MARSOL

Demonstrating Managed Aquifer Recharge as a Solution to Water Scarcity and Drought

Methodology for Probabilistic Risk Evaluation Linked to MAR Activities

Deliverable No.	Deliverable 16.1
Version	Version 1
Version Date	23.02.2015
Author(s)	Xavier Sanchez-Vila, Daniel Fernàndez-Garcia, Albert Folch, Paula Rodríguez-Escales, Erica Siirila, Christopher Henri, Carme Barba (UPC) Contact: Xavier Sanchez-Vila (xavier.sanchez-vila@upc.edu)
Dissemination Level	PU
Status	Final

The MARSOL project has received funding from the European Union's Seventh Framework Programme for Research, Technological Development and Demonstration under grant agreement no 619120.

Contents

Figure captions

EXECUTIVE SUMMARY

WP Technology Assessment and Risk (WP16) deals with a proper methodology to combine in a single tool the analysis of the potential risk associated to a Managed Aquifer Recharge facility. Here risk is defined in a global multidisciplinary and integrated approach.

We present a fault tree methodology to evaluate failure in engineered systems and its application to MAR facilities characterized by the presence of surface infiltration ponds. The main idea consists of breaking complex problems into individual (less complex) events, and evaluating the probabilities of failure of such individual events. Probabilities are then computed based on the rules provided in Boolean algebra. This method allows the introduction of multidisciplinary problems into a single common framework. With this approach one can identify the events that contribute most to the final risk estimate or those that propagate the highest degree of uncertainty throughout the system. This can then be used to invest further resources to specific events. Assessing risk in hydrological systems is an interdisciplinary field; as a result, communicating the information across interfaces between different fields in a comprehensible and efficient manner is needed, allowing decision makers to better visualize the components leading to system failure, as well as the uncertainty emerging from each subsystem.

Managing risk during remediation efforts is technically cumbersome due to the difficulty of combining a large number of uncertainties associated with subsurface heterogeneities, and the existence of multiple potential sources, receptors, and pathways of exposure. The probabilistic risk analysis (PRA) framework allows updating risk evaluations in a full stochastic framework, from a pre-designed and operational monitoring system. By integrating a Bayesian framework into a PRA framework based either on event trees or else on fault trees, measurements from observation wells can be used to update the probability of system failure over time. As information is added, the Bayesian interpretation of the problem allows to automatically recalibrating the probability of system failure.

The theoretical framework is finally converted into a practical tool, where the main steps are presented here. First, a list of all the potential individual events that could be leading to system failure in a general MAR facility is presented. Then, the individual components are combined to construct potential failure paths that are specific of each site. Emphasis is placed both on general as well as on site-specific failure paths. The initial steps toward the development of a generic tool are hinted. The initial point is a technical questionnaire that will bring to the surface the main relevant events to be considered in each specific site. This will allow constructing the fault tree for each specific site, and to use it as a driver for failure probability evaluation. This simplifies greatly the work of the decision maker since all potential failure paths are recognized by the system, allowing also a direct comparison of the results from all MARSOL sites.

1 Introduction

1.1 Objectives

This is the first deliverable in **WP 16 Technology Assessment and Risk**. Deliverable 16.1 can be synthesized in one main objective: Defining a strategy for risk evaluation linked to MAR activities.

In this document, we provide the basis for a rigorous treatment of risk analysis in an MAR facility, specific for surface ponds. The variety of objectives and facility types in MAR leads to a need to define a list of potential hazards associated to the general activity of artificial recharge from surface recharging ponds. This list of hazards can then be reduced to a subset, specific to each individual case. The Deliverable provides both the general framework as well as a specific methodology to deal with individual hazards, in an integrated and transferable (from site to site) framework.

1.2 Outline

The document incorporates different chapters. In Chapter 2, we present a fault tree methodology, aimed at breaking complex infiltration problems into individual events and accounting for uncertainty in the different components, always in a pluridisciplinar approach, meaning that it formally merges uncertainties in hydrological, legal, geotechnical, and environmental problems. The integration involves the correct operations involved in deterministic Boolean logic, as well as the creation of a new stochastic Boolean operator that weakens the need for the conventional assumption of independent events. With this approach one can identify the events that contribute most to the final risk estimate or those that propagate the highest degree of uncertainty throughout the system, allowing decision makers to better visualize the components leading to system failure, as well as the uncertainty emerging from each subsystem into the combined final risk assessment.

Managing risk during remediation efforts is technically cumbersome due to the difficulty of combining a large number of uncertainties associated with subsurface heterogeneities, and the existence of multiple potential sources, receptors, and pathways of exposure. In Chapter 3, we developed a probabilistic risk analysis (PRA) framework to evaluate the risk of failure of a generic remediation or contamination project that links system components with the transport phenomena and the monitoring system. By integrating a Bayesian interpretation of the system into the PRA, measurements from observation wells can also be incorporated to update the probability of system failure over time. As information is added, the Bayesian interpretation of the problem allows to automatically recalibrating the probability of system failure.

Finally, Chapter 4 is the core contribution in this deliverable. This task starts by defining all the potential individual components that could be leading to system failure in a general MAR facility. Then the individual components can be used to construct failure paths that are specific of each site. While potential paths are site specific, we focus on a number of combinations of paths that might be considered as general and so appear in a number of sites. The next step is the development of a tool, to be applied to the MARSOL sites in subsequent deliverables. From the responses to a technical questionnaire, the tool will directly construct a fault tree for any site. This simplifies greatly the work of the decision maker since all potential failure paths are recognized by the system.

2 Risk assessment and decision making in MAR facilities

2.1 Introduction, goals, and approach

All engineered systems are subject to uncertainty caused by a combination of epistemic and aleatory uncertainties. In hydrological problems, these include, in a non-extensive list, lack of characterization data, inadequate conceptual models, and occurrence of natural variability (Maxwell and Kastenberg, 1999). Moreover, when other aspects are included in a multidisciplinary framework, uncertainties arise in topics such as stability of hydraulic works, legal aspects, catastrophic events, …

Given such uncertainties, the traditional concept of single deterministic predictions for a given scenario has little practical purpose, as recognized by government regulatory bodies (EPA, 2001). As a consequence, regulatory agencies insist on the use of approaches that include estimates of uncertainty in risk analysis (Persson and Destouni, 2009).

In this work, we propose a formal probabilistic risk analysis (PRA) approach that relies on the use of fault trees and can address heterogeneity and uncertainty in a multidisciplinary framework. Fault trees have been used in risk assessment of engineered systems (Bedford and Cooke, 2003). Since hydrosystems comprise a mixture of natural and engineered components, this approach has received some attention in the hydrological community (e.g., Bolster et al., 2009).

The basic idea of the approach consists of taking a complex system, difficult to be handled as a whole, and divide it into a series of quasi-independent simpler modules that are manageable individually through simplified methods or existing codes. Once probabilities of occurrence of each small problem are computed, they are recombined in a systematic manner to provide the global risk assessment of a given system.

A rigorous PRA based on fault tree should consist of six steps (Bedford and Cooke, 2003): (1) Define failure of the system under study; (2) Identify the key events that would potentially result in failure; (3) Construct a fault tree depicting the combination of events; (4) Develop a probabilistic representation of the fault tree using Boolean algebra; (5) Compute the individual probabilities of event occurrence using conservative approaches and Occam's razor; and (6) Use individual probabilities to calculate the global risk (i.e., probability of system failure).

The advantage of the approach is that, for a well-developed fault tree, all key events should be weakly interdependent. Actually, each one can be evaluated without (or with little) explicit knowledge of all others, meaning that it can be addressed by different experts within a multidisciplinary group. Additionally, a decision maker can use the fault tree to visually understand where risk and uncertainty emerge in this system, without having to enter into the complexities of each component. In some sense, the fault tree enables better decision making as it acts as translator of information between experts in different fields.

Furthermore, the approach enables optimal allocation of resources and the incorporation of new theories and data sets as they become available. Moreover, it can be used for rational allocation of resources for further data acquisition (Nowak et al., 2010) within a dynamical and adaptive framework.

In this chapter we illustrate the fault tree approach to a groundwater pollution problem with emphasis in the potential hazardous effect of polluted water reaching the water supply system, and the effect in humans. This is just one of the many paths that will be later (in successive chapters) included in the global framework for hazards associated to MAR facilities, and will be used to show the potential of the method based on fault tree analysis. Thus, here we define system failure as risk exceeding a threshold value, the latter often given by environmental regulation bodies for the sensitive target at stake (EPA, 2001).

2.2 Problem formulation

Subsurface water can be polluted by the presence of different chemicals (organic or inorganic) and pathogenic microorganisms (bacteria, protozoa, and viruses) (Molin and Cvetkovic, 2010). Exposure of humans to polluted water through ingestion, inhalation, or skin contact may result in a number of diseases. Whether one of these potential diseases is developed in a given individual depends on the toxicity of the pollutant, but also on the metabolism, personal habits of an individual's water-related practices, and a combination of consumption and exposure habits.

Diseases can be either caused by accumulation over the years or by acute exposure. For a given hazardous substance, the potential (risk) of developing a disease increases with concentration and with time of exposure. When several hazardous substances are assumed to coexist, risk can be assumed additive (Speek, 1981).

Health risk model

Depending on the particular contaminant, there are a number of models to evaluate the risk for a single substance. We use here the one suggested by EPA (1989) for carcinogenic compounds

$$r(\mathbf{x},t) = \beta_G \left(C(\mathbf{x},t) - C_G^* \right)^{m_G} + \beta_H \left(C(\mathbf{x},t) - C_H^* \right)^{m_H} + \beta_S \left(C(\mathbf{x},t) - C_S^* \right)^{m_S} \qquad (2.1)$$

where the concentration $C(\mathbf{x},t)$ is an outcome of all the relevant flow, transport and transformation processes within the system, the subscript *G* stands for ingestion, *H* for inhalation, and *S* for skin contact, and the different *β* values are pathway related coefficients that account for behavioural and physiological parameters (ingestion rate, body weight, exposure duration, and frequency); exponents *m* determine the non-linearity of dose-response curves; C^* are pollutant dependent threshold values, below which no adverse effects are expected for a given individual.

In (2.1) $C(\mathbf{x},t)$ can represent a point or a flux-averaged concentration. In most health risk applications, it corresponds either to the highest value or to an averaged concentration over the exposure period at the sensitive target. All parameters included in β relate to an individual from the exposed population, and thus contain some level of uncertainty.

In equation (2.1), the total risk is the result of different uncertain quantities. For example, β depends on toxicity, and vary according the population cohort (age groups and gender; Yu et al., 2003).

Stochastic representation of health risk

The aim is to evaluate the pollutant concentration at any particular point within an environmentally sensitive target over a period of time, t_p , and to quantify its uncertainty. Spatial variability and uncertainty in concentration is due to the ubiquitous heterogeneity in physical and biochemical processes, boundary conditions and contaminant release patterns. The processes involved are solute- and soil-dependent, and might include advection, diffusion, dispersion, sorption, precipitation/dissolution, redox processes, cation exchange, evaporation/condensation, microbial or chemical transformation and decay.

The specified uncertainty in a number of parameters results in risk r regarded as a random function R . It is convenient to formulate risk in terms of exceeding probabilities after defining r_{crit} , an environmentally regulated value:

$$P(R > r_{crit}) = 1 - F_R(r_{crit}) \qquad (2.2)$$

Uncertainty in the concentration can be reduced by conditioning on measurements of either the dependent variables (e.g. concentrations, groundwater heads, river discharges, etc) or the parameters themselves (through field or laboratory tests).

2.3 Fault-tree (FT) analysis

The initial step in an FT analysis is to define the system setup. We consider several sources of contamination (SO_i), a mean flow field and a target region. Sources of contamination could be natural sources, spills, industrial regions, or agricultural practices. The target sensitive location could be an abstraction well, a lake or a residential area. Based on this generic system, we will follow the first four steps already outlined. Steps 5 and 6 are straightforward and need no further explanation.

Step 1: Defining System Failure (SF). We define SF as risk exceeding a critical regulatory value ($r > r_{crit}$), with exceedance probability given by Eq. (2.2).

Step 2: Identifying Key Events. We divide the problem into two components: A hydrological contamination scenario and its consequences on human health. This is an important distinction because only the joint effect can culminate in adverse health effects.

The first key component is the “Critical Concentration of Exposure” (CCE_i), defined as the event that the concentration of a contaminant *i* arrives at some sensible zone exceeding some critical concentration value. Only if such an event occurs, decision makers must be alerted and should become concerned about the consequences on human health. The lower-level events associated with this key event are:

- SO (Source Occurrence). In many possible scenarios, the existence of a contaminant source is not deterministic, and the probability of its occurrence must be quantified.

- $P_{2,i}$ (Plume Path) refers to the event that the pollution plume hits the target zone. If such a path does not exist, there is no reason for concern.

- NA (Natural Attenuation) represents the event that reactions, dispersion and dilution combined lead to concentration peaks below a defined threshold value.

The second component relates to health risk considerations. The basic event is:

- BPC_i (Behavioral Physiological Component) corresponds to the event that an exposed individual (or cohort) has a characteristic β value.

Note here that CCE is conditioned on a value of β coming from BPC, which is not a single value and it varies within the population based on several uncertain parameters. This fact will require, in a later stage of our analysis, an extension of the conventional fault tree approach to account for all possible values from the distribution of β, introducing a new Boolean operator.

Step 3: Building the Fault Trees. The first branch of the fault tree addresses the hydrogeological contamination scenario, leading to the key event CCE_i. The very simple fault tree is shown in Figure 2.1.

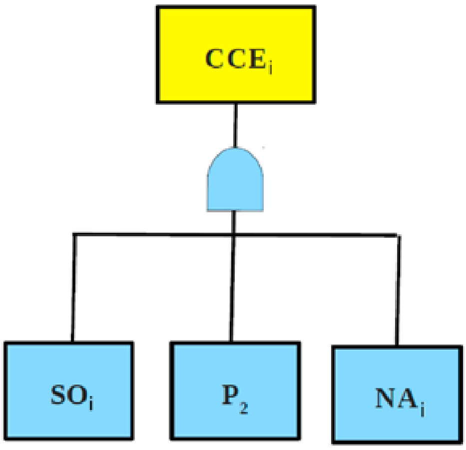

Figure 2.1. Fault tree for CCE_i

The combination with the second branch yields the main fault tree and represents the novelty of this work. This tree replicates for all species and sources (see Figure 2.2). It illustrates visually the link between contamination and human health risk. The system failure (risk exceedance) for contaminant *i* is the joint occurrence of the events CCE_i and BPC_i.

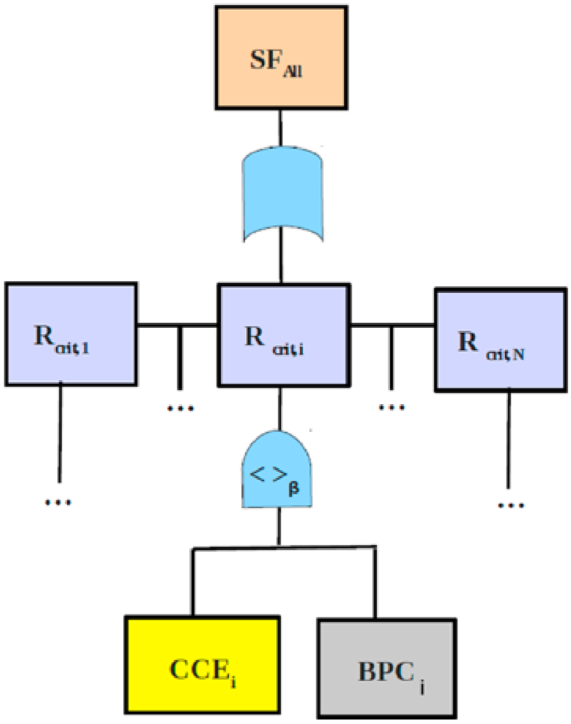

Figure 2.2. Fault tree for the total system failure

Figure 2.2 includes a gate Boolean operator $<>_\beta$. We define it as an "ENSEMBLE AND" gate, and reflects that the R_{crit} event must be calculated based on all possible values of β and of the concentration values reaching the target zone. The operator indicates that the risk should be averaged over the ensemble of β, because $P[R] = < P[R/\beta] >_\beta$. In other words, the fault tree here is generalized for every individual of the exposed population. This process represents the internal loop from the nested Monte Carlo approach proposed by Maxwell and Kastenberg (1999). Accounting for $<>_\beta$ within the fault tree implies that one needs to account for the variability in its description such that one can assign the probability of occurrence for the event R_{crit}.

Step 4: Translation to Boolean Logic. This is the final stage in a risk assessment problem. The subsequent steps (5 and 6) involve the actual calculations of probabilities of all basic events and its combination. First, we write a Boolean logic expression for the probability of event CCE_i occurring: $CCE_i = SO_i$ {AND} $P_{2,i}$ {AND} NA_i, with probability of occurrence:

$$P(CCE_i) = P(SO_i / P_{2,i}) P(NA_i) \quad (2.3)$$

since SO_i, $P_{2,i}$, and NA_i are independent events. Similarly, for the main fault tree depicted in Figure 2.2, the Boolean expression for system failure associa-ted with each source $R_{crit,i}$ can be written as $R_{crit,i}$ = CCE_i {AND} BPC_i, with the corresponding probability of occurrence given by

$$P(R_{crit,i}) = P(CCE_i) P(BPC_i) \quad (2.4)$$

If N contaminants were present, then the total system failure would be given as SF_{all} = $R_{crit,1}$ {OR} $R_{crit,2}$ {OR} … {OR} $R_{crit,N}$, and the corresponding probability could be computed (under the assumption that all probabilities are small) by

$$P(SF_{all}) = P(SF_1) + P(SF_2) + ... + P(SF_N) \quad (2.5)$$

The methodology presented would be extended in a subsequent chapter to analyse other possible causes for failure of MAR activities.

3 Probability Updating in Risk Assessment Efforts

3.1 Introduction

Probabilistic Risk Assessment (PRA) efforts involve the study of subsurface pollution and the possibility that it leads to undesired contamination at the comply surface (either an ecosystem, or a well, used for irrigation or drinking). Uncertain estimations regarding the concentration value of a given species (or a combination of them) reaching the surface of interest are driven by the inherent uncertainty in subsurface heterogeneity mapping, combined with the presence of unidentified processes such as rate-limited mass transfer. Understanding the impact of uncertainty on both the characterization of polluted plumes and the performance of remediation strategies is key to decision making (Cardiff et al., 2010).

Nonlinearities in the equations used to predict the subsurface migration and fate of organic or inorganic compounds along with the large number of uncertain parameters entering these equations complicate both uncertainty quantification and decision-making. A standard approach for quantifying uncertainty in complex phenomena is to treat relevant flow and transport parameters as correlated random fields, rendering the corresponding governing equations stochastic (Dean et al., 2005). Solving these equations typically yields an ensemble mean and variance of concentration in the subsurface, rather than providing deterministic single values at each point and any given time.

Such solutions are often inadequate, since many environmental regulations are formulated in terms of probabilities of a contaminant concentration exceeding a certain predefined mandated value. The first two moments (in the statistical sense) of a solute concentration do not provide such information, because the concentration statistics are typically non-Gaussian (Fernàndez-Garcia et al., 2005). The presence of a large number of uncertainties about parameter values (in particular when different scales are involved), locations and strength of multiple sources, potential receptors and pathways of exposure, can render standard Monte Carlo simulations computationally unfeasible. For the same reason, equations aimed at directly obtaining the full PDF of the variables of interest (Sanchez-Vila et al., 2009) can become impractical due to their high dimensionality. Probabilistic risk analyses (PRAs) alleviate the high-dimensionality of the problem by invoking a system approach to decontamination efforts.

In this document we develop a PRA framework to evaluate the risk of a generic contamination project with emphasis on the possibility of remediation, linking the various system components (exposure, sources and pathways) with their fate and transport phenomena. The method allows one to update the probability of failure given the observed real-time measurements of concentrations. This is achieved by integrating a Bayesian interpretation of the system into the PRA. The continuous update in the risk of failure of remediation actions can then be used systematically for decision making.

3.2 General Pollution Problem & Fault Tree Analysis

When conducting a probabilistic risk assessment (PRA) of a system there are a multitude of methods that can be used, e.g. failure mode and effect analysis (FMEA), multi-barrier analysis (MBA), event tree analysis (ETA), and fault tree analysis (FTA) among others. In each of these, while the goal is the same, the philosophy is slightly different.

For example, ETA and MBA methods are built on the concept of forward logic where for a system to fail an initiating event must occur and then be followed by a series of consequent events. In order for system to fail under an ETA an accident sequence must occur. On the other hand FTA is built on backward logic. In this case one identifies a particular failure of the system and defines it as the top event. Then one seeks the combination of all possible events that may contribute to system failure.

In this chapter we present the general methodology to construct fault trees for pollution problems of simple inorganic or organic compounds (no reaction included, except for natural attenuation). In order to do so we must first define the meaning of system failure (the definition is not unique) and the events that may lead to it. In general, a contamination problem involves a potentially hazardous substance, moving towards several receptors $\{\Omega_1,...,\Omega_m\}$. Failure of the system can be defined as the event that the concentration of the pollutant surpasses some critical value in any given receptor within a legally mandated time interval T.

Defining the event CC_j as the surpass of the critical contaminant concentration C_j^* at receptor *j*

$$CC_j = \{C(\mathbf{x} \in \Omega_j, t \leq T) > C_j^*\} \quad (3.1)$$

System failure is then formally equal to any of these events happening (thus equivalent to an “OR” operator in Boolean logic),

$$SF = \bigcup_{j=1,...,m} CC_j \ .$$

For a given event to occur the following basic sub-events must occur simultaneously (thus equivalent to an “AND” operator in Boolean logic)

- Contaminant Source: In some cases we cannot be certain that a contaminant source exists, but rather suspect its potential presence and we must deal with a probabilistic approach.

- Potential receptors: A receptor must be susceptible to adverse impact by any of the contaminant. Receptors can include individual people, wildlife, water reservoirs, or environmentally sensitive zones.

- Pathways: A path connecting contaminant source and receptor must potentially exist. This includes natural flow fields (in general heterogeneous), preferential flow paths, diffusive paths, capture zones, and so on.

- Fate: Mechanisms of natural attenuation or remediation might not be capable of reducing pollutant concentration sufficiently along a given pathway.

The inclusion of the previous elements allow constructing a generic fault tree for any number of receptors and pathways that could then be individualized for a given pollution setup.

Once the fault tree is constructed, the probability of each individual event can be evaluated. This is no easy task and can rely on a variety of alternative methods, from analytical models, simple or complex numerical approaches, surveys, expert opinion, or reliable data bases. Altogether, this would allow identifying a link of subevents that lead to the highest occurrence of an undesired event (in general, hazardous). More resources can then be allocated to these specific events and their probability of occurrence can be refined. Alternative remediation efforts can also be focused on the subevents leading to the largest risk.

3.3 Monitoring a pollutant plume

Observation wells are typically put in place to monitor and control an existing plume. This is an important step for monitored natural attenuation projects, and a necessary component for many others. Let us consider a monitoring system consisting of n_w observation wells. For the monitoring system to fail, the pollutant needs to go unnoticed, either because it bypasses an observation point and afterwards reaches the receptor, or simply because there is a device failure or insufficient sampling periodicity at the well. If pollution was detected, corrective actions could be taken (e.g., closure of the drinking well) and the system will not fail.

In this work we illustrate how Bayesian methods can be integrated into a probabilistic risk analysis to converge any initial estimate of probability of failure to its optimal value.

3.4 Computation of Probabilities

In order to compute probabilities for each set of basic components *jk* (receptor, pathways), one needs to solve a conservative or reactive transport model. The ensemble of models provides a mathematical description of all transport phenomena included in the fault tree. These models are typically expressed in terms of stochastic partial differential equations, conditioned to the presence of receptors. This is formally written by incorporating a vector of system parameters $\boldsymbol{\theta}$, as

$$F_{jk}\left\{C(\mathbf{x},t);\boldsymbol{\theta}\right\}=0 \tag{3.2}$$

The randomness of (3.2) stems from uncertainties associated with subsurface processes. These uncertainties can be structural, which are those that account for errors in the conceptual model, or parametric, associated to the imperfect knowledge of $\boldsymbol{\theta}$. Structural uncertainty can be implemented in PRA studies by evaluating several competing models instead of a single one, thus weighting models to reflect their different reliability. Parametric uncertainty involves the use of stochastic methods and tools: perturbation approaches, moment equations, PDF equations, Monte Carlo methods, etc.

3.5 Integration of Bayesian Methods

In subsurface hydrology, the Bayes formalism has been used in geostatistical inverse methods to account for uncertainty in the stochastic model (e.g., the mean and the covariance function of the natural log of transmissivity). Here we integrate a Bayesian formalism into PRA studies to update the probability of system failure given some real-time concentrations observations.

We denote $f(\boldsymbol{\theta})$ as the prior multivariate distribution of the parameters involved in the transport model; i.e., before the measurements of concentrations or any other data are accounted for. As concentration measurements of contaminant at different locations, $\{\mathbf{c}(\mathbf{x},t)\}$, become available, these prior estimates are updated into a posterior multivariate distribution by Bayes theorem

$$f(\boldsymbol{\theta}|\mathbf{c}) = \frac{f(\mathbf{c}|\boldsymbol{\theta})\, f(\boldsymbol{\theta})}{\int f(\mathbf{c}|\boldsymbol{\theta})\, f(\boldsymbol{\theta}) d\boldsymbol{\theta}} \tag{3.3}$$

where $L(\boldsymbol{\theta}) = f(\mathbf{c}|\boldsymbol{\theta})$ is the likelihood function. Essentially, this relationship expresses that the posterior distribution is proportional to the likelihood function.

The prior distribution is often subjectively estimated based on the assessment of an expert panel. A variety of prior distributions exist: informative or uninformative. The former are those that provide the system with external information on the parameters, based on past data and/or estimated from theoretical foundations. Uninformative priors are those that only incorporate some vague information about the parameters (e.g., the principle of parsimony that assigns equal probabilities within a range, implying that the prior pdf is uniform).

Measurements may result in posterior distributions that may clearly depart from the priors. The discrepancy between distributions is marked by the number and the relevance of data. After a sufficient number of data values are incorporated, the posterior distribution basically becomes independent of the prior. For this reason, when uncertainties are large, it is advisable to use priors displaying a sufficiently large variance.

The likelihood function $L(\boldsymbol{\theta})$ can be estimated using either parametric or non-parametric approaches. The former describe the unknown likelihood function with a finite set of parameters. The multinormal distribution is often adopted for this purpose. In the non-parametric approach the likelihood function can take any form and it is not limited to belong to a family of a priori known density functions. Although this can be computationally demanding, one can actually identify the true underlying structure more easily using such approaches.

Once the distribution function of the parameters is updated, we can generate many possible realizations of $\boldsymbol{\theta}$, and simulate the concentration values that can eventually be compared with new data as it becomes available, checking the predictability of the model, or else the need to be further updated.

As a final remark, it is relevant to state that a PRA framework is of potential use as a tool for decision-making processes, since it allows identifying the most critical events of a system. Once identified, this information can be used to prioritize future developments or redesign the system, allowing the continuous incorporation of data as it becomes available.

4 Risk assessment of MAR facilities involving infiltration ponds

4.1 Introduction

In this chapter we analyse the potential causes that may lead to failure in a MAR facility involving infiltration ponds. We will define events and subevents, classifying them into an integrated framework, and providing ideas about the potential ways to evaluate the probability of individual events.

The initial consideration, thus being also the first row in the fault tree, is to define all the potential generic causes that may lead to failure of an infiltration pond recharge facility. In fault tree format, it is shown as (see Figure 4.1):

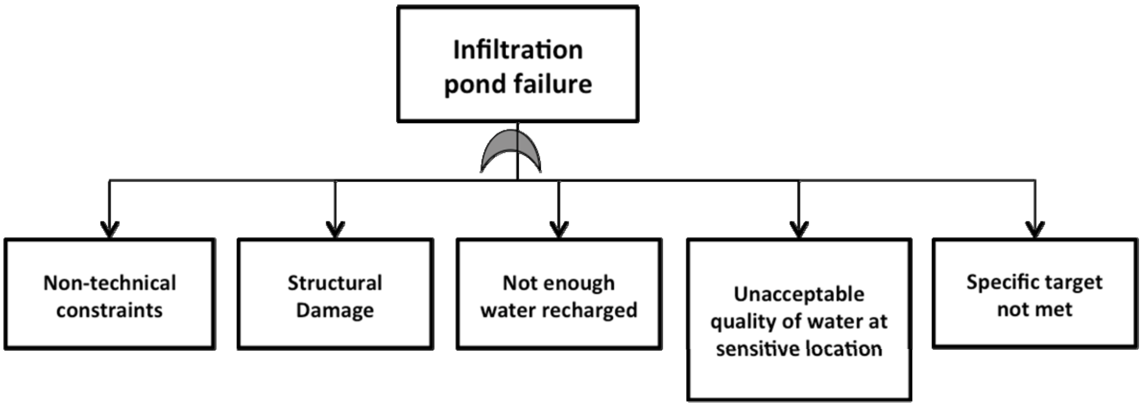

Figure 4.1. Fault tree for the total system failure

In this context, **we say that the MAR facility fails if it cannot function properly and has to cease operations for a prolonged time, t, larger than a prefixed one, T (i.e., t>T)**. Actually, a number of conditions can lead to failure, either having to abandon the facility completely, or temporally as the conditions for proper operation of the facility are not met (meaning it does not reach the objectives it was built for). We can consider 5 different alternatives for that, noticing that if any one of them actually happens, failure will occur. We thus consider the following five events:

1) Administrative/Management problem (Non-technical constraints):
- changes in the social or political interest for artificial recharge, maybe caused by the perception of society toward reutilization
- legislative initiatives
- economic problems, such as a crisis or the potential appearance of new techniques that are more efficient and less costly, but also changes in water pricing that may lead to eventual abandonment of the facility.

2) Structural damage in the civil works
- catastrophic events caused by natural disasters or terrorist activities
- structural damages due to construction problems

3) The quantity recharged does not reach some target value that makes it economically feasible
- not enough water available
- water available does not reach the quality standards needed to allow it to be used in the recharge facility

4) The water finally resulting in the aquifer does not meet some quality standards once it reaches some sensitive location (river, supply well, wetland, …)

5) Specific target not met,
- Seawater intrusion is not sufficiently contained
- A protected water body is reached by polluted water
- Water levels at target surface water bodies (river, spring, wetland) are not reached

Figure 4.1 displays an “OR” Boolean gate, precisely indicating that if any of the events occur, then system will fail. Furthermore, all events can be considered independent or weakly dependent. Thus, each event can then be addressed individually, and we will address now all of them albeit to a different degree of insight. The complete fault tree is shown in Appendix 1.

4.2 Non-technical constraints

We start by analysing the "non-technical constraints" event, including administrative or managerial problems that may result in MAR practices being discontinued in a given facility. The fault tree including subevents is presented in Figure 4.2.

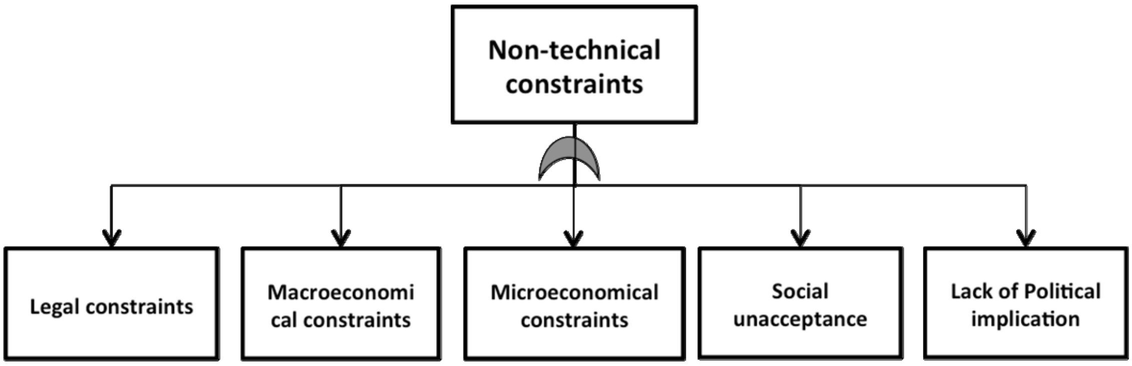

Figure 4.2. Fault tree for the event Non-technical constraints

Again we assume that this event is composed by a number of quasi-independent events. Obviously this is an oversimplification. Yet, it would be possible to treat non-independent events using probability theory. We illustrate it with a simple example, considering that social and political implications are strongly correlated, and disregarding the remaining 3 boxes in Figure 4.2.

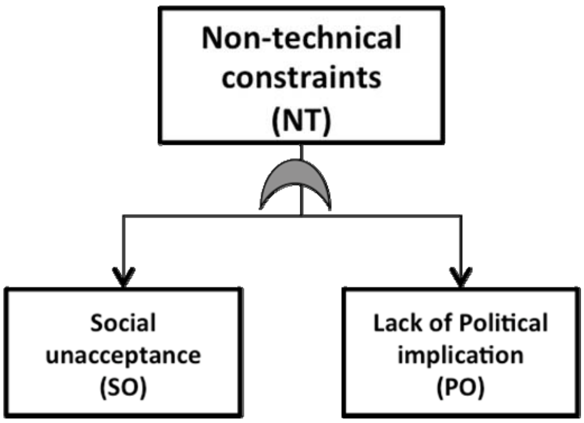

Figure 4.3. Simplified illustrative case for Non-technical constraints involving only sociopolitical constraints

From Figure 4.3, now we will obtain P[NT] from those of the subevents

$$P(NT) = P(SO \cup PO) = P(SO) + P(PO) - P(SO \cap PO) \quad (4.1)$$

where the last term is given by conditional probability assuming that politics respond directly to social concern (which is not necessarily true).

$$P(SO \cap PO) = P(SO) \cdot P(PO / SO) \quad (4.2)$$

4.3 Structural damage

Without being exhaustive, there are a number of hazards that, if ever occurring, may result in catastrophic disasters for the facility. Most of them have a very low probability of occurrence, but would lead to a high cost of repair maybe leading to the total abandon of the facility.

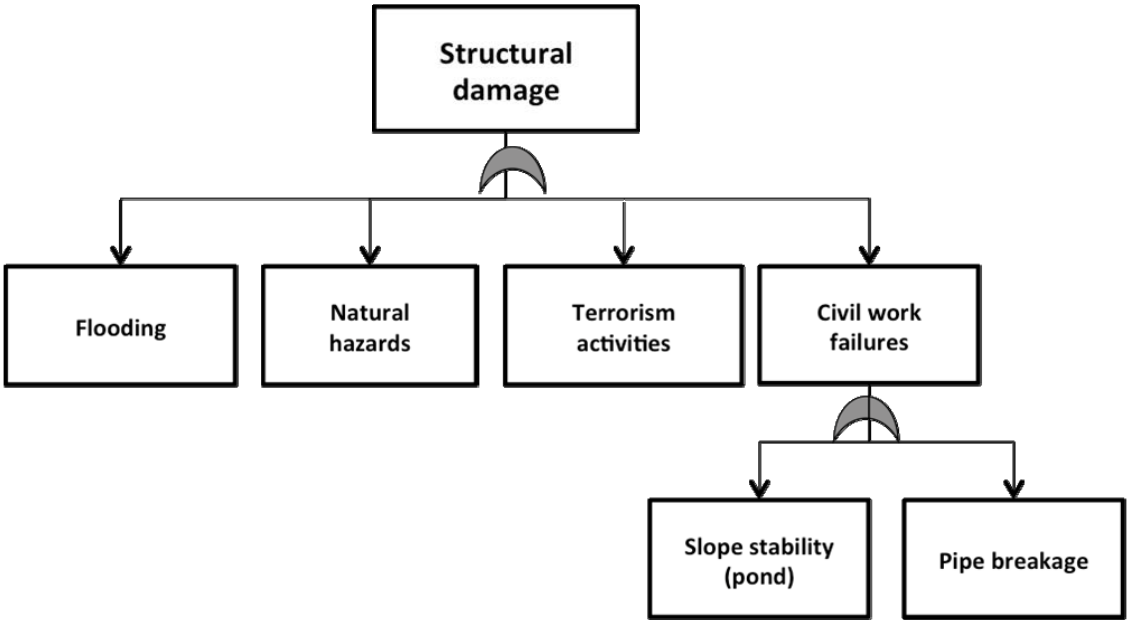

Figure 4.4. Simplified fault tree for the Structural Damage subevent

Several considerations arise from Figure 4.4. We combine in principle catastrophic (the system will be too expensive to repair, or will take very long time) and non-catastrophic events. In the latter case (e.g. the pipes bringing water to the system break), repairing time is the issue, and if this is acceptable then it should not be included. The separation of these events is somewhat arbitrary, as flooding has been separated from other natural hazards. Here we consider landslides, earthquakes, volcanoes, tsunamis,… Obviously the list is completely site-dependent, and the managers of the facility should make a detailed list of the ones that may potentially affect any specific site.

Probabilities may or may not be difficult to obtain. In some particular cases it would be possible to act to reduce this probability. An example would be that of a terrorist action, where probability of the facility being affected can be reduced if deemed necessary by increasing surveillance.

In Figure 4.4 we want to stress the benefit of a fault tree approach, by looking at the last block. Here we consider all structural damages that may take place in the facility. We could use a simple block to be evaluated by construction experts as a whole, or else break it down into subevents that might be evaluated individually. Thus, we can study in detail the stability of the slopes before the facility is put into operation (first filling of the pond would be the most critical time), while another team can take care of evaluating the probability of breakage of the pipes bringing the water to the facility. The more independent the event is with respect to the combination of all subevents, the more we can take advantage of breaking large problems into small pieces.

Moreover, we can start by doing a very preliminary analysis of slope stability (SS), using for example Bishop's method (thus involving a 2D analysis). This will result in a given probability of failure, let us say $P(SS)$. Now we can decide whether we can proceed further or not. This probability value would lead to a total probability of failure once combined with the remaining probabilities by the Boolean operators within the complete fault tree (Appendix 1). If the contribution of $P(SS)$ to total probability is small or results in an acceptable risk, then we can end up the process of risk evaluation here. Alternatively we can find that this probability is too large, and then we can, for example, re-evaluate the probability of failure using a more sophisticated numerical code (involving for example a 3D analysis by means of finite elements approach). Being a less conservative, the new probability will possibly be smaller than the previous one, leading to a smaller overall risk.

This is precisely the strength of an approach based on fault trees. Simple evaluations of the individual events can be done in a preliminary phase, then check for the paths leading to the highest risk (based on the combination of subevents), and finally spend the efforts into the proper re-evaluation of such subevents. This may involve the need to spend money in additional monitoring, or re-evaluation of existing data. One question is whether the additional expenses needed for a re-evaluation of risk are worth.

Similar approach is considered for flooding or other natural hazards. Again, simple models based on statistics can be used for preliminary estimation of probability of failure. If such values are considered unacceptable, it is possible to reconsider them by using more sophisticated modelling approaches. An example would be the use of code HEC-RAS to compute the probability of flooding of any given facility located close to the river.

4.4 Quantitative problems

One of the main reasons for failure of an MAR facility relies on quantitative issues. Failure can be defined as the lack of recharging enough water to meet the target (based either on economical or on environmental conditions). Here fault trees are quite complicated, as there are quite a large number of events and subevents that may lead to failure in the system based on quantitative aspects.

The main tree leading to failure in terms of quantitative aspects is presented in Figure 4.5. In short, there are three events that may lead to failure, being that: (1) the available water to be supplied to the system is not adequate in quality (event AQ) and thus cannot be used for recharge; (2) there is simply not enough water available for recharge (event NE), or (3) the system is not capable of recharging a sufficient amount of water due to insufficient vertical permeability, mainly caused by the development of surface or deep clogging (event CL).

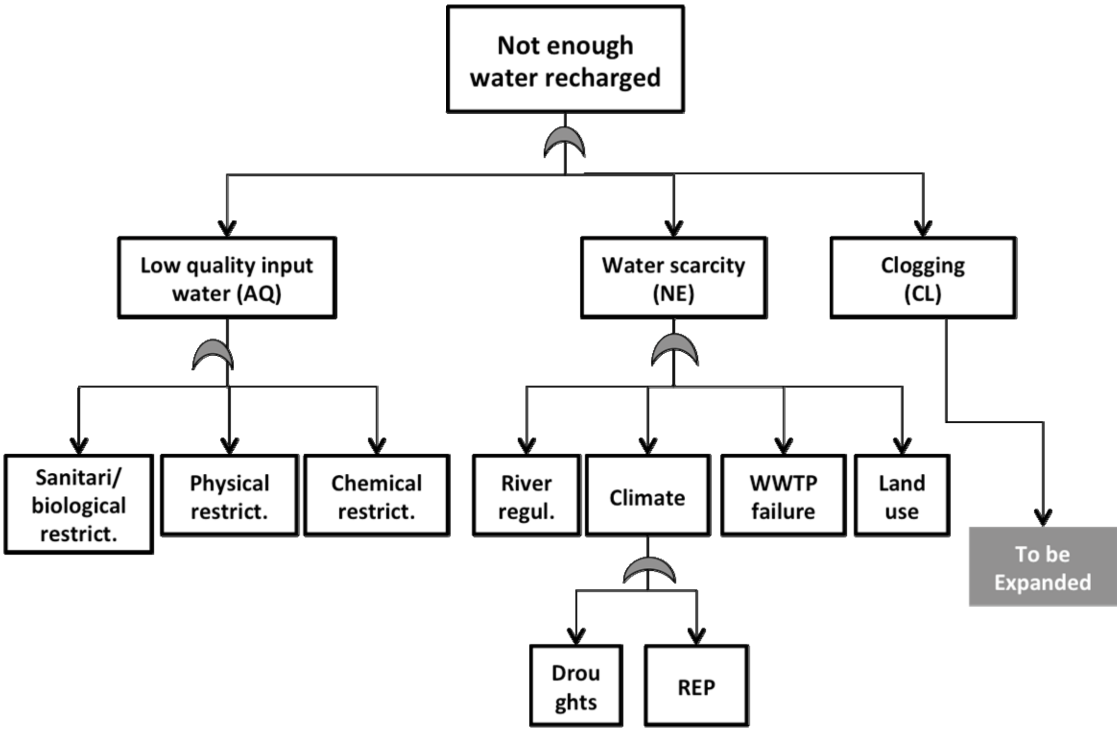

Figure 4.5. Fault tree for Infiltration Pond System Failure based on quantitative restrictions

Regarding event AQ, we should indicate the potential restrictions in the input water. In principle this depends whether the water sent to the pond comes from river water, reclaimed, desalinated, or stormwater. The quality restrictions may include: (1) physical aspects, such as turbidity; (2) chemical species, such as ammonia (NH_4^+) or chloride (Cl^-) content, or the presence of emerging contaminants; or (3) biological signature, including the undesired presence of pathogens, bacteria or viruses. Such restrictions depend on local or national legislation, are subject to social and political restrictions, and may change with time.

The next event to address is NE. Water scarcity is associated to a number of subevents depending on the input source. In the case of river water, the amount of water could be scarce related to the extension of droughts, or the increase in river regulation, both resulting in the reduction in river flow. Global change issues might be included here. Depending on the source of recharge water, the

NE event may be caused by a number of subevents. For example, a failure in WWTP may lead to water of insufficient quality to be supplied to the infiltration pond that could be solved by the inclusion of new treatment processes within the plant. Also, changes in land use may lead to a different spatial or temporal pattern of run-off, so that the quantity of storm water reaching the facility could be put in jeopardy.

Lastly, clogging (CL) can become a very significant event leading to MAR failure. When the clogging process reaches a certain value, the infiltration rates may be not enough to guarantee a correct MAR operation. The fault tree corresponding to clogging (Figure 4.6) involves its different causes (physical, biological, chemical, irreversible soil compaction, or due to gas generation). Each type of clogging has different causes, in principle quasi-independent with respect to the others, and could be further developed. It is significant to state that in most cases clogging is reversible by means of surface operations such as periodic surface scrapping; that is clogging does not necessarily imply system failure. Here we consider only the case when clogging exceeds expectations, resulting in the impossibility of recharging some target flow rate despite all remediation actions.

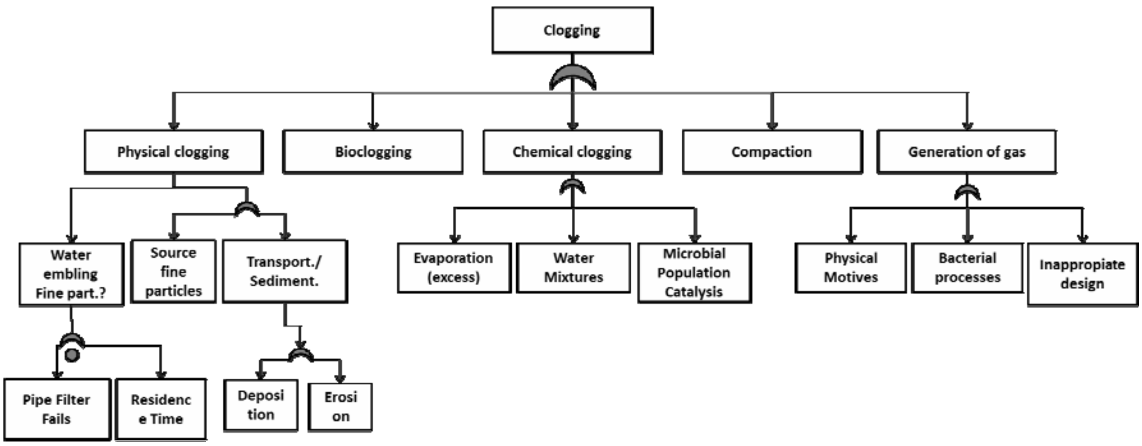

Figure 4.6. Fault tree including events that may lead to clogging

4.5 Qualitative problems

Another main reason for failure of MAR relies on qualitative issues. Failure is defined as the presence of hazardous compounds above predefined thresholds reaching some surface of interest (drinking supply well, ecosystems of interest,…). The compounds are divided in three main families: compounds from inefficient degradation of MAR (Inefficient Attenuation event), compounds produced during the MAR process or during the transport from the pond to the comply surface (event Metabolites), and mobilization of toxic metals within the aquifer due to spatial and temporal changes in the geochemical setup (event Mobilization). Then, the fault (sub-)tree is based on these three main events (see Figure 4.7).

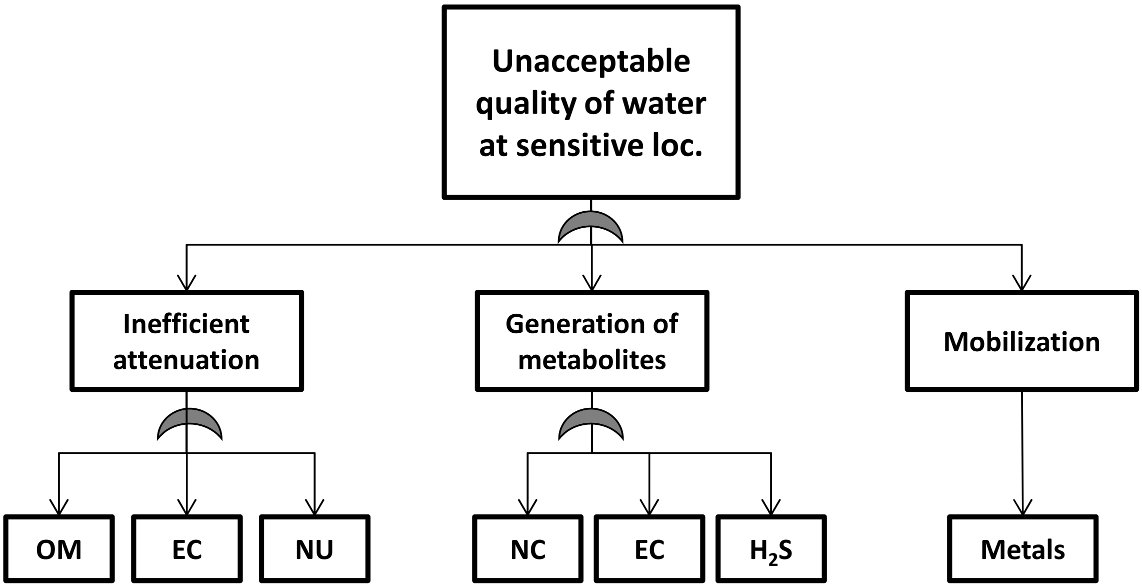

Figure 4.7. Fault tree for events that may result in unacceptable water quality at some sensitive/target location

The potential accumulation of unsafe compounds can be due to inefficient attenuation. In this case, the hazardous compounds to be considered are the ones that are already present in the recharged water, including, among others, total organic matter (OM), nutrients (NU) such as nitrate, ammonia or sulphate, and emerging compounds (EC). The EC encompassed all the chemicals that are present in consumer products such as pharmaceutical and personal care compounds. Although its concentration is between nanograms to micrograms per litre, the attention to EC is escalating due to their potential undesirable health effects in human and ecosystems.

The Insufficient Attenuation event can be studied by means of simple analytical expressions (exponential decay like), or else be included in sophisticated 3D stochastic subsurface reactive transport models. The actual approach can again be done in successive steps, going from the simplest to the more sophisticated ones, always depending on the relevance of this particular event into the total probability computed from the full fault tree.

Metabolite production is related to 1) Reduction-Oxidation processes linked to MAR activities, and to 2) general organic and specific emerging compounds potentially present in the recharge water. Here, redox metabolites refer to end or intermediate products that may arise from the nitrogen cycle (NC) or the sulphate cycle depending on the existing redox states. For example, in the case of the nitrogen cycle, nitrite is an intermediate compound which is substantially more hazardous than nitrate or ammonia, and then its accumulation during MAR could create an undesirable hazard. Another hazardous metabolite is H_2S, produced during sulphate reduction. Related to the emerging pollutants (EC), both direct degradation processes and cometabolism can generate metabolites, defined as compounds produced directly from parents by a breakage in the molecule or the exchanges with free radicals.

Lastly, the biogeochemical activity that degrades the OM and EC can modify both the pH and the Eh of the system. Depending on the mineralogy of the MAR site, some metals can be mobilized (Mobilization event). For example, changes in pH and Eh can produce arsenic, manganese, iron or uranium mobilisation.

4.6 Specific targets

The specific targets involve all the situations that have not been considered previously. Obviously, this is quite an open event, and each MAR facility can come up with individual situations that may potentially occur. A few of such events are specified in Figure 4.8.

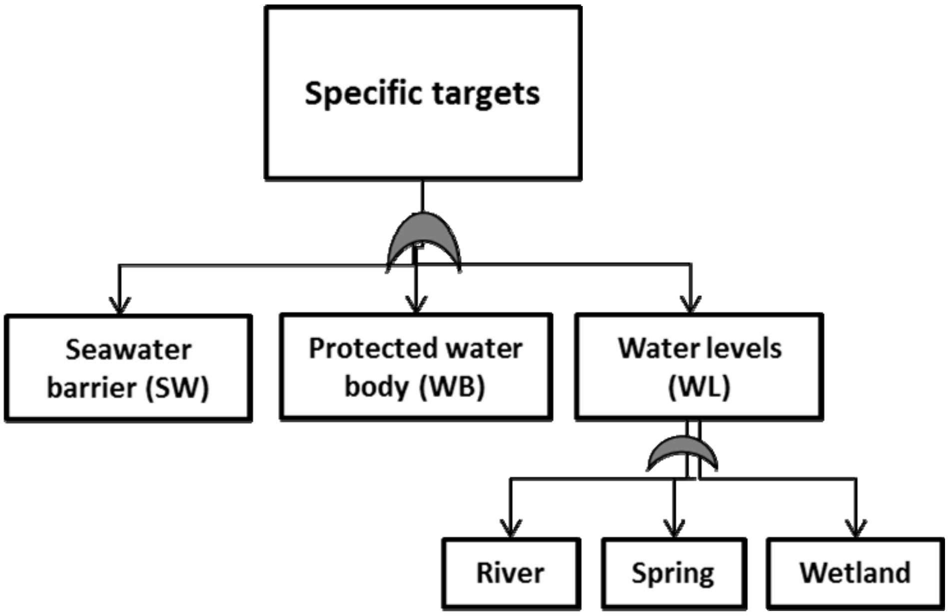

Figure 4.8. Fault tree for failure caused by not meeting specific targets

The objective of MAR facilities is most often the increase in water resources of good quality. Nevertheless, in some cases the facilities are designed as a way to control the regional hydrological regime. Thus, by artificially recharging an aquifer by means of infiltration ponds, (fresh-)water levels are raised. Then, by manipulating these levels, a number of elements can be controlled: such as the rate and direction of flow in order to protect subsurface water bodies (event WB), to prevent seawater intrusion (event SW), or to control the impact of groundwater to the land or to surface water bodies (e.g., spring discharge, land subsidence, and seepage to or out of rivers and lakes).

Each one of these events can be evaluated by means of local models to evaluate the potential for failure.

References

Barbieri M, Carrera J, Sanchez-Vila X, Ayora C, Cama J, Köck-Schulmeyer M, López de Alda, M, Barceló D, Tobella J, Hernández M (2011) Microcosm experiments to control anaerobic redox conditions when studying the fate of organic micropollutants in aquifer material, J. Contam. Hydrol. 126: 330-345.

Bedford, T., R. Cooke, *Probabilistic Risk Analysis: Foundations and Methods* (2003), Cambridge Univ. Press, Cambridge, U.K.

Bolster D., M. Barahona, M. Dentz, D. Fernàndez-Garcia, X. Sanchez-Vila, P. Trinchero, C. Valhondo, D. M. Tartakovsky, Probabilistic risk analysis of groundwater remediation strategies, Water Resour. Res. 45 (2009) W06413, doi:10.1029/2008WR007551.

Cardiff M., X. Liu, P. K. Kitanidis, J. Parker, U. Kim, Cost optimization of DNAPL source and plume remediation under uncertainty using a semi-analytic model, J. Contam. Hydrol. 113 (1-4) (2010) 25-43.

Dean D.W., T. H. Illangasekare, T. F. Russell, A stochastic diferential equation approach for modeling of NAPL flow in heterogeneous porous media, Tech. rep., University of Colorado at Denver, Denver, CO (2005).

Environmental Protection Agency (EPA), Risk Assessment Guidance for Superfund, Vol. 1, Human Health Evaluation Manual (Part A) (1989), Rep. 1074 EPA/540/1-89/002, Washington, D.C.

Environmental Protection Agency (EPA), Risk Assessment Guidance for Superfund, Vol. III - Part A: Process for conducting probabilistic risk assessment (2001), Rep. EPA 540/R-02/002, Washington, D.C.

Fernàndez-Garcia D., H. Rajaram, T. H. Illangasekare, Assessment of the predictive capabilities of stochastic theories in a three-dimensional laboratory test aquifer: Effective hydraulic conductivity and temporal moments of breakthrough curves, Water Resour. Res. 41 (2005) W04002, doi: 10.1029/2004WR003523.

Maxwell R.M., W.E. Kastenberg, Y. Rubin, A methodology to integrate site characterization information into groundwater-driven health risk assessment, Water Resour. Res. 35 (9) (1999) 2841-2855.

Molin, S., V. Cvetkovic (2010), Microbial risk assessment in heterogeneous aquifers: 1. Pathogen transport, Water Resour. Res., 46, W05518, doi:10.1029/2009WR008036.

Nowak W., F. de Barros, Y. Rubin, Bayesian geostatistical design: Task-driven optimal site investigation when the geostatistical model is uncertain, Water Resour. Res. 46 (W03535) (2010) doi:10.1029/2009WR008312.

Persson, K., G. Destouni, Propagation of water pollution uncertainty and risk from the subsurface to the surface water system of a catchment, J. Hydrol., 377, 434-444 (2009).

Sanchez-Vila X., A. Guadagnini, D. Fernàndez-Garcia, Conditional probability density functions of concentrations for mixing-controlled reactive transport in heterogeneous aquifers, Math. Geosci. 41 (2009) 323-351, doi:10.1007/s11004-008-9204-2.

Speek, A.J. (1981), Lifespan oral toxicity study of vinyl chloride in rats. Food Cosmet. Toxicol. 19(3), 317-333.

Yu, W., C. Harvey, C. Harvey, Arsenic in groundwater in Bangladesh: A geostatistical and epidemiological framework for evaluating health effects and potential remedies, Water Resour. Res., 39(6), (2003), doi:10.1029/2002WR001327.

APPENDIX 1. COMPLETE FAULT TREE

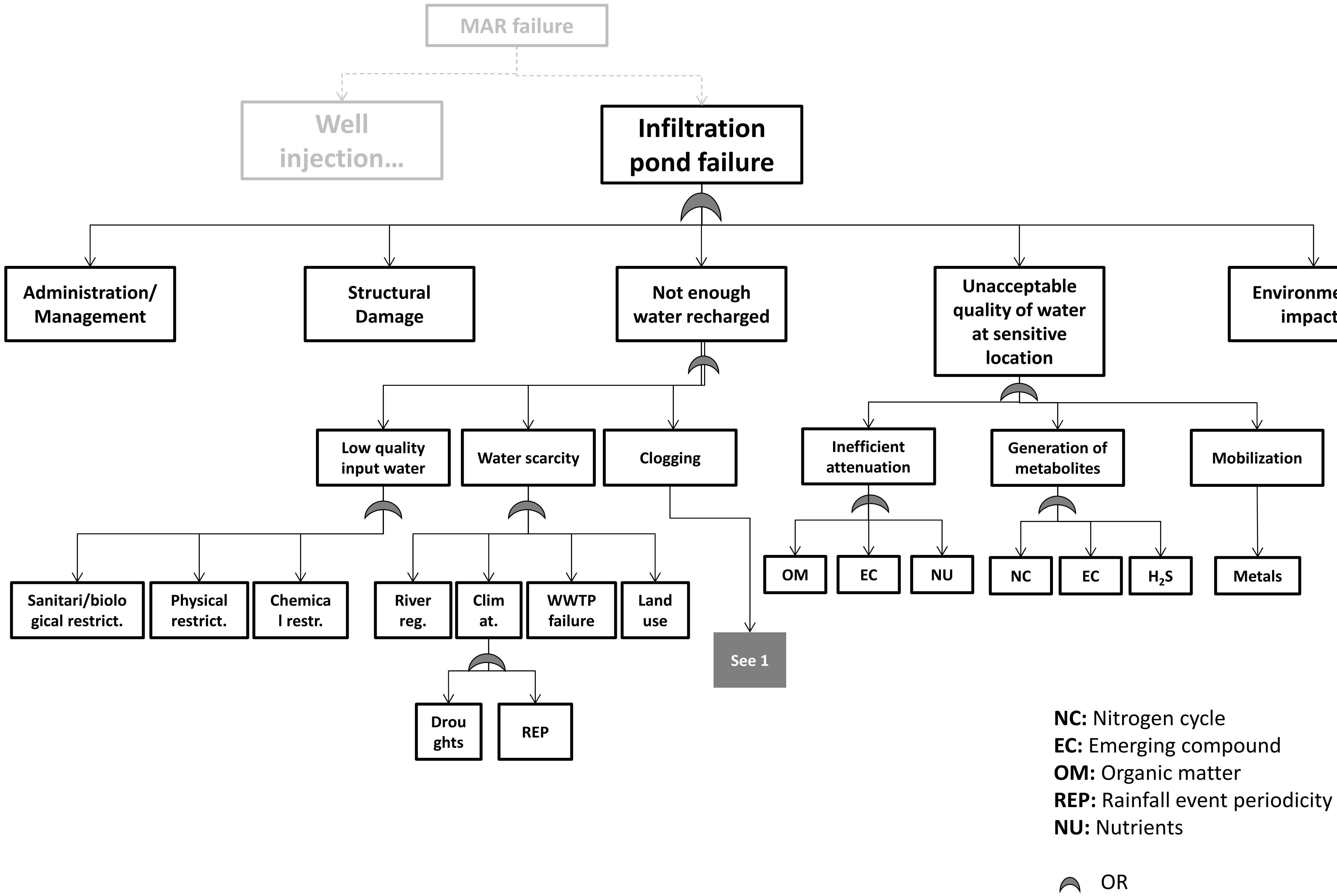

Figure AP1. Complete fault tree of MAR involving infiltration ponds.