ leads to its own values δ_{ijk} , τ_{ijk}^{det} , $\tau_{ik,well}$ and Δt_{ijk}^{vis} for the quantities defined in Sect. 3.1. Then, we replace the affected objective functions f_{det} and f_{warn} in Eq. 10 by adequate statistics over the objective function values $f_{det}^{(k)}$ and $f_{warn}^{(k)}$ obtained per realization:

$$\mathbf{d}_{opt} = \arg \min_{\mathbf{d} \in \mathbf{D}} \left[\tilde{f}_{det}, \tilde{f}_{warn}, f_{cost} \right], \tag{12}$$

where \tilde{f} could denote, e.g., the largest value or a high percentile (since the objective functions are formulated for a minimization problem) or the arithmetic mean. The expected value is

a risk-neutral approach to optimization under uncertainty (e.g., Faber et al. 2004), while working with extremes or percentiles is a risk averse and even more robust approach.

3.3 Extending the Multi-objective Optimization for Risk Categories

The optimization problems formulated above do not yet prioritize according to the severity of risk sources. A risk source can be more or less severe depending on the type of the hazardous activity, stored or handled contaminants and contaminant masses, compliance with the applying regulations and so forth. A common approach to prioritizing risk sources is to use weights. Provided that all required data are available to calculate these weights through quantitative and probabilistic risk estimations (e.g., Enzenhoefer et al. 2012, 2015), this approach would be statistically rigorous. Unfortunately, it is often impractical, or sometimes even impossible to get all relevant data concerning the risk sources (e.g., the probability of failure of a nuclear power plant).

Therefore, following concepts of qualitative risk assessment, we categorize all risk sources into the classes *severe*, *medium* and *almost tolerable* (Cox 2008). This categorization can be visualized through the color scheme of traffic lights, i.e., as red, yellow, and green. It expresses the prioritization preferences for monitoring in the sense that red risk sources should be monitored with first priority, yellow ones later, and green ones with the smallest priority.

In order to accommodate for this categorization in the optimization problem, we introduce separate objective functions for the detection probability and for the early-warning time of each risk category. Because individual monitoring wells can cover risk sources from several risk categories, the cost function cannot be separated. Thus, we obtain a total of seven objectives:

$$d_{opt} = \arg \min_{d \in D} \left[\tilde{f}_{det}^{(red)}, \tilde{f}_{warn}^{(red)}, \tilde{f}_{det}^{(yellow)}, \tilde{f}_{warn}^{(yellow)}, \tilde{f}_{det}^{(green)}, \tilde{f}_{warn}^{(green)}, f_{cost} \right]. \quad (13)$$

3.4 Extending the Multi-objective Optimization for Remaining Risk

The last extension of the optimization problem addresses the fact that the list of known risk sources is very likely to be incomplete. Reasons for this include the limited existence of information on land use in the private or industrial sector, restricted access to existing information due to data privacy policies, known risk sources with unknown locations (from vague historical records) or unforeseeable hazards that fall under the classical category of black swans (Taleb 2007).

Although the remaining risk is unknown in existence and location, we can only achieve robustness against its possible existence by representing it through an additional risk model within the risk inventory. The approach presented in the following addresses the total lack of knowledge about locations of potential risk sources. Uncertainty in the exact location of a known risk source can also be handled with this approach, but in a slightly different manner that is not subject to the discussion below.

As a possible modeling approach, we represent the remaining risk as a dense fence of point sources, called *line of attack* in the following, that encircle the drinking water well. The line of attack is located at a distance in travel time coordinates that can be freely chosen to fit the respective application context (e.g., the minimal time needed to install countermeasures). This is a plausible model because any possible contaminant spill beyond this line will have to pass through the line on the way to the drinking water well. For all these possible spills, it is a conservative (worst-case) representation in the following aspects (compared to the

Table 2 Transport-relevant parameters

Average conductivity K (m/s)	Porosity n (-)	Molecular diffusion coefficient D_m (m ² /s)	Longitudinal dispersivity α_ℓ (m)	Transversal dispersivity α_t (m)
$3.15 \cdot 10^{-4}$	0.35	$1 \cdot 10^{-9}$	3	0.3

The hydraulic conductivity value provided here is the arithmetic mean of all zone-wise values

actual properties of the unknown risk source represented by this model): (1) The modeled travel time from the line of attack to the well is always shorter; (2) the detectable width of the modeled plumes is always smaller; (3) the contaminant impact on the well is always larger. All residual risks between the line and the well cannot be controlled with the resulting monitoring networks. Hence, the choice of travel time to the well is a compromise between residual risk coverage and achievable early-warning time.

A different approach to representing the remaining risk would be to assume risk sources at locations randomly chosen from a uniform distribution over the entire catchment area. This technique is intuitive, but cannot guarantee to really cover all potential positions of unknown risk sources. A guaranteed control of the remaining risk can only be ensured by assuming risk sources to be everywhere in the investigated catchment, but would lead to an explosion of calculation time.

Following the same approach as in Sect. 3.3, we introduce a new color for the remaining risk (e.g., blue) and extend the optimization problem accordingly:

$$\mathbf{d}_{\text{opt}} = \arg \min_{\mathbf{d} \in \mathbf{D}} \left[\tilde{f}_{\text{det}}^{(\text{red})}, \tilde{f}_{\text{warn}}^{(\text{red})}, \tilde{f}_{\text{det}}^{(\text{yellow})}, \tilde{f}_{\text{warn}}^{(\text{yellow})}, \tilde{f}_{\text{det}}^{(\text{green})}, \tilde{f}_{\text{warn}}^{(\text{green})}, \tilde{f}_{\text{det}}^{(\text{blue})}, \tilde{f}_{\text{warn}}^{(\text{blue})}, f_{\text{cost}} \right] \tag{14}$$

4 Application Scenario and Numerical Implementation

In order to demonstrate and discuss our proposed approach, we set up a synthetic application scenario. The scenario with its flow and transport model is inspired by typical models used for well catchment management. The model describes a single geological layer of an aquifer as a quasi-three-dimensional domain (15,000m × 7000m × 10m) and is divided into few zones with different hydraulic conductivity values. All relevant model parameters are provided in Table 2, and a system sketch is shown in Fig. 3. Near the eastern boundary, there is a gallery of 15 pumping wells with 50-m spacing between neighboring wells. The figure also shows the location and classification of the considered risk sources. For simplification reasons, all risk sources are assumed to be point sources with a pulse release of their contaminant.

The regional flow direction is from east to west, defined through Dirichlet conditions at all domain boundaries. The prescribed boundary values follow geometrically from a uniform gradient that is specified through its absolute value and its orientation relative to the east–west axis. For the robustness aspect mentioned in Sect. 3.2, we work with four different hydraulic scenarios (c.f., Table 3) that differ in the strength of the regional head gradient, in its orientation angle, and in the overall pumping rate of the well gallery. These scenarios serve to represent aspects of hydrological uncertainty. They do not consider the other sources of uncertainty listed in Sect. 3.2, because we prefer to keep the application scenario straight-

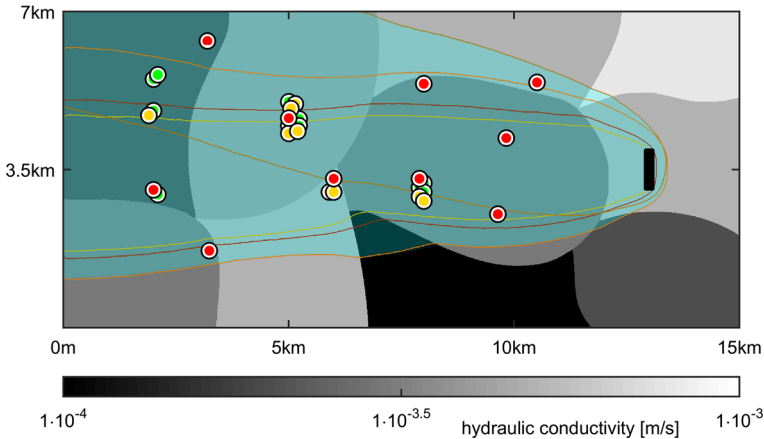


Fig. 3 Model domain for the well catchment. *Gray scale* zonation and values of hydraulic conductivity; *black rectangle* position of the well gallery; *colored lines*: catchment outlines obtained from reverse transport simulations in the four different hydraulic scenarios (see Table 3); *blue overlay* union of all scenario-wise well catchments; *colored circles* risk sources classified according to severe (*red*), medium (*yellow*), and almost tolerable (*green*)

Table 3 Definition of four different hydraulic scenarios

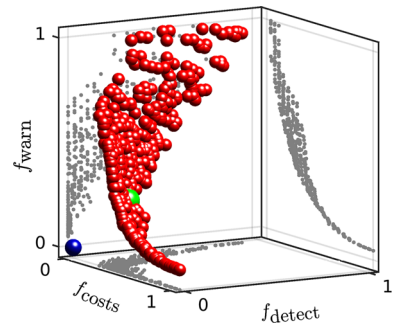
Hydraulic scenario	Gradient strength (–)	Gradient angle (°)	Pumping rate (l/s)
1	0.01	0	15
2	0.01	15	15
3	0.015	0	15
4	0.01	0	7.5

forward. The different scenarios result in travel time values from the most distant risk sources to the well of up to 80 years. Again for the sake of simplicity, we assume conservative tracer transport. This assumption may seem crude, but it is justifiable as a worst-case scenario for contaminant impact on the well.

We discretize the domain with rectangular, equispaced cells sized $10\text{ m} \times 10\text{ m} \times 10\text{ m}$ and simulate groundwater flow with the standard Galerkin finite element code already used in Nowak et al. (2008). For all subsequent transport simulations, we rely on the particle-tracking random walk (PTRW) code used earlier by Enzenhoefer et al. (2014) and by Koch and Nowak (2014). We chose PTRW due to its ease of implementation and its absence of numerical dispersion (e.g., Kinzelbach 1988; LaBolle et al. 1996; Salamon et al. 2006).

In Sect. 3, we decided to discretize the well catchment by a fine grid of candidate positions for monitoring wells. We use the same grid as for the flow simulations. Thus, the decision vector \mathbf{d} in Eqs. 10–14 includes a Boolean (yes/no) variable for each of the 1,050,000 candidate locations. For the sake of simplicity, we do not optimize but fix the sampling interval $\Delta t_j^{\text{sample}}$ to a value of one year for all monitoring wells. Choosing Boolean variables results in a discrete optimization problem. Regardless of that choice, the resulting optimization problem is multi-objective, nonlinear and high-dimensional. For our problem, we tailored a problem-specific fusion of the well-known multi-objective optimization algorithms NSGA-II (Deb et al. 2002) and Borg (Hadka and Reed 2013). Both algorithms are evolutionary algorithms. We refrain from reporting details here, because the problem-specific fusion yields

Fig. 4 Pareto front for the basic multi-objective optimization without uncertainty and without risk categorization (see Eq. 10 and Sect. 5.1). *Red spheres* Pareto-optimal solutions; *blue sphere* idealistic (unattainable) optimum; *green sphere* performance of the selected monitoring network shown in Fig. 5. *Gray dots* projections of the Pareto front onto the planes of the coordinate system



the same results as the individual algorithms, but merely does so in smaller computational time. All codes and additional implementations were written in MATLAB. The simulations and optimization runs reported below were performed on a contemporary desktop computer with an Intel Core 2 Quad CPU (2.83 GHz) and 16 GB RAM within 20 h.

5 Results and Discussion

In the following, we present the results obtained by applying our proposed method to the application scenario presented in Sect. 4. In specific, we investigate the four steps listed in Sect. 2 one by one (see Sect. 5.1 through 5.4) through corresponding modifications of the application scenario.

5.1 Multi-objective Optimization is the Appropriate Approach

The design of early-warning monitoring networks is subject to competing objectives. The first step of our approach was to formulate this problem as a multi-objective optimization problem. For demonstrating and discussing this step, we perform an optimization according to Eq. 10. In this formulation, we do not yet consider robustness against uncertainty and hence work only with hydraulic scenario 1. Also, we do not apply any risk classification.

Figure 4 shows the Pareto front obtained during this optimization variant. In comparison with Figs. 2, 4 does not include dominated solutions, but instead shows the projections of the three-dimensional Pareto front onto the planes of the coordinate system. Figure 5 shows a selected Pareto-optimal monitoring network.

Figure 4 clearly reveals that the three considered objectives are competing. One can achieve a full coverage of all risk sources at maximal early-warning utility, but only with an extreme budget that allows installing 22 monitoring wells (bottom right tip of the Pareto front). At decreasing costs, one can either maintain maximal detection probability (front/left edge of the Pareto front) while losing early-warning time, or one can try to lose less early-warning functionality but restrict that functionality to a smaller number of risk sources (back/right edge of the Pareto front). In general, there is a list of five geometric aspects in the Pareto front that will hold for any catchment: (1) The idealistic goal of benefits without costs (blue sphere in Fig. 4) can never be reached. (2) Maximizing detection probability and yet minimizing costs is only possible by directly monitoring the mixed water of the pumping well gallery (upper back/left corner of the Pareto front). Since we did not consider this case in our application, there is no solution at this part of the Pareto front. (3) The solution with almost no costs and a very late and poor detection of the risk sources (upper back/right edge of the

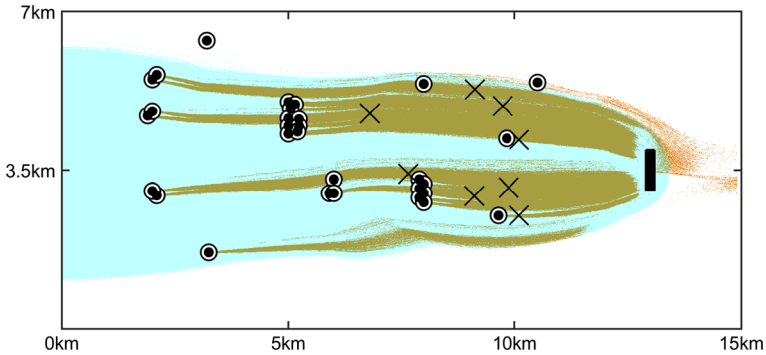


Fig. 5 One selected Pareto-optimal monitoring network from the basic multi-objective optimization without uncertainty and without risk categorization (see Eq. 10 and Sect. 5.1). *Blue area* well catchment from hydraulic scenario 1; *Black rectangle* position of the well gallery; *orange point clouds* visualization of the simulated plumes from each risk source; *black circles* risk sources without categorization; *cross markers*: optimized positions of monitoring wells. The goal attainment levels of this specific solution are shown as *green sphere* in Fig. 4

Pareto front) is to place one individual monitoring well at the cheapest individual drilling position and depends only on the complexity of f_{cost} . This selection defines the upper/back right corner of the Pareto front. (4) The 45° inclination of the back/right edge of the Pareto front is caused by a correlation between attainable early-warning time and detection probability. It originates from the detailed formulation of the objective functions (see Sect. 3.1 and “Appendix”), where we count an early-warning utility of zero for all risk sources for which the potential plumes are predicted to be undetectable. (5) The lower/front tip of the Pareto front is given by the most expensive solution that places one monitoring well just downstream of each risk source. These five facts define the fundamental geometry of the Pareto front. Only the curvature of the Pareto front toward the idealistic unattainable optimum depends on the properties of the investigated catchment and the positions of its risk sources. Risk sources arranged in a line transverse to the dominant flow direction would lead to a strong competing (and hence to a low curvature of the Pareto front), because non-overlapping plumes prohibit the monitoring of several risk sources with only a few monitoring wells. Risk sources arranged in a longitudinal line would lead to one overlapping plume, with almost no competition between costs and detection. Clustered risk sources would fully remove the competing character of all objective functions, because these risk sources could be monitored with good early-warning time by just one or a few monitoring wells very close to the spill location.

Given this information basis, the decision maker can investigate the trade-offs in performance under budgetary constraints and find a situation-specific good compromise. Hence, it is legitimate to say that multi-objective optimization is an adequate approach to finding proper early-warning monitoring networks. The selected monitoring network shown in Fig. 5 is the so-called best compromise solution (e.g., Talbi 2009), which is defined as the one Pareto-optimal solution with the smallest distance to the idealistic unattainable optimum in normalized axes (such as it is the case in Fig. 4). Through optimized positioning of only eight monitoring wells (34 % of the cost maximum), it attains a detection probability of 91 % and an early-warning utility of 0.73. This relatively high performance at relatively low costs can be achieved by two strategies: (1) monitoring the well catchment at positions where several predicted contaminant plumes coincide and neglecting risk sources that have clearly separated plumes; (2) by avoiding to attain early-warning functionality through covering the most remote risk sources.

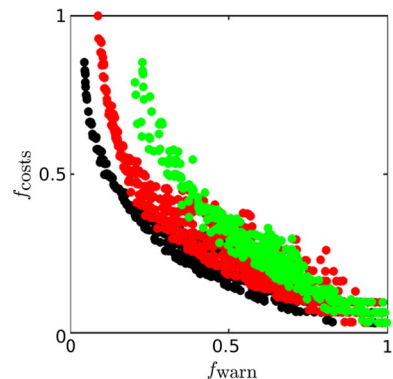
5.2 Robustness of Monitoring Networks Can Be Improved by Considering Uncertainties

When looking at the relatively thin predicted plumes in Fig. 6, it quickly becomes clear that the actual performance of optimized early-warning monitoring networks (compared to the performance predicted during the optimization) is highly sensitive to predictive uncertainties. Therefore, the second step of our approach was to include a corresponding representation of uncertainty during the optimization in order to achieve robustness. We call a Pareto-optimal monitoring network robust against considered uncertainties, when its performance is not sensitive to these uncertainties. Hence, a suitable measure of robustness is the difference between the optimized predicted performance and the particular performance obtained after actual installation. For demonstrating and discussing this step, we perform an optimization according to Eq. 12. Compared to the basic application scenario considered in Sect. 5.1, we now work with a modified scenario that uses as robust objective functions the average over the respective goal attainment levels across the four hydraulic scenarios listed in Table 3. To demonstrate the effect of uncertainty and the benefit of robustness, we consider three different uncertainty scenarios for the optimization:

1. We have an idealistic situation without any uncertainty.
2. We assume that we have an idealistic situation, but in reality the used model is subject to uncertainties.
3. We have uncertainty in modeling our system and also consider it during the optimization procedure.

Figure 6 shows the obtained Pareto fronts. For better comparison within a single figure, we show only projections onto the cost/time plane. Scenario one leads to the same results to those from Sect. 5.1 (black). Compared to the other fronts, these solutions perform the best. However, in practice, the scenario without uncertainty does not exist. Ignoring uncertainty (scenario two) leads to differences between the expected performance (black) and the actual performance in reality (green), an average over all hydro(geo)logic scenarios. In theory, the most expensive solution (from the black front) has a close-to-perfect benefit in early-warning time. Compared to that, the same solution performs poorly in reality (green). Considering uncertainty (third scenario), the results (red) are more modest in their claims of performance during the optimization compared to these without uncertainty (black). The results are less optimistic, because it is harder to guarantee a reliable and early detection of contaminant plumes when admitting that one does not know their exact locations and mutual overlaps.

Fig. 6 Projection of the Pareto front for the robust multi-objective optimization (see Eq. 12 and Sect. 5.2). *Black* Pareto front from an optimization that does not consider uncertainty (same as in Sect. 5.1). *Red* Pareto front obtained when optimizing the performance on average over four hydraulic scenarios (see Table 3). *Green* performance of the non-robust optimization results on average over the four hydraulic scenarios



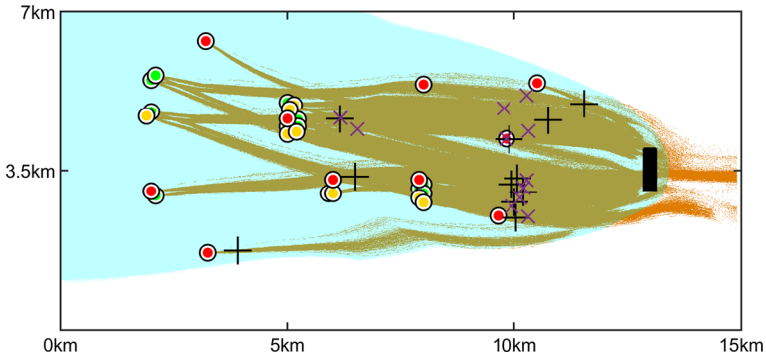


Fig. 7 Comparison of a risk-prioritized solution from the robust multi-objective optimization (see Eq. 13 and Sect. 5.3) to a solution without risk categorization (see Eq. 12 and Sect. 5.2). *Plus marks* prioritized monitoring network; *cross marks* network without prioritization; all other *symbols and colors* same as in Fig. 5

Compared to the green front, one could either find monitoring networks with the same performance in early-warning time but with less costs, or monitoring networks with a much better performance with identical costs. Obviously, the black solutions are only optimal for one hydraulic scenario, but not for the considered uncertainties.

The difference in early-warning time between the black front and the red front at given cost values can be interpreted as the costs of having (and considering) uncertainty. The difference between the black front and the green front at given cost values can be interpreted as costs of not considering uncertainty, and these costs are larger than the costs of considering uncertainty. Without any surprise, all scenarios agree that one cannot expect any performance at the limit of no costs.

We can conclude that considering uncertainties in the optimization is in fact required, because otherwise one will obtain overly optimistic model-based results that do not perform as desired when installed in reality. The scenario analysis performed here is very crude and serves mainly for illustration. In practice, it would be advisable to also account for the other sources of uncertainty mentioned in Sect. 3.2, and to work with a much larger number than four hydraulic scenarios or realizations.

5.3 Risk Prioritization Can Help to Find Good But Low-Cost Monitoring Networks

The third step of our approach was to categorize risks according to severity. Such a categorization can be used to prioritize the monitoring of risk sources during the optimization. This leads to practically affordable early-warning monitoring networks that cover the risk sources in proportion to their (perceived) risk. For demonstrating and discussing this step, we now perform an optimization according to Eq. 13, i.e., working with a risk categorization. In this optimization, we categorize the same inventory of risk sources used above into severe, medium, and tolerable risk sources, and compare the results to those obtained without risk categorization according to Eq. 12 (see Sect. 5.2).

In Fig. 7, we compare a selected Pareto-optimal monitoring network from the risk-prioritized formulation (plus marks) with a selected one obtained from Sect. 5.2, i.e., without risk prioritization. The prioritized solution was selected to achieve a 100% detection probability and a large early-warning time utility (here: 86%) for all severe risk sources while maintaining the number of monitoring wells at eleven, i.e., using the same number as the non-prioritized solution. This results in a reduced detection probability and early-warning

Table 4 Key data of the different monitoring networks illustrated in Fig. 7

Monitoring network	Performance for risk sources...	Detection probability (–)	Early-warning time utility (–)
Non-prioritized	All	0.97	0.79
Non-prioritized	Severe	0.88	0.72
Prioritized	All	0.80	0.66
Prioritized	Severe	1	0.86

utility across the remaining risk inventory. The key performance data of both networks can be found in Table 4. By comparison of these two networks, one can see that two additional severe risk sources (red) are now monitored [one at (4 km, 1.5 km) in domain coordinates and one at (7 km, 3 km)]. One monitoring well roughly at (7 km, 4 km) has been removed, because it mainly served to monitor medium and almost tolerable (yellow, green) risk sources. Other significant changes appear around (11 km, 4.5 km), where a group of three monitoring wells was replaced by two monitoring wells that coincide better with the possible plume paths of only severe risk sources (red).

The decision maker has several strategies to screen through the list of Pareto-optimal results. One option is to screen through the possible compromises between overall coverage and prioritized coverage at fixed costs. A second option is to start with perfect coverage of severe risk sources, and then to investigate the additional costs and willingness to pay for additional coverage of medium and almost tolerable risk sources. With this, decision makers are able to make the necessary compromises between the performance for different risk categories under budgetary constraints in an informed and transparent manner.

5.4 Remaining Risk Can Be Managed

The fourth step of our approach was to introduce an appropriate representation of unknown risk sources in the risk inventory in order to make the early-warning system robust against residual risks. For demonstrating and discussing this step, we now perform an optimization according to Eq. 14 (see Sect. 5.4), i.e., working with a model for residual risks, with a respective risk category, and with two corresponding additional objective functions. For simplicity of the analysis, we drop the classification of the known risk sources.

Figure 8 shows one selected Pareto-optimal network, which can be separated into two parts with different tasks: The first part specializes only on good coverage of the residual risks (*line of defense* close to the line of attack), and the second part is an augmentation of the line of defense in order to improve the performance in monitoring all the known risk sources. The model for residual risks is the representation by the line of attack (blue circles) that lies between the well gallery and the bulk area of the catchment. Again, the line of attack is a line of hypothetical point sources placed at a travel time of two years away from the well gallery.

The performance data for both parts are shown in Table 5. The strategy for covering the residual risk is a line of defense that directly matches the line of attack. In the solution shown, this line is composed of 17 monitoring wells. This strategy achieves a detection probability of 81 % and claims an early-warning utility of 68 %. The detection probability number is accurate for all residual risks (behind the line of attack). The early-warning time, however, is strongly suboptimal, because most residual risks could be detected earlier when knowing their location.

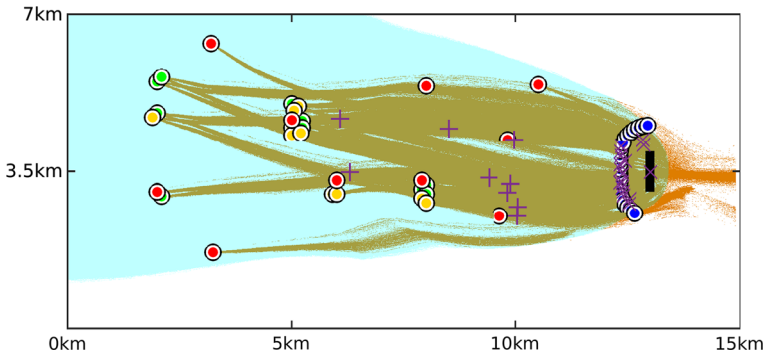


Fig. 8 Comparison of two early-warning network from the optimization with residual risks (see Eq. 14 and Sect. 5.4). *Cross marks* monitoring network specialized only on residual risks; *plus marks* augmentation to improving the performance for the known risk sources; *blue circles*: model of residual risk (line of attack); all other *symbols* and *colors* same as in Fig. 5

Table 5 Key data of the different monitoring networks illustrated in Fig. 8

Monitoring network	Performance for risk sources ...	Detection probability (–)	Early-warning time utility (–)	Number of monitoring wells
Only residual	Residual	0.81	0.68	17
Only residual	All known	0.65	0.07	17
Augmented	Residual	0.81	0.68	26
Augmented	All known	0.96	0.66	26

The line of defense provides a good coverage for almost all possible residual risk sources anywhere in the catchment. Only 19% of the entire catchment area cannot be controlled by the line of defense. Hence, it is clear that it should also provide a good detection probability for the known risk sources. However, the detection probability of the line of defense for all known risk sources is only 65%. This means that 35% of the known risk sources lay inside the area which is not controlled by the line of defense, which is more than the 19% predicted during the optimization. This can be explained by considering the known risk inventory as only a *small* sample out of all potential risk source positions (the remaining risk). The early-warning utility for the known risk sources, however, is unsatisfactory (7%). In order to improve this performance, additional monitoring wells in the more remote parts of the catchment are required. The plus marks show a selected corresponding augmentation of the current network by nine monitoring wells. These additional wells increase the detection probability for the known risk sources from 65 to 96%, and the corresponding early-warning utility from 7 to 66%.

Depending on her/his risk aversion and risk perception, a decision maker could either start with coverage of residual risks and then buy more early-warning time for known risk sources (possibly for the severe ones first), or start with network solutions that offer a satisfactory performance for known risk sources at affordable costs, and then investigate the cost–performance trade-offs for additional coverage of residual risks. The provided information is valuable decision support in situations where decision makers are highly risk averse, or where the residual risks are known to exist at unknown locations. In situations penetrated with deep uncertainty in flow conditions, we expect the residual risk solution to provide a large degree of robustness.

6 Conclusion and Outlook

Finding optimal early-warning monitoring networks for well catchments is challenging because the desired qualities of monitoring networks are multiple and mostly competing. In our study, we proceed from the assumption that risk control in drinking water well catchments requires monitoring networks to detect all possible future contaminant spills at a maximum reliability level and that they should detect the corresponding contaminant plumes as early as possible in order to allow for a maximum time to install counter measures. Nevertheless, the installation and operation costs should be as low as possible. We proposed a multi-objective and risk-prioritized optimization approach for early-warning monitoring networks that is based on four steps and leads to four corresponding key findings:

1. The considered objectives are in fact competing, and multi-objective optimization is an adequate approach to solving the problem.
2. The actual performance of optimized early-warning networks is highly susceptible to predictive uncertainty of contaminant transport toward the well. Therefore, mechanisms for robust optimization under uncertainty (e.g., based on Monte Carlo simulation of scenario analysis) should be applied.
3. A prioritization of risk sources and a corresponding augmentation of the multi-objective optimization offers options for cost-efficient focus on only the most severe risks, and provides valuable insights into the trade-offs for additional coverage of medium or almost tolerable risks.
4. Early-warning monitoring networks can even be optimized to cover risk sources that are unknown in existence or location through an adequately chosen residual risk model.

As an outlook for future studies, we plan (1) to extend the scenario-based uncertainty approach to a Monte Carlo approach that considers a more comprehensive list of uncertainties and (2) to apply this method to real catchments.

Acknowledgments We gratefully acknowledge the financial support of the German Research Foundation (DFG) through the Cluster of Excellence in Simulation Technology (EXC310/1) and through the International Research Training Group NUPUS (IRTG 1398). We also acknowledge the financial support by the German Technical and Scientific Association for Gas and Water (DVGW) through the Project W 1-01-10-F Risk-based groundwater monitoring for wellhead protection areas.

Appendix

Here, we provide the mathematical definition of our objective function for early-warning time and the link to the objective function detection probability.

For any given design \mathbf{d} , there is a list $L_{\mathbf{d}}$ of monitoring wells to be installed. From that list, several monitoring wells M_j , $j \in L_{\mathbf{d}}$ could have nonzero detection probabilities P_{ij}^{det} for any given risk source R_i . These monitoring wells could differ in early-warning time and detection probability, and we need to define a rule how to identify a unique value for detection probability and early-warning utility for each risk source R_i . We need such a rule to avoid that the monitoring wells have to be sampled for a large list of contaminations (alternatively, this could also be solved by optimization). There is no general rule that the monitoring well with the best early-warning time for a risk source R_i will automatically have the best value in detection probability for the same risk source. As rule to identify a unique monitoring well, we use a method of best compromise, i.e., we select the monitoring well position M_j for the risk source R_i out of the list $L_{\mathbf{d}}$ that maximizes the product of detection probability and early-warning time t_{ij} :

$$\ell_i = \arg \max_{j \in L_d} \left[P_{ij}^{\det} t_{ij} \right]. \tag{15}$$

Finally, we average the values of P_{ij}^{\det} with $j = \ell_i$ over all risk sources R_i to obtain the overall detection probability P_d^{\det} of the monitoring network and define:

$$f_{\det}(\mathbf{d}) = 1 - \underbrace{\frac{1}{n_R} \sum_{i=1}^{n_R} P_{i\ell_i}^{\det}}_{P_d^{\det}}. \tag{16}$$

As argued in Sect. 3.1, the achievable early-warning time values for different risk sources may vary between none and several decades, such that a simple linear relation between individual early-warning time values and the overall early-warning performance of the monitoring network is not suitable. Instead, we work with a nonlinear utility function $U_i(t_{ij})$ that expresses (on a normalized scale) the early-warning time achieved by a monitoring well M_j for a risk source R_i . Thus, in analogy to f_{\det} , we define:

$$f_{\text{warn}}(\mathbf{d}) = 1 - \underbrace{\frac{1}{n_R} \sum_{i=1}^{n_R} U_i(t_{i\ell_i})}_{U_d^{\text{warn}}}, \tag{17}$$

where U_d^{warn} is the overall early-warning utility of the monitoring network defined by \mathbf{d} . The utility function $U_i(t)$ is defined by a case-dependent piecewise linear function. The respective sections of piecewise linearity are separated by predefined values \hat{t}^{\min} and \hat{t}^{\max} . \hat{t}^{\min} is the desirable minimal early-warning time, and \hat{t}^{\max} is a maximal useful early-warning time beyond which there is no additional utility. $\hat{t}_i^{\max} = \hat{t}^{\max}$ is a user-defined value, but is set to $\hat{t}_i^{\max} = \tau_{i,\text{well}}$ whenever $\tau_{i,\text{well}}$ is smaller than the user-defined value. This modification serves to ensure that a monitoring well that achieves the maximal possible early-warning time (defined through $\tau_{i,\text{well}}$) is rated with the maximal possible utility value. Then, we assign $U(t = 0) = 0$, $U(t = \hat{t}^{\min}) = \hat{u}^{\min}$ and $U(t \geq \hat{t}^{\max}) = 1$. This yields as utility function:

$$U_i(t) = \begin{cases} 1 & \forall t : t \geq \hat{t}^{\max} \\ \frac{\hat{u}^{\min} \cdot (t - \hat{t}_i^{\max})}{\hat{t}^{\min} - \hat{t}_i^{\max} + 1} + 1 & \forall t : \hat{t}^{\min} \geq t \geq \hat{t}_i^{\max} \wedge t \leq \hat{t}_i^{\max} \hat{u}^{\min} \\ \frac{t}{\hat{t}_i^{\max}} & \forall t : t \leq \hat{t}^{\min} \vee t > \hat{t}_i^{\max} \hat{u}^{\min}. \end{cases} \tag{18}$$

The utility function $U_i(t)$ is illustrated in Fig. 1.

References

Allen, R.M., Kanamori, H.: The potential for earthquake early warning in southern California. *Science* **300**(5620), 786–789 (2003)

Aller, L., Lehr, J., Petty, R., Bennett, T.: DRASTIC: a standardized system to evaluate groundwater pollution potential using hydrogeologic setting. *J. Geol. Soc. India* **29**(1), 23–37 (1987)

Carrera, J.: An overview of uncertainties in modelling groundwater solute transport. *J. Contam. Hydrol.* **13**(1), 23–48 (1993)

Chamberland, J.F., Veeravalli, V.V.: Decentralized detection in sensor networks. *IEEE Trans. Signal Process* **51**(2), 407–416 (2003)

Cox, L.A.T.: What’s wrong with risk matrices? *Risk Anal.* **28**(2), 497–512 (2008)

- Cushman, D.J., Driver, K.S., Ball, S.D.: Risk assessment for environmental contamination: an overview of the fundamentals and application of risk assessment at contaminated sites. *Can. J. Civil Eng.* **28**(S1), 155–162 (2001)
- Davison, A., Howard, G., Stevens, M., Callan, P., Fewtrell, L., Deere, D., Bartram, J.: Water safety plans: managing drinking-water quality from catchment to consumer. Technical report, World Health Organization, Geneva, Water, Sanitation and Health Protection and the Human Environment, <http://www.who.int/iris/handle/10665/42890> (2005)
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **6**(2), 182–197 (2002)
- Enzenhofer, R., Binning, P., Nowak, W.: Stakeholder-objective risk model (storm): determining the aggregated risk of multiple contaminant hazards in groundwater well catchments. *Adv. Water Resour.* **83**, 160–175 (2015)
- Enzenhofer, R., Bunk, T., Nowak, W.: Nine steps to risk-informed wellhead protection and management: a case study. *Groundwater* **52**(S1), 161–174 (2014)
- Enzenhofer, R., Nowak, W., Helmig, R.: Probabilistic exposure risk assessment with advective-dispersive well vulnerability criteria. *Adv. Water Resour.* **36**, 121–132 (2012)
- Faber, M., Maes, M., Maes, M., Huysse, L.: Modeling of risk perception in engineering decision analysis. *Proc IFIP WG 7*, 113–122 (2004)
- Frind, E., Molson, J., Rudolph, D.: Well vulnerability: a quantitative approach for source water protection. *Groundwater* **44**(5), 732–742 (2006)
- Hadka, D., Reed, P.: Borg: an auto-adaptive many-objective evolutionary computing framework. *Evol. Comput.* **21**(2), 231–259 (2013)
- Hoeting, J.A., Madigan, D., Raftery, A.E., Volinsky, C.T.: Bayesian model averaging: a tutorial. *Statistical Science* **14**(4), 382–401 (1999)
- Jacobson, E., Andricevic, R., Morrice, J.: Probabilistic capture zone delineation based on an analytic solution. *Groundwater* **40**(1), 85–95 (2002)
- Kinzelbach, W.: The random walk method in pollutant transport simulation. In: Custodio, E., Gurgui, A., Lobo Ferreira, J.P. (eds.) *Groundwater Flow and Quality Modelling*, pp. 227–245. Springer, Netherlands (1988)
- Koch, J., Nowak, W.: A method for implementing Dirichlet and third-type boundary conditions in PTRW simulations. *Water Resour. Res.* **50**(2), 1374–1395 (2014)
- LaBolle, E.M., Fogg, G.E., Tompson, A.F.: Random-walk simulation of transport in heterogeneous porous media: local mass-conservation problem and implementation methods. *Water Resour. Res.* **32**(3), 583–593 (1996)
- Loaiciga, H.A., Charbeneau, R.J., Everett, L.G., Fogg, G.E., Hobbs, B.F., Rouhani, S.: Review of ground-water quality monitoring network design. *J. Hydraul. Eng.* **118**(1), 11–37 (1992)
- MacFarlane, D., Cherry, J., Gillham, R., Sudicky, E.: Migration of contaminants in groundwater at a landfill: a case study: 1. groundwater flow and plume delineation. *J. Hydrol.* **63**(1), 1–29 (1983)
- Marler, R.T., Arora, J.S.: Survey of multi-objective optimization methods for engineering. *Struct. Multidiscip. Optim.* **26**(6), 369–395 (2004)
- Meyer, P.D., Valocchi, A.J., Eheart, J.W.: Monitoring network design to provide initial detection of groundwater contamination. *Water Resour. Res.* **30**(9), 2647–2659 (1994)
- Michalak, A.M., Kitanidis, P.K.: Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling. *Water Resour. Res.* **40**(8), W08302 (2004)
- Moutsopoulos, K.N., Gemtzi, A., Tsihrintzis, V.A.: Delineation of groundwater protection zones by the backward particle tracking method: theoretical background and GIS-based stochastic analysis. *Environ. Geol.* **54**(5), 1081–1090 (2008)
- Neupauer, R., Wilson, J.: Backward probabilistic model of groundwater contamination in non-uniform and transient flow. *Adv. Water Resour.* **25**(7), 733–746 (2002)
- Nowak, W., Bode, F., Loschko, M.: A multi-objective optimization concept for risk-based early-warning monitoring networks in well catchments. *Proc. Environ. Sci.* **25**, 191–198 (2015)
- Nowak, W., Schwede, R.L., Cirpka, O.A., Neuweiler, I.: Probability density functions of hydraulic head and velocity in three-dimensional heterogeneous porous media. *Water Resour. Res.* **44**(8), W08452 (2008)
- Overheu, N.D., Tuxen, N., Flyvbjerg, J., Aabling, J., Andersen, J.A., Pedersen, J.K., Thyregod, T., Binning, P.J., Bjerg, P.L.: Risk-based prioritization of ground water threatening point sources at catchment and regional scales. *Sci. Total Environ.* **485**, 769–775 (2014)
- Reed, P.M., Minsker, B.S.: Striking the balance: long-term groundwater monitoring design for conflicting objectives. *J. Water Resour. Plan. Manag.* **130**(2), 140–149 (2004)

- Saenton, S., Illangasekare, T.: Determination of DNAPL entrapment architecture using experimentally validated numerical codes and inverse modeling. *Dev. Water Sci.* **55**, 767–778 (2004)
- Salamon, P., Fernández-García, D., Gómez-Hernández, J.J.: A review and numerical assessment of the random walk particle tracking method. *J. Contam. Hydrol.* **87**(3), 277–305 (2006)
- Sampat, P.: *Deep Trouble: The Hidden Threat of Groundwater Pollution*, vol. 154. Worldwatch Institute, Washington DC (2000)
- Scheidegger, A.E.: Statistical hydrodynamics in porous media. *J. Appl. Phys.* **25**(8), 994–1001 (1954)
- Schwede, R.L., Cirpka, O.A.: Stochastic evaluation of mass discharge from pointlike concentration measurements. *J. Contam. Hydrol.* **111**(1), 36–47 (2010)
- Stauffer, F., Guadagnini, A., Butler, A., Franssen, H.J.H., Van De Wiel, N., Bakr, M., Riva, M., Guadagnini, L.: Delineation of source protection zones using statistical methods. *Water Resour. Manag.* **19**(2), 163–185 (2005)
- Talbi, E.G.: *Metaheuristics: From design to Implementation*, vol. 74. Wiley, Hoboken (2009)
- Taleb, N.N.: *The Black Swan: The Impact of the Highly Improbable Fragility*. Random House, New York (2007)
- Tarantola, A.: *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics, Philadelphia (2005)
- Tartakovsky, D.M.: Assessment and management of risk in subsurface hydrology: a review and perspective. *Adv. Water Resour.* **51**, 247–260 (2013)
- Troldborg, M., Lemming, G., Binning, P.J., Tuxen, N., Bjerg, P.L.: Risk assessment and prioritisation of contaminated sites on the catchment scale. *J. Contam. Hydrol.* **101**(1), 14–28 (2008)
- USEPA (1993) Guidelines for delineation of wellhead protection areas. EPA-440/5-93-001
- Varljen, M.D., Shafer, J.: Assessment of uncertainty in time-related capture zones using conditional simulation of hydraulic conductivity. *Groundwater* **29**(5), 737–748 (1991)
- Yenigül, N., Elfeki, A., Van den Akker, C., Dekking, F.: A decision analysis approach for optimal groundwater monitoring system design under uncertainty. *Hydrol. Earth Syst. Sci. Discuss. Discuss.* **3**(1), 27–68 (2006)
- Zwahlen, F.: *Vulnerability and risk mapping for the protection of carbonate (karst) aquifers, final report (COST action 620)*. Technical report, European Commission, Directorate-General XII, Science, Research and Development (2003)