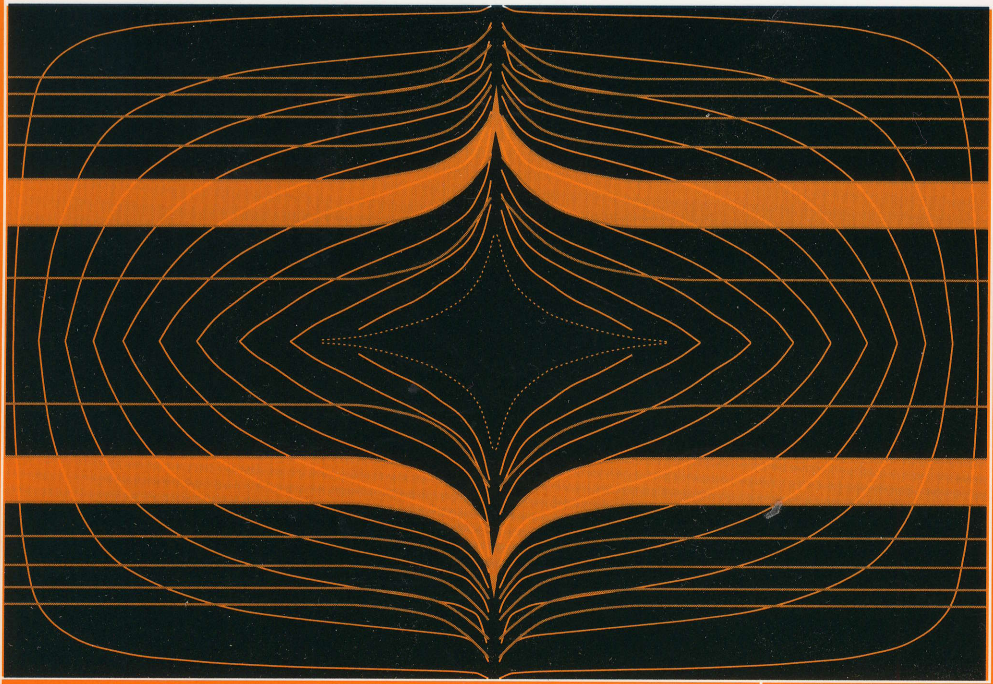
Strategies for regional groundwater quality monitoring

Hans Peter Broers



Nederlandse Geografische Studies
Netherlands Geographical Studies

NGS
306

Strategies for regional groundwater quality monitoring

ISSN 0169-4839

The research presented in this thesis was carried out at the Centre of Hydrology of the University of Utrecht and the Netherlands Institute of Applied Geoscience TNO - *National Geological Survey.*

This publication also serves as the thesis submitted for the title of Doctor, which was defended in public 16 October 2002 at Utrecht University

Promotors: Prof. dr. P. A. Burrough
Faculty of Geographical Sciences
Prof. dr. C. J. Spiers
Faculty of Earth Sciences

Co-promotors: dr. ir. F. C. van Geer
TNO-NITG
dr. J. Griffioen
TNO-NITG

Examination committee:
Prof. dr. ir. C. van den Akker
Prof. dr. S. E. A. T. M. van der Zee
Prof. dr. Ph. van Cappellen
Dr. J. J. B. Bronswijk
Dr. P. F. M. van Gaans

ISBN 90-6809-342-8

Printed in the Netherlands by Labor Grafimedia b.v. - Utrecht

Strategies for regional groundwater quality monitoring

Nederlandse Geografische Studies 306

Strategies for regional groundwater quality monitoring

Hans Peter Broers

Utrecht 2002

Koninklijk Nederlands Aardrijkskundig Genootschap/
Faculteit Ruimtelijke Wetenschappen, Universiteit Utrecht

Table of contents

Figures

Tables

1 Introduction

1.1 Scope and motivation of this study

The contamination of groundwater resources by diffuse (or nonpoint) sources is a serious problem in the EU and especially the Netherlands. Diffuse sources include agrochemicals, such as pesticides and nutrients derived from fertiliser and animal manure, and the atmospheric deposition of NO_x, SO_x and metals. In the Netherlands, contamination by agricultural sources has increased markedly during the last 40 years due to the increase of intensive livestock farming and the use of pesticides. Human impact on groundwater quality has been assessed and monitored at the national scale (van Duijvenbooden et al. 1985, 1993), at the regional scale (for example Broers 1996, Frapporti 1997) and at the scale of phreatic well fields (Baggelaar 1996). The results of these monitoring efforts have been used to define national and regional policy on soil protection and water management and in the prediction and protection of groundwater composition at drinking water well fields.

This thesis focuses on regional and local scale monitoring of diffuse contaminants in groundwater. The central hypothesis is that improvement of monitoring effectiveness is possible at these scales when using hydrological and hydrogeochemical information plus concepts of advective and reactive transport to steer the monitoring design and data analysis. The thesis presents approaches that incorporate hydrological information, such as travel times and groundwater ages, plus hydrogeochemical information into the analysis of regional and local scale monitoring network data, and into the design, evaluation and optimization of such networks. Its importance lies in the integration of knowledge on monitoring statistics, hydrology and hydrogeochemistry. It will be shown that the identification and understanding of spatial and temporal patterns in groundwater quality is strongly improved if groundwater age distribution and reactive processes are considered at all stages of monitoring, including the design and data analysis stages.

1.2 Previous work

During the last 10 years many contributions have been made to designing groundwater quality monitoring networks. Most contributions have concentrated on contamination problems caused by point sources and focused strongly on the statistical aspects of the choice of monitoring locations and sample size (Woldt & Bogardi 1992, Hudak & Loaigica 1993, Meyer & Brill 1988). Relatively little work has been published on the design or associated data analysis of regional networks or local scale networks around well fields. A brief overview of previous work on the following topics is subsequently discussed: (1) general background and monitoring network terminology, (2) regional groundwater quality monitoring, (3) local scale monitoring at phreatic well fields, (4) groundwater age and (5) reactivity of subsurface sediments.

In the thesis, a distinction is made between regional scale monitoring programs and local scale programs. The regional scale is defined as dealing with areas of typically 500-10,000 km^2. The local scale corresponds to the scale of a typical contributing area of a phreatic well field, which is about 50-100 km^2.

General background and monitoring network terminology
Eight main phases are distinguished in most monitoring studies, which typically proceed as an

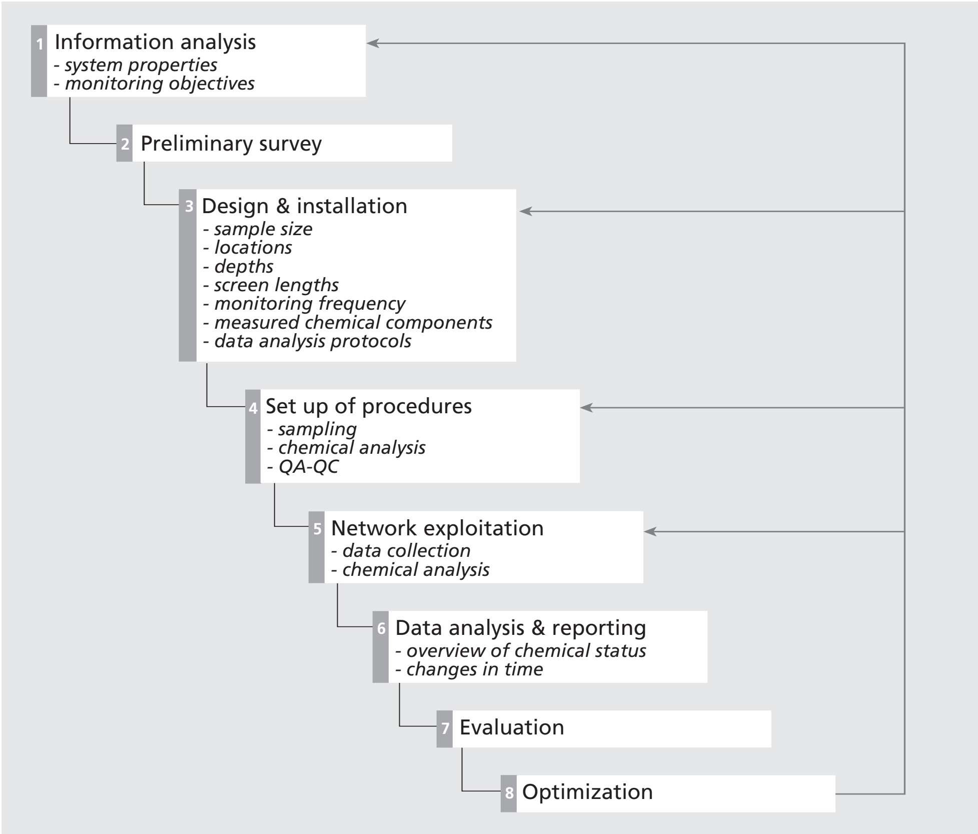


Figure 1.1 - The eight main stages in the operation of a monitoring network (modified from Ward et al. 1990)

iterative process; (1) information analysis, (2) preliminary surveys, (3) design and installation, (4) set up of procedures, (5) network exploitation, (6) data analysis and reporting, (7) evaluation and (8) optimization (Figure 1.1).

Stage 1

The first stage of information analysis includes an evaluation of the system and the definition of the objectives of monitoring (Figure 1.2). Three main factors determine the groundwater composition and are used to describe the system properties.

The first factor is the input of solutes into the groundwater. These inputs are derived from atmospheric sources, diffuse surface sources and point and line sources of contamination. Atmospheric inputs include sea spray and dispersed airborne compounds emitted by industry, vehicles and agricultural practices. Diffuse surface sources are often related to agriculture and may include manure, fertiliser and pesticides. Point sources include waste dumps and organic spills released at factories and petrol stations, for example. Line sources often relate to pesticide use at railways and the application of road salts. Diffuse sources require different approaches for assessment and monitoring than point and line sources because the contaminants are spatially dispersed and frequent in occurrence (Alley 1993).

Second, groundwater composition is determined by the hydrologic pathways through the

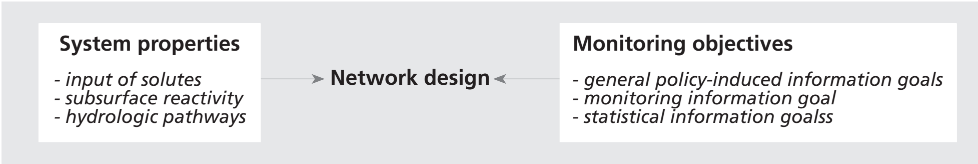


Figure 1.2 - The design of a monitoring network is tuned to the properties of the system being studied and the objectives for conducting monitoring. Terminology is explained in the text.

unsaturated and saturated zones. These pathways are determined by the regional and local geomorphologic patterns, the position of surface watercourses and by the hydrogeological properties of the subsurface. The hydrologic pathways influence the age distribution of the groundwater, which is of primary importance to the advective transport of contaminants that are introduced in recharging groundwater. The hydrologic pathways also determine which reactive sediments or rocks are encountered (Engelen 1981, Appelo & Postma 1993).

The third factor is the reactivity of the porous medium through which the groundwater passes. Hydrogeochemical reactions between the groundwater and the reactive portions of the porous media may alter the original groundwater quality by adding and/or removing solutes via processes such as mineral dissolution and precipitation, cation-exchange and redox reactions.

Ideally, networks for monitoring groundwater quality are set up to account for all three of the above mentioned factors, although it might not be possible to address all adequately during the design stage of the network. The weight of the individual factors in the design process is dependent on the objectives for conducting monitoring.

The definition of monitoring objectives and specific information requirements is probably the most important part of monitoring network design. Three subsequent levels of information goals are distinguished (Adkins et al. 1995): (1) *general, policy-induced or regulatory information goals*, (2) *monitoring information goals*, and (3) *statistical information goals* (Figure 1.2). The first two information goals are normally qualitative statements that describe the motives for conducting monitoring. For example, *a general information goal* might be: determine the effects of the excessive use of manure to groundwater resources. The corresponding *monitoring information goal* is more specific and would be: determine the concentrations of nutrients in sandy areas with agricultural land use in the first 25 m of the subsurface. The *statistical information goals* indicate the statistics that will be derived including a measure of desired precision. For example, *a statistical information goal* could be defined as: determine the percentage of nitrate contaminated groundwater in a specific area, with a relative precision of 10% at 95% confidence level. A proper monitoring design already defines the data analysis methods in fixed protocols. In practice however, the statistical information goals are often not stated specifically, which may lead to the 'data-rich but information-poor' syndrome described by Ward et al. (1986) where a lot of data is collected without anyone knowing how to handle the data.

Stage 2

When designing a new network, a first indication of the variation of the concentrations of the chemical components can be assessed using a preliminary survey (Figure 1.1). Such a reconnaissance survey provides groundwater quality characteristics in the study area and helps to frame hypotheses about the magnitude and spatial patterns of groundwater contamination. The results can be used to define the required sample size, well locations, well completion and monitoring frequency (Alley 1993).

Stage 3

The design of a groundwater quality monitoring network includes answering the following questions (Figure 1.1):

1. how many wells are needed? (sample size)
2. at which locations?
3. at what depth?
4. what screen lengths?
5. how frequently will the wells be sampled?
6. which chemical components will be analysed?
7. which data analysis procedures will be used?

These seven design options are tuned to the properties of the system, which include the three determining factors for groundwater quality, and to the monitoring objectives (Figure 1.2). For example, the sample size is to be adapted to the expected amount of variation in the area under consideration and the desired precision of the results. The amount of variation is a function of spatial variations in the input of solutes, the hydrological pathways and the reactive properties of the traversed formations. Similarly, monitoring depth and frequency should be tuned to the horizontal and vertical flow velocities of groundwater and the corresponding residence times. In regional monitoring, the well locations are often selected using a stratified approach. Usually, the strata represent areas with similar properties concerning the input of solutes, the hydrogeological conditions and the chemical vulnerability of the subsurface. Monitoring objectives of regional networks often emphasize the estimation of contaminant concentrations under specific land use classes in different hydrogeological settings (next section).

Stages 4 and 5

Stages 4 and 5 represent the set up of procedures and the actual exploitation of the network. Procedures need to be specified for sampling, chemical analysis and quality assurance and quality control (QA-QC).

Stage 6

Data analysis and reporting normally includes two main issues: (1) the assessment of spatial patterns or characteristics of areas at a specific moment in time, and (2) the assessment of changes in groundwater quality in time. For example, the EU Water Framework Directive distinguishes: (1) reporting of an overview of groundwater chemical status and (2) detecting the presence of long-term anthropogenically induced upward trends in the concentration of pollutants (EU, 2000). Ideally, protocols for data analysis and reporting are already established in the design stage of the network.

Stages 7 and 8

Data analysis methods and information goals may change after new information and new insights evolve during monitoring. In fact, groundwater quality monitoring is often a sequential approach, and network operation should be evaluated and optimized regularly. Each sampling round provides new information about variations in groundwater composition that can be used to evaluate the success of the design or to reframe or refine monitoring objectives and information goals (arrows in Figure 1.1). Multiple sampling rounds are required to evaluate the monitoring frequency in order to acquire adequate information on temporal variations. Given the small flow velocity of groundwater, long monitoring periods are needed for a proper evaluation of temporal variations.

Regional groundwater quality monitoring

Alley (1993) provides an excellent overview of all aspects of regional groundwater quality assessment and monitoring. Regional groundwater quality assessment programs are usually based on existing observation points, such as domestic wells, public supply wells or existing observation wells for groundwater heads (Wolter et al. 2001, Nolan et al. 1997, Cain et al. 1989, Chilton & Foster 1996, Kolpin et al. 1998). Many countries have set up assessment programs for groundwater quality (Chilton & Milne 1994, Jedlitschka 1996). For example, the USA National Water Quality Assessment Program (NAWQA) was set up to determine the quality of recently recharged groundwater beneath specific land use and for different hydrogeological settings (Leahy et al. 1993, Mueller et al. 1995). Nolan et al. (1997) assessed nitrate contamination in the NAWQA using information on nitrogen input, population density, soil properties and woodland/cropland ratios. They conclude that well type (domestic/ public supply) largely influences the measured nitrate contents. Pumping wells generally have large screen lengths and attract water from various depths. Samples from those are mixed samples of the entire pumped aquifer and cannot be compared with samples of observation wells which have small screen lengths and are pumped only for sampling. Alley (1993) states that 'as a general rule, existing and newly constructed wells should be considered as sampling different subpopulations'. National groundwater quality assessment programs in other large states, such as the UK and Germany, also make use of existing observation wells and pumping wells.

The Netherlands has installed a national monitoring network which consists of new observation wells with standardised dimensions and well completion, which is solely used for groundwater quality monitoring (Van Duijvenbooden 1993). The monitoring depths of the monitoring network were carefully chosen to be 10 and 25 m, using a hydrological concept of the age of the groundwater following Ernst (1973). Compared to the previous mentioned assessment programs, this network has the advantage of acquiring data from similar depths and age and of acquiring samples of similar representative sample volume.

The original objectives of the Dutch national network were (1) to investigate the quality of the groundwater in the upper aquifer in relation to land use, soil type and geohydrological conditions, (2) to determine the extent of human influence on groundwater quality, (3) to identify the changes of groundwater quality over time and (4) to provide data for good management of groundwater resources (Van Duijvenbooden et al. 1985, Van Duijvenbooden 1993). The network focuses on groundwater pollution from diffuse sources.

Land use and soil type have long been used as strata for groundwater sampling, but the Dutch geohydrological situation has neither been mapped nor used in the data analysis (Snelting et al. 1990, Reijnders et al. 1998). For national scale data analysis, this choice might be jusitified because the combination of land use and soil types also defines the major geo-hydrological situations in the Netherlands, as the land-use/soil type strata reflect different geomorphological positions in the Dutch landscape. For instance, the geohydrological situation in the grassland-fluvial clay areas is very different from the situation in grassland in the higher sandy areas of the Netherlands. However, at the regional scale the geohydrological situation shows large variation, especially in the vulnerable sandy areas. Here, the data analysis will benefit from extra geohydrological stratification.

Various authors have developed different approaches for the data analysis of the Dutch network because no data analysis protocol with specific statistical information goals was defined in the early stages of the network (Van Drecht et al. 1996, Reijnders et al. 1998, Pebesma 1996, Frapporti et al. 1993, Frapporti 1994). Three main approaches are distinguished that have different objectives.

The first approach is used for the regular, four-yearly reports on groundwater quality in the Netherlands. The approach aims at the estimation of the statistical distribution of concentrations of targeted contaminants for 14 land-use/soil-type strata (Van Drecht et al. 1996, Reijnders et al. 1998). The method has proven successful for delineating general patterns of targeted contaminants in groundwater in the Netherlands and to indicate proportions of contaminated groundwater within the land-use/soil-type strata. Later, results have been aggregated for larger connected areas, such as physical geographical regions and ecological districts. These areas have been used for maps that showed large differences in proportions of contaminated water over the country (for example Reijnders et al. 1998). Especially large differences exist between the Holocene and Pleistocene parts of the Netherlands.

The second approach aims at acquiring 'location-specific' estimates for 4x4 km² blocks, using kriging interpolation within 8 major land-use/soil-type strata (Pebesma 1996, Pebesma & De Kwaadsteniet 1997). The spatial interpolation yielded maps of groundwater quality for a large number of targeted chemical components using the 95 percent confidence interval on the estimated median for each 4x4 km² block. Consequently, these maps can be used directly for obtaining specific estimates for areas of 16 km². The results of both approaches show large variation and skewness of groundwater concentrations within the strata or km² blocks, especially in the vulnerable sandy areas of the Netherlands.

Frapporti (1994) argued that part of the large variation of targeted contaminants within the land-use/soil-type categories is due to geochemical processes. He proposed a third approach based on fuzzy c-means clustering of the whole set of chemical components measured in the network, yielding water types that correlate with the hydrogeochemical processes that occurred. Compared to the other two approaches, his work concentrated on general water quality and general indications of groundwater pollution emphasizing interrelationships between the chemical components, rather than regarding them as individuals. The distribution of his water types over the Netherlands is shown as point data. His approach was not aimed at estimating the statistical distribution within areas or 4x4 km² blocks, but concentrated on observations of groundwater quality in individual wells.

On the whole, the land-use/soil-type stratification was effective at the national scale in assessing large scale patterns of groundwater quality over the Netherlands. Pebesma (1996) suggested that his approach could benefit from the use of geohydrological information in defining the strata, thus reducing part of the observed variation in the land-use/soil-type strata in the sandy areas. A further improvement will be obtained if hydrogeochemical knowledge is incorporated in the data analysis of the land-use/soil-type/geohydrology strata, combining information of several related chemical components to assess hydrogeochemical processes and conditions that determine the fate and distribution of the targeted contaminants.

Local scale monitoring at phreatic well fields

Groundwater quality monitoring networks of phreatic well fields used for public water supply are established for other purposes than the regional monitoring networks. Monitoring objectives for these local networks are: the short-term safeguarding of public water supply and the signalling and prediction of future quality changes in the extracted groundwater (Jedlitchska 1996, Baggelaar 1992, 1996). In general, little work has been published internationally on the subject of monitoring of phreatic well fields. Baggelaar (1996) designed an overall monitoring strategy for the Dutch water supply companies which includes the monitoring of (1) the pumped groundwater, (2) the groundwater that is 10 or 15 years from the pumping well, and (3) the shallow groundwater in the protection area. The monitoring lay-out was determined mainly by the travel time distribution of the extracted groundwater. The lay-outs (2) and (3) were recommended only if the minimal travel times are less than 25 years. Baggelaar's network protocols included the design and installation and procedures for sampling, chemical analysis, quality assurance, data storage and statistical procedures for the analysis of the data. The travel time distribution within the contributing area of the well has been used to predict the groundwater quality evolution of the pumping well (Raats 1978, Beugelink & Mühlschlegel 1989, Van Brussel 1990, TCB 1991, Laeven 1997). Predictions of concentrations in the pumping well were compared with measured concentrations in the pumping wells, but seldom with measured concentrations in the observation network in the contributing area.

Stuyfzand (1996) proposed an additional approach based on geohydrochemical knowledge and argued that considerable cost reductions can be achieved if this knowledge is used to tune the set of chemical components to be measured to known processes occurring at the well field.

Groundwater age

The assessment of groundwater ages and isochrones is an essential part of the understanding of convective flow of contaminants in the subsurface (Ernst 1973, Raats 1978). Ernst and Raats demonstrated that isochrones are horizontal and groundwater age increases with depth for simple geohydrological situations in flat areas. This concept was used to define the monitoring depths of the Dutch national monitoring network (Van Duijvenbooden 1985). Monitoring depths of 10 and 25 m were chosen, assuming an average vertical groundwater velocity of about 1 m/yr in the shallow groundwater. Thus, groundwater ages of about 10 and 25 year can be expected in regional recharge areas and thick aquifers.

Several environmental tracers, including ^{3}H, ^{3}He, ^{14}C, ^{36}Cl, ^{39}Ar, ^{85}Kr and specific chlorofluorocarbons, are suitable for groundwater dating (Plummer et al. 1993, Coplen 1993). Tritium (^{3}H) is specifically suitable for dating of groundwater less than 50 years old (Plummer et al. 1993, Robertson & Cherry 1989, Engesgaard et al. 1996) and has been used successfully in interpreting monitoring results in the USA (Hallberg & Keeney 1993). Tritium measurements have been carried out in most of the wells of the national monitoring network and some of the provincial networks in the Netherlands. Meinardi (1994) used these tritium data to acquire information on the average groundwater recharge rates for regions in the Netherlands. Frapporti et al. (1993) used the tritium data of the national network to indicate the age of the samples in his water types. The tritium data of the regional networks of Noord-Brabant and Drenthe were used to test the geohydrological subdivision of the monitoring wells (Broers & Griffioen 1992, Broers 1993, Broers 1996a, Venema et al. 2000) and to interprete temporal trends (Broers & Buijs 1996). Recently, Bronswijk & Prins (2001) used tritium data to relate the proportion of nitrate contaminated groundwater in the national network to the infiltration year of the groundwater.

Little attention has been paid to regional scale variations in groundwater ages due to aquifer heterogeneity or complicated superficial drainage patterns of groundwater. An exception is found in the work of Modica et al. (1997), who simulated the groundwater age distribution below streams in a coastal plain aquifer. Their aim was to assess the complex age distribution of groundwater that discharges into surface water at different locations in the catchment. They concluded that steep lateral gradients in groundwater age were present, especially in downstream parts of the streams. In the downstream parts, relatively old groundwater from the regional flow system moved upward toward the centre of the stream. Lateral inflow of young groundwater discharged adjacent to the older water, and caused large variation of groundwater age over short distance perpendicular to the stream. These kind of variations in groundwater age might also explain much of the observed variation in contaminant concentrations in the vulnerable sandy regions of the Netherlands.

Reactivity of subsurface sediments

The role of reactive sediments controlling the quality of shallow groundwater has been recognised since the 1980's (Van Duijvenbooden & Waeghening 1987). In the Netherlands, the assessment of reactive properties of the soil zone (< 1.2 m depth) was integrated into standard soil mapping. These data have been used to derive vulnerability maps of the shallow subsoil of the Netherlands (Van Duijvenbooden & Breeuwsma 1987). However, there is virtually no information on the reactivity of deeper parts of the subsurface. The geochemical characterisation of the Dutch subsurface focused on the geological origin of the sediments and the total contents of elements in the sediments (for example Moura & Kroonenberg 1990, Veldkamp & Kroonenberg 1993, Huisman 1998).

Sampling of subsurface sediments to characterise their reactivity with respect to groundwater quality started only after 1985 (Broers 1988, Beekman 1991, Griffioen 1992, Geochem 1992, Von Gunten & Zobrist 1993, Matsunaga et al. 1993, Broers & Griffioen 1994). Methods derived from soil science, such as sequential extraction and selective extraction techniques, were used and adapted for the characterisation of reactive phases of subsurface sediments (Postma et al. 1991, Heron 1994, Griffioen & Broers 1993, Brown et al. 1999) and were used to relate groundwater composition to the reactivity of the sediments. Methods for characterising sediment reactivity have improved since then and kinetic experiments have been used to quantify the reactive capacity of sediments (Hartog et al. 2001). Others focused on the assessment of the spatial heterogeneity of geochemical properties in order to predict the effects on transport of contaminants in groundwater (Robin et al. 1991, Barber et al. 1992, Davis et al. 1993, Fuller et al. 1996, Allen-King et al. 1998).

The improved knowledge on subsurface reactivity and hydrogeochemical reactions has not yet been integrated in design and data analysis of regional monitoring networks. Moreover, there is an increasing need for specific information goals and standardised procedures for the sampling of sediment reactivity, to provide model input for prognoses of the evolution of groundwater quality at the regional scale and the scale of well fields.

1.3 Research issues and aims

The preceding discussion indicates that monitoring studies often focus on one of the individual fields of monitoring statistics, hydrology or hydrogeochemistry. The central hypothesis of this thesis is that more effective monitoring is achieved when using hydrological and hydrogeochemical information plus concepts of advective and reactive transport to steer the monitoring design and data analysis. In line with the hypothesis, the aim of the study is to integrate statistical, hydrological and hydrogeochemical methods and information in the design, the data analysis, the evaluation and the optimization of regional and local scale monitoring networks. The work is aimed at integrating methods and information from the fields of hydrology, hydrogeochemistry and monitoring statistics, rather than at adding new contributions in the individual fields.

The thesis focuses on the following six specific research issues that are related to regional groundwater quality monitoring.

Research issue 1 - Groundwater age distribution in regional monitoring
The work of Modica et al. (1987) has shown that groundwater age might vary substantially in drained areas and variations in groundwater age should be considered when designing monitoring networks in drained areas. Large variations in groundwater age are especially expected in areas with complicated geology or with a dense drainage network. For example, local and regional groundwater quality surveys have shown that groundwater age in regional discharge areas in the Netherlands could have ages up to 30,000 years (Stuurman et al. 1990) whereas groundwater ages of 10 years are foreseen at about 10 m depth in regional recharge areas.

For research issue 1, the aim is to investigate the influence of the drainage network and aquifer heterogeneity on the groundwater age distribution and to test whether regional mapping can be used to predict the age distribution in a regional groundwater quality monitoring network.

Research issue 2 - Integrating groundwater age and reactive processes in the design and data analysis of regional monitoring networks
Frapporti (1994) demonstrated the importance of geochemical reactions in the analysis of the Dutch national monitoring network data. He argued that these reactions could overrule the effects of land use, but made no attempt to translate the results to areas or land-use/soil-type/geohydrology strata. Pebesma (1996) and Reijnders et al. (1998) presented methodologies for characterising and mapping targeted contaminants. However, they neither interrelated the patterns of individual chemical components using geochemical knowledge, nor considered the hydrological position of the monitoring wells. Data analysis in the framework of the national and regional monitoring networks would obviously benefit from an integrated approach that combines hydrogeochemical and hydrological knowledge with the estimation of typical values and proportions of contaminated groundwater in the sampled areas.

For research issue 2, the aim is to integrate information on groundwater age and hydrogeochemical processes in the design and data analysis of regional monitoring networks and to investigate if such an approach yields extra value in the identification of groundwater quality patterns. Two hypotheses are examined: (1) extra geohydrological stratification helps to reduce the variation in the data, and (2) extra indicators based on geochemical knowledge provide a better identification and understanding of contamination patterns than the sole statistical analysis of concentrations of targeted contaminants.

Research issue 3 - Evaluation and optimization of regional monitoring networks
Two types of monitoring information goals are often present in the operation of regional monitoring networks: (1) the desire to monitor all areas or strata in a specific region thereby affording a spatial overview of groundwater quality, and (2) the desire to monitor targeted contaminants in areas with high vulnerability to the contamination of deeper groundwater, satisfying strict precision criteria. In the present design of the Dutch regional networks neither the monitoring information goals were clearly defined, nor were the corresponding statistical information goals specified. Monitoring efficiency will greatly benefit from defining a framework for evaluation and optimization, which includes specific information goals for areas with different vulnerability and pollution loading.

The aim for research issue 3 was to design a framework for the evaluation and optimization of a regional groundwater quality monitoring network, using specific information goals and ambition levels for areas with low, moderate and high risks for the contamination of deeper groundwater.

Research issue 4 - Detection and understanding of temporal changes in regional monitoring
Changes in agricultural practices are expected to affect groundwater quality by changing the loads of nutrients and salts in the recharging groundwater, but regional monitoring networks installed to register the changes often fail to detect them and interpretation of trend analysis results is difficult. For example, temporal trends for nitrate in the Dutch monitoring networks could not be detected in any of the land-use/soil-type strata (Van Drecht et al. 1996, Reijnders et al. 1998) although time series of 10 years were available and use of animal manure increased markedly in recent decades. Probable reasons of not detecting groundwater quality changes are: (1) the long travel times from the groundwater recharge locations to the well screens, (2) the

obscuring, attenuating or retarding effect of physical and chemical processes on solute breakthrough, (3) the spatial variability of contaminant concentrations in recharging groundwater, in hydrologic residence time, and in reactive properties of the aquifer sediments, and (4) the short-term natural temporal variability of groundwater composition at the monitoring depths. Following reasoning (1) and (2), changes will not be detected if the polluted groundwater has not yet arrived at the well screens or because chemical reactions retarded or transformed the chemical species of interest. The design of most networks is based on conservative transport, and monitoring depths and frequencies are based on this assumption. This leads to the hypothesis that trends in reactive systems will be identified earlier when chemical properties are studied that behave conservatively under specific conditions and when groundwater data from similar geohydrological situations and groundwater age are used in the analysis.

Normally, trend detection is limited to the analysis of time series of individual wells or groups of wells. Hallberg & Keeney (1993) showed that trends in groundwater quality can also be detected by using concentration-depth information, because depth and groundwater age are interrelated, which is most pronounced in recharge areas. Combining time series information, concentration-depth information and groundwater age dating will probably help to identify and explain changes of groundwater quality in time. A further improvement in the understanding of trends is anticipated by comparing detected trends at specific depths with concentration-depth prognoses based on information on the input history of solutes that were introduced in recharging groundwater.

Hence, the aim of research issue 4 is to improve the detection and understanding of quality changes in time in reactive groundwater systems, combining time series information, concentration-depth profiles, age dating and concentration-depth prognosis based on data on the historical input of solutes.

Research issue 5 - Monitoring configurations at phreatic well fields

Observation networks around vulnerable phreatic well fields are often installed with three main monitoring objectives: early warning, prediction of future groundwater quality and evaluation of protection measures (for example, Baggelaar 1996). Monitoring configurations were often based on a horizontal two-dimensional concept of groundwater flow, focusing on early warning using wells at 10 or 15 years transit time to the pumping well. Interestingly, the design of the Dutch regional monitoring networks (research issues 1 to 4) emphasised the vertical flow component and the residence time from the earth surface to the observation screens (Van Duijvenbooden 1985). Monitoring at phreatic well fields will benefit from combining these two concepts by assessing the three-dimensional travel time distribution.

The aim of research issue 5 is to judge the effectiveness of monitoring configurations for phreatic well fields, using a three-dimensional travel time approach and scenarios with advective and simple reactive transport.

Research issue 6 - Sampling reactivity

Sediment reactivity is one of the largest unknowns in the interpretation of groundwater quality data and the largest impediment for prediction of groundwater quality changes. Methods have been developed to characterise sediment reactivity of aquifer sediments. These methods are generally applied to samples collected for research projects at local scale. This includes the sampling of very small sample volumes and very high vertical resolution of measurements (for example Postma et al. 1991, Davis et al. 1993). For prognoses of the evolution of groundwater quality at the scale of well fields, there is an increasing need for specific information goals and standardised procedures for the sampling of reactive sediments.

The aim of this study was to formulate sampling objectives and initiatory data analysis protocols for the reactive properties of drinking water well fields, in order to produce input for transport models that are used to predict the evolution of the groundwater quality.

1.4 Structure of the thesis

Figure 1.3 shows how the six research issues fit into the general scheme for groundwater quality monitoring as introduced in section 1.2. Research issues 1 to 4 are covered in Chapters 2 to 5. Together, these 4 research issues address the 8 stages of regional groundwater quality monitoring, using the regional networks of Noord-Brabant and Drenthe as case studies. Research issues 5 and 6 relate to the monitoring of drinking water well fields and are tackled in Chapters 6 and 7. These research issues emphasize the information analysis and design and installation stages of the network operation. General conclusions and suggestions for further work are presented in the Chapter 8.

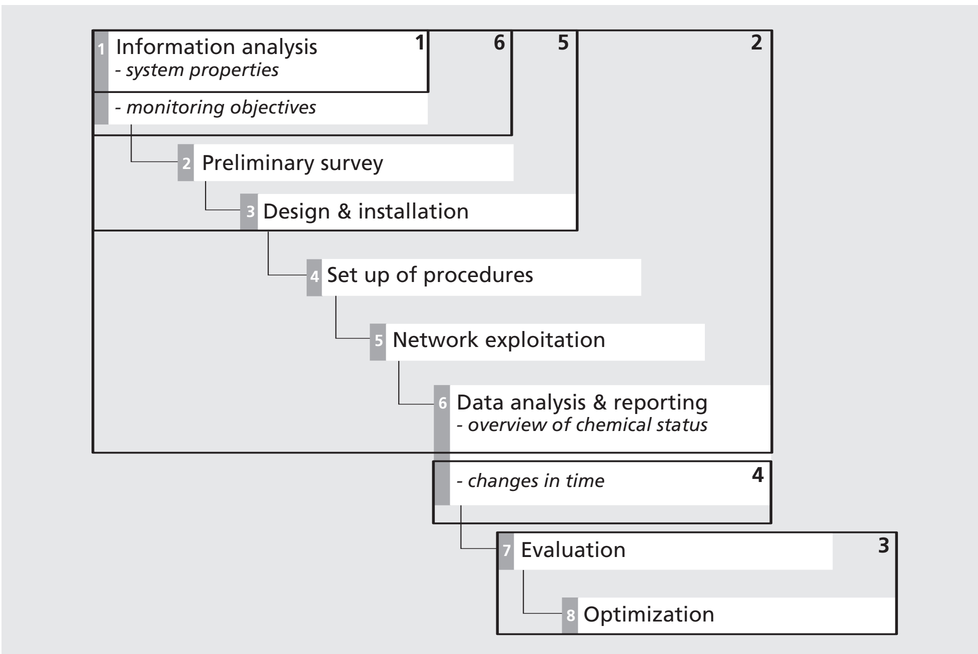


Figure 1.3 - Aspects of groundwater quality monitoring that are addressed in the research issues 1 to 6 (bold) which are elaborated in Chapters 2 to 7

2 The distribution of groundwater age for different geohydrological situations in the Netherlands: implications for groundwater quality monitoring at the regional scale

2.1 Introduction and background

Rationale and objectives

Contamination of groundwater resources by diffuse, surface related sources is a serious problem in the Netherlands. Contamination by agricultural sources has especially increased during the last 40 years due to intensive livestock farming and the use of pesticides. In order to assess and quantify the human impact on groundwater quality in time and space, a national monitoring network for groundwater quality was established between 1979 and 1992 (Van Duijvenbooden et al. 1985, 1993). Since 1989, regional monitoring networks have also been installed as an addition to the national network. The monitoring wells of the national and provincial networks were installed using standardised dimensions and well completion. The wells were screened at about 10, 15 and 25 m depth . The shallow screens (10 m) and the deep screens (25 m) are sampled annually and analysed for inorganic macro and micro constituents.

The screen depths of the monitoring wells were chosen using an elementary concept of the groundwater flow and groundwater age distribution in an aquifer characterized by groundwater recharge due to precipitation (Van Duijvenbooden et al. 1985, Snelting et al. 1990). For groundwater flow to a fully penetrating drain or watercourse, the following travel time distribution is used (Eldor & Dagan 1972, Ernst 1973, Raats 1977):

$$t_z = \frac{\varepsilon D}{N} \ln\left(\frac{D}{D - z}\right) \qquad (2.1)$$

where t_z = age at depth z [day], D = aquifer thickness [m], ε = porosity, N = groundwater recharge [m day^{-1}] and z = depth below land surface [m]. This equation is valid under the following assumptions:

1. the aquifer is homogeneous, isotropic and has constant thickness
2. groundwater flow is steady
3. the rise of groundwater table is small compared with the aquifer depth
4. the horizontal fluxes are constant over depth z (Dupuit assumption).

Equation (2.1) yields a horizontal pattern of *isochrones* (lines of equal groundwater residence time) which is shown in Figure 2.1a. The equation has proved useful for a range of Dutch conditions, because the Netherlands has a flat topography and thick, permeable aquifers (see also Chapter 6). For typical Dutch conditions, the equation predicts groundwater ages of 12-13 year and 33-40 years at 10 and 25 m depth, respectively, assuming N = 300 mm/year, ε = 0.35 and D = 50 to 100 m, (e.g. Meinardi 1994). Thus, the established monitoring depths seem to be suitable to determine the effects of diffuse groundwater contamination that was introduced during the last 40 years. The locations of the observation wells were chosen to guarantee homogeneous land use in the upstream catchment area (Figure 2.1b). In this way, the observation well should yield a vertical pattern of groundwater quality of increasing age with depth for a specific land-use.

Although the elementary concept seems suitable for the overall design of the monitoring locations and depths, deviations in the groundwater age distribution are to be expected due to aquifer heterogeneity and when a more complicated superficial drainage system exists. For example, one would expect that groundwater ages in regional discharge areas to deviate from the concept.

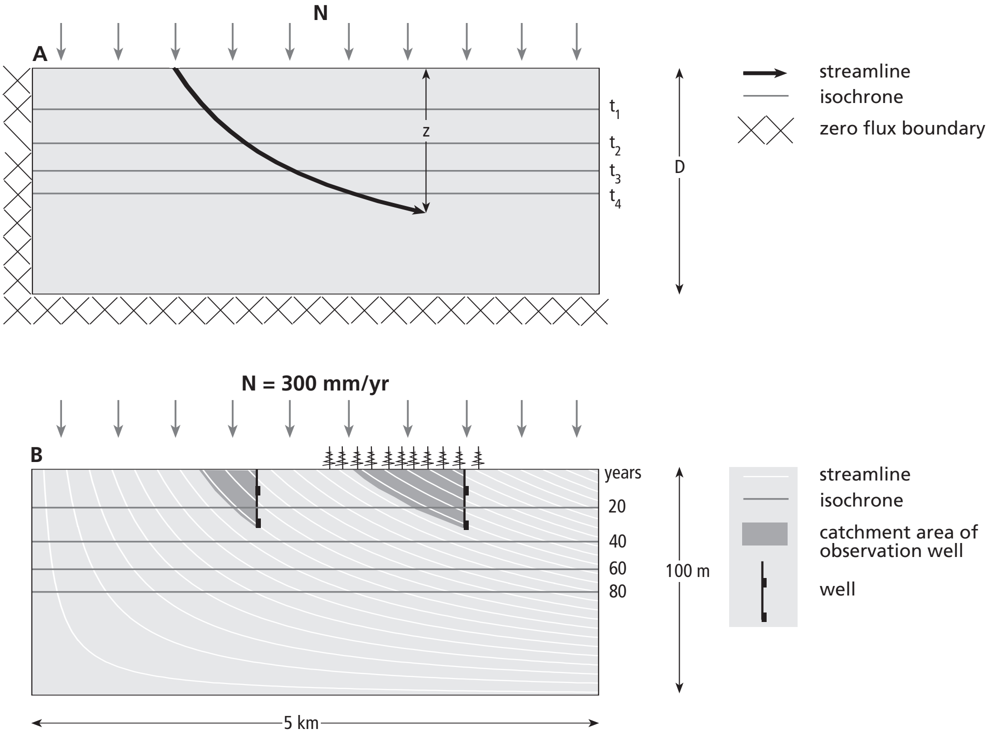


Figure 2.1 - Groundwater flow and isochrone patterns in a homogeneous aquifer with constant groundwater recharge drained by parallel fully penetrating ditches (after Ernst 1973) A. Elementary concept, B: Concept used for the set-up of the monitoring networks.

This chapter investigates the effects of a superficial drainage network and aquifer heterogeneity on the groundwater age distribution in aquifers in flat areas and presents the consequences for the monitoring of contaminants from diffuse sources. First, the effects are assessed using simulations of groundwater flow and groundwater age in different geohydrological situations. Second, the groundwater age distribution is evaluated for the two regional monitoring networks of Noord-Brabant and Drenthe using tritium measurements.

Geology and hydrogeology

Figure 2.2 summarises some relevant information on the geology and hydrology of the Netherlands and shows the positions of the monitoring wells of the Drenthe and Noord-Brabant regional networks. Geologically, the Netherlands is subdivided into a Holocene and a Pleistocene part. The low western part of the Netherlands comprises shallow Holocene marine and peri-marine deposits as well as fluvial deposits from the Rhine and Meuse rivers. The Pleistocene part of the Netherlands comprises older fluvial deposits and glacial and peri-glacial deposits at or near the surface. The provinces of Noord-Brabant and Drenthe are mainly located within the Pleistocene part of the Netherlands.

The altitude of Noord-Brabant ranges from 30 m above MSL (Mean Sea Level) in the south-east to 0 m above MSL in the north and west. The topography is determined by a buried horst-and-graben structure. The subsurface consists of older fluvial sand and gravel deposits from the Meuse river, overlain by a 2-30 m thick cover of fluvio-periglacial and eolian

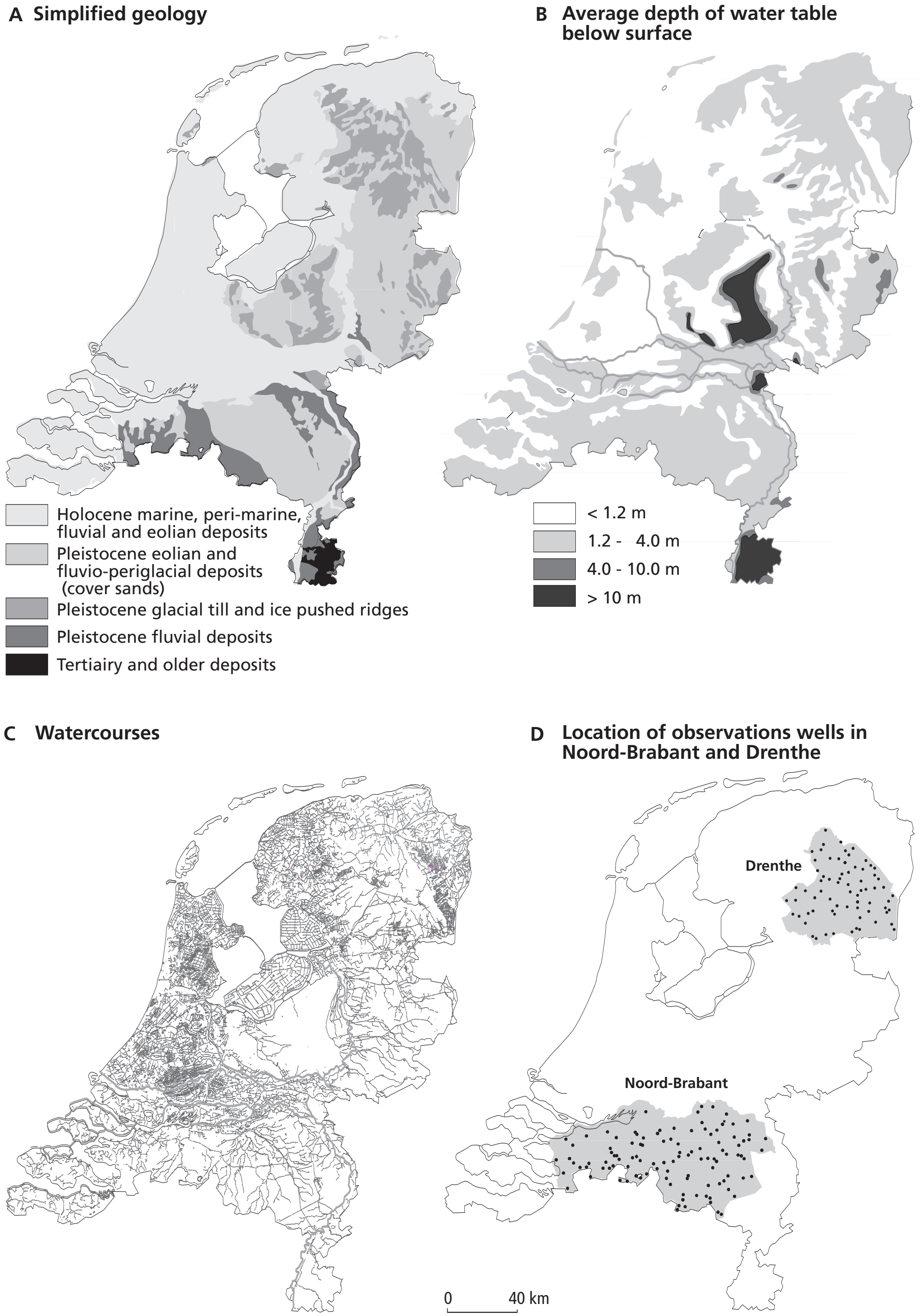


Figure 2.2 - Maps of (a) simplified geology, (b) depth of groundwater level, (c) watercourses in the Netherlands and (d) monitoring networks of Noord-Brabant and Drenthe

deposits consisting of fine sands and loam. In the western part of Noord-Brabant the top 20 m consists of estuarine clay and sand deposits from the Schelde estuary.

The province of Drenthe is situated on a glacial plateau, which is drained by several brooks. The topography of Drenthe ranges between 0 and 15 m above MSL. The subsurface mainly consists of heterogeneous glacial till (Figure 2.2a) underlain by a 200 m thick series of sandy fluvial deposits.

The provinces of Noord-Brabant and Drenthe are drained by a series of brooks. The extent and the position of these natural surface drainage networks are strongly related to the presence of local and regional groundwater flow systems (de Vries 1977, 1994, 1995). De Vries has shown that the stream systems in the Netherlands developed in equilibrium with the groundwater systems, and that the stream spacing and channel dimensions are controlled by subsurface permeability, rainfall characteristics and large-scale topography. Thus, the drainage network density partly reflects the permeability of the subsurface, and areas with shallow low-permeable layers have denser drainage networks than areas with shallow permeable layers. In both Noord-Brabant and Drenthe the original natural drainage network was artificially extended during the 20^{th} century, to allow for agricultural use of the poorly drained areas. This resulted in a dense network of ditches, drains and small watercourses (Figure 2.2c).

2.2 Simulation of the effects of drainage and heterogeneity on the groundwater age distribution

Model set-up and model scenarios

In order to gain insight in the effects of the drainage network and heterogeneity on the groundwater age distribution, model simulations were carried out. Groundwater flow and isochrones were simulated in a cross-sectional model, which describes groundwater flow in an aquifer of 100 m thickness and 5 km length. These dimensions are characteristic for the studied Pleistocene sandy regions of the Noord-Brabant and Drenthe, where the main brook valleys are separated by an average distance of about 10 kilometres (see Figure 2.2c). Different scenarios were evaluated to study the combined effects of drainage, aquifer thickness, variable conductivity and a spatially varying groundwater recharge.

The cross-section was modelled using FLOWNET (Van Elburg et al. 1993). FLOWNET is based on a 2D stream function approach (Bear 1972, Fogg & Senger 1985). The model contains 20 layers (rows) and 100 columns. The model cells had the resulting dimensions length x width x depth of 50 x 50 x 5 m. The left boundary represents the water divide between two watercourses and was modelled as a no-flow boundary. Groundwater recharge at the top boundary was introduced using a constant flux. Table 2.1 lists the parameters used in the base case model and the model scenarios.

The base case model was homogeneous and isotropic with conductivity (k) equalling 30 m day^{-1}. Figure 2.3 shows the streamlines and isochrones for the base case simulation (upper case). In the base case model, outflow occurred only in the drain in the upper right cell, which represents the major watercourse with water level of 0.0 m. This model scenario conforms almost completely to the analytical solution of equation (2.1) (depicted in Figure 2.1a). Contrary to the assumption in the analytical solution, the drain in the base case model was partially penetrating, which caused radial flow and upward bending of the isochrones in the direct surroundings of the drain. Equation (2.1) still appeared to be valid for 98 percent of the flow domain of the base case model. In the model, the drain affected the streamlines and the isochrones over a horizontal distance of about 100 m.

Changes were made to this base case model to identify the effects of heterogeneity and

Table 2.1 - Parameterisation model scenarios

Scenario		Aquifer thickness D (m)	Horizontal conductivity k_h (m day^{-1})	Anisotropy k_v/k_h	Groundwater recharge N (m day^{-1})
Scenarios without drains					
	Base case	100	30	1	8.2 e-4
A1	Reduced aquifer thickness	50	30	1	8.2 e-4
B1	Continuous layer with low permeability	30 (upper aquifer)	30	1	8.2 e-4
		10 (layer)	10	0.0025	
		60 (lower aquifer)	30	1	
C1	Moderately permeable shallow cover layer	20 (cover layer)	10	0.005	8.2 e-4
		80 (aquifer)	30	1	
D1	Discontinuous layer with low permeability	30 (upper aquifer)	30	1	8.2 e-4
		10 (layer)	10/10	0.0025/0.01	
		60 (lower aquifer)	30	1	
E1	Variable recharge	100	30	1	*
Scenarios with drains					
A2	Reduced aquifer thickness	50	30	1	8.2 e-4
B2	Continuous layer with low permeability	idem scenario B1	idem B1	idem B1	8.2 e-4
C2	Moderately permeable shallow cover layer	idem scenario C1	idem C1	idem C1	8.2 e-4
D2	Discontinuous layer with low permeability	idem scenario D1	idem D1	idem D1	8.2 e-4

* variable groundwater recharge alternating each 500m: 5.5e-4 and 8.2e-4 m day^{-1}

drainage. First, scenarios without drainage were used to identify the sole effects of an inhomogeneous subsurface (Figure 2.4 and Table 2.1, scenarios A1-E1). The scenarios include reduction of the aquifer thickness to 50 m (A1), continuous and discontinuous low-permeable layers (B1, C1 and D1) and spatially varying groundwater recharge (E1). Together, the scenarios reflect different hydrogeological configurations encountered in the two provinces. For example, scenario C1 presents a moderately permeable, anisotropic cover layer of 20 m thickness. This scenario is representative for large areas in Noord-Brabant where fluvio-periglacial and eolian deposits make up the upper part of the subsurface.

Subsequently, drains were added to the cross-sections to identify the effects of drainage. Outflow of the model was made possible at 6 extra drains in the columns 31, 46, 61, 76, 91, 96

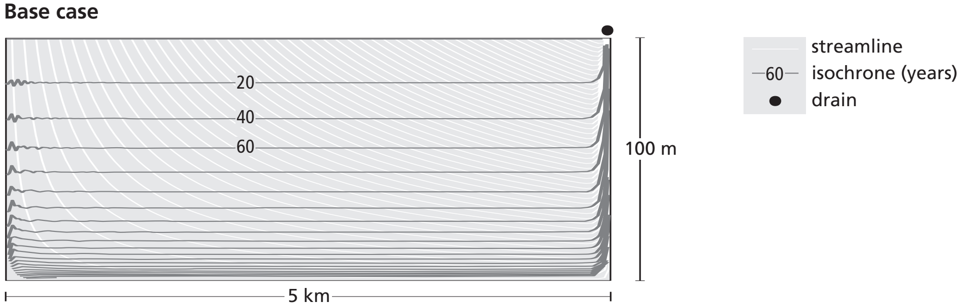


Figure 2.3 - Simulation of groundwater streamlines and isochrones for the base case scenario

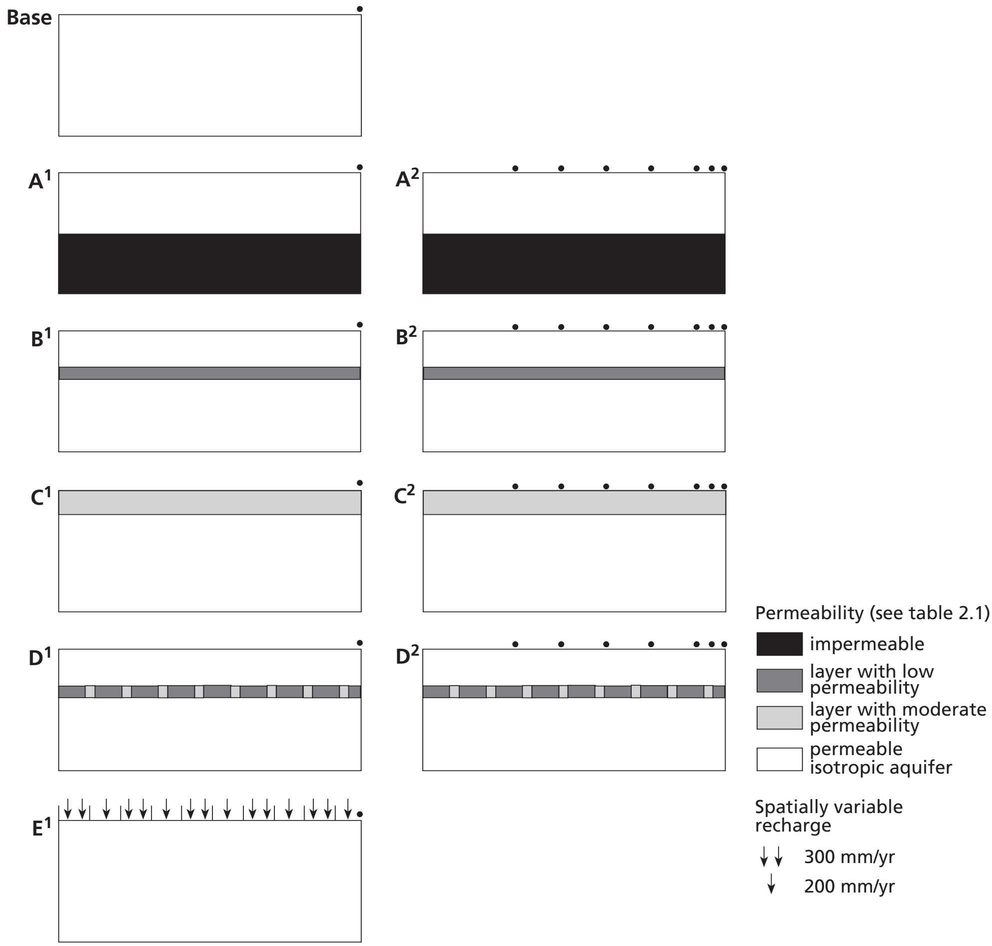


Figure 2.4 - Dimensions of inhomogeneities, variable groundwater recharge and drain positions in the base case and the scenarios A1-E1 and A2-D2. Scenarios: A1: reduced aquifer thickness, B1: continuous low permeability layer between 30 and 40 m depth, C1: moderately permeable cover layer, D1: discontinuous low-permeable layer between 30 and 40 m depth, E1: spatially variable groundwater recharge. Scenarios A2-D2 are similar to A1-D1 but have 6 extra drains.

in the scenarios A2 to D2 (Figure 2.4). Drain conductance was chosen to represent the dimensions of typical Dutch ditches and small watercourses and equalled 500 m^2 day^{-1} for the first six drains and 1000 m^2 day^{-1} for the drain in the upper right cell. The outflow level of the individual drains was defined by the groundwater heads calculated in the homogeneous base case model which are approximated by (Bear 1972):

$$h_x = \frac{N}{2kD}(l^2 - x^2) \quad \text{for } h = 0 \text{ at } x = l \tag{2.2}$$

where h_x = groundwater head at x m from the water divide[m], l = length between water course and water divide [m] and x = distance from water divide [m]. These yield drain levels of 3.13, 2.67, 2.23, 1.54, 0.70 and 0.39 m for the six drains in columns 31 to 96. In this way, the six smaller drains did not contribute to outflow of groundwater from the base case model, because the drain levels equalled the head in the aquifer.

In the scenarios A1-D1 and A2-D2, the overall transmissivity (*kD*) of the model was lowered relative to the base case scenario. As a result, the drains in the models A2-D2 started to contribute to the outflow of water from the model. The drains also reduced the groundwater head increase that was calculated in the scenarios A1-D1. For example, a maximum groundwater head of 11.8 m was simulated in scenario C1, which is unrealistic in the flat Dutch landscape. Using drains in scenario C2, the maximum head reduces to 3.9 m, which is a moderate increase compared with the maximum of 3.4 m in the base case scenario.

Isochrone patterns

Figure 2.5 presents the streamlines and isochrones for the scenarios without drains (A1-E1). Table 2.2 summarises the groundwater age at 10 and 20 m depth for the scenarios without drains.

Reduction of the aquifer thickness to 50 m in scenario A1 resulted in vertical compression of the isochrones; which means older groundwater at a similar depth. The groundwater age in scenario A1 conformed to equation (2.1) for an aquifer of 50 m thickness, except for the vicinity of the main watercourse. A continuous low-permeable layer (between 30 and 40 m depth, scenario B1) yielded vertical compression of the isochrone pattern in the upper aquifer. The 20 years isochrone was found at a shallower level than in the base case scenario. A shallow, anisotropic and moderately permeable cover layer had the opposed effect (Figure 2.5: C1). Here, the 20 years isochrone was found at deeper level compared with the base case and groundwater was younger at a similar depth (Table 2.2).

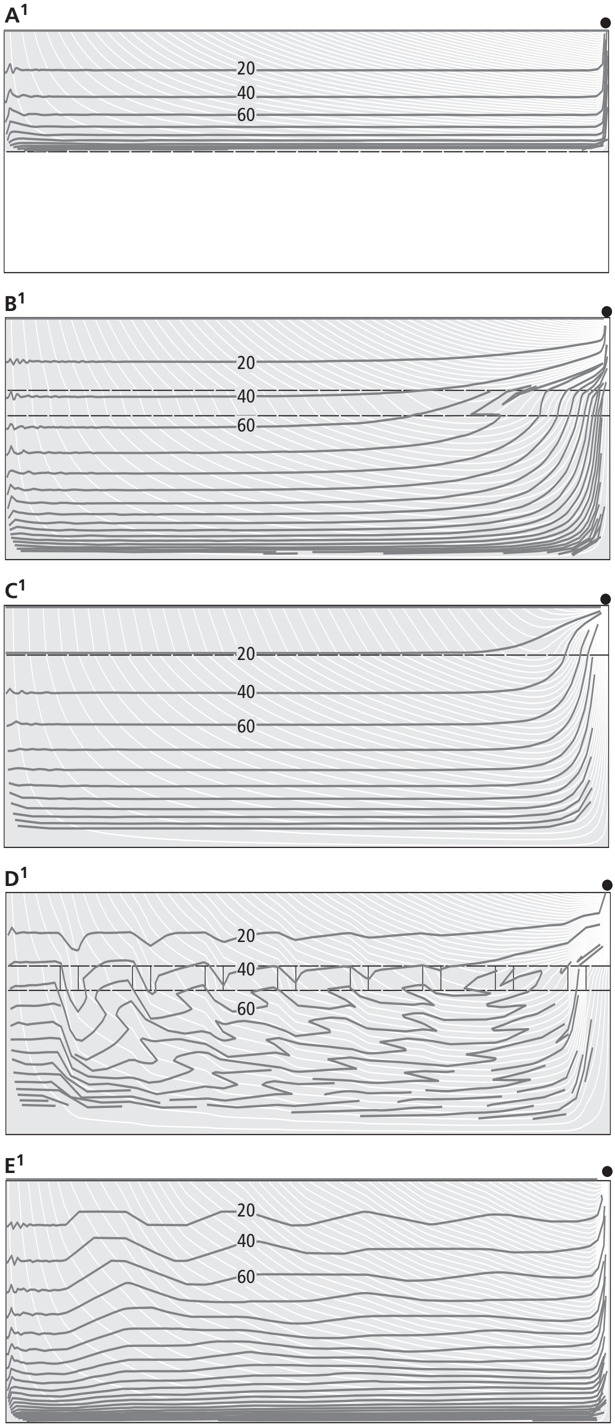
This leads to the following conclusions for laterally continuous inhomogeneities. When a high-permeable layer is found above a low-permeable layer, vertical compression of the isochrones pattern is observed. The vertical flow velocity decreases faster with depth when compared with a homogeneous aquifer. For a low-permeable layer above a high-permeable aquifer, the vertical flow velocity decreases more slowly with depth and a vertical extension of the isochrone pattern is found. For the laterally continuous inhomogeneities, the isochrones remain horizontal, especially for the upper part of the aquifer.

However, deviations from the horizontal position of the isochrones were observed for discontinuous layers and for spatially varying groundwater recharge (scenarios D1 and E1, Figure 2.5).

A discontinuous, low-permeable layer (scenario D1) had a large local impact on the isochrone pattern. This effect was due to preferential vertical flow through the more permeable parts of the confining layer, which lead to the increase of the vertical groundwater velocity at

Table 2.2 - Median groundwater ages and groundwater age variation (indicated by the 25 and 75 percentiles) for the model scenarios without drains at 10 and 20m depth

Scenarios		Groundwater age (yrs) 10m depth		20m depth	
		P50	(P25-P75)	P50	(P25-P75)
	Base case	10.5	(10.5-10.5)	22	(22-22)
A1	Reduced aquifer thickness	11	(11 -11)	26	(26-26)
B1	Continuous low-permeable layer	11	(11 -12)	23	(23-25)
C1	Moderately permeable shallow cover layer	10	(10 -10)	21	(21-21)
D1	Discontinuous low-permeable layer	11	(11 -11)	24	(22-25)
E1	Variable recharge	13	(10 -16)	27	(24-28)



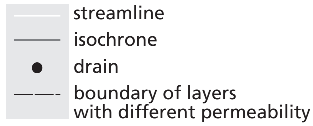


Figure 2.5 - Simulation of groundwater streamlines and isochrones for the scenarios A1-E1. Scenarios: reduced aquifer thickness (A1), continuous low permeability layer between 30 and 40 m depth (B1), moderately permeable cover layer (C1), discontinuous low-permeable layer between 30 and 40 m depth (D1) and spatially variable groundwater recharge (E1)

those locations. This increase was caused by the increase of the horizontal flux above the confining layer, upward from the openings in the layer. The effects of the discontinuities were largest near the groundwater divide at the left part of the model. Thus, irregular discontinuous clay layers causes variations in groundwater age at a specific depth, with locally areas of younger groundwater at a larger depth. However, a strong effect was only observed when a large conductivity contrast is set between the low permeability layer and the aquifer and when the openings in the layer have smaller horizontal extension than the low permeability lenses.

Deviations from the horizontal isochrone pattern might also arise from variations in groundwater recharge rates (scenario E1) which are caused by variations in evapotranspiration for different land use units. In the model, groundwater recharge rates of 200 and 300 mm year^{-1} were used. These numbers resemble ranges of groundwater recharge in forest areas and agricultural areas (Meinardi 1994, Gehrels 1999). Figure 2.5 shows that the effects of variations in groundwater recharge on the isochrone pattern were most clearly found near the ground water divide and that they smoothed out in the direction of the major watercourse.

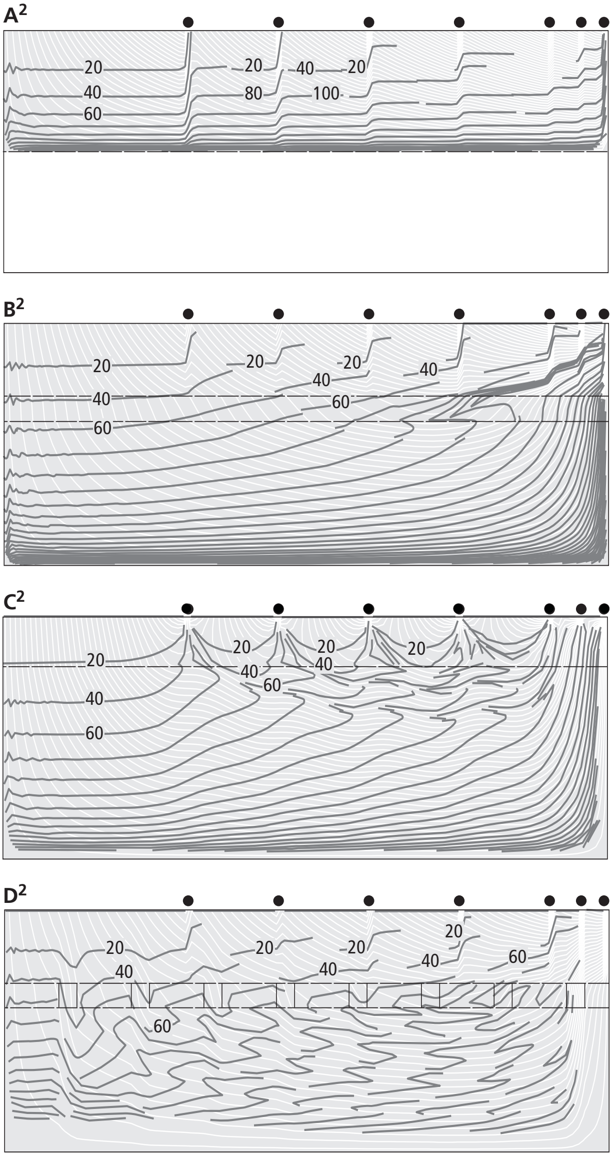
In general, fluxes associated with lateral inhomogeneities and spatially varying groundwater recharge cause variations in groundwater age at a specific depth in the aquifer. However, the variations are not very large, except for the local effects of discontinuous confining layers, which cause local bodies of young groundwater at a relatively large depth relative to a homogeneous case.

Much larger variations in the groundwater age distribution were observed when a drainage network was present. Figure 2.6 shows streamlines and isochrones for scenarios A2-D2. In all scenarios, local flow systems existed, that are superposed at the regional scale system. The drains mainly collected groundwater from the local flow systems, except for the two drains at the right part of the model, which are the regional outlets.

The drains did not influence the isochrone pattern at the left part of the model where drains were absent. This part of the model was considered representative for regional recharge areas without a superficial drainage network. Here, the variations in groundwater residence times were minor and almost similar to the results of the scenarios A1-E1, where no drains were present (Table 2.3, columns 0-30).

Table 2.3 - Median groundwater ages and groundwater age variation (indicated by the 25 and 75 percentiles) for the model scenarios with drains at 10 and 20m depth

Scenario		Groundwater age (yrs)	
		Columns 0-30 (upstream area) P50 (P25-P75)	Columns 30-95 (drained areas) P50 (P25-P75)
10 m depth			
A2	Reduced aquifer thickness	11 (11-11)	21 (11-25)
B2	Continuous low-permeable layer	11 (11-11)	24 (11-30)
C2	Moderately permeable shallow cover layer	10 (10-12)	20 (13-37)
D2	Discontinuous low-permeable layer	11 (10-11)	11 (11-21)
20 m depth			
A2	Reduced aquifer thickness	26 (26-26)	49 (39- 59)
B2	Continuous low-permeable layer	23 (23-23)	37 (32- 55)
C2	Moderately permeable shallow cover layer	21 (21-23)	50 (36-135)
D2	Discontinuous low-permeable layer	24 (22-26)	34 (28- 49)



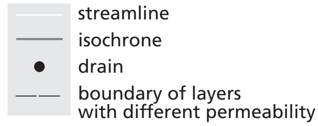


Figure 2.6 - Simulation of groundwater streamlines and isochrones for the scenarios with drains: A2: reduced aquifer thickness), B2 (continuous low-permeable layer), C2 (moderately permeable cover layer and-D2 (discontinuous low-permeable layer)

Strong distortion of the isochrone pattern arose in the drained areas where local flow systems exist. The upward flow near the drains resulted in upconing of older groundwater, especially directly under and downstream of the drain locations. Remnants of the horizontal isochrone pattern were still present within the local flow systems (Figure 2.6). The net effect of the superficial drainage network was that relatively young groundwater was extracted from the system. This resulted in older groundwater at shallower depth. The second effect was an increase in the variation of groundwater age at a specific depth. Both effects are summarised in Table 2.3 (columns 30-95) and are discussed below.

The median groundwater age at 10 m depth for the four scenarios varied between 11 and 24 years (Table 2.3, columns 30-95). The 25 percentile varied less (11-13 years). This represents the groundwater in the middle of the local flow systems, where the horizontal isochrone pattern is still visible. The large variations in the 75 percentile (21-37 years) reflect the groundwater age in the surroundings of the drains.

The effects of drainage were larger at 25 m depth. Median groundwater ages varied between 34 and 50 years for the drained parts of the four scenarios (columns 30-95) versus median groundwater ages of 21-26 years in the parts without drains (columns 0-30).

In conclusion, the simulations show that the isochrone pattern becomes distorted in areas with a drainage network, resulting in older groundwater at shallow depth and greater spatial variation in groundwater age. These effects are large compared with the effects of an inhomogeneous subsoil and spatially varying groundwater recharge.

Extent of young groundwater

The groundwater age distribution has direct consequences for the advective transport of dissolved solutes. In the Netherlands, the concentration of dissolved solutes in recharging groundwater has increased during the last 40 years due to the intensified use of manure and fertilizer in agricultural practices. Details on the increase of agricultural pollution loadings are given in the next chapter. Therefore, it is relevant to investigate the effects of inhomogeneities and the drainage network on the spatial and vertical extent of young groundwater. Here, young groundwater was defined as being less than 40 years old, which for samples taken in 1992 coincides with the boundary between groundwater with or without tritium (see next section for details).

Figure 2.7 shows the extension of young groundwater in the base case scenario and the scenarios A2 to D2, using the 40 years isochrone. In the upstream areas without drains, the young groundwater has infiltrated to about 28 to 36 m depth, depending on the specific scenario. At both 10 and 20 m depth, 95-100% of the groundwater was younger than 40 years in all scenarios (Table 2.4, columns 0-30).

In the drained areas, however, young groundwater has infiltrated to shallower depth. At 10 m depth, predominantly young groundwater was observed; more than 90 percent for scenarios A2, B2 and D2 and 75% for scenario C2. At 20 m depth, the proportion of young groundwater was reduced to 26-67 percent in the four scenarios (Table 2.4, columns 30-95). Hence, the probability of finding young groundwater at 20 m depth was much smaller for the drained areas than for the recharge areas. This shows that the groundwater age in the drained areas depends strongly on the

Table 2.4 - Proportion of young groundwater (less than 40 years old) in the model scenarios

Scenario		Proportion of groundwater younger than 40 years	
		Columns 0-30 (%)	Columns 30-95 (%)
10 m depth			
A2	Reduced aquifer thickness	100	97
B2	Continuous low-permeable layer	100	91
C2	Moderately permeable shallow cover layer	100	75
D2	Discontinuous low-permeable layer	100	94
20 m depth			
A2	Reduced aquifer thickness	100	26
B2	Continuous low-permeable layer	100	61
C2	Moderately permeable shallow cover layer	97	29
D2	Discontinuous low-permeable layer	100	67

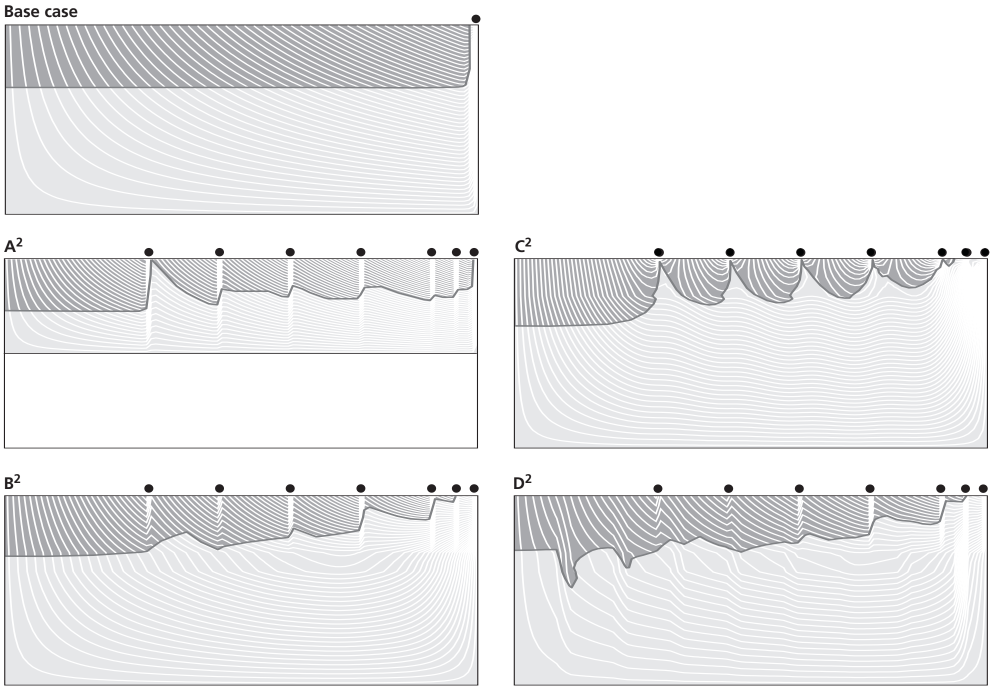


Figure 2.7 - Distribution of young groundwater (less than 40 years old) in the base case scenario and the scenarios A2 (reduced aquifer thickness), B2 (continuous low-permeable layer), C2 (moderately permeable cover layer) and D2 (discontinuous low-permeable layer)

hydrogeological situation and the position relative to the drainage network. In general, the spatial extent of young groundwater decreases going from recharge to discharge areas.

In summary, the presence of a drainage network tends to increase the spatial variation in groundwater age for a specific depth. This is caused by the presence of local flow systems and local upward flow. Accordingly, the proportion of young groundwater at a specific depth decreases when compared with areas that lack a superficial drainage network. The largest effects are found for shallow aquifers and areas with shallow layers of low permeability.

2.3 Evaluation of the groundwater age distribution in two regional networks

Methods

Tritium measurements were used to assess the distribution of old and young groundwater for two regional monitoring networks in the Netherlands. Proportions of young groundwater were determined for three geohydrological situations: recharge areas, discharge areas and intermediate areas.

Mapping recharge, discharge and intermediate areas

The geohydrological situation was mapped to identify recharge, discharge and intermediate areas during the design stage of the regional networks of Noord-Brabant and Drenthe (see below). This subdivision was subsequently used to determine the vulnerability of the areas for contamination of deeper groundwater resources (see Chapter 3).

The geohydrological subdivision in recharge, discharge and intermediate areas was based on the 1:50,000 groundwater map of the Netherlands, the 1:50,000 soil map of the Netherlands (STIBOKA) and 1:50,000 topographical maps of the Netherlands, where the latter indicate the position of watercourses. In general, the distinction between recharge, intermediate and discharge areas coincides with geomorphological and hydrogeological factors. The absence of a drainage network, the occurrence of deep groundwater levels, relatively permeable soils and a high topographical position characterized the regional *recharge areas*. Regional *discharge areas* were classified by their low elevation, the shallow groundwater depth and the dense drainage network. Local field studies and existing groundwater modelling studies were used to determine the position of the regional discharge areas (Uil &Vlot 1990, Stuurman et al. 1990).

The *intermediate areas* have an intermediate topographical position between the regional recharge and discharge areas and included all areas that were not mapped as recharge or discharge areas. These areas were distinguished from the regional recharge areas on the basis of the existence of a superficial drainage network of ditches, drains and small watercourses. For example, the distinction between regional recharge and intermediate areas in Drenthe was made using a digital map of the main watercourses in the province (Figure 2.2c). Areas within 250 m of a watercourse were classified as intermediate, using GIS buffer techniques (Broers 1996). Large parts of the *intermediate areas* were only brought under cultivation during the 20th century after the artificial creation of the drainage network. The aim of this artificial network was to create sufficiently dry circumstances for agriculture. Today, a dense drainage network is present in relatively flat and low areas and in areas with layers with low permeability within the first 30 m. In both provinces, the intermediate areas have a large spatial extension relative to the regional recharge and discharge areas (Figure 2.8, 36 and 44% of Drenthe and Noord-Brabant, respectively).

During the design stage of the monitoring networks, the geohydrological map was combined with a land use map and a simplified soil map using GIS overlay techniques. The spatial units on the resulting overlay were called *homogeneous areas* (for details, see Chapter 3). Monitoring wells were chosen within the largest of the homogeneous areas, using stratified sampling. Figure 2.8 shows the position of the monitoring wells relative to the distinguished recharge, discharge and intermediate areas for the two provinces.

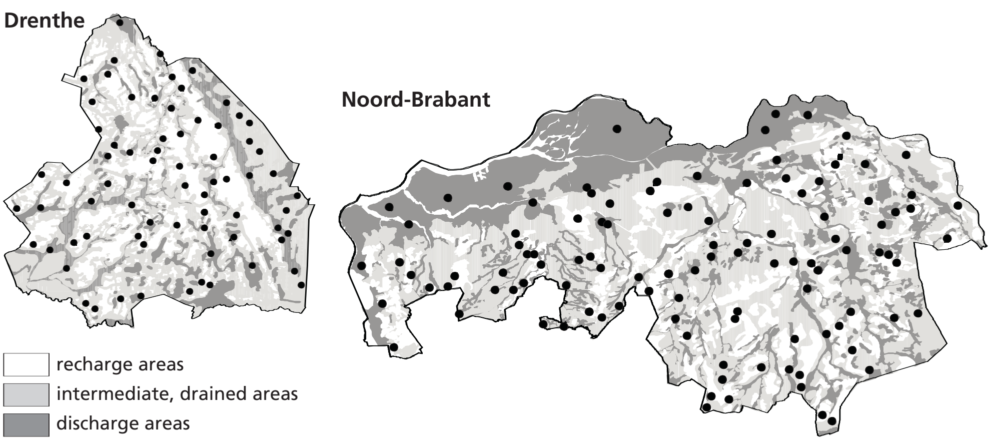


Figure 2.8 - Recharge, intermediate and discharge areas in the provinces of Drenthe and Noord-Brabant and the corresponding monitoring wells

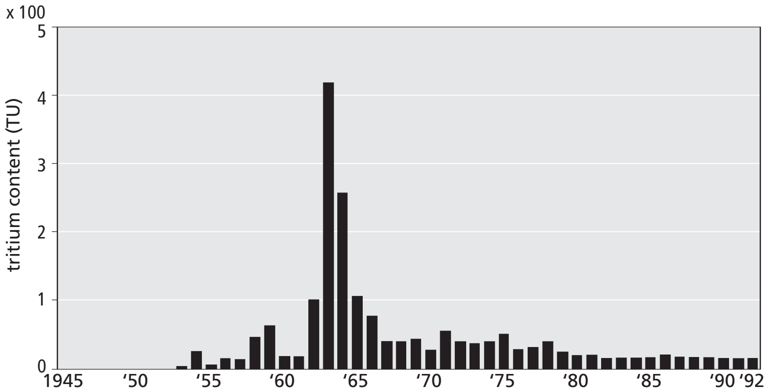


Figure 2.9 - Tritium-input in recharging groundwater in Noord-Brabant (corrected for radioactive decay up to 1992). Data from Meinardi (1994)

Determination of proportion of post-1950 groundwater

Tritium is used as an environmental tracer for the occurrence of young groundwater (Fritz & Fontes 1980, Robertson & Cherry 1989). Large amounts of tritium were introduced into the atmosphere during the 1950-1965 nuclear tests and a tritium peak in the precipitation of 1963 was observed globally. Figure 2.9 shows the average tritium peak for precipitation in the province of Noord-Brabant based on data from Meinardi (1994). He used the peak position to determine the amount of groundwater recharge for many local situations in the Netherlands using mini-screen observation wells and monitoring wells of the national monitoring network.

The Noord-Brabant and Drenthe regional monitoring networks consist of 122 and 79 monitoring wells, respectively (Figure 2.8). The observation wells of the two monitoring networks were sampled for tritium at two depths: about 9 m and 24 m below surface. The objective of the sampling was to acquire information on the age of the groundwater in order to improve the interpretation of the groundwater quality data. Here, the screens between 5 and 15 m depth were denoted as *shallow screens*, the screens between 15 and 30 m depth as *deep screens*. Average screen depths and standard deviations are listed in Table 2.5. The length of the monitoring screens is 2 m. The part of the wells that belong to the national monitoring network were sampled in 1983 (Meinardi 1994). The part of the wells that was installed by the provincial authorities were sampled in 1992 (Broers & Griffioen 1992, Broers 1993, Broers 1996). Detection limits were 5 Tritium Units (1 TU = 1 ^{3}H atom per 10^{18} atoms of H) for the 1983 data and 0.6-1.6 TU for the 1992 data.

In this study, tritium concentrations were only used as an indication of the presence of young groundwater, distinguishing between groundwater with and without tritium. Hallberg

Table 2.5 - Average screen depths for the 5-15 m (shallow) and 15-30 m (deep) depth intervals in the regional networks of Noord-Brabant and Drenthe

	Average screen depths and standard deviation (m)	
	Noord-Brabant	**Drenthe**
Shallow screens	8.7 ± 2.1	9.3 ± 1.2
Deep screens	23.5 ± 2.1	24.3 ± 2.0

Table 2.6 - Water table classes used on the Dutch 1:50,000 soil map (van der Sluijs & de Gruijter 1985)

Code	Average highest groundwater level (cm-surface level)	Average lowest groundwater level (cm-surface level)
I	(<20)	<50
II	(<40)	50- 80
III	<40	80-120
IV	>40	80-120
V	<40	>120
VI	40-80	>120
VII	>80	(>160)

& Keeney (1993) used a similar method to distinguish old and 'modern' water, in order to detect and explain nitrate contamination patterns in Devonian carbonate aquifers in Iowa. Here, groundwater was classified as *post-1950* if the measured concentrations exceed 5 TU (1983) or 2 TU (1992). Using the 1992 and 1983 tritium measurements, *post-1950* groundwater coincides with *young groundwater* of less than 40 years old in the model simulations. This enables comparison of the modelling results with the tritium ages.

The proportion of post-1950 groundwater was assessed for the Noord-Brabant and Drenthe monitoring wells in the recharge, intermediate and discharge areas. The proportion was defined as the amount of wells with post-1950 water divided by the total amount of wells. A 95% confidence interval (α = 0.05, two-sided) for the estimated proportion was calculated using the methods of Blyth & Still (in: Gilbert 1987, see appendix III.5).

Additionally, to test the specific relations between local drainage density and *water table class* and the age of the groundwater, the local *water table class* and the *drain length per square kilometre* were assessed from soil maps and topographical maps for each well of the Drenthe network. In the Netherlands, the *water table class* was classified in 7 main classes and mapped at 1:50,000 scale (Table 2.6, Van der Sluijs & de Gruijter 1985). The *drain length per square kilometre (DLSK)* was determined by measuring the total length of watercourses within a square kilometre around each observation well at the 1:50,000 topographical map. Proportions of post-1950 groundwater were estimated for three classes of *water table class* and three classes of *drain length per square kilometre* for the province of Drenthe.

Results

Young groundwater in recharge, intermediate and discharge areas

Figure 2.10 shows the proportion of post-1950 groundwater for the recharge, intermediate and discharge areas as determined for the provinces of Noord-Brabant and Drenthe. The proportions were determined separately for the shallow and deep screens.

Recharge areas exhibited post-1950 water in all shallow screens for both networks (confidence interval 86-100%). The deep screens contained young groundwater in 96% of the cases (Drenthe) or 70% (Noord-Brabant). In general, the tritium measurements in the recharge areas confirmed the prior expectations from the model scenarios; at about 24 m depth predominantly young groundwater was present.

Discharge areas had small proportions of young groundwater in the shallow screens (18 and 30% for Drenthe and Noord-Brabant, respectively) and the deep screens (0 and 17%, respectively). This confirmed the expectations regarding the upward seepage of pre-1950 groundwater in discharge areas. The larger proportion of tritiated water in the discharge areas

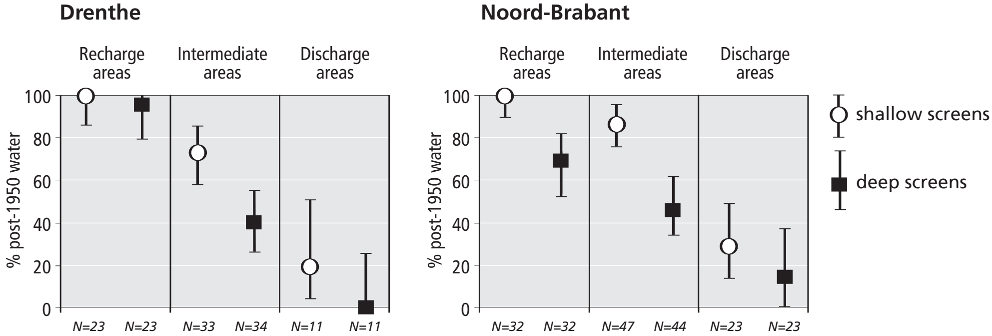


Figure 2.10 - Proportion of young, post-1950 groundwater in recharge, intermediate and discharge areas in the regional networks of Drenthe and Noord-Brabant

of Noord-Brabant was due to wells in the Northern fluvial clay area, and was probably caused by bank infiltration of the rivers Meuse and Rhine, which have higher average water levels than the surrounding clay areas.

The intermediate, drained areas comprise an estimated proportion of 73 and 87% young groundwater for the shallow screens, and 38 and 47% for the deep screens. The overall pattern for intermediate areas showed that shallow groundwater is predominantly post-1950 and deep groundwater is often pre-1950. The presence of 38-47% young groundwater and a complementary 53-62% of old groundwater in the deep screens points to large variation of the groundwater age in those areas. Thus, the expected effects of the drainage network on the groundwater age distribution in drained areas were confirmed by the tritium data.

The relationships were most clear for the Drenthe regional network. The larger variations of deep groundwater in the recharge areas of Noord-Brabant were probably due to the more complicated hydrogeological structure of Noord-Brabant with shallow low-permeable formations in parts of the province. The general pattern, however, is comparable for both provincial networks.

For both the shallow screens and the deep screens in Drenthe, the 95 percent confidence intervals on the estimated proportion of post-1950 water did not overlap (Figure 2.10). Thus, a good discrimination of the proportions of young and old groundwater was achieved when using the geohydrological subdivision. The strong relationship was confirmed by a non-parametrical test which indicates that the proportion of post-1950 groundwater increases significantly from discharge to recharge areas (Kendall τ_b, $p<0.001$, Helsel & Hirsch 1992).

Relations with drainage density and water table class

The proportion of post-1950 groundwater was also determined for three classes of *water table class* and for three classes of *drain length per square kilometre (DLSK)* in the province of Drenthe that were derived from the 1: 50,000 topographical and soil maps. (Figure 2.11). These two types of maps are readily available in the Netherlands, whereas the geohydrological map was custom-made using a large number of maps and required the acquisition of field data and the use of regional groundwater modelling studies.

For the factor *drain length per square kilometre (DLSK)*, a significant relationship was found with the proportion of young groundwater in both shallow and deep screens ($p<0.001$) using three classes (Figure 2.12b). The class *< 1500 m km*$^{-2}$ showed more than 80% young groundwater at both depths. This is in accordance with the model simulations: a higher probability of young groundwater was expected at larger distances from watercourses. However, a large

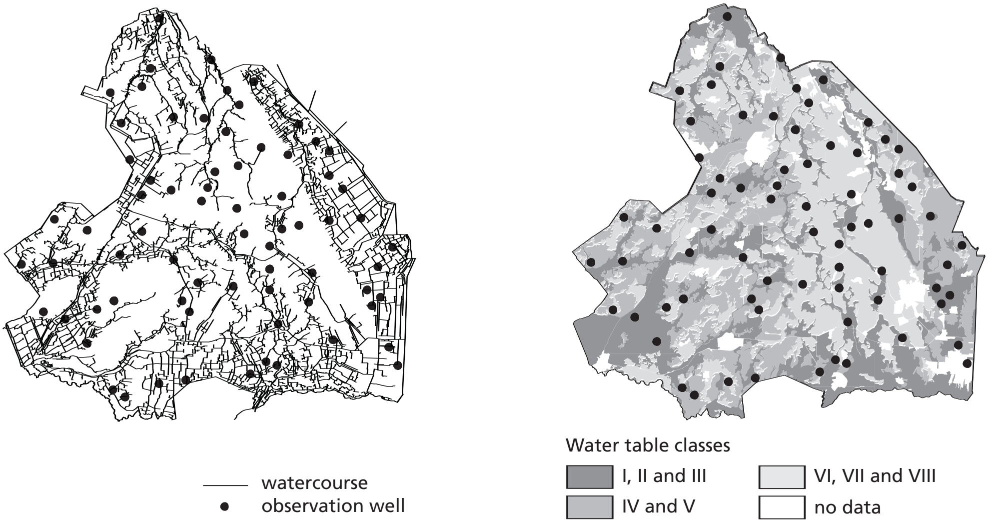


Figure 2.11 - The drainage network of Drenthe and the distinguished water table classes

overlap in the estimated proportion of post-1950 water existed for the classes *1500-3000 m km*$^{-2}$ and > *3000 m km*$^{-2}$. This suggests that the regional discharge areas were not effectively distinguished from the intermediate, drained areas when using *DLSK* as the sole indicator.

In an earlier stage, the *water table class* maps were used successfully for the delineation of discharge and recharge areas (Engelen 1984, Stuurman et al. 1990). Three classes of water table

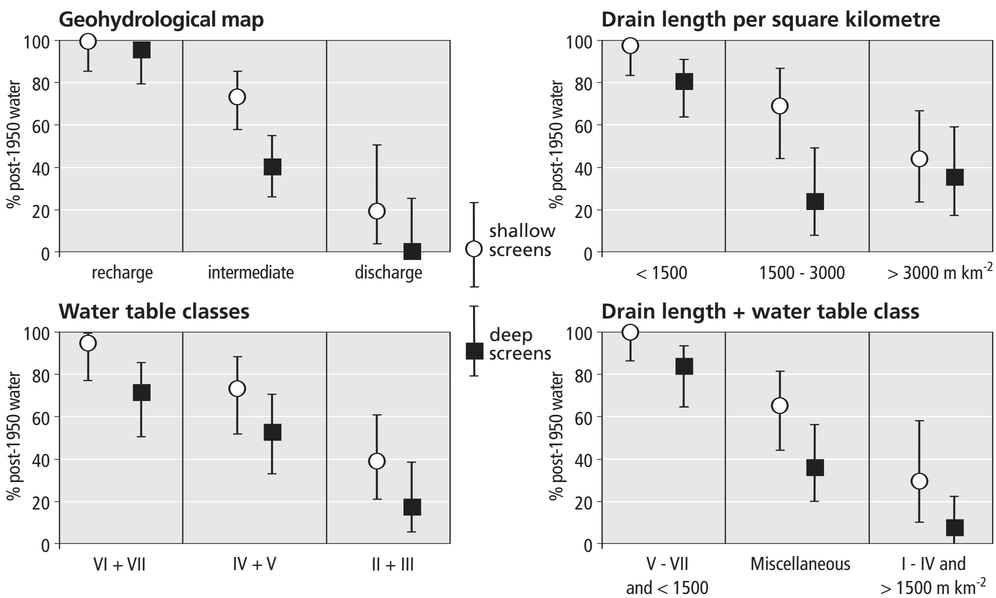


Figure 2.12 - Relations between the proportion of young groundwater and the geohydrological classification, drain length per square kilometre, the water table class and a combination of drain length and water table class for wells in the regional network of Drenthe

classes were distinguished: *II+III*, *IV+V* and *VI+VII* (see Table 2.6). Normally, the shallow water table classes II and III are found in brook valleys, wetlands or areas with peat soils, whereas regimes VI and VII are found at the higher ridges in the landscape. A significant relationship with the proportion of young groundwater was also found for the three water table classes ($p<0.001$) for both the shallow and deep screens. Large proportions of young groundwater were found at both depths in the class *VI+VII* and the shallow screens of class *IV+V* (Figure 2.12). A small proportion of young groundwater was found for the deep screens of class *II+III*. The confidence intervals showed larger overlap than the intervals for the recharge, intermediate and discharge areas of the geohydrological map. This indicates that the groundwater age distribution was not distinguished unambiguously using solely the three water table classes.

In conclusion, the separate use of the two factors *DLSK* and *water table class* was less successful in predicting the groundwater age distribution than the geohydrological subdivision in recharge, intermediate and discharge areas. However, a better separation was obtained when the two types of information were combined into three new classes, using a threshold value of 1500 m km^{-2} for DLSK and three classes for water table class (Figure 2.12). The overlap between two neighbour classes decreased and the significance of the relationships increased. The combination of *DLSK < 1500 m km*$^{-2}$ plus *water table classes V*, *VI* or *VII* separated the areas with predominantly young groundwater at shallow and deep level. The combination of *DLSK > 1500 m km*$^{-2}$ plus *water table classes II*, *III* and *IV* separated the areas with less than 25% young groundwater at about 24 m depth.

The Drenthe example shows that a combination of water table class and drainage density yields a first prediction of areas with young groundwater at the established monitoring depths.

2.4 Implications for groundwater quality monitoring

In the Netherlands, threats to groundwater quality have increased because agricultural practices have become more intensive after the Second World War. The introduction of new maize varieties in the early 70's and the simultaneous fast increase of intensive livestock farming caused a large increase of manure loads (see also Chapter 3). Thus, diffuse contamination of young groundwater is likely in agricultural areas.

The effects of the groundwater age distribution on the proportion of contaminated groundwater were illustrated using the 1995 monitoring results of the Drenthe network. Based on local and regional groundwater studies in the Netherlands, groundwater pollution indices were defined to distinguish groundwater that shows signs of anthropogenic pollution. Two indices were used: a general pollution index (POLIN) that was proposed by Stuyfzand (1993) and the MANURE index that indicates agricultural pollution (appendix I).

Table 2.7 compares the proportions of young groundwater and the proportions of groundwater with indications of pollution for agricultural areas in Drenthe. The table shows that 60-100% of young groundwater in agricultural areas showed signs of pollution. Especially for deep groundwater, a striking difference between the recharge and the intermediate areas was observed for both pollution indices, which coincides with the proportion of young groundwater. Although the intermediate areas have an overall lower risk for groundwater pollution compared with recharge areas; still a considerable proportion of young groundwater with signs of anthropogenic pollution was present due to the groundwater age variations that characterize them.

The simulated isochrone patterns were used to evaluate the advective transport of a block front

Table 2.7 - Proportion of post-1950 groundwater and proportions of two indicators of anthropogenic pollution in the Drenthe regional monitoring network

	Number of observations	Post-1950 water %	Indications for agricultural pollution (MANURE) %	General indications for pollution (POLIN) %
Shallow screens (5-15 m depth)				
agriculture/recharge	15	100	100	85
agriculture/intermediate	26	75	51	34
discharge areas	11	18	18	0
Deep screens (15-30 m depth)				
agriculture/recharge	16	100	65	83
agriculture/intermediate	26	39	20	23
discharge areas	11	0	9*	0

* due to one well with brackish water (Cl>50)

contamination input. Figure 2.13 shows the advective propagation of an imaginary pollution block front input of 20 year for scenarios with and without drainage (scenarios C1 and C2). The block front pollution was introduced homogeneously in the whole model. In the scenario without drains, the pollution front moved down vertically. The volume of contaminated groundwater increased during the first 20 years and decreased gradually because contaminated groundwater was removed at the major drain in the upper right quarter of the model. After 100 years a considerable amount of contaminated groundwater is still present at depth in the aquifer.

In the scenario with drains (C2) the contaminated groundwater is also removed from the local flow systems. This resulted in shallower contamination of the aquifer in the drained areas. Because of the shorter transit times in the local flow systems, the upper part of the aquifer became gradually decontaminated after 60 to 100 years in the drained areas. The pollution that originated from the regional recharge areas, however, was still in transit in deeper parts of the aquifer.

The model simulations and the proportions of young, contaminated groundwater in the Drenthe network both showed that the potential contamination of deep groundwater resources is concentrated in the regional recharge areas. Ultimately, these contamination fronts proceed laterally under the local flow systems, but these are long-term effects that are less relevant for the time scales of groundwater quality monitoring.

However, the shallow groundwater in the intermediate, drained areas is also vulnerable for diffuse contamination and large spatial variations in groundwater age and risks for contamination of deeper groundwater were simulated and observed. In part of the intermediate areas, the groundwater contamination patterns resemble those in recharge areas, whereas in the other part old, uncontaminated groundwater dominates. Given the large spatial variability of the groundwater age in the intermediate, drained areas, the use of spatially averaged groundwater recharge fluxes as proposed by Meinardi (1994) is not appropriate for predicting the depth of groundwater contamination and the proportion of young and potentially contaminated groundwater. Spatially averaged values of groundwater recharge rates do not explain the observed variations in the proportions of young and old groundwater in the drained parts of provinces of Drenthe and Noord-Brabant.

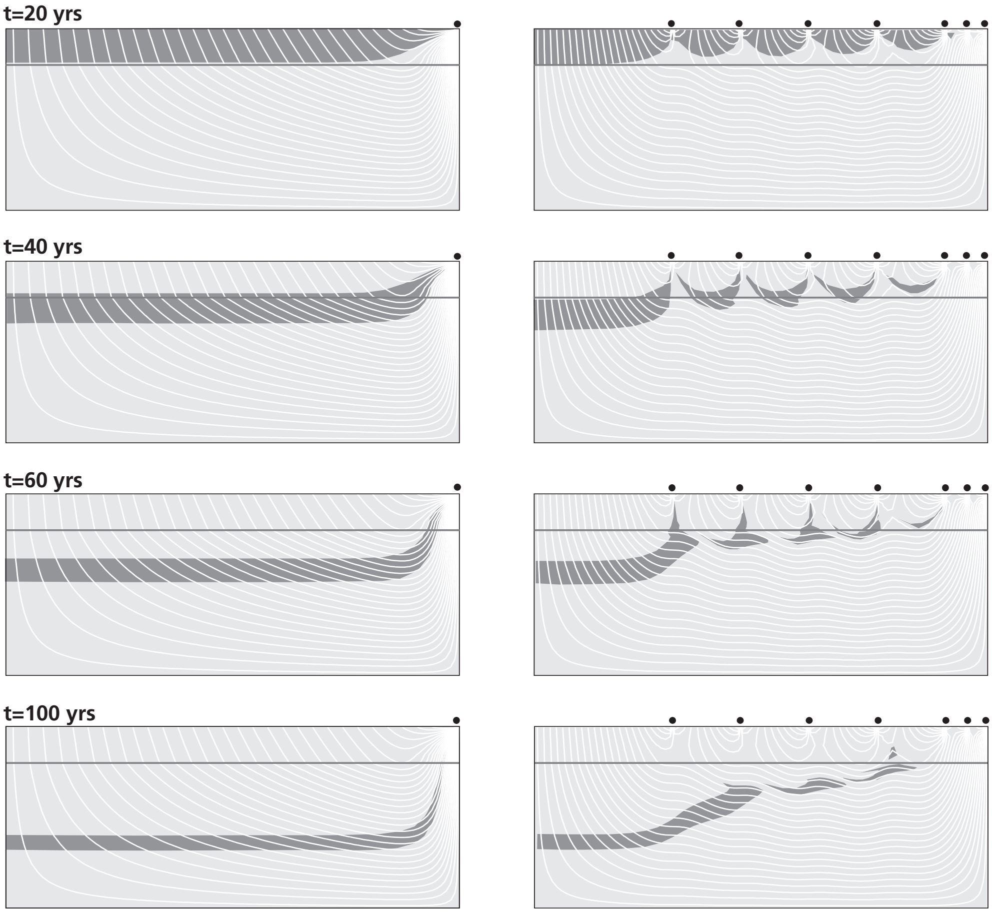


Figure 2.13 - Propagation of a 20 year contaminant block front in the model scenarios C1 (moderately permeable cover layer without drains) and C2 (idem, with drains) after 20, 40, 60 and 100 years

A groundwater monitoring program which aims to quantify the recent human impact on groundwater quality must account for the differences in groundwater ages and contamination risks for deep groundwater between the recharge areas and the intermediate, drained areas. It is advisable to differentiate sample size, monitoring depth and monitoring frequency to account for these differences. A proposition for a risk-based concept and area-specific monitoring objectives is presented in Chapters 3 and 4. Given the large spatial variability of groundwater ages in the drained, intermediate areas at the relevant monitoring depths, large variations in the concentration of dissolved solutes are anticipated. Contaminated and uncontaminated groundwater will both be present at similar monitoring depth, but the locations of contamination are difficult to predict without modelling the local groundwater flow patterns. As a result, a relatively large sample size will be necessary to acquire precise statistics of contaminant concentrations for those areas with high variability (Chapter 4). This is especially important in the Netherlands, where drained, intermediate areas have large spatial extension, relative to the recharge and discharge areas.

2.5 Conclusions

At the regional scale, large spatial variations in groundwater age exist in flat areas with a dense drainage network. Superficial drainage limits the depth of recently infiltrated young groundwater and increases the spatial variability of groundwater ages at the typical Dutch monitoring depths of 10 and 25 m. The large impact of superficial drainage on the groundwater age distribution is caused by local flow systems that remove a substantial part of the young groundwater and prevent the feeding of deeper groundwater resources. The effects of superficial drainage on groundwater age are large compared with the effects of inhomogeneous aquifers or spatially varying groundwater recharge.

Tritium measurements in two regional monitoring networks showed that the proportion of young groundwater decreases from recharge areas, via intermediate, drained areas to discharge areas. Recharge areas showed 70-100% post-1950 groundwater at the typical monitoring depths of 10 and 25 m. The large variability in the drained areas is indicated by the large proportions of post-1950 (38-47%) and pre-1950 groundwater (53-62%) at 25 m depth. The proportions of young groundwater could be predicted using maps of the drainage network and water table classes in the province of Drenthe.

The variations in groundwater age cause large variations in the concentration of contaminants that were introduced in the last decades. For example, 60 to 100% of young groundwater in agricultural areas in the province of Drenthe, showed signs of anthropogenic pollution. The proportions of contaminated groundwater in recharge and intermediate, drained areas agree well with the proportion of young, post-1950 groundwater. Therefore, groundwater monitoring programs that aim at quantifying recent human impacts on groundwater quality must account for the differences in groundwater age and contamination risk for deep groundwater between the recharge areas and the intermediate, drained areas. Given the large spatial variability of groundwater ages in the drained, intermediate areas at the relevant monitoring depths, a relatively large sample size will be necessary to acquire precise statistics of contaminant concentration for those areas.

3 Regional monitoring of agricultural pollution and acidification of groundwater in two Dutch provinces

1 - Network design and data analysis

3.1 Introduction

Rationale and objectives

Contamination of groundwater resources by diffuse sources is a serious problem in the Netherlands. The contamination by agricultural sources has especially increased during the last 40 years due to the increase of intensive livestock farming and the use of pesticides. A national monitoring network for groundwater quality was established between 1979 and 1992 to assess and quantify the human impact on groundwater quality in time and space (Van Duijvenbooden et al. 1985, 1993). Since 1989, regional monitoring networks have also been installed as an addition to the national network. Results are used for environmental policy reports on national and provincial scale (for example RIVM 2001, Province of Noord-Brabant 2000). The design strategy and monitoring objectives of the regional monitoring networks differ from the national monitoring network.

This chapter presents the design and data analysis strategy of two provincial networks in the Netherlands. The first network is in the province of Noord-Brabant, which is known for large inputs of manure due to intensive livestock farming (Menke 1992). The other network is in the province of Drenthe, which is less affected by agricultural practices. The chapter describes the network design and the monitoring objectives and compares the monitoring results from 1995-1998.

The study aims at integrating information on groundwater age and hydrogeochemical processes in the design and data analysis of regional monitoring networks and to investigate if such an approach yields extra value in the identification of groundwater quality patterns. Two main hypotheses are examined: (1) geohydrological stratification helps to reduce the variation in the data, and (2) extra indicators based on geochemical knowledge provide a better identification and understanding of contamination patterns than the sole analysis of concentrations of targeted contaminants.

Monitoring network design: general risk concept

The first stage in monitoring network design covers the information analysis and includes the evaluation of properties of the system studied and the definition of objectives for conducting monitoring (Figure 3.1, see also chapter 1).

The design of the Dutch national and regional networks was based on a preceding assessment of risks for groundwater contamination of deeper groundwater resources. In general, such a risk assessment includes factors of pollution loading and aquifer vulnerability

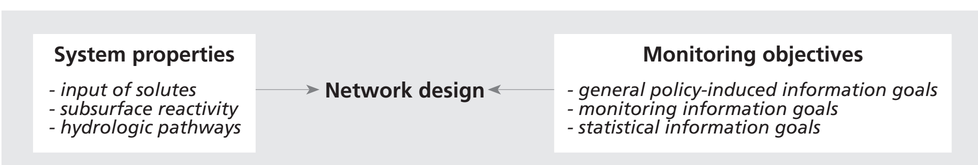


Figure 3.1 - The design of a monitoring network is tuned to the properties of the system studied and the objectives for conducting monitoring

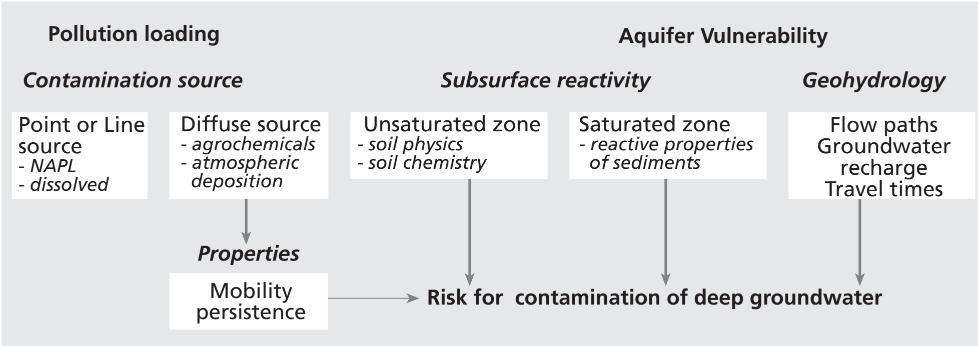


Figure 3.2 - General risk concept for the design of regional monitoring networks

that reflect the physical and chemical properties of the studied system ((Foster 1987, Figure 3.2). For diffuse contaminants, the pollution loading is a function of land use, and sources are atmospheric deposition and agricultural inputs such as manure, fertiliser and pesticides. Vulnerability reflects the subsurface sensitivity to the leaching of contaminants to deeper groundwater. Vulnerability is a function of the geohydrological situation and the physical and chemical properties of sediments in the saturated and unsaturated zones (Figure 3.2). Often, the risks for groundwater contamination are assessed using geographical information systems (GIS) to map the vulnerability and pollution loadings in the study area. The resulting risk maps are then used as a basis for monitoring design and data analysis. For example, the US-EPA's DRASTIC method uses complex weighting of several geographical factors that determine risks for deep groundwater contamination (Aller et al. 1987). Spatial information about the reactive properties of the saturated zone is generally not available, and simplifications to the risk factors of Figure 3.2 are normally made when designing monitoring programs.

3.2 The national groundwater quality monitoring network

The original objectives of the Dutch national network were: (1) to investigate the quality of the groundwater in the upper aquifer in relation to land use, soil type and geohydrological conditions, (2) to determine the extent of human influence on groundwater quality, (3) to identify the changes of groundwater quality over time and (4) to provide data for good management of groundwater resources (Van Duijvenbooden et al. 1985, Van Duijvenbooden 1993).

The national network was designed before GIS methods became widely used. The locations of the newly installed observation wells were chosen using readily available topographical maps and soil maps. The 380 wells were evenly distributed over the Netherlands (Van Duijvenbooden 1993). Part of the monitoring wells was installed in areas of specific interests, such as groundwater protection areas of drinking water well fields and areas with riverbed infiltration. The wells have screens at three depths, about 9, 15 and 24 m depth, and screen lengths of 2 m. The wells were all newly drilled using a cable tool drilling system (Van Duijvenbooden 1985, Snelting 1990).

After digital soil maps and remote sensing derived land use maps became available, the wells were classified into 16 land-use/soil type combinations; the so called land-use/soil-type strata (Reijnders et al. 1988). In fact, the land-use/soil-type strata represent the risk classes for groundwater contamination for the national monitoring network (Figure 3.3).

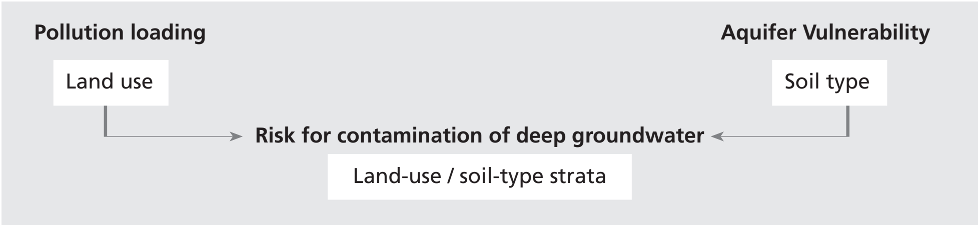


Figure 3.3 - Risk concept used for monitoring design of the national monitoring network.

The geohydrological situation has neither been mapped nor used for data analysis (Snelting et al. 1990, Reijnders et al. 1998). At the national scale, this data analysis choice might be useful because the combination of land use and soil types also defines the major geohydrological situations in the Netherlands. The land-use/soil-type strata reflect different geomorphological positions in the Dutch landscape. For instance, the geohydrological situation in the grassland-fluvial clay areas is very different from the situation in grassland in the higher sandy areas of the Netherlands. However, at the regional scale the geohydrological situation shows large variation, especially in the vulnerable sandy areas.

Nevertheless, the national monitoring network database contains a geohydrological classification of the wells, which is based on the head difference that is measured between the two monitoring screens at about 9 and 24 m depth. Wells with a positive head difference of > 3 cm are classified as *recharge*, with a negative head difference as *discharge*, and with less than plus or minus 3 cm as *stagnant*. This classification should be considered to be *a posteriori*, because it makes use of measured data, instead of using independent spatial information like the land use and soil type maps. Moreover, the classification is erroneous, because a head difference only exists if low-permeable layers are present between the two well screens. Low-permeable layers often hamper the deep infiltration of contaminants. Likewise, in highly permeable recharge areas a negligible head difference is to be expected, because of the small vertical hydraulic resistance of the permeable formations. Therefore, the absence of a head difference is a strong indicator for recharge conditions, and not for stagnant conditions.

Various authors have developed different approaches for the data analysis of the Dutch network. Section 1.2 (Chapter 1) provides a brief overview of these approaches. Pebesma (1996) and Reijnders et al. (1998) presented methodologies for characterising and mapping targeted contaminants, using the land-use/soil-type stratification. However, they neither interrelated the patterns of individual chemical components using geochemical knowledge, nor considered the hydrological position of the monitoring wells. The data analysis results showed large variation and skewness of groundwater concentrations within the land-use/soil-type strata (Reijnders et al. 1998) or km^2 blocks (Pebesma 1996) especially in the vulnerable sandy areas of the Netherlands. Pebesma suggested that his approach would benefit from the use of hydrological information in defining the strata, thus reducing part of the observed variation in the land-use/soil-type strata.

Frapporti (1994) demonstrated the importance of geochemical reactions in the analysis of the national monitoring network data. He argued that these reactions overrule the effects of land use and soil type, but made no attempt to translate the results to proportions of contaminated groundwater in areas or in land-use/soil-type strata. An improvement of the data analysis is anticipated when hydrogeochemical knowledge is incorporated in the data analysis of the strata, combining information of several related chemical components to assess hydrogeochemical processes and conditions that determine the fate and distribution of the targeted contaminants.

3.3 Design of the monitoring networks of Noord-Brabant and Drenthe

In this section the design of the two regional networks of Noord-Brabant and Drenthe is discussed. A brief overview of the geological and hydrogeological situation in the two provinces was given in Chapter 2. The network design was based on a land-use/soil-type/geohydrology stratification that is obtained using independent spatial information. The present study was restricted to the Pleistocene sandy areas of the two provinces, since these are most vulnerable for groundwater contamination.

Information analysis
The general, policy-induced monitoring information goals of the provincial networks are similar to the national network. The main objectives are to assess the human impact on groundwater quality and to identify groundwater quality changes in time. The provincial networks additionally aim at the monitoring of the effects of groundwater protection policy. The networks should yield information at the more detailed regional scale to be useful for regional water management and soil protection strategies. In Noord-Brabant, a preliminary groundwater quality survey of existing observation wells revealed regional groundwater quality patterns (Stuurman et al. 1990). Groundwater quality in the regional discharge areas was determined by the seepage of deep groundwater that was affected by carbonate dissolution at depth and sulphate reduction. Groundwater contamination under topographically high agricultural areas was detected up to 40 m depth. Therefore, the new network was focused on the risks for contamination of deep groundwater. The risks were assessed using overlays of maps on land use, soil type and geohydrological situation (Figure 3.4). Spatial information about the reactive properties of the saturated zone was not available at the time of the design of the regional networks. Accordingly, subsurface reactivity could not be used for stratification.

The necessary soil maps were readily available, and were simplified to distinguish sand, clay and peat soils (de Vries & Dennekom 1992). In this chapter only results for the more vulnerable sandy soils will be presented. Differences in sandy soils were not further distinguished.

The land use was classified using a 1986 LANDSAT image (Peeters 1990, Uil & Vlot, 1992) and yielded a 500 x 500 m grid with the predominant land use type. In the province of Drenthe the existing land use pattern still largely reflects the original subdivision between arable land and permanent grassland used for dairy farming. Land use classes were: grassland, arable land and forests. Large extents of grassland and arable land in the province of Noord-Brabant have been turned into maize land since the early seventies because of the increase in intensive livestock farming. Therefore, intensive livestock farming was identified as an

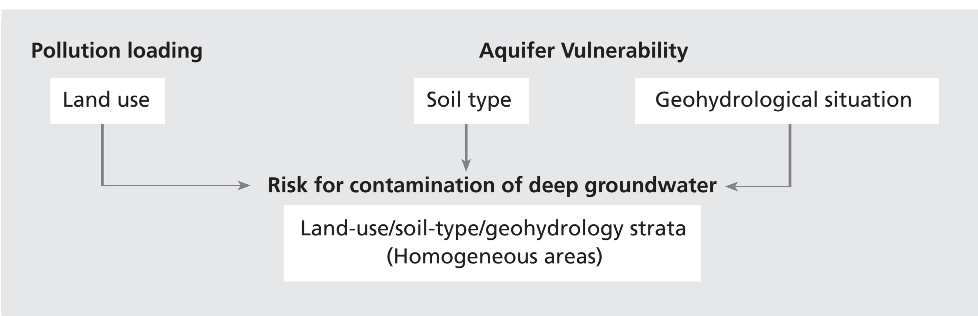


Figure 3.4 - Risk concept used for monitoring design of the regional networks

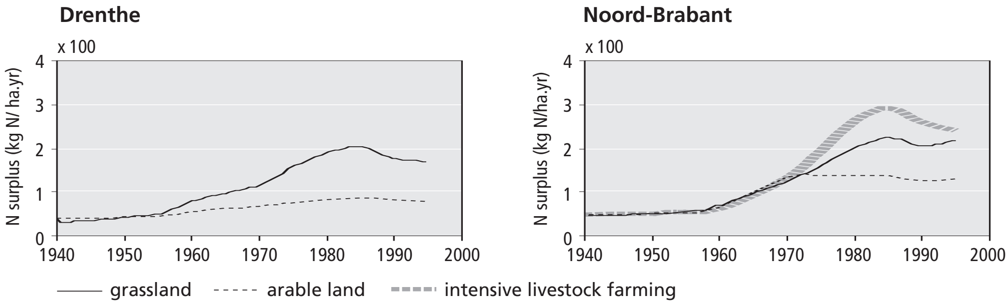


Figure 3.5 - The nitrogen-surplus over the period 1940-1995 for grassland, arable land and intensive livestock farming areas (alternating maize and grassland) in Drenthe and Noord-Brabant

additional land use unit which typically shows both maize and grasslands occupying >30% of the 500 x 500 m grid cells.

Figure 3.5 shows the surplus of nitrogen for grassland, arable land and intensive livestock farming for Drenthe and Noord-Brabant, respectively. The nitrogen surplus was calculated from the input of atmospheric deposition, fertiliser and animal manure minus the nitrogen uptake by crops and the removal by harvesting (data of Menke 1992, van der Grift & van Beek 1996, Beekman 1998, see also Chapter 5). The large differences in the nitrogen surplus between Drenthe and Noord-Brabant are caused by the large amounts of animal manure that is produced in the intensive livestock areas in Noord-Brabant. The excess manure is used both on maize land, grassland and arable land.

The geohydrological situation was mapped to identify recharge, discharge and intermediate areas (for details see Chapter 2). Briefly, the absence of a drainage network, the occurrence of deep groundwater levels, relatively permeable soils and a high topographical position characterised the regional *recharge areas*. Regional *discharge areas* were classified by their low elevation, the shallow groundwater depth and the dense drainage network. The *intermediate areas* have an intermediate topographical position between the regional recharge and discharge areas and include all areas that were not mapped as recharge or discharge areas. These areas were distinguished from the recharge areas on the basis of the existence of a superficial drainage network of ditches, drains and small watercourses (Chapter 2).

Definition of strata for sampling

The risk approach of Figure 3.4 was implemented using a concept of stratified sampling from areas with homogeneous land use, soil types and geohydrological situation. Following the risk concept of Figure 3.4, each stratum represents a specific combination of vulnerability and pollution loading, and a different groundwater quality was anticipated for each of them. The strata were formed by GIS overlay of maps of soil type, land use and geohydrological maps, where all layers had equal weight (Figure 3.6).

The largest difference in the design of the Noord-Brabant and Drenthe regional networks compared with the national network is the added geohydrological stratification. The extra stratification was meant to increase the internal homogeneity, by decreasing one source of variation within the original *land-use/soil-type* strata used in the national monitoring network (Van Duijvenbooden 1993, Pebesma 1997, Reijnders et al. 1998). Land use in discharge areas was not further distinguished because groundwater quality was expected to be controlled by upward seepage and not by land use.

The resulting land-use/soil-type/geohydrology strata were called *homogeneous areas*, because

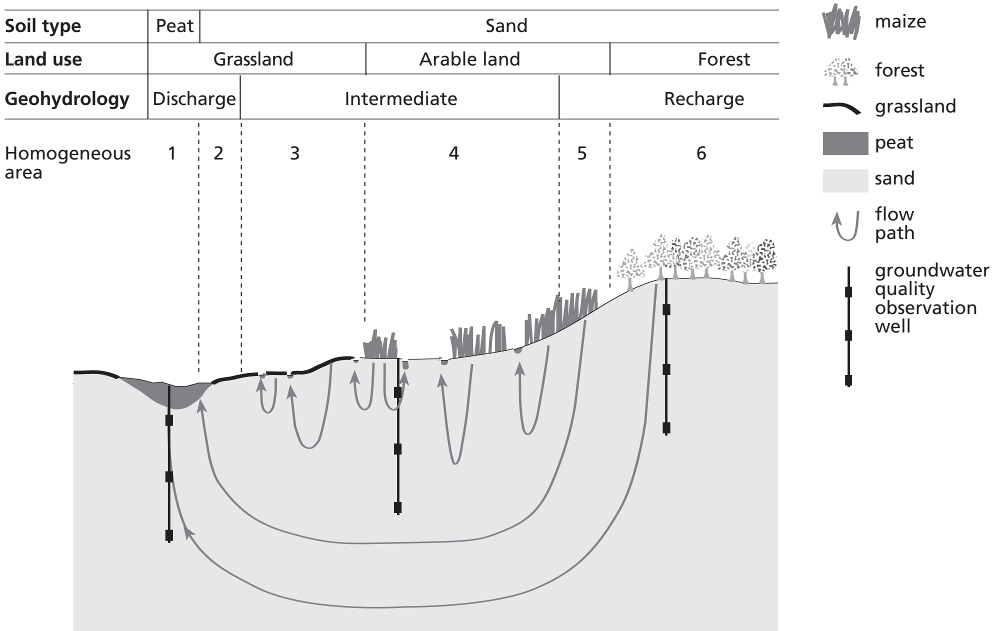


Figure 3.6 - The concept of homogeneous areas that are created by overlay of geohydrology, soil and land use maps. Regional recharge areas and sandy soils are considered vulnerable for diffuse groundwater contamination. Intermediate areas are characterized by the presence of local flow systems

the variation between those areas was expected to be larger than the variation within them. Eventually, the GIS overlays resulted in 8, respectively 6, large homogeneous areas in the Pleistocene sandy parts of Noord-Brabant and Drenthe (Table 3.1). Appendix II shows the spatial delineation of the areas. The homogeneous areas are in fact hydrogeomorphologial units with a specific land use. The land use will not be a constant in time, but will change continuously, for example because of the Dutch national policy that intends to reduce manure inputs. The monitoring networks are meant to register the effects of those changes on groundwater quality in the homogeneous areas.

Selection of well locations

Adding extra wells to the national monitoring network in the province created the regional networks (Broers 1990, Uil & Vlot 1991). Figure 3.7 presents the number of wells that was added for each of the selected homogeneous areas in the sandy Pleistocene parts of Noord-Brabant and Drenthe. The amount of wells was determined according to the areal extent of the homogeneous area and the a priori presumed risks for groundwater contamination. The preliminary risk assessment of Table 3.1 was made using knowledge obtained from earlier studies on groundwater quality and preliminary surveys on regional groundwater quality in the regions itself (Stuurman et al. 1989, 1990). A larger sample size was implemented for recharge areas and agricultural land use relative to intermediate areas, discharge areas and forests. This resulted in a larger monitoring density in areas with high risks for agricultural pollution of deep groundwater (Table 3.1). For example, a relatively large number of wells was located in the area *intensive livestock farming - recharge* in Noord-Brabant and *arable land - recharge* in Drenthe. The relatively small monitoring density in the homogeneous area *intensive livestock*

Table 3.1 - Areal extent, sample size, monitoring density and presumed risks for homogeneous areas in sandy regions of Drenthe and Noord-Brabant

Abbre-viation	Homogeneous area	Spatial % in province	Number of wells	Network density (km2/well)	Risk for agricultural pollution	Risk for acidification
Drenthe						
a-r	Arable land-recharge	7.9	10	19.9	high	moderate/high
g-r	Grassland-recharge	8.3	7	30.1	high	moderate/high
a-i	Arable land-intermediate	13.7	13	26.8	moderate	moderate
g-i	Grassland-intermediate	21.9	16	34.7	moderate	moderate
f-r	Forests-recharge	9.9	7	35.9	low	high
dis	Discharge areas	19.7	11	40.0	low	low
Noord-Brabant						
a-r	Arable land-recharge	0.4	4	4.5	high	moderate/high
il-r	Intensive livestock-recharge	4.8	14	16.6	high	moderate/high
a-i	Arable land-intermediate	4.7	9	25.2	moderate	moderate
g-i	Grassland-intermediate	8.3	12	33.5	moderate	moderate
f-r	Forests-recharge	8.8	12	35.5	low	high
f-i	Forests-intermediate	9.0	11	39.7	low	moderate
dis	Discharge areas	12.3	13	45.7	low	low
il-i	Intensive livestock-intermediate	22.2	20	53.0	moderate	moderate

farming-intermediate in Noord-Brabant is due to the large areal extent of the area type. For this area, a sample size of 20 wells was considered adequate to estimate typical concentrations, because precision of the estimates in a stratified sampling approach is determined by sample size and not by network density (Cochran 1977, see Chapter 4).

Because the areal extent of the areas was taken into account, the newly designed networks are better balanced than the original national network in the two provinces (Figure 3.7). For

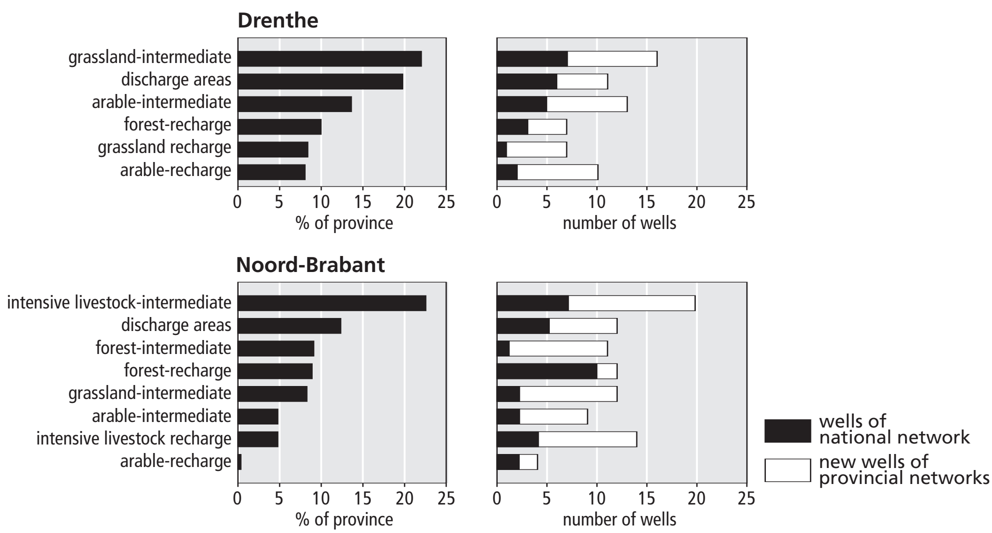


Figure 3.7 - Spatial percentage of the homogeneous areas in the sandy Pleistocene areas of Drenthe and Noord-Brabant and the numbers of national wells and added provincial wells

example, the national network in Noord-Brabant emphasizes forests in recharge areas, whereas the provincial network also covers the forests in intermediate areas that have similar areal extent.

Digital maps were used to help select the well locations in the larger delineations of a homogeneous area. This was done to assure homogeneous land use and geohydrological conditions around the well locations. Digital maps of drinking water protection zones and point sources of contamination were used in choosing well locations to prevent the local influence of specific regulations or local contamination. Groundwater flow directions were read from groundwater maps and topographical maps, to estimate the upstream catchment area of the planned well in order to ensure a homogeneous land use upstream (Figure 3.8). In this way, the observation wells yield a vertical pattern of groundwater quality because the groundwater age increases with depth. This is especially true for the regional recharge areas where a superficial drainage network is absent. Much larger variations in groundwater age are expected for the drained, intermediate areas (Chapter 2).

The local position of the selected well locations was checked in the field to avoid influence of local sources of groundwater pollution and to assure a representative local land use, soil type and geohydrological situation. The position of the well in relation to the superficial drainage network was carefully judged for the recharge and intermediate geohydrological situations. Wells in recharge areas were positioned in the downstream parts of the recharge areas. Much attention was paid to field accessibility: locations at minor roads and paths were preferred to prevent local effects of the use of road salts at larger roads. No wells were installed on military terrain because accessibility was not guaranteed.

The resulting regional networks of Noord-Brabant and Drenthe consist of 122 and 79 monitoring wells, respectively. About half of these are part of the national monitoring network of the Netherlands, the rest were added to the monitoring network by the provinces (Figure 3.7). The focus of the present study was on the Pleistocene sandy areas in the provinces which contain 95 and 64 wells in Noord-Brabant and Drenthe, respectively.

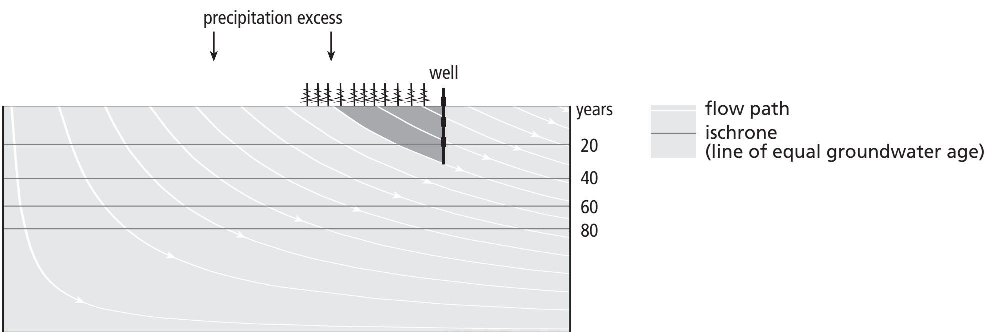


Figure 3.8 - Well locations were selected downstream of areas with homogeneous land use. Deeper screens collect older groundwater from the same land use unit.

Well completion, monitoring procedures and network exploitation

The extra provincial wells were all newly made in the period 1990-1992 using a well design similar to the national wells (Snelting, 1990, Van Duijvenbooden 1993, Broers & Bergsma 1992). The wells contain two 2″ screens of 2 m length at about 9 and 24 m depth and one 1″ screen at 15 m depth. The exact depth of the well screens depends on the hydraulic properties at the well locations. Table 3.2 lists the average screen depths and standard deviations for the complete networks of Noord-Brabant and Drenthe. Some wells have an additional 2″ screen directly below the average lowest groundwater table.

Table 3.2 - Average screen depths for the 5-15 m and 15-30 m depth intervals in the regional networks of Noord-Brabant and Drenthe

	Average screen depths and standard deviation (m)	
	Noord-Brabant	**Drenthe**
Shallow screens	8.7 ± 2.1	9.3 ± 1.2
Deep screens	23.5 ± 2.1	24.3 ± 2.0

The screens at 9 and 24 m depth are sampled annually using a submersible Grundfoss pump. Electrical Conductivity (EC), pH, alkalinity and oxygen are measured in the field. Sampling and analysis of the national wells is done at the RIVM. The provincial wells are sampled and analysed by contractors. To assure the quality and comparability of the national and provincial data, standard operation procedures and quality assurance and quality control (QA-QC) procedures have been developed by RIVM and are used by the contractors. Table 3.3 lists the measured chemical components.

Assessing information on the reactivity of the subsurface sediments

Local field studies of groundwater quality in Noord-Brabant indicated the presence of pyrite in the subsurface (for example, Van Beek 1988, 1989). Soil samples were taken from 27 boreholes during the drilling of the provincial wells in 1991, to acquire information on the reactivity at the regional scale. The samples were analysed for a range of chemical elements, including S, using a total destruction method and ICP analysis (Geochem 1992). Sulphur contents were used to calculate pyrite contents, assuming pyrite to be the only source of sulphur. This is justified since gypsum, another possible major source of sulphur, is not present in the shallow subsoil of the Netherlands.

Specifying statistical information goals

Since 1995, suggestions have been made for more specific monitoring information goals and statistical information goals for the joint provincial monitoring networks (Baggelaar & Van Beek 1997, Jousma et al. 1997, Broers & Peeters 2000). Broers & Peeters proposed to determine specific information goals for each of the homogeneous areas, accounting for the presumed risk for contamination of deep groundwater. Hence, specific information goals and different ambition levels were established for high, intermediate and low-risk areas. The definition of different information goals for high-risk and low-risk areas corresponds to the original design concept that already used differentiated monitoring densities in the homogeneous areas (Table 3.1).

Table 3.3 - Chemical components analysed in the monitoring networks (situation 1995)

Frequency	**Chemical properties**
Annually, lab	EC, pH, Temp, O_2 (field), Ca, Mg, Na, K, NH_4, NO_3, SO_4, HCO_3, Cl, PO_4, P-tot, Fe, Mn, DOC, Al, Zn,
	As, Cd, Cr, Ni, Cu, Ba, Sr, Pb (national wells)
Once every 5 year	As, Cd, Cr, Ni, Cu (provincial wells)
Once	tritium (1983 national wells, 1992 provincial wells)
Ad hoc	groups of pesticides
	industrial organic contaminants

The EU Water Framework Directive (EU 2000) distinguishes two types of monitoring objectives: (1) ‘to present an overview of groundwater chemical status’ and (2) ‘to detect the presence of long-term anthropogenically induced upward trends in the concentration of pollutants’. These objectives were refined and yielded the following two monitoring information goals for the regional networks in the Netherlands:

1. determination of time averaged groundwater quality characteristics in homogeneous areas at two depths for a specific monitoring period
2. determination of quality changes in time in homogeneous areas at two depths for a specific period.

For the second monitoring information goal, the trend definition of Loftis (1996) was adopted who defines a trend as ‘a change in groundwater quality over a specific period in time, which is related to land use or water quality management’.

The two monitoring information goals formed the basis for the definition of area-specific statistical information goals for high-risk, moderate-risk and low-risk areas (Table 3.4). The assessment of typical values was considered a basic requirement for each of the homogeneous areas (statistical information goal A). For *high-risk* areas the further ambition is to determine the magnitude of the contamination and to identify the magnitude of temporal trends (statistical information goals D, F and G). The ambition is to estimate typical values, proportions of contaminated groundwater and temporal trends with a predefined precision. For *moderate-risk* areas the focus is on the signalling of groundwater contamination (statistical information goals C and E). For these areas it is adequate to determine if any contamination occurs or if any trend is present. For *low-risk* areas the aim is to identify base line concentrations and to demonstrate differences with high-risk areas (statistical information goals A and B).

These statistical information goals serve also as a framework for the evaluation and optimization of the regional networks (see Chapter 4). In this chapter the emphasis is on the

Table 3.4 - Specific statistical information goals for high, moderate and low-risk homogeneous areas for the Dutch provincial networks

	Monitoring information goal	**High-risk areas**	**Moderate-risk areas**	**Low-risk areas**
	Ambition level	**High**	**Moderate**	**Low**
	Overview of chemical status			
	Determination of time averaged characteristics of homogeneous areas at 5-15 and 15-30 m depth for a specific monitoring period			
A	Determine typical values (medians/percentiles)	x	x	x
B	Identify differences between areas	x		x
C	Signal the exceeding of environmental standards		x	
D	Determine the proportion of contaminated groundwater *	x		
	Temporal trends			
	Determination of changes in time in homogeneous areas at 5-15 and 15-30 m depth for a specific monitoring period			
E	Signal temporal trends		x	
F	Determine the median temporal trend	x		
G	Determine the temporal trend in the proportion of contaminated groundwater	x		

* i.e. the proportion of groundwater exceeding environmental or drinking water standards

statistical information goals A, B and D which involve the time-averaged characteristics of homogeneous areas over 1995-1998.

3.4 Methods of data analysis

The Dutch environmental policy focuses on specific pollution issues, such as agricultural pollution, acidification and trace element dispersion (including both organic and anorganic contaminants). Indicators are used for each environmental issue to define the state of groundwater contamination. Indicators for agricultural pollution are nitrate, potassium and total-phosphate. For acidification, the pH and the aluminum concentration are normally evaluated (for example, Reijnders 1998, Pebesma 1997). The disadvantage of the sole use of these indicators is that the masking influence of geochemical reactions on groundwater quality is not easily recognised. Four extra indicators were used for the data analysis of the two regional networks, which are based on geochemical knowledge of probable subsurface reactions. The oxidation capacity (OXC) was used as an extra indicator for agricultural pollution (see Appendix I for details). The hardness/alkalinity ratio and the calcite- and siderite-saturation indices are used as extra indicators for acidification (Appendix I). These indicators provide indications of important subsurface buffering mechanisms, such as the reduction of nitrate by the oxidation of pyrite and the neutralizing of acidification by carbonate dissolution.

The monitoring results of the annual sampling rounds of 1995 to 1998 were used for data analysis. The dataset of 1995-1998 was chosen after quality checks on the monitoring data, such as checks on electro-neutrality and a comparison of field EC and field pH versus lab EC and lab pH, respectively. The concentrations of each monitoring screen were averaged over the 4 monitoring years and then used for further data-analysis. This reduces the effect of outliers in the time series of the individual monitoring wells. All concentrations below the detection limit have been given the value of 0.5 times the detection limit to enable the evaluation of summary statistics for all the monitoring data. This is sensible because the monitoring information goals have no special focus on the very low part of the frequency distribution.

As a first step in the data analysis, the frequency distribution of the indicator is presented in box plots using Tukey's hinges for the top and bottom of the boxes (Helsel & Hirsch 1992). The box plots show the complete range of the frequency distribution, including outliers and extreme values. Because outliers and extremes are common in the groundwater quality data sets, non-parametrical methods were preferred for the estimation of typical values (statistical information goal A), the evaluation of differences between areas (statistical information goal B) and proportions of contaminated groundwater (statistical information goal D). Outliers have only been removed from the data set if there was strong evidence that the data were not representative for the respective homogeneous area.

The median was used as a measure for the typical value of an indicator for a homogeneous area (statistical information goal A). The uncertainty of the estimated median value was assessed by computing a 95% two-sided non-parametric confidence interval, using the binomial distribution (Helsel & Hirsch, 1992, appendix III.1). These confidence intervals are non-symmetric if the frequency distribution of the sample is skewed.

To evaluate statistical information goal B, the differences between the median values in the homogeneous areas were evaluated using a multiple-comparison test on ranks (Tukey method, two-sided, $\alpha < 0.05$, appendix III.2). The multiple comparison test was only performed if a Kruskall-Wallis one-way Analysis of Variance (ANOVA) test indicated that the within-group variance was smaller than the inter-group variance. Significant differences between the concentrations of the homogeneous areas are only found if sample sizes in both the compared areas are large enough and limited variation within the areas is observed.

For statistical information goal D, the proportion (or percentage) of contaminated groundwater was defined by comparing the measured concentration with a relevant water quality standard. For instance, the EU drinking water standard of 50 mg NO_3 l^{-1} was used as threshold value for nitrate contamination. The estimated proportion $\hat{p}_{x_c}$ is defined as:

$$\hat{p}_{x_c} = \frac{u}{n} \tag{3.1}$$

where n is number of observations in the specific homogeneous area and u is the number of observations above the environmental standard. The precision of the estimates was determined using a 95% confidence interval for the estimated proportion (α=0.05, two-sided) using the method of Blyth & Still (Gilbert 1987, see Appendix III.5 for details).

Estimates and confidence intervals of the proportion of young, *post-1950* groundwater have been calculated similarly. Groundwater was classified as post-1950 if the measured concentrations exceed 5 TU (1983) or 2 TU (1992) (section 2.3, Chapter 2). The proportions were determined for each homogeneous area.

The estimates and the calculated uncertainty of the estimates were made on the assumption of random sampling from the homogeneous areas. However, this condition is not strictly fulfilled by the national and provincial networks. Although a first selection of the monitoring locations was made with GIS, the locations had to fulfil conditions of field accessibility and avoidance of upstream point sources. Subsequently, a hydrological expert guess was made on the contributing area of the well to assure homogeneous land use in the upstream area. The presented estimates and precision should therefore be considered to be indicative and not absolute.

Data from several homogeneous areas were combined to enable the comparison of concentrations in agricultural areas in the provinces of Noord-Brabant and Drenthe. The proportion of contaminated groundwater in those combined areas must be corrected for the diverging monitoring density in the individual areas (Table 3.1). The original homogeneous areas are considered to be strata with different weights. The proportion of the combined area was calculated using the spatial percentage of the h original areas as weight-factors W_i calculating:

$$\hat{p}_{x_c(st)} = \sum_{i=1}^{h} W_i . \hat{p}_i \tag{3.2}$$

where W_i = weight of homogeneous area i, $\hat{p}_i$= proportion in homogeneous area i and $\hat{p}_{x_c(st)}$ is the estimated proportion in the combined strata and:

$$\sum_{i=1}^{h} W_i = 1 \tag{3.3}$$

The confidence interval on the estimated proportion $\hat{p}_{x_c(st)}$ must also be corrected to account for the non-proportional allocation of observations in the individual strata. This correction is explained in Appendix III.6.

The proportions of contaminated groundwater were mapped to show the spatial differentiation of groundwater contamination. The method described by Reijnders et al. (1998) was adopted, which was used for the national monitoring network. They defined classes for the proportion of contaminated groundwater, taking the 95% confidence intervals into account. The advantage of the method is that areas are only highlighted where the estimates have a specified precision. The following classes were defined: *very high*; whole confidence interval above 20%, *high*; whole confidence interval above 10%, *low*; whole confidence interval entirely

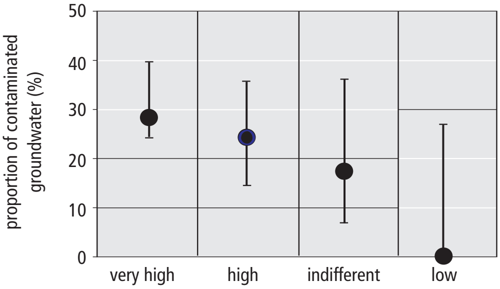


Figure 3.9 - Classification used for mapping of areas with very high, high and low proportions of contaminated groundwater, using the 95% confidence interval on the estimated proportion (criterion for class 'low' changed after van Drecht 1996)

under 30%, and *indifferent*; miscellaneous confidence intervals (Figure 3.9). Here, the 30% threshold, instead of 20%, is used for classification *low*. The 30% threshold better reflected the statistical information goals for low-risk areas than the 20% threshold; estimating the exact proportion of contaminated groundwater was not considered as a relevant statistical information goal in low-risk areas (Table 3.4). A large number of wells, and high costs, would be required to achieve the confidence interval entirely under 20%. For instance, in an area with 0% contamination, 17 wells would be needed to meet this criterion (Gilbert 1987). The class *indifferent* corresponds to areas with estimated proportions near the thresholds of 10 and 20%, or with too few observations to be classified *low* or *high* (Figure 3.9).

3.5 Results and interpretation of 1995-1998 monitoring

This section provides an overview of the data analysis of the regional networks. Two hypotheses are examined:

1. extra geohydrological stratification helps to reduce the variation within the strata,
2. extra indicators based on geochemical knowledge provide a better identification and understanding of contamination patterns than the sole use of the targeted contaminants.

The monitoring results of the two regional networks of Noord-Brabant and Drenthe are used to illustrate the approach and to test these hypotheses. First, the potential for recent groundwater contamination is examined using the proportion of post-1950 groundwater. Second, agricultural contamination of groundwater is assessed using four indicators, including nitrate. Third, the acidification of groundwater is studied using the pH, the hardness/alkalinity ratio and two carbonate saturation indices.

Proportion of post-1950 groundwater

The proportion of post-1950 groundwater in the homogeneous areas was determined as a first step in the data analysis of the monitoring networks. The proportion of post-1950 groundwater decreases from recharge areas, via drained, intermediate areas to discharge areas for both provinces (Chapter 2). This general tendency is also visible in the data of the homogeneous areas in Noord-Brabant and Drenthe (Figure 3.10). In all recharge areas, 100% of groundwater was post-1950 at 5-15 m depth. Intermediate areas showed slightly lower proportions of about 70-90% at this depth. The discharge areas had 20-30% of post-1950 groundwater. At 15-30 m depth the proportion in recharge, intermediate and discharge areas showed larger

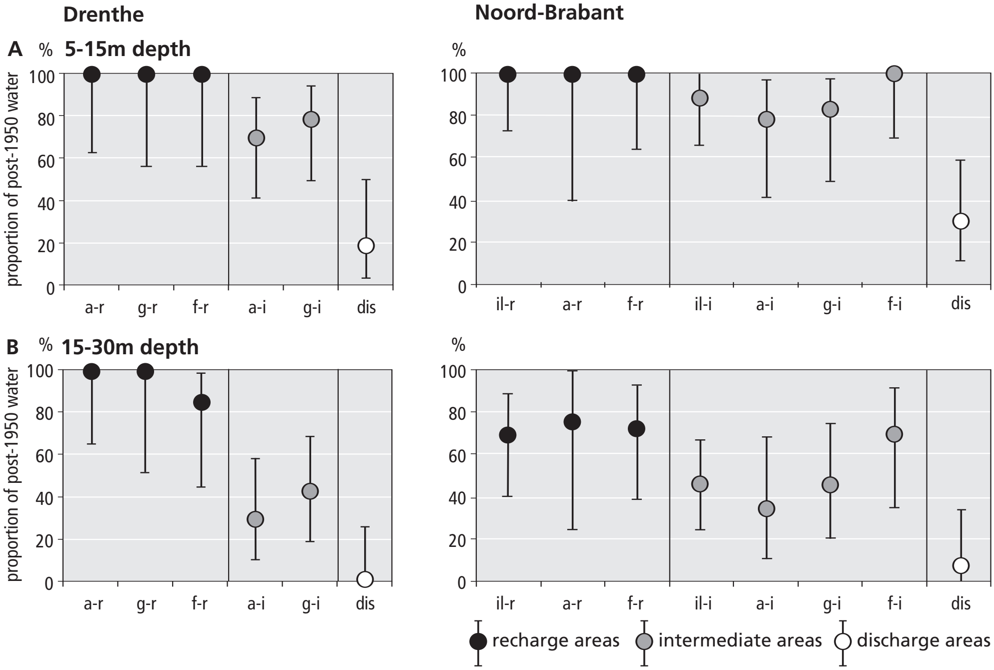


Figure 3.10 - Proportion of post-1950 groundwater for homogeneous areas in Drenthe and Noord-Brabant at 5-15 m depth (a) and 15-30 m depth (b)

differentiation. Recharge areas still had proportions of 85-100% and 70-80% post-1950 groundwater in Drenthe and Noord-Brabant, respectively. Only 30-50% of groundwater was post-1950 in the intermediate areas, except for forest-intermediate *(f-i)* in Noord-Brabant. Discharge areas showed small proportions of post-1950 groundwater at 15-30 m depth.

From these data, we expect that the impact of land use from the period 1950-1990 will be evident in the groundwater in recharge areas and in the shallow groundwater in intermediate areas. The potential of post-1950 groundwater contamination at the 15-30 m depth interval in the intermediate areas is limited to about 30-50% in the agriculturally used areas. In the next sections, these findings have been used to interpret the degree of groundwater contamination. In conclusion, the extra geohydrological stratification used for the regional networks works well to differentiate between areas with larger and smaller potential for recent groundwater contamination.

Agricultural pollution

Nitrate

Nitrate is the most important indicator for agricultural pollution. Figure 3.11 summarises the nitrate concentrations using box plots, median concentrations and proportions of nitrate-contaminated groundwater for the homogeneous areas in Drenthe and Noord-Brabant. Here, nitrate contamination is defined as exceeding the EU threshold value of 50 mg l^{-1}. The homogeneous areas have been ordered according to their presumed risks in the Figures 3.11 to 3.14: high risks for agricultural land in recharge areas, moderate risks for agricultural land in intermediate areas, and low risks for forest areas and discharge areas (see also Table 3.1).

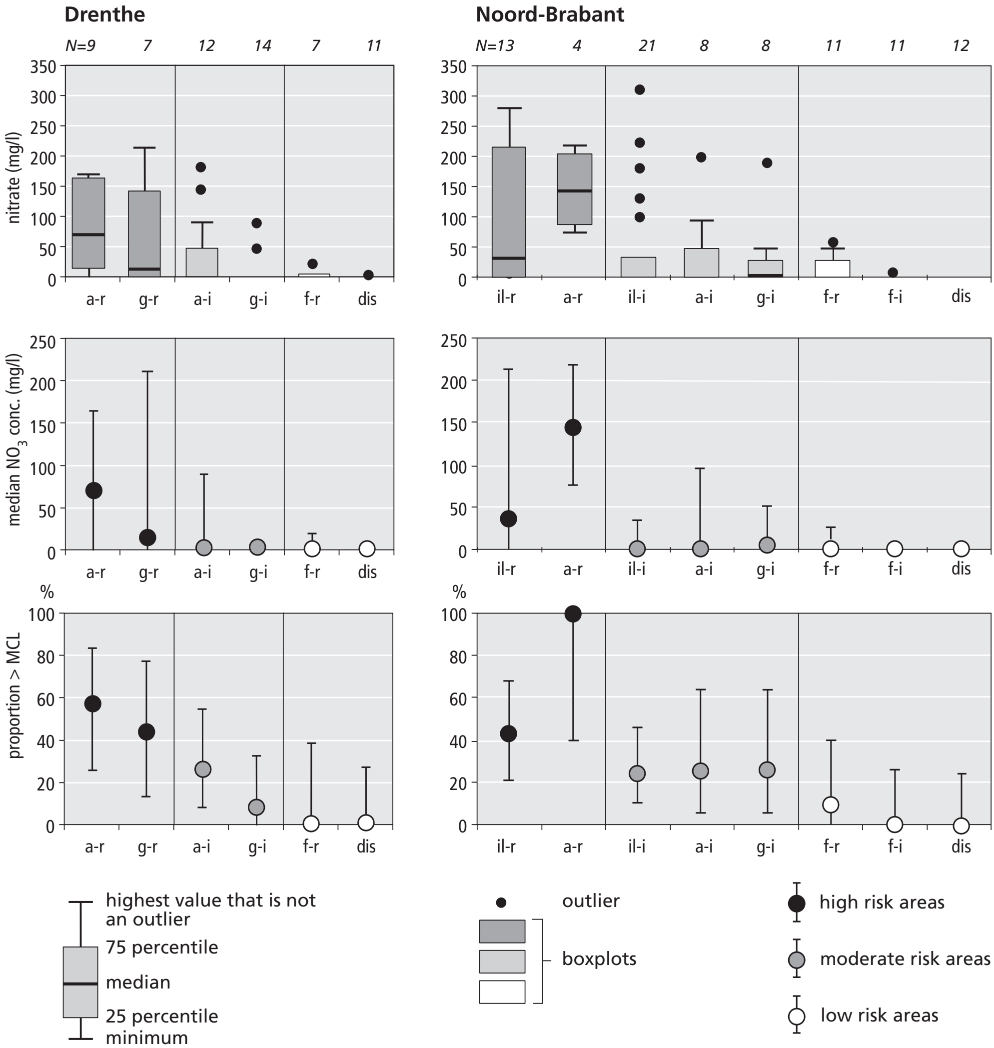


Figure 3.11 - Nitrate concentrations at 5-15 m depth in homogeneous areas in Drenthe and Noord-Brabant for 1995-1998: (a) box plots of nitrate concentrations, (b) median nitrate concentrations + 95% confidence intervals and (c) proportion of nitrate-contaminated groundwater + 95% confidence interval. Acronyms explained in Table 3.1.

The box plots of Figure 3.11 show that nitrate contamination is common in *shallow groundwater*, especially in the high-risk areas where the 75 percentile frequently exceeded 150 mg l^{-1}. Median nitrate concentrations and the proportion of contaminated groundwater showed a similar pattern for the provinces of Noord-Brabant and Drenthe. For example, median nitrate concentrations above the detection limits only occurred in the high-risk areas. Further, the proportion of nitrate-contaminated groundwater generally decreases with the risk order; > 40% in high-risk areas, 5-25% in intermediate-risk areas and about 0% in the low-risk areas, except for forest-recharge areas in Noord-Brabant (<10%).

The nitrate concentrations in the high-risk and moderate-risk areas varied greatly and

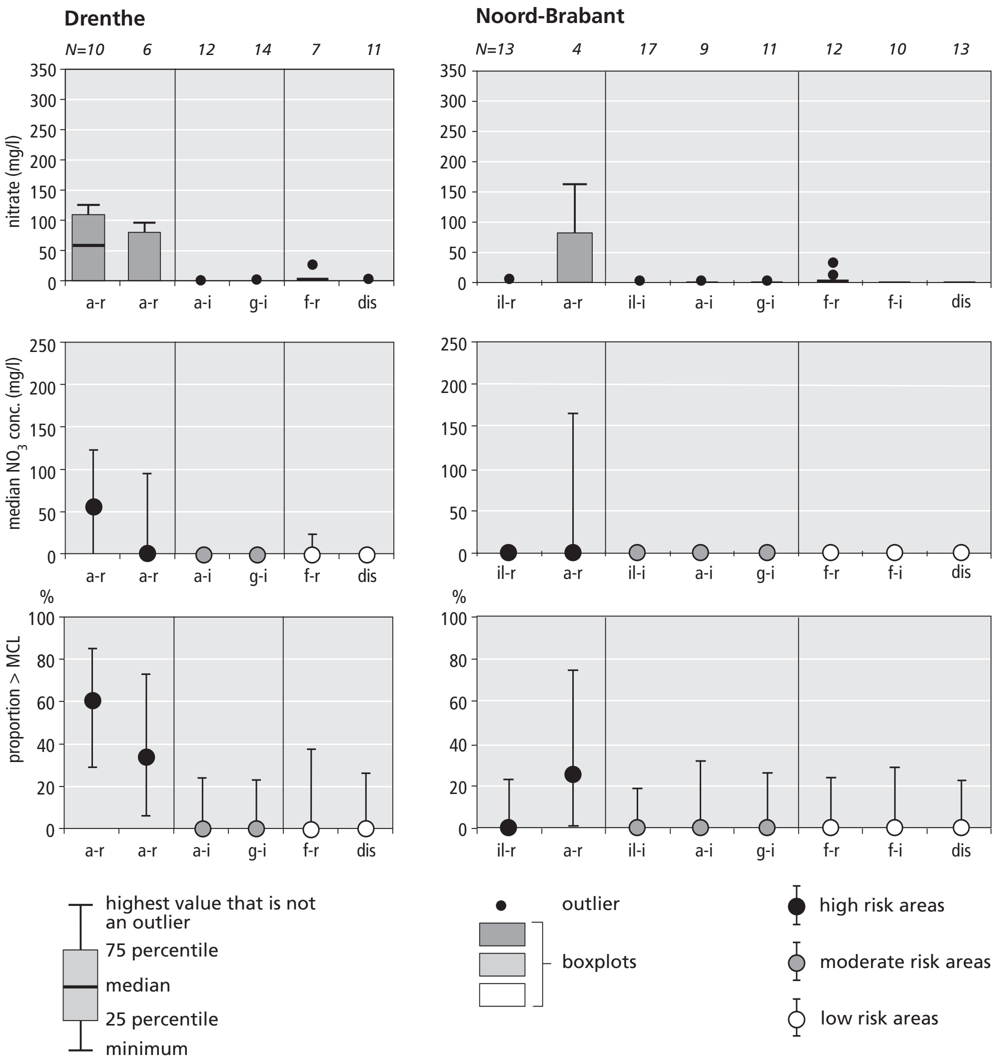


Figure 3.12 - Nitrate concentrations at 15-30 m depth in homogeneous areas in Drenthe and Noord-Brabant for 1995-1998: (a) box plots of nitrate concentrations, (b) median nitrate concentrations + 95% confidence intervals and (c) proportion of nitrate-contaminated groundwater + 95% confidence intervals. Acronyms explained in Table 3.1.

showed a strongly skewed bimodal distribution. Usually, a large number of concentrations below the detection limits was present and also a large number of concentrations far above the EU standard. For example, the nitrate concentrations in the 7 wells in grassland-recharge in Drenthe are <0.5, <0.5, 0.71, 11.8, 133, 147, 211 mg l^{-1}. This yields a median of 11.8 mg NO_3 l^{-1} and a proportion of nitrate-contaminated groundwater of 43%. Non-parametrical 95% confidence intervals around these estimates are given in the figures. Because of the highly variable concentrations and limited sample size, the confidence intervals were wide in the high-risk areas. The medians and the 75 percentiles in the moderate-risk areas were below the EU standard of 50 mg NO_3 l^{-1}. However, outliers and extreme values were present in all of these

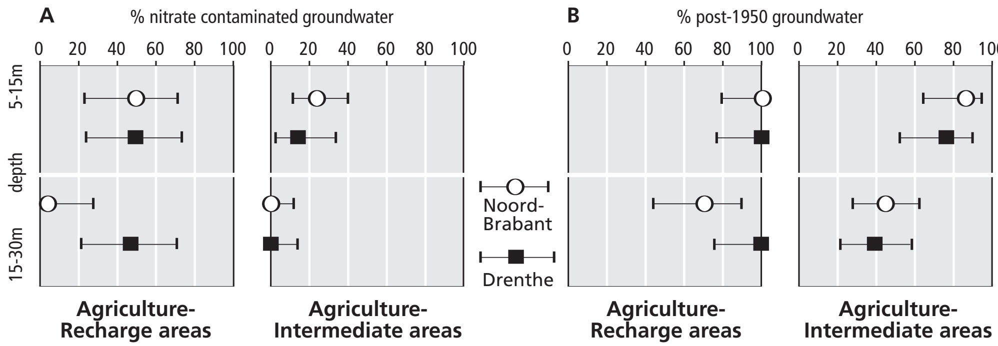


Figure 3.13 - Proportion of nitrate-contaminated groundwater (a) and the proportion of post-1950 groundwater (b) in the combined agricultural areas for two depth intervals in Noord-Brabant and Drenthe for 1995-1998

homogeneous areas. The large variability and bimodal distributions are characteristic for large nitrate data sets, which typically show a major mode at concentrations below detection limits and a secondary mode at higher concentrations (Hallberg & Keeney 1993). This limits the use of nitrate as a general indicator for agricultural pollution.

Differences between the agricultural land use classes grassland, arable land and intensive livestock farming (with similar hydrology) were minor compared with differences between recharge and intermediate areas (with similar land use classes). For example, *il-i*, *a-i* and *g-i* (abbreviations see Table 3.1) showed comparable proportions of nitrate-contaminated groundwater, whereas *il-r* showed a substantial larger proportion than *il-i*. Although this trend is clear from the estimates, the sample sizes were insufficient to significantly demonstrate the differences between high and moderate-risk areas. This is clear from the overlapping 95% confidence intervals. Significant differences were found for *a-r* versus *f-r* and *dis* in Drenthe and *il-r* versus *f-i* and *dis* for Noord-Brabant (multiple comparison test, $p<0.05$).

Nitrate concentrations were much lower at *15-30 m depth* (Figure 3.12). The presence of nitrate was almost restricted to the high-risk recharge areas. A median concentration larger than the detection limit is found only in one homogeneous area in Drenthe *(a-r)* and a significant proportion of contaminated groundwater was only found in *a-r* and *g-r* in Drenthe and in *a-r* in Noord-Brabant. The general conclusion for the 15-30 m depth interval is that nitrate contamination is rare, except for agriculturally used recharge areas in Drenthe.

Since no important differences appeared between the different agricultural land use classes, the classes were grouped to enable comparison between the two depth intervals and the two provinces. Figure 3.13 compares the proportions of nitrate-contaminated groundwater in agricultural recharge areas (left part) with the proportions in agricultural intermediate areas (right part). The proportions were weighted according to the areal extent of the contributing homogeneous areas, using the method described in the previous section. The proportion of nitrate-contaminated groundwater decreased from 5-15 m depth to 15-30 m depth, except for recharge areas in Drenthe. In the Noord-Brabant recharge areas, the estimated proportion of nitrate-contaminated groundwater decreased from about 50% to 5%, whereas no clear decrease was found in Drenthe recharge areas (49 versus 47%). Concluding, nitrate leaching to deeper groundwater is more common in Drenthe compared with Noord-Brabant for high-risk areas with agricultural land use.

The proportion of nitrate-contaminated groundwater was mapped to show the spatial

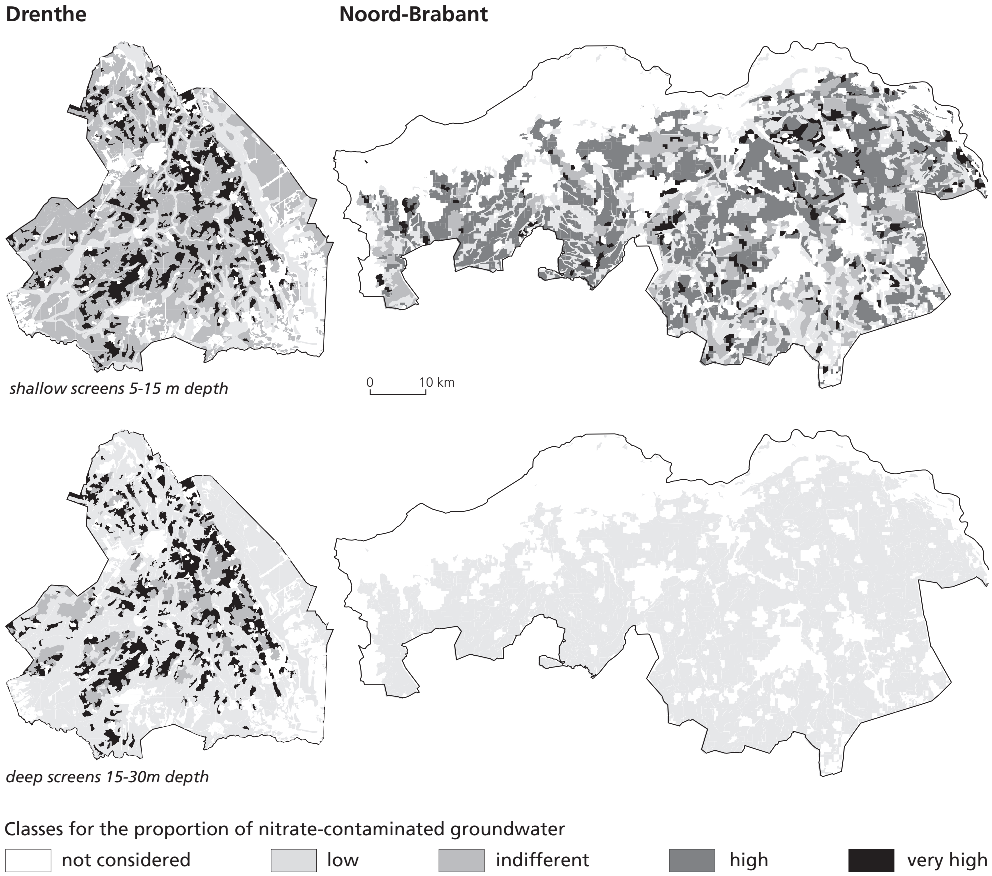


Figure 3.14 - Estimated proportions of surface areas above the drinking water standard for nitrate in Drenthe and Noord-Brabant at 5-15 m depth (a) and 5-30 m depth (b) for 1995-1998. Classification is based on the 95% confidence interval on the estimated proportion.

differentiation of groundwater contamination following the classes *very high*, *high*, *indifferent* and *low* that were described in the previous section. The agricultural areas were classified using the proportions in the combined homogeneous areas, which were shown in Figure 3.13. For the 5-15 m depth interval a clear spatial differentiation was present, both in Noord-Brabant in Drenthe (Figure 3.14). The maps show that areas that were classified as *very high* are restricted to a small percentage of the provinces Noord-Brabant and Drenthe. Table 3.5 shows that the class *very high* comprised approximately 5% and 16% of Noord-Brabant and Drenthe, respectively. The areal extent of areas with classes *high* or *very high* in shallow groundwater was much larger in Noord-Brabant than in Drenthe, where large areas were classified as *indifferent*. At 15-30 m depth, the entire Pleistocene part of the province of Noord-Brabant was classified as *low*, whereas approximately 16% of Drenthe was classified as *very high* and 10 percent as *indifferent* (Figure 3.14 and Table 3.5).

The similar patterns that were found for the proportions of nitrate-contaminated groundwater in the shallow screens in Drenthe and Noord-Brabant may look sensible at first sight, but

Table 3.5 - Percentage of province for 4 classes of nitrate contamination for shallow and deep screens in the regional networks of Noord-Brabant and Drenthe (peat and clay areas and build-up areas not considered)

Class	**Percentage of province**	
	Noord-Brabant	**Drenthe**
Shallow screens		
Very high	5.2	16.2
High	35.2	0
Indifferent	8.8	45.5
Low	21.3	19.7
Deep screens		
Very high	0	16.2
High	0	0
Indifferent	0	9.9
Low	70.5	55.3

become peculiar if one compares the nitrogen pollution loading of Noord-Brabant with Drenthe (Figure 3.5). Based on the pollution input, one would expect larger proportions of contaminated groundwater and higher median concentrations in Noord-Brabant than in Drenthe. Because the concentrations and proportions were similar and proportions in deep groundwater were even higher in Drenthe, the explanation should be found in a different vulnerability, either geohydrologically or chemically (Figure 3.2).

Geohydrological vulnerability was evaluated using the proportion of post-1950 groundwater in homogeneous areas (Figure 3.13). The differences between Noord-Brabant and Drenthe cannot be explained by comparing the proportions of post-1950 groundwater (Figure 3.13b) with the proportions of nitrate-contaminated groundwater (Figure 3.13a). Although the proportion of post-1950 groundwater is smaller in Noord-Brabant recharge areas compared with Drenthe (70 and 100%, respectively) this does not explain the differences in nitrate concentrations.

This leaves differences in subsurface reactivity as the most likely explanation. Figure 3.15 summarises the pyrite contents of 24 wells of the Noord-Brabant monitoring network for 9 depth intervals between 0 and 27 m below the surface. The wells were selected from intermediate and recharge areas. The pyrite contents exhibit a very skewed distribution and a large number of outliers. However, the general tendency of increasing pyrite contents up to 9-12 m depth is visible for the median, the 25 and 75 percentiles and the average pyrite content. Median pyrite concentrations were below 0.05 weight % above 9 m depth and around 0.07 weight.% between 9 and 27 m depth. The average contents were larger, but sensitive to outliers. The large number of outliers is due to the small volume of the sediment samples, which increases the probability of encountering large pyrite contents. This phenomenon is known as the 'support effect'. Average concentrations of sample sizes of larger support would have reduced the variation (see also Chapter 6).

The measured pyrite contents strongly suggest the availability of potential nitrate-reduction capacity of the subsoil, especially below 10 m depth (for example, Postma et al. 1991). Local studies in Noord-Brabant and adjacent areas indeed showed pyrite oxidation as a removing mechanism for nitrate (Van Beek et al. 1989, Broers et al. 1994, Broers & Buijs 1997). The measured amount of pyrite in the 24 wells in Noord-Brabant will effectively slow down the vertical movement of a nitrate front. For example, the vertical velocity of a nitrate front would

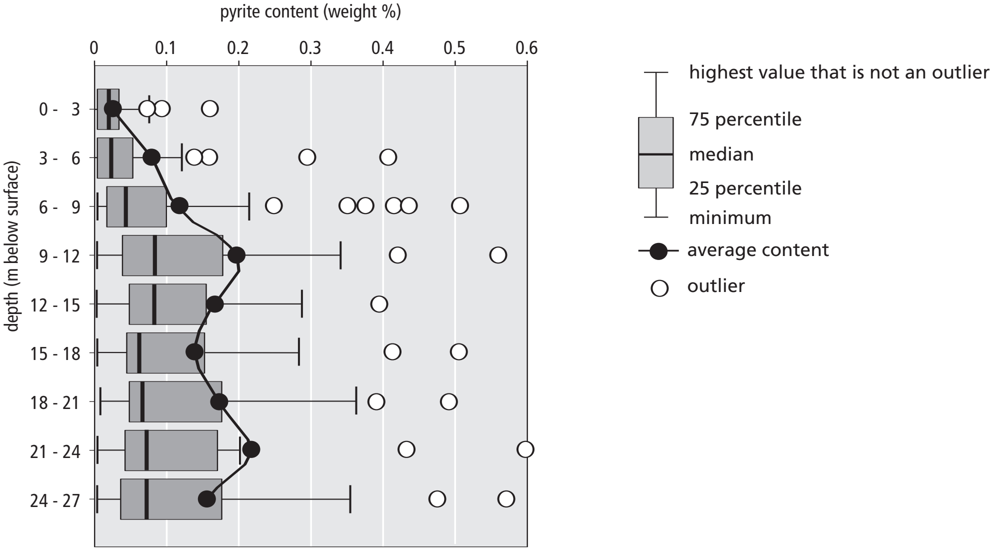


Figure 3.15 - Average pyrite contents (weight %) and frequency distribution (box plots) of 433 soil samples from 24 observation wells in the province of Noord-Brabant.

be reduced from 1 m per year for conservative transport to 11 mm per year, using pyrite contents of 0.1 weight %, a porosity of 0.3 and a constant input concentration of 100 mg NO_3 l^{-1}. This explains the absence of nitrate at larger depths in almost all wells in Noord-Brabant.

The presence of reactive solids in the subsurface is also the most probable explanation for the simultaneous occurrence of concentrations below the detection limits and concentrations far above the EU standards. If nitrate reduction capacity is present in the shallow subsoil, nitrate disappears completely from the groundwater. However, if no reduction capacity is present, nitrate behaves conservatively and is transported at the vertical groundwater velocity. This results in the bimodal frequency distributions that are often observed for nitrate (Hallberg & Keeney, 1993). Even in the high-risk areas, highly variable nitrate concentrations were observed, which hamper the interpretation of the monitoring data and the assessment of typical values.

Oxidation capacity

The observed differences in nitrate contamination must be due to reactive processes in the deeper subsoil, such as the reduction of nitrate by organic matter or by pyrite. The reduction of nitrate by organic matter does not yield reaction products that can be identified unambiguously in groundwater (Korom 1992):

$$5CH_2O + 4NO_3^- \rightarrow 4HCO_3^- + H_2CO_3 + 2H_2O + 2N_2 \qquad (3.4)$$

The reaction product HCO_3 is difficult to identify, because of the complex carbonate chemistry and the large effects of pH on the aqueous distribution of carbonate species (Griffioen & Hoogendoorn 1993). However, the reduction with pyrite oxidation yields the production of sulphate, as follows:

$$5FeS_2 + 14NO_3^- + 4H^+ \rightarrow 5Fe^{2+} + 10SO_4^{2-} + 2H_2O + 7N_2 \qquad (3.5)$$

Thus, if pyrite oxidation has occurred, the agriculturally polluted groundwater is still identifiable after the reaction by increase of the SO_4 concentration. This groundwater is characterised using the oxidation capacity, which was defined as (Postma et al. 1991):

$$OXC = 7\,[SO_4^{2-}] + 5\,[NO_3^-]\ \ mmol\ l^{-1} \qquad (3.6)$$

where brackets indicate molar concentrations and the constants 7 and 5 refer to the number of electrons transferred in the redox half-reactions. The advantage of OXC is that the amount is equal before and after the pyrite oxidation reaction of equation (3.5). This implies that OXC is transported conservatively, under the condition that no other reactions such as nitrate reduction with organic matter or sulphate reduction occur.

Figures 3.16 and 3.17 show the OXC concentrations in Noord-Brabant and Drenthe for 5-15 m depth and 15-30 m depth. The box plots of OXC show distinct differences with the box plots of nitrate. The frequency distribution was less skewed and more normal and the number of outliers and extreme values was much smaller. Moreover, the lower box boundaries almost always deviated from 0 (detection limit). This implies that samples that lacked nitrate did contain sulphate in substantial concentrations. On the whole, the data sets of the homogeneous areas were much more homogeneous for OXC than for nitrate.

At *5-15 m depth*, the medians and 25 and 75 percentiles of OXC decreased from high-risk to low-risk areas (Figure 3.16). The OXC concentrations conformed much better to the presumed risk order than nitrate did: median OXC concentrations in Noord-Brabant recharge areas were above 15 mmol l^{-1}, in intermediate areas between 10 and 15 mmol l^{-1} and in the low-risk areas below 5 mmol l^{-1}. Groundwater was defined as OXC-contaminated if OXC exceeds 15.1 mmol l^{-1}. Using this threshold value, either nitrate was above the EU standard of 50 mg l^{-1} or sulphate was above the standard of 150 mg l^{-1}. The proportion of OXC-contaminated ground-water decreased from high-risk areas to low-risk areas for both Drenthe and Noord-Brabant for 5-15 and 15-30 m depth (Figures 3.16 and 3.17).

The median OXC concentrations were much larger in Noord-Brabant compared with Drenthe for similar risk classes, and better reflected the different input loading in the two provinces. This is especially true for the high and moderate-risk areas, but also for forest areas. Proportions of OXC-contaminated groundwater were also substantially larger in high and moderate-risk areas in Noord-Brabant compared with Drenthe.

The low OXC concentrations in forested recharge areas indicate that atmospheric deposition of sulphate did not contribute significantly to the OXC concentrations in the agricultural areas. OXC concentrations above 7 mmol l^{-1} in Noord-Brabant and 3 mmol l^{-1} in Drenthe must be attributed to inputs of manure or fertiliser. The larger OXC concentrations under forests in Noord-Brabant were explained by the larger dry deposition of SO_4 and NH_3 from surrounding agricultural areas (Erisman & Bobbink 1997).

For groundwater between *15-30 m depth*, the OXC concentrations were less than for 5-15 m depth (Figure 3.17). OXC concentrations above the 15.1 mmol l^{-1} threshold were almost completely limited to the high-risk recharge areas. In these areas about 30% of the ground-water in Noord-Brabant appeared to be OXC contaminated. This contamination must be contributed to sulphate levels above the EU standard, because nitrate was almost absent at this depth interval. The OXC threshold is seldomly exceeded in Drenthe.

The large differences in input of manure between Noord-Brabant and Drenthe became visible when correcting for the effects of pyrite oxidation using the OXC. OXC indicated that the amount of agriculturally contaminated groundwater is much larger in Noord-Brabant compared with Drenthe. This was not observed using nitrate as indicator for agricultural pollution, because pyrite oxidation has a large effect on nitrate transport in Noord-Brabant.

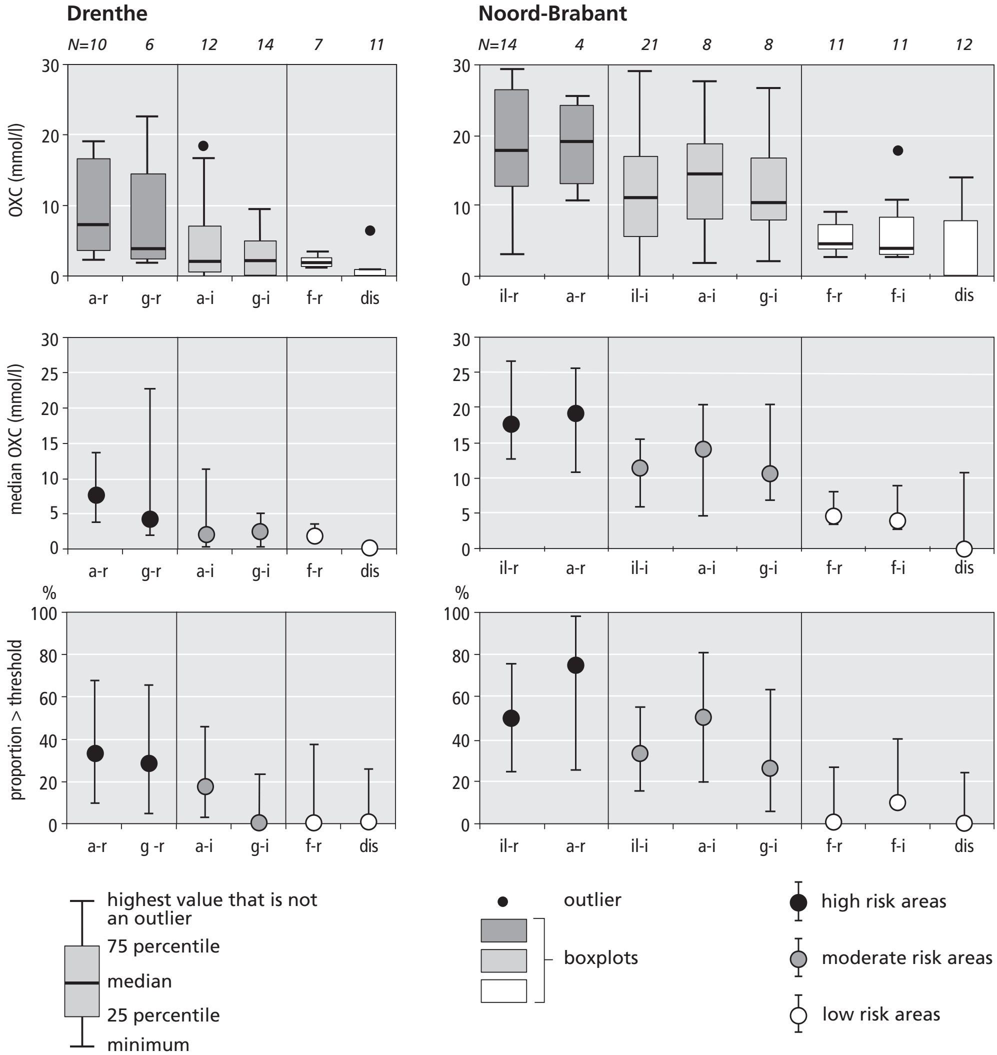


Figure 3.16 - OXC concentrations at 5-15 m depth in homogeneous areas in Drenthe and Noord-Brabant for 1995-1998: (a) box plots of OXC concentrations, (b) median OXC concentrations + 95% confidence intervals and (c) proportion of OXC contaminated groundwater + 95% confidence intervals. Acronyms explained in Table 3.1.

Apparently the pyrite redox buffer is absent in the shallow subsoil of the Drenthe recharge areas, which enables the leaching of nitrate to deeper levels than in Noord-Brabant.

General indications of agricultural pollution

The results of OXC indicate that agricultural pollution of groundwater is more widespread than is suggested by nitrate concentrations alone. Up to now groundwater was defined to be agriculturally polluted if rather large thresholds were exceeded, such as the EU standards for nitrate or sulphate. However, signals for agricultural pollution of groundwater can be detected earlier if smaller threshold concentrations are used or general pollution indicators. For

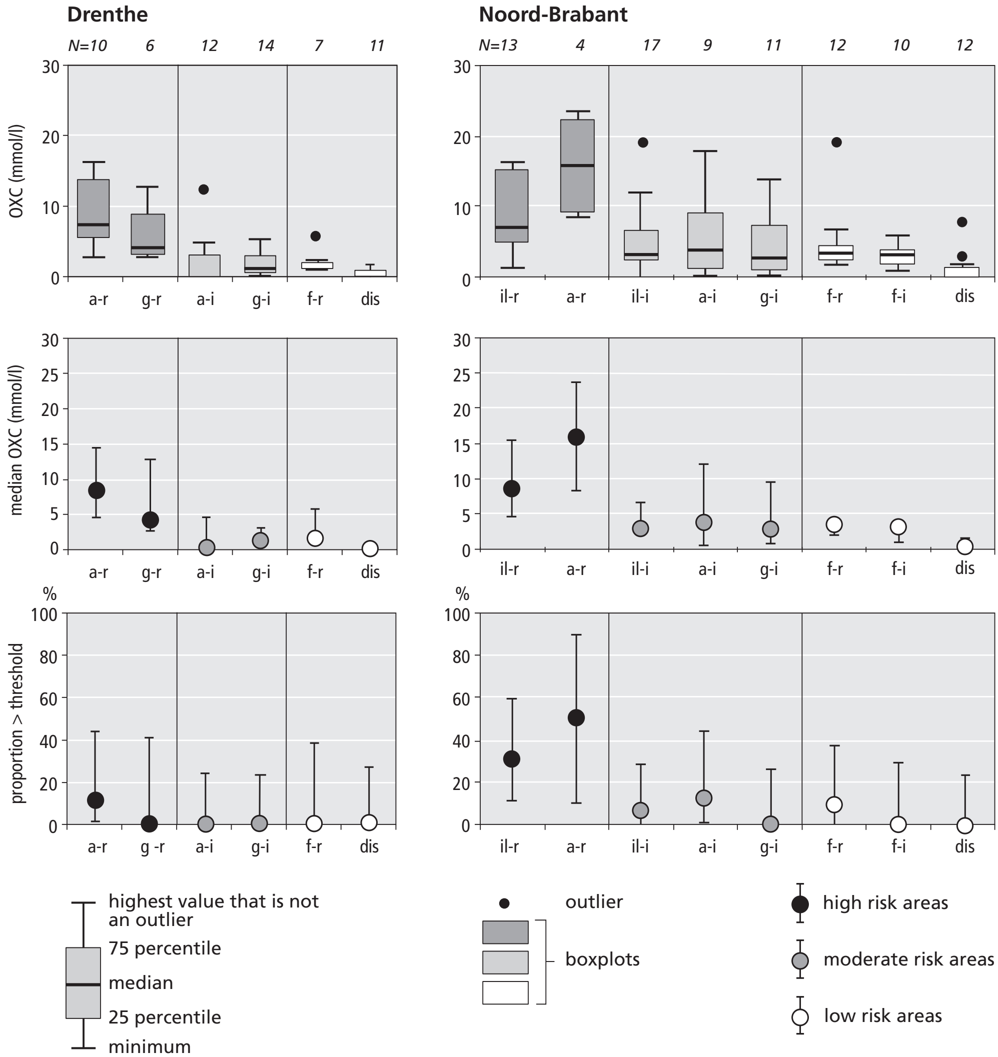


Figure 3.17 - OXC concentrations at 15-30 m depth in homogeneous areas in Drenthe and Noord-Brabant for 1995-1998: (a) box plots of OXC concentrations, (b) median OXC concentrations + 95% confidence intervals and (c) proportion of OXC contaminated groundwater + 95% confidence intervals. Acronyms explained in Table 3.1.

instance, groundwater could be defined as being agriculturally contaminated if OXC exceeds 7 mmol l^{-1}, if we consider maximum observed concentrations under forests as a reference. Table 3.6 summarises the proportions of agriculturally contaminated groundwater using 2 general pollution indicators and smaller thresholds for OXC and nitrate concentrations.

The POLIN index of Stuyfzand (1993) includes three quality aspects of inorganic pollution sources: A. acidification/eutrophication, B. excessive application of fertilisers and manure spreading and C. pollution with environmental hazardous inorganic trace elements (details in Appendix I). In our analysis, groundwater was considered polluted if POLIN > 1.5, which corresponds to Stuyfzand's classes 'slightly polluted' to 'extremely polluted'.

The pollution index MANURE uses the Potassium Adsorption Ratio (PAR), the oxidation capacity (OXC) and the chloride concentration to identify groundwater that is influenced by nutrients and salts from animal manure or fertiliser (details in Appendix I). Using MANURE, groundwater is considered to be agriculturally polluted if one of the following conditions was fulfilled: (1) PAR > 0.1 mmol$^{0..5}$ l$^{-0.5}$, (2) OXC > 7 mmol l^{-1} or (3) Cl > 50 mg l^{-1}. These thresholds were chosen using information from local studies in the Netherlands (Griffioen & Hoogendoorn 1993, Frapporti et al. 1995, Broers et al. 1994, Broers & Buijs 1997, Griffioen 2001). The indicators MANURE and POLIN are less sensitive to specific hydrogeochemical reactions, because they make use of several chemical properties, including conservative species like chloride.

In Table 3.6 the proportions of contaminated groundwater are compared with the proportion of post-1950 groundwater for agriculturally used areas. The table indicates two main trends; a trend in rows and a trend in columns. First, the proportion of contaminated groundwater generally decreases from recharge via intermediate to discharge areas. Second, the proportion of polluted groundwater decreases from POLIN and MANURE via OXC to NO_3.

The first trend shows the effects of *groundwater age*; the proportion of contaminated groundwater (using MANURE or POLIN) is close to the proportion of post-1950 groundwater (Table 3.6). Thus, almost any post-1950 groundwater in the agricultural areas showed signs of agricultural pollution in both Drenthe and Noord-Brabant. The second effect, the decrease in rows, should be attributed to *reactive processes*. Indicators using a range of components, including conservative ones such as chloride, showed larger proportions of contaminated groundwater than indicators based on one or two reactive components.

These effects occur in both the shallow and deeper screens of Noord-Brabant and Drenthe. In Noord-Brabant, a strong decrease is visible between OXC and nitrate, which indicates the

Table 3.6 - Proportion of post-1950 groundwater and proportions of four indicators of agricultural pollution in the Noord-Brabant and Drenthe regional monitoring networks

	Post-1950 water (%)	POLIN >1.5 (%)	MANURE >1 (%)	OXC >7 (%)	NO_3 >25 (%)
			Reactive processes		
Noord-Brabant					
5-15 m depth					
Agriculture-recharge	100	87	93	87	60
Agriculture-intermediate	85	72	75	66	33
Discharges areas	31	17	33	25	0
15-30 m depth					
Agriculture-intermediate	70	62	64	57	2
Discharges areas	43	29	34	27	0
Agriculture-recharge	8	0	8	8	0
Drenthe					
5-15 m depth					
Agriculture-recharge	100	85	100	52	52
Agriculture-intermediate	75	34	51	18	18
Discharges areas	18	0	18	0	0
15-30 m depth					
Agriculture-recharge	100	83	65	46	51
Agriculture-intermediate	39	23	20	3	0
Discharges areas	0	0	9	0	0

Groundwater age

reactive, attenuating effect of pyrite oxidation on the nitrate concentrations. In Drenthe the proportions of OXC>7 and NO_3>25 are almost similar and pyrite oxidation does not seem to affect the nitrate concentrations.

Concluding, the proportion of agriculturally polluted groundwater should not be determined by using only one, potentially reactive, chemical component if the general monitoring objective is to assess the impact of human influence on groundwater quality.

Acidification

In the Netherlands, pH and aluminum are generally used as indicators for acidification of groundwater (Reijnders et al 1998, Pebesma 1997). The pH of the groundwater is important because it controls the mobility of trace metals. The mobility of trace metals such as Zn, Ni, Cd and Cu increases dramatically below pH=5. Therefore, groundwater was defined to be acidified if pH is less than 5.0. For acidification, the homogeneous areas were ordered according to the presumed risks following Table 3.1. The highest risk for acidification of deep groundwater was expected in the forests in recharge areas, because of adsorption of NH_3, HNO_3 and H_2SO_4 from dry and wet deposition on the leaves of the trees and the subsequent washing off into the soil (Van Breemen et al. 1982, de Vries & Breeuwsma 1987, de Vries 1993). Especially in Noord-Brabant, the supply of NH_3 that is volatilized from intensively manured fields causes large inputs of acidifying compounds into the forest areas (Erisman & Bobbink 1997). High risks for acidification of deep groundwater were also presumed for groundwater under agricultural recharge areas, because the large inputs of NO_3 and NH_3 are an indirect acidifying factor (Jacks 1993, de Vries 1988). Nitrification of NH_3 causes extra acidification to the atmospheric sources. However, liming of agricultural soils is done to prevent acidification of the root zone and might reduce the overall acidifying effect, but increases the hardness of the groundwater (Broers & Griffioen 1992).

Figure 3.18 presents the median pH and the proportion of groundwater with pH<5 for the *shallow* screens for Drenthe and Noord-Brabant, following the risk order described before. The pH patterns for Drenthe and Noord-Brabant are very similar: increasing pH from high-risk recharge areas to discharge areas. The lowest median pH was found for forest-recharge areas: 5.0 and 4.7 for Drenthe and Noord-Brabant, respectively. This confirms the hypothesis of acidification of groundwater under forests. In these areas, the estimated proportion of pH<5 is 40-60%. Surprisingly low pH's were also found in the agriculturally used recharge areas in Drenthe and Noord-Brabant. Proportions of acidified groundwater were less than 30% in those areas, except for *a-r* in Noord-Brabant. In general, the pH in Noord-Brabant recharge areas was slightly less than in Drenthe recharge areas. The pH in discharge areas was neutral: 6.8-7.0.

The pH in *deeper* groundwater follows the same pattern as in shallow groundwater; increasing pH from recharge via intermediate to discharge areas (Figure 3.19). However, median pH's in recharge areas and intermediate areas were higher and a pH < 5 was only found incidentally in recharge areas. The pH of groundwater in recharge areas was substantially lower than in intermediate areas (median pH 5-6 versus pH 6-7). These results indicate that the pH<5 acidification front has not yet reached the deeper monitoring interval in the recharge areas because of retardation. The retardation is due to ion-exchange of aluminum and protons and the dissolution of aluminum hydroxides and silicates (Sverdrup &Warfvinge 1988, Appelo et al. 1982, Appelo & Postma 1993, Böttcher et al. 1985).

The geohydrological stratification was effective to show differences in pH for the forest areas in Noord-Brabant. *Forest-recharge* shows substantial lower pH than *forest-intermediate* (median pH 4.8 versus 5.8 in the shallow groundwater, respectively). These forests have a different hydrogeomorphological position in the landscape and a different history. The areas *f-r* are

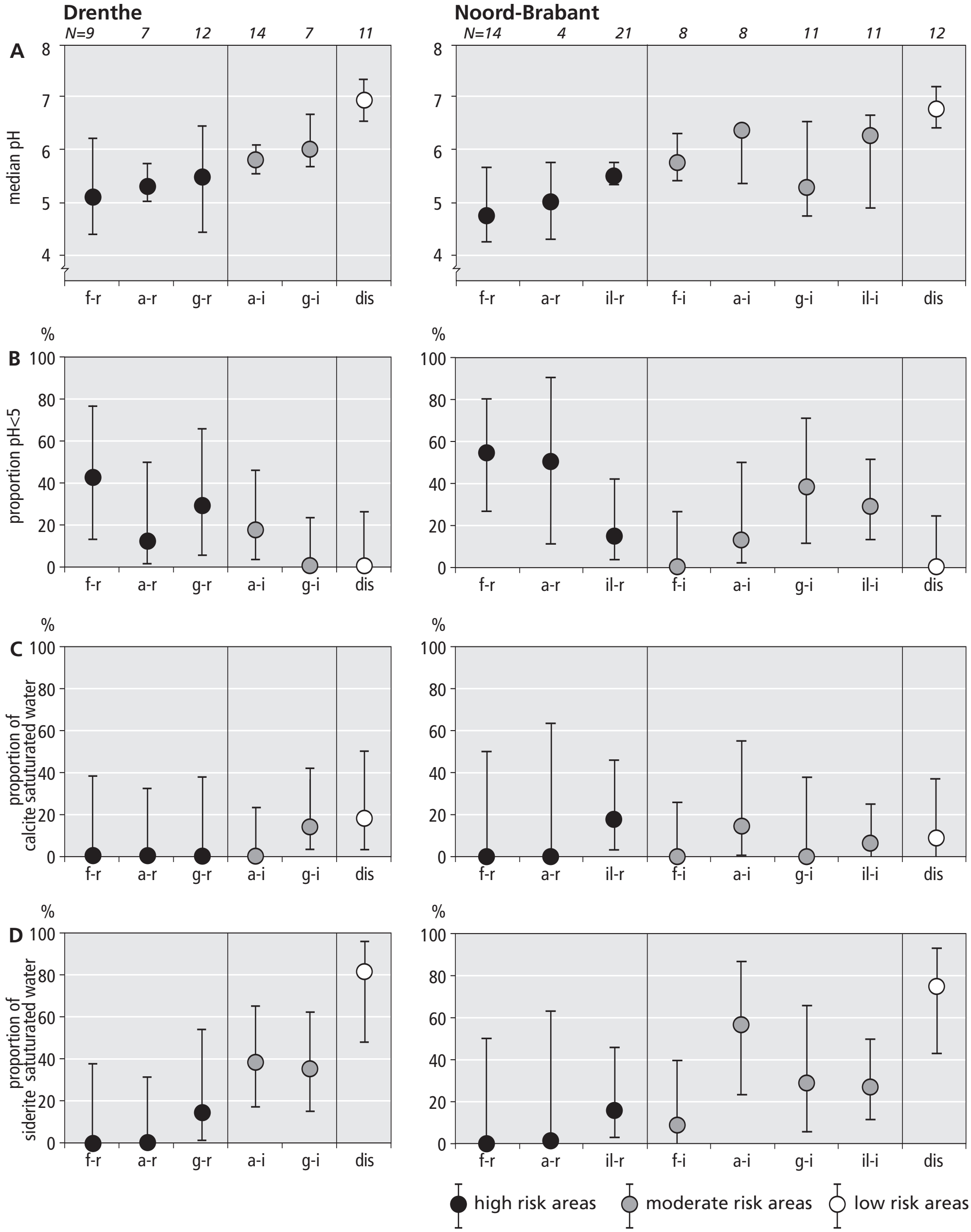


Figure 3.18 - Acidification of groundwater at 5-15 m depth in Drenthe and Noord-Brabant for 1995-1998: (a) median pH, (b) proportion of pH<5, (c) proportion of calcite saturated groundwater and (d) proportion of siderite saturated groundwater. Acronyms explained in Table 3.1.

situated on dry soils at younger dunes. The forests in the intermediate areas are situated in wet areas that were drained artificially to enable wood production. In these forests even fertiliser was used to stimulate wood production. The differences between the two forest types are explained

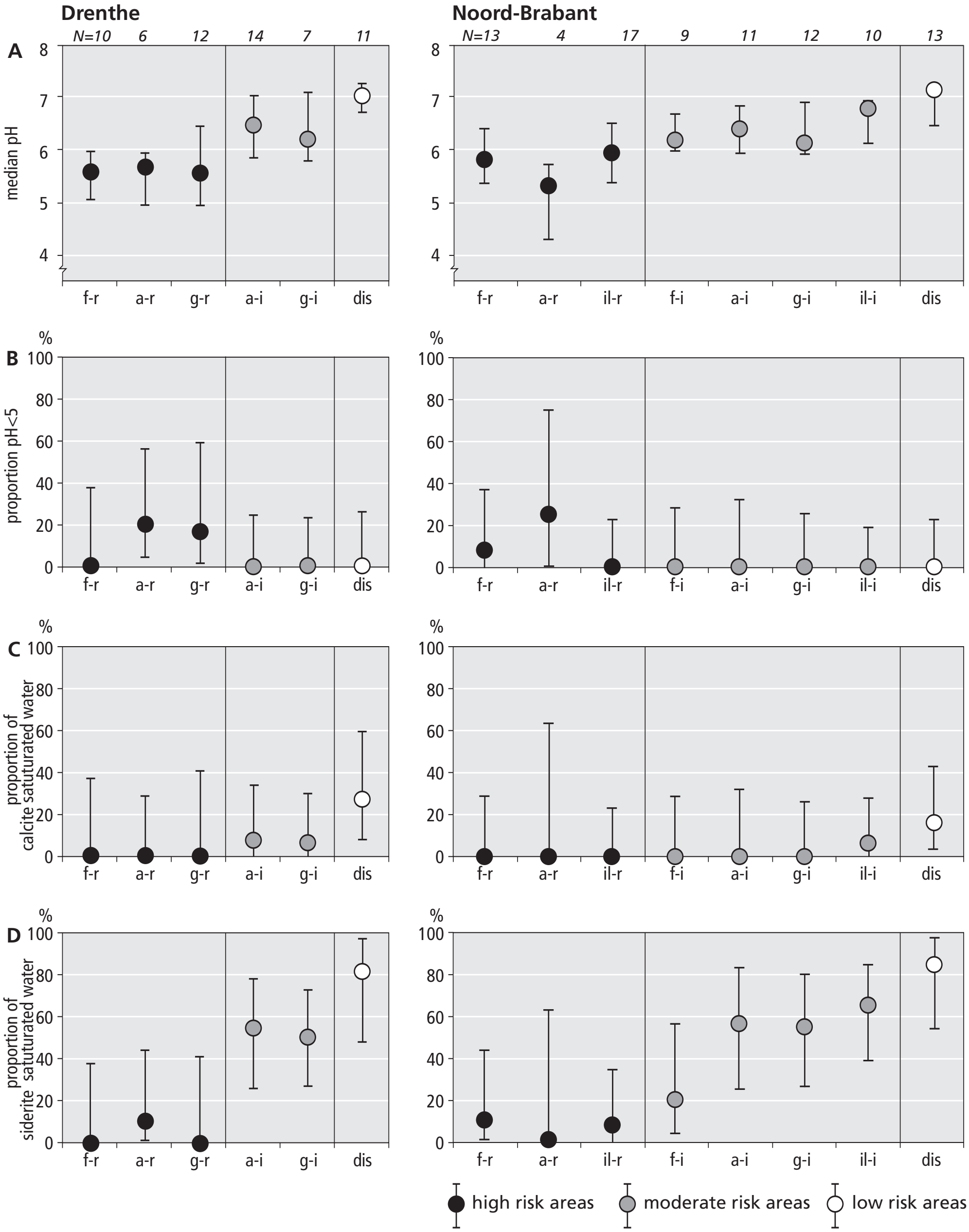


Figure 3.19 - Acidification of groundwater at 15-30 m depth in Drenthe and Noord-Brabant for 1995-1998: (a) median pH, (b) proportion of pH<5, (c) proportion of calcite saturated groundwater and (d) proportion of siderite saturated groundwater. Acronyms explained in Table 3.1.

by different chemical vulnerability of the subsoil, for example differences in cation exchange capacity (CEC). Groundwater age cannot explain the differences, because both homogeneous areas have similar proportions of post-1950 groundwater at shallow depth (Figure 3.10).

The chemical vulnerability for acidification is characterised by the acid-neutralizing capacity of the subsoil, including the cation exchange capacity. The strongest acidification buffer in Dutch subsoils is calcite dissolution, following:

$$CaCO_3 + H^+ \rightarrow Ca^{2+} + HCO_3^- \tag{3.7}$$

or, for calcite dissolution by carbonic acid:

$$CaCO_3 + H_2O + CO_2 \rightarrow Ca^{2+} + 2HCO_3^- \tag{3.8}$$

Dissolution of siderite ($FeCO_3$) can also control the pH of the groundwater. The reactions are similar to the above if Ca is replaced with Fe. The presence of siderite in the Dutch subsurface has been confirmed by geochemical studies (for example: Huisman 1998, Broers & Griffioen 1994, Broers et al. 1994). Hoogendoorn (1983) suggested that regional recharge areas are depleted in carbonate because of the continuous input of carbonic acid in infiltrating rainwater over hundreds to thousands of years (Appelo et al. 1982, Appelo & Postma 1993).

Subsurface carbonate dissolution was evaluated using the calcite and siderite saturation indices (details see Appendix I). Groundwater was considered saturated or supersaturated if the saturation index SI is larger than -0.2. Figures 3.18c+d and 3.19c+d present the proportion of groundwater that is saturated or supersaturated with respect to calcite and siderite. Calcite saturation did not appear frequently in Noord-Brabant and Drenthe, and only discharge areas exhibited substantial proportions of calcite-saturated groundwater (between 10% and 30%). Hence, calcite dissolution does not play a major role in the buffering of groundwater acidification at the depths considered. However, siderite saturation occurred regularly, especially in intermediate and discharge areas. Discharge areas showed > 75% siderite-saturated or supersaturated groundwater at both 5-15 and 15-30 m depth. This points to overall iron-reducing conditions and the presence of carbonate in the subsurface. Since groundwater flow is upward in those areas, carbonates are probably present below the current investigation depth of 5-30 m. In agriculturally used intermediate areas generally 40-60% of groundwater was saturated or supersaturated with respect to siderite in both provinces. In the Noord-Brabant and Drenthe recharge areas, carbonates seem to be absent in the first 30 m of the subsoil and do not control the pH. This conforms to the carbonate depletion theory of Hoogendoorn and explains the overall pH differences between recharge and intermediate areas that are used for agriculture.

The results show that agricultural recharge areas frequently have a substantial proportion of groundwater with pH < 5. Apparently, the large inputs of ammonia through manure and fertiliser are not neutralized by the dissolution of lime that is applied to the soils. This means that acidification and agricultural pollution are strongly interrelated in the Dutch agricultural areas. This finding is further illustrated in Figure 3.20, which shows the relations of OXC, total hardness and the hardness/alkalinity ratio with pH (details see Appendix I). The oxidation capacity showed a clear negative correlation with pH; high OXC concentrations were especially frequent below pH = 6. The correlation between pH and OXC demonstrates the acidifying effect of excess manure on groundwater.

The total hardness shows a more complicated pattern, with a concentration minimum at pH 5.5 to 6 for both provinces. Above pH = 6 the hardness has increased, which was explained by the dissolution of carbonates following reactions (3.7) and (3.8) as the main process. Increased hardness was also observed below pH = 6. Three effects explain these concentrations: (1) the leaching of inputs of Ca and Mg that are introduced by manure and fertiliser, (2) the

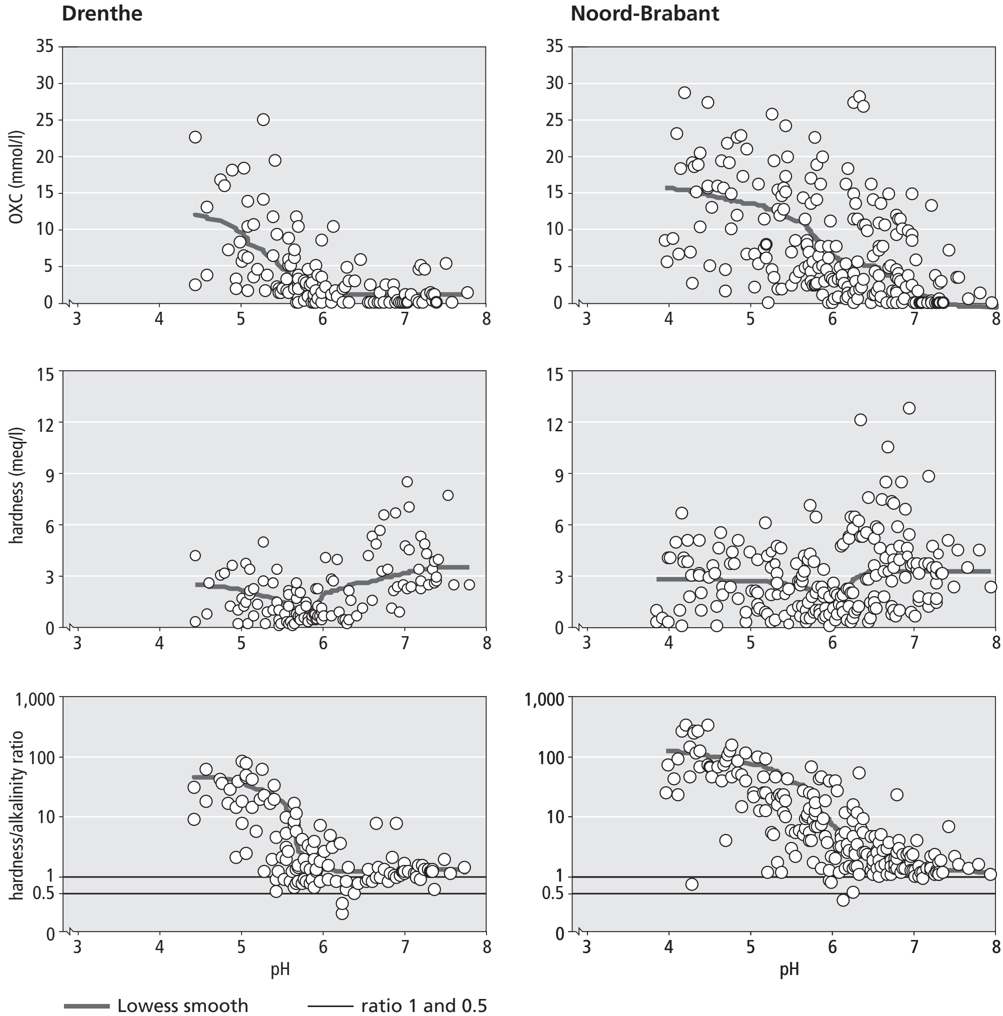


Figure 3.20 - Relations between OXC (a), total hardness (b) and hardness/alkalinity ratio (c) with pH for Drenthe and Noord-Brabant for 1995-1998. The black line gives a LOWESS smooth (f=0.75) which indicates the centre of the data scatter. A hardness/alkalinity ratio of 1.0 indicates the process of calcite dissolution with natural carbondioxyde.

dissolution of lime, and (3) the exchange of Ca and Mg against Al and protons at the acidification front.

The hardness/alkalinity molar ratio is a useful, alternative indicator for acidification because it combines the effects of pH and the increase of total hardness in one variable. The hardness/alkalinity ratio shows a strong correlation with pH because of the decrease of bicarbonate with pH (see for example Appelo & Postma 1993). When dissolution of Mg-calcite controls the pH and the concentrations of Ca, Mg and HCO_3 following the equations (3.7) and (3.8), the hardness/alkalinity ratio equals 0.5 or 1.0, respectively. Ratios of 0.5 to 1.0 were indeed frequently observed above pH = 6 for Drenthe and Noord-Brabant. This indicates the dissolution of carbonate as the major process (Figure 3.20). The strong increase of the

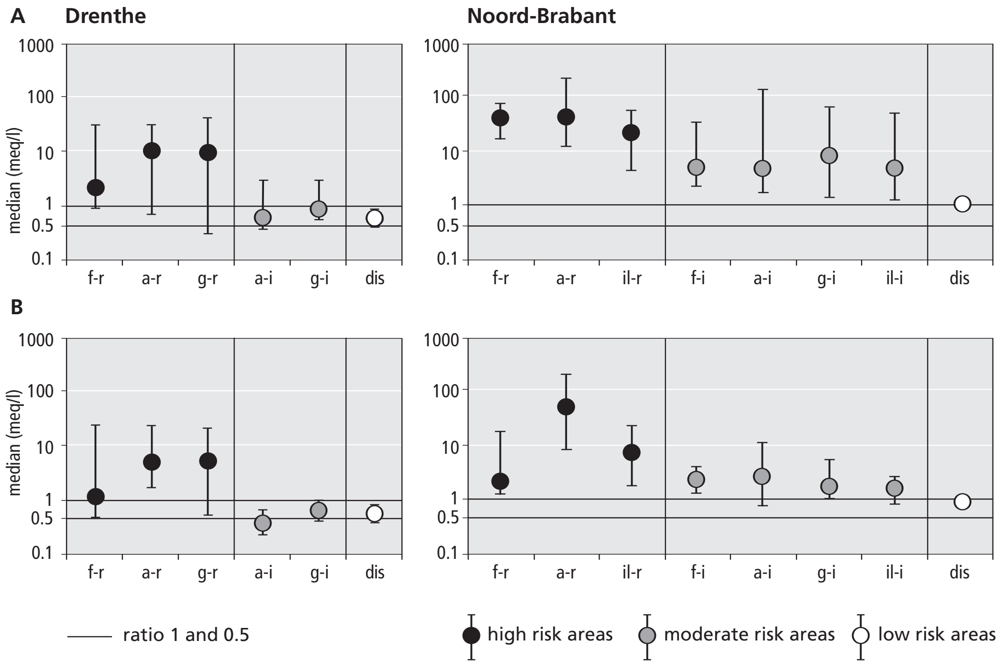


Figure 3.21 - Median hardness/alkalinity ratio for groundwater at 5-15 m depth (a) and 15-30 m depth (b) in Drenthe and Noord-Brabant for 1995-1998. Acronyms explained in Table 3.1.

hardness/alkalinity ratio below pH = 6 demonstrates the effects of extra leaching of Ca and Mg from manure and the release of these alkaline cations by ion-exchange with Al^{3+} and H^{+}.

Figure 3.21 shows the median values of the hardness/alkalinity ratio for the homogeneous areas in Drenthe and Noord-Brabant. In Drenthe, the occurrence of hardness/alkalinity ratios > 1 is most pronounced in the agriculturally used recharge areas, both in the shallow and the deeper screens. In Noord-Brabant, however, the forested recharge areas also showed hardness/alkalinity ratios > 1, as did the intermediate areas. The large hardness/alkalinity ratios in the Noord-Brabant forests and agricultural areas reflect the combined effects of large inputs of manure and acidifying substances in this province.

Concluding, acidification appears to be most pronounced in forest-recharge areas, closely followed by agriculturally used recharge areas. Almost no pH buffering by carbonate dissolution is present in these areas, even though liming is applied in the agriculturally used areas. The pH<5 front has not yet reached the deeper monitoring interval in the recharge areas because of pH-buffering by the sediments. However, median hardness/alkalinity ratios > 1 at this depth interval point to the first arrival of the combined agricultural pollution and acidification front in Drenthe recharge areas and Noord-Brabant recharge and intermediate areas.

General conclusions on the data analysis

Two main conclusions are drawn from the data analysis of the Noord-Brabant and Drenthe monitoring networks. First, a stratification using land use and geohydrology is effective in determining the magnitude of the effects of agricultural pollution and acidification at the regional scale. Using the stratification, similar contamination patterns were found for the two provinces that do agree with the a priori expected risks for groundwater contamination.

Second, subsurface reactivity causes another important source of variation in the data. This type of variation is partly reduced using secondary chemical properties that better describe geochemical processes than the measured individual components themselves. The reduction of variation helps to identify the human impact on the groundwater resources. Reactive processes are responsible for the main differences in the contamination patterns of Drenthe and Noord-Brabant. The two aspects will be discussed separately.

Effects of geohydrological stratification

The use of both land use and geohydrology for the stratification of the sandy areas in the two provinces revealed groundwater quality patterns that would not be identified using land use as the sole factor. Median concentrations of pollutants and the proportion of contaminated groundwater generally agreed well with the a priori assumed risk order that was based on geohydrology and land use. The geohydrological factor appeared to be crucial, because the proportion of post-1950 groundwater contrasts largely between recharge, intermediate and discharge areas. In both provinces the recharge areas exhibit large proportions of post-1950 groundwater and are most affected by agricultural pollution and acidification. These areas occupy only 5-16% of the total area in the two provinces. The shallow groundwater in the intermediate, drained areas is predominantly post-1950 and is also affected by agricultural pollution and acidification. However, a large proportion of pre-1950 groundwater and a limited proportion of contaminated groundwater was found at 15-30 m depth in those areas. Discharge areas are least affected by agricultural pollution and acidification. Using nitrate as an example, the proportion of contaminated groundwater at 5-15 m depth decreases from about 50% in the Noord-Brabant and Drenthe agricultural recharge areas, via 5 to 25 percent in agricultural intermediate areas to 0% in discharge areas.

Effects of reactive processes

Reactive processes in the subsurface have large effects on the groundwater contamination patterns, because they have a buffering effect. Redox processes, such as denitrification and pyrite oxidation, buffer nitrate transport. Acidification is buffered by carbonate dissolution or, in more acid conditions, by dissolution of aluminum hydroxides, cation-exchange, proton buffering and silicate weathering. The extent to which the redox and acid buffers are present in the Netherlands subsurface is a priori not known when designing a regional monitoring network. For example, data on organic matter and carbonate contents are only available for the uppermost 1.2 m of the soil. These shallow data are of little use in the prediction of buffering effects below the soil zone and for vulnerability mapping of threats to deeper groundwater. Unfortunately, reactive properties such as the reactivity of organic matter, pyrite and carbonate, cannot be linked unambiguously to geological formations, because many reactive properties have evolved or have been changed after deposition (for example Moura & Kroonenberg 1990). As a result, no reliable spatial information was available in the design stage of the networks that could be used for vulnerability mapping and stratification. Alternatively, the effects of reactivity should be dealt with during the data analysis of the groundwater quality data themselves. Geochemical calculations help to some extent to unravel hydrogeochemical processes from the monitoring data and to explain observed variations.

Using geochemical calculations, the data analysis of Noord-Brabant and Drenthe presents clear indications for reactive processes that changed the groundwater quality. In Noord-Brabant, reduction processes limited the depth of nitrate transport in the recharge and intermediate areas. Pyrite oxidation is at least partly responsible for the disappearance of nitrate. Using the indicator ‘oxidation capacity’, it was demonstrated that agricultural pollution does influence deeper groundwater below agricultural areas in Noord-Brabant as well. The

OXC concentrations in agricultural areas in Noord-Brabant are much larger than in Drenthe. and reflect the differences in manure input loads between the two provinces. Thus, OXC is a useful indicator for the monitoring of the effects of excessive manure loads. The OXC results furthermore indicate that sulphate concentrations in Noord-Brabant groundwater have already increased above drinking water standards. This shows that the reaction products of nitrate reduction have generated new groundwater contamination problems. The Drenthe recharge areas are much more vulnerable for leaching of nitrate to deeper groundwater, despite the lower inputs of manure. These results point to the absence of the redox buffer in the form of reactive organic matter or pyrite.

The use of the hardness/alkalinity ratio and the calcite and siderite saturation indices helps to identify the processes that determine the effects of acidification on the shallow and deeper groundwater. Buffering of acidification by calcite dissolution is generally not present in the Noord-Brabant and Drenthe recharge and intermediate areas. The hardness/alkalinity ratio revealed the combined effects of agricultural pollution and acidification in Drenthe recharge areas and Noord-Brabant recharge and intermediate areas. These effects are not apparent from the use of pH as the single indicator for acidification.

3.6 General conclusions

This study aimed at integrating information on groundwater age and hydrogeochemical processes in the design and data analysis of regional monitoring networks. The results show that the extra geohydrological stratification, based on hydrogeomorphological characteristics of areas, helps to reduce the variation in the strata and to detect general contamination patterns. Differences in groundwater age in recharge, intermediate and discharge areas explain a large part of the variation between and within the land-use/geohydrology strata. Hydrogeochemical processes explain another part of the variation. Extra indicators, based on geochemical knowledge, help the detection of the impact of these processes and the identification and understanding of contamination patterns.

4 Regional monitoring of agricultural pollution and acidification of groundwater in two Dutch provinces

2 - Evaluation and optimization of the networks

4.1 Introduction

In Chapter 3, the design and data analysis of regional groundwater monitoring networks was presented, using the provincial networks of the Dutch provinces of Drenthe and Noord-Brabant for illustration. The objective of the study presented in this chapter is to design a framework for the evaluation and optimization of regional groundwater quality monitoring networks, accounting for area-specific information goals for areas with low, moderate and high risks for the contamination of deeper groundwater. The use of the methodology is illustrated using the monitoring results of the province of Noord-Brabant and Drenthe, but is generally applicable in evaluating regional monitoring networks that focus on diffuse contaminants in groundwater.

The presented framework builds on the proposal for statistical information goals and evaluation criteria of Baggelaar & Van Beek (1997). They introduced several statistical information goals that are relevant for the provincial monitoring objectives and provided criteria for the evaluation and optimization of the existing networks. However, no differentiation of the statistical information goals has been made for high and low-risk areas in the provinces, although different sample sizes and monitoring network densities are present in them.

The criteria of Baggelaar & Van Beek are suitable in a stratified sampling strategy that aims to estimate characteristics of spatial units, such as median concentrations and proportions of contaminated groundwater in land-use/geohydrology strata. The presented approach differs from geostatistical approaches for optimization that are especially useful for other monitoring objectives, including spatial interpolation to areas that have not been measured (Loaigica 1989, Spruill & Candela 1990, Dixon & Chiswell 1996, Brus & de Gruijter 1996, Pebesma & de Kwaadsteniet 1997).

4.2 Methodology

For the Dutch regional monitoring networks, a risk concept of vulnerability and pollution loading has been used to differentiate the sample size and monitoring network density in land-use/soil-type/geohydrology strata (Chapter 3). These strata were called homogeneous areas to emphasize the expectation that the variation between the homogeneous areas is larger than the variation within them. In this chapter, the collected monitoring data in the homogeneous areas serve as the basis for the evaluation and optimization of the networks.

The underlying design strategy of the regional monitoring networks was to have higher ambitions for high-risk homogeneous areas than for low-risk areas. A high ambition means less desired uncertainty in the estimated characteristics of areas. The low-risk areas were sampled to serve as a reference for high-risk areas in order to achieve an overview of groundwater chemical status over the entire province. In evaluating the effectiveness of such a monitoring set up, these area-specific ambitions must be accounted for. The information collected during several monitoring years is preferably used to obtain knowledge about the statistical characteristics of targeted chemical components in the homogeneous areas. This way, the evaluation/optimization procedure builds on the results of the initial network, improving it step-by-step.

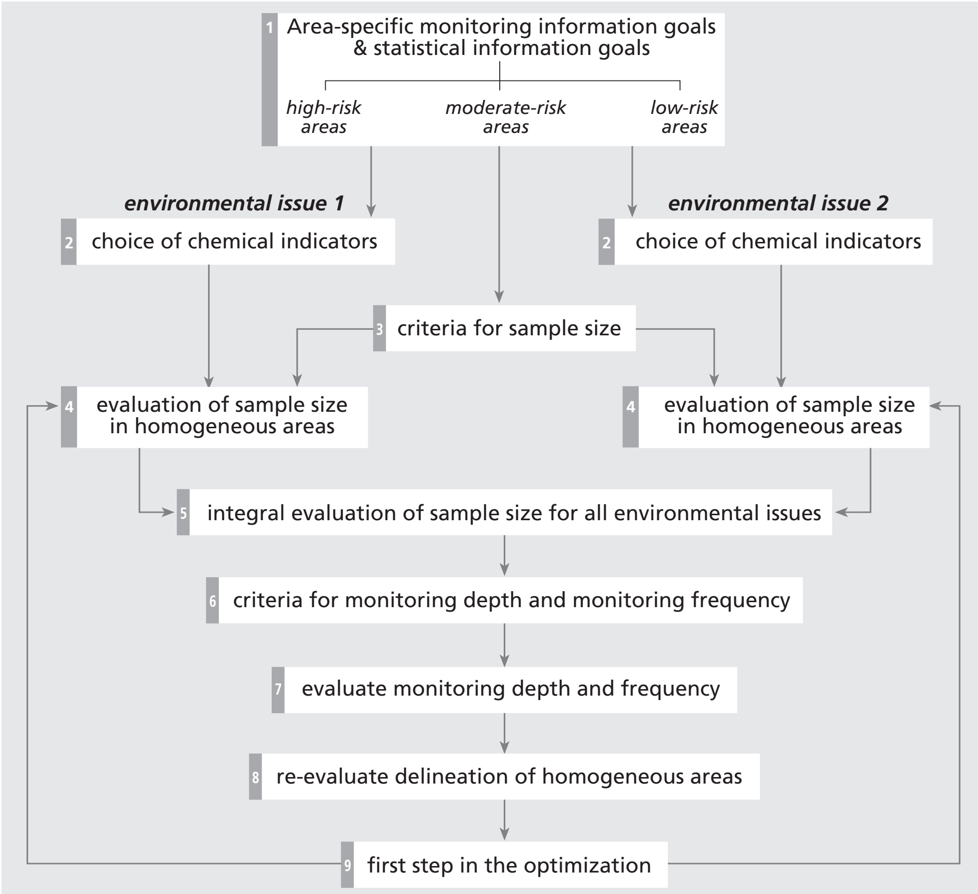


Figure 4.1 - The 9 steps of the framework for the evaluation and optimization of regional groundwater quality monitoring networks

The presented evaluation methodology includes the following steps (Figure 4.1, modified from Baggelaar and Van Beek 1997):

1. specify area-specific monitoring information goals and statistical information goals for areas with high, moderate and low risks for the contamination of deep groundwater
2. identify chemical indicators that suit the monitoring information goals
3. specify criteria to evaluate sample size in the homogeneous areas, based on the statistical information goals
4. evaluate the sample size within the existing network separately for each environmental issue
5. integrally evaluate sample size for the selected environmental issues
6. specify criteria for monitoring depth, monitoring frequency and sets of measured chemical components
7. evaluate monitoring frequency, monitoring depth and sets of measured chemical components in the existing network
8. re-evaluate the delineation of homogeneous areas if necessary
9. first step of the optimization

The subsequent steps are discussed below.

Step 1 Defining area-specific statistical information goals

Two types of general, policy-induced information goals are often present in regional groundwater quality monitoring:

1. the desire to monitor all areas or strata in a specific region to establish an overview of groundwater chemical status, and
2. the desire to effectively monitor targeted contaminants in areas with the greatest risk of contamination of deep groundwater resources, both in time and space.

In an attempt to address these two general information goals, many Dutch regional networks have been designed with relatively high monitoring density in high-risk areas and low monitoring density in low-risk areas (Chapter 3, Table 3.1). In line with these two general information goals, area-specific statistical information goals have been specified that reflect the different ambitions for homogeneous areas, according to their risk of the contamination of deep groundwater (Chapter 3, Table 3.4).

In the design stage of the networks, the statistical information goals C and E of Table 3.4 were called the *signal* function of the network. The *signal* function is especially relevant for the moderate-risk areas to detect contamination if present and to detect quality changes with time. The information goals A and B of Table 3.4 reflect the so-called *reference* function of the network. The *reference* function aims at an overview of the groundwater chemical status over the entire province, by providing typical concentrations of low-risk, moderate-risk and high-risk areas. The reference function does not aim at comparing recent with historical groundwater quality in a specific area. Statistical information goals D, F and G of Table 3.4 aim at the determination of the magnitude of contamination in high-risk areas, by providing information on the proportion of contaminated groundwater (D) and the magnitude of temporal trends (F and G, see Chapter 5).

Step 2 Choice of chemical indicators

In the Netherlands, several specific contaminants are distinguished in relation to a number of well-defined *environmental issues*. Relevant *environmental issues* for groundwater quality monitoring are the leaching of nutrients, including nitrate, the acidification of groundwater, the distribution of trace metals and arsenic and the leaching of pesticides (RIVM 2001). In this chapter, the evaluation concentrates on the leaching of nutrients from agricultural origin and on acidification. The choice of indicators differs for different provinces or areas, depending on the hydrogeochemical or hydrological situation (Baggelaar and Van Beek 1997). Here, the oxidation capacity (OXC, Appendix I) and nitrate were chosen to represent 'agricultural pollution' and pH to represent 'acidification' for the sandy provinces of Noord-Brabant and Drenthe. Especially in Noord-Brabant, the OXC was a better indicator of agricultural pollution than nitrate, because of subsurface geochemical reactions (Chapter 3). The pH was considered the most important indicator of acidification, because it also controls the mobility of other elements such as trace metals and aluminum (Dzombak & Morel 1990).

Step 3 Criteria for the evaluation of sample size

The approach uses available monitoring data and non-parametric methods to evaluate the sample size in the networks. Specific evaluation criteria were coupled to each of the statistical information goals of Table 3.4. The choice of the evaluation criteria and the corresponding threshold values is argued below. Table 4.1 lists the evaluation criteria for sample size in the homogeneous areas. The criteria correspond to the statistical information goals A to D in Table 3.4.

The evaluation of the sample size in the strata of the provincial networks was not based on a preliminary, theoretical relationship between sample size and predefined desired precision. Such relationships require assumptions about the frequency distribution of the population. For

Table 4.1 - Criteria for the evaluation of sample size in high-risk, intermediate-risk and low-risk areas

Criterion		High-risk areas	Moderate-risk areas	Low-risk areas
Ambition level		**High**	**Moderate**	**Low**
Typical concentrations				
A	Precision of the median	x	x	x
B	Significance of differences between high-risk and low-risk areas	x		x
Comparison with critical concentrations				
C	Probability of detection of contamination		x	
D	Precision of the proportion of contaminated groundwater	x		

example, Baggelaar & Van Beek (1997) estimated required sample sizes with predefined precision for the estimate of the median, using the log-normal distribution following Hale (1972). They realised that it is very uncertain that the data fulfil this assumption and also preferred non-parametric methods for data-analysis.

In the presented methodology no frequency distribution was assumed and non-parametric methods were used for the evaluation. The criteria are discussed below. Thresholds for the evaluation criteria are presented at step 4.

Criterion A

Groundwater quality data are rarely normally distributed and often exhibit a positively skewed frequency distribution. Therefore, the median was used to measure the central tendency of the concentrations in the homogeneous areas. The median robustly represents typical values in skewed data sets. The average concentration was not considered to be an alternative because it is more sensitive to outliers and extreme values that occur regularly in the groundwater quality data sets. The 75 percentile is another choice, but requires larger sample size for similar precision. The precision of the median was used to evaluate the sample size for the assessment of typical concentrations (statistical information goal A).

The median and the corresponding 95% confidence interval were calculated non-parametrically following Helsel and Hirsch (1992, see Appendix III.1). The 95% confidence interval is typically non-symmetrical because the frequency distributions of the samples are often positively skewed. The largest part of the 95% confidence interval on the estimated median is used to evaluate the sample size (Figure 4.2, see step 4).

Criterion B

The overall effectiveness of the sample size in the networks was evaluated by testing for differences between the homogeneous areas. A multiple-comparison test was used to test which homogeneous areas are significantly different at the 95% confidence level (Figure 4.2, see Appendix III.2). The procedure is similar to the one used by Nolan et al. (1997) except that ranks were used instead of a log-normal transformation of concentrations.

In the Dutch provincial monitoring networks, significant differences between homogeneous are all relevant for water quality managers, given the small number of samples that are available (approximately 8-20) and the amount of variation that was observed (Chapter 3). Therefore, a large number of differences between the homogeneous areas indicated a successful delineation of homogeneous areas, especially if high-risk, moderate-risk and low-risk areas were distinguished significantly. In general, the more differences between homogeneous areas that were detected, the better the networks performed for the 'reference function'.

Criterion C

Criterion C (Table 4.1) applies to the probability of not detecting a specific percentage of contaminated area within the homogeneous area using the current sample size (Figure 4.2). This criterion is especially important for the moderate-risk areas in order to signal the occurrence of contamination. If the sample size is small, it is likely that purely by chance no contamination is detected. The criterion is based on the conditional probability of encountering no samples exceeding the critical concentration under the condition that a known percentage p_{x_c} of the homogeneous area is above the critical concentration, using sample size n (Appendix III.3).

Criterion D

Statistical information goal D is relevant for high-risk areas to determine precisely the magnitude of groundwater contamination. The question was 'how well does the network perform when estimating the proportion of contaminated groundwater in the high-risk areas'. Appendix III.5 provides details on the calculation of the proportions and confidence intervals. The precision of the estimated proportion was used as the criterion for the evaluation (Figure 4.2). The precision is given by the two-sided non-parametric 95% confidence interval of the estimates, calculated following Gilbert (1987).

Step 4 Evaluation of sample size in the homogeneous areas

Criterion A

The 95% confidence interval on the estimated median was used to evaluate the sample size for statistical information goal A (Figure 4.2). The desired precision of the median must be determined separately for each indicator using environmentally relevant criteria. Table 4.2 lists the criteria that were used to evaluate criterion A for pH, nitrate and OXC.

For pH, the entire two-sided 95% confidence interval was used to evaluate this criterion. For nitrate and OXC, the network was judged effective if the largest half of the confidence interval on the median is within the specified concentrations. The thresholds of Table 4.2 were based on hydrochemical knowledge and relevance for water quality managers. Differences of 1 pH-unit (for example between pH=5.0 and pH=6.0) are largely relevant for geochemical processes, such as desorption, cation-exchange and redox transformations. The threshold values for nitrate were based on drinking water standards (50 mg l^{-1}) and the EU target value (25 mg l^{-1}). The network was judged to be effective if the largest half of the confidence interval on the

Table 4.2 - Judgement of monitoring effectiveness based on precision of medians for pH, nitrate, OXC

	Complete interval	Largest half part of 95% confidence interval	
	pH within (pH-unit)	Nitrate (mg l^{-1})	OXC (meq l^{-1})
Very effective	< 0.5	< 25	< 4
Effective	0.5-1.0	25- 50	4 - 7.5
Moderately effective	1.0-1.5	50-100	7.5-15
Not-effective	> 1.5	> 100	> 15

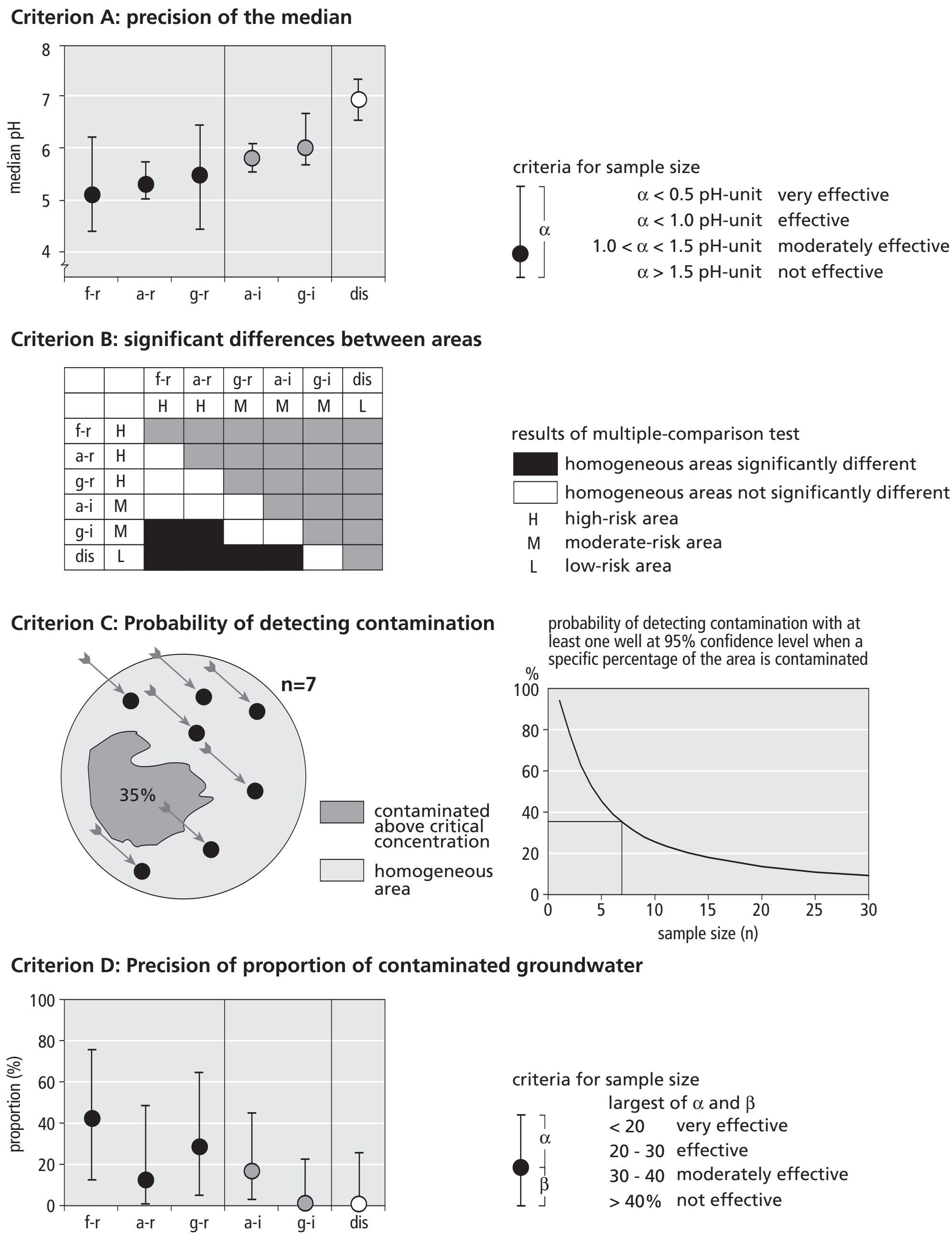


Figure 4.2 - Overview of the 4 criteria used to evaluate the sample size in the homogeneous areas. Abbreviations of the homogeneous areas are explained below Table 4.7

median is less than 50 mg NO_3 l^{-1}. Using this threshold, the probability that the population median is above the drinking water standard is less than 2.5% for a low-risk area with an estimated median equal to the detection limit. For OXC a threshold of 7.5 meq l^{-1} was chosen, which equals half of the critical concentration of 15.1 meq l^{-1} where either sulfate or nitrate is above the drinking water standard (see Chapter 3).

Criterion B

The overall effectiveness of the networks was evaluated by testing for relevant differences between high and low-risk areas (criterion B). A multiple-comparison test was used to test whether the medians of the homogeneous areas differ significantly at the 95% confidence level (Figure 4.2, Appendix III.2). The sample size in the monitoring network was considered effective if the high-risk and low-risk areas were distinguished significantly. This was a minimum requirement for the reference function. The risk assessment failed if no differences could be demonstrated between these two end-members in the risk assignment for all indicators of an environmental issue. In that case, major modifications to the monitoring set up are considered necessary.

Criterion C

For criterion C, the sample size was evaluated by the probability of detection of a specified percentage of contaminated area at the 95% confidence level (Figure 4.2, Appendix III.3). The Figure shows the required sample size to signal a specified percentage of contaminated area at the 95% confidence level. For example, 14 wells are required to detect at least one observation above the critical concentration, if 20% of the area is actually contaminated.

For statistical information goal C, the sample size was judged effective if the probability of not detecting excursions above the critical concentration is less than 5%, when 25% of the area is actually contaminated above this concentration. Table 4.3 explains the criteria used to evaluate the sample size for statistical information goal C. Following Figure 4.2 at least 11 observations are necessary to meet this criterion. A sample size of 11 wells was thus considered as a minimum for the signal function of the moderate-risk areas.

Criterion D

The precision of the estimated proportion $\hat{p}_{x_c}$ was used as the criterion for the evaluation (Figure 4.2). This precision depends on the sample size n and number of excursions above the critical concentration u (Appendix III.5). This is illustrated in Figure 4.3 which shows the 95% confidence interval as a function of estimated proportion ($\hat{p}_{x_c}$ = 0, 0.2, 0.5 and 0.8, respectively) and the sample size n. For $\hat{p}_{x_c}$ =0.5, the interval is symmetric. The curves become very flat above sample size n = 20, which indicates that a precision better than plus or minus 15% is only feasible by increasing the sample size to n = 48. The confidence interval is asymmetric for $\hat{p}_{x_c} \neq 0.5$. For $\hat{p}_{x_c}$ = 0, 0.2 and 0.8, the boundary farthest from the estimated percentage $\hat{p}_{x_c}$ decreases faster with sample size n than for $\hat{p}_{x_c}$ = 0.5.

The largest of the two distances between the estimate of $\hat{p}_{x_c}$ and the upper or lower boundary of the interval was used as a measure for the precision in the evaluation of sample

Table 4.3 - Criteria for the evaluation of the sample size for statistical information goal C: probability of contamination of excursions above the critical concentration

Percentage detected with 95% probability (%)	Effectiveness
0-10	very effective
10-25	effective
25-40	moderately effective
> 40	not effective

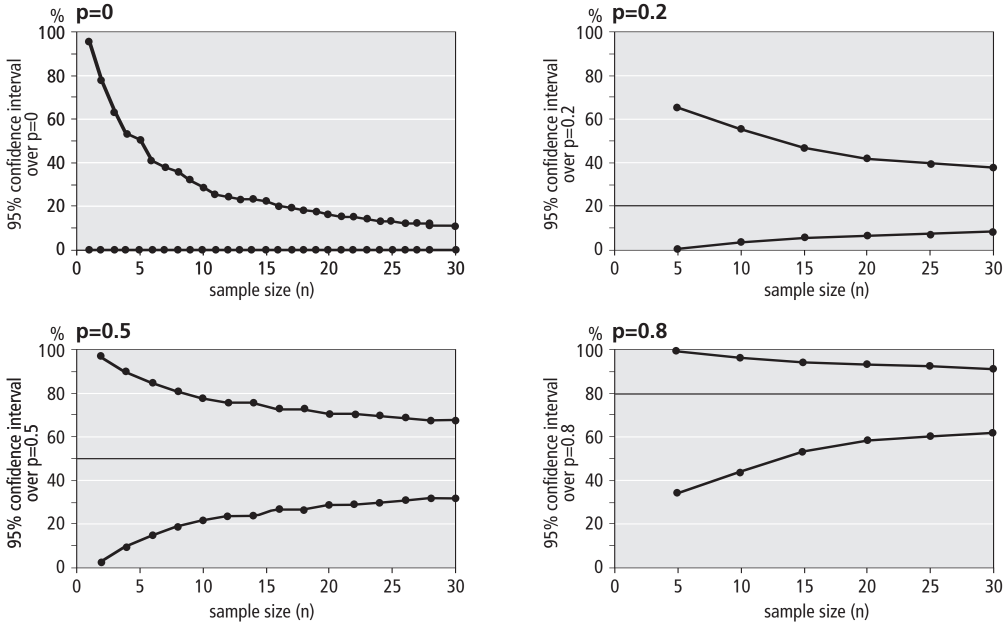


Figure 4.3 - Precision of the estimated proportion of contaminated groundwater as a function of sample size (n) for various values of the true proportion in the population.

size. Thus, for $\hat{p}_{x_c}$ = 0.2, the difference between the upper boundary and the estimate is used as the precision criterion. For $\hat{p}_{x_c}$ = 0.8, the difference between the lower boundary and the estimate is used (Figure 4.3).

Table 4.4 explains the criteria used to judge the effectiveness of the sample size. The sample size in a homogeneous area was judged effective if the precision of the estimated proportion is within 30%. (Figure 4.2). The criterion implies that a sample with an estimated $\hat{p}_{x_c}$ = 0.5 was significantly distinguished from the threshold value p_{x_c} = 0.2. The p_{x_c} = 0.2 threshold was considered relevant because it was used to map areas with a high degree of contamination (Chapter 3, Figure 3.9)

Step 5 Integral evaluation of sample size for the selected environmental issues
Step 4 was performed separately for the indicators of each relevant environmental issue (Figure 4.1). The evaluation results might contradict between different environmental issues or indicators. For example, using the criteria of step 4, the sample size in a homogeneous area

Table 4.4 - Criteria for the evaluation of the sample size for statistical information goal D: precision of the proportion of contaminated groundwater

Precision of proportion of groundwater exceeding the critical concentration (%)	**Effectiveness**
0-20	very effective
20-30	effective
30-40	moderately effective
> 40	not effective

might have been judged effective for acidification based on pH, but ineffective for agricultural pollution using nitrate or OXC. In step 5 the overall effectiveness of the sample size was evaluated according to the highest risk assignments of the two issues. In this way, the monitoring effort was focused on the high-risk areas of all relevant environmental issues.

Step 6 Specifying criteria for monitoring frequency, monitoring depths and sets of measured chemical components

In the evaluation methodology for the joint Dutch provincial monitoring networks, the monitoring depths and screen lengths were considered fixed. Changing these would require a completely new network set-up and high installation costs. However, more effective monitoring and cost reduction were achieved using an area-specific differentiation of monitoring frequency at the different monitoring depths.

The existing monitoring frequency, monitoring depths and screen lengths of the network are tuned to the average vertical flow velocities of recharging groundwater. The Dutch networks consist of screens of 2 m length at about 9 and 24 m depth that are sampled annually (for details see Chapter 3). Using a downward velocity of 1 m per year and a screen length of 2 m and assuming strictly horizontal inflow during sampling, the sample contains the water of approximately two years of recharge (Meinardi 1994). Since flow velocity decreases with depth (Chapter 2) and concentric flow will occur during sampling, one expects a sample to be a mix of about three or four years of recharge. Mixing 3 or 4 years recharge water averages out short term fluctuations in groundwater quality. Consequently, a yearly sampling frequency is sensible and gradual changes in the sampled groundwater are anticipated (Meinardi 1994, Baggelaar & van Beek 1997).

Following the monitoring information goals listed in Table 3.4 (Chapter 3), the detection of temporal changes in groundwater quality was considered relevant only for high and moderate-risk areas. Corresponding statistical information goals for high-risk, moderate-risk and low-risk areas are given in Table 4.5. For the moderate-risk areas, the signalling of trends was sufficient, whereas higher ambitions are pursued for high-risk areas. Therefore, a differentiation of monitoring frequency according to monitoring ambitions is quite possible.

The information about typical concentrations and proportions of contaminated ground-water is required in reports on the state of the environment in order to provide an overview of groundwater chemical status. These reports have a typical publication frequency of about once every 4 years. Therefore, using a monitoring frequency of once every four years was considered sufficient to meet the monitoring information goals A to D of Table 3.4. The detection of changes of groundwater quality with time (statistical information goals E, F and G of Table 4.5) requires a higher, annual monitoring frequency (see below).

Table 4.5 - Criteria for the evaluation of monitoring frequency in high, intermediate and low-risk areas

	Monitoring information goal	High-risk areas	Moderate-risk areas	Low-risk areas
E	Probability of detection of temporal trends		x	
F	Precision of median temporal trend	x		
G	Precision of trend in proportion of contaminated groundwater	x		

Criterion E

Criterion E refers to the probability of detection of temporal trends in *individual wells*. Here, a trend was defined as a change in groundwater quality over a specific period in time which is related to land use or water quality management (Loftis 1996). The probability of detecting such a trend depends on a large number of unknowns: the type of trend (monotonic, step trend), the magnitude of the trend, the natural temporal variations of the observations, the monitoring period and the type of statistical trend test used (Loftis 1996, Baggelaar & Van Beek 1997, Burn & Hag Elnur 2002, Yue et al. 2002).

In general, trend detection is more difficult if the trend is small relative to the natural temporal variations and the number of independent observations is small. In trend analysis of groundwater quality data a monotonic trend type is often assumed (for example: Pebesma 1996, Frapporti 1993, Reijnders et al. 1998, Baggelaar & Van Beek 1997). Because the probability distribution of the concentrations is unknown, Baggelaar and Van Beek (1997) recommended the use of non-parametric tests for the trend analysis. However, they used parametric methods to evaluate the probability of detection using examples from the Dutch regional networks. Using linear regression and the assumption of log-normal distributed data in a time series, they showed that the probability of detection decreases strongly when the number of observations decreases from 9 to 5 in a ten-year monitoring period. This corresponds with the experience that trend detection in individual wells using non-parametric methods has never shown significant trends for annual time series with less than 7 years (Broers 1996, see also Chapter 5). Therefore, time series of at least 10 observations were considered the minimum for significant trend detection, which requires 10 years of monitoring in the present set-up. Accordingly, decreasing the monitoring frequency from once every year to once every two years is no option if the aim is to detect trends over a 10 year monitoring period.

Increasing the monitoring frequency was also not considered sensible, because serial correlation between the observations will increase (Loftis 1996, Baggelaar & Van Beek 1997). Already with a monitoring frequency of once every year, substantial serial correlation is expected because of the slow downward velocity of groundwater relative to the length of the well screens (maximum 1 m per year versus 2 m screen length).

Criteria F and G

The criteria F and G of Table 4.5 refer to temporal trends in *homogeneous areas*. The trend definition of Loftis (1996) was adopted, which defines a temporal trend as a change in groundwater quality over a specific period in time, over a given region, which is related to land use or water quality management. Note that this trend definition is different from that commonly used in geostatistics. The probability of detecting trends in areas is dependent on the five factors mentioned under criterion E, plus the spatial variability in the area under consideration.

The statistical information goals F and G were considered relevant only for the high-risk areas for which the aim is to determine the median trend and the trend in the proportion of contaminated groundwater in the homogeneous areas. These information goals strongly depend on the precision of the estimates of the median concentrations and the proportions of contaminated groundwater in the homogeneous areas. These were evaluated using criteria A and

Table 4.6 - Differentiation of monitoring frequency into once every year and once every 4 years

Risk assignment	**High**	**Moderate**	**Low**
Ambition level	**High**	**Moderate**	**Low**
Frequency shallow screens	1 yr^{-1}	1 yr^{-1}	1 $(4\ yr)^{-1}$
Frequency deep screens	1 yr^{-1}	1 $(4\ yr)^{-1}$	1 $(4\ yr)^{-1}$

D, which were considered required first steps in the evaluation of the networks. In Chapter 5, examples of trend detection are given for two homogeneous areas in Noord-Brabant.

Step 7 Evaluating monitoring frequency, monitoring depths and sets of chemical components to be measured

Trend detection is relevant only for areas where changes in groundwater quality are expected that result from recent changes in land use or input of solutes. According to this argumentation, a differentiation in monitoring frequency is presented in Table 4.6. Annual monitoring is done only in the high-risk areas (statistical information goals F and G) and the shallow screens of the moderate-risk areas (statistical information goal E).

Further opportunities to increase the monitoring efficiency include the differentiation of the sets of measured chemical components for the 1 (4 yr)$^{-1}$ sampling rounds and 1 yr^{-1} sampling rounds. For example, the annual sampling could focus solely on chemical indicators that are directly relevant for environmental policy such as nitrate and pH. Although some cost-reduction is feasible, two disadvantages are present:

1. trends in chemical components that are not targeted at this moment will not be signalled, and
2. opportunities for quality assurance decrease.

Opportunities for quality assurance are important, since an evaluation of monitoring data in the Noord-Brabant network revealed large deviations from electro-neutrality in some monitoring years (Figure 4.4). Here, the electro-neutrality was evaluated using the cation surplus as:

$$Cation\ surplus = \frac{\Sigma cations - \Sigma anions}{\Sigma cations + \Sigma anions} \qquad (cation\ and\ anions\ in\ meq\ l^{-1}) \tag{4.1}$$

A cation surplus of 0% indicates electro-neutrality. The kind of quality check used for Figure 4.4 is no longer possible if not all major chemical species are measured in the annual sampling rounds. Quality checks are especially important in the Dutch regional networks because the national and provincial wells are sampled and analysed by different contractors and laboratories. Consequently, analysing all major cations and anions was advised to maintain the quality assurance possibilities.

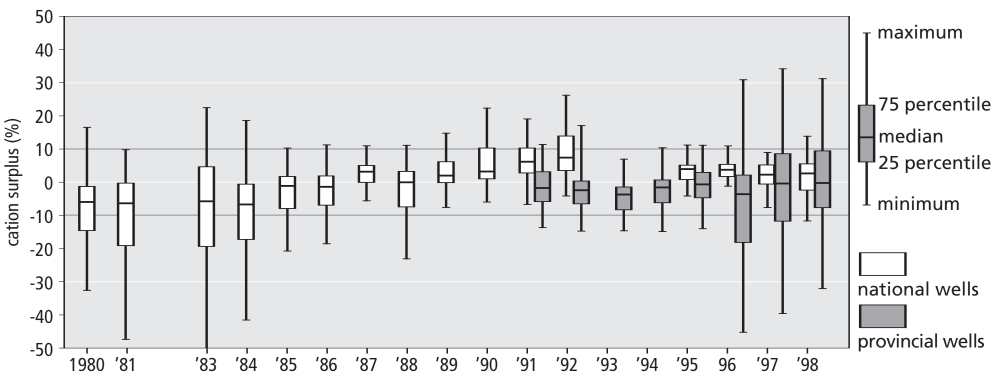


Figure 4.4 - Evaluation of electro-neutrality of the groundwater samples from the national and provincial monitoring wells in Noord-Brabant. Changes in laboratories are recognised from the jump in the size of the boxes (1984/1985 national wells, 1995/1996 provincial wells)

Step 8 Re-evaluating the delineation of homogeneous areas
In steps 1 to 7, the sample size and monitoring frequency were evaluated for the homogeneous areas. The contaminant concentrations in the homogeneous were assumed to be statistically stationary (no spatial trend). The delineation of homogeneous areas should be changed if indications exist for important regional spatial trends within a homogeneous area. For example, a spatial trend in concentrations could be due to a spatial trend in the pollution loading within a land-use type. Therefore, spatial trends are assessed in step 8 by checking the correlation of concentrations with spatial co-ordinates or maps. Changes in the policy-induced, general information goals would be another reason for changing the delineation of homogeneous areas. Examples are given in the application of section 4.3.

Step 9 First step of the optimization
When the delineation of homogeneous areas is accepted, recommendations for optimized sample size, monitoring frequencies and sets of measured chemical components are to be made for the first step in the optimization. The optimization aims at:
1. improving the precision of the estimates in areas that are monitored ineffectively
2. reduction of costs in areas that are monitored effectively.

Aim 1 requires adaptations to the sample sizes in the homogeneous areas. Aim 2 is best accomplished by area-specific reduction of monitoring frequencies as was described in step 7. Sample size reduction would be another option to accomplish aim 2, but is not sensible for the Dutch regional monitoring given the small sample sizes in the homogeneous areas and the large installation costs of the existing network.

The required sample size to obtain an effective network was derived from the relations between sample size and desired precision for the criteria C and D (Figures 4.2 and 4.3). For criterion D, the assumption was made that the proportion in the sample $\hat{p}_{x_c}$ is a good estimate of the true proportion in the population. No specific relation between sample size and desired precision of the median concentration is available for criterion A, without making assumptions about the frequency distribution. The same applies for criterion B: no a priori sample size is known that guarantees significant differences between areas at the 95% significance level. A sequential approach is used for the criteria A and B in which sample size is increased in several steps and the increase in precision is evaluated between the steps. Thus, after the first optimization round, a new evaluation round is recommended to judge the effectiveness of the improved networks and make further adaptations if necessary.

Table 4.7 - Criteria used for evaluation of sample size in the homogeneous areas

	Drenthe						Noord-Brabant							
Homogeneous area	f-r	a-r	g-r	a-i	g-i	dis	f-r	a-r	il-r	f-i	a-i	g-i	il-i	dis
Risk assignment	H	H	H	M	M	L	H	H	H	M	M	M	M	L
Ambition level	H	H	H	M	M	L	H	H	H	M	M	M	M	L
A precision median	X	X	X	X	X	X	X	X	X	X	X	X	X	X
B significant differences	X	X	X			X	X	X	X					X
C %area not detected at α=0.05				X	X					X	X	X	X	
D precision proportion (%)	X	X	X				X	X	X					

f-r	forest-recharge	a-i	arable- intermediate	H	high ambition/risk
a-r	arable-recharge	g-i	grassland-intermediate	L	low ambition/risk
g-r	grassland recharge	il-i	intensive livestock farming-intermediate	M	moderate ambition/risk
il-r	intensive livestock farming-recharge	f-i	forest-intermediate	X	criterion used for evaluation
dis	discharge areas				

4.3 Application to the Noord-Brabant and Drenthe monitoring networks

The 1995-1998 monitoring data from the networks of Drenthe and Noord-Brabant are now used to illustrate the evaluation procedure and to discuss opportunities for optimization. The general design of the two networks was described in Chapter 3. Table 3.1 summarizes the areal extent, the sample size, the network density and the relative risk assignments for the homogeneous areas in the sandy regions of the provinces. Evaluation of the network in the peat and clay areas is discussed elsewhere (Van Vliet 2000, Broers & Peeters 2000).

Step 1 Area-specific information goals
The application focuses on the evaluation of sample size in the homogeneous areas using criteria A to D (Table 4.1). Evaluation of the criteria F and G of Table 4.5 was considered to be a second step in the evaluation/optimization procedure, and is only sensible if criteria A to D are met effectively. Table 4.7 shows the criteria that were used to evaluate sample size in the individual homogeneous areas, based on the area-specific monitoring information goals.

Steps 2 to 4 Evaluation of sample size for separate environmental issues
Following the methodology described in section 4.2, steps 2 to 4 were performed separately for the each of the environmental issues. Subsequently, the issues acidification and agricultural pollution are discussed. These results are evaluated integrally for both environmental issues in step 5.

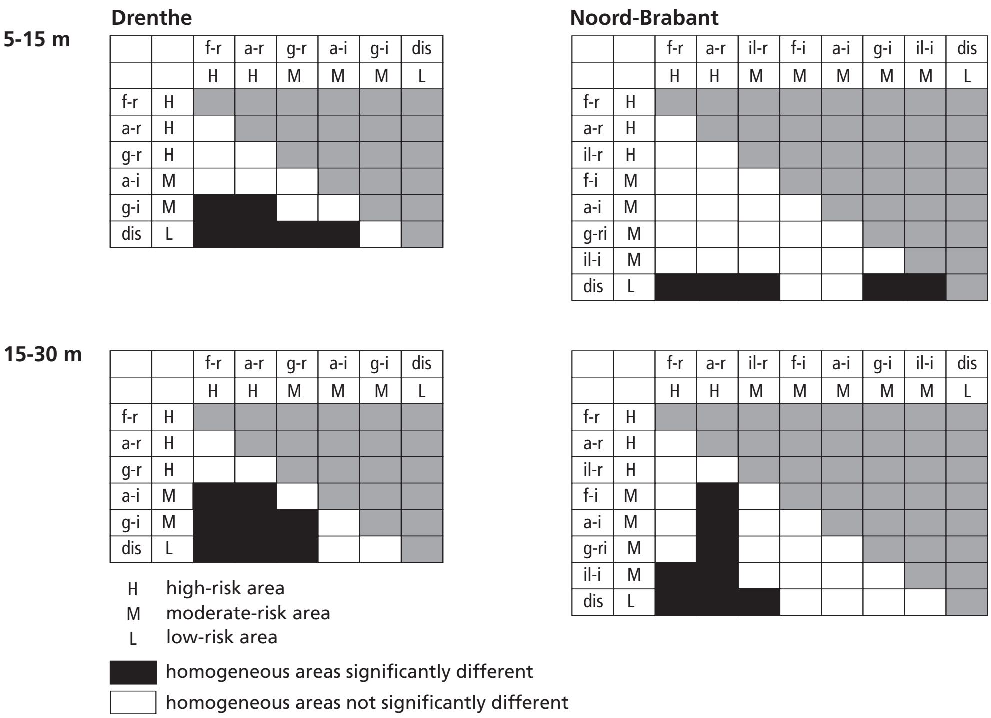


Figure 4.5 - Results of the multiple-comparison test for pH in Drenthe and Noord-Brabant. Black squares indicate that the median concentrations of the two homogeneous areas differ significantly (p<0.05)

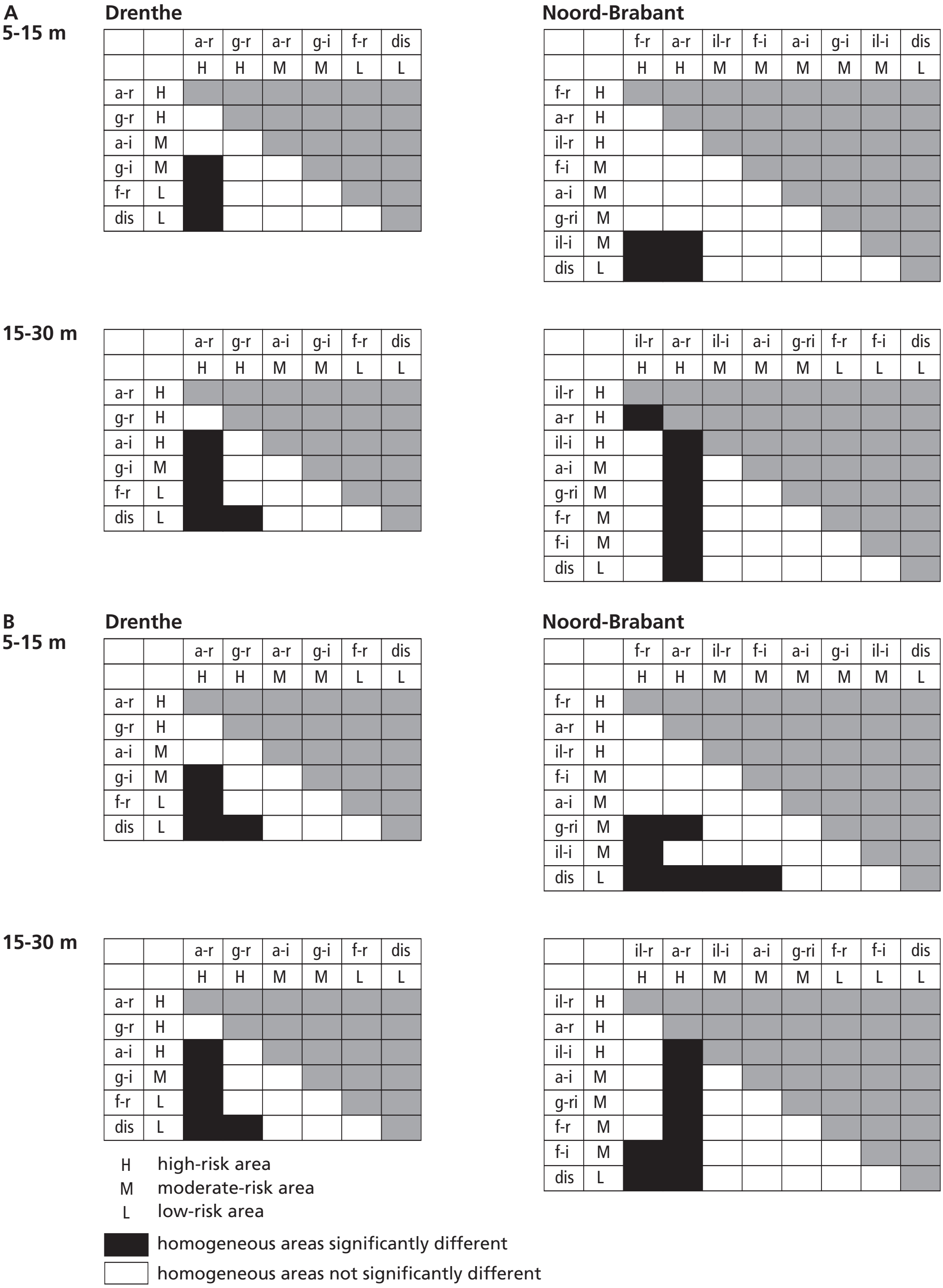


Figure 4.6 - Results of the multiple-comparison test for nitrate (a) and OXC (b) in Drenthe and Noord-Brabant. Black squares indicate that the median concentrations of the two homogeneous areas differ significantly (p<0.05)

Table 4.8 - Results of the evaluation of sample size for pH

	Drenthe						Noord-Brabant							
Homogeneous area	f-r	a-r	g-r	a-i	g-i	dis	f-r	a-r	il-r	f-i	a-i	g-i	il-i	dis
Risk assignment	H	H	H	M	M	L	H	H	H	M	M	M	M	L
Ambition level	H	H	H	M	M	L	H	H	H	M	M	M	M	L
Shallow screens														
Sample size	*7*	*8*	*7*	*12*	*14*	*11*	*11*	*4*	*14*	*11*	*8*	*8*	*20*	*12*
A Precision median	1.8	0.7	2.0	0.6	1.0	0.8	1.4	1.5	0.4	0.9	1.1	1.8	1.8	0.7
B Significant differences (%)	100	100	100	-	-	100	100	100	100	-	-	-	-	100
C % area not detected at α=0.05	-	-	-	22	19	-	-	-	-	24	31	31	13	-
D Precision proportion (%)	34	38	37	-	-	-	29	40	27	-	-	-	-	-
Deep screens														
Sample size	*7*	*10*	*6*	*12*	*14*	*11*	*12*	*4*	*13*	*10*	*9*	*11*	*17*	*13*
A Precision median	0.9	1.0	1.5	1.2	1.3	0.6	0.8	1.4	1.2	0.7	0.9	1.0	0.8	0.7
B Significant differences	100	100	100	-	-	100	100	100	100	-	-	-	-	100
C % area not detected at α=0.05	-	-	-	22	19	-	-	-	-	26	28	24	16	-
D Precision proportion (%)	38	36	42	-	-	-	28	50	23	-	-	-	-	-

f-r forest-recharge
a-r arable-recharge
g-r grassland recharge
il-r intensive livestock farming-recharge
dis discharge areas

a-i arable- intermediate
g-i grassland-intermediate
il-i intensive livestock farming-intermediate
f-i forest-intermediate

H high ambition/risk
L low ambition/risk
M moderate ambition/risk
\- criterion not relevant for ambition level

effective; moderately effective; not effective

Acidification
The pH was chosen as indicator to evaluate the sample size for the acidification issue . Table 4.8 gives an overview of the evaluation of sample size using the criteria A-D for the two monitoring depths.

The methodology was explained Figure 4.2. The evaluation of the criteria A and D was based on the Figures 3.18 and 3.19, which show the median pH and the proportion of groundwater with pH<5 including their 95% confidence intervals (Chapter 3). Tables 4.2, 4.3 and 4.4 show the threshold values used for the criteria A, C and D. Evaluation results for criterion B were based on the results of the multiple-comparison test that are shown in Figure 4.5. A 100% score for criterion B was assigned when all combinations of high-risk and low-risk areas were significant. The results for criterion C were evaluated by comparing the number of observations in Table 4.8 with Figure 4.2.

Table 4.8 shows that sample sizes in both existing monitoring networks were effective for *low-risk* discharge *areas*. The median pH was estimated within 1.0 pH unit for shallow and deep screens. Moreover, differences between the high-risk recharge areas and low-risk discharge areas were all significant for pH (Figure 4.5). The relatively small pH variation in the discharge areas and the large difference between high and low-risk areas allowed for this result.

Figure 4.5 shows the results of the multiple-comparison test for pH. Especially for Drenthe, a relatively large percentage of significant differences was observed; 40 and 53% of all possible combinations were significant for shallow and deep screens, respectively. This result indicates that sample sizes in the networks were effective in fulfilling the desired reference function.

However, the networks in the individual *high-risk areas* were less effective. In only one high-

Table 4.9 - Results of the evaluation of sample size for nitrate

	Drenthe						Noord-Brabant							
Homogeneous area	a-r	g-r	a-i	g-i	f-r	dis	il-r	a-r	il-i	a-i	g-i	f-ri	f-i	dis
Risk assignment	H	H	M	M	L	L	H	H	M	M	M	L	L	L
Ambition level	H	H	M	M	L	L	H	H	M	M	M	L	L	L
Shallow screens														
Sample size	*9*	*7*	*12*	*14*	*7*	*11*	*14*	*4*	*20*	*8*	*8*	*11*	*11*	*12*
A Precision median	95	199	88	0.2	17	0.1	180	74	32	93	46	25	3	0.4
B Significant differences (%)	100	0	-	-	50	50	67	67	-	-	-	0	100	100
C % area not detected at α=0.05	-	-	22	19	35	-	-	-	13	31	31	-	-	-
D Precision proportion (%)	31	30	-	-	-	-	25	60	-	-	-	-	-	-
Deep screens														
Sample size	*10*	*6*	*12*	*14*	*7*	*11*	*13*	*4*	*17*	*9*	*11*	*12*	*10*	*13*
A Precision median	67	95	0.5	0.1	24	0.2	0.4	161	0.3	0.3	0.2	2	0.2	0.2
B Significant differences (%)	100	50	-	-	50	100	0	100	-	-	-	50	50	50
C % area not detected at α=0.05	-	-	22	19	35	-	-	-	16	28	24	-	-	-
D Precision proportion (%)	31	39	-	-	-	-	23	50	-	-	-	-	-	-

risk area was the median pH estimated precisely enough at both shallow and deep level using the current sample size. The sample size was insufficient to effectively detect the proportion of groundwater with pH<5 within 30% in the high-risk areas in Drenthe (Figures 3.18b and 3.19b). In Noord-Brabant, two high-risk areas met this criterion: *f-r* and *il-r*. The small number of wells in the *a-r* area did not allow for precise determination of the proportion of acidified groundwater. Overall, only one high-risk area was monitored effectively: *il-r* in Noord-Brabant.

The network in the *moderate-risk areas* aimed at assessing typical concentrations and signalling acidification of groundwater. At shallow depth in Noord-Brabant, three homogeneous areas were not effective for estimating the median pH, even though a relatively large sample size was present. For example, in the area *il-i*, 20 wells were not sufficient to estimate the median within 1.5 pH-units. In general, the sample size in the moderate-risk areas was sufficient to detect contamination if the actual contaminated percentage is above 25% following criterion C. Signals for acidification of shallow groundwater were found for *g-i* and *il-i* in Noord-Brabant. The variations of pH were the largest in these homogeneous areas.

In summary, for acidification the networks fulfil their reference function and a clear contamination pattern was found with increasing pH from high-risk areas to low-risk discharge areas (Chapter 3). However, evaluating the individual areas, in three *moderate-risk areas* typical concentrations were not assessed precisely and in four *high-risk areas* the sample size was insufficient for estimating the proportion of contaminated groundwater within 30% precision.

Agricultural pollution

Nitrate and OXC were chosen as indicators for agricultural pollution. Tables 4.9 and 4.10 list the evaluation results for nitrate and OXC, respectively. Estimates of the medians and proportions and the corresponding confidence interval were presented in Figures 3.11, 3.12, 3.16 and 3.17 (Chapter 3).

For the *low-risk areas*, the median concentrations of OXC and nitrate were estimated precisely enough, except for OXC in discharge areas in Noord-Brabant (Tables 4.9 and 4.10 and Figures 3.11, 3.12, 3.16 and 3.17). The results of the multiple-comparison tests for OXC and nitrate are shown in Figure 4.6. The number of significant differences was larger for OXC than

Table 4.10 - Results of the evaluation of sample size for OXC

	Drenthe						Noord-Brabant							
Homogeneous area	a-r	g-r	a-i	g-i	f-r	dis	il-r	a-r	il-i	a-i	g-i	f-ri	f-i	dis
Risk assignment	H	H	M	M	L	L	H	H	M	M	M	L	L	L
Ambition level	H	H	M	M	L	L	H	H	M	M	M	L	L	L
Shallow screens														
Sample size	*9*	*7*	*12*	*14*	*7*	*11*	*14*	*4*	*20*	*8*	*8*	*11*	*11*	*12*
A Precision median	12	19	9.0	2.8	1.6	0.9	8.7	8.5	5.4	11	10	3.4	5.1	10
B Significant differences (%)	100	50	-	-	50	100	100	67	-	-	-	100	50	100
C % area not detected at α=0.05	-	-	22	19			-	-	13	31	31	-	-	-
D Precision proportion (%)	34	37	-	-	-	-	26	50	-	-	-	-	-	-
Deep screens														
Sample size	*10*	*6*	*12*	*14*	*7*	*11*	*13*	*4*	*17*	*9*	*11*	*12*	*10*	*13*
A Precision median	6.2	8.3	4.4	1.7	4.3	1.0	7.1	8.0	3.5	8.2	6.9	1.3	2.2	1.5
B Significant differences (%)	100	50	-	-	50	100	67	100				50	100	100
C % area not detected at α=0.05	-	-	22	19	-	-	-	-	16	28	24	-	-	-
D Precision proportion (%)	34	41	-	-	-	-	28	40	-	-	-	-	-	-

for nitrate because OXC concentrations showed less variation within the individual homogeneous areas (see also Chapter 3). Differences between the high-risk areas and the low-risk discharge areas were all significant for OXC (Figure 4.6). Concluding, for OXC the overall sample size of the network was regarded effective for the reference function.

However, sample size in the *high-risk areas* was not sufficient to estimate median concentrations or proportions of contaminated groundwater precisely. This was especially true for the shallow screens, where variations in nitrate and OXC concentrations were large (Figures 3.11 and 3.16, Table 4.9 and 4.10). The network in the area *il-r* in Noord-Brabant performed best in estimating the proportion of agriculturally polluted groundwater. Even in this area, the median concentrations of nitrate and OXC were only estimated moderately effective for the shallow screens. Sample sizes in the homogeneous areas *a-r* and *g-r* in Drenthe and *a-r* in Noord-Brabant were too small to effectively estimate medians and proportions.

Two *moderate-risk areas* performed well in the estimation of typical concentrations and the signalling of contamination: *g-i* in Drenthe and *il-i* in Noord-Brabant (Tables 4.9 and 4.10). The relatively large sample size (14 versus 20) apparently paid off in these two areas. In the shallow groundwater, signals for groundwater contamination with nitrate or OXC were present in *a-i* (Drenthe) and *il-i*, *a-i* and *g-i* (Noord-Brabant). From these homogeneous areas only the first two are monitored effectively. At greater depths, the OXC concentrations were lower and showed less variation. As a result, the networks performed much better at greater depth. Therefore, optimization of the network should first focus on the shallow monitoring level.

Step 5 Integral evaluation of environmental issues

In step 4, the sample size was evaluated separately for the two environmental issues acidification and agricultural pollution. Here, the results have been integrated for the two environmental issues, using the indicators pH, nitrate, OXC (Table 4.11). For some homogeneous areas, different ambitions for the two separate issues were assigned. For example, high ambitions were assigned to *f-r* for the monitoring of acidification, against low ambitions for the monitoring of agricultural pollution. This resulted in different evaluation results for the two issues in the same area.

For the integral evaluation of the two issues, the homogeneous areas were sorted in descending order following sample size (Table 4.11). An area was considered to be sampled

Table 4.11 - Integral evaluation for the two environmental issues considered and a proposal for the first step of optimization indicating sample size and monitoring frequency

	Drenthe						Noord-Brabant							
Homogeneous area	g-i	a-i	dis	a-r	g-r	f-r	il-i	il-r	dis	f-r	f-i	a-i	g-i	a-r
Sample size	14	12	11	9	7	7	20	14	12	11	11	8	8	4
Integral evaluation														
Ambition agricultural pollution	M	M	L	H	H	L	M	H	L	L	L	M	M	H
Nitrate	e	e	e	me	ne	me	e	ne	e	ne	e	me	e	ne
OXC	e	e	e	ne	ne	me	e	e	e	e	e	me	me	ne
Ambition acidification	M	M	L	H	H	H	M	H	L	H	M	M	M	H
pH	e	e	e	me	ne	ne	me	e	e	e	e	me	ne	ne
First step optimisation														
Ambition agricultural pollution	M	M	L	H	H	L	<u>H</u>	H	L	L	L	M	M	H
Ambition acidification	M	M	L	H	H	H	M	H	L	H	M	M	M	H
Optimisation necessary	no	no	no	yes	yes	yes	?	?	no	no	no	yes	yes	yes
Proposed optimised sample size (first step)				12	12	12	25	18				11	11	4
Monitoring frequency shallow	1/1	1/1	1/4	1/1	1/1	1/1	1/1	1/1	1/4	1/1	1/1	1/1	1/1	1/1
Monitoring frequency deep	1/4	1/4	1/4	1/1	1/1	1/1	1/1	1/1	1/4	1/1	1/4	1/4	1/4	1/1

H high ambition/risk; M moderate ambition/risk; L low ambition/risk
e effective; me moderately effective; ne not effective
1/1 annual sampling; 1/4 four yearly sampling; <u>H</u> changed ambition level

effectively when all relevant criteria for the ambition level were met for both the shallow and deep screens, or if only one criterion was moderately effective in the columns of tables 4.8, 4.9 and 4.10. The sample size was judged to be ineffective if one criterion was judged not-effective (ne) and another moderately effective (me), or if several criteria were not-effective or moderately effective.

For Drenthe, the sample size in three homogeneous areas was effective for both environmental issues (Table 4.11). These areas belong to the ambition levels low and moderate and have sample sizes larger than 11. However, sample size in the high-risk areas for agricultural pollution (*a-r* and *g-r*) and for acidification (*f-r*) was too small to meet the higher ambitions on the proportion of contaminated groundwater and the assessment of typical concentrations.

For Noord-Brabant, the five homogeneous areas with the largest sample size score effective for OXC, but one failure was present for pH and two for nitrate (Table 4.11). The sample sizes in the homogeneous areas *dis* and *f-i* were sufficient for the reference and signal function in the two areas. The 11 wells in the area *f-r* were sufficient to monitor pH at the high ambition level, but not effective as a low-risk reference for agricultural pollution. Eight wells in *a-i* and *g-i* were not sufficient to estimate typical concentrations for pH, nitrate and OXC. Sample size in *a-r* was ineffective.

Steps 6 and 7 Evaluation of monitoring frequency, monitoring depths and sets of measured chemical components

Following the argumentation of section 4.2, trend detection was limited to areas where changes in groundwater quality are expected that result from recent changes in land use or input of solutes. This excludes homogeneous areas with a high percentage of old, pre-1950 groundwater, which corresponds to the low-risk discharge areas in the two networks and to the deep screens of the moderate-risk areas for which more than 50% of groundwater is pre-1950 (Chapter 3).

A differentiation in monitoring frequency is presented in Table 4.11. Annual monitoring is done only for the purpose of trend detection in the high-risk areas and in the shallow screens of moderate-risk areas. Monitoring frequency is reduced to once every four years in the low risk areas and the deeper level of the moderate risk areas. An exception was made for the area *il-i* in Noord-Brabant because a high ambition level was assigned to this area in the first step of the optimization (see step 9).

Step 8 Re-evaluating the delineation of homogeneous areas

There are two possible reasons to change the delineation of homogeneous areas:

1. when there are important regional spatial trends within the homogeneous areas
2. when monitoring information goals change in time.

Both aspects are considered below.

In the presented methodology, the contaminant concentrations in homogeneous areas were assumed to be statistically stationary. This assumption is presumably met, given the large distances between the wells with an average minimum well distance of about 6-9 km. Local studies on groundwater quality indicated that very large variation in groundwater quality is present over distances less than one km, even within areas with 'homogeneous' land use (Griffioen & Hoogendoorn 1993, Frapporti et al. 1995, Broers & Buijs 1997). A large part of the variations is probably due to the variation in geohydrological conditions and is primarily determined by the position of watercourses that define the extent of local and regional groundwater flow systems (see Chapter 2). Typical distances between the large brooks in Noord-Brabant and Drenthe are 5-10 km. Thus, the wells in one homogeneous area were distributed over many local and regional flow systems and no direct spatial correlation is to be expected.

Spatial trends within the homogeneous areas were assessed by analysis of the correlation of concentrations with spatial co-ordinates in order to check the assumption that the data were spatially stationary. For example, Figure 4.7 shows the 1995-1998 concentrations of OXC and nitrate at shallow depth in three homogeneous areas in a WSW-ENE section through Noord-Brabant. The section is perpendicular to the main geological structure of Noord-Brabant and the well locations were projected on the section.

The observations in the areas *intensive livestock farming - recharge* and *intensive livestock farming - intermediate* showed large variation over distances of less than 5 km in the WSW-ENE direction. Wells without nitrate and with nitrate concentrations above 100 mg l^{-1} were present over the whole section. The graphs present no indications for a spatial trend in OXC or nitrate concentrations that would call for a different network set-up. No spatial trend is present either in the *forest - recharge* areas.

Changing monitoring information goals is another possible reason for adaptations to the delineation of homogeneous areas. For example, the preceding evaluation of sample size and monitoring frequencies within homogeneous areas was based on the implicit assumption that water quality managers still desire to monitor all areas individually. However, a reasonable choice is to cluster or aggregate homogeneous areas to increase the monitoring precision for the price of losing differentiation in land use or geohydrology classes. Clustering of the geohydrological classes is not sensible, given the pronounced differences in the proportion of contaminated groundwater between recharge, intermediate and discharge areas for both the acidification and agricultural pollution issues. Clustering the agricultural land use classes is sensible, especially for Noord-Brabant. No distinct differences between the land use classes arable land, grassland and intensive livestock farming were observed for the intermediate areas. This probably resulted from the spreading of the large manure surplus across all land uses in

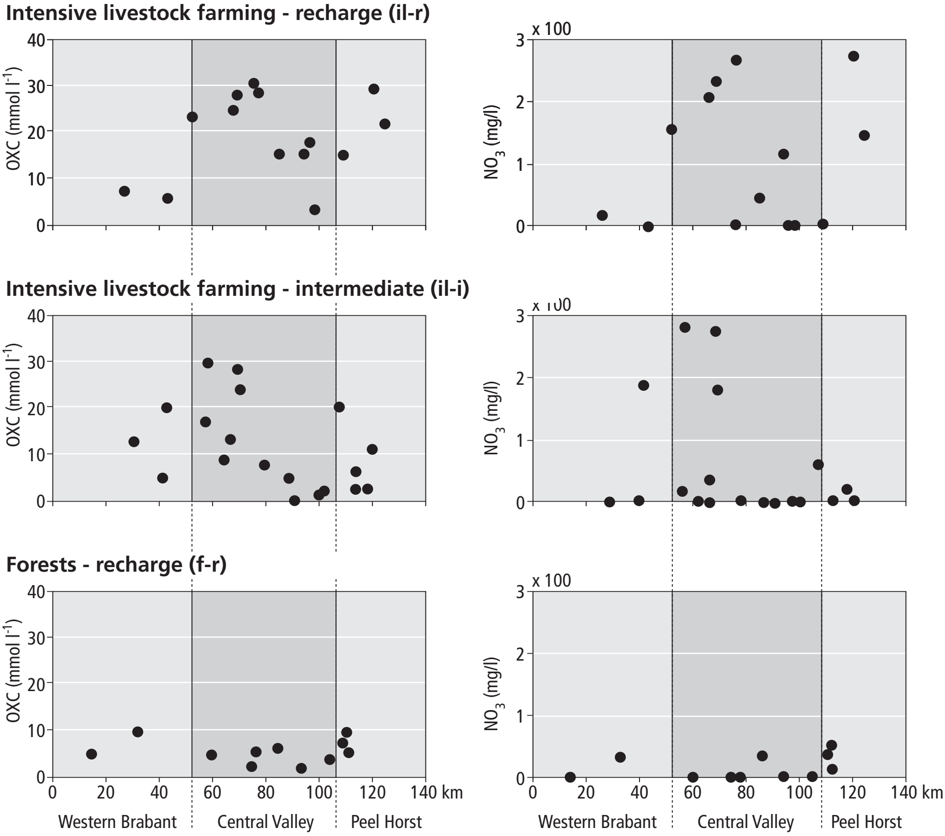


Figure 4.7 - Assessment of spatial trends of concentrations of OXC and nitrate for three homogeneous areas in Noord-Brabant using a projection on a WSW-ENE section that is perpendicular to the main geological structure in the province. 1995-1998 monitoring data for 5-15 m screen depth.

this province. Under these conditions, the monitoring of individual land use classes is not sensible.

Clustering the homogeneous areas might result in the wells not being allocated proportionally over the new combined area, and a correction must be made to estimate the proportion of contaminated groundwater and the corresponding 95% confidence interval (see Appendix III.6). Using such a correction, the larger sample size in the combined area *agricultural-intermediate* allowed for more precise estimation of the proportion of nitrate-contaminated groundwater at shallow depth. The proportion was estimated to be 23% with a confidence interval between 12 and 40% (Figure 3.13, Chapter 3). Assigning a high ambition level to the clustered area is a rational choice after clustering.

Step 9 First step in the optimization

In the first step of the optimization as presented here, no changes in the delineation of homogeneous areas were made. The integral evaluation of step 5 resulted in six homogeneous areas where increase of sample size was required to obtain an effective network: *a-r*, *g-r* and *f-r* in Drenthe and *a-i*, *g-i* and *a-r* in Noord-Brabant. Enlargement of sample size could first be limited to the shallow screens where variations in groundwater quality were largest. Table 4.11

presents a proposal for new sample sizes in those areas, based on the relations between sample size and desired precision for signalling (Figure 4.2) and for estimating the proportion of contaminated groundwater (Figure 4.3). The argumentation follows below.

For the high-risk areas in Drenthe, the initial network indicated proportions of nitrate-contaminated groundwater of 0.6 and 0.4 for *a-r* and *g-r*, respectively, and a proportion of groundwater with pH<5 of 0.4 for *f-r*. For these areas, the optimization aims at more precise determination of these proportions. Assuming that these proportions are good estimates of the proportion in the populations, a minimum of 12 wells is required to achieve a 30% precision, interpolating from Figure 4.3.

An increase of sample size to 11 wells is required to address the signal function in the the moderate-risk areas *a-i* and *g-i* in Noord-Brabant. Sample size enlargement in the area *a-r* in Noord-Brabant was not recommended, because too large investments would be required for an area that occupies only 0.4% of the provincial area (Appendix II). The current sample size was continued because the homogeneous area is the most vulnerable in the province according to the available monitoring results.

The evaluation results were straightforward for the six previously mentioned areas. For the homogeneous areas *il-i* and *il-r* in Noord-Brabant, however, a dilemma was present (Table 4.11). The 20 wells in *il-i* were not sufficient to estimate the pH within 1.5 pH units, but the sample size was sufficient to effectively monitor agricultural pollution. The 14 wells in *il-r* were not sufficient to precisely estimate the median nitrate concentration, but were effective in estimating the OXC concentrations. The strongly non-normal distributions of pH and nitrate concentrations in the two homogeneous areas did not allow for the specification of the required sample size without assuming a log-normal distribution.

Alternatively, bootstrapping was used to acquire an indication of the sample size needed to obtain precise estimates of the median pH and nitrate concentrations, assuming that the available samples give a good indication of the frequency distribution in the entire area. Thousand independent samples of sub-sample size 3, 5, 7 up to 60 were drawn with replacement out of the 20 pH values, using the procedure that is described in Appendix IV. The results indicate that the median pH in *il-i* cannot be estimated within 1 pH-unit, even with 60 samples. A similar result was obtained for the nitrate concentrations in the area *il-r*. If the 14 samples give a good indication of the population frequency distribution; a precision within 50 mg l^{-1} is not feasible. These results show that estimation of the median is not sensible when bimodal frequency distributions exist in the homogeneous areas (see section 4.4). Alternatively, the typical concentrations of OXC instead of the typical concentration of nitrate can be estimated in order to assess the impact of agricultural pollution (section 4.4).

A small increase in sample size was recommended for the two homogeneous areas in an attempt to improve the precision of typical concentrations (Table 4.11). For the homogeneous area *il-i* a decision was made to assign a high ambition level for agricultural pollution, because the homogeneous area is the largest in the province and signals for temporal trends at greater depth are present (Chapter 5). Using the high ambition level, the emphasis changes from signalling into determining proportions of contaminated groundwater, and determining median temporal trends (statistical information goals D and F).

The application of area-specific monitoring information goals yields opportunities for cost-reduction. For example, by decreasing the monitoring frequency in the low-risk homogeneous areas from once every year to once every four years, the reference function is still obtained for a quarter of the present costs (step 6). Every 4 years a complete overview of groundwater chemical status will be presented for the two monitoring depths. This overview will be used in the policy reports on the state of the environment that are published every 4 years. The 4-year sampling rounds may also be used to assess specific contaminants, for instance a reconnaissance survey of selected pesticides. This enables comparison of routinely measured chemical indicators with the selected specific contaminants for the homogeneous areas.

Decreasing the monitoring frequency is also recommended for the deep screens of the moderate-risk areas where no signals for groundwater quality deterioration were observed (Table 4.11). The reduced monitoring frequency should later be adapted when the four-year sampling rounds indicate an increase of typical concentrations or proportions of contaminated groundwater in the moderate or low-risk areas.

The reduction of monitoring frequency in low and moderate-risk areas renders an overall reduction of the exploitation costs of the network, even though the sample size is enlarged for some areas. For Drenthe, an average yearly amount of 110 samples has to be sampled and analysed, instead of 120 in the original set-up. For Noord-Brabant the amount of samples is reduced from 178 $year^{-1}$ to 136 $year^{-1}$. The indicated exploitation costs for Noord-Brabant reduce from € 56.500,- to € 47.300,- per year. Installation costs for the shallow wells needed to increase the sample size requires a one-time investment of about € 30.000,- for each province.

The evaluation procedure presented in this chapter should be considered as a first step in the optimization of the networks, using the collected monitoring data to refine information requirements, sample size and monitoring frequency. After the first optimization round, a new evaluation round is recommended to judge the effectiveness of the improved networks and make adaptations if necessary.

4.4 Discussion

The following issues are discussed in this section: (1) the effects of random sampling, (2) decisions about area-specific ambition levels and desired precision, and (3) the observed variation and multi-modal frequency distributions in the two regional networks.

Random sampling

The monitoring data were assumed to be sampled randomly from the homogeneous areas. This assumption has not been strictly fulfilled in the Dutch national and regional networks as explained in Chapter 3 (sections 3.3 and 3.4). This is probably the largest flaw in the design of these networks. Therefore, the confidence intervals should be regarded as indications, rather than absolute values. New wells should be chosen randomly within the homogeneous areas to help improve the estimates for the homogeneous area as a whole.

If a new network is to be designed in other areas, random sampling should be applied within the strata chosen for sampling. However, the sampling locations should be carefully chosen with respect to field accessibility and operational constraints, such as the avoidance of local point and line sources contamination, including road salts. This requires robust procedures to combine the local conditions with random sampling. Given the results of the evaluation of the two Dutch regional networks, an initial sample size of at least 10 wells in each stratum should be advised to address the reference and signal functions appropriately.

Decisions about ambition levels and desired precision

The evaluation procedure contains decisions concerning ambition levels and required precision:

1. the ambition level that is assigned to a certain area
2. the level of precision that is desired for effective monitoring.

The ambition level or the desired precision level for specific areas can change during the evaluation procedure. For example, ambitions in moderate-risk homogeneous area *il-i* were raised because serious signals for groundwater contamination were observed. In this first evaluation of the networks the desired precision levels were not very ambitious for signalling contamination and determining the magnitude of groundwater contamination. Though, sample size in 8 out of 12 moderate-risk and high-risk areas was insufficient when evaluating these precision levels and sample size increase was recommended as the first step in the optimization. Further enlargement of sample size is necessary when higher precision of the estimates is desired for water quality management.

In this chapter no special attention has been paid to the precision of the analysis of temporal trends in homogeneous areas (statistical information goals F and G, Table 4.5). This was considered to be the second step in the evaluation/optimization procedure. If no precise estimates of median concentrations and proportions of contaminated groundwater are assessed using the current sample size, then median trends or trends in proportions will not be detected easily either. Therefore, the first optimization step aimed at improving the precision of estimated medians and proportions before evaluating statistical information goals F and G. Nevertheless, recommendations for suitable monitoring frequencies for individual wells were given that should enable the detection of trends following statistical information goal E. An appropriate monitoring frequency is also a prerequisite to meet the statistical information goals F and G.

Observed variations and multi-modal frequency distributions

In general, a larger variation of concentrations was encountered than previously expected on the basis of preliminary surveys, especially in high-risk areas and the shallow screens of moderate-risk areas. Especially the assessment of typical concentrations in these areas, which was a basic function of the network, was more difficult than anticipated. Three options are now possible:

1. increase the sample size in the individual areas
2. aggregate some homogeneous areas to obtain larger sample size, losing differentiated information for specific land use or geohydrological classes, and/or
3. redefine the delineation of homogeneous areas, for instance using spatial information on the reactivity of sediments.

Opportunities for evaluation using the options (1) and (2) were given in the previous section. Option 3 is not yet feasible because of the present lack of spatial information on subsurface reactivity. Moreover, option 3 increases the number of strata and the required sample size of the complete network accordingly.

Part of the observed large variation is determined by specific behaviour of the selected chemical components. For example, nitrate showed a strongly bimodal frequency distribution even in the high-risk areas. This is caused by the hydrogeochemical behaviour of nitrate; the nitrate concentration drops below the detection limit if small amounts of reactive organic matter or pyrite are encountered (Chapters 3 and 5). If no reactive material is encountered, nitrate moves conservatively and the high concentrations in recharging groundwater flow unhampered through the subsoil. Because of this behaviour, the assessment of the proportion of nitrate-contaminated groundwater is a much better measure for nitrate contamination than the median or a high percentile of the nitrate concentration. Because the frequency distribution of OXC was much closer to normal than for nitrate, OXC is preferably also used to assess the impact of agricultural pollution in the homogeneous areas. Precise estimates of typical concentrations are easier obtained for OXC than for nitrate.

Similar problems as for nitrate also occur with aluminum and heavy metals like zinc (Broers

& Peeters 2000, Venema et al. 2000, van Vliet 2000). Aluminum shows a strong pH-dependency because of the dissolution of aluminum hydroxides and silicates and desorption at low pH (Appelo & Postma 1993). High aluminum concentrations are observed below pH = 5 in the Dutch national and regional monitoring networks (for example, Van Duijvenbooden 1993). Sorption of trace metals is also strongly dependent on pH and causes steep desorption gradients with pH (Dzombak & Morel 1990). As a result, the assessment of typical concentrations of aluminum and trace metals is seriously hampered if the stratification in homogeneous areas does not result in a narrow range of pH values within these areas. In this study, narrow ranges of pH were only observed in some specific homogeneous areas.

Proportions of nitrate, aluminum and trace metal concentrations above critical concentrations are much easier to estimate precisely than medians, means or percentiles, even if large variations are observed within the homogeneous areas. For these kinds of chemical species with multi-modal frequency distributions it is advised to focus on proportions above critical concentrations in the evaluation procedure, rather than focusing on typical concentrations or high percentiles of the frequency distribution.

Alternatively, the samples of a bimodal distribution can be split in two parts. The typical concentration of groundwater is then assessed separately for the subgroup of wells with the high part of the bimodal distribution. However, larger sample sizes are then required to obtain precise estimates of the high part of the frequency distribution.

4.5 Conclusions

The aim of the present study was to provide a framework for the evaluation and optimization of the Dutch regional networks for groundwater quality. The underlying design strategy of those networks was to have higher ambitions for areas with high risks for contamination of deep groundwater than for areas with moderate or low-risks. The framework assigns area-specific statistical information goals to high-risk, moderate-risk and low risk areas. The presented evaluation strategy distinguishes three main functions:

1. the reference function, which aims at an overview of groundwater chemical status in the region
2. the signal function, which aims at the detection of groundwater contamination and the detection of temporal trends
3. the determination of the magnitude of groundwater contamination in high-risk areas both in time and space.

The approach uses the available monitoring data and non-parametric methods to evaluate the sample size in the networks and to optimize the network step-by-step. The advantages of the methodology are:

1. a strong focus on the high-risk areas, which results in less uncertainty in the estimates of typical concentrations and the proportion of contaminated groundwater
2. feasibility of cost-reduction in low-risk and moderate-risk areas, while maintaining the opportunities for the detection of groundwater contamination and for the assessment of a regional overview of groundwater chemical status.

The demonstration of the evaluation of the two Dutch provincial networks indicated that the sample size of the networks is sufficient to effectively assess general groundwater quality patterns and to address the reference function. The groundwater quality patterns in the two provinces reflected the original risk assignment that was used to differentiate the sample sizes in the homogeneous areas. However, judging the monitoring effectiveness in the individual homogeneous areas, especially the sample size in many high-risk areas was insufficient to monitor typical concentrations of targeted contaminants or the proportion of contaminated groundwater. More effective monitoring and lower exploitation costs are feasible by increasing the sample size in the high-risk areas and reduction of the monitoring frequency in the low and moderate-risk areas.

5 Regional monitoring of temporal changes in groundwater quality

(with Bas van der Grift)

5.1 Introduction

The EU Water Framework Directive (EU 2000) distinguishes two types of monitoring objectives for groundwater quality: (1) 'to present an overview of groundwater chemical status' and (2) 'to detect the presence of long-term anthropogenically induced upward trends in the concentration of pollutants'. The Dutch groundwater quality monitoring networks fulfil their task by providing the overview of the time-averaged chemical status of groundwater for areas with specific land use, soil types and geohydrological situations (Chapter 3, Reijnders et al. 1998). However, temporal changes in concentrations of targeted components in larger areas are not detected easily, even though data have been collected for 10 to 15 years. Moreover, there is a lack of understanding of the trends detected and the reasons why trends are not found. The general lack of understanding is illustrated by using the trend detection results for nitrate in the Netherlands, which is an important component in groundwater monitoring.

Temporal changes in nitrate concentrations over the period 1984-1993 have been found for only one specific physical geographical region in the Netherlands, namely the central sand district (Reijnders et al. 1998, Van Drecht et al. 1996). Significant temporal trends for nitrate have not been detected in any of the individual *land-use/soil-type* strata of the national monitoring network. This is an unexpected result, because of the overall increase in the production and use of animal manure and fertilizer salts in the Netherlands in the last decades. In contrast, Pebesma & de Kwaadsteniet (1995, 1997) using stratified kriging interpolation of the 1987-1993 data detected large areas with increasing nitrate concentrations. A closer look at these results indicates that much of the increasing nitrate concentrations occur in sandy areas with semi-natural vegetation such as forest, and only few occur in the grassland or arable land areas where fertilizer and manure are used. The maps of Pebesma & de Kwaadsteniet indicate increasing nitrate concentrations under grassland and arable land in some specific parts of the eastern and southeastern sand districts of the Netherlands.

Frapporti (1994) assessed temporal trends within specific water types. He demonstrated a significant increase in the nitrate concentration in the *polluted* water type, which is generally rich in nitrate. However, his results cannot directly be translated to areas, although he indicates that this water type is most often found in the higher sandy parts of the Netherlands. Both Frapporti and Reijnders et al. related the detected increasing nitrate trends to lowering of groundwater tables that limits the denitrification in the upper soil (Steenvoorden 1987, Van Drecht et al. 1991, Van Drecht 1993). No attempts were made to relate the trends to the input history of manure and fertiliser or to the travel time of the groundwater to the monitoring screens.

Recently, Bronswijk & Prins (2001) used an alternative method to demonstrate temporal changes in nitrate concentrations. They used groundwater age dating to interpret nitrate concentrations of wells under agricultural land and sandy soils in the Netherlands and found a relationship between infiltration year and the proportion of nitrate contaminated groundwater.

An earlier, short version of this chapter was published in *Impact of Human Activity on Groundwater*, IAHS Publication no. 269, pp. 247-253 (2001)

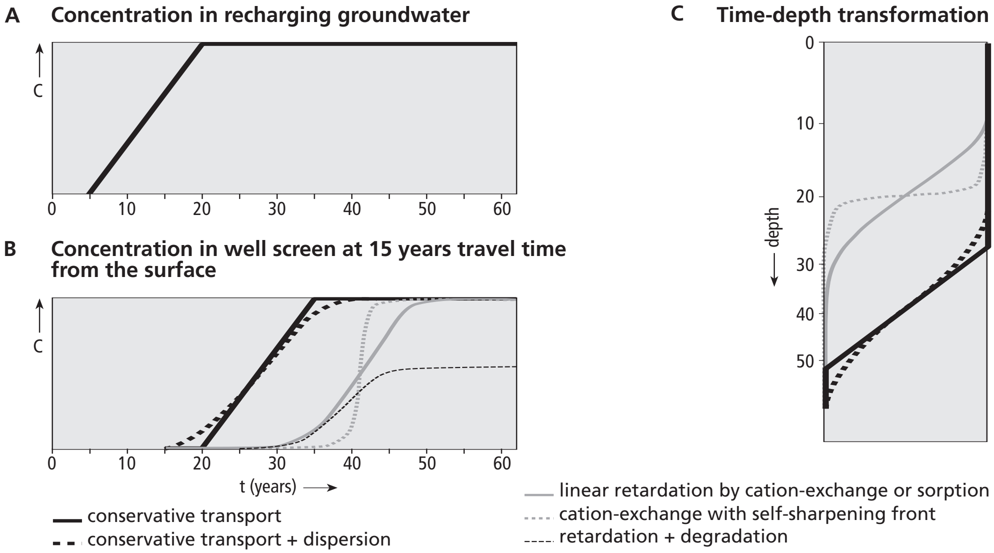


Figure 5.1 - The concentration response of groundwater at a specific depth (B) to a linear increase of the manure loads (A) given a 15 year hydrologic residence time from the recharge point. A concentration-depth transformation indicates the position of conservative and reactive contamination fronts at a specific moment in time (C)

Their results showed good correspondence with estimated agricultural nitrogen inputs and indicated increasing nitrate concentrations over the last 30 years.

Overall, fewer trends were detected than previously expected and results of the trend analysis differed for different methods. Probable reasons of not detecting groundwater quality changes are:

1. the long travel times of groundwater to the well screens
2. the obscuring, attenuating or retarding effect of physical and chemical processes on solute breakthrough
3. the spatial variability of contaminant concentrations in recharging groundwater, in hydrologic residence time, and in reactive properties of the aquifer sediments
4. the short-term natural temporal variability of groundwater composition at the monitored depths (Chapter 4).

Figure 5.1 presents a theoretical example of the way the factors (1) and (2) influence the detection of temporal trends. Figure 5.1a reflects an input history with linearly increasing concentrations in recharging groundwater. Figure 5.1b presents the concentration response in a monitoring well, given a 15 year hydrologic residence time from the recharge point. A linear trend will be observed beginning 15 years later for a conservative chemical component. Retarding or dispersive processes tend to extend the response time and the associated monitoring period needed to detect the changes in concentration for reactive, non-conservative chemical components. Reactive processes may even transform the linear trend into a step trend. Many targeted chemical components, including the nutrients N, P and K, are known to behave non-conservatively in groundwater. Different hydrogeochemical processes determine the breakthrough of those nutrients (Griffioen & Broers 1999, Griffioen 2001).

The example shows that trends in conservative solutes will be detected earlier than trends in reactive components and indicates that groundwater travel times and likely subsurface reactions should be evaluated to explain the trends in the monitoring data. In regular trend analysis studies, only concentrations of the chemical components that are of direct interest, like nutrients N, P and K, are investigated (for example, Reijnders et al. 1998, Bronswijk & Prins 2001). These trends become easily obscured by chemical reactions. Therefore, in this study the combined use of conservatively behaving indicators and potentially reactive components is recommended.

Trend detection is normally limited to the analysis of time series of individual wells or groups of wells. Hallberg & Keeney (1993) showed that trends in groundwater quality can also be detected using concentration-depth information, because depth and groundwater age are interrelated, a phenomenon that is most pronounced in groundwater recharge areas (Raats 1978, Raats 1981, Dillon 1989; see Chapter 2). Environmental tracers, such as tritium, are useful for age dating of the groundwater. The input history of Figure 5.1a can be transformed into the concentration-depth profile of Figure 5.1c after an age-depth relation is established for a specific area. Figure 5.1c shows the theoretical position of the conservative and reactive contaminant fronts at a specific moment in time. The example shows that a comparison of measured front positions with a conservative breakthrough yields indications of retarding and attenuating processes in the aquifer. This kind of information is useful to explain observed quality changes and concentration-depth changes.

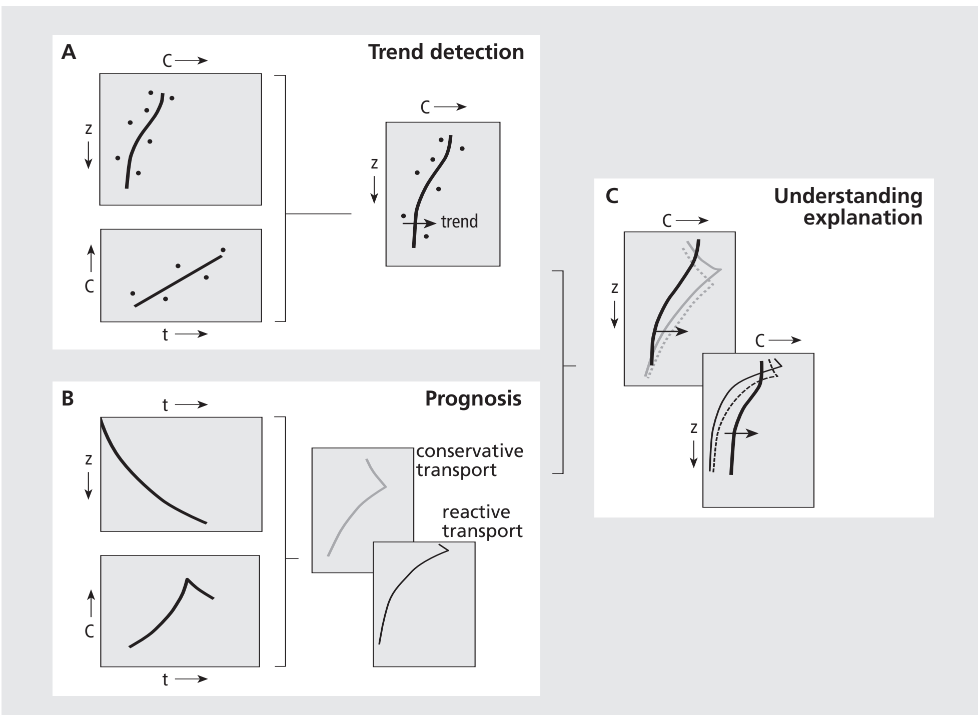


Figure 5.2 - Overview of the presented approach. Trends are detected using concentration-depth profiles and concentration-time series for homogeneous areas (A). Prognoses for conservative and reactive transport are based on a theoretical groundwater age distribution and time series of solute input in recharging groundwater (B). Prognoses and detected trends are compared to increase understanding of trends and processes that prevent quality changes in time (C)

Based on the above-mentioned argumentation, this study aims at improving the detection and the understanding of groundwater quality changes in time. We adopted the trend definition of Loftis (1991, 1996), which defines a temporal trend as a change in groundwater quality over a specific period in time, over a given region, which is related to land use or water quality management.

An overview of the presented approach is given in Figure 5.2. For *trend detection*, we combine the analysis of time series data from a specific depth, and concentration-depth profiles for a specific time (Figure 5.2a). Time series analysis and concentration-depth analysis are done both for conservative and potentially reactive components. Prognoses are made of conservative and reactive transport to aid the *understanding* of the detected trends. The prognoses are based on groundwater age determinations using tritium and historical time series of the inputs of components in recharging groundwater (Fig 5.2b). The prognoses are compared with the detected trends and observed depth profiles to explain the observed quality changes (Figure 5.2c). The approach is illustrated by an analysis of the effects of increasing and decreasing agricultural contaminant loads in two areas in the Dutch province of Noord-Brabant.

Section 5.2 introduces the monitoring network in the investigated areas. In the sections 5.3 to 5.7, the following issues are discussed: (1) the age-depth relationship, (2) concentration-depth profiles, (3) trend analysis on time series, (4) prognosis for conservative transport and (5) prognosis for reactive transport. Methods are presented in the individual sections. Results are illustrated for the two areas in Noord-Brabant.

5.2 Monitoring network in the investigated areas

Noord-Brabant is one of the areas in Europe that is most affected by agricultural pollution, because of intensive livestock farming which causes a large surplus of manure. The monitoring data used originate from the provincial groundwater quality monitoring network, which assesses groundwater quality using wells in so-called *homogeneous areas* that have homogeneous land use, soil type and geohydrological situation (Chapter 3). Two homogeneous areas were selected in the Pleistocene sandy areas of Noord-Brabant: *intensive livestock farming-recharge* and *intensive livestock farming-intermediate*. The homogeneous areas cover 230 and 1100 square kilometers and contain 14 and 21 wells, respectively (Chapter 3, Table 3.1). The extent of the homogeneous areas and the positions of the available monitoring wells are shown in Figure 5.3.

The selected homogeneous areas are characterized by a large proportion of maize land (> 25%) which is grown alternating with grassland and arable crops. Originally, those areas were used as arable land or as grassland for dairy farming, but the amount of maize strongly increased after 1970. The two homogeneous areas have a different position in the hydrological system (Chapter 2). The geological and hydrogeological situation in Noord-Brabant have been discussed in Chapter 2.

The *intensive livestock farming-recharge* areas correspond with the higher part of the landscape and show relatively deep groundwater levels and the absence of a superficial drainage system of ditches (Chapter 3). The areas belonged to the oldest agricultural land in the province and contain well-drained, often fertile and humic soils (Dutch soil types zEZ21/23).

The area *intensive livestock farming-intermediate* corresponds to areas that were brought in culture after the making of ditches in the beginning of the 20^{th} century. Shallow groundwater levels and wet podzolic and humic soils (Dutch soil types Hn21/23, pZg21/23) characterize the area. The intermediate areas are characterized by the presence of local flow systems that discharge in the local ditches and small brooks (Chapter 2).

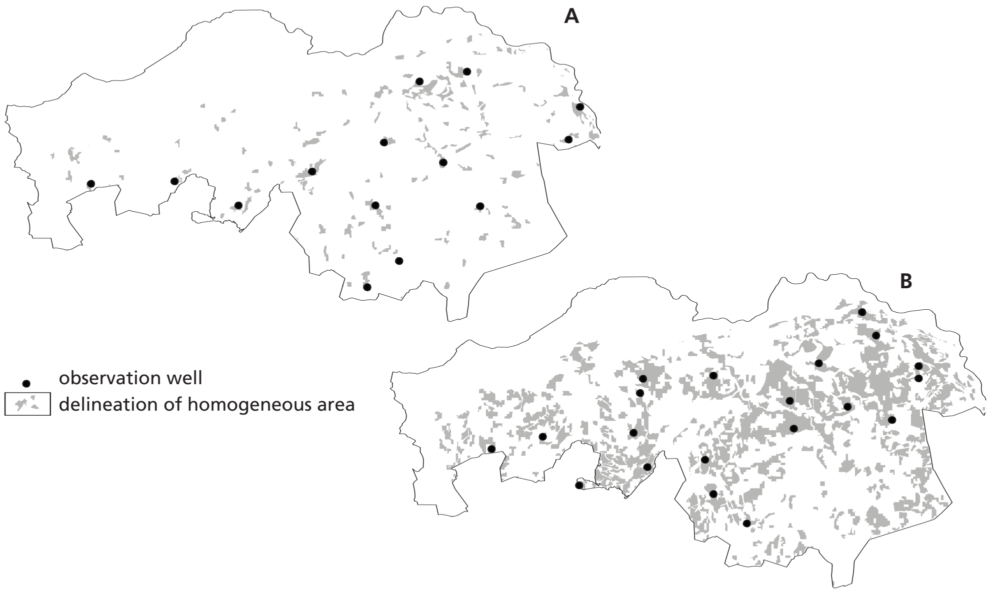


Figure 5.3 - The spatial extent of the homogeneous areas intensive livestock farming - recharge (A) and intensive livestock farming - intermediate (B) and the position of the monitoring wells.

Time series over the period 1984-1998 with annual data were available for the wells that are part of the national monitoring network for screens at about 9 and 23 m depth. The provincial wells were installed later and have time series of the period 1991-1998 for similar depths. Some of the provincial wells contain an additional screen directly below the average lowest groundwater level, which comprise time series of the same period. We consider the data to be a random sample of groundwater quality in the homogeneous areas (Chapters 3 and 4).

5.3 Age-depth relationships for homogeneous areas

Tritium concentrations were used to assess the age-depth relationships for the two homogeneous areas. Tritium was measured in the national monitoring wells in 1983, and in the provincial monitoring wells in 1992 (Chapters 2 and 3). First, groundwater ages were assessed for individual wells by fitting the tritium concentrations at a time series record of tritium in precipitation. Second, age-depth relations for the homogeneous areas were derived, using the median and 10 and 90 percentiles of the groundwater ages in individual wells. This way, not only an average of the age distribution was obtained, but the uncertainty was also indicated using the percentiles. This is important, because large variations in groundwater age were expected, especially in the drained, intermediate areas (Chapter 2).

The method for age dating of the individual screens is similar to the one used by Meinardi (1994) who assessed the annual amount of groundwater recharge, using the tritium concentrations of 9 and 23 m deep screens of the national monitoring wells. To derive the time passed since the recharge year of the groundwater, the measured tritium concentrations of the two screens were compared with the 5-year averaged tritium concentrations in precipitation (data from Meinardi 1994 and Stuyfzand 1991). The tritium input of precipitation in Noord-Brabant was corrected for radio-active decay up to the sampling years 1983 and 1992 (Figure

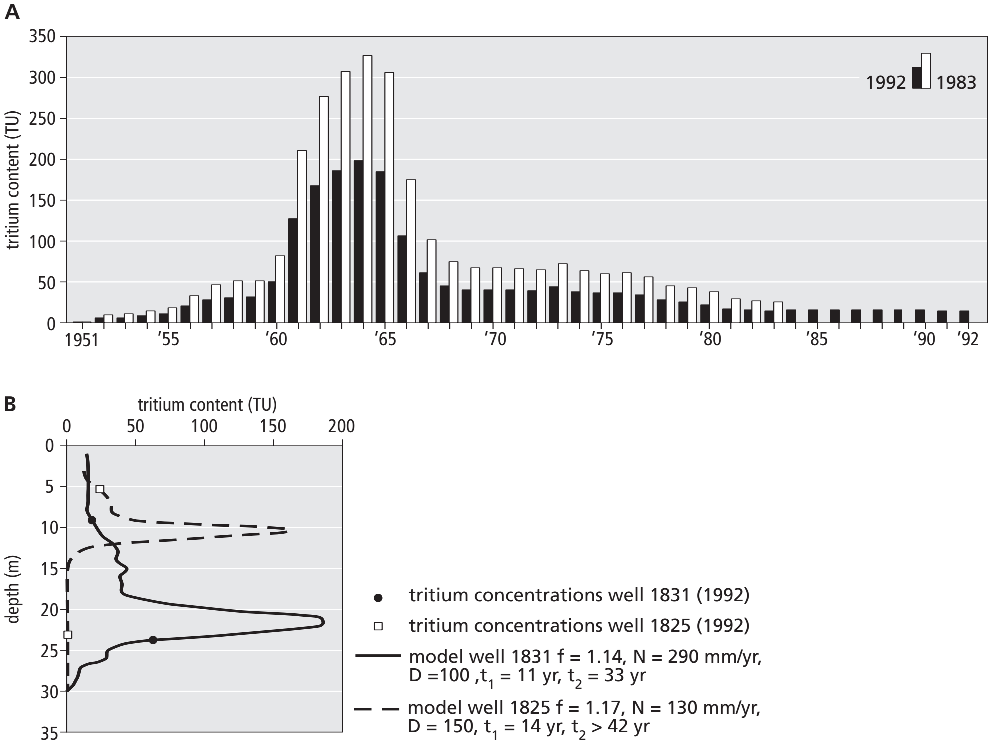


Figure 5.4 - Five year averaged tritium concentrations in Noord-Brabant precipitation (A) and example of the fitting method of the measured tritium data for two provincial wells (B). Tritium concentrations were corrected for radio-active decay up to 1983 and 1992. Data from Meinardi (1994) and Stuyfzand (1991).

5.4a). Five-year averages were used for age dating to correct for the sampling out of the two metre long screens, because sampling from the long screens induces mixing of several years of recharging groundwater (Meinardi 1994). The tritium concentrations of the two screens of one monitoring well were fitted simultaneously, using the theoretical age-depth of relationship that is valid for a homogeneous aquifer of finite thickness (Raats 1978, 1981, see also Chapter 2):

$$t_z = \frac{\varepsilon D}{N} \ln\left(\frac{D}{D - z}\right) \tag{5.1}$$

In this equation t_z is the groundwater age at depth z, N is the groundwater recharge rate (m day^{-1}), ε = porosity, and D = aquifer thickness (m). A porosity of 0.35 was used for all calculations (Meinardi 1994). Aquifer thickness D was taken from Meinardi (1994) who distinguished 6 geohydrological regions in Noord-Brabant with different aquifer thickness. The tritium input curves were individually corrected for the spatial regional trend in the yearly averaged tritium levels in precipitation over the Netherlands, using the factors of Meinardi (1994, p.51). The groundwater recharge rate N was varied until a reasonable fit was obtained for the tritium concentrations of the two monitoring screens (Figure 5.4b).

The fit results in groundwater ages for the shallow screens (t1) and the deep screens (t2). Results for groundwater recharge N, aquifer thickness D and groundwater ages t_1 and t_2 are presented in Tables 5.1 and 5.2. The results for t_1 and t_2 are less uncertain than results for N

Table 5.1 - Groundwater ages derived from tritium measurements in 14 wells of the homogeneous area intensive livestock farming - recharge

	Shallow screens		Deep screens						
	Depth (m)	Tritium (TU)	Depth (m)	Tritium (TU)	f	N (m yr^{-1})	D (m)	t_1 (yr)	t_2 (yr)
National wells 1983									
(det. limit 5 TU)									
107	7.6	66	24	18	1.2	0.36	50	8	32
108	8	57	24	<5	1.15	0.30	100	10	>33
122	9	76	24	42	1.2	0.41	50	8	28
125	12.1	78	23	29	1.2	0.33	80	14	29
Provincial wells 1992									
(det. limit 0.6-1.6 TU)									
1806	5	11.1	23	43.9	1.12	0.27	150	7	32
1810	4	20.5	12.5	1.6	1.12	0.10	150	14	40
1823	8	19.6	18	40.6	1.11	0.20	150	14	34
1833	8	29.7	19	<1.6	1.15	0.17	100	17	>42
1840	10	17.5	24.5	7.8	1.17	0.26	80	14	39
1843	8	21.5	23	16.9	1.17	0.25	80	12	38
1851	9	31	23	33.9	1.19	0.28	80	12	34
1853	10	25	24	7.6	1.16	0.26	80	14	38
1863	10	34	22	3.8	1.15	0.25	50	16	41
1866	9	12.9	23	-	1.16	0.32	100	10	-
median						0.27	80		
10 percentile						0.18	50		
90 percentile						0.35	150		

f = correction factor for regional spatial trend of tritium in precipitation (Meinardi 1994)

because different combinations of N and D yield equivalent fits, which results in a wide range of possible values of N. A good estimate for aquifer thickness D is difficult to obtain without local geohydrological information. However, the range of combinations of N and D yield only small differences in t_1 and t_2.

Figure 5.4b shows an example of a well with post-1950 water in both screens and one with pre-1950 in the deep screen. In the second case, the model is fitted on the tritium concentration of the shallow screen. A groundwater recharge rate N was also estimated for two wells with pre-1950 water in both screens. This does not yield valuable results for t_1 and t_2, but the determined low values of N are used for the calculation of percentiles of the age-depth relations (see below). The resulting age-depth data were plotted as data points in Figure 5.5. A range is given for screens with pre-1950 water.

The medians and the 10 and 90 percentiles of the groundwater recharge rate N and aquifer thickness D were used to derive the continuous age-depth relations of Figure 5.5 using equation (5.1). The middle line reflects the median age-depth relation. The outer lines represent the 10 and 90 percentiles of the age-depth relations the homogeneous area.

The results for the homogeneous area *intensive livestock farming-recharge* indicate that groundwater at 9 m depth is 9 to 19 years old, with a median of 12 years. The median for groundwater depth at 23 m depth is 36 years, with 10 and 90 percentiles 25 and 60 years.

A larger variation of groundwater ages and generally older groundwater is detected in the area *intensive livestock farming - intermediate* compared with *intensive livestock farming - recharge* (Table 5.2 and Figure 5.5). Median, 10 and 90 percentiles are 15 years and 11 and 39

Table 5.2 - Groundwater ages derived from tritium measurements in 20 wells of the homogeneous area intensive livestock farming - intermediate

	Shallow screens		Deep screens						
	Depth (m)	**Tritium** (TU)	**Depth** (m)	**Tritium** (TU)	**f**	**N** ($m\ yr^{-1}$)	**D** (m)	t_1 (yr)	t_2 (yr)
National wells 1983									
(det. limit 5 TU)									
102	9.5	<5	24	<5	1.15	0.08	100	>33	>33
111	10.1	<5	32.9	<5	1.15	0.08	100	>33	>33
112	7	54	38	<5	1.15	0.31	80	8	>33
116	11	47	25	41	1.18	0.44	50	10	28
140	10	46	24	<5	1.12	0.14	150	26	>33
147	7.5	37	24	<5	1.12	0.27	150	10	>33
151	6.2	43	19.5	<5	1.13	0.20	150	11	>33
Provincial wells 1992									
(det. limit 0.6-1.6 TU)									
1806	8	20.1	24	1.7	1.11	0.22	150	13	41
1817	8	20.7	24.5	5.2	1.12	0.23	150	13	41
1822	11	43	23	7.5	1.15	0.22	150	18	40
1825	5	29.3	23	<1.6	1.17	0.13	150	14	>42
1831	9	19.3	24	62.9	1.14	0.29	100	11	33
1835	7	19	18.5	2.3	1.17	0.18	80	14	41
1837	9.5	35.4	25	2.5	1.16	0.25	80	14	41
1841	9	26	22	<1.6	1.17	0.22	80	15	>42
1855	9	25.8	23	17.4	1.18	0.25	80	13	38
1859	8	18.5	19	<1.6	1.18	0.20	50	15	>42
1860	5	19	24	10.7	1.15	0.28	50	7	41
1861	9	10	24	<1.6	1.16	0.09	50	39	>42
1862	8	16	24	<1.6	1.16	0.26	50	12	>42
median								0.22	90
10 percentile								0.09	50
90 percentile								0.29	150

f = correction factor for regional spatial trend of tritium in precipitation (Meinardi 1994)

years at 9 m depth, and 42, 30 and 121 years at 23 m depth. This corresponds with the findings of Chapter 2, which showed large spatial variability in the intermediate areas, especially at the depth of the deep monitoring screens. This result was explained by the presence of local groundwater flow systems that cause distortion of the horizontal position of isochrones.

The established age-depth relations are used to qualitatively evaluate the concentration-depth data of chemical components in the next section. The median, 10 and 90 percentile distributions are subsequently used for the prognoses of conservative and reactive transport (sections 5.6 and 5.7).

5.4 Concentration-depth profiles for homogeneous areas

Once the groundwater age-depth relationship has been established for a homogeneous area, we interpret the concentration-depth profiles as time-related profiles. Figure 5.6 shows concentration-depth profiles for a number of chemical components in the homogeneous area *intensive livestock farming-recharge*. The points reflect the average concentrations over the monitoring period 1995-1998 for individual monitoring screens. The upper scatter plots show

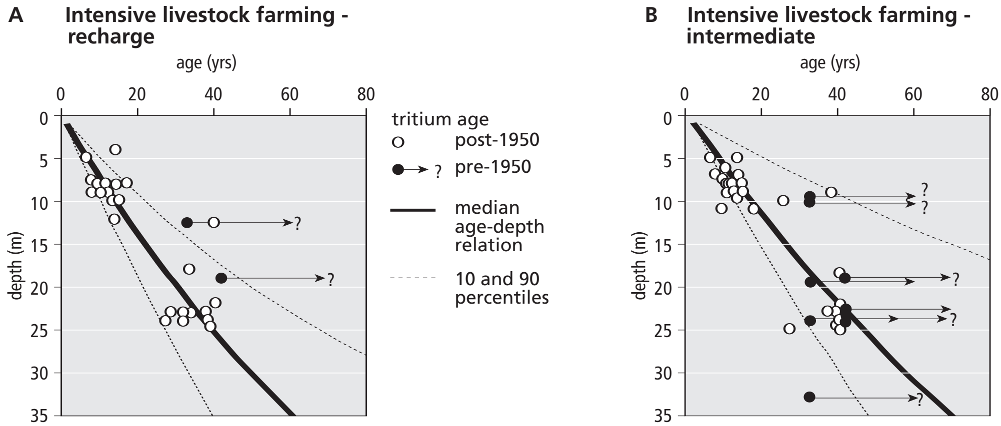


Figure 5.5 - Age-depth relations for the homogeneous areas intensive livestock farming-recharge (A) and intensive livestock farming-intermediate (B). Lines represent the median age distribution and the 10 and 90 percentile age distributions that were derived from the groundwater age data of all wells in the homogeneous area. Symbols indicate the groundwater age in individual monitoring screens for samples with tritium (open symbols) and samples without tritium (closed symbols)

concentration profiles for nitrate, aluminum and potassium. These are targeted contaminants, which are normally used to determine the magnitude of contamination and changes of concentrations with time. The lower scatter plots present three chemical indicators that are not reported on a regular basis: the oxidation capacity (OXC), the chloride concentration and the sum of cations (SUMCAT). The oxidation capacity was defined as the sum of NO_3 and SO_4 after Postma et al.(1991):

$$OXC = 5^*(NO_3) + 7^*(SO_4) \qquad (5.2)$$

where brackets denote molar concentrations. These chemical indicators are used because they are insensitive to specific subsurface reactions. We refer to them as *conditionally conservative* indicators. For example, oxidation capacity (OXC) behaves conservatively during the process of nitrate reduction by pyrite oxidation under the condition that no other reactions occur:

$$5FeS_2 + 14NO_3^- + 4H^+ \rightarrow 5Fe^{2+} + 10SO_4^{2-} + 2H_2O + 7N_2 \qquad (5.3)$$

The OXC concentration remains constant during pyrite oxidation: 10/14 mmol l^{-1} SO_4 is released when 1 mmol l^{-1} NO_3 reacts with pyrite. Before and after this reaction the OXC equals 5 mmol l^{-1} and OXC transport is conservatively.

The sum of cations (SUMCAT) is useful as a conditionally conservative indicator when cation-exchange processes dominate the transport of the cations and mineral dissolution processes do not occur. Moreover, SUMCAT is an indicator of the total load of solutes in the groundwater. The use of SUMCAT for the studied two homogeneous areas is sensible because no indications for mineral dissolution are present. For example, the proportions of calcite saturated groundwater are low (Chapter 3).

Chloride is a conservatively transported ion under normal pH conditions. The conditionally conservative indicators were used in the time-series and concentration-depth analysis.

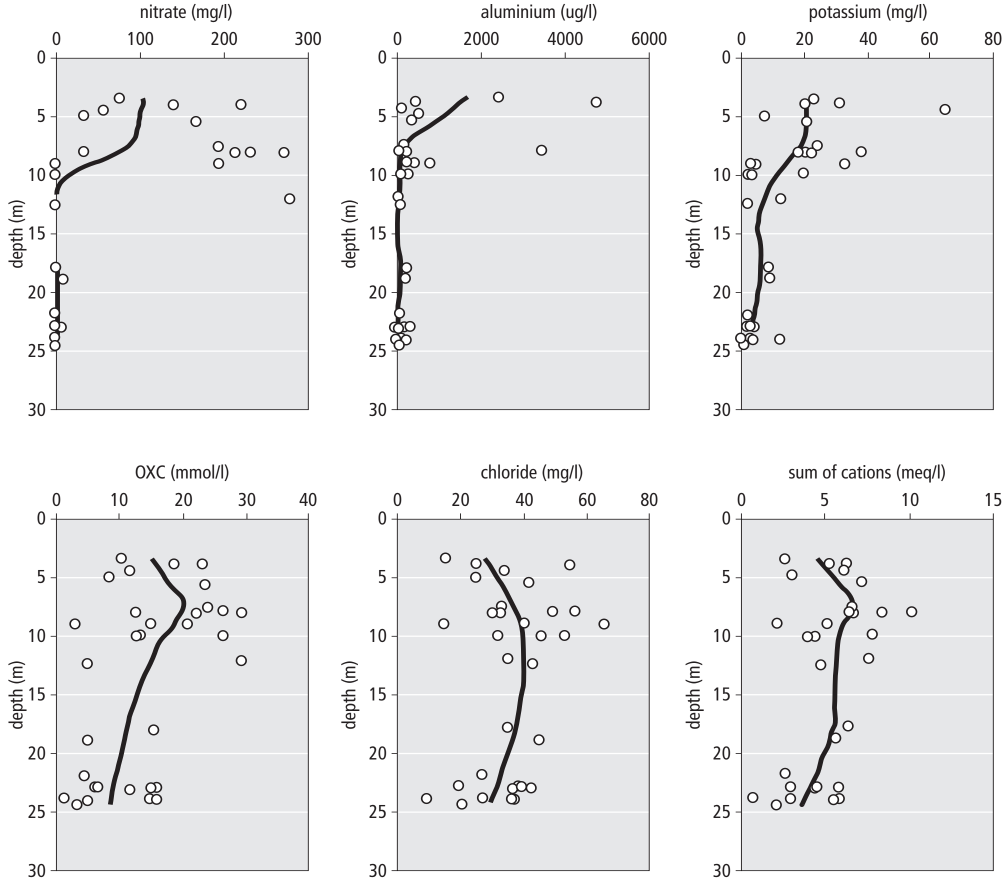


Figure 5.6 - Concentration-depth relations for the homogeneous area intensive livestock farming-recharge. Data from 14 wells, averaged for the 1995-1998 monitoring period. The LOWESS smooth indicates the centre of the scatter (span = 0.75)

A LOWESS smooth (locally weighted scatter plot smoothing; Cleveland & Devlin 1988) was applied to indicate the overall shape of the concentration-depth relationship. The LOWESS smooth is a non-parametric tool for exploratory data analysis, which is preferred in data analysis of water resources data because it is robust and insensitive to outliers in the data (Helsel and Hirsch 1992). Here, the LOWESS smooth was used to indicate the centre of the data scatter.

Intensive livestock farming - recharge

Figure 5.6 presents the concentration-depth profiles and LOWESS smooths for the homogeneous area *intensive livestock farming - recharge*. A striking difference appears for the concentration-depth relations of nitrate, aluminum and potassium, compared with OXC, chloride and SUMCAT. Nitrate, aluminum and potassium show a sharp decrease of concentrations with depth, whereas OXC, chloride and SUMCAT show a more gradual decrease with depth. For OXC and SUMCAT, the shape of the LOWESS smooths indicates a concentration maximum at about 8 m depth.

The median concentrations of OXC, chloride and SUMCAT over the whole depth range

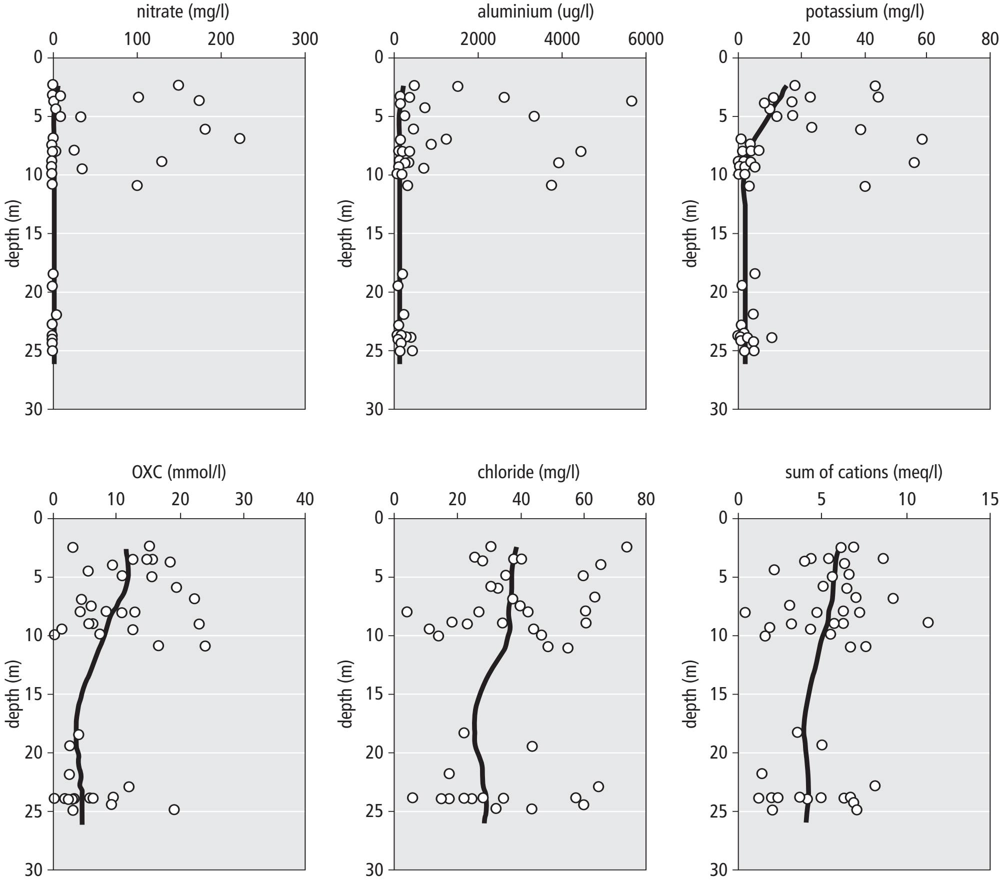


Figure 5.7 - Concentration-depth relations for the homogeneous area intensive livestock farming-intermediate. Data from 20 wells, averaged for the 1995-1998 monitoring period. The LOWESS smooth indicates the centre of the scatter (span = 0.75)

are much higher than under forest areas in similar hydrologic situations (Chapter 3). This indicates that these indicators are mainly introduced by agricultural practices. Therefore, the profiles of chloride, OXC and SUMCAT indicate that agriculturally polluted groundwater already reached to the depth of the deeper monitoring screens. Concentrations in the shallower groundwater are even higher. According to the determined age-depth relations of Figure 5.5, the groundwater at the deep screens is approximately 25 to 40 years old. The elevated concentrations of OXC, SUMCAT and chloride should thus be attributed to agricultural inputs during the period 1958-1973. Similarly, the concentration maximum at about 8 m depth corresponds to the recharge period 1982-1990.

Elevated concentrations of nitrate, aluminum and potassium are limited to the upper parts of the profiles. Clearly, the contamination fronts of these components have not yet reached deeper layers, whereas the conditionally conservative indicators show that the agricultural pollution front already has. Comparing the concentration-depth profiles of the targeted contaminants with the conditionally conservative indicators, we conclude that the pollution fronts of nitrate, aluminum and potassium have been retarded. Probably, different reaction mechanisms explained retardation of the three targeted components. The nitrate front is retarded by reduction processes in the aquifer (see also Chapter 3). The similar shapes of the OXC and the SUMCAT fronts suggests that pyrite oxidation is a probable reduction process for nitrate. The aluminum front is probably retarded by cation-exchange with calcium and magnesium. Aluminum exchange fronts in groundwater depth profiles have earlier been identified by for example Dahmke et al. (1986) and Broers & Griffioen (1993). The most probable mechanism for potassium retardation is cation-exchange for calcium and magnesium (Griffioen 2001, see also section 5.7).

Intensive livestock farming - intermediate

Figure 5.7 presents the concentration depth profiles for *intensive livestock farming-intermediate*. In this area, no clear contamination fronts of nitrate and aluminum are encountered; the centre of the data scatter is below 5 mg l^{-1} nitrate and 200 ug l^{-1} aluminum, respectively, according to the LOWESS smooths. Nevertheless, a substantial number of data points indicate higher concentrations of nitrate and aluminum in specific wells. Less variation is observed for potassium; the LOWESS smooth indicates a clear potassium front above 7 m depth. The front is found at shallower depth than in the recharge areas.

For OXC, chloride and SUMCAT, a larger variation is found than in the recharge areas, which is caused by a larger number of low concentrations. The large number of low concentrations agrees well with the large number of screens with pre-1950 groundwater in the area (Table 5.2). We explain the larger variations in OXC, chloride and SUMCAT by the larger variations of groundwater age at the monitoring depths (Chapter 2, see Figure 5.5b).

The LOWESS smooth indicates that the concentration maximum of OXC is found at shallower depth than in the recharge areas. However, the shape of the LOWESS smooth suggests the presence of an OXC pollution front. The shape of the LOWESS smooths for chloride and SUMCAT shows a very gradual decrease with depth and the absence of concentration maximums. The concentration-depth profiles for chloride and SUMCAT indicate that pollution fronts have arrived at the deeper level, at least in part of the area.

5.5 Trend analysis of time series data

Trend analysis of time series data was applied to detect temporal changes in groundwater quality in the homogeneous areas. First, trends were detected for time series data of individual monitoring screens. Second, the trends have been aggregated for homogeneous areas.

Trend analysis on individual time series

Time series of the annual data over the period 1984-1998 were available for the national monitoring wells. Time series for the provincial wells cover the period 1991-1998 or 1992-1998. The time series were checked for changes in the detection limits and the sampling and analysis procedures. Detection limits were adapted to the value of the highest limit in the time series. Significant trends in the time series of the individual well screens were identified using the non-parametric Mann-Kendall trend test (Appendix III.7). This test is robust and rather insensitive to outliers in the time series, compared with ordinary linear regression. The test is recommended for use in large data sets where the normality assumption cannot be checked for all individual time series (Helsel & Hirsch 1992). The Mann-Kendall procedure tests for monotonic trends in the data. No corrections for seasonal fluctuations were made, because of the annual monitoring frequency and the time-averaging that was introduced by sampling from the 2 m long screens (see also section 5.3).

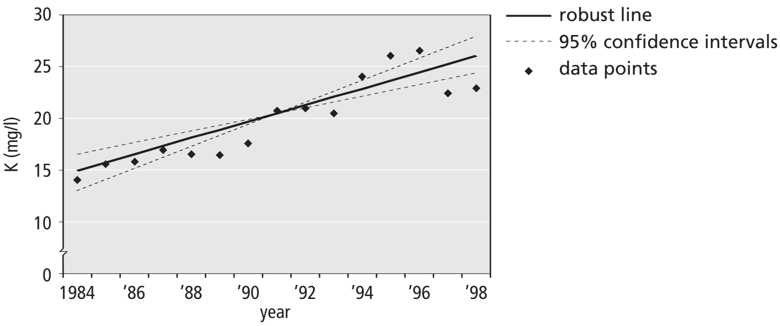


Figure 5.8 - Kendall-Theil robust line and corresponding non-parametric confidence intervals for the time series of OXC of well 125 at 7.5 m depth. Trend is significant at < 0.05, two-sided. Kendall-correlation coefficient equals 0.79.

The trend slope and 95% confidence intervals were computed using the Kendall-Theil robust line method (Helsel & Hirsch 1992, Appendix III.7). Additionally, the Kendall τ correlation coefficient was determined for each time series, which is a robust rank-based measure for correlation in the time series and resistant to the effect of a small number of unusual values (Appendix III.7).

Figure 5.8 shows the Kendall-Theil robust line and the corresponding 95% confidence intervals for the potassium concentrations in the shallow screen of well 107. For each individual screen, the concentration change over the period 1992-1998 was determined using the Kendall-Theil median slope. Table 5.3 lists the Kendall-Theil concentration changes over 1992-1998 and the Kendall τ correlation coefficients for the homogeneous area *intensive livestock farming-recharge* for potassium, nitrate, OXC and SUMCAT. Trends that were significant at α <0.05 are indicated in the table.

For *intensive livestock farming - recharge*, a large number of significant trends were found for OXC (12), SUMCAT (10) and potassium (11) out of 31 time series. For nitrate, only 2 significant trends were detected.

Aggregation of trends per homogeneous area

Trends were aggregated for three depth intervals: *phreatic* (3-6 m), *shallow* (7-13 m) and *deep* (18-25 m). For each interval, the median concentration change over 1992-1998 and the median Kendall τ correlation coefficients were calculated together with their non-parametric two-sided 95% confidence interval. The confidence intervals of the medians were obtained following the procedure explained in Appendix III.1. The results are listed in Table 5.3.

An aggregated trend was considered significant if both the 95% confidence intervals of the concentration changes and the 95% confidence intervals of the correlation coefficients were completely above or completely below 0. This procedure is similar to the procedure used for the national monitoring network (Van Drecht et al. 1996) except that non-parametric methods were used for the determination of the concentration change, the correlation coefficient and the confidence intervals. Furthermore, Kendall -Theil slopes were used to estimate the concentration changes (Appendix III.7) instead of subtracting the first and last concentrations in the time series (Van Drecht at al. 1996). Consequently, the presented procedure is more robust and less sensitive to outliers in the time series data.

Following Table 5.3, significant increasing trends over 1992-1998 for *intensive livestock farming-recharge* were detected for potassium (shallow and deep interval), OXC and SUMCAT (deep

Table 5.3 - Kendall- τ correlation coefficients and Kendall-Theil slopes (concentration change over 1992-1998) for potassium, nitrate, OXC and SUMCAT for individual times series in the wells of the area intensive livestock farming recharge (Significant trends bold italic).

Well	Depth	No. obs.	K τ	K slope (mg l^{-1})	NO3 τ	NO3 slope (mg l^{-1})	OXC τ	OXC slope (mmol l^{-1})	Sumcat τ	Sumcat slope (meq l^{-1})
Phreatic										
1806	5	7	0.38	2.3	0.10	0.8	-0.52	- 3.5	***-0.71***	***-1.7***
1810	4	8	-0.07	0.0	-0.29	- 33.8	***-0.79***	***- 3.5***	-0.50	-1.3
1843	4	8	-0.25	-1.5	-0.50	-183.8	-0.50	-14.2	-0.50	-2.2
1853	5.5	7	0.24	7.0	0.14	83.0	0.14	4.9	-0.05	-0.0
1863	3.5	7	0.24	3.0	-0.33	-131.7	-0.24	-11.3	-0.43	-2.6
95% lower			-0.25	-1.5	-0.50	-183	-0.79	-14.2	-0.71	-2.6
median			0.24	2.3	-0.29	- 34	-0.50	- 3.5	***-0.50***	***-1.7***
95% upper			0.38	7.0	0.14	82	0.14	4.8	-0.05	-0.0
Shallow										
107	7.5	15	***0.79***	***4.8***	0.05	4.3	-0.16	-0.9	***-0.42***	***-0.4***
108	8	8	***0.61***	***1.9***	***-0.93***	***-103.2***	***-0.79***	***-5.9***	***-0.36***	***-0.4***
122	9	15	***0.70***	***6.8***	-0.10	- 27.4	-0.30	-4.1	-0.32	-1.3
125	12	15	0.24	0.9	***0.87***	***55.8***	***0.77***	***4.1***	***0.76***	***0.7***
1810	12.5	7	-0.43	-0.4	-0.76	- 4.2	-0.52	-0.7	-0.05	-0.2
1823	8	7	0.10	3.0	-0.24	- 30.0	-0.24	-4.1	-0.05	-0.2
1833	8	8	0.39	30.9	0.18	13.4	-0.29	-1.1	0.07	0.7
1840	10	8	***0.79***	***15.5***	-0.25	0.0	0.32	5.5	-0.14	-0.4
1843	8	8	0.39	5.1	-0.18	- 13.7	0.00	0.3	0.14	0.2
1851	9	7	***0.90***	***1.5***	0.00	0.0	-0.33	-2.9	***0.71***	***3.0***
1853	10	7	0.67	1.2	-0.05	0.0	0.24	1.1	0.14	0.2
1863	10	7	0.52	1.2	-0.52	- 1.2	***0.71***	***5.4***	0.62	0.6
1866	9	7	0.33	1.5	0.00	0.0	***0.71***	***1.0***	-0.05	0.0
95% lower			0.24	1.2	-0.52	-27	-0.33	-4.1	-0.32	-0.43
median			***0.52***	***1.9***	-0.1	0.0	-0.16	-0.7	-0.05	-0.02
95% upper			0.79	6.8	0.05	4.2	0.71	4.1	0.62	0.69
Deep										
107	24	15	***0.47***	***0.2***	-0.13	0.0	***0.43***	***0.3***	***0.84***	***0.5***
108	24	15	***0.69***	***0.4***	-0.13	0.0	***0.57***	***4.1***	***0.50***	***1.4***
122	24	15	***0.66***	***5.4***	-0.13	0.0	0.16	1.8	0.16	0.7
125	23	15	***0.60***	***0.5***	0.00	0.0	***0.45***	***1.7***	***0.41***	***0.5***
1806	23	7	0.43	0.8	-0.10	0.0	***0.86***	***3.3***	***0.81***	***1.4***
1823	18	7	-0.13	-9.0	-0.47	-0.4	0.27	1.5	0.33	0.6
1833	19	8	***0.79***	***4.8***	0.00	0.0	0.21	0.6	0.21	0.6
1840	24.5	8	***0.75***	***0.4***	0.00	0.0	***0.82***	***2.6***	***0.71***	***0.8***
1843	23	8	0.25	0.8	-0.04	-0.5	-0.14	0.0	0.21	0.4
1851	23	7	0.19	1.2	0.00	0.0	0.48	1.9	0.24	0.6
1853	24	7	-0.38	-0.4	0.00	0.0	***-0.62***	***-1.8***	-0.52	-0.4
1863	22	7	0.10	0.1	-0.14	0.0	0.00	0.0	-0.43	-0.4
1866	23	7	0.19	0.6	0.00	0.0	***0.76***	***7.0***	0.14	1.2
95% lower			0.10	0.15	-0.13	0.0	0.0	0.0	0.14	0.38
median			***0.43***	***0.49***	-0.04	0.0	***0.43***	***1.7***	***0.24***	***0.61***
95% upper			0.69	1.2	0.0	0.0	0.76	3.3	0.71	1.17

Table 5.4 - Median Kendall-Theil concentration changes over 1992-1998 and median Kendall- τ correlation coefficients for K, NO3 , OXC and SUMCAT, aggregated for the area intensive livestock farming - intermediate (Significant trends bold italic).

	K		NO_3		OXC		Sumcat	
Well	**τ**	**slope** (mg l^{-1})	**τ**	**slope** (mg l^{-1})	**τ**	**slope** (mmol l^{-1})	**τ**	**slope** (meq l^{-1})
Phreatic (n=11)								
95% lower	-0.36	-3.0	-0.14	- 7.8	-0.14	-0.3	-0.14	-0.3
median	0.38	4.0	0.07	0.0	0.14	1.6	0.24	1.3
95% upper	0.54	6.7	0.14	22.5	0.43	5.1	0.36	1.7
Shallow (n=18)								
95% lower	-0.05	0.0	-0.10	0.0	-0.32	-3.8	-0.30	-1.2
median	0.10	0.1	0.00	0.0	-0.06	-0.1	-0.06	-0.2
95% upper	0.52	0.8	0.07	0.0	0.14	0.6	0.21	0.3
***Deep** (n=15)*								
95% lower	0.11	0.0	0.00	0.0	0.29	0.2	0.14	0.1
median	***0.43***	***0.3***	0.00	0.0	***0.50***	***0.6***	***0.47***	***0.4***
95% upper	0.61	0.9	0.00	0.0	0.64	2.6	0.72	1.1

interval). A significant decreasing trend was detected for SUMCAT at the phreatic interval.

Significant trends in the area *intensive livestock farming-intermediate* were only found for the deep interval. Potassium, OXC and SUMCAT showed increasing concentrations over the period 1992-1998 (Table 5.4).

Combining concentration-depth profiles and time series analysis

Figure 5.9 provides an overview of the indications for temporal changes in groundwater quality for the 4 indicators in the area *il-r*. The figure shows the LOWESS smooths for the monitoring periods 1995-1998 and 1991-1994. Arrows indicate the aggregated significant trends that were derived from time series analysis (see Tables 5.3 and 5.4).

For nitrate, the LOWESS smooths indicate decreasing concentrations at the phreatic and the shallow depth interval (Figure 5.9a). However, these changes were not demonstrated using time series analysis because large short-term temporal variations were present in the time series (Table 5.3). For potassium, the general small increases indicated by the LOWESS smooth, especially between 7 and 12 m depth, has been demonstrated using time series analysis. A median increase of 6 mg l^{-1} over the period 1992-1998 was detected. This increase was interpreted as the slow downward movement of the retarded potassium front. For OXC, a median increase of 0.5 mmol l^{-1} was detected at larger depth. This increase is attributed to slow downward movement of the OXC front, which is probably transported conservatively. The LOWESS smooths indicate a concentration decrease for OXC above 7 m depth. This decrease was not detected using time series analysis. For SUMCAT however, a corresponding decrease was demonstrated together with the increasing trend at larger depth.

Concluding, both OXC and SUMCAT present indications of a general increase at larger depth and a concentration decrease above 7 m depth. The increase at larger depth is probably due to the conservatively moving downward movement of the agricultural pollution front. The decrease in the upper parts is probably caused by the decreasing use of manure in the recent agricultural practices after 1986. This hypothesis will further be investigated in the next section using a conservative prognosis for these indicators.

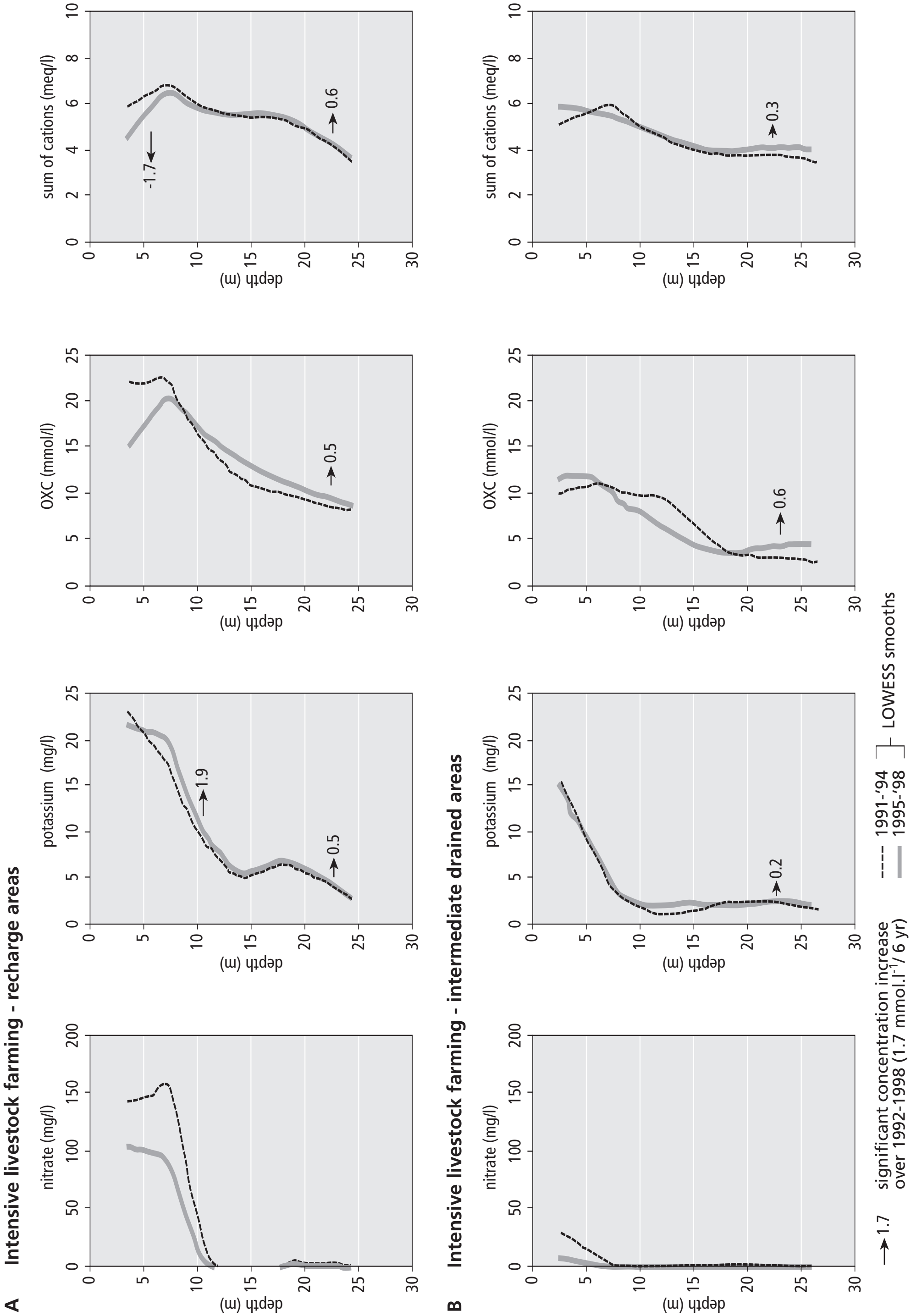


Figure 5.9 - Significant aggregated trends from time series analysis (arrows) and LOWESS smooths for the periods 1991-1994 and 1995-1998 for homogeneous areas il-r (A) and il-i (B)

Concentration changes in the area *intensive livestock farming-intermediate* are less pronounced. A 5-10% increase of concentrations was demonstrated between 18 and 25 m depth for potassium, OXC and SUMCAT (Figure 5.9b). According to the age-depth relations, the groundwater at this depth interval is at least 30 years old and changes would reflect the period before 1968. For OXC and SUMCAT, the increase suggests the first arrival of the agricultural pollution front at this depth. The demonstrated increase for potassium is not easily explained, because of the apparent absence of a concentration gradient with depth.

Overall, the trends in this areas are less pronounced than in the recharge areas. This is explained by the larger spatial variability which is present in the area (according to reason 3 of section 5.1). Part of this variation is due to variations in groundwater age in the intermediate areas (section 5.3 and Chapter 2).

5.6 Prognoses for conservative transport

Up to now, the monitoring data has been evaluated without applying models for contaminant transport. Only tritium data were used to acquire indications of groundwater ages at several depth that support the interpretation of the data. However, the data analysis and the understanding of the changes in groundwater quality can benefit from the use of simple models for conservative and reactive transport based on information on the historical input of solutes in recharging groundwater.

The aim of the application of the conservative prognoses is to explain and understand the temporal changes of the chemical components. Intentionally, attempts were not made to create a better fit between the prognosis and the measurements. Instead, deviations of measurements

Table 5.5 - Sources and sinks of nitrogen, sulfur, potassium, calcium and magnesium for grassland (g), arable land (a) and maize land (m) for the year 1995. Sources of information and empirical relations are indicated below the table.

	N (kg ha^{-1} yr^{-1})			S (kg ha^{-1} yr^{-1})			K (kg ha^{-1} yr^{-1})			Ca (kg ha^{-1} yr^{-1})			Mg (kg ha^{-1} yr^{-1})		
Wet deposition[1]	15.6			6.5			1.6			6.8			2.7		
Dry deposition[1]	38.5			13			1.6			6.8			2.7		
Lime[2]										59			7		
	g	a	m	g	a	m	g	a	m	g	a	m	g	a	m
Fertiliser[3]	290	100	100	4	5	5	11	45	45	76	29	29	4	3	3
Animal manure[1,2,4,5]	352	250	614	90	50	90	389	230	434	93	78	220	48	30	73
Crop uptake & harvesting[1]	480[6]	277[6]	292[6]	48	23	20	277	180	172	68	37	12	14	27	9
Surplus	216	127	476	66	51	95	126	98	310	173	143	310	51	18	79

[1] Stuyfzand (1991), van der Grift & Van Beek (1996), Bleeker & Erisman (1996), Mylona (1996)
[2] Menke (1992)
[3] Pronk (1994)
[4] CBS (1998, 1999)
[5] Statline (1999)
[6] N balance of SPREAD (Beekman 1998)

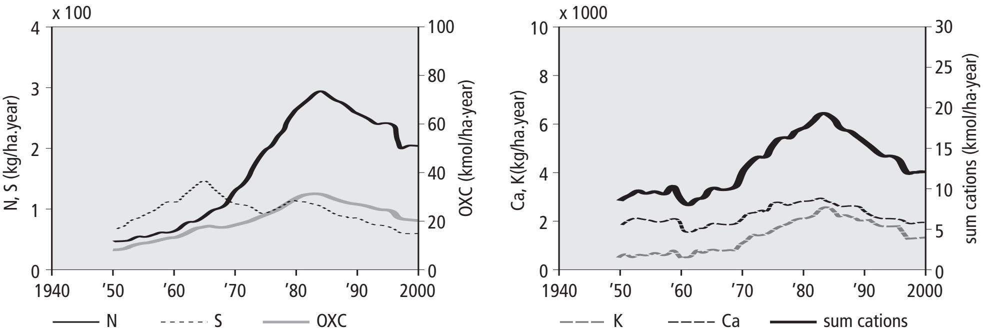


Figure 5.10 - Mineral surplus for nitrogen, sulphur, potassium and calcium and derived surplus of OXC and SUMCAT for areas with intensive livestock farming in Noord-Brabant between 1940 and 2000.

and prognoses are used as indications for the occurrence of reactive processes. This approach is comparable to approaches used in trend analysis for groundwater heads in the vicinity of well fields, where time series results were compared with simple models that describe the head drawdown (van Geer et al. 1987, Broers et al. 1991).

The historical inputs of nitrogen, sulphur, potassium, calcium and magnesium were estimated for the period 1950-2000. The input includes the following sources: atmospheric deposition, fertilizer salts, lime and animal manure. Table 5.5 summarizes the average input from those sources in the intensive livestock farming areas for the year 1995 and indicates the sources of information used. The table indicates that animal manure is the main source of all species. Fertilizer is an important source for nitrogen and calcium as well. The animal manure was distributed across grassland, arable and maize land according to Dutch agricultural advice (Menke 1992). This includes a favored order of the use of manure produced by cattle, pigs and poultry on grassland, arable land and maize land and limitations for amounts of animal manure for grassland (maximum 400 kg K_2O ha^{-1} yr^{-1}) and arable land (maximum 250 kg N ha^{-1} yr^{-1}). The surplus of manure was distributed on maize land, because maize is able to tolerate high inputs of nitrogen. Crop uptake and harvesting was calculated using records on the production of crops and hay and average mineral contents of the crops (de Jong et al. 1986, Griffioen & Keijzer 2001). Crop uptake and harvesting of nitrogen was calculated using SPREAD (Beekman (1998). The surplus amounts of Table 5.5 were used for the prognoses.

Figure 5.10 shows the resulting surplus of nitrogen, sulphur, calcium and potassium for the two homogeneous areas with intensive livestock farming in Noord-Brabant. The resulting concentrations of OXC and SUMCAT are also given. The figures show overall increasing loads for all indicators between 1950 and 1985. A large increase especially occurred since 1970, corresponding with the expansion of intensive livestock farming that caused a large increase in the production of animal manure. Simultaneously, the amount of maize land increased because of the introduction of new maize varieties that were adapted to Dutch climatological circumstances. Since 1985, an overall decrease of the mineral surplus is observed, due to the regulations on the use of manure that aim at the reduction of livestock and the leaching of minerals. For example, since 1985 there have been restrictions on the application of manure in the winter months.

The nitrogen surplus of Figure 5.10 was further corrected for denitrification in the unsaturated zone, using denitrification factors that are based on the average groundwater level

Table 5.6 - Parameters for conservative prognoses

	Recharge areas	Intermediate areas
Porosity	0.35	0.35
Denitrification factor unsaturated zone*	0.7	0.5
Median prognosis		
Groundwater recharge (m yr^{-1})	0.27	0.22
Aquifer thickness (D)	80	90
P10 prognosis		
Groundwater recharge (m yr^{-1})	0.18	0.09
Aquifer thickness (D)	50	50
P90 prognosis		
Groundwater recharge (m yr^{-1})	0.35	0.29
Aquifer thickness (D)	150	150

* reduction of N losses in the unsaturated zone (after Steenvoorden 1987, Boumans & Fraters 1995)

regime in the areas (Steenvoorden et al. 1986, Steenvoorden 1987, Van Drecht 1993, Boumans & Fraters 1995). Denitrification in the intermediate, drained areas is higher than in the recharge areas because of the more shallow groundwater levels (see also Chapter 2). Table 5.6 summarizes the data that were used for the concentration-depth prognosis in the recharge areas and the intermediate drained areas.

The conservative transport of the mineral surplus into the subsurface was modeled using the age-depth relationships that were established in section 5.2, using the median and 10 and 90 percentile age distributions of Figure 5.5. The conservative prognoses for nitrate, sulphate, OXC and SUMCAT in the two homogeneous areas are presented in Figure 5.11a and b. The thick line indicates the median prognosis for 1997 and the two dashed lines reflect the uncertainty of the prognoses using the 10 and 90 percentiles of the age-depth relations. These lines only indicate the part of the uncertainty that arises from the age-depth assessment and not the uncertainty in the solute inputs. The LOWESS smooths indicate the centre of the measured concentrations over the period 1995-1998.

Intensive livestock farming - recharge

The prognosis for nitrate shows a large misfit with the measured data (Figure 5.11). Contrary, an overall agreement is present for SUMCAT, and partial agreement for sulphate and OXC. Details are discussed below.

Sum of cations was considered as the most conservative indicator that was investigated. The overall shape of the LOWESS smooth and the conservative prognosis for SUMCAT agree well and the modelled and measured concentrations are of the same order of magnitude. The conservative prognosis was used to evaluate the transport of the SUMCAT front between 1986 and 1998 (Figure 5.12). The figure shows the conservatively downward transport of the 1985 manure peak and the resulting concentration changes between 1992 and 1998. Concentrations are expected to increase below the 1985 concentration maximum and to decrease above it. Indeed, the trends actually observed for SUMCAT indicate an increase of concentrations between 18 and 25 m depth, and a decrease between 3 and 6 m depth. No trends were found for the interval 7-13 m depth, which corresponds with the opposed directions of concentration changes above and below the infiltrating SUMCAT peak. Following Figure 5.12, a future

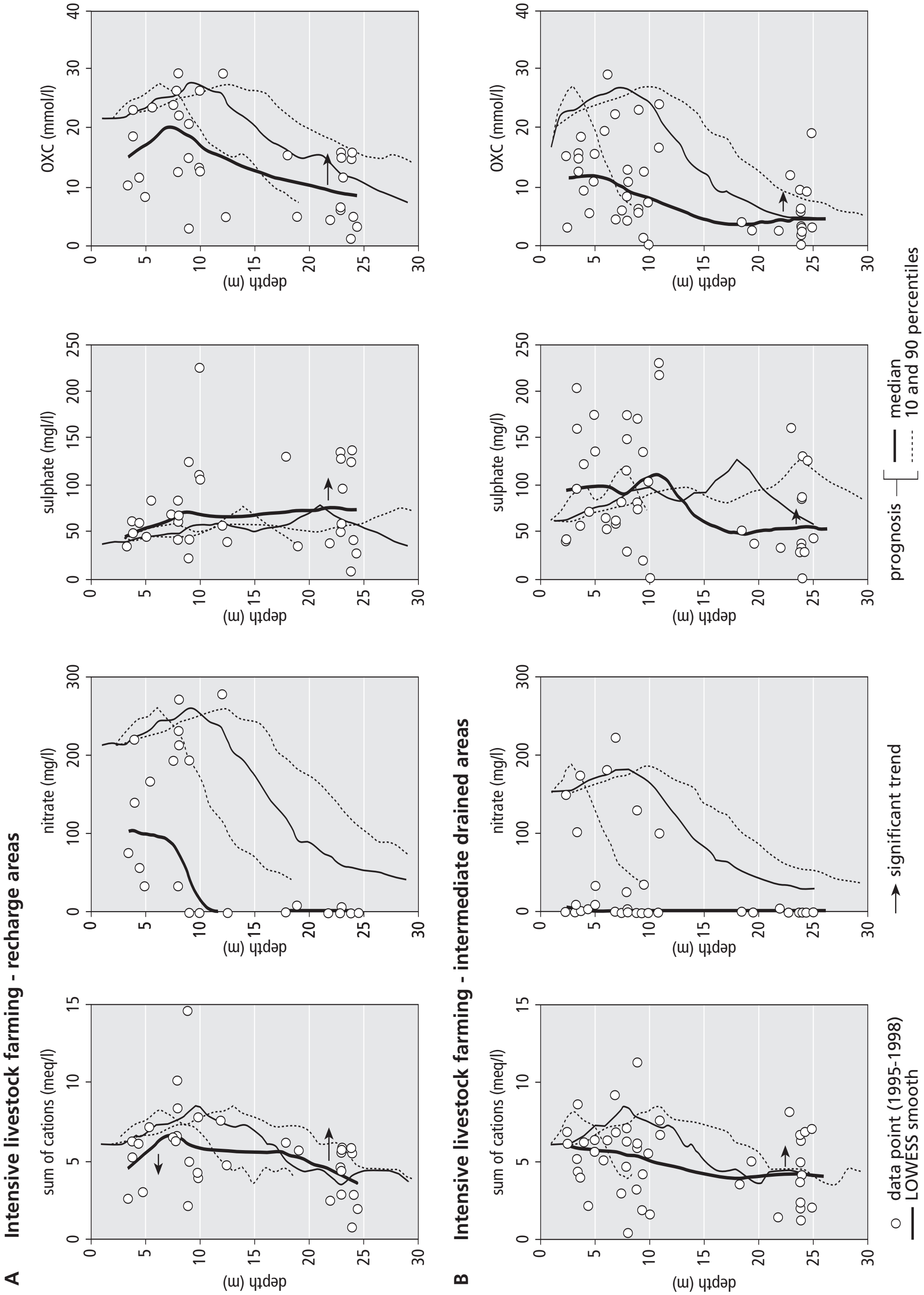


Figure 5.11 - Measured concentrations, LOWESS smooths and significant trends over the period 1995-1998 and conservative prognoses for the year 1997 for SUMCAT, nitrate, sulphate and OXC in the homogeneous areas il-r (upper row) and il-i (lower row)

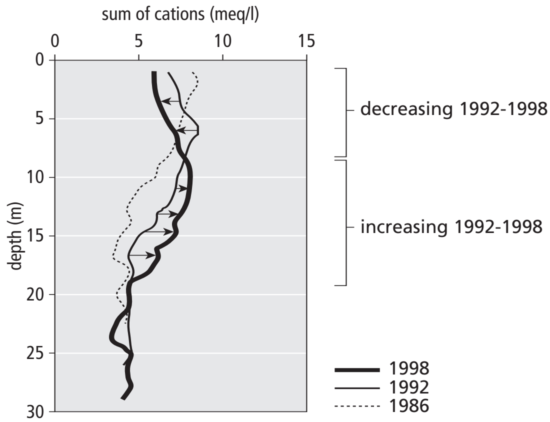


Figure 5.12 - Model reconstruction of the downward movement of the conservative transported front of SUMCAT between 1986 and 1998 based on the conservative prognosis. Arrows indicate the resulting trends over the period 1992-1998.

concentration increase is expected for the regular monitoring interval 18-25 m depth for indicators that behave conservatively.

The example suggests that concentration changes in SUMCAT are useful to acquire information on the movement of the agricultural pollution fronts. The use of SUMCAT is limited to areas were mineral dissolution or precipitation processes do not occur. SUMCAT is less useful in areas were carbonate dissolution is frequent, because it would increase the overall dissolved solutes. In such areas, the assumption of conservative transport of SUMCAT is not valid.

The concentration-depth profiles of nitrate, sulphate and OXC are discussed simultaneously to unravel the geochemical processes and to understand the observed trends over the period 1992-1998. For nitrate, the prognosis predicts much higher concentrations and a larger depth of the 1985 concentration maximum. At 20 m depth, nitrate concentrations between 50 and 200 mg l^{-1} would result from conservative transport, whereas the actual concentrations are below the detection limit. This strongly indicates the occurrence of attenuating processes for nitrate. Pyrite oxidation is a probable reaction mechanism for the disappearance of nitrate with depth, because pyrite is abundant in Noord-Brabant sediments (Postma et al. 1991, see Chapter 3). The hypothesis is plausible, given the large number of sulphate concentrations above 100 mg l^{-1} (Figure 5.11a). These concentrations cannot be explained by agricultural inputs, which are indicated by the conservative prognoses for sulphate.

The OXC is transported conservatively under the condition that pyrite oxidation is the only process that occurs in the aquifer. The conservative OXC prognosis is consistent with the higher part of the measured OXC concentrations and the general shape of the LOWESS smooth and the prognosis are similar. Both the LOWESS smooth and the prognosis indicate an OXC peak, which reflects the 1985 maximum in the OXC input. Because the upper half of the observations is quite well predicted for OXC, we conclude that pyrite oxidation is the dominant mechanism for this part of the data.

However, the OXC prognosis overestimates the median concentrations that are indicated by the LOWESS smooth. Thus, pyrite oxidation cannot explain the complete range of measured OXC concentrations. Two reasons are possible for the discrepancy between the lower part of the concentrations and the prognoses: (1) other reactions lower the concentrations of nitrate or sulphate, and (2) the historical input of N and S is overestimated. The most probable reaction that lowers the OXC is nitrate reduction by organic matter. Contrary to the reduction by sulphides, the reduction by organic matter does lower the OXC because no sulphate is formed.

$$5CH_2O + 4NO_3^- \rightarrow 4HCO_3^- + CO_2 + 3H_2O + 2N_2 \qquad (5.4)$$

For the lower part of the data, nitrate reduction with organic matter is apparently relevant. The denitrification happens in the unsaturated zone or deeper in the subsurface. Since no unambiguous reaction products are formed during denitrification, further analysis requires models for reactive transport.

Concluding, we attribute the increasing trends for OXC and sulphate at deeper level to the part of the data where nitrate is reduced by pyrite oxidation. A further increase of OXC and sulphate concentrations is expected, as the 1985 concentration maximum has not yet reached the deeper monitoring screens.

Intensive livestock farming-intermediate

For this area, the diverging median, 10 and 90 percentile prognoses reflect the observed large variations in groundwater age and the corresponding variations in advective vertical transport (Figure 5.11b). For example, the depth of the peak of the conservative prognoses of nitrate varies between 3 and 10 m.

Still, a reasonable correspondence between the prognoses and measured concentrations is found for SUMCAT. The increasing trend at larger depth is attributed to the wells in areas with the largest vertical groundwater velocities and indicates the first arrival of the agricultural pollution front. However, the high SUMCAT concentrations in some wells at about 25 m depth might have been caused by dissolution of siderite, given the high proportion of siderite-saturated groundwater at this depth (Chapter 3).

A smaller input of N was taken for the intermediate areas to account for the shallower groundwater levels, using an unsaturated zone denitrification factor of 0.5 (Table 5.6). The older age of the groundwater and the smaller input of N results in lower predicted concentrations at similar depths and shallower infiltration of nitrate and OXC compared with the recharge areas (Figure 5.11). The high sulphate concentrations between 4 and 12 m depth indicate that pyrite oxidation occurs already at shallow depth in the aquifer. As for the recharge

Table 5.7 - Input of PHREEQC model for potassium transport

	K (mmol l^{-1})	Ca (mmol l^{-1})	Mg (mmol l^{-1})	X (meq l^{-1})
Initial solution	0.13	1.42	0.24	0.053
1950-1955	0.41	1.42	0.24	0.053
1955-1960	0.45	1.41	0.26	0.053
1960-1965	0.44	1.27	0.25	0.053
1965-1970	0.61	1.33	0.34	0.053
1970-1975	0.99	1.67	0.51	0.053
1975-1980	1.35	1.95	0.67	0.053
1980-1985	1.69	2.01	0.78	0.053
1985-1990	1.59	1.81	0.70	0.053
1990-1995	1.37	1.54	0.58	0.053
1995-2000	1.06	1.40	0.50	0.053

areas, the nitrate and OXC prognoses are consistent with highest part of the measured concentrations, but overestimate the centre of the data. Thus, the overall absence of nitrate in the groundwater cannot be completely explained by pyrite oxidation. Denitrification with organic matter is a probable extra mechanism for nitrate removal in part of the area. The significant trends in OXC and sulphate between 18 and 25 m depth are attributed to the downward movement of the agricultural pollution front in areas where pyrite oxidation is the dominant mechanism.

5.7 Prognoses for reactive transport

For many reactive chemical components, the observed trends are only understood using prognoses for reactive transport that evaluate the most probable reactions. Here, potassium was used to illustrate how a reactive prognosis supports the interpretation of the observed trends, using a 1D geochemical transport model. The transport of agriculturally derived potassium is dominated by cation exchange against calcium and magnesium at circumneutral pH (Griffioen 2001).

$$({}^{1}/{}_{2})\,Ca\text{-}X_2 + K^+ \leftrightarrow K\text{-}X + ({}^{1}/{}_{2})\,Ca^{2+} \qquad (5.5)$$

where X^- represents the exchanger with charge -1. Cation exchange effectively results in the retardation of the potassium front. This probably causes the large misfit between the conservative prognoses for potassium and the measured concentrations (Figure 5.13a).

The transport of potassium was modeled with PHREEQC (Parkhurst & Appelo 1999) using the average inputs of potassium, calcium and magnesium for five-year intervals over the period 1950-2000 (Table 5.7). In PHREEQC, the equilibrium between the exchanger and the aqueous solution is described following the Gaines-Thomas convention:

$$K_{GT} = \frac{\beta_K}{[K^+]}\sqrt{\frac{[Ca^{2+}]}{\beta_{Ca^{2+}}}} \qquad (5.6)$$

where K_{GT} is the Gaines-Thomas selectivity coefficient, $[i]$ is the aqueous activity of the

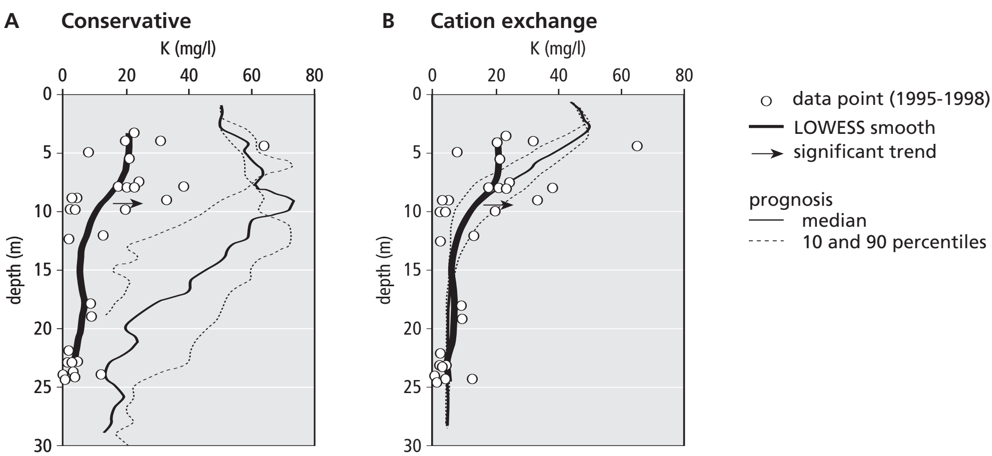


Figure 5.13 - Measured concentrations, LOWESS smooths and significant trends over the period 1995-1998 and the reactive prognosis for the year 1997 for potassium in the area il-r

cation and β_i is the adsorbed equivalent fraction of cation *i*. A Gaines-Thomas selectivity constant K_{GT}= 2.0 between K and Ca was assumed. A cation-exchange capacity (CEC) of 10 meq kg^{-1} was used as an average for fine sandy deposits in Noord-Brabant (Broers & Griffioen 1994). This corresponds to a concentration of X of 0.053 meq l^{-1}, assuming porosity = 0.3 and bulk density of 1.8 kg dm^{-3} (Table 5.7).

Figure 5.13b shows the resulting position of the potassium front for the year 1997. The reactive prognosis corresponds much better with the measured concentrations in this area than the conservative prognosis. The concentration maximum of 1985 is found at about 3.5 m depth, versus 10 m depth in the conservative prognoses. Elevated concentrations of potassium occur only in the first 15 m of the subsurface. Figure 5.14 shows the retarded transport of the potassium front between 1986 and 1998. The depth interval over which increasing concentrations are predicted over the period 1992-1998 is indicated. The depth of the observed concentration increase (7-12 m) corresponds well with the predicted depth.

Concluding, the observed trends in the time series indicate the retarded vertical movement of the potassium front that infiltrated between 1950 and 1985. Contrary to the conservatively transported indicators, the 1985 concentration maximum has not yet reached the average depth of the shallow screens. Consequently, increasing trends in the retarded potassium front are found at more shallow level than trends of the conservatively transported indicators.

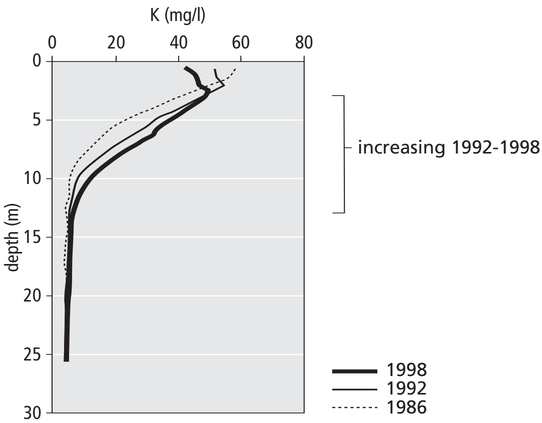


Figure 5.14 - Model reconstruction of the downward movement of the reactive front of potassium between 1986 and 1998 based on the reactive prognosis. Depth interval with increasing trends over the period 1992-1998 is indicated.

5.8 Discussion and conclusions

This study aimed at improving (1) the detection and (2) the understanding of groundwater quality changes in time in large areas with homogeneous land use, soil types and geohydrological situation. In the presented approach, a distinction is made in trend detection and trend understanding (Figure 5.2). The trend detection uses measured data of time series and time-averaged concentration-depth profiles. The interpretation of trends was aided by prognoses of conservative and reactive transport using tritium age dating and information on the historical loads of solutes in recharging groundwater. The two aspects are discussed separately, and implications for groundwater quality monitoring strategies are indicated briefly.

Trend detection

The complementary use of time-series information and time-averaged concentration-depth profiles yields additional value compared to the individual methods. Time series analysis alone might fail to detect groundwater quality changes reliably because non-conservative behaviour limits the breakthrough of the solutes. Concentration-depth profiles reveal this non-conservative behavior and indicate the position of contamination fronts. The detection of groundwater quality changes is supported by comparing significant trends in time series at the monitoring depths with the vertical concentration profiles.

Especially, the use of conditionally conservative indicators helps to identify groundwater quality changes, because their pollution front moves through the aquifer with groundwater velocity. The use of conditionally conservative indicators is advantageous because trends are detected earlier, directly after the advective travel time has passed. This was demonstrated for the two homogeneous areas, where many trends in the time series of OXC and SUMCAT were found, which indicate the downward movement of the agricultural pollution front. No temporal trends in nitrate have been detected, because nitrate disappears with depth and the pollution front of nitrate probably moves too slowly to be detected using the regular monitoring depths.

Understanding of observed trends

The understanding of quality changes is strongly supported by comparison of the position of the pollution fronts and the temporal trends from time series with prognoses of conservative and reactive transport. These prognoses require information on the historical inputs of the contaminants and an indication of the groundwater age.

An unknown uncertainty is inherent to the calculated historical inputs. The uncertainty mainly arises from the estimates of crop uptake, the distribution of the produced animal manure over the various land uses and spatial variations in the use of manure. However, the uncertainty is reduced by the fact that we used regionally averaged data on manure production, fertiliser use and atmospheric deposition for the whole province of Noord-Brabant. This results in the averaging out of local deviations in the mineral surplus. No large uncertainties in the order of magnitude and the period of increasing and decreasing inputs are expected for the investigated components. These two aspects are the most important for understanding the depths and magnitudes of temporal trends. The uncertainty in the groundwater age distribution was evaluated using the median, 10 and 90 percentiles of ages derived from the tritium measurements. This part of the uncertainty of the prognoses was visualised and used for the data analysis.

For the Noord-Brabant case, the prognoses for conditionally conservative indicators correspond much better with the measured concentrations than the prognoses for individual targeted contaminants, such as nitrate, sulphate and potassium. In the Noord-Brabant homogeneous areas, the downward movement of the agricultural pollution front was demonstrated by the increasing trends between 18 and 25 m depth for OXC and SUMCAT. The 1985 peak of OXC and SUMCAT has not yet arrived at the deeper screens and further increase of concentrations is expected. No trends for these indicators were found at 7-12 m depth, which is plausible because the contaminant peak was predicted to be within this interval, with decreasing concentrations above it and increasing concentrations below it.

The large misfit between the predicted conservative position of the nitrate and potassium front and the actual position of the fronts indicate retarding and/or attenuating processes in the aquifer. For potassium the retarded front was modelled using a simple 1D geochemical model of cation exchange between potassium and calcium and magnesium. The model indicates that the increasing concentrations of potassium between 7 and 12 m depth are well explained by the movement of a retarded potassium front. This indicates that concentration changes of the retarded species are expected and demonstrated at shallower depth than concentration changes of the conservatively transported indicators.

No significant quality changes have been demonstrated for nitrate, due to nitrate reduction by oxidation of pyrite and organic matter. The significant changes of OXC in the deep groundwater that lacks nitrate are due to increasing sulphate concentrations, which is a reaction product of the pyrite oxidation. A future increase in sulphate concentrations above the drinking water standard of 150 mg l^{-1} is anticipated, given the loads of nitrogen and sulphate that are still underway.

Implications for groundwater quality monitoring

Often, the design of groundwater quality monitoring networks implicitly assumes conservative transport, and sampling depths and monitoring frequency are adapted to advective transport velocities. In the Netherlands' monitoring networks, regular monitoring depths of approximately 10 and 25 m depth, and an annual frequency were chosen. Using data from these networks, we were able to demonstrate increasing concentrations at both depths for chemical indicators that behave conservatively. However, large concentration gradients with depth are present for many targeted contaminants, which indicate retarding or attenuating processes within the first 20 m of the subsurface. Groundwater quality changes of these components are expected to occur at more shallow depth. For potassium, this was indeed demonstrated in the study.

The concentration-depth profiles of the two homogeneous areas show a lack of data between 12 and 20 m depth, because the monitoring screens at this depth are not sampled frequently. This results in larger uncertainty of the LOWESS smooth at this depth. The concentration-depth profiles would clearly benefit from sampling these intermediate screens, because this would better define the vertical position of the contamination fronts (section 5.8). Since many concentration changes are expected between the two regular monitoring depths, one could consider to sample all available screens in the wells annually to obtain time series information. Alternatively, and cheaper, we advise the sampling of the complete concentration-depth profile every 4 years, including the shallow screens and the screens between 12 and 20 m depth. This would allow for a four yearly evaluation of the vertical propagation of the conservative and reactive contamination fronts.

General conclusions

The combined use of time series information and time-averaged concentration-depth profiles helps detect temporal changes in groundwater, especially when both potentially reactive and conditionally conservative indicators are used in the data analysis. Interpretation of the trend analysis results is aided by comparing the depths of the demonstrated temporal trends with prognoses of conservative and reactive transport, using information on the historical inputs of the contaminants and age dating using tritium.

6 Evaluating monitoring strategies for groundwater quality at phreatic well fields: a 3D travel time approach

6.1 Introduction

Phreatic well fields used for public water supply are vulnerable to groundwater contamination originating from diffuse sources, such as agricultural pesticides and nutrients. Usually protection zones are established to regulate the use of pesticides, fertiliser and manure and to protect the groundwater resource from calamities of other origin. The effectiveness of the groundwater protection and the threats of contaminated groundwater from outside the protection zone can be assessed using an adequate groundwater quality monitoring network. However, groundwater quality monitoring is expensive because it requires the frequent sampling and analysis of many samples and installation costs of wells are high. An effective monitoring strategy should therefore yield valuable information at justifiable costs.

Groundwater quality monitoring by water supply companies is done for five distinct objectives (after Baggelaar 1996):

1. to fulfil the legal monitoring necessity
2. to reassure customers
3. to signal unexpected or new threats to the quality of extracted groundwater
4. to support operational decisions by the prediction of future quality changes, and
5. to evaluate protection measures in the protection zone.

In the Netherlands the first two monitoring objectives are covered by the three-monthly analysis of the macro chemistry of the extracted groundwater at the pumping well and the monthly analysis on Coli-bacteria (Baggelaar, 1996). Especially the last three monitoring objectives require a specific observation network, apart from the pumping wells.

In the following the monitoring objectives 3, 4 and 5 will be referred to as *early warning*, *prediction* and *protection*, respectively. The design of an adequate monitoring network requires the evaluation of the hydrogeology of the area, the expected threats from agricultural or atmospheric origin and the possible hydrogeochemical and microbial processes that attenuate or worsen the groundwater quality evolution (Baggelaar 1992, Stuyfzand 1996).

Baggelaar (1992, 1996) designed an overall monitoring strategy for the Dutch water supply companies. For phreatic well fields the recommended monitoring configurations were:

1. the monitoring of the extracted groundwater at the pumping well
2. the monitoring of groundwater in the pumped aquifer that is 10 to 15 years travel time from the well, and
3. the monitoring of shallow groundwater in the protection zone and in a control area.

Several adaptations and improvements to this strategy have been proposed during the last five years (Baggelaar 1996, Zhang 1996, Foppen & Kremers 1997, Hetterschijt & Foppen 1998). None of these strategies have evaluated the 3D travel time distribution in the aquifer in order to design an appropriate network.

The objective of the present study was to judge the effectiveness of different network configurations, using a 3D travel time approach, evaluating advective and simple reactive transport. The selected monitoring configurations are relevant for typical Dutch conditions, which include shallow groundwater tables and unconsolidated aquifers in a flat landscape. The study concentrated on partially penetrating wells in an unconfined situation, since most phreatic well fields conform to this situation. Since the concern was the monitoring of the quality of the groundwater, the water flow and degradation processes in the unsaturated zone were neglected, although processes in the unsaturated zone may be important for the breakthrough of specific solutes such as pesticides at the pumping well (Beltman et al. 1996).

Numerical simulations of groundwater flow and travel times were made to evaluate seven monitoring configurations for pollution scenarios with advective transport, first-order degradation and linear sorption, using a block input. Existing concepts and analytical solutions for travel time and solute breakthrough of fully and partially penetrating wells are introduced first.

6.2 Analytical solutions for the travel time distribution and solute breakthrough

Residual transit times and residence times

An analytical solution for the travel time distribution in the saturated zone around a well that fully penetrates a phreatic aquifer of infinite dimensions is obtained (Bear & Verruijt 1987) by assuming that 1. the aquifer is homogeneous, isotropic and has constant thickness, 2. groundwater flow is steady, 3. the lowering of groundwater table due to pumping is small compared with the aquifer depth, and 4. the horizontal fluxes are independent of depth z (Dupuit assumption). The flux through an aquifer of thickness D at distance r of the well is:

$$Q = \pi.N(r_e^2 - r^2) \tag{6.1}$$

where Q = well discharge flux (m³ day⁻¹), N = groundwater recharge rate (m day⁻¹) and r_e = maximum radius of the catchment area (Figure 6.1a). Then, the radial flow velocity v_r at distance r of the well is (Van Ommen 1986):

$$v_r = -\frac{dr}{dt} = \frac{\pi.N(r_e^2 - r^2)}{2\pi.r\varepsilon D} = \frac{N}{\varepsilon D}\frac{(r_e^2 - r^2)}{2r} \tag{6.2}$$

where ε = effective porosity. Hence, the travel time for water recharging at distance r of the well towards the well equals:

$$t_r = -\int_0^r \frac{2\varepsilon D}{N}\frac{r}{r_e^2 - r^2}\,dr = -\frac{\varepsilon D}{N}\ln\left(1 - \frac{r^2}{r_e^2}\right) \tag{6.3}$$

Following equations (6.1) and (6.3), the travel time to the well is determined mainly by the thickness D of the aquifer and the ratio between well discharge flux Q and groundwater recharge rate N. The factor $\varepsilon D/N$ is also known as *characteristic travel time* (van Ommen 1986):

$$T = \frac{\varepsilon D}{N} \tag{6.4}$$

and

$$t_r = -T\ln\left(1 - \frac{r^2}{r_e^2}\right) \tag{6.5}$$

Equation (6.5) results in cylindrical patterns of isochrones of equal time t_r for a specific location in the aquifer to the pumping well. These travel times are referred to as *residual transit times* (t_ω) using the definition of Raats (1977). If the residual transit time represents the complete flow path from the water table to the pumping well, it is called *transit time*. Figure 6.1b shows the position of some t_ω-isochrones as fraction of the characteristic time T.

Raats also defined the *residence time* t_α that reflects the time passed from the water table to the present position in the aquifer. For the system studied, the residence time t_α was found by evaluating the distribution of the vertical flow velocity in the aquifer. Using the Dupuit

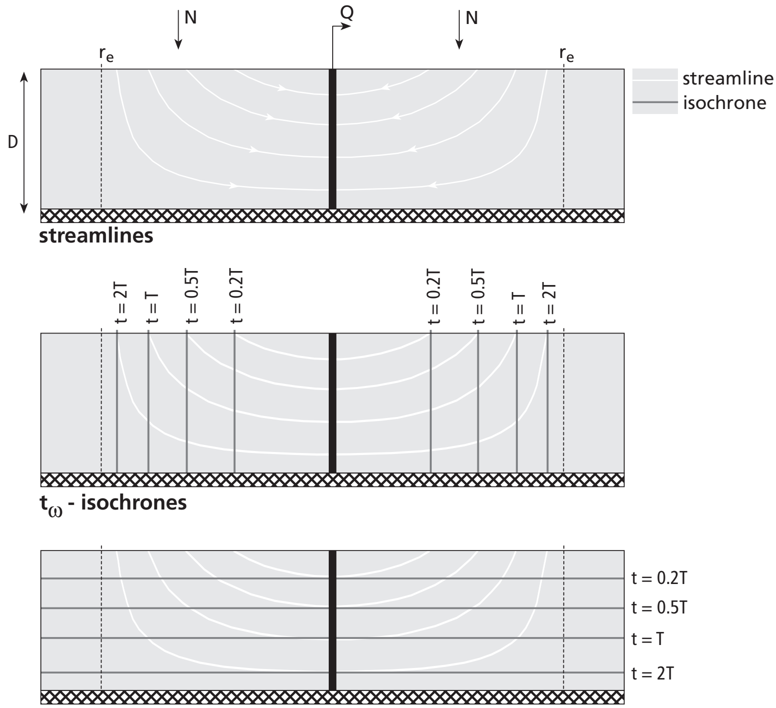


Figure 6.1 - Streamlines, isochrones of residual transit time (t_ω) and isochrones of residence time (t_α) for a fully penetrating well in an aquifer with constant transmissivity

assumption, the vertical groundwater velocity decreases linearly with depth according to (Raats 1977, Beltman 1995):

$$v_z = \frac{dz}{dt} = \frac{N}{\varepsilon}\left(\frac{D - z}{D}\right) \tag{6.6}$$

Accordingly, the travel time from the surface to a position at depth z equals (Raats 1977):

$$t_z = \frac{\varepsilon D}{N} \ln\left(\frac{D}{D - z}\right) = T \ln\left(\frac{D}{D - z}\right) \tag{6.7}$$

Equation (6.7) shows that the residence time distribution is determined by the depth of the aquifer D, the groundwater recharge rate N and the porosity ε. Note that the residence time at depth z is independent of the distance r from the well field and isochrones are horizontal over the contributing area of the well (Figure 6.1c). In the following, the lines of equal residence time t_α are referred to as t_α-isochrones according to Raats (1977). Figure 6.1c shows the position of some t_α-isochrones for fractions of characteristic time T. Equation (6.7) is equal to equation (2.1) that is valid for groundwater flow between two parallel ditches (Chapter 2).

Using equation (6.5) and defining the fraction f of extracted groundwater with *transit time* less than T_f, the next equation was derived for the cumulative transit time distribution of the pumping well:

$$T_f = T\ \ln\left(\frac{1}{1 - f_r}\right) \text{ with } f_r = \left(\frac{r}{r_e}\right)^2 \tag{6.8}$$

The cumulative transit time frequency distribution of equation (6.8) is also known as the response function of the pumping well, because the shape of the curve determines solute breakthrough in the pumping well, as shown below.

Solute breakthrough in the pumping well

An analytical solution for the breakthrough in the pumping well is obtained for a step input with concentration $C = C_{initial}$ for $t < 0$ and $C = C_{input}$ for $t \geq 0$ (van Ommen 1986):

$$\overline{C}_t = \frac{C_{input}\,\pi r^2 + C_{initial}\,\pi (r_e^2 - r^2)}{\pi r_e^2} \tag{6.9}$$

Rearranging yields:

$$\overline{C}_t = C_{initial} + (C_{input} - C_{initial})\ \frac{r^2}{r_e^2} \tag{6.10}$$

Combining equations (6.5) and (6.10) yields:

$$\overline{C}_t = C_{input} + (C_{initial} - C_{input})\ e^{-\frac{t_r}{T}} \tag{6.11}$$

The solute breakthrough for the step input is similar to that of an ideally mixed reservoir, although the system is based on piston displacement and determined by the flow pattern and travel time distribution shown in Figure 6.1 (Raats 1981, van Ommen 1986). Figure 6.2 illustrates how the concentration distribution in the aquifer and the concentration in the pumping well are interrelated, for $C_{initial} = 0$ and $C_{input} = 1$. The figure shows the concentration distribution in the aquifer for t = 0.5 T, t = T and t = 2T, based on the t_α-isochrones. The contamination front moves vertically through the aquifer, corresponding to the horizontal position of the t_α-isochrones. At a specific moment in time a specific fraction $1\text{-}e^{-t_r/T}$ of the aquifer is contaminated and contributes to the outflow. The concentration of the extracted groundwater is equal to the depth-averaged concentration of the aquifer. The shape of the concentration-time relationship of equation (6.11) equals the shape of the cumulative transit time frequency distribution of equation (6.8) for the step input. Analytic solutions for the step input and a Dirac-delta pulse input which include linear sorption and first-order degradation, were given by Beltman et al. (1995).

The presented analytical solutions assume a homogeneous subsoil and no dispersion. Duffy & Lee (1992) studied an aquifer-stream setting which is equal to the one depicted in Figure 2.3 (Chapter 2). They demonstrated that small-scale dispersion and aquifer hydraulic heterogeneity have negligible effects on the concentration response of the base flow of the stream for a non-point source of contamination. Given the equivalence of the concentration breakthrough in a fully penetrating pumping well (equation 6.11) and the concentration breakthrough in a stream, their conclusions are also valid here. The breakthrough is insensitive to heterogeneity if the correlation scale of the heterogeneities is smaller than 2.5% of the maximum horizontal travel distance, which in our case equals to r_e.

Partially penetrating wells

The Dupuit assumption is not valid in the surroundings of the well when the well is partially penetrating. To solve this problem, Van der Eem (1991) assessed minimal vertical transit times for partially penetrating wells, assuming that drawdown is negligible compared with aquifer thickness. He varied the depth and length of the pumping well screen and the well discharge.

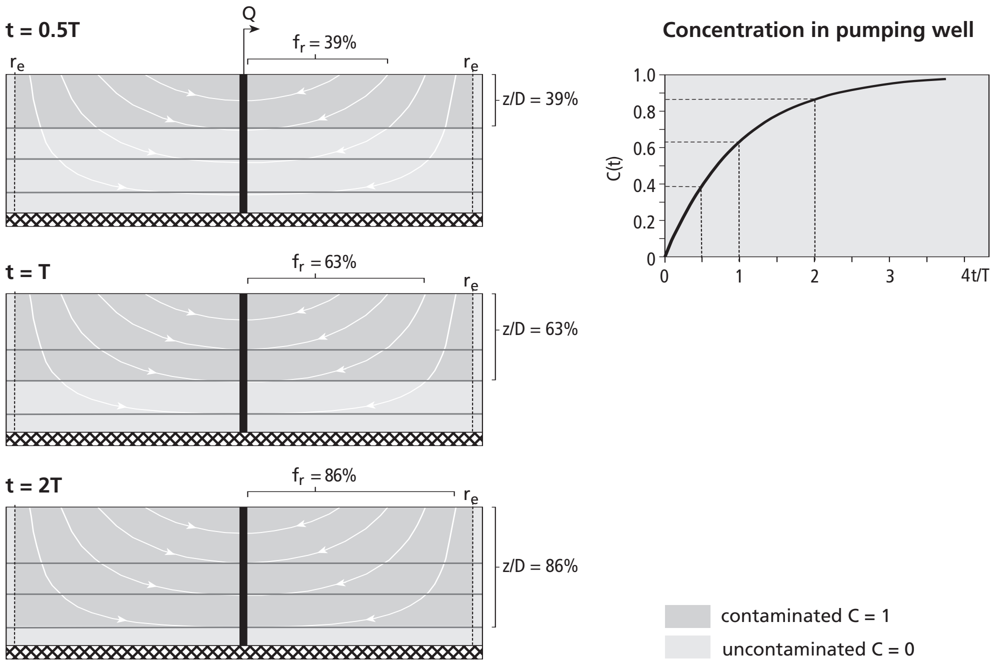


Figure 6.2 - Concentration distribution for a step input of C = 1 in a case of a fully penetrating well in an aquifer with constant transmissivity at t = 0.5 T, t = T and t = 2T and the resulting concentration breakthrough in the pumping well.

He concluded that the screen position has insignificant effects on the cumulative transit time frequency distribution, as long as the radius of the contributing area r_e is larger or equal to aquifer thickness D. Wells used for public drinking water supply always fulfil this condition and equation (6.5) is considered applicable to those wells. Consequently, vertical transit times at the pumping well location are negligible.

When the drawdown is large relative to the aquifer thickness D and the saturated aquifer thickness is reduced, a distortion of the horizontal pattern of t_α-isochrones should be expected in the vicinity of the well field. A simple numerical model was applied to assess the position of t_α and t_ω isochrones and to evaluate the monitoring strategies for this situation.

6.3 Methods

A numerical approach was used to assess the t_α and t_ω isochrone patterns and the solute breakthrough in observation and pumping wells. Some advantages of the numerical approach are:

1. any shape of concentration input is possible
2. heterogeneous systems may be considered, including spatially variable concentrations and hydraulic heterogeneity of the aquifer
3. the effects of reduced saturated aquifer thickness can be assessed.

Here, a relatively simple system was studied, using a block concentration input, a homogeneous subsoil and a spatially constant input of contaminants. The simple system was

considered most appropriate for evaluating the monitoring configurations and interpretation of the results.
The travel time distribution and solute breakthrough was simulated for a case of cylindrical flow towards a pumping well in an aquifer of 40 m thickness. The simulations assumed steady-state flow in the saturated zone, no dispersion, and linear sorption and first-order degradation. The modelled situation is typical for Dutch conditions with shallow groundwater tables, unconsolidated deposits and a flat landscape.

Simulation of flow to a partially penetrating well
Groundwater flow towards a partially penetrating well was simulated using MODFLOW96 (McDonald & Harbaugh 1986). The aquifer modelled has dimensions 10 km x 10 km x 40 m, using 64 columns and rows in x and y directions and 20 layers in z direction. The grid was refined in the x and y directions, using a minimum cell size of 20 x 20 m in the vicinity of the well and a maximum cell size of 100 x 100 m at the boundaries of the model. The layers have constant thickness of 2 m. The model was homogeneous and slightly anisotropic with hydraulic conductivity $k_v = k_h = 30$ m day^{-1} and $k_z = 10$ m day^{-1}. Porosity was 0.35 and the groundwater recharge rate equaled 300 mm yr^{-1}. The outside borders were modelled using constant-head boundary conditions with $h = 0$.

The well discharge flux of 8 Mm^3 year^{-1} was distributed evenly over 4 partially penetrating wells screened between 30 and 40 m depth in the middle four model cells. Two flow situations were modelled: 1. flow with constant transmissivity k_hD, and 2. flow with variable saturated aquifer thickness due to drawdown of the water table around the pumping well. The first situation was denoted as 'confined' using the terminology used in MODFLOW, the second situation as 'unconfined'. The confined situation corresponds largely to the analytical solution of equations (6.1) to (6.6), except that the well is partially instead of fully penetrating. Groundwater recharge is applied to the uppermost active cell in the unconfined situation. The model set-up results in radial groundwater flow towards the pumping well. The maximum radius of the contributing area, r_e, is 2915 m, which agrees with the analytical solution of equation (6.1).

Streamlines and isochrones were simulated using MODPATH (Pollock 1994). Streamline and transport calculations were investigated using only one quarter of the model, since the model is symmetric relative to the x=0 and y=0 axes. A fine vertical grid with 2 m thick model

Table 6.1 - Monitoring configurations used for the evaluation

Configuration	Well code	Relative distance to centre of well field (r/r_e)	Depth centre of screen (m)	Screen length (m)
Extracted groundwater from Pumping well	P	0	35	10
Shallowest groundwater inside protection zone	S	< 0.44 (a)	0.5 (b)	1
Shallowest groundwater in Control area (outside protection zone)	C	> 0.44 (a)	0.5 (b)	1
Vertical monitoring within the protection zone	V	0.30	2.5, 5.0, 7.5, 10 etc.	1
Vertical monitoring on the 10 year Isochrone	I	0.44	2.5, 5.0, 10, 15, 20	2
Long-screen monitoring on the 10 year isochrone	L	0.44	10, 20, 30	5
Monitoring along Flow path	F	0.3, 0.44, 0.6	2.5, 5.0, 10, 15, 20	0.5-2

(a) r/r_e = 0.44 is boundary protection zone
(b) meters below water table

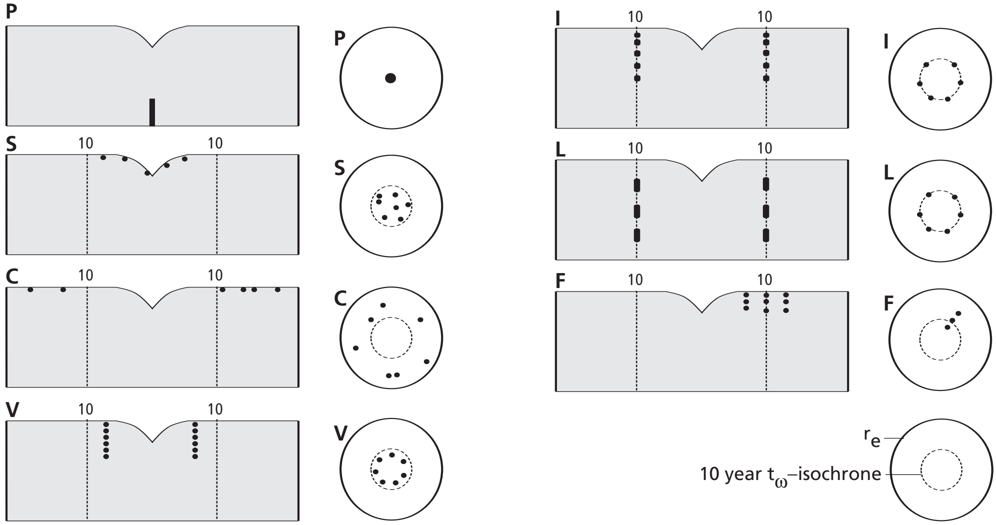


Figure 6.3 - Monitoring lay-outs of the seven monitoring configurations in cross-sections and plane view.

layers was chosen to ensure accurate interpolation of the vertical velocity field for the calculation of the streamlines. The error criterion for groundwater heads was chosen to be 1e-8 m for accurate vertical flow calculations. The particle starting locations for forward tracking were distributed proportionally to the amount of groundwater recharge, according to Barlow (1994). Thus, streamlines represent an equal amount of groundwater flux. Streamlines were calculated from the highest active cell to the well using forward tracking. Residence time t_α-isochrones in the cross-sections were drawn using the forward calculated residence times. Backward tracking from the well screen to the recharge location was used to locate t_ω-isochrones. The position of advective pollution fronts in the cross-sections was determined directly from the t_α isochrone positions.

Configurations of monitoring networks

Seven design configurations, which are relevant for Netherlands conditions and monitoring conventions, were evaluated using the numerical model. Each configuration includes a specific type of observation wells with specific screen depths and lengths and well locations. The configurations and well types are summarized in Table 6.1 and shown in Figure 6.3. The configurations P, S, C, I and L are based on current monitoring strategies used in the Netherlands overall strategy as given by Baggelaar (1992, 1996).

Configuration P reflects the monitoring of the extracted water at the pumping well, which is done to fulfil the legal monitoring necessity and to assure customers (section 6.1). Sampling of the shallowest groundwater in the protection zone (S) and a control area (C) was recommended to evaluate protection measures (Figure 6.3, Baggelaar 1996). Typically, the sampling locations of configurations S and C were found using a stratified random sampling scheme and the sampling is often done using temporary open boreholes (Fraters et al. 1998). Configurations L and I represent the monitoring of groundwater in the pumping aquifer at 10 year residual transit times t_ω to the pumping well (Figure 6.3). These configurations show circular designs with wells distributed over the 10 year t_ω-isochrone to account for differentiated regional inputs of contaminants. The well types I and L are normally sampled

Table 6.2 - Pollution scenarios used for the evaluation of the monitoring strategies

Scenario	Time of solute input (yr)	Half life time τ (yr)	Retardation factor R	Protection area
1a Advective block front	0-10	∞	none	no
1b Advective block front	0-10	∞	none	yes
2a First-order degradation	0-10	5	none	no
2b First-order degradation	0-10	5	none	yes
3a Linear sorption	0-10	∞	3	no
3b Linear sorption	0-10	∞	3	yes

from permanent screens installed in wells. Screen lengths of well type I were 2 metres, which is typical for most Dutch observation wells (for example Chapters 2 to 5).

Monitoring configuration F is occasionally used for identifying chemical processes in aquifers (for example Van Beek et al. 1989, Broers & Buijs 1997) but is not regularly applied in the Dutch monitoring strategy (Baggelaar 1996). This configuration was ordered in a transect along the main flow direction to the well. Wells of configuration V were situated on a circle inside the t_ω= 10 year isochrone to evaluate monitoring inside the 10 year protection zone.

The effectiveness of each of the seven configurations was evaluated for the monitoring objectives *early warning*, *prediction* and *protection* that were introduced in section 6.1.

Pollution scenarios

Five pollution scenarios were defined with relevance for the time scales of diffuse contaminant transport in groundwater. The scenarios are listed in Table 6.2. The pollution front was introduced over the whole contributing area of the well in the scenarios 1a, 2a and 3a. The pollution front in the scenarios 1b, 2b and 3b was only introduced outside the protection zone. The 10 year protection zone was based on the t_ω = 10 year isochrone. The last mentioned scenarios were referred to as *optimal protection*, to stress that no contamination was introduced in the protection zone.

Solute breakthrough in pumping and observations wells

Solute breakthrough was calculated using numerical methods, using the simulated transit times t_ω and residence times t_α.

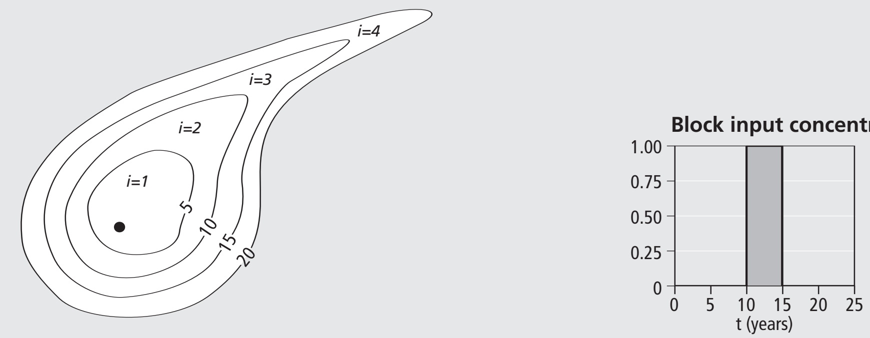
Solute breakthrough in the pumping well was calculated using the cumulative transit time distribution, which was derived from the forward flow path analysis. The resulting well concentration was calculated numerically by (van Brussel, 1990):

$$C_{well,\,t} = \sum_{i=1}^{n} f_i \,.\, C_{recharge,\; t-(i-1/2)\Delta t} \tag{6.12}$$

for n fractions f_i of the contributing area with fixed time step Δt between the t_ω-isochrones (Figure 6.4). The fraction was defined as:

$$f_i = \frac{A_i\, N_i}{Q} \tag{6.13}$$

where A_i is the surface area between two t_ω-isochrones with time step Δt, N_i is the groundwater recharge in the area A_i and Q = the total well discharge. Equations (6.12) and (6.13) enable the calculation of the concentration response for contributing areas with irregular shapes, varying recharge amounts and time-variant recharge input concentrations (Van Brussel



$$C_{well,t} = f_1\, C_{recharge,\ t-1/2\,\Delta t} + f_2\, C_{recharge,\ t-1\,1/2\,\Delta t} + f_3\, C_{recharge,\ t-2\,1/2\,\Delta t} + f_4\, C_{recharge,\ t-3\,1/2\,\Delta t}$$

for example

$$C_{well,t=20} = f_1\, C_{block,\ t=17.5} + f_2\, C_{block,\ t=12.5} + f_3\, C_{block,\ t=7.5} + f_4\, C_{block,\ t=2.5} = f_1 .0 + f_2 .1 + f_3 .0 + f_4 .0 = f_2$$

Figure 6.4 - Numerical calculation of solute breakthrough in a pumping well with an irregularly shaped contributing area and 4 fractions f_i with equal Δt of 5 years for a block input in recharging groundwater between t = 10 and t = 15 years.

1990, TCB 1991, Beugelink & Muhlschlegel 1989). Figure 6.4 gives an example of the numerical calculation for a block input.

For radial flow towards the pumping well, the fractions f_i are concentric rings with equal time step Δt, which were obtained from equation (6.8):

$$f_i = e^{(i-1)\Delta t/T} - e^{i\Delta t/T} \tag{6.14}$$

Equation (6.12) was extended to include first-order degradation and linear sorption

$$C_{well,\ t} = \sum_{i=1}^{n} f_{i,R} . C_{recharge,\ t-(i-1/2)\Delta t} . e^{-\mu(i-1/2)\Delta t} \tag{6.15}$$

where R = retardation factor and μ = first-order degradation constant (day^{-1}). The retardation factor is defined as (Bear 1979)

$$R = 1 + \frac{\rho\Delta S}{\varepsilon\Delta C} \tag{6.16}$$

where $\rho\Delta S$ = fraction sorbed and $\varepsilon\Delta C$ = fraction in solution and $\Delta S/\Delta C$ is constant. First-order degradation is:

$$\frac{dC}{dt} = -\mu C \tag{6.17}$$

with μ = first-order degradation constant. The half-life time τ of the solute is defined as:

$$\tau = \frac{\ln 2}{\mu} \tag{6.18}$$

The factor $f_{i,R}$ corrected the fractions f_i of equation (6.14) for retardation, according to:

$$f_{i,R} = e^{(i-1)\Delta t/RT} - e^{i\Delta t/RT} \tag{6.19}$$

Time steps Δt of 1 year were used for the calculations of concentrations in the pumping well and the observation wells.

Solute breakthrough in the *observation wells* was calculated using the residence times t_α to the upper and lower part of the well screen. The mixing of uncontaminated and contaminated water over the screen was calculated accordingly, assuming horizontal inflow during sampling (Figure 6.5). Concentrations in the observation wells were calculated by integrating the concentrations of the block front over the screen length, including the vertical concentration gradient in the case of degradation.

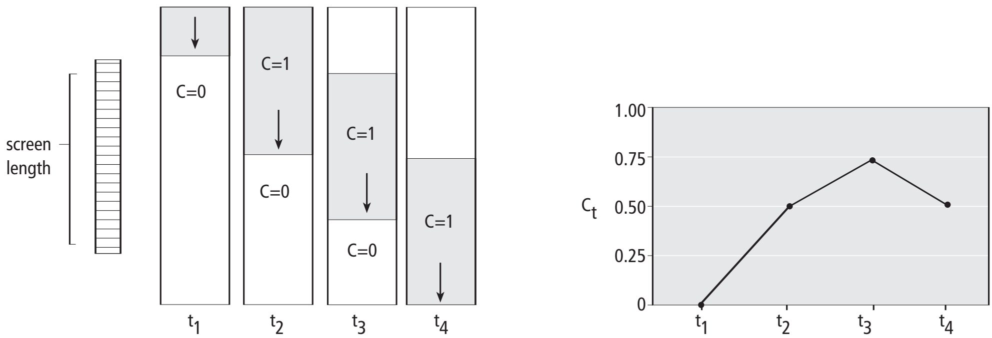


Figure 6.5 - Numerical calculation of the concentration response of an observation well by integrating concentrations over the vertical length of the well screen.

6.4 Results

Position of t_α and t_ω-isochrones

Figure 6.6 shows the streamlines and isochrones in cross-sectional view for partially penetrating wells in the confined (a) and unconfined situations (b and c). The t_α-isochrones of the confined simulation remained horizontal up to a short distance from the partially penetrating well (Figure 6.6a). The flow pattern was similar to the situation with a fully penetrating well, with the exception of the innermost 100 m, where the vertical flow component has increased. Consequently, equation (6.7) is valid for 98% of the model domain.

The flow pattern and t_α-isochrones in the unconfined situation differed significantly from the confined situation. In the area where drawdown of the water table was observed, the vertical flow component was enhanced due to the reduction of the thickness of the saturated zone (Figure 6.6b). The t_α-isochrones in the unconfined simulations showed an increased downward bending up to about 1 km distance from the well.

The cumulative transit time distribution of the unconfined situation differed insignificantly from the cumulative transit time distribution of a fully penetrating well in a confined situation, which is described by the analytical solution of equation (6.8) (Figure 6.2). The observed small differences are mainly due to discretisation errors. Using the same particle starting locations and comparing a fully penetrating well in a confined situation with a partially penetrating well in the unconfined situation, maximum differences of 0.25 years

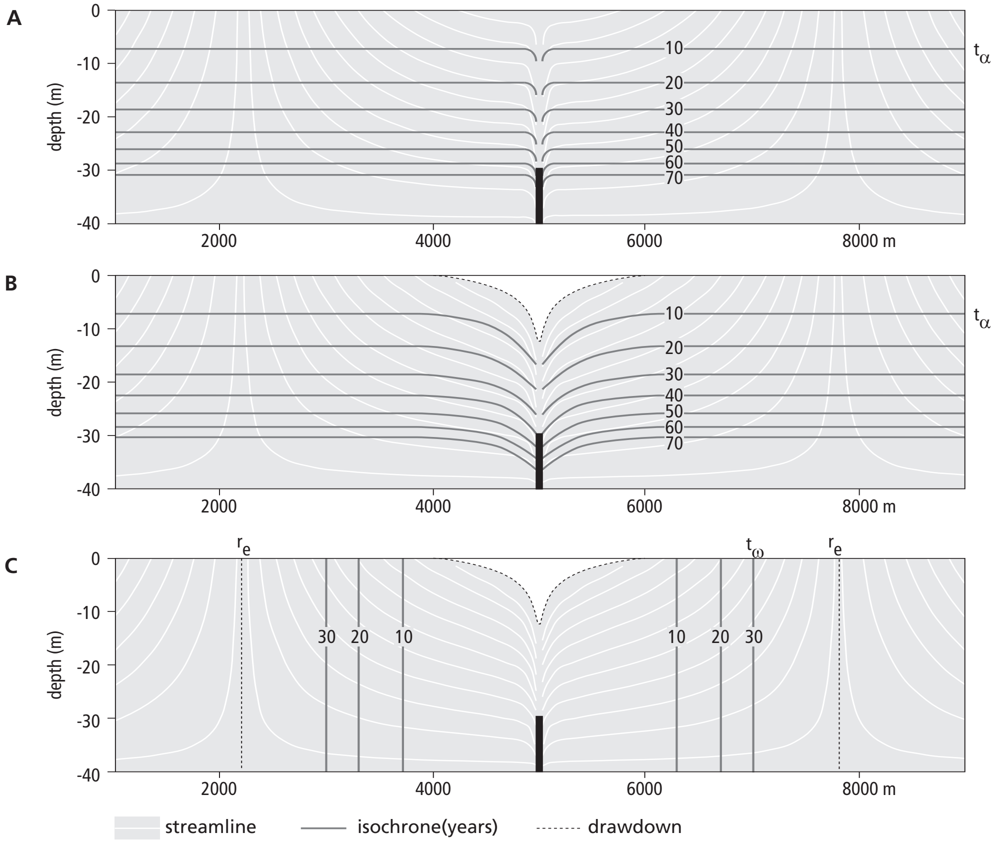


Figure 6.6 - Simulated streamlines and t_α and t_ω-isochrones: (A) confined flow with constant transmissivity and (B and C) unconfined flow with variable saturated aquifer thickness caused by drawdown of the water table. Residence times isochrones are horizontal (B) and residual transit time isochrones are vertical (C).

transit time were observed. These small differences are negligible when evaluating contaminant breakthrough at the pumping well. Correspondingly, the position and shape of the t_ω residual transit time isochrones differed insignificantly from the confined situation (Figure 6.6c). For example, the 10 year t_ω-isochrone, which was used as the boundary of the protection zone, was located at position $r/r_e = 0.44$, which corresponds to equation (6.3).

However, the effects of the downward bending of the t_α-isochrones is worth considering when designing a groundwater quality observation network in the vicinity of a well field. The flow field and isochrone patterns of Figures 6.6b and 6.6c were used to evaluate the monitoring configurations.

Contributing areas and travel times of the monitoring configurations

Figure 6.7 shows the contributing areas for the seven monitoring configurations for the unconfined groundwater flow field. Configurations S and C monitored the most recently recharged groundwater inside and outside the protection zone. Configuration V monitored the groundwater inside the protection zone and groundwater that horizontally entered the protection zone at a relatively shallow level. Configuration I monitored the groundwater from

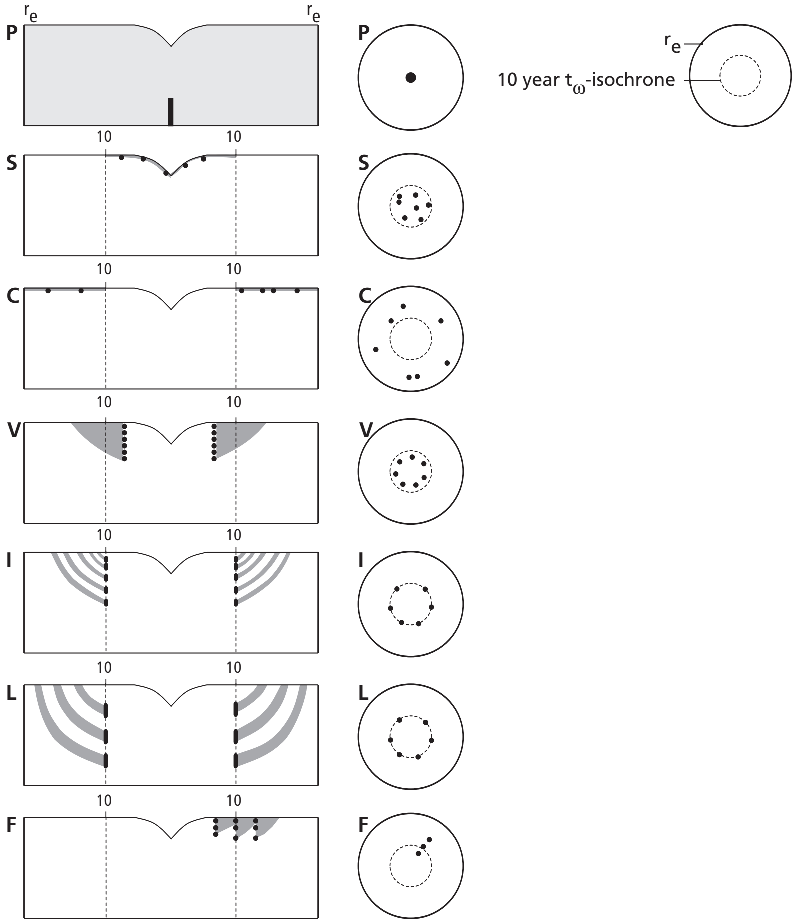


Figure 6.7 - Contributing areas for the seven monitoring configurations (cross-sections + plane views)

outside the protection zone at a shallow level (0-20 m), whereas the long screens of configuration L monitored deeper groundwater originating outside the protection zone. Configurations I, L and V monitored water from all directions in the catchment area. Configuration F monitored the groundwater quality along a transect and does not account for regional differences in the upstream area.

Table 6.3 lists the residence times and residual transit times for the individual screens of the seven monitoring configurations. The listed values are averaged over the screen length. The characteristic transit time T of the system equals 46.7 yr, which corresponds to the average transit time from the surface to the pumping well (configuration P). Residence times in the saturated zone were less than one year for the monitoring wells of configurations S and C. The residence times in the unsaturated zone of configuration S will be larger than in configuration C due to the lowering of the groundwater table inside the protection zone.

Residence times for the vertical wells in the protection zone (configuration V) varied between 2 and 22 years, depending on the depth of the well screens (Table 6.3). Significant differences in the residence times were found between the confined and the unconfined

Table 6.3 - Residual transit times and residence times for the seven monitoring configurations

Configuration	r/r_e	Depth	Screen length	Residual transit time t_ω	Average residence time t_α	
					Partially penetrating well	Fully penetrating well
		(m)	(m)	(yr)	(yr)	
P	0-1	30-40	10	0	46.7(a)	46.7(a)
S	<0.44	0.5	1	<10	0.5	0.6
C	0.44-1.0	0.5	1	>10	0.5	0.6
V	0.3	2.5	1	4.4	2.5	3.0
	0.3	5	1	4.4	4.9	6.2
	0.3	7.5	1	4.4	8.3	9.7
	0.3	10	1	4.4	12.0	13.4
	0.3	12.5	1	4.4	15.8	17.5
	0.3	15	1	4.4	20.9	21.9
I	0.44	2.5	2	10	2.8	3.0
	0.44	5	2	10	5.9	6.2
	0.44	10	2	10	13.0	13.4
	0.44	15	2	10	21.5	21.9
	0.44	20	2	10	31.7	32.3
L	0.44	10	5	10	13.0	13.4
	0.44	20	5	10	32.0	32.3
	0.44	30	5	10	64.0	64.7
F	0.3	2.5	1	4.4	2.5	3.0
	0.3	5	1	4.4	4.9	6.2
	0.3	7.5	1	4.4	8.3	9.7
	0.44	2.5	1	10	2.8	3.0
	0.44	5	1	10	5.9	6.2
	0.44	10	1	10	13.0	13.4
	0.6	2.5	1	21	3.0	3.0
	0.6	5	1	21	6.2	6.2
	0.6	10	1	21	13.4	13.4

(a) average transit time of the system (equals characteristic time T)

simulation for well screens of configuration V at the circle $r/r_e = 0.3$. The residence times for the unconfined simulation were about 10-20 percent less than the residence times in the confined situation. Thus, for the unconfined situation one has to account for shorter residence times and higher vertical groundwater velocities in the monitoring configuration V, due to the reduction of the saturated thickness of the aquifer. However, the drawdown of the phreatic water table will increase the residence times in the unsaturated zone to some extent.

Residence times t_α for the monitoring wells of configuration I on the 10-years t_ω isochrone varied between 3 and 33 years for depths between 2.5 and 20 m (Table 6.3). At this distance $r/r_e = 0.44$, the differences between the confined and the unconfined situation were small (< 10% and < 0.5 yr.). Average residence times to the long screens of configuration L varied between 13 and 64 years for screens between 10 and 30 m depth.

Residence times for configuration F were similar to screens at equal depth in the

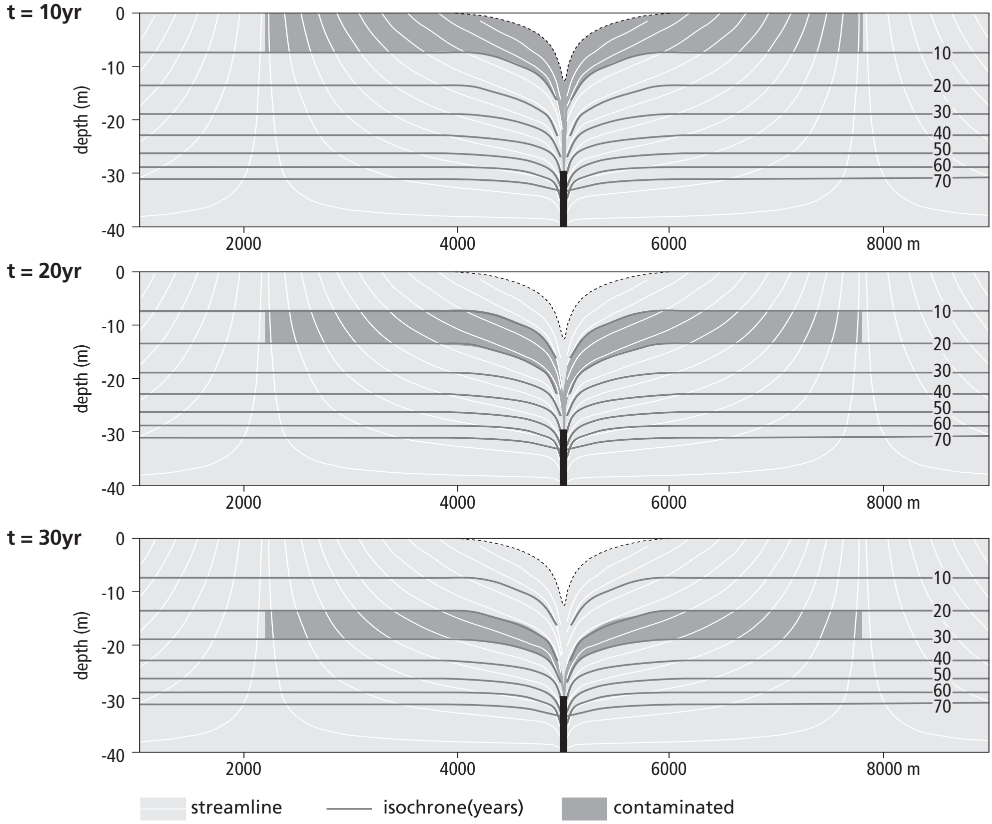


Figure 6.8 - Position of the contaminated groundwater 10, 20 and 30 years after the start of the ten-year block front for scenario 1a without a protection zone

configurations V and I. Residence times of the three screens at r/r_e = 0.3 were equal to the three shallowest screens of configuration V and residence times at r/r_e = 0.44 were equal to the three shallowest screens of configuration I (Table 6.3). At r/r_e = 0.6 the effects of drawdown on the residence times had become negligible and results were similar for the confined and unconfined situations.

The evaluation of the monitoring configurations for advective and reactive scenarios in the following sections is based on the unconfined flow field with water table drawdown due to pumping.

Advective transport scenarios

Front position and breakthrough in pumping wells

Two advective transport scenarios were evaluated, introducing a 10-year block input of contaminants into the system (Table 6.1). Figures 6.8 and 6.9 show the position of the contaminants in the system after 10, 20 and 30 years for the scenario without and with a protection zone. The position of the block front was based on the position of the t_α-isochrones. The figures show that the contamination front moved vertically through the aquifer and was accelerated in the drawdown area.

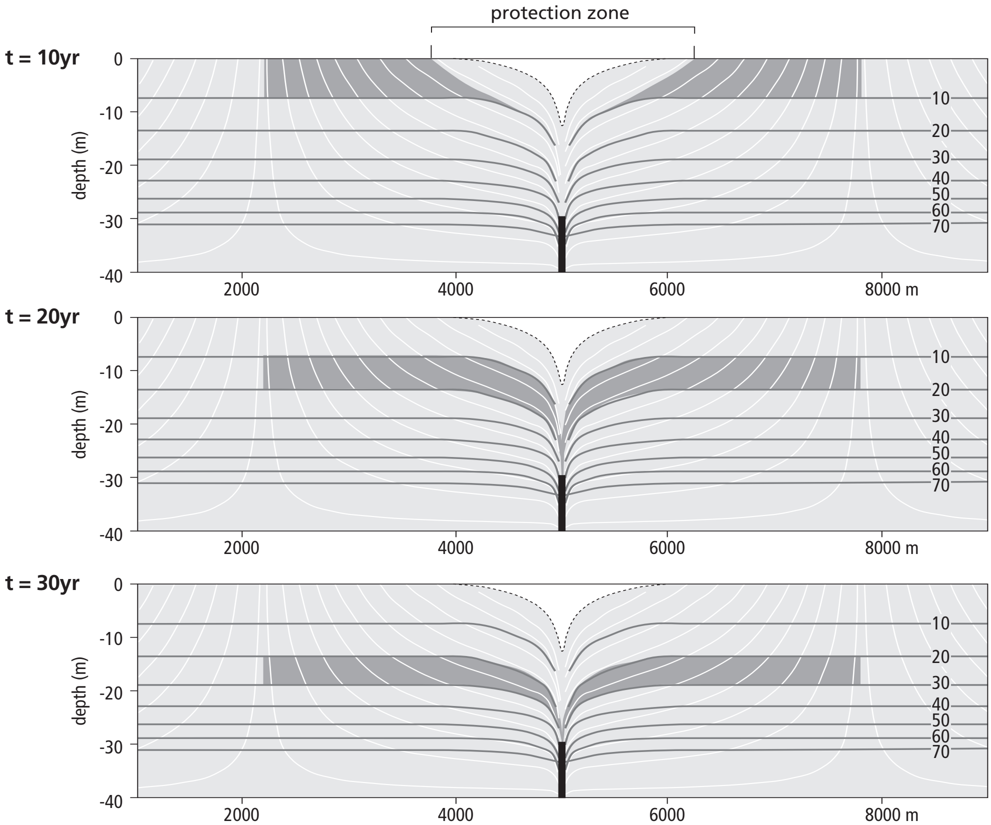


Figure 6.9 - Position of the contaminated groundwater 10, 20 and 30 years after the start of the ten-year block front for scenario 1b with a protection zone

The total contaminant load in the system increased during the first 10 years and decreased subsequently because it was removed gradually by the pumping well (Figure 6.10a). Without a ten-year protection zone the contaminants immediately started to enter the pumping well (see also Beltman 1996). The maximum concentration of 19.3% of the input concentration was reached after 10 years and a gradual decrease follows afterwards (figure 6.10a).

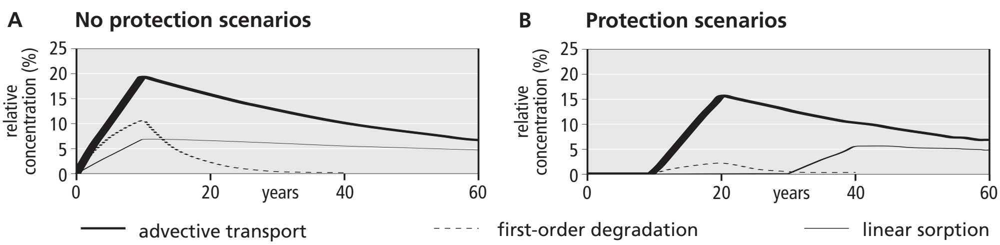


Figure 6.10 - Breakthrough of the contaminants in the pumping well for scenarios without a protection zone (A) and with a 10-year protection zone (B). Scenarios: advective block front, first order degradation (τ = 5 yrs), retardation (R=3).

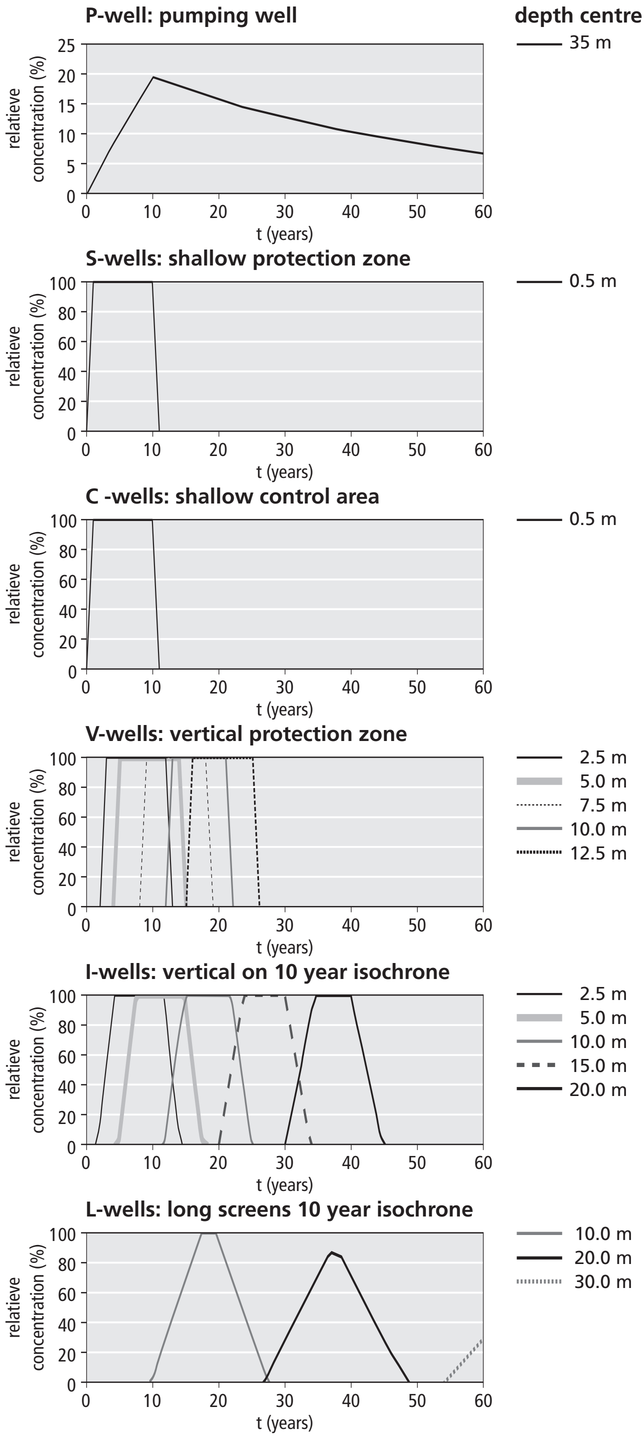


Figure 6.11 - Breakthrough of the contaminants in the observation wells for six monitoring configurations. Advective transport, no protection inside protection zone

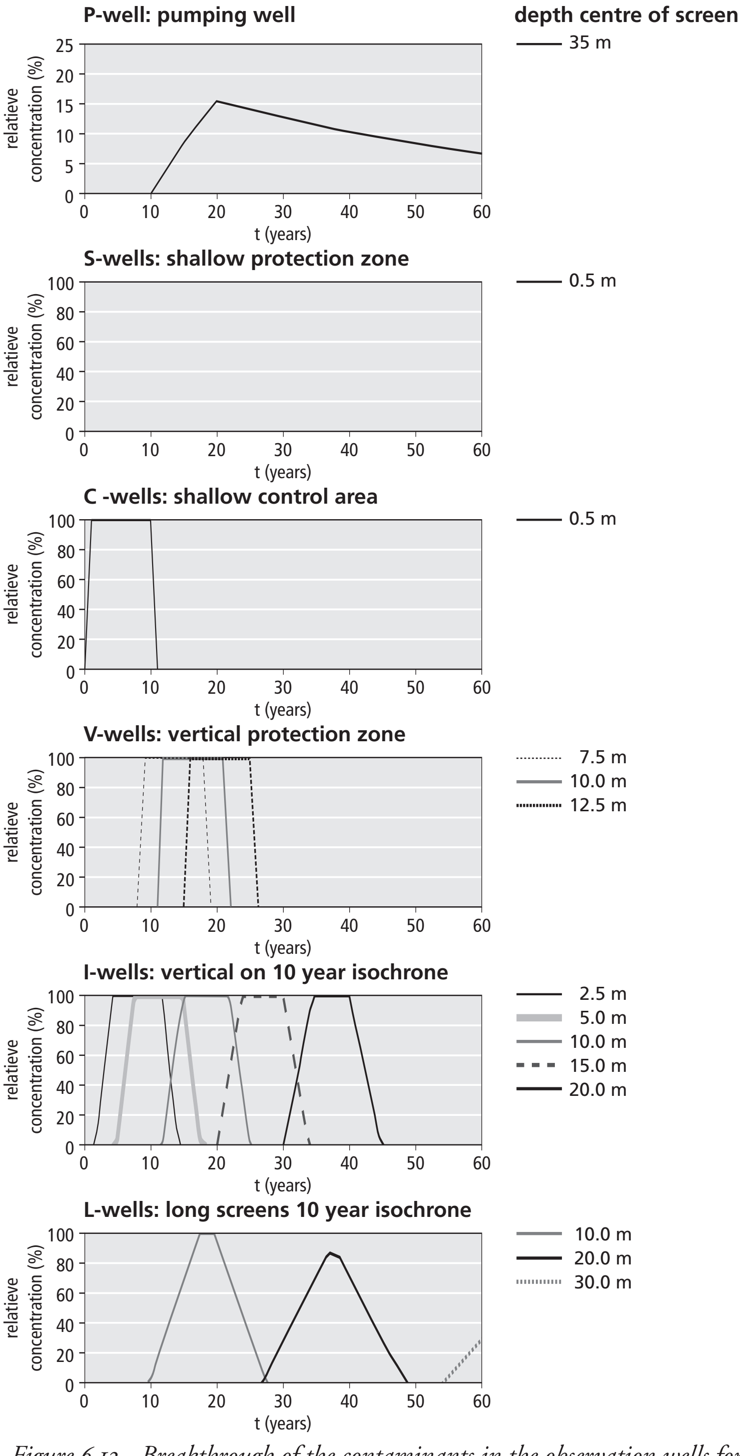


Figure 6.12 - Breakthrough of the contaminants in the observation wells for six monitoring configurations. Advective transport, optimal protection inside protection zone

No contaminants arrived at the pumping well in the first 10 years when a protection zone was applied (scenario 1b, Figure 6.10b). However, the contaminants did start to enter the protection zone laterally at a shallow level (Figure 6.9). After 10 years, the first contaminants started to enter the pumping well. Ten years later the maximum concentration of 15.6% of the input concentration was reached in the pumping well (Figure 6.10b). The contaminant load in the system and pumping well concentrations were equal for the protection and no protection scenario's after 20 years (Figure 6.10, a and b).

Breakthrough in the observation wells

Figures 6.11 and 6.12 present the breakthrough in the observation wells of six configurations for the advective transport scenarios with and without a protection zone (scenarios 1a and 1b). Breakthrough for configuration F is not shown because the residence times were equal to those of configurations V and I, and results are similar (Table 6.3).

The breakthrough in the observation wells is determined by the residence time t_α to the screen and the mixing of water over the length of the well screen. The smallest mixing took place in the short well screens of configurations S, C and V, yielding only slight deformation of the shape of the block front. The longer screens of the configurations I and L (respectively 2 and 5 m) resulted in a stronger deformation of the shape of the block front. The slope of the concentration increase at an observation well depends on the vertical flow velocity of the contaminant front. The larger the screen length, the larger the time the front needs to move from the top of the screen to the bottom, resulting in a slower increase of the concentration (Figure 6.5).

If no protection is applied inside the protection zone, an immediate breakthrough is observed in configurations S and C and within 2.5 year in the shallow screens of configurations V and I (Figure 6.11). Breakthrough in the deeper screens of the configurations V, I and L happened simultaneously or later than the time at which the concentration maximum in the pumping wells is reached. This was especially true for well screens deeper than 10 m depth.

Figure 6.12 shows the breakthrough for scenario 1b with a protection zone. Logically, the breakthrough in the observation wells outside the protection zone equals that of scenario 1a. No contaminants arrived at the shallow screens of configurations S and V within the protection zone. However, at depths larger than 5 m in configuration V the breakthrough equals the breakthrough in scenario 1a because the contaminants laterally move into the protection zone (see Figure 6.9 at t = 10 yr).

Effectiveness of the monitoring configurations

The study was designed to evaluate the effectiveness of the seven monitoring configurations for early warning, prediction and protection. Table 6.4 summarizes the results, which are discussed below.

The *early warning monitoring* is meant to signal unexpected threats to the quality of the extracted groundwater. The time of breakthrough in an observation well relative to the breakthrough in the pumping well is the most important criterion for the evaluation of this objective. Since concentrations in the pumping well started to increase directly in the 'no protection' scenario, only configurations S and C and the shallow screens of configurations V, I and F are considered useful for early warning purposes (Figure 6.11). The concentration increase in the pumping wells started later when a protection zone was present. In that situation, both the monitoring configuration C and the shallow screens of configuration I provide an early warning signal for the concentration increase in the pumping wells (Figure 6.12 and Table 6.4). The shallow screens of configuration F at r/r_e = 0.44 and r/r_e = 0.6 are also useful for early warning. The shallow screens of configuration V are less effective: the shallowest screens remain uncontaminated because of the protection applied. Consequently,

Table 6.4 - Effectiveness of the seven configurations for early warning, prediction and protection

Objective/Scenario	P	S	C	V	I	L	F
Early warning							
1a Advective, no protection	-	X	X	X^{A}	X^{A}	-	X^{A}
1b Advective, protection zone	-	-	X	-	X^{A}	-	$X^{A,B}$
2a Degradation, no protection		X	X	X^{A}	X^{A}	-	X^{A}
2b Degradation, protection zone	-	-	X	-	X^{A}	-	$X^{A,B}$
3b Retardation, protection zone	-	-	X	-	X^{C}	-	X^{C}
Prediction							
1a Advective, no protection	C	I	I	C	C	-	C
1b Advective, protection zone	C	-	I	-	C	-	C
2a Degradation, no protection	C	I	I	C,P	C,P	-	C,P
2b Degradation, protection zone	C	-	I	-	C,P	-	C,P^{B}
3b Retardation, protection zone	-	-	I	-	C,P	-	C,P^{B}
Protection							
1a Advective, no protection	-	X	X	-	-	-	-
1b Advective, protection zone	-	X	X	-	-	-	-
2a Degradation, no protection	-	X	X	-	-	-	-
2b Degradation, protection zone	-	X	X	-	-	-	-
3b Retardation, protection zone	-	X	X	-	-	-	-

X = suitable for objective
^{A} = only shallow screens (< 10 m)
^{B} = only screens outside protection zone
^{C} = only shallowest screens (< 5m)
C = suitable for calibration purposes
I = suitable to obtain information on input
P= suitable to obtain information on processes

the early warning monitoring is preferably done at shallow depth outside the protected area.

Prediction monitoring aims to support operational decisions on future groundwater quality changes. In contrast to early warning monitoring, the focus is on the prediction of contaminants that have been earlier identified at the well field. Three types of information are considered relevant for this information goal:

1. information on the input of solutes in the recharging shallow groundwater (Pd-I; input)
2. information on the type and rate of chemical processes (Pd-P; processes)
3. information on the concentrations in the aquifer and in the extracted groundwater to calibrate prediction models (Pd-C; calibration).

The criterion for evaluating information goal Pd-I is that the monitoring data should yield a time series of concentrations in the recharging groundwater under relevant land use types in the contributing area of the well. The time series should not be influenced by transformation or retardation processes in the saturated zone, but attenuating processes in the unsaturated zone are preferably already incorporated. Configurations C and S best suit the criterion for information goal Pd-I. When a protection zone is applied, configuration C is most appropriate (Table 6.4).

Two types of information are relevant for the calibration of prediction models (information goal Pd-C):

1. breakthrough curves of pumping wells and observation wells. For example, calibration can be established by comparing observed pumping well concentrations with model results such as depicted in Figures 6.10a and b.
2. concentration gradients in the aquifer. This type of information can be obtained from concentration-depth gradients in observation wells.

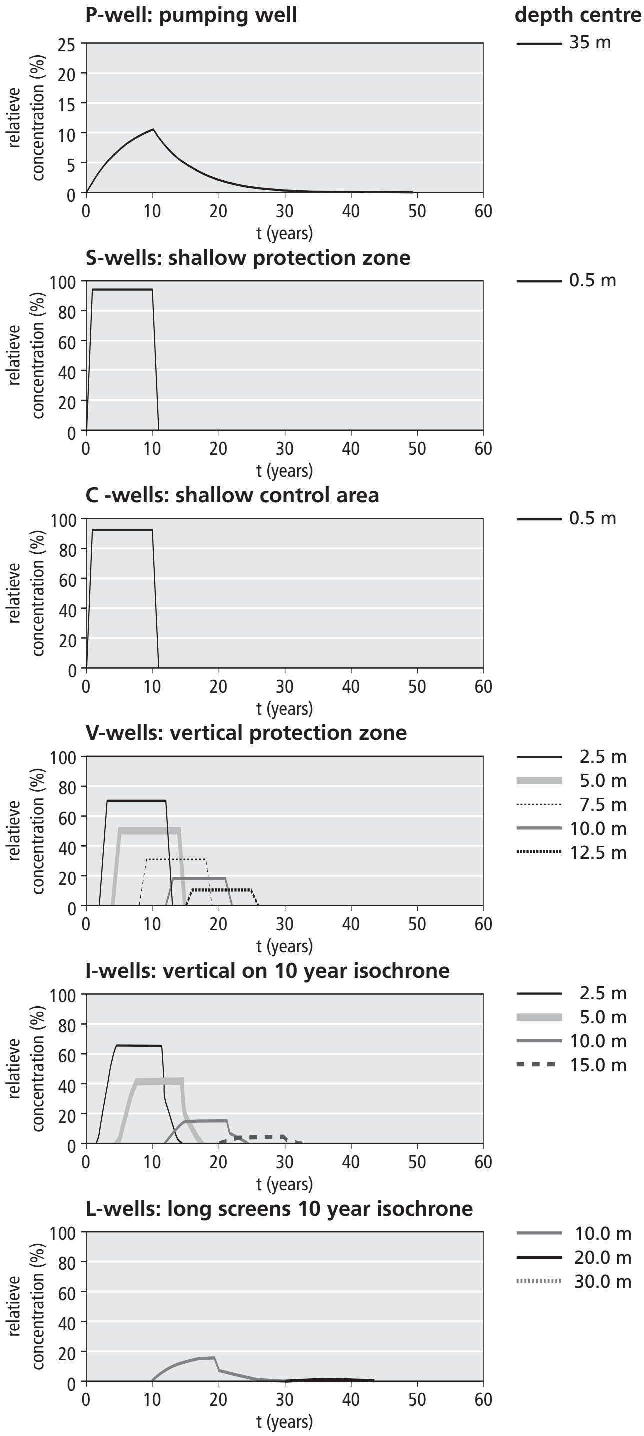


Figure 6.13 - Breakthrough of the contaminants in the observation wells for six monitoring configurations. First-order degradation (τ = 5 yr), no protection inside protection zone

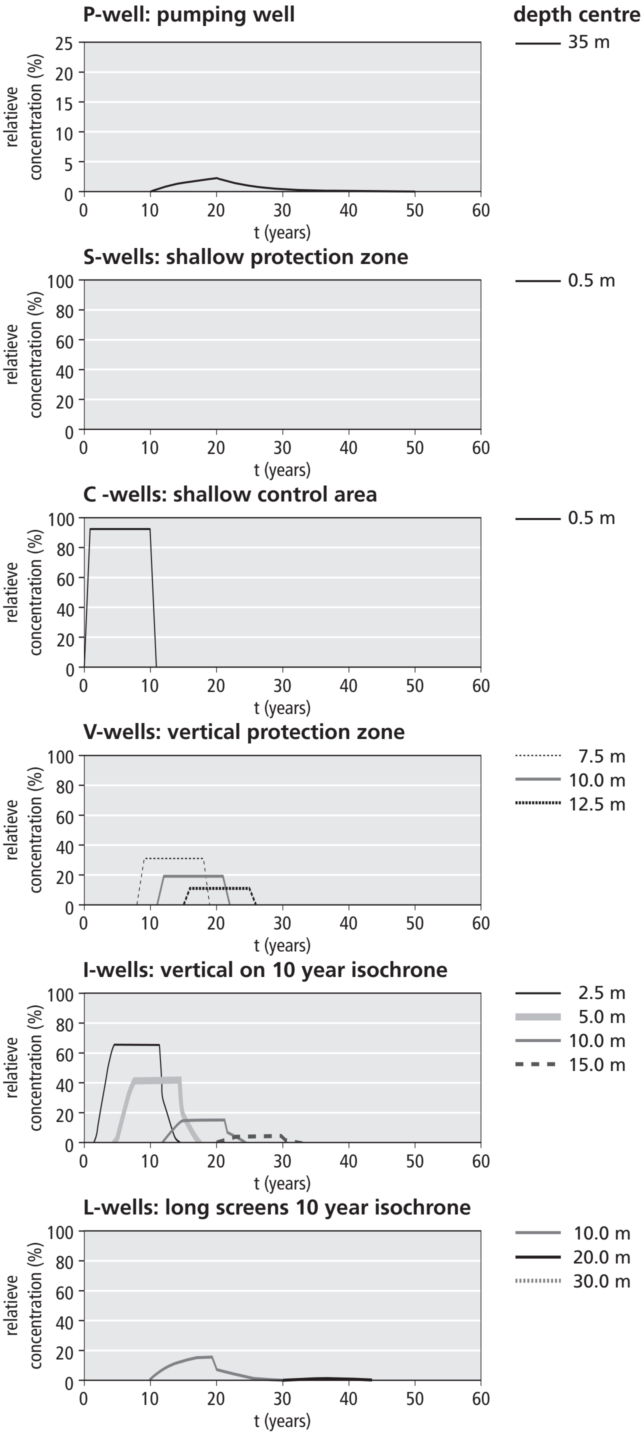


Figure 6.14 - Breakthrough of the contaminants in the observation wells for six monitoring configurations. First-order degradation (τ = 5 yr), optimal protection inside protection zone

Breakthrough in the pumping well (configuration P) and the observation wells of the configurations V, I and F yield the required information for information goal Pd-C for advective transport without a protection zone (Table 6.4). When a protection zone was applied only the configurations I and F are suitable for the calibration task. Configuration F has little extra value compared with configuration I since breakthrough at $r/r_e = 0.6$ is almost simultaneous with breakthrough at $r/r_e = 0.44$. Because of the long residence times to the deeper screens of configuration L, these screens are inadequate for calibration purposes if the modelling goal is to predict future changes. Other configurations, including measuring at the pumping well, provide the required information in a much earlier stage of monitoring (Figure 6.12).

For advective transport, no information on attenuating processes is required, which leaves objective Pd-P irrelevant.

The aim of the *protection* objective is to evaluate protection measures that are applied within the protection zone. This requires information on concentrations of recharging groundwater inside and outside the protection zone, in order to identify differences between those areas. This information goal is best achieved by the comparison of the monitoring results of the configurations S and C, which are based on random sampling.

In summary, only the configurations S and C and the shallow screens of V and I were suitable for early warning monitoring of advectively transported solutes. Shallow monitoring just outside the protection zone was most useful for both the early warning and prediction objectives. Combinations of the configurations S and C performed best for evaluating the effect of protection measures.

First-order degradation scenarios

Breakthrough in observation and pumping wells

Two scenarios were designed to investigate the effect of adding first-order degradation, using a half-life time τ of 5 years for the entire aquifer (Table 6.2). Figure 6.10a shows the breakthrough in the pumping well. The first-order degradation leaded to a slower concentration increase, a smaller maximum and a sharper decrease afterwards when no protection zone was applied.

Optimal protection within $r/r_e = 0.44$ also leaded to slower increase and faster decrease of concentrations relative to the advective case (scenario 2b, Figure 6.10b). The concentration reduction was much larger (about seven times) than without protection. This result agrees with findings of Beltman et al. (1996) that protection zones reduce the pesticide transport to wells substantially if the pesticide half-life τ is much smaller than the characteristic transit time of the aquifer T. The ratio τ/T equalled 0.11 in the scenarios used.

Figures 6.13 and 6.14 present the breakthrough in the observation wells for the first-order degradation scenarios. In general, the degradation process does not influence the moment of breakthrough, but only the concentrations. The residence time t_α to the well screen and the half-life time τ of the solute determine the concentration decrease with depth. Because the residence time is a function of depth, a vertical concentration gradient results with lower concentrations in deeper screens.

For the monitoring of the shallowest groundwater (configurations S and C) only a small reduction of concentration occurred relative to the advective scenario. Much larger reduction of concentrations resulted for the deeper well screens of configurations C, I and L. The degradation also affected the shape of the breakthrough curves of the longer well screens of configuration and I and L, yielding a slower increase of concentrations and sharper decrease.

Effectiveness of the monitoring configurations
The effectiveness of the monitoring configurations for the *early warning* objective is judged equal to the advective scenario because the degradation does not influence the time of breakthrough (Table 4.6). However, the deeper screens of configurations I and especially L become even less effective compared with the advective scenarios, because the concentrations in the observation wells decrease with depth (Figure 6.15b). Thus, for early warning the focus must be on the shallow screens. The higher the degradation rates, the shallower the monitoring effort should be for effective early warning. This is especially true for pesticides which are harmful in concentrations that are orders of magnitude lower than the input concentrations (Beltman et al. 1995).

For the reactive scenarios, it is important to obtain information about the degradation processes in the aquifer in order to *predict* the future concentration development in the pumping wells. Two types of information are useful for information goal Pd-P:

1. information on the decrease or increase of concentrations with depth at a fixed moment in time (for example Figure 6.15b), and
2. information on the concentration changes in time at different depths (Figures 6.13 and 6.14).

These types of information are best provided by the well configurations V and I for the first-order degradation scenario without a protection zone (Figure 6.13, Table 6.4). When a protection zone is applied, only configuration I is capable of presenting the requested information, because the shallow screens of configuration V do not render a complete concentration-depth profile (Figure 6.14).

The effectiveness of the configurations for the *prediction* information goals Pd-C (calibration) and Pd-I (input) is judged equal to the corresponding advective scenarios. Accordingly, configurations S and C are found to be most appropriate to achieve the *protection* objective (Table 6.4).

Linear sorption scenarios

Breakthrough in observation and pumping wells
Scenarios 3a and 3b describe the effect of linear sorption, using a retardation factor $R = 3$ for

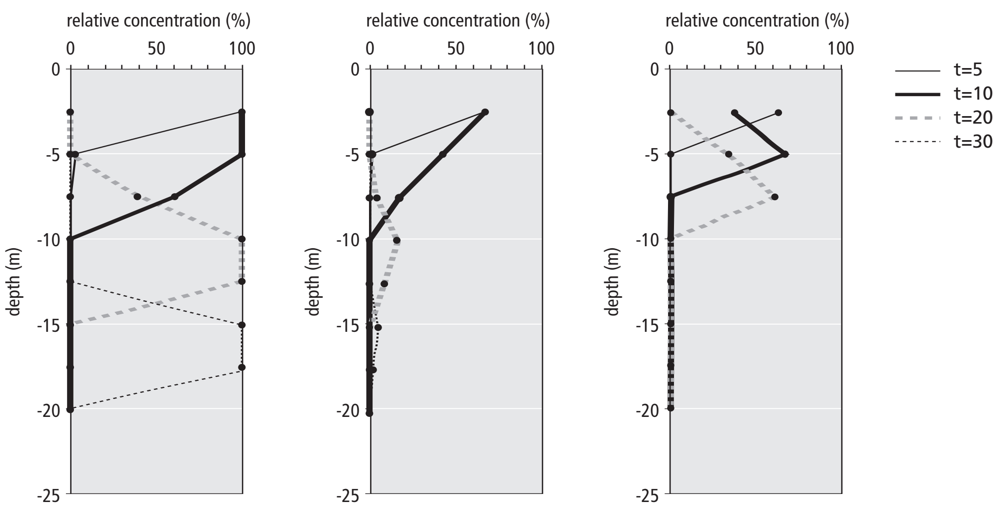


Figure 6.15 - Concentration-depth profiles for configuration I at t= 10, 20 and 30 years after start of the block input: (a) advective transport, (b) first-order degradation (c) linear sorption.

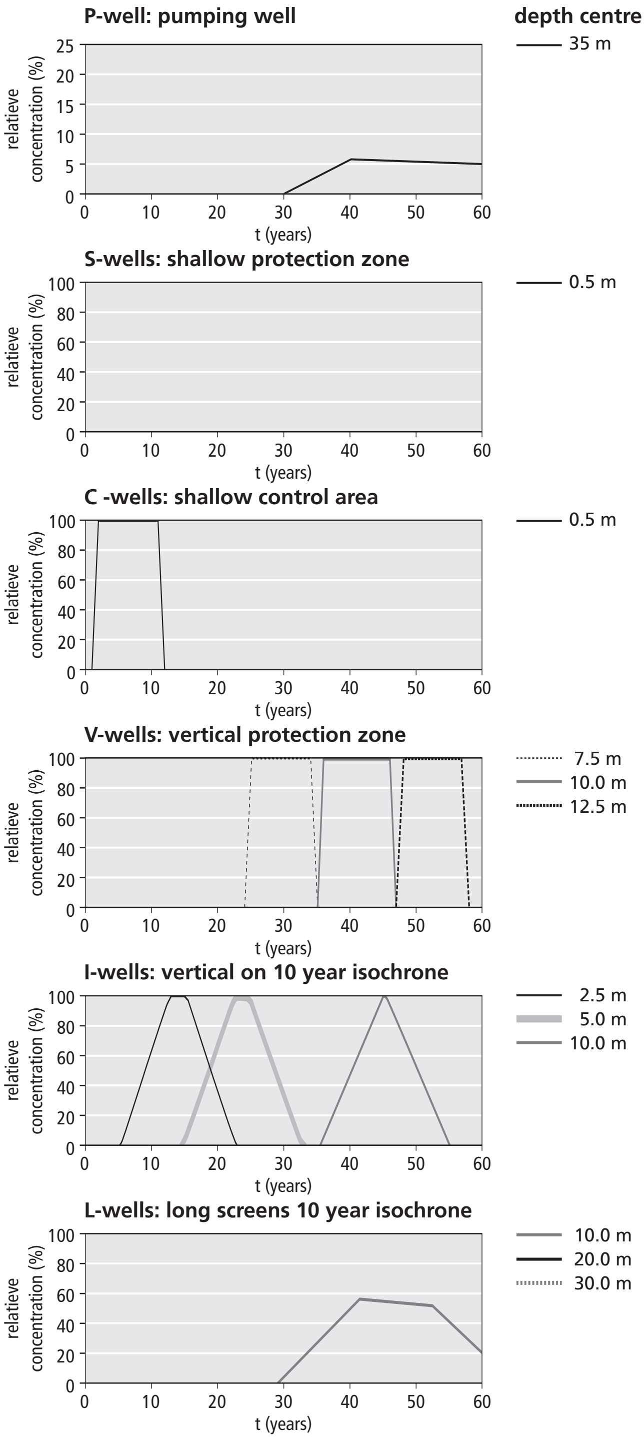


Figure 6.16 - Breakthrough of the contaminants in the observation wells for six monitoring configurations. Linear sorption (R=3), optimal protection inside protection zone

the 10-year block input (Table 6.2). Figures 6.10a and 6.10b show the effect on the breakthrough curves of the pumping well. When no protection zone was applied, the retardation resulted in an approximately threefold decrease of the concentration maximum compared with the advective case (figure 6.10a). Even with retardation, an immediate breakthrough was observed in the pumping well. The smaller vertical velocity of the retarded pollution front causes the lower concentrations in the pumping well because a smaller volume of contaminated water in the aquifer contributes to the outflow in the well at a specific moment in time. For example at t = 20 year, the volume of contaminated water occupied the aquifer volume between the 3.3 and 6.6 year t_α-isochrones, instead of the 10 and 20 t_α-isochrones as depicted in Figure 6.8. The slow concentration decrease after t = 10 years is caused by the three times slower movement of the front through the aquifer.

The concentrations in the protection zone scenario 3b only started to rise thirty years after the start of the contamination. After the maximum at t = 40 yr, a gradual decrease of concentrations was calculated which equals scenario 3a without a protection zone.

Figure 6.16 shows the effects of linear sorption on the concentration breakthrough in the observation wells for scenario 3b with a protection zone. For the shallow screens of configuration C only a small retarding effect was present because of the short residence time t_α to the screen. The effects of retardation are clearer for the observation screens with long residence times t_α. Breakthrough in the short screens of configuration V started only after 25 years at the depth of 7.5 m. In configurations I and L, the linear sorption affected the breakthrough in two ways:

1. breakthrough was retarded compared with the advective case, and
2. stronger deformation of the shape of the block front occurred.

The second effect resulted from the smaller vertical velocity of the block front and the longer time needed to pass the screen vertically. As the total thickness of the block front became less than the screen length, the maximum concentration of 100% was no longer reached (Figure 6.16, configuration L, see also Figure 6.5).

Effectiveness of the monitoring configurations

The results indicate that the use of a protection zone is effective for retarding solutes. In the pumping well, the ten-year block front broke through after thirty years. This leaves time to take countermeasures when an appropriate configuration is chosen for early warning. The configuration C is most effective for *early warning*, because the concentration maximum is reached long before the concentration in the pumping well start to rise (Table 6.4). Monitoring at the ten-year isochrone (configuration I) is an alternative but must be done at a depth smaller than 5 or 10 m to ensure enough time left for countermeasures (Figure 6.16).

The following results were found for the information goals Pd-C, Pd-I and Pd-P. Calibration using the pumping well concentrations is considered inadequate since the water remained uncontaminated for a long time. Therefore, calibration of prediction models is best done using the concentration-depth and concentration-time data obtained from configuration I (Figure 6.15c and Figure 6.16). Especially, the time delay of the breakthrough between two screens (for instance 2.5 and 5 m) is useful to recognize retardation processes (information goal Pd-P). Configuration C is most appropriate to determine the input of solutes. Again, the *protection* objective is best covered using a comparison of data of configurations S and C.

Effective monitoring configurations

The results of the scenarios showed that the monitoring configurations S, C and I are most functional for a variety of objectives and conditions. The random sampling of shallowest groundwater inside and outside the protection zone (configurations S and C) is effective for early warning and for the estimation of the input of contaminants from the unsaturated zone. Additionally, the comparison of sample statistics from configurations S and C enables the evaluation of the effectiveness of protection measures. Shallow monitoring at the boundary of the 10-year protection zone (configuration I) is also useful for the early warning objective, even if no protection zone is applied. The successive vertical screens of configuration I are most appropriate for calibration and prediction purposes. The concentration-depth profiles and the concentration-time series of those wells can then be used to unravel attenuating or retarding processes in the aquifer, to estimate the contaminant loads that are already underway and to calibrate prediction models. The larger the degradation rates or the retardation, the shallower should the monitoring effort be for effective early warning, prediction and calibration.

The configurations V, L and F are of little use for monitoring. The results of the vertically screened wells inside the protection zone (configuration V) are difficult to interpret because water from outside the protection zone is moving in laterally (Figure 6.9). Moreover, residence times t_α to the screens of configuration V are influenced by the increase of the downward vertical velocity due to the reduction of the saturated aquifer thickness, which results from the drawdown of the groundwater table. This complicates the estimation of the age of the sampled groundwater. Monitoring at depths below 20 m (configuration L) is ineffective because of the long forward residence times t_α and the late arrival of contaminants relative to the arrival in the pumping well. Other configurations, including measuring at the pumping well, provide the required information in a much earlier stage of monitoring. Monitoring along a flow path (configuration F) adds little extra value to configuration I, since breakthrough at a certain depth occurs simultaneously in all wells along the transect, except for small differences due to the effect of the reduction of saturated aquifer thickness on the residence times t_α.

6.5 Discussion

The following issues are discussed in this section: (1) implications for the Dutch monitoring practice, (2) effects of spatially heterogeneity, (3) the monitoring of the shallowest groundwater, and (4) the identification of hydrogeochemical processes.

Implications for Dutch monitoring practice

The current monitoring strategy of the Dutch water companies focuses on the pumping wells (comparable to configuration P), the shallowest groundwater (configuration S and C) and the groundwater in the aquifer at 10 to 15 years transit time towards the pumping wells (comparable to configurations I and L).

The underlying monitoring concept of the last mentioned configuration was implicitely based on a 2D concept of groundwater flow, assuming predominant horizontal transport of groundwater and solutes. The resulting monitoring design emphasized the residual transit times t_ω rather than the residence times t_α. Observation wells are located at the 10 or 15 years t_ω isochrone, but little attention has generally been paid to the depths of the observation screens. Assuming predominantly horizontal transport, the depth of the observation wells did not seem relevant, and many observation wells have screens at considerable depth.

The evaluation results indicate that monitoring screens at relatively large depth are ineffective for the objectives of early warning and prediction. Groundwater at greater depth in the aquifer remains uncontaminated for a long time, while solute breakthrough in the pumping well starts immediately after the start of contamination at the surface (Figure 6.17). Hence, upstream monitoring at greater depth in the aquifer (deeper screens of configurations L

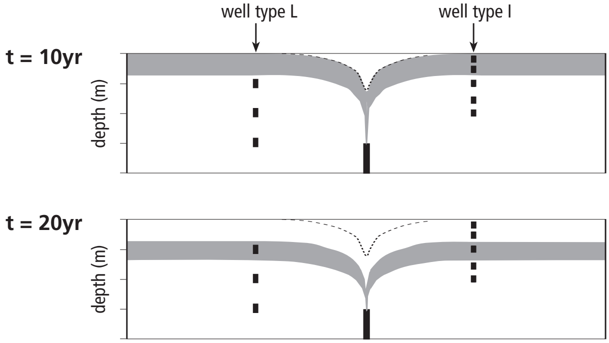


Figure 6.17 - Deep screens of configurations L and I are ineffective for early warning monitoring.

and I) is unable to signal quality changes induced outside the protection zone. This result was also shown for a real-world case by Van Vught & Van der Eijnden (1998) who used a transport model to evaluate monitoring configurations at the Noord-Bargeres well field (Drenthe, the Netherlands). Introducing first-order degradation and retardation resulted in even less effectiveness of the deeper screens in those configurations.

The presented results show that for early warning and prediction, one should monitor shallow screens with limited screen lengths at the 10 years t_ω isochrone. This conclusion was obtained by assuming that the aquifer was uncontaminated before the start of the monitoring. However, diffuse contamination by agricultural sources in the Netherlands increased during the last 40 years and a contamination load is already underway. Results from the regional monitoring networks of Drenthe and Noord-Brabant (Chapters 2 to 5) indicated that most pollution fronts are still limited to the first 25 m of the saturated zone and many contaminants are found at more shallow level because of retardation and transformation processes. Given the current depth of contamination, the use of configuration I is still the best choice, but screens should have depths up to 30 m in order to identify the complete pollution history of the last decades and to predict future changes in extracted groundwater.

Spatially heterogeneous inputs, subsurface heterogeneity and irregular shaped contributing areas

The results of the study are valid for contaminants of diffuse origin that infiltrate with the recharge water. Examples are pesticides and nutrients used on agricultural land. Point sources of contamination (such as NAPLs) and line sources (recharging rivers, rail roads) require other monitoring configurations. A laterally homogeneous input of contaminants was assumed in the contributing area of the pumping wells. In reality, a more heterogeneous input is to be expected, resulting from parcels with different land use. Beltman et al. (1995) argued that the use of spatially averaged input concentrations is justified when calculating the concentration breakthrough in the *pumping well*. They refer to Duffy and Lee (1992) who showed that stationary spatial variations in the solute inputs have negligible effects on solute breakthrough when the correlation scale of the variations is less than 10% of the maximum horizontal travel distance. Thus, pumping well breakthrough is insensitive to spatial variations in recharging groundwater with correlation scales < 0.1 r_e.

However, spatially variable inputs will disturb the vertical sequence of quality patterns that are used to evaluate solutes breakthrough in the *observation wells* (Griffioen 2001). In that case, a number of well locations should be installed in order to assess the spatial variability of concentrations at different depths in the aquifer. Wells on the 10 year t_ω-isochrone should

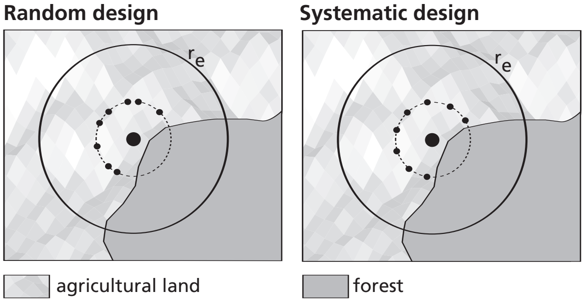


Figure 6.18 - Selection of well locations on the 10 year t_ω isochrone, downstream of agricultural areas with high risks for groundwater contamination using a random (a) or a systematic design (b)

preferably be chosen downstream of areas with high-risks for groundwater contamination, such as agricultural lands (Figure 6.18). Both systematic or random designs on the t_ω-isochrone, using wells with multiple screens at fixed monitoring depths, are appropriate to evaluate the average contaminant concentration and its variance at different depths in the aquifer.

The selected scenarios assume a homogeneous subsoil and no dispersion. Duffy & Lee (1992) demonstrated that small-scale dispersion and aquifer hydraulic heterogeneity have negligible effects on the concentration response of the base flow of the stream for a non-point source of contamination. The breakthrough is insensitive to heterogeneity if the correlation scale of the heterogeneities is smaller than 2.5% of the times the maximum horizontal travel distance, which in our case equals to r_e.

However, the concentration response of the *observation wells* is easily influenced by the hydraulic heterogeneities in their direct surroundings. At the aquifer scale this results in uncertainty of residence times t_α at certain depth. An appropriate number of wells should be chosen to acquire reliable sample statistics of the concentration breakthrough at specific depths in the aquifer. For example, when configuration I is chosen, a number of observation well locations on the 10 years t_ω isochrone could be chosen using random or systematic sampling designs to ensure representative sample statistics (Figure 6.18).

The design and evaluation of monitoring configurations in a real world case should include the calculation of flow paths and t_α and t_ω-isochrones in the heterogeneous subsurface, using relevant boundary conditions for the system studied. Equation (6.15) also applies for irregular shapes of the contributing area and spatially variable inputs for different land use types can be introduced easily (Beugelink & Mühlschlegel 1989, van Brussel 1990, TCB 1989, Laeven 1997). The general principles and evaluation results given in this chapter will still hold for hydraulic heterogeneous subsoils, irregularly shaped contributing areas and spatially variable contaminant inputs when both t_α and t_ω-isochrones are assessed.

Monitoring of the shallowest groundwater

The monitoring of the shallowest groundwater in the monitoring configurations S and C seems promising for early warning and quantification of input of solutes from the unsaturated zone. However, many practical problems have been observed during recent years. Shallow monitoring suffers from large temporal variability. The variability is caused by seasonal and more-yearly fluctuations of precipitation and evaporation, resulting in variable leaching of contaminants during the time frame of several years (Fraters et al. 1998). Since the uppermost groundwater is sampled, the monitoring is also complicated by the changing of the monitoring depth due to the temporal fluctuations of the water table. Moreover, a large spatial variation is observed in the monitoring of shallow groundwater, probably caused by local spatial variations in the contaminant input and soil characteristics. These problems can partly be solved by using large sample sizes and a correction procedure for meteorological fluctuations (Fraters et al. 1998).

Monitoring frequencies for such a monitoring set-up should be carefully tuned to the monitoring objectives. Low-frequency monitoring is suitable when the monitoring objective is to evaluate the effects of *protection* measures, because the main aim is to compare monitoring results from the protection area and the control area at a specific moment in time. However, short-term temporal variations of groundwater quality in the first metres below the water table probably hamper a sound interpretation of the collected data when the information goal is to acquire information on the time-dependent *input* of contaminants (Pd-I). A regular sampling frequency of two or four times a year seems inevitable to obtain reliable estimates of the average contaminant concentrations on the spatial scale of a protection zone or control area.

Identifying chemical processes in the saturated zone

The concentration-depth profiles at a fixed time are useful to unravel the character and rate of degradation and retardation processes when an age-depth transformation is applied (Figure 6.15). The advantage of this approach is that environmental tracers such as tritium, tritium/helium and CFCs can be used for age dating of the shallow groundwater in order to quantify the advective component of solute breakthrough (Solomon et al. 1992, Ekwurzel et al. 1994, Johnston et al. 1998, Zoellmann et al. 2001). Subsequently, retardation and degradation rates can be estimated from the concentration-depth profiles, assuming that the solute input history is known. In Chapter 5 examples were given of the use of concentration-depth profiles for advectively and reactively transported components. For example, the retarded behaviour of potassium was deduced from the monitoring data using age dating and a geochemical transport model. The observed retardation was comparable with the R=3 that was used in this chapter's model simulations.

6.6 Conclusions

The effectiveness of different monitoring network configurations was evaluated for advective and simple reactive transport scenarios. The scenarios were judged for three monitoring objectives: (1) signalling of unexpected threats to the quality of extracted groundwater *(early warning)*, (2) supporting operational decisions by the prediction of future quality changes *(prediction)*, and (3) evaluating protection measures in the protection zone *(protection)*. It appeared essential to determine the travel time distribution in three dimensions.

The results indicate that the location and especially the depth of the observation wells should be carefully chosen, in accordance with the residual transit time to the pumping well (t_ω), the residence times from the surface to the observation screen (t_α) and the degradation and retardation rates. In general, the larger the degradation rates or the retardation, the shallower should the monitoring effort be for effective early warning, prediction and calibration. Therefore, shallow monitoring is frequently the best option for a variety of objectives and conditions.

Shallow monitoring at the boundary of the 10-year protection zone (configuration I) was found most useful for early warning for scenarios with and without a protection zone. The successive vertical screens of configuration I are also appropriate for calibration and prediction purposes. Concentration-depth gradients and concentration time changes measured in those wells are useful to unravel attenuating or retarding processes in the aquifer, to estimate the contaminant loads that are already underway and to calibrate prediction models.

The random sampling of shallowest groundwater inside and outside the protection zone (configurations S and C) is effective for early warning and for judging the effectiveness of protection measures. However, short-term temporal variations of groundwater quality in the first meters below the water table probably hamper a sound interpretation of the collected data when the information goal is to acquire information on the time-dependent input of contaminants.

Deep monitoring at the boundary of the 10 year protection zone (configuration L) was ineffective because of the long forward residence times to the observation wells and the late arrival of contaminants relative to the arrival at the pumping well.

7 A strategy for sampling reactive aquifer sediments in drinking water well fields

7.1 Introduction

Solute transport models are becoming widely used for predicting the evolution of groundwater composition in aquifers used for drinking water production (Griffioen et al. 1998, Brun et al. 1998, Saaltink et al. 1998, Van Breukelen et al. 1998). The models require input on the hydrogeochemical reactivity of the aquifer, but sediment reactivity is one of the largest unknowns in the interpretation of groundwater quality data (Glynn & Brown 1996, Hartog et al. in press, Chapter 1). The spatial variability of sediment reactivity is also an important source of variation encountered in data analysis of regional groundwater quality monitoring programs (Chapters 3 and 5) and local scale monitoring programs at drinking water production sites (Chapter 6).

There is an increasing need for specific monitoring information goals and standardised procedures for sampling sediment reactivity, in order to provide model input for prognoses of the evolution of groundwater quality at the scale of well fields. A sampling strategy should yield a selection of aquifer samples that represents the reactive subsurface properties at the relevant spatial scale of a drinking water production site. It should preferably account for the hydraulic and the chemical heterogeneity of the aquifer sediments, since these have serious implications for the predicted breakthrough of solutes.

This chapter describes procedures for setting up a sampling strategy that was developed to collect reactivity data for use in transport modelling in phreatic well fields and deep-well recharge systems. The aim is to formulate sampling objectives and initiatory data analysis protocols for the reactive properties of sediments at drinking water well fields in order to predict the evolution of the quality of the extracted groundwater. The sampling strategy was designed for the joint Dutch water supply companies, but the results are generally applicable for reactivity sampling at the spatial scale of well fields.

A sampling strategy should provide answers on the following questions:
1. what sampling objectives are suitable and which are the specific information goals?
2. how many samples are required?
3. which sampling depths and sampling volumes are suitable?
4. which sampling methods should be chosen?
5. which geochemical analyses are needed?
6. how are sampling results to be used in transport models?

The presented study focuses on the definition of sampling objectives and specific information goals (question 1) and the definition of initiatory data analysis protocols (question 6). Questions 2 and 3 are addressed briefly. More information on these subjects and on questions 4 and 5 is available in Broers (1999) and Stuyfzand & Meima (2000). The strategy presented in this chapter concentrates on conditions that are relevant to the Netherlands situation, such as unconsolidated deposits, shallow groundwater levels and a flat landscape.

An earlier, short version this chapter was published in *Impact of Human Activity on Groundwater*, IAHS Publication no. 269, 2001. pp. 247-253

7.2 Origin and scales of chemical heterogeneity

The geochemical and reactive properties of sediments are generally dependent on: (1) the provenance of the sediments, (2) the depositional environment, and (3) post-depositional processes, including diagenesis and weathering (Moura & Kroonenberg 1990, Huisman 1998). The depositional environment determines the sorting of grain size and the availability of chemical components during diagenesis. For example, deposits originating from marine environments are generally more reactive than deposits from fluvial environments, because of the presence of shell fragments and a large availability of seawater-ions such as sulphate during sedimentation and diagenesis. Decalcification of carbonate-rich marine sediments is an example of post-depositional changes that affect the reactivity of the sediments. Often, enrichment of reactive mineral phases is observed at the boundary of the sedimentary facies or lithology (Huisman 1998). As a result, typical contents and variations of geochemical and reactive properties are often related to sedimentary facies and structures (Allen-King et al. 1998).

Variation in reactive properties is to be expected at several scales that relate to sedimentary facies. Figure 7.1 illustrates this, using an example of fluvial sedimentary environments (Weber 1986, in Weerts 1996). At the scale of well fields, the scale levels (b), (c), (d) and (e) are relevant. Information on the reactive properties is demanded at scale level (b). The model discretisation of the groundwater transport models is at about scale level (c). However, samples obtained during drilling usually have scale level (d) and geochemical analysis is done on samples of about 1 gram, which corresponds to scale level (e). Geochemical analysis is often done by the sampling of each individual layer that is observed in the drilling core (for example, Huisman 1998). This approach is sensible for understanding geochemical processes and patterns in a scientific study at the local scale. However, such detailed sampling is unnecessary to determine the typical contents of reactive components and their variability at the scale level of a drinking water well field.

Small sample volumes are even disadvantageous for use in transport models, because of the large variability which is inherent to data sets with small sample supports. The probability of

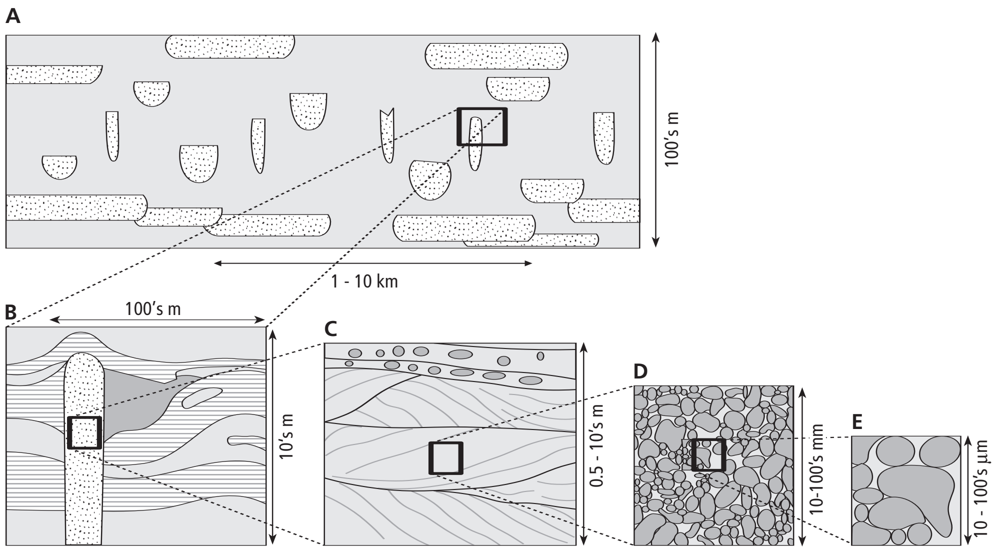


Figure 7.1 - Scale levels of heterogeneity in a fluvial sedimentary environment (from Weerts 1996)

detecting high contents of a reactive component is much larger for sample volumes of 1 gram, compared with sample volumes of several kilograms that are derived from drilling cores. For example, 25 weight percent of pyrite is much more probable using spoon-sized samples than using the homogenized sample of 1 m of drilling core. In mining geostatistics, this effect is called the support effect (Isaaks & Srivastava 1989). Increasing sample volumes generally causes a more symmetric frequency distribution and less variability. Concluding, the sampling strategy should preferably produce samples that are relevant for the scale of the size of the cells of the transport models.

7.3 Effects of hydraulic and geochemical heterogeneity on solute breakthrough

The solute breakthrough in a geochemically heterogeneous system depends on the direction of solute transport relative to the main direction of reactive geochemical variation. This is illustrated for an aquifer volume containing an average content of 100 mmol l^{-1} reactive sediment phase which is flushed with 1 mmol l^{-1} reactive solute (Figure 7.2). For a homogeneously distributed reactive phase, the solute breaks through after 100 pore volumes, assuming instantaneous equilibrium and a reaction stoichiometry of 1 to 1 (Figure 7.2a).

The breakthrough in a geochemically stratified aquifer volume depends on the flow direction. If groundwater flow is perpendicular to the geochemical stratification, the breakthrough is similar to the homogeneous case and occurs after all reactive solid is exhausted (Figure 7.2c). For parallel flow relative to the geochemical stratification, the reactive phase is depleted earlier in the less reactive middle section than in the more reactive upper and lower parts of the aquifer volume (Figure 7.2b). Thus, an estimate of the *average* reactive solid phase content is sufficient to predict solute breakthrough for perpendicular flow. The vertical *variation* of reactive properties should be quantified to predict the solute breakthrough for parallel flow.

For parallel flow the hydraulic heterogeneity becomes important as well (Figure 7.2d). Preferential flow through the less reactive and more permeable parts of the aquifer volume results in earlier breakthrough compared with the hydraulic homogeneous cases. In the example, approximately 93% of transport is through the middle layer and breakthrough

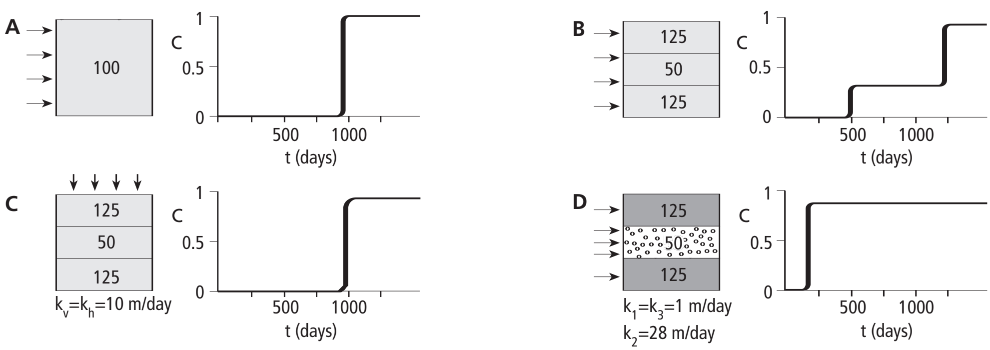


Figure 7.2 - Solute breakthrough in layered reactive systems: (a) homogeneous distribution of reactive phase; (b) geochemically layered, parallel flow; (c) geochemically layered, perpendicular flow; (d) geochemically and hydraulically layered, parallel flow. Dimensions: $\Delta x = \Delta y = 3$ m, hydraulic gradient 0.01, porosity 0.33, initial solute concentration 1 mmol l^{-1}. Reactive solid phase contents are expressed in mmol per litre pore water.

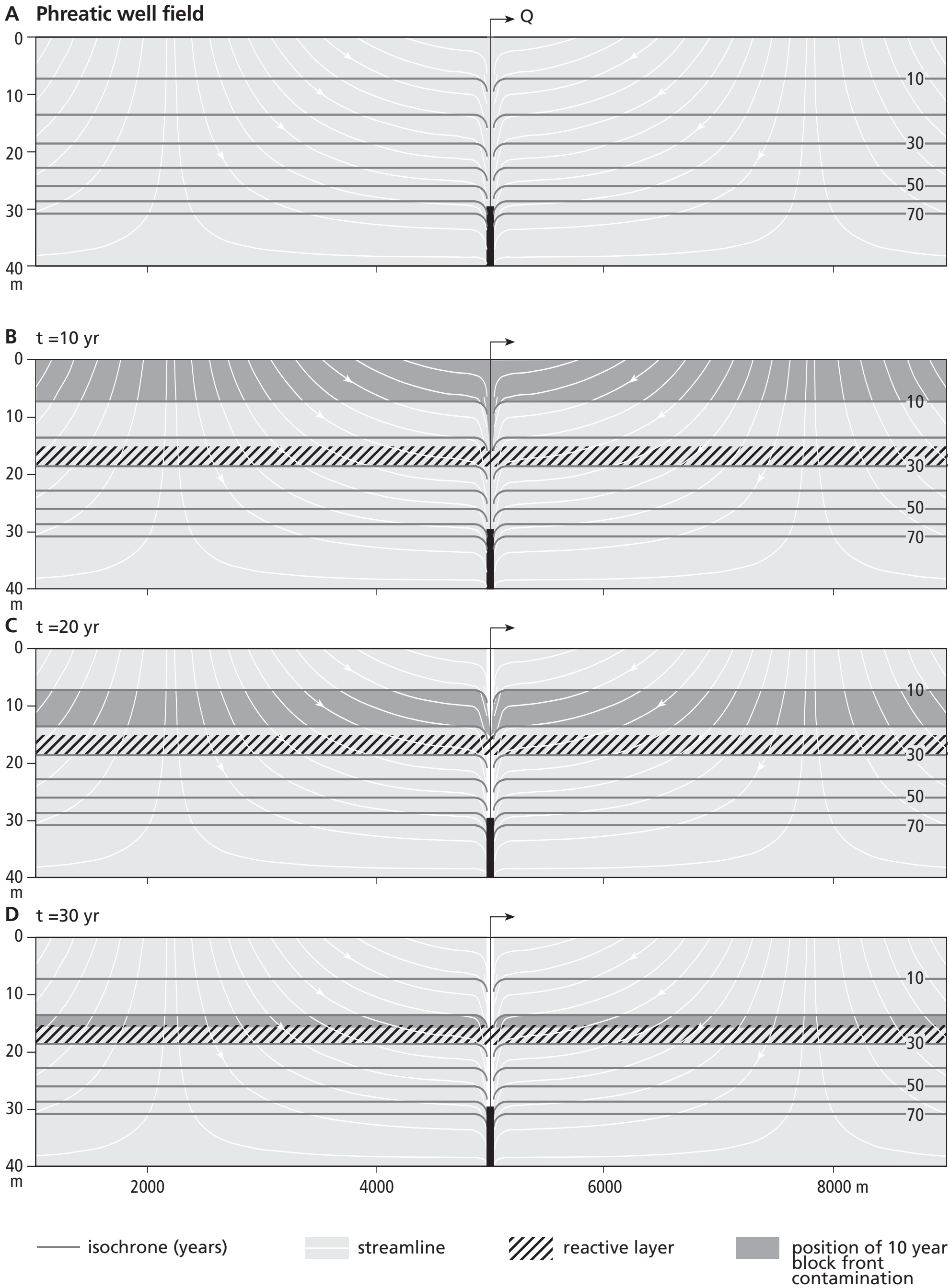


Figure 7.3 - Groundwater flow patterns and movement of a pollution front at a phreatic well field. (A) streamlines and isochrones of equal residence time, (B) + (C) advective propagation of a 10 year block front input at t = 10 year and t = 20 year, (D) retarded propagation after entering a reactive layer.

concentration after t = 200 days equals 0.93. This concentration continues until the last 7% contributes to the outflow concentration (not shown). Concluding, a sampling strategy should account for the direction of the propagation of reaction fronts relative to the main direction of geochemical variation.

7.4 Specification of sampling objectives

Several model assumptions and practical constraints were adopted when designing a sampling strategy for the Dutch water supply companies. First, a predominant horizontal layering of the geochemical, reactive properties was assumed, which relates to the main horizontal layering of the unconsolidated deposits that is common for most Dutch aquifers (for example Hartog et al. in press). This simplification is sensible, because current modelling practice requires a considerable simplification of the hydrogeological and hydrogeochemical situation in order to use complex hydrogeochemical transport models (Griffioen et al. 1998, Stuyfzand 1998b). Lateral variations in hydraulic and geochemical properties *within* the model layers are currently neglected in the model schematization. In deep-well recharge studies, a 1D simulation is often carried out (Brun et al. 1998, Saaltink et al. 1998). Because the presented sampling strategy should yield input data for the simplified transport models, a basic framework of geochemical and hydraulic layering was used to define the sampling objectives.

The second constraint was that data collection is restricted to existing drilling programs for planned observation wells and pumping wells. This implies that the selection of well locations cannot be tuned to the purpose of the sampling of reactivity data and predefined locations should be used. The two constraints imply that the sampling design is mainly directed to assess vertical variations in the reactive properties. Horizontal variations in sediment reactivity can only be assessed when sufficient boreholes become available.

The transport models are normally used for two types of problems: 1. the prediction and monitoring of quality evolution at phreatic well fields and 2. the prediction and process control at deep-well recharge systems. Groundwater flow patterns, the movement of reaction fronts and the input of solutes are significantly different for the two kinds of problems, which is discussed below.

Phreatic well fields
For a phreatic well field (PWF) the aim is to be able to predict the propagation of reaction fronts from diffuse pollution sources, such as pesticides and nutrients. These are introduced as solutes in infiltrating groundwater. Figure 7.3 provides an overview of the groundwater flow patterns at a phreatic well field for a situation where drawdown is negligible (see Chapter 6). In this situation, the groundwater travel times and the vertical position of advective pollution fronts are determined by the groundwater recharge rates, the aquifer thickness and the porosity (Raats 1981, Chapter 6). Assuming spatially averaged inputs of solutes in recharging groundwater, the pollution fronts move vertically, perpendicular to the reactive geochemical layering (Figure 7.3). The assumption of a homogeneous input is reasonable because a spatially heterogeneous input averages out in the breakthrough of the pumping well when the correlation scale is less than 10% of the radius of the contributing area of the well (Duffy & Lee 1992, Beltman 1995, see Chapter 6). The shifts of the pollution fronts are to be measured in the vertical. Figures 7.3 b-d show how an infiltrating pollution front becomes retarded when a reactive layer is encountered. This situation is similar to the one shown in Figure 7.2c and the estimation of the average content of the reactive component will suffice to predict the front movement. Therefore, the chosen sampling objective is to determine the *average* content of the

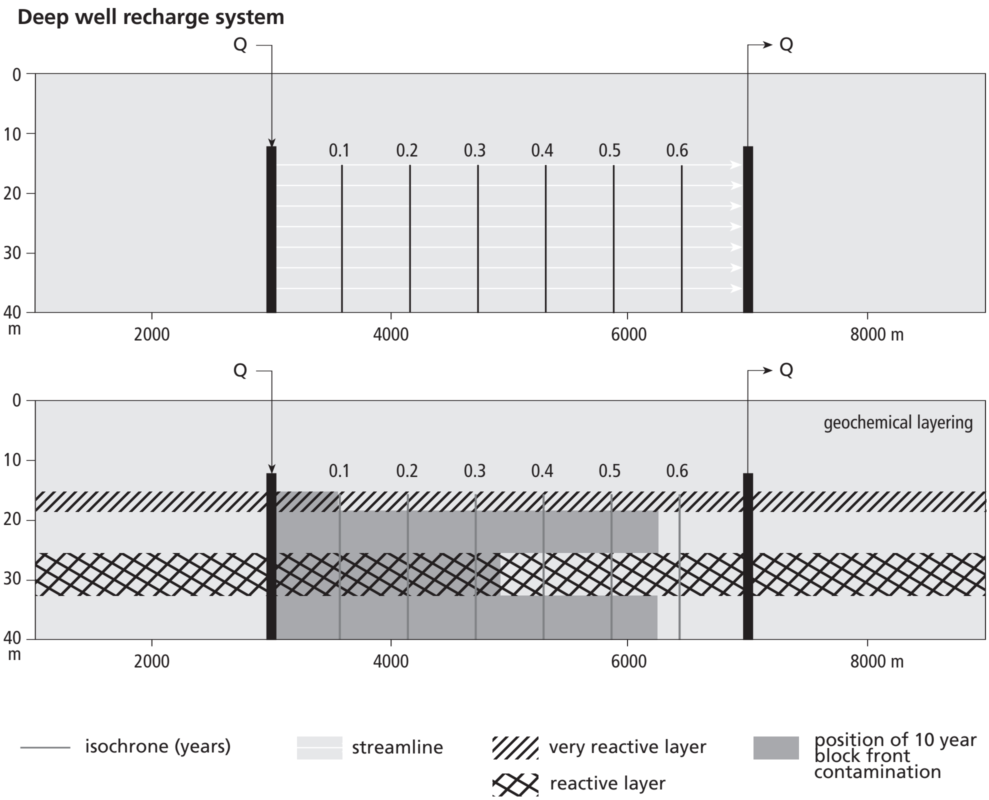


Figure 7.4 - Groundwater flow patterns and movement of a pollution front at a deep-well recharge system. (A): streamlines and isochrones of equal residence time, (B): propagation of a block front input with retardation in two reactive layers.

reactive components and to determine the depth and thickness of the reactive subsurface layers.

Deep-well recharge systems

Deep-well recharge systems (DWR) are generally used to improve water quality using subsurface passage (Peters 1998). Surface water or groundwater is injected into the aquifer and abstracted downstream (Figure 7.4). The injected water normally has different redox, pH or dissolved solid content than the native water in the aquifer. For example, water with dissolved oxygen may be pumped into an anoxic aquifer.

In a deep-well recharge system, groundwater transport is merely horizontal, parallel to the assumed reactive geochemical layering (Figure 7.4). Transit times in the deep-well recharge systems are short compared with transit times at phreatic well fields (months versus years or tens of years). The breakthrough of solutes will proceed differently at different depths, due to the vertical variations in the reactive aquifer properties (Hartog et al. 2002). Figure 7.4b shows the position of the reactive fronts in a hypothetical case with two reactive layers. The situation is similar to the one of Figure 7.2b and comparable breakthrough is likely to be observed.

For deep-well recharge systems, sampling objectives should therefore include the quantification of the *vertical variability* of the reactive properties. Average values of reactive

Table 7.1 - Specific information goals for the reconnaissance and quantification stages

Stage	Phreatic well fields	Deep-well recharge systems
Reconnaissance	first estimate of average content and coefficient of variation	first estimate of percentiles and coefficient of variation
Quantification	estimates of average content and quantified uncertainty (confidence interval)	estimates of percentiles and quantified uncertainty (confidence interval)

contents are not sufficient to predict solute breakthrough. Preferably, also the vertical variability of the hydraulic conductivity is quantified.

In summary, different objectives may be adopted for phreatic well fields and deep-well recharge systems, to account for the different directions of groundwater transport relative to the geochemical stratification. Table 7.1 lists the information goals for phreatic well fields and deep-well recharge systems. Two subsequent sampling stages are foreseen with different information goals. The reconnaissance stage is used to obtain first indications about average contents (PWF) and percentiles (DWR). The quantification stage is subsequently used to improve the precision of the estimates and to quantify the uncertainty using confidence intervals.

7.5 Sampling stages

Phreatic well fields

The reconnaissance stage consists of systematic sampling, using a fixed vertical sampling distance, of one available borehole (Figure 7.5). At this stage systematic sampling is preferred because the depths of reactive layers are not yet known. The disadvantage of systematic sampling is that the fixed distance might by chance correspond to a regular vertical sequence of reactive layering. Therefore, random sampling is advised for the second quantification stage. The reconnaissance sampling yields first indications on the position of the most reactive layers and a first indication of the average content of the reactive component ($\bar{x}$) of the major reactive layer (Figure 7.5).

The second, quantification stage aims at improving the precision of the estimate of the average content of the reactive component and quantifying the uncertainty using a 95% confidence interval. As above, random sampling is recommended to obtain unbiased estimates of the average content of the reactive layer (Figure 7.5). For sample size n less than 30, the parametric 95% confidence interval is given by:

$$\bar{x} - t_{0,975,n\text{-}1} * \frac{s}{\sqrt{n}} \leq u \leq \bar{x} + t_{0,975,n\text{-}1} * \frac{s}{\sqrt{n}} \qquad (7.1)$$

where $\bar{x}$ = average content of sample, s = standard deviation of sample, u = expected value of the population average, and $t_{0,975,n\text{-}1}$= two-sided critical value for the Student's t-distribution with *n-1* degrees of freedom. The random sampling is preferably carried out on several boreholes to ensure that the estimates are spatially representative (Figure 7.5). However, a practical constraint of the presented sampling strategy is the restriction to the use of planned drilling programs (section 7.4).

The sampling results of the reconnaissance stage are used to determine the required sample size in the quantification stage. The sample size is adapted to the amount of variation, based on

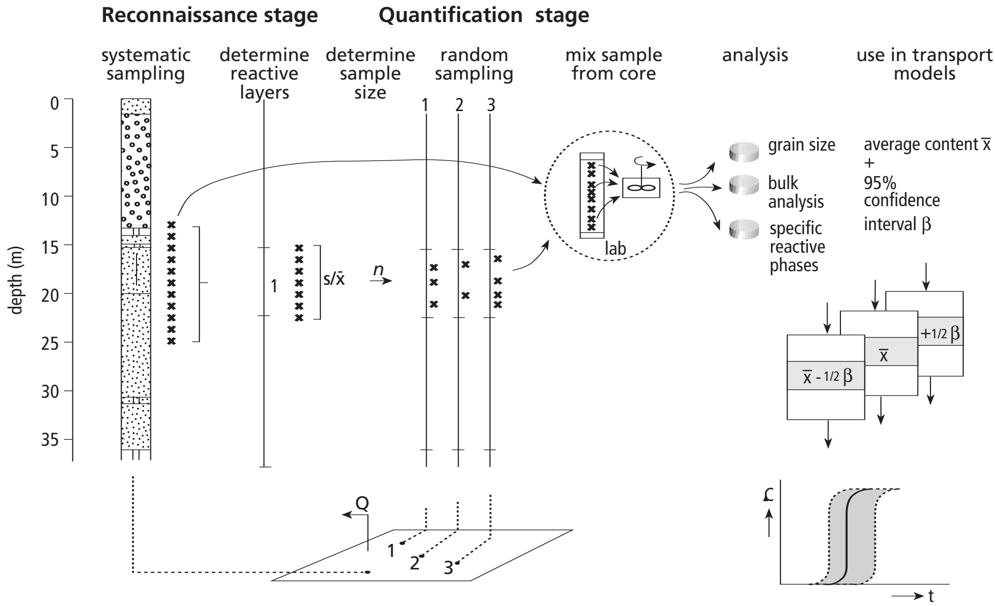


Figure 7.5 - Overview of the sampling strategy at phreatic well fields with indication of the reconnaissance stage (systematic sampling of one borehole) and the quantification stage (random sampling from available boreholes), sample treatment, chemical and physical analysis and data use.

the *coefficient of variation* ($s/\bar{x}$), measured during the reconnaissance stage (Figure 7.5). For a normally distributed data set, the following equation describes the relation between the coefficient of variation and the relative precision as a percentage of estimated average content:

$$r = \frac{t}{\sqrt{n}} * \frac{s}{\bar{x}} * 100\% \qquad (7.2)$$

where r is relative precision, n is sample size and t is critical value for the t-distribution for $n < 30$. A relative precision of 50% means that the probability is less than 5% that the population average u is outside the range $\bar{x} - 0.5\,\bar{x} < \mu < \bar{x} + 0.5\,\bar{x}$.

Following equation (7.2), a larger sample size is required for large $s/\bar{x}$ to obtain similar precision of the estimate. Figure 7.6 shows the required sample size to obtain 30% and 50% relative precision for different values of $s/\bar{x}$. Appendix IV discusses the use of the normal distribution for the determination of the sample size and shows that when using Figure 7.6 the required sample size is likely to be overestimated for $n < 15$.

As discussed in section 7.2, the sampling should provide data that are useful at the scale of model cells of the transport model (scale level (c) of Figure 7.1). Transport models for use at phreatic well fields need model layers of 1 to 5 m thickness in order to accurately predict the downward movement of the pollution fronts and to control the effects of numerical dispersion (Van Vught & van der Eijnden 1999, Uffink et al. 2001, Uffink 2001). Using this vertical discretisation scale, the model results can also be compared with monitoring results from mixed samples of well screens of 2 m length (Chapter 6).

In the Netherlands, geochemically undisturbed samples are usually taken using thin wall tube samplers of about 0.5 m length. A decision was made to acquire average values for a 0.5 m depth interval of the aquifer by homogenizing the whole content of the sampling device or by mixing of a number of random samples from the core (Figure 7.5). The sample volumes

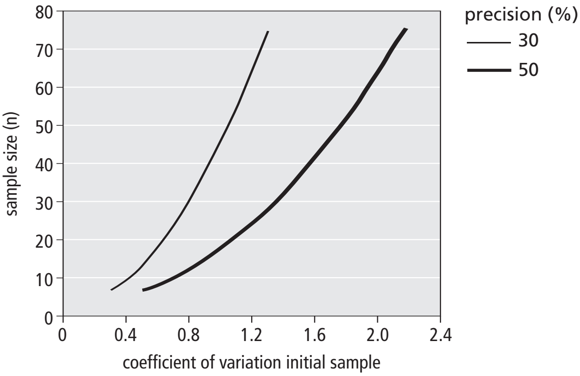


Figure 7.6 - Required sample size to achieve 30% and 50% relative precision of the average content as a function of coefficient of variation $s/\bar{x}$.

obtained have a similar support as the model cells, and no further correction for the support effect is necessary in the data analysis.

Deep-well recharge systems

The sampling stages at deep-well recharge systems correspond largely to the description for the phreatic well fields. However, the aim is to estimate the percentiles of the frequency distribution in order to quantify the variability of the reactive properties with depth. Systematic sampling of the entire infiltration aquifer is recommended for the reconnaissance stage in order to obtain information over the complete depth range (Figure 7.7). The reconnaissance sampling yields a first indication of the frequency distribution in the infiltration aquifer, including an initial estimate of the percentiles and the coefficient of variation $s/\bar{x}$.

This information is subsequently used to determine sample size for the quantification stage. The sampling in this stage aims at quantifying the precision of percentiles of the frequency distribution, using 95% confidence intervals. Non-parametric methods are preferred for the assessment of the confidence intervals, because parametric confidence intervals on percentiles depend strongly on the assumed frequency distribution (Helsel & Hirsch 1992). This effect is less for confidence intervals on the average contents at phreatic well fields as explained in Appendix IV.

In the example discussed in the next section, the 12.5, 37.5, 62.5 and 87.5 percentiles are used for illustration. The precision may be assessed using a non-parametric 95% confidence interval on the estimated percentile, using a normal approximation to the binomial distribution for sample size $n > 20$ (Helsel & Hirsch 1992). The confidence interval is calculated as:

$$R_l = np + z_{0.025} \sqrt{np(1-p)} + 0.5 \qquad (7.3)$$

$$R_u = np + z_{0.975} \sqrt{np(1-p)} + 0.5 \qquad (7.4)$$

where p is the percentile (0.75 for the 75 percentile) and $z_{0.025}$ and $z_{0.975}$ are the critical values of the normal distribution. The data are ordered according to rank and R_u and R_l represent the ranks that correspond to the upper and lower limits of the confidence interval.

No parametric methods are available to determine the required sample size for obtaining an estimate with a pre-specified relative precision without assuming a distinct probability

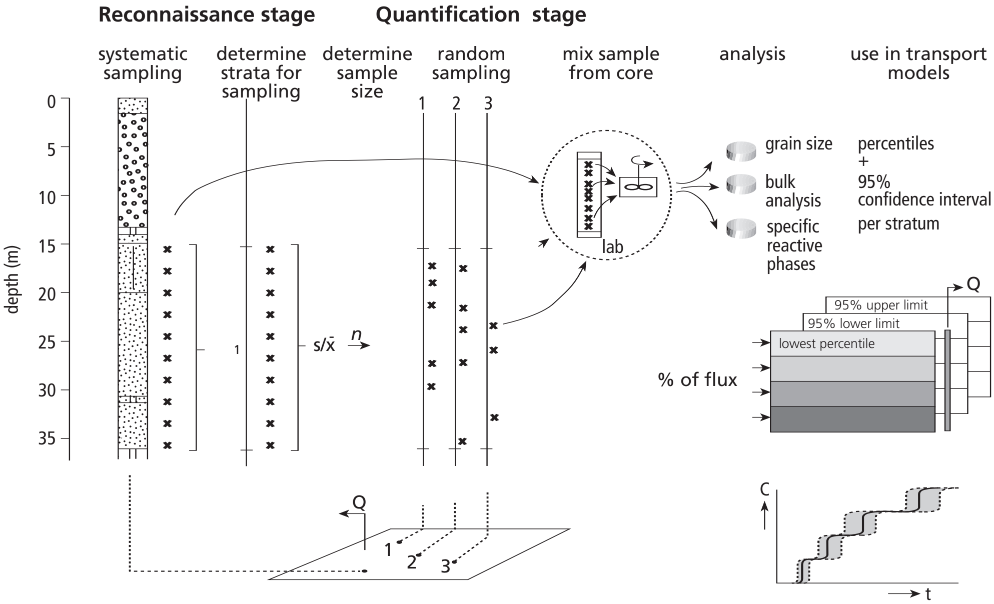


Figure 7.7 - Overview of the sampling strategy at deep-well recharge systems with indication of the reconnaissance stage (systematic sampling of one borehole) and the quantification stage (random sampling from available boreholes), sample treatment, chemical and physical analysis and data use.

distribution. Instead, bootstrapping may be used to get a general idea of the required sample size. In appendix IV, data from the Oostrum aquifer are used for bootstrapping (see next section for details about the Oostrum aquifer). The bootstrapping results indicated that a larger sample size is required to attain accurate estimates of a percentile than to attain accurate estimates of the average content. A suggestion for sample size is given in Appendix IV.

Transport models for use at deep-well recharge systems usually have model layers of 5-10 m thickness (for example Stuyfzand 1998). Therefore, mixing the contents of the 0.5 m thin wall tube samplers was considered reasonable to attain sample volumes of the largest possible support as input for the model cells.

7.6 Use of the sampling results in transport models

Specification of data analysis protocols is considered an essential step in designing sampling and monitoring programs (Adkins et al. 1995). Here, data of the Oostrum aquifer have been used to illustrate how the collected data can be used in transport models, using the estimates and the 95% confidence intervals to quantify uncertainty. The Oostrum aquifer was chosen, because of the available large set of reactivity data.

The Oostrum aquifer

The Oostrum aquifer has been studied in order to trace the origin of heavy metals and arsenic in pumped groundwater (Broers & Buijs, 1998, van Helvoort et al. 2000, Schipper et al. 2000). The aquifer is situated in the northern part of the Dutch province of Noord-Limburg (Figure 7.8). The pumped aquifer consists of a 30 m thick series of coarse fluvial sands of Pliocene age.

A total of 57 samples has been collected from 8 bore holes that are situated within one

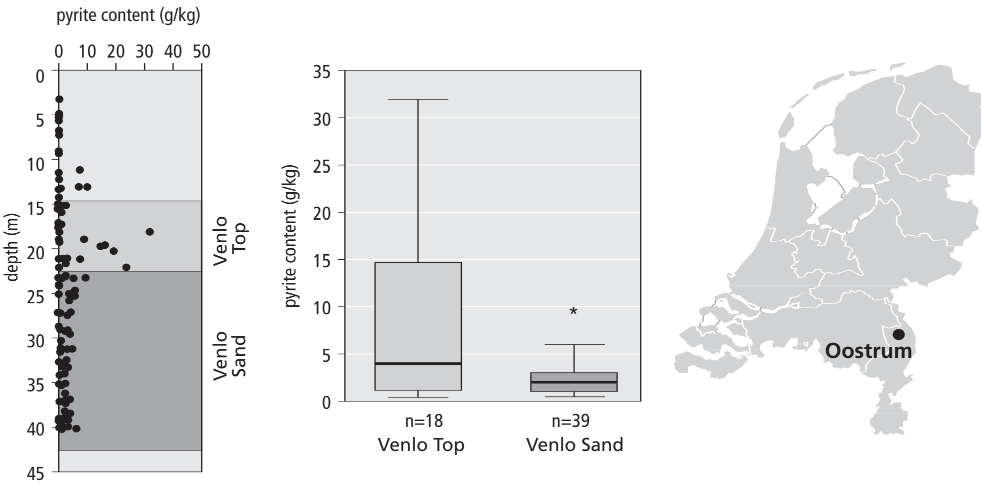


Figure 7.8 - Pyrite contents in the Oostrum aquifer. (A) Pyrite profile with depth, indicating the Venlo Top and Venlo Sand strata. (B) Boxplots of pyrite content for Venlo Top and Venlo Sand.

square kilometre. The aquifer sediments were sampled using thin wall tube samplers and the grain size distribution, total element content and specific reactive phases were analyzed. Two main reactive strata with pyrite are present of about 7 and 20 m thickness: Venlo Top and Venlo Sand (Figure 7.8). Both strata consist of the coarse fluvial sands, but the Venlo Top stratum has larger organic matter contents and contains thin clay lenses. Total-sulphur was measured with a Strohlein CS-MAT 5500, and was recalculated into pyrite contents. Table 7.2 shows summary statistics for the pyrite content measured in the two strata. The Venlo Top stratum has larger pyrite contents, and larger variation and skewness than the Venlo Sand stratum.

Table 7.2 - Summary statistics for pyrite content (g kg^{-1}) in two strata. The parametric 95% confidence interval was obtained using equation (7.1)

	n	Average $\bar{x}$	95% conf. interval of $\bar{x}$	Coeff. of variation $s/\bar{x}$	Percentiles		
					P25	P50 (median)	P75
Venlo Sand	39	2.5	2.5 ± 0.6	0.75	1.0	2.1	3.0
Venlo Top	18	8.0	8.0 ± 4.5	1.16	1.3	3.9	13.1

Calculation of solute breakthrough

In the example, the solute breakthrough has been calculated for a step input of C = 0 mmol NO3 l^{-1} for and C = 2.8 mmol l^{-1} for $t > 0$. The nitrate reacts simultaneously with pyrite in the aquifer according to:

$$5FeS_2 + 14NO_3^- + 4H^+ \rightarrow 5Fe^{2+} + 10SO_4^{2-} + 2H_2O + 7N_2 \qquad (7.5)$$

Using the step input, instantaneous equilibrium and a finite amount of reactive solid phase, the propagation of the nitrate front can be described as a retardation reaction (following Dria et al. 1987):

$$v_{front} = \frac{v_{water}}{R}, \text{ where } R = 1 + \frac{\rho_b \Delta S}{\varepsilon \Delta C} \tag{7.6}$$

where R = retardation factor, ρ_b = bulk mass density, ε = porosity, ΔS= content of solid phase reactant (mol kg^{-1}) and ΔC = concentration of reactant in water (mol l^{-1}). Recalculating the content of the solid phase reactant into the aqueous concentration of the solid phase reactant, and including the reaction stoichiometry, the front velocity becomes:

$$v_{front} = \frac{v_{water}}{1 + \left(\frac{\Delta S_{solid}}{\Delta C} . r_r\right)}$$

where r_r is the reaction stoichiometry (mol of solid reactant leached by 1 mol of reactant in water). For the reaction between pyrite and nitrate, the velocity of the nitrate front becomes:

$$v_{front} = \frac{v_{water}}{1 + \left(\frac{FeS_2}{NO_3} . \frac{14}{5}\right)} \tag{7.7}$$

using reaction stoichiometry 5:14.

Phreatic well fields

The results from Table 7.2 can directly be used as input for the transport models at phreatic well fields. The retardation of the downward velocity of the nitrate front has been calculated following equation (7.8), assuming a porosity of 0.3 and a bulk mass density of 1.75 kg dm^{-3}. Retardation of the front equals 133 ± 32 for transport in Venlo Sand and 423 ± 237 for transport in Venlo Top, where ± refers to the upper and lower limits of the 95% confidence interval. Using the step input and the pyrite and retardation data of Venlo Sand, it takes between 800 and 1310 years to exhaust all pyrite from the reactive layer of Figure 7.3, assuming aquifer thickness is 40 m, a reactive layer between 15 and 19 m depth, porosity is 0.3 and groundwater recharge rate is 0.3 m year^{-1}.

Breakthrough at the pumping well depends on the configuration of reactive layers in the subsurface and the 3D flow field. An example of solute breakthrough at the pumping well is given for the flow field and reactive layer configuration given in Figure 7.3, using the previously used step input and the reactivity data of the Venlo Sand stratum.

The concentration breakthrough for this situation is best understood by examining the breakthrough for a well that *fully penetrates* the aquifer, as depicted in Figure 6.2 (Chapter 6). For that situation, nitrate contaminated water in the upper part of the aquifer above the reactive layer contributes immediately to the concentration in the well, according to the front movement of Figure 6.2. Using equation (6.7) (Chapter 6), the reactive layer is reached by the advectively moving front at t = 22 years and further breakthrough is retarded after this time (Figure 7.9). Using the retardation factor R = 133 ± 32 for Venlo Sand, the concentrations in the pumping well only change marginally after t = 22 years, and the upper and lower limits of the confidence intervals cannot be distinguished during the first 80 years. The effects of the uncertainty of the estimated reactivity is better visible when using a retardation of R = 3 ± 1 (Figure 7.9). The figures show that the uncertainty in the solute breakthrough is directly determined by the uncertainty of the estimated reactivity.

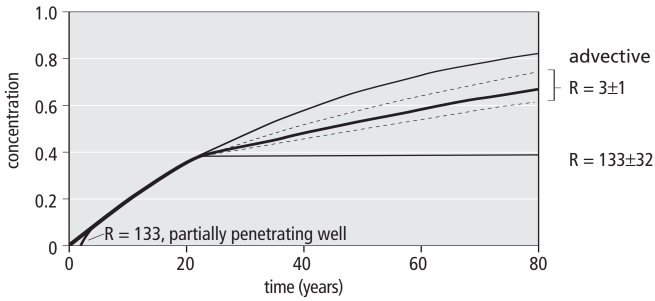


Figure 7.9 - Solute breakthrough at the pumping well of a phreatic well field for the following situations: (1) advective transport, (2) reactive layer between 15 and 19 m depth for retardation factor R = 3 ± 1 and R = 133 ± 32. Dashed lines indicate the 95% upper and lower confidence limits of solute breakthrough for retardation factor R = 3.

When a partially penetrating well is situated *below* a reactive layer (Figure 7.3), the pyrite in the reactive layer has first to be exhausted before breakthrough occurs. In the flow field of Figure 7.3, the vertical flow of contaminated water is concentrated in a cylinder with radius r_o of approximately 50 m. Thus, a cylindrical pore volume $\pi r_o^2 \varepsilon d$ of the reactive layer must be exhausted, with radius r_o = 50 m, thickness d = 4 m and porosity ε = 0.3. In the example with R = 133, the cylindrical pore volume directly above the well was exhausted after 3 years (Figure 7.9). The example shows that the results are very similar to the situation with a fully penetrating well, except for the first breakthrough. The results further indicate that both the average reactive content and the *depth* of the reactive layers determine the eventual concentrations in the pumping well.

Deep well recharge

For a deep-well recharge system, the observed vertical variations in pyrite content in the transport model must be included in the transport model.

The estimated percentiles for the Venlo sand and Venlo Top strata are listed in Table 7.3. The 95% confidence interval for the estimated percentiles was computed non-parametrically using equations (7.3) and (7.4). Figure 7.10 shows the estimated percentiles and 95% confidence intervals for Venlo Sand and Venlo Top. The estimates of the skewed data of Venlo Top had much larger uncertainty than those of Venlo Sand.

In the example, a geochemical layered, but hydraulically homogeneous subsoil was assumed. The measured variation of the pyrite contents is modelled in a four-layer schematization (Figure 7.11a). The model layers do not describe the physical structure of the subsurface, but

Table 7.3 - Percentiles and 95% confidence limits of pyrite contents (g kg^{-1}) used for the four-layer transport model

	n	P12.5 (95% interval)	P37.5 (95% interval)	P62.5 (95% interval)	P87.5 (95% interval)
Venlo Sand	39	0.8 (0.5-0.9)	1.8 (0.9-2.2)	2.5 (2.0-3.2)	3.9 (3.2-5.9)
Venlo Top	18	0.8 (0.4-1.2)	2.2 (0.8-4.9)	6.2 (2.0-15.7)	18.4 (8.5-31.7)

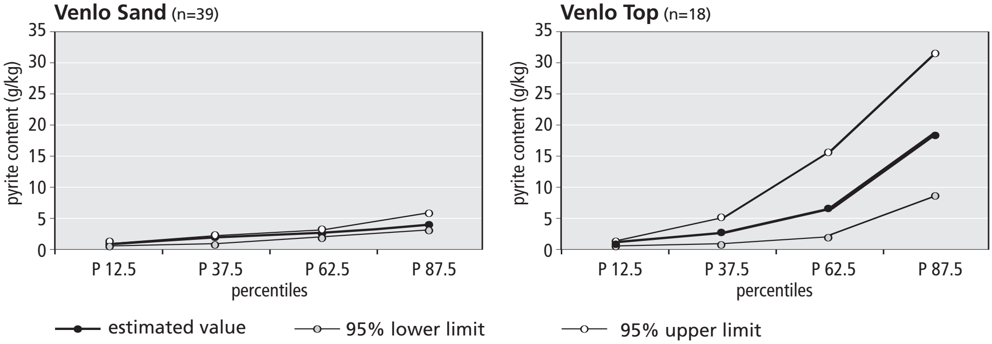


Figure 7.10 - Estimates of 4 percentiles and corresponding non-parametric 95% confidence intervals for Venlo Sand and Venlo Top.

represent the statistical characteristics of the aquifer. Each model layer represents one-quarter of the total variation in pyrite content (P0–P25, P25–P50, P50–P75 and P75–P100) and is characterized by the typical value of the pyrite content for that quarter (Figure 7.11). The estimated 12.5th percentile (P12.5), which is the median of the lowest 25% of the frequency distribution, characterizes the least reactive parts of the aquifer (Figure 7.11a). Each model layer receives 25% of the horizontal flux to achieve a hydraulically homogeneous model. In this way, the four-layer model assumes that the least reactive zones of the sampled sedimentary stratum are mutually connected. This represents the extreme condition where the geochemical reactivity has the largest effects on solute breakthrough. The other extreme condition is where the geochemical variation is randomly distributed within the flow domain.

The four-layer model may be used to calculate the breakthrough of a continuous input of 2.8 mmol l^{-1} nitrate in the vertically heterogeneous deep-well recharge system for Venlo Sand (Figure 7.11b) and Venlo Top (Figure 7.11c). The first step of the breakthrough is determined by the least reactive sublayer of the aquifer, where pyrite is exhausted after about 40 pore volumes.

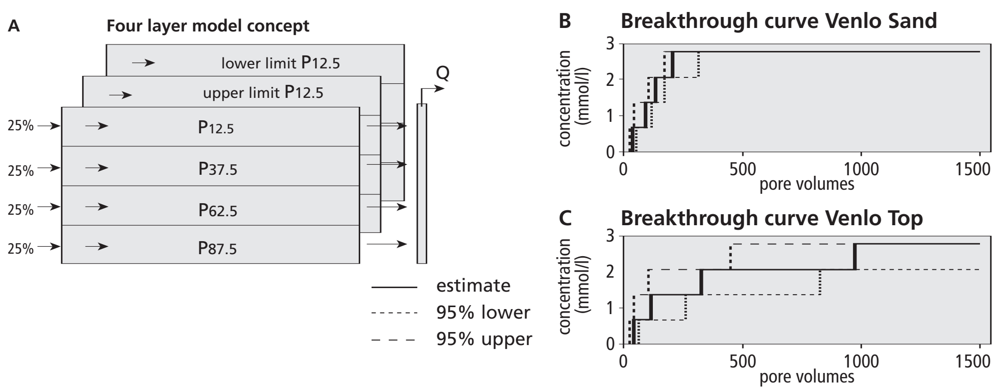


Figure 7.11 - Hydraulically homogeneous four-layer model concept and estimated breakthrough (solid lines) and 95% confidence limits around the estimate (dashed lines) for a deep-well recharge system in a geochemically layered aquifer for the Venlo Sand and Venlo Top strata.

The first breakthrough occurs simultaneously for both strata, despite the larger average pyrite content of Venlo Top. Ultimately, the larger average pyrite content and variability of Venlo Top result in a more retarded breakthrough relative to Venlo Sand. The confidence interval boundaries of Figure 7.10 may be used to determine upper and lower limits for the calculated breakthrough in the sensitivity analysis (Figure 7.11 b+c). The heterogeneous Venlo Top stratum exhibits a larger uncertainty on the breakthrough curve compared with the relatively homogeneous Venlo Sand stratum.

In the four-layer model a connection is assumed between the least reactive parts of the aquifer, which results in early breakthrough and tailing at a later stage. Assuming a random distribution of pyrite in the aquifer, the average content is used to calculate the breakthrough after 132 ±32 and 423 ±237 pore volumes for Venlo Sand and Venlo Top, respectively. The most appropriate model may be selected on the basis of the observed sedimentary structures in the sampled stratum and monitoring results from observation and pumping wells.

Furthermore, in this example a *hydraulically homogeneous* subsoil was assumed. Hydraulic heterogeneity is easily added to the four-layer models by specifying different fluxes in each of the layers (Figure 7.12) similar to the implementation in Easy-Leacher (Stuyfzand 1998). However, this requires information on the correlation between hydraulic and geochemical heterogeneity, which can be partly obtained from the grain size distribution of the samples (Christiansen et al. 1998). Therefore, standard measurements of the grain size distribution are advised in the sampling strategy (Figure 7.7).

The method requires a choice on the number of model layers. Because more layers yield a higher resolution, but the precision of the result decreases, this choice is a balance between the resolution and the precision of the results. The number of model layers should be tuned to the available number of samples and the observed variation. Developing models consisting of more than four layers is only sensible when accurate estimates of the low and high percentiles of the frequency distribution can be obtained from the data. This requires a large number of samples or a stratum with little variation in reactivity.

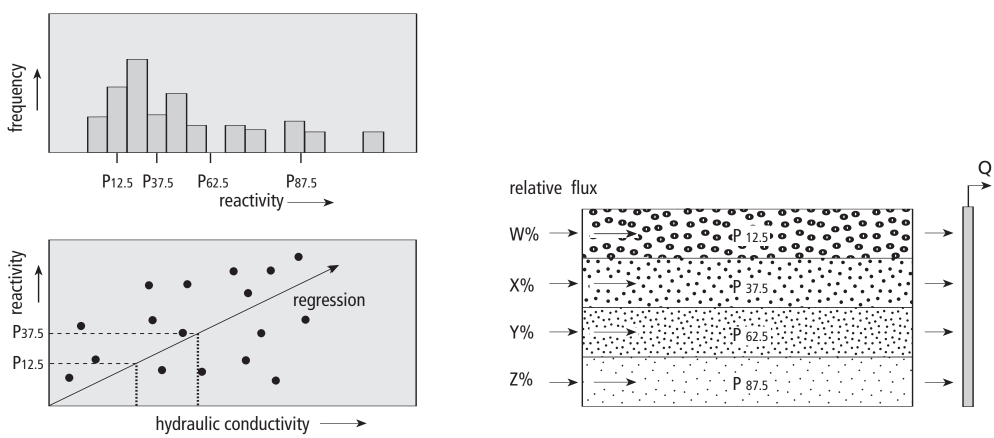


Figure 7.12 - Four-layer model concept for hydraulically heterogeneous subsoils, using specific fluxes and reactivity for each model layer, based on regression between hydraulic conductivity and reactivity of the sediment samples

7.7 Discussion

The lack of data on subsurface reactivity is one of the largest obstacles to the prediction of changes in groundwater quality at the regional and local scale. Monitoring programs at phreatic well fields that aim at the prediction of future groundwater quality (Chapter 6) should preferably include sampling programs for reactivity data. Combining the groundwater quality monitoring results of observation wells and pumping wells with results from the reactivity sampling will improve the predictions and help to interpret the monitoring data.

The present study focused on the definition of information goals for the sampling of reactivity data. A concept of geochemical layering was considered most appropriate for application using the currently used transport models. The complete, vertical and lateral characterization of reactivity, such as done by Davis (1993) or modelled by Christiansen et al. (1998) was not considered feasible for practical use in prediction models at drinking water well fields. However, the presented sampling strategy needs further testing, especially for deep-well recharge systems where the vertical and lateral variations in reactivity may be exploited. For example, the tailing of the solute breakthrough that was observed in the conceptual four-layer models, is easily interpreted as kinetic behavior of reactions when using a 1D modelling approach (Brun et al. 1998, Saaltink et al. 1998). For phreatic well fields, the use of vertically averaged data of the reactive layers suffices to predict the movement of the pollution fronts and the solute breakthrough in pumping wells. However, horizontal variations in reactivity will probably influence solute breakthrough in the pumping wells, especially when they are correlated with hydraulic heterogeneities at a large scale compared with the radius of the contributing area of the well (Duffy & Lee 1992).

7.8 Conclusions

The study shows that different sampling objectives and specific information goals are necessary for phreatic well fields and deep-well recharge systems to account for the different direction of the propagation of reaction fronts relative to the main direction of geochemical variation. The examples revealed that the uncertainty of the estimated sample statistics directly affects the uncertainty of the calculated breakthrough. Hence, a proper design of the sampling strategy helps to reduce the uncertainty of the transport model predictions.

8 General conclusions and suggestions for further work

8.1 General conclusions

This study has shown that the effectiveness of monitoring is significantly improved by integrating statistical, hydrological and hydrogeochemical methods and information in the design, the data analysis and the evaluation and optimization of groundwater quality monitoring networks. The results indicate that an effective monitoring strategy for diffuse contaminants needs well-focused monitoring objectives and statistical information goals and substantial understanding of the hydrological and hydrogeochemical system properties (Figure 1.2).

Hydrological and hydrogeochemical system properties
The most important hydrological and hydrogeochemical *system properties* for the monitoring of diffuse contaminants are: 1. the travel time distribution in the aquifer, and 2. the reactivity of the introduced solutes and the subsurface sediments which control retardation and transformation processes and propagation of contamination fronts. Some general conclusions on both aspects are given below.

Travel time distribution
The travel time distribution in the aquifer determines the advective transport of contaminants. In monitoring studies a distinction should be made between (1) *residence time* - the travel time between the earth's surface and a specific point in the aquifer, which is a feasible location of an observation well, and (2) *residual transit time* - the travel time between the observation well and the outflow point, which could be a pumping well or a surface water body.

The calculation and evaluation of *residence times* should be the first step in monitoring network design. In the examples given in previous chapters, residence times were evaluated for typical Dutch conditions, such as shallow groundwater tables, permeable unconsolidated deposits and a flat landscape. Under these conditions, a horizontal pattern of residence times isochrones and a gradual increase of groundwater age with depth is observed in recharge areas without a superficial drainage network. In areas with an extensive superficial drainage network, however, local flow systems cause large variations in residence times at a specific depth in the aquifer and relatively old groundwater at shallow depth in the vicinity of drains. This residence time pattern was confirmed by tritium age dating in two regional networks. The residence time distribution had large consequences for the typical concentrations and the proportion of contaminated groundwater in recharge and drained areas. Residence time isochrone patterns in areas with complex subsurface geology, hilly terrain or strongly variable groundwater recharge will differ from the studied situation and adaptations to the presented monitoring strategies will obviously be necessary. Nevertheless, assessment of residence time patterns, using for example groundwater modelling and age dating, should be advised as a general first step in monitoring design.

When the objectives of the monitoring study are to predict water quality in pumping wells or in the base flow to surface water, the *residual transit times* between the observation wells and the outflow points should also be considered in the monitoring design. The position and depths of observations wells must be based on both the residence time isochrones and the residual transit time isochrones which preferably are deduced from a 3D groundwater flow field.

Reactive processes
The reactive behaviour of infiltrating solutes is often difficult to predict in the design stage of monitoring networks because information on the reactivity of the subsurface is seldomly known a priori. Spatial information on soil types or geological formations is often used in the design stage as surrogate information for reactivity, but relations between geology and reactivity have not been studied well. Therefore, information on the reactive processes must be collected after the initial design or during a preliminary survey.

The use of hydrogeochemical calculations, such as mineral saturation indices and process-based geochemical indicators, is strongly advised in the data analysis stages of the monitoring networks. Preferably, the use of hydrogeochemically based water quality indicators is integrated in the regular, statistical methods that are used to match the monitoring objectives and specific information goals. The regional case studies presented in Chapters 3, 4 and 5 showed that the use of these indicators helps to identify and to interpret spatial and temporal contamination patterns. Especially the use of the indicator 'oxidation capacity' (OXC) helps to identify the impact of agricultural pollution on groundwater quality and the detection and interpretation of quality changes in time. Distinctive groundwater quality patterns between similar areas in the provinces of Noord-Brabant and Drenthe would have remained unnoticed without the additional geochemical indicators. These indicators are also valuable for the evaluation and optimization of the monitoring networks, because they often show less variation and less skewed concentration frequency distributions.

Adaptations to the monitoring set-up are necessary when there are indications of retardation or transformation processes in the studied aquifers. The design of most monitoring networks for groundwater quality implicitly assumes advective transport of solutes and advective travel times are used to select screen depths and monitoring frequencies. However, the regional examples show that many contaminants do not behave conservatively and are delayed relative to the groundwater itself. Given the general increase in residence times with depth, the monitoring of retarded solutes requires shallower screen depths for the timely detection of quality changes compared with advectively transported solutes. Identification of retardation and transformation processes is best achieved using wells with multiple screens in the shallow groundwater (less than 20 m depth under the studied Dutch conditions).

An alternative to adapting monitoring depths is the use of indicators that behave conservatively under specific conditions. The indicator OXC performed well for this purpose in the regional network of Noord-Brabant, where it was assumed that pyrite oxidation was the most important transformation process in the aquifers.

Monitoring objectives and statistical information goals
Monitoring efficiency is best when the monitoring information goals and statistical information goals are defined before the actual installation and operation of the monitoring networks and data analysis protocols are established during the design stage. Often, a range of monitoring objectives is present. For regional monitoring these include:

1. presenting an overview of groundwater chemical status over the whole region (the reference function)
2. signaling groundwater contamination and changes in groundwater quality (the signal function), and
3. determining the magnitude of groundwater contamination in vulnerable areas, both in time and space.

In the Dutch regional networks, these general monitoring information goals were successfully integrated in a framework with area-specific statistical information goals for areas with low, moderate and high risks for contamination of deep groundwater. Four information goals refer

Table 8.1 - Specific monitoring information goals for high, moderate and low-risk homogeneous areas for the Dutch regional networks

Monitoring information goal	**High-risk areas**	**Moderate-risk areas**	**Low-risk areas**
Ambition level	**High**	**Moderate**	**Low**
Overview of chemical status			
A Determine typical values (medians/percentiles)	x	x	x
B Identify differences between areas	x		x
C Signal the exceeding of environmental standards		x	
D Determine the proportion of contaminated groundwater *	x		
Temporal trends			
E Signal temporal trends		x	
F Determine the median temporal trend	x		
G Determine the temporal trend in the proportion of contaminated groundwater	x		

* i.e. the proportion of groundwater exceeding the allowed thresholds

to the assessment of the time-averaged chemical status of groundwater at specified depths (Table 8.1). Three information goals emphasize the assessment of temporal changes in groundwater quality (E to G).

Figure 8.1 gives an overview of the monitoring strategy and the use of these information goals in data analysis protocols for regional networks. In this example, high-risk, moderate-risk and low-risk areas were defined on the basis of the geohydrological subdivision into recharge, intermediate and discharge areas. However, land use or geological factors can additionally be used for the delineation of the areas. The position of a 20 year advective pollution block front in the conceptual flow field is indicated in this Figure (scenario C2, Chapter 2). The position of the front was based on the 10 and 30 years residence time isochrones and represents the situation at 30 years after the start of the pollution.

The example shows the differentiation of sample size, monitoring depths and monitoring frequencies for high-risk, moderate-risk and low-risk areas that was used to match the area-specific information goals of Table 8.1. The two regular monitoring depths of the Dutch regional and national monitoring networks of approximately 10 and 25 m depth were chosen as the basis for monitoring. Monitoring frequency was reduced in the deeper screens of the moderate-risk areas and in the low-risk areas. Especially, this differentiation in monitoring frequencies yielded opportunities for cost-reduction in the Dutch networks. Extra screens in the high-risk areas enable the making of concentration-depth profiles. This configuration yields information on retardation and transformation processes which can be used to interpret temporal trends (Chapter 5). A four-yearly sampling frequency is sufficient for the extra screens in order to combine the concentration-depth information with the time series data from the annual sampling of the regular monitoring depths.

In such an approach, the sample size in the high-risk, moderate-risk and low-risk areas should be tuned to the desired precision of the results and on the amount of variation within the areas. Results from the Noord-Brabant and Drenthe networks indicate that a relatively large sample size is needed in the drained moderate-risk areas to account for the variations in residence times (Chapter 4).

In summary, the advantage of the use of area-specific information goals and area-specific designs in regional monitoring are:

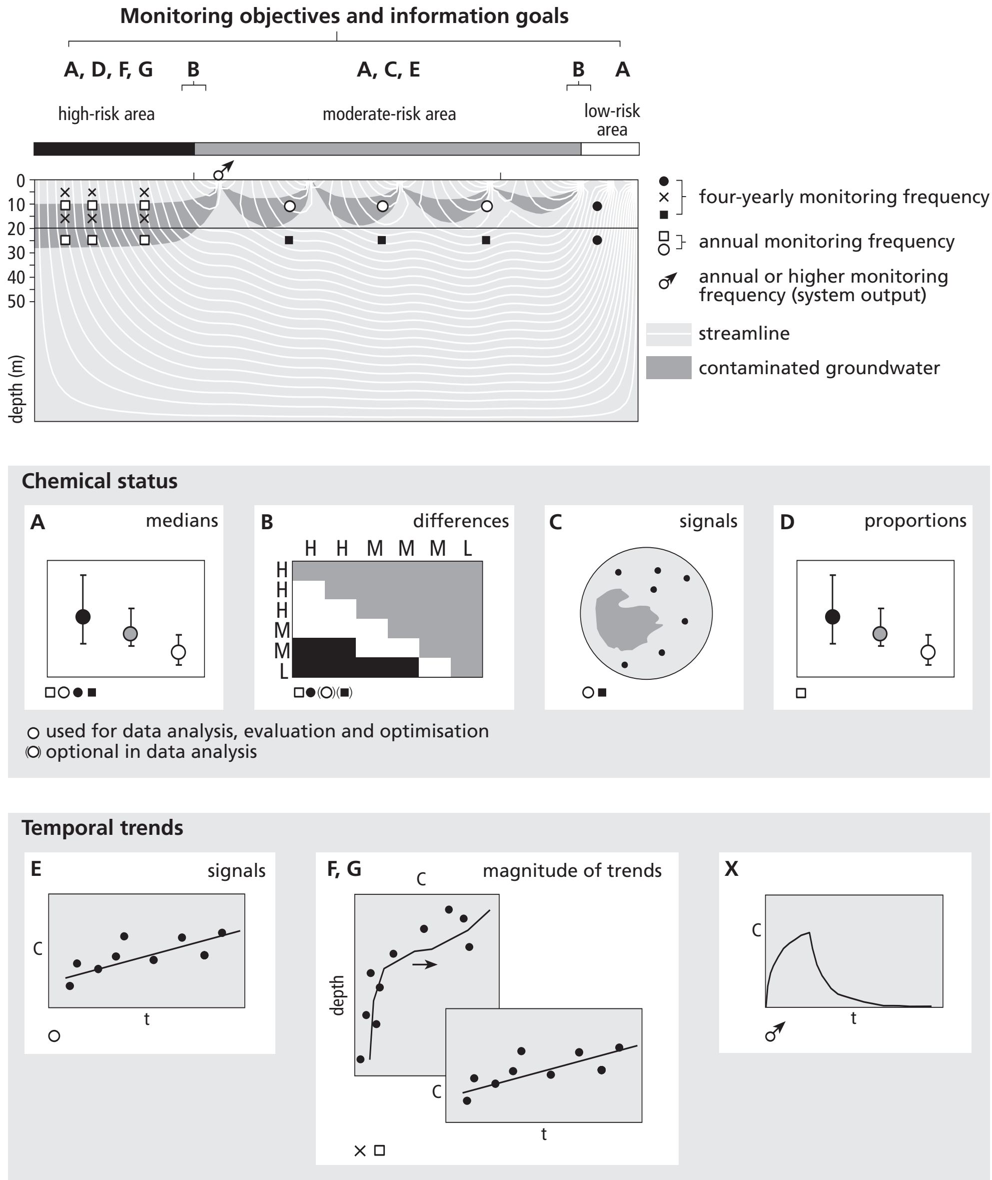


Figure 8.1 - Overview of a regional monitoring strategy, using area specific information goals for low-risk, moderate-risk and high-risk areas (A-G, see Table 8.1) and differentiated monitoring frequencies, monitoring depths and sample size. The position of the 20 year advective block front was based on 10 and 30 year residence time isochrones in the flow field of scenario C2 (Chapter 2). Data analysis protocols for assessment of chemical status and temporal trends are indicated for the well types and observations screens that are shown in the flow field. Details on the data analysis methods are given in Chapters 3, 4 and 5.

1. opportunities to integrate geohydrological and hydrogeochemical system properties in the delineation of areas and to adapt monitoring frequencies and depths to the a priori expected advective and reactive propagation of pollution fronts
2. strong focus on the high-risk areas, which results in less uncertainty in the estimates of typical concentrations and the proportions of contaminated groundwater
3. feasibility of cost-reduction in low-risk and moderate-risk areas by reducing monitoring frequencies, while maintaining the opportunities for the detection of groundwater contamination and for the assessment of a regional overview of groundwater chemical status
4. improved opportunities to detect the effects of retardation and transformation processes, combining concentration-depth and concentration-time information in the areas with high risks for contamination of deep groundwater.

Monitoring objectives for phreatic well fields used for public water supply are completely different from the objectives of regional monitoring. They include (Baggelaar 1996):
1. to fulfil the legal monitoring necessity
2. to reassure customers
3. to *signal* unexpected or new threats to the quality of extracted groundwater
4. to support operational decisions by the *prediction* of future quality changes
5. to evaluate *protection* measures in the protection zone.

Effective monitoring strategies for these 5 information goals are presented in Figure 8.2, based on residence times and residual transit times and scenarios of simple reactive transport. The figure shows the position of a 10 year block front, 10 years after the start of the contamination for a scenario with a protection zone (scenario 1b, Chapter 6). Effective monitoring configurations for phreatic well fields have relatively shallow screen depths to meet the early warning and prediction objectives. The effective monitoring configurations (Figure 8.2) differ from the configurations for regional monitoring (Figure 8.1) because the main monitoring concern is the concentration in outflowing groundwater in the pumping well and not the contamination risks for deep groundwater. Similar elements in the monitoring strategies are the use of vertically screened wells to obtain concentration-depth profiles that can be linked with time-series analysis at specific depths in these profiles.

The examples of regional and phreatic well field monitoring show the opportunities of combining diverging specific information goals into an overall monitoring strategy. A priori analysis of hydrological and hydrogeochemical system properties is a prerequisite for effective

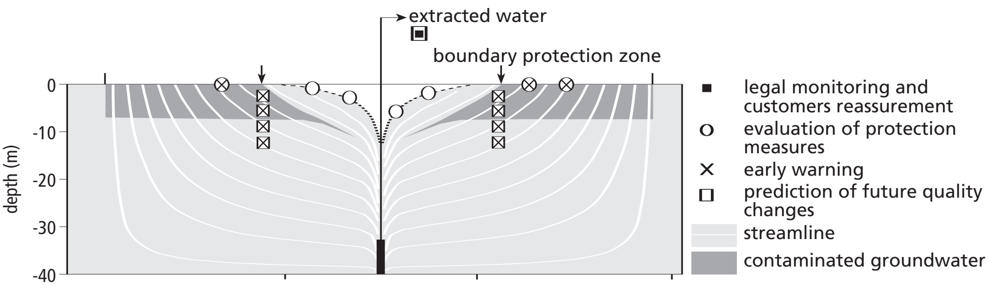


Figure 8.2 - Overview of effective monitoring configurations at a phreatic well field. The position of a 10 year block front is indicated for a scenario with a protection zone (scenario 1b, Chapter 6) at 10 years after the start of the contamination. Details on the evaluation of the configurations are given in Chapter 6.

monitoring and creates the basis for effective set-up of sample size, monitoring depths, frequencies and data analysis protocols. The examples have shown that different monitoring objectives lead to different monitoring systems, even when hydrological and hydrogeochemical system properties are similar.

8.2 Suggestions for further work

Some specific subjects need further attention in regional groundwater quality monitoring:

1. The most fundamental issue in groundwater quality monitoring relates to the temporal and spatial scales of heterogeneity and variability of the factors that determine groundwater quality.
   Spatial heterogeneity of groundwater recharge, the input of solutes, the hydraulic properties, the hydraulic boundary conditions and the reactive properties of the subsurface determine the spatial variability of groundwater quality and the sample size and monitoring set-ups that are required for effective monitoring. Some work has been done on the effects of the first 4 factors on groundwater base flow to a stream (Duffy & Lee 1992, Chapter 2). The results of Duffy & Lee indicate that heterogeneity at a scale smaller than 2 to 10% of the problem scale has limited effects on transit time distributions and solute breakthrough of groundwater base flow. Given the similarity of transit time distribution for drains and pumping wells, these results can be extended to phreatic well fields (Beltman et al. 1995, Chapter 6).
   Little is known, however, about how the spatial scales of heterogeneous aquifer reactivity affect solute breakthrough from diffuse contaminants in pumping wells or surface water outflow. Research is recommended on this subject, emphasizing both the effects on breakthrough in pumping wells and the effects on spatial variations at specific monitoring depths in the aquifer. The second aspect is relevant to specify the required number of observation wells. Such research should also address the interrelationship between hydraulic and reactive variability at several scales.
   Temporal variability of concentrations in shallow groundwater is predominantly caused by time-changing input of solutes and the combined effects of meteorological conditions and unsaturated zone flow (for example van Ommen et al. 1989, Fraters et al. 1998). In evaluating time series data of the regional networks, it was assumed that large temporal variations were tempered using the chosen sample support (2 m screen lengths) and monitoring depths (> 5 m below the water table). However, this tempering effect has not been studied well because of the lack of long-term time series of groundwater quality. The temporal variability at time scales of 0.5 - 10 years needs further attention in the analysis and interpretation of temporal trends in order to distinguish effects of meteorological conditions and changing input loads. The time series of the national wells, which have lengths of about 15 years, are suitable for a reconnaissance survey on this subject.
2. *Age dating* of young groundwater is essential to couple temporal trends in groundwater to the input history of contaminants in recharging groundwater. Promising new age dating techniques have become available during the last decade (Solomon et al. 1992, Ekwurzel et al. 1994, Johnston et al. 1998, Zoelmann et al. 2001). These have not yet been used at the scale of regional monitoring networks or the scale of phreatic well fields. Given the importance of the advective residence times in monitoring design and data analysis, these techniques should be tested to improve knowledge on system properties and the interpretation of monitoring results.
3. The lack of data on *aquifer reactivity* is the largest impediment for prediction of

groundwater quality changes. Information analysis on the acquisition of reactivity data (Chapter 6) indicates that average contents of reactive properties will suffice for the regional scale prediction of the vertical propagation of reaction fronts. A further assessment of subsurface reactivity is recommended to improve predictions on the future developments in groundwater quality and the propagation of pollution fronts. The assessment should preferably be directed to establishing relations between sediment reactivity and sedimentary facies or geological formations in order to delineate the spatial extent of reactive zones and layers. Following the argumentation in Chapter 7, it is recommended to collect sediment samples of standardized volumes that are relevant for the discretisation scale of regional or local scale prediction models.
Random sampling of sediment reactivity enables later quantification of the uncertainty of the estimates. Random sampling of reactivity could proceed simultaneously with the stratified random selection of well locations for regional groundwater quality monitoring. Combining the groundwater quality and reactivity data will improve the interpretation of the monitoring data. Understanding of the effects of aquifer reactivity on the groundwater quality patterns can subsequently be used to adapt monitoring depths and frequencies for groundwater quality (for example, Stuyfzand 1996).

4. Information goals of the regional monitoring networks were based on the risk of contamination of deep groundwater. Another risk approach would be possible which emphasizes the risk for the contamination of *surface waters* by groundwater base flow (information goal X in Figure 8.1). Conceptually, the calculation of solute breakthrough in base flow is similar to breakthrough in wells (Chapter 6). However, integrated monitoring of surface waters and groundwater is complicated because of the different time scales involved. Further research is necessary to accomplish the integration of information on surface water quality and groundwater quality at the scale of regional and local catchments.

Appendix I Hydrochemical methods

I.1 Hydrogeochemical indicators

Extra indicators were introduced in the data analysis to evaluate geochemical processes, such as the oxidation capacity (OXC), the hardness/alkalinity ratio, saturation indices for calcite and siderite and the potassium adsorption ration (PAR). These indicators are described below.

Oxidation Capacity

The oxidation capacity of the groundwater (OXC) is equal to the amount of electrons consumed by redox processes and was defined after Postma et al. (1991), as:

$$OXC = 7\,[SO_4^{2-}] + 5\,[NO_3^-] \quad (mmol\ l^{-1}) \tag{I.1}$$

The OXC is advantageous because it sums two important indicators of agricultural pollution and does not decrease after the process of pyrite oxidation. Following the definition above, the OXC is only useful for fresh groundwater. Adaptations exist for brackish and saline groundwater (Griffioen et al. 1997, Graf-Pannatier 2000).

Hardness/alkalinity ratio

The hardness/alkalinity ratio was used as an alternative indicator of acidification. The ratio was defined as:

$$HH\,/\,alkalinity = \frac{[Ca] + [Mg]}{[HCO_3]} \quad (meq\ l^{-1}) \tag{I.2}$$

where brackets indicate molal concentrations.

Calcite and siderite saturation indices

Calcite and siderite saturation was calculated using WATEQ4F (Ball & Nordstrom 1991). The saturation index *SI* for calcite is:

$$SI_{calcite} = \log\left(\frac{IAP}{K}\right) = \log\left(\frac{[Ca^{2+}] + [CO_3^{2-}]}{K_{calcite}}\right) \tag{I.3}$$

where *IAP* refers to the ion activity product, [ ] denotes molar concentrations, and *K* is the solubility coefficient of the mineral, which is temperature and pressure dependent. For siderite saturation, Fe replaces Ca in the equation. All measured iron (Fe_T) was assumed to be present as Fe(II) for the calculation of siderite saturation.

Potassium Adsorption Ratio

The Potassium Adsorption Ratio (PAR) describes the relative amount of potassium in groundwater. PAR is a useful indicator of potassium contamination, since the ratio describes the cation exchange, which is a major process determining potassium concentrations in groundwater. PAR was defined by Griffioen (2001) as:

$$PAR = \frac{K}{\sqrt{Me^{2+}}} \quad (mmol^{0.5}\ l^{-0.5}) \tag{I.4}$$

where Me^{2+} is the sum of divalent cations, such als Ca^{2+}, Mg^{2+} and Fe^{2+}.

1.2 Pollution indices

The POLIN index

The POLIN index of Stuyfzand (1993) was developed to identify a variety of pollution sources in the Netherlands dune area. The first three quality aspects of his index indicate inorganic pollution sources. The parameters of his index are a measure for A. acidification/ eutrophication, B. excessive application of fertilisers and manure-spreading and C. pollution with environmentally hazardous heavy metals and arsenic. The calculation of POLIN proceeds as:

$$POLIN_A = 1.33 * |pH - 7| \tag{I.5}$$

$$POLIN_B = \frac{\ln\{10 * ([NO_3] + 0.67[SO_4] - 0.0232[Cl])\}}{\ln 2} \quad (mmol\ l^{-1}) \text{ or}$$

$$POLIN_B = \frac{\ln 10 * [NO_3]}{\ln 2} \text{ if } (0.67[SO_4] - 0.0232[Cl]) < 0 \tag{I.6}$$

$$POLIN_C = \frac{1}{\ln 2} * \ln\left(\frac{100}{6} * \frac{As+Cr+Ni}{50} + \frac{Cu+Zn}{100} + \frac{Cd}{5}\right) \quad (ug\ l^{-1}) \text{ or}$$

$$POLIN_C = 0 \quad \text{if} \quad \left(\frac{100}{6} * \frac{As+Cr+Ni}{50} + \frac{Cu+Zn}{100} + \frac{Cd}{5}\right) < 1 \tag{I.7}$$

$$POLIN = \frac{POLIN_A + POLIN_B + POLIN_C}{3} \tag{I.8}$$

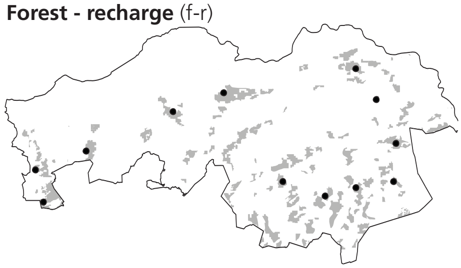
where [ ] indicate molar concentrations. In this thesis, groundwater was considered polluted if POLIN > 1.5. This corresponds to Stuyfzand's classes 'slightly polluted' to 'extremely polluted'.

The MANURE index

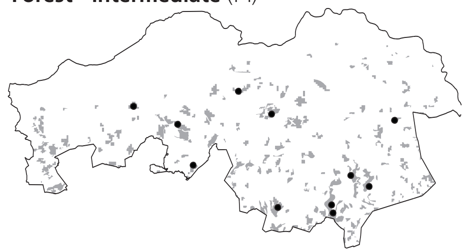
The pollution index MANURE uses PAR, OXC and the chloride concentration to identify groundwater that is influenced by nutrients and salts from animal manure or fertiliser. In this study, groundwater was considered to be agriculturally polluted if one of the following condition is fulfilled: (1) PAR > 0.1 $mmol^{0.5}\ l^{-0.5}$, (2) OXC > 7 mmol l^{-1}, or (3) chloride > 50 mg l^{-1}. These thresholds were chosen using information from local studies in the Netherlands (Griffioen & Hoogendoorn 1993, Frapporti et al. 1995, Broers et al. 1994, Broers & Buijs 1997, Griffioen 2001).

Appendix IIa Homogeneous areas in Noord-Brabant

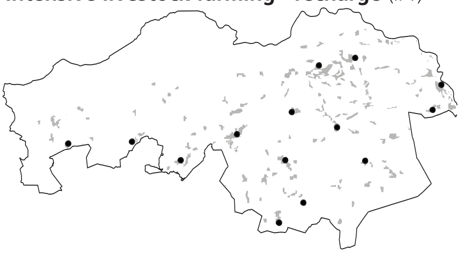
Forest - recharge (f-r)



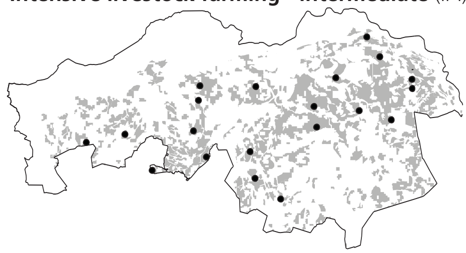
Forest - intermediate (f-i)



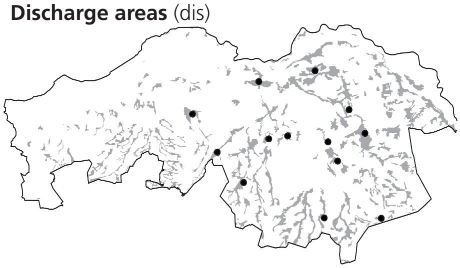
Intensive livestock farming - recharge (il-r)



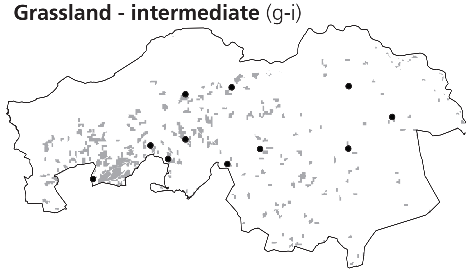
Intensive livestock farming - intermediate (il-i)



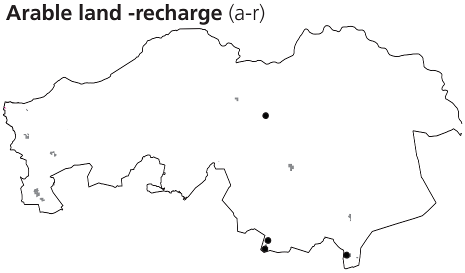
Discharge areas (dis)



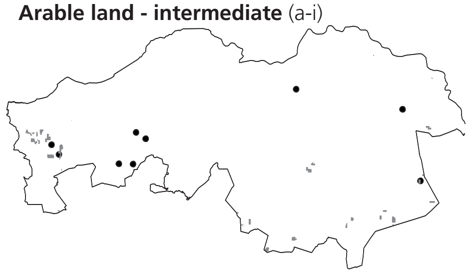
Grassland - intermediate (g-i)



Arable land -recharge (a-r)



Arable land - intermediate (a-i)



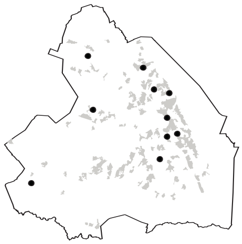
delineation of homogeneous area

• monitoring well

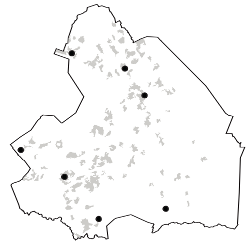
0 25 km

Appendix IIb Homogeneous areas in Drenthe

Arable land - recharge (a-r)

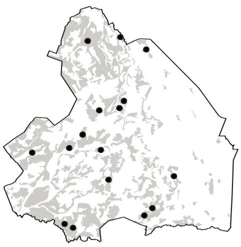


Grassland - recharge (g-r)

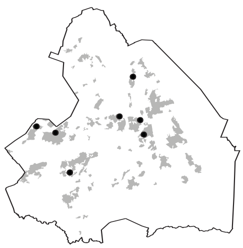


Arable land - intermediate (a-i)

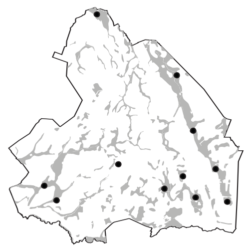
Grassland - intermediate (g-i)



Forest - recharge (f-r)



Discharge areas (dis)



delineation of homogeneous area

• monitoring well

0 25 km

Appendix III Statistical methods

III.1 Non-parametric confidence intervals for the median

The median and the corresponding 95% confidence interval were calculated non-parametrically following Helsel and Hirsch (1992, p. 70). Here, the significance level 5% was chosen as the acceptable risk of not including the median. One-half of assigned to the upper and lower end of the confidence interval. The critical values x' and x were read from a Table of the binomial distribution at for p = 0.5 (median) and $\alpha/2$ (Table B5, Helsel & Hirsch 1992). These critical values were transformed in the lower rank R_l and upper rank R_u following:

$$R_l = x' + 1 \quad \mathit{and} \quad R_u = x \tag{III.1}$$

The concentration data were sorted and the concentrations that correspond to R_l and R_u define the lower and upper limits of the 95% confidence interval. Table III.1 lists the ranks R_l and R_u for different values op sample size n. The 95% confidence interval is non-symmetrical when the frequency distribution of the samples is positively or negatively skewed.

Table III.1 - Values for R_l and R_u for sample size n between 7 and 20

Sample size (n)	**Lower rank** R_l	**Upper rank** R_u
7	1	7
8	2	7
9	2	8
10	2	9
11	3	9
12	3	10
13	3	11
14	4	11
15	4	12
16	4	13
17	5	13
18	5	14
19	5	15
20	6	15

III.2 Significant differences between areas using a multiple-comparison test

The overall effectiveness of the sample size in the networks was evaluated by testing for significant differences between the homogeneous areas. A multiple-comparison test was used to test whether the medians of the homogeneous areas were significantly different at the 95% confidence level.

The multiple-comparison test was only performed if a Kruskall-Wallis one-way ANOVA test rejects the null hypothesis that all medians of the homogeneous areas are identical. This indicates that the variance *within* the homogeneous areas is smaller that the variance *between* the homogeneous areas.

Tukey's method was used to compute the multiple-comparison test because sample sizes in the homogeneous areas are unequal (Helsel & Hirsch 1992, p. 196). Ranks of the concentration data were used instead of the concentrations themselves to obtain a non-parametric test. The procedure is similar to the one used by Nolan et al. (1997) except that ranks were used instead of a log-normal transformation of concentrations. Significant differences between homogeneous areas were detected at a two-sided significant level α=0.05. Significant differences between the medians in the homogeneous areas are only demonstrated if sample sizes in both the compared areas are large enough and limited variation within the areas is observed.

III.3 Probability of detecting contamination

The probability of not detecting a specific percentage of contaminated area within the homogeneous area using the current sample size was computed using the binomial distribution. If the sample size is small, there is a large probability that no contamination is detected although part of the investigated area is contaminated (Criterion C, Figure 4.2, chapter 4). The criterion was based on the conditional probability of encountering no samples exceeding the critical concentration under the condition that a known proportion p_{x_c} of the area is above the critical concentration, using sample size n (Baggelaar & Van Beek 1997). The proportion (or percentage) of contaminated groundwater was defined by comparing the measured concentration with a relevant water quality standard (Appendix III.5). The estimated proportion $\hat{p}_{x_c}$ is defined as (Gilbert 1987):

$$\hat{p}_{x_c} = \frac{u}{n} \tag{III.2}$$

where n is number of observations in the specific homogeneous area and u is the number of observations exceeding the critical concentration x_c.

The probability of detecting u observations above the critical concentration in sample of size n, given the true proportion p_{x_c} in the population, equals (Cochran 1977):

$$P[u|(n,p_{x_c})] = \frac{n!}{u!(n-u)!}\, p_{x_c}^{u}\,(1-p_{x_c})^{n-u} \tag{III.3}$$

Hence, the probability of detecting 0 observations above the concentration in a sample of size n, given the true proportion p_{x_c} equals (Baggelaar & Van Beek 1997)

$$P[0|(n,p_{x_c})] = (1-p_{x_c})^{n} \tag{III.4}$$

The required sample size n to detect at least one observation above the critical concentration x_c at 95% confidence level (α=0.05) is found for:

$$P[0|(n,p_{x_c})] = (1-p_{x_c})^{n} < 0.05 \tag{III.5}$$

Figure III.1 shows the required sample size to detect contamination at α=0.05 level as a function of the true proportion p_{x_c} of contaminated area in the population. For example, 14 wells are required to detect at least one observation above the critical concentration, if 20% of the area (p_{x_c} = 0.2) is actually contaminated at 95% confidence level.

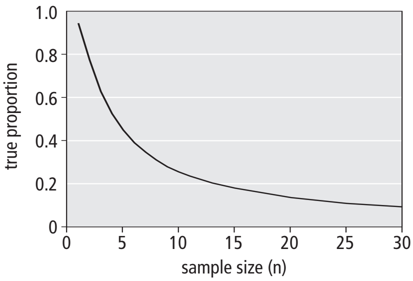


Figure III.1 - Required sample size to detect contamination at α=0.05 level as a function of the true proportion p_{x_c} of contaminated area in the population

III.4 Obtaining LOWESS smooths

A LOWESS smooth (LOcally WEighted Scatter-plot Smoothing, Cleveland & Devlin 1988) is a non-parametric tool for exploratory data analysis, which is often used for water resources data because of its robustness and the insensitivity of outliers in the data. The LOWESS smooth indicates the centre of the data scatter. The calculation of LOWESS is computationally intensive, but LOWESS smooth codes are available in statistical packages such as S-Plus and SPSS. A window width of 75% (span =0.75) was used for all data analysis.

III.5 Non-parametrical confidence intervals on the proportion

The proportion (or percentage) of contaminated groundwater was defined by comparing the measured concentration to a relevant water quality standard. The estimated proportion $\hat{p}_{x_c}$ is defined as (Gilbert 1987):

$$\hat{p}_{x_c} = \frac{u}{n} \tag{III.6}$$

where n is number of observations in the specific homogeneous area and u is the number of observations exceeding the critical concentration x_c.

The corresponding two-sided non-parametrical 95% confidence interval was obtained using the method of Blyth & Still (1983) using Table A4 of Gilbert (1987) for $n < 30$. For $n > 30$ the upper and lower limits of the confidence intervals are computed as follows:

$$\text{Lower limit} = \frac{1}{n + Z_{0.975}^2} \cdot \left\{ (u - 0.5) + \frac{Z_{0.975}^2}{2} - Z_{0.975} \left[(u - 0.5) - \frac{(u - 0.5)^2}{n} + \frac{Z_{0.975}^2}{4} \right]^{1/2} \right\} \tag{III.7}$$

$$\text{Upper limit} = \frac{1}{n + Z_{0.975}^2} \cdot \left\{ (u + 0.5) + \frac{Z_{0.975}^2}{2} + Z_{0.975} \left[(u + 0.5) - \frac{(u + 0.5)^2}{n} + \frac{Z_{0.975}^2}{4} \right]^{1/2} \right\} \tag{III.8}$$

except that the upper limit equals 1 if $u = n$ (Gilbert 1987). Here, $Z_{0.975}$ refers to the standard normal variable that cuts off 2.5% of the upper and lower tail of the standard normal distribution and equals 1.96. Hence, the probability that the population proportion contaminated groundwater is outside interval is less than 5%.

The precision of the estimate depends on the sample size (n) and number of excursions above the critical concentration (u). This is illustrated in Figure III.2 which shows the 95% confidence interval as a function of estimated proportion ($\hat{p}_{x_c}$ = 0, 0.2, 0.5 and 0.8, respectively) and the sample size n. The figure shows the intervals for $n < 30$ and is based on Table A4 of Gilbert (1987). For $\hat{p}_{x_c}$ = 0.5, the confidence interval is symmetric around $\hat{p}_{x_c}$ = 0.5. The curves become very flat above sample size n = 20, which indicates that a precision better than plus or minus 15% is only feasible for a large sample size. The confidence interval is asymmetric for $\hat{p}_{x_c} \neq 0.5$. For $\hat{p}_{x_c}$ = 0, 0.2 and 0.8, the boundary farthest from the estimated percentage $\hat{p}_{x_c}$ decreases faster with sample size n than for $\hat{p}_{x_c}$ = 0.5 (Figure III.2).

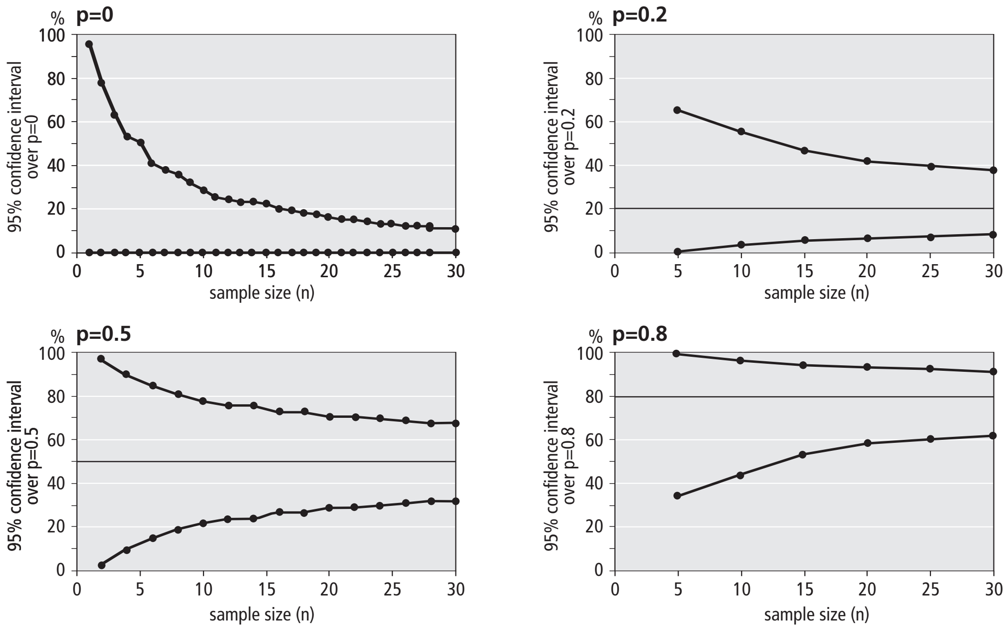


Figure III.2 - Precision of the estimated proportion of contaminated groundwater as a function of sample size (n)

III.6 Obtaining proportions for combined strata

Proportions of contaminated groundwater in individual homogeneous areas were combined to determine proportions in larger areas that comprise several homogeneous areas. The proportion of contaminated groundwater in those combined areas must be corrected for the non-proportional allocation of wells over the homogeneous areas. The homogeneous areas were considered to be individual strata, with weights W_i that reflect their spatial extent. For h homogeneous areas the weights sum to 1:

$$\sum_{i=1}^{h} W_i = 1 \qquad \text{(III.9)}$$

The proportion of the combined area was then calculated following Cochran (1977):

$$\hat{p}_{x_c(st)} = \sum_{i=1}^{h} W_i . \hat{p}_i \qquad \text{(III.10)}$$

where W_i = weight of homogeneous area i, $\hat{p}_i$ = estimated proportion in homogeneous area i and $\hat{p}_{x_c}$ is the estimated proportion in the combined strata.

The confidence interval on the estimated proportion $\hat{p}_{x_{c(st)}}$ must be corrected to account for the non-proportional allocation of observations in the individual strata. No general equations are available to calculate these confidence intervals, because non-parametrical methods were used to obtain the confidence intervals, and because the intervals are non-symmetric around the estimated proportion. Cochran (1977, p. 109) argued that only a slight gain in precision is to be expected from stratified sampling, even if the proportions of the individual strata differ largely. Therefore, the confidence intervals of the proportions in the combined strata were calculated by correcting the number of observations in the individual strata to establish a proportional allocation. This is explained below.

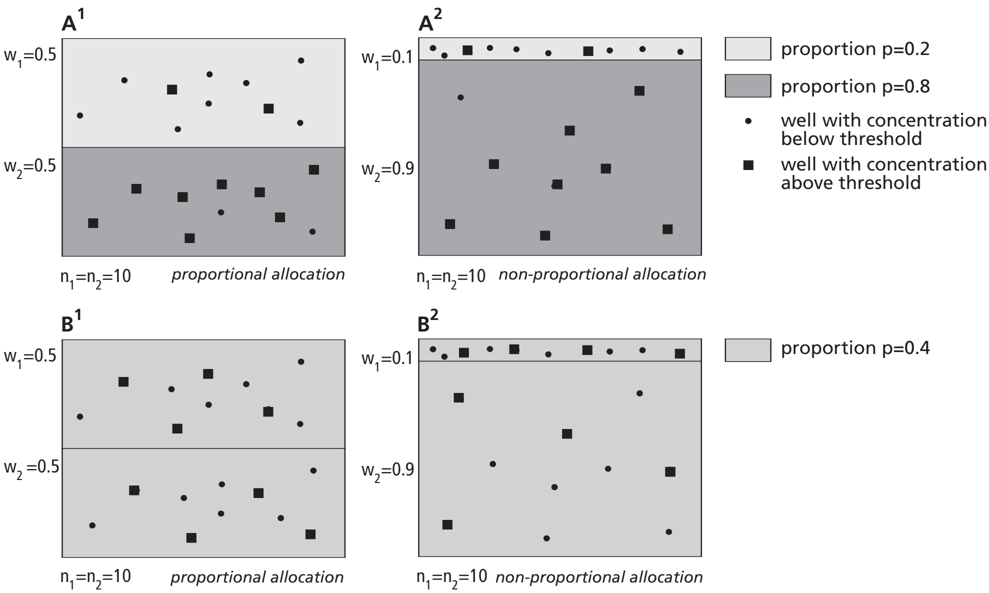


Figure III.3 - Example of the correction of the confidence interval on the estimated proportion of the combined strata (1) en (2). Both strata contain 10 samples.
A1 and A2: proportions in strata 1 and 2 differ: p_1=0.2 and p_2=0.8.
B1 and B2. proportions in both strata are 0.4
Proportional allocation is achieved for w_1=w_2=0.5 (A1 and B1).
Non-proportional allocation (w_1=0.1, w_2=0.9) shows superfluous sampling in stratum 1 (A2 and B2).

Calculation of the confidence intervals proceeds simply if a proportional allocation exists in both strata (examples A1 and B1 of Figure III.3). For example A1, the proportion of contaminated groundwater is calculated following equation (III.10): $\hat{p}_{x_{c(st)}}$ = 0.5 * 0.2 + 0.5 * 0.8 = 0.5. The number of wells exceeding the critical concentration = 2 + 8 = 10. The confidence interval is then found from Table A4 of Gilbert (1987) for u = 10 and n = 20 and lower and upper limits are 0.29 and 0.71, respectively.

The sampling is much less efficient if the samples are not allocated proportionally to the spatial extent of the individual strata. In this case, not all the observations from the superfluously sampled stratum are included in the calculation of the confidence interval. For example A2 of Figure III.3, the proportion $\hat{p}_{x_{c(st)}}$ equals 0.1 * 0.2 + 0.9 * 0.8 = 0.74. To calculate the corresponding confidence interval, a new number of observations $n_{(st)}$ and a new number of samples $u_{(st)}$ exceeding the critical concentration x_c is calculated as follows:

$$n_{(st)} = n_1 \left(1 + \frac{w_2}{w_1} + \frac{w_3}{w_1} + \ldots\right) \tag{III.11}$$

$$u_{(st)} = n_{(st)} \cdot \hat{p}_{x_{c(st)}} \tag{III.12}$$

where n_1 is the number of observations in the strata with the largest weight, and w_1, w_2 and w_3 are weigths of homogeneous areas with decreasing order of spatial extent. For example A2, $n_{(st)}$ = 10(1+0.1/0.9) = 11, and $u_{(st)}$ = 11 * 0.74 = 8. The newly calculated $n_{(st)}$ and $u_{(st)}$ are then used to calculate the confidence interval following the regular procedure of Gilbert (1987). The asymmetric confidence interval is found using Table A4 for u=8 and n=11 and lies between 0.40 and 0.92. Thus, for this case of strongly non-proportional allocation, only one of the observations of stratum 1 is used for the calculation of the combined confidence interval.

Figure III.4 shows how the confidence interval changes with the weights of the contributing strata. For two strata with p_1=0.2 and p_2=0.8 (example A1 and A2) the stratified proportion decreases with increasing weight w_1. The confidence intervals are asymmetric for w_1>0.5 and

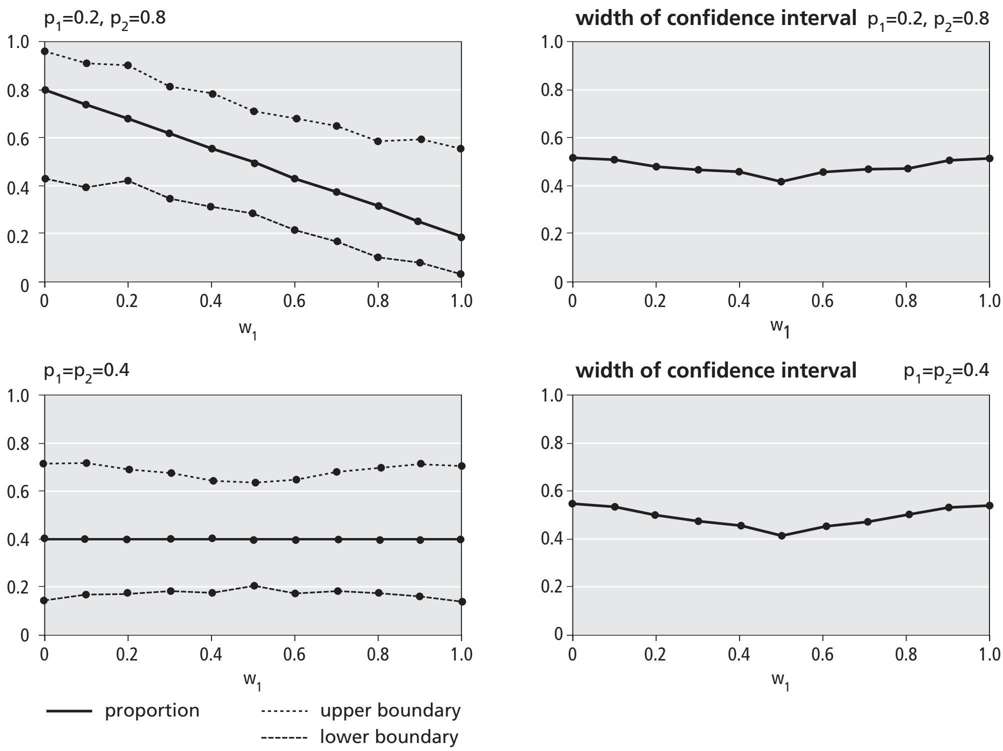


Figure III.4 - Confidence intervals over the estimated proportion as a function of strata weight w_1 for two strata with p_1=0.2 and p_2=0.8 (upper part, example A) and for two strata with equal proportion p_1=p_2=0.4 (lower part, example B). The width of the confidence intervals decreases toward proportional allocation (w_1=w_2=0.5).

w_1<0.5, which reflects the asymmetry in the confidence intervals of the original strata that is shown in Figure III.2 (right figures; sample size = 10). The smallest width of the confidence interval of the combined strata is found for the case of proportional allocation (w_1=w_2=0.5, Figure III.4).

The width of the confidence interval shows a larger decrease towards w_1=w_2=0.5 for two strata with equal proportions of contaminated groundwater (example B1 and B2, p_1=p_2=0.4).

Concluding, the width of the confidence interval of the estimated proportion in the combined strata is a function of the proportion $\hat{p}_{x_{c(st)}}$, which in turn depends on the proportions in the original strata, the spatial extent of the strata and the number of observations in each of them.

III.7 Trend analysis of time series

Mann-Kendall trend test

Significant trends in time series were identified using the non-parametric Mann-Kendall trend test. This test is robust and rather insensitive to outliers in the time series, compared with ordinary linear regression. The test is recommended for use in large data sets where the normality assumption cannot be checked for all individual time series (Helsel & Hirsch 1992). The Mann-Kendall procedure tests for monotonic trends in the data.

The Mann-Kendall test statistic S is calculated from all $n(n\text{-}1)/2$ data pairs in a time series:

$$S = \sum_{i=1}^{n-1} \sum_{j=i-1}^{n} Sgn(X_j - X_i) \qquad \text{(III.13)}$$

where X_j and X_i are sequential data values, n is the number of data in the time series, and

$$Sgn(\theta) = \begin{cases} +1 & \theta > 0 \\ 0 & \theta = 0 \\ -1 & \theta < 0 \end{cases} \qquad \text{(III.14)}$$

The null hypothesis of no change is rejected when S is significantly different from zero. For n > 10 the test statistic S is approximately normally distributed with mean is zero and variance:

$$\sigma_s = \sqrt{\frac{n}{18}(n-1)(2n+5)} \qquad \text{(III.15)}$$

For n > 10 a large sample approximation is used to test this hypothesis:

$$Z_s = \begin{cases} \dfrac{S-1}{\sigma_s} & S < 0 \\ 0 & S = 0 \\ \dfrac{S+1}{\sigma_s} & S > 0 \end{cases} \qquad \text{(4)}$$

The null hypothesis is rejected at significance level α=0.05 if $Z_s > Z_{crit}$ where Z_{crit} is the value of the standard normal distribution with a probability of exceedence of $\alpha/2$. Z_{crit} equals 1.96 for α=0.05.

Kendall-Theil robust line
The trend slope and 95% confidence intervals were computed using the Kendall-Theil robust line method (Helsel & Hirsch 1992). The slope is computed as the median of all $n(n\text{-}1)/2$ slopes between the (X,Y) data pairs:

$$\beta = \frac{(Y_j - Y_i)}{(X_j - X_i)} \quad \text{for all } i < j \text{ and } i = 1,2,\ldots(n\text{-}1) \;\; j = 2,3,\ldots,n \tag{III.17}$$

Confidence intervals for the Theil slope are obtained by computing the upper and lower ranks of all the ranked slopes between the $N = n(n\text{-}1)/2$ data pairs:

$$R_u = \frac{N + Z_{crit}\sqrt{n/_{18}(n - 1)(2n + 5)}}{2} + 1 \tag{III.18}$$

$$R_l = \frac{N - Z_{crit}\sqrt{n/_{18}(n - 1)(2n + 5)}}{2} + 1 \tag{III.19}$$

The median slope and the 95% confidence interval slopes join at the point (median X, median Y).

Kendall tau correlation coefficient
The Kendall τ correlation coefficient is a robust rank-based measure for correlation in the time series and resistant to the effect of a small number of unusual values. The kendall τ correlation coefficient equals (Helsel & Hirsch 1992):

$$\tau = \frac{S}{n(n - 1)/2} \tag{III.20}$$

Appendix IV Bootstrapping the Oostrum aquifer

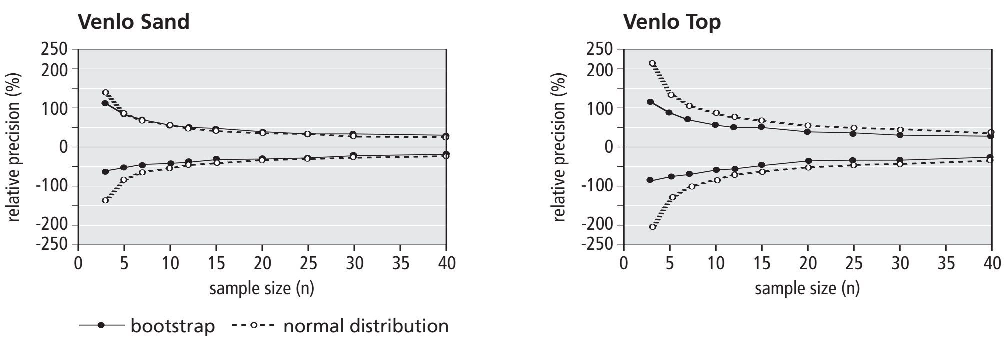
Pyrite data from the Oostrum aquifer were used to test relations between the sample size and the relative precision of estimated average contents and estimated percentiles. The strata Venlo Sand and Venlo Top were described in section 7.6 of Chapter 7. Confidence intervals for a specific sample size were evaluated non-parametrically using bootstrapping. This implies the drawing with replacement of 1000 random samples of sub-sample size n out of the available data of the two strata (n_1=39 and n_2= 18, respectively). The assumption is that the available samples give a good indication of the frequency distribution of the entire population. In the presented results, sub-sample sizes n of 3, 5, 7,10, 12, 15, 20, 25, 30 and 40 were used.

IV.1 Confidence intervals on average contents

For each random sample of sub-sample size n, the average content ($\bar{x}$) was determined from the obtained 1000 averages. The non-parametrical 95% confidence interval was obtained using the 2.5 and 97.5 percentiles of the 1000 averages. The obtained confidence intervals were compared with the parametric 95% confidence intervals that were calculated using the assumption of normally distributed data (equation (7.1), section 7.5).

The parametric and non-parametric 95% confidence intervals are shown in Figure App IV.1 as a function of sub-sample size n. The non-parametric confidence intervals show the effects of the skewness of the Venlo Sand and Venlo Top data sets. The parametric confidence intervals are symmetrical. The asymmetry of the non-parametric confidence intervals decreases with sub-sample size n, because the distribution of averages becomes more symmetrical with increasing n according to the Central Limit Theorem.

For n>7, the non-parametric and parametric confidence intervals were similar for the homogeneous Venlo Sand data sets ($s/\bar{x}$ = 0.75). For the skewed, largely variable data of Venlo Top ($s/\bar{x}$ = 1.17) similar results were obtained for n > 15. In general, the parametric confidence intervals gave a good indication of the required sample size to attain a specific precision above n = 12. A relative precision smaller than 50% was obtained for n > 12 at Venlo Sand and n > 22 for Venlo Top. The results of the Oostrum aquifer indicate that the parametric approximation overestimates the required sample size for a specific desired precision. Concluding, the use of the parametric equation for the relative precision of the average content is on the safe side.

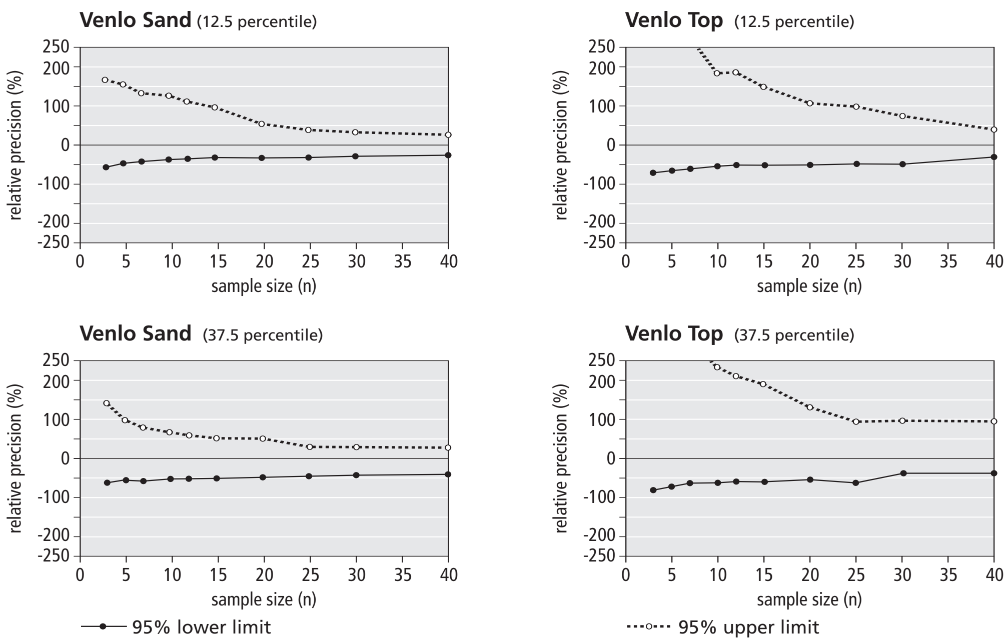


Appendix IV.1 - Relative precision of the estimated average content as a function of sample size for A. Venlo Sand and B. Venlo Top. Open symbols represent the normal distribution, closed symbols were obtained by bootrapping.

IV.2 Confidence intervals of percentiles

The similar procedure was carried out to evaluate the precision of the estimated percentiles. Figure App. IV.2 shows the non-parametric 95% confidence intervals for the 12.5 and 37.5 percentiles of Venlo Sand and Venlo Top. The confidence intervals were strongly asymmetric and the relative precision of the upper boundary is less than the relative precision of the lower boundary. For a distinct sub-sample size, the relative precision of the estimated percentiles is much less than the relative precision of the estimated average content (Figure App. IV.1). Accordingly, larger sample sizes are required for precise estimation of percentiles. Even in the relatively homogeneous data set of Venlo Sand ($s/\bar{x}$ = 0.75), a desired relative precision of 50% requires 15-25 samples.

The results of the bootstrapping should only be considered as an indication, because of the limited number of samples available in the two strata (n_1 = 39 and n_2 = 18, respectively) and the limited number of 1000 random drawings. However, the results can be used to obtain an indication of sample size for the estimation of percentiles for strata of different variability. The experiments showed that the lower boundary of the confidence interval can be determined within 50% using a reasonable sample size, but the upper boundary was not easily estimated within 200%, because of the positively skewed data. These thresholds were considered acceptable for the use of the sample statistics in a sensitivity analysis with the transport model. Using these thresholds, the Table IV.1 was extrapolated to determine an initial sample size for the estimation of percentiles.



Appendix IV.2 - Relative precision of the estimated 12.5 and 37.5 percentiles as a function of sample size for A. Venlo Sand and B. Venlo Top obtained by bootstrapping.

Table Appendix IV.1 - Suggestions for the sample size for determining percentiles within -50% and +200% relative precision. The coefficient of variation is estimated from the reconnaissance stage data

Coefficient of variation (s/x)	**Sample size**
0.4-0.6	10-12
0.6-0.8	12-15
0.8-1.0	15-20
1.0-1.2	20-25
1.2-1.5	25-30
> 1.5	30

References

ADKINS, N.C., H.F. BELL & R.C. WARD (1995). The role of data analysis protocols in obtaining groundwater quality management information. In: Models for Assessing and Monitoring Groundwater Quality (Proceedings of a Boulder Symposium, July 1995). IAHS Publication 227, pp. 99-106.

ALLEN-KING, R.M., R.M. HALKET, D.R. GAYLORD & M.J.L. ROBIN (1998). Characterizing the heterogeneity and correlation of perchloroethene sorption and hydraulic conductivity using a facies-based approach. Water Resources Research 34, pp. 385-396.

ALLER, L., T. BENNETT, J.H. LEHR, R.J. PETTY & G.HACKETT (1987). DRASTIC: A standardized system for evaluating groundwater pollution potential using hydrogeologic settings. Report U.S. Environmental Protection Agency: Washington, DC: EPA/600/2-87/035.

ALLEY, W.M. ed. (1993). Regional Ground-Water Quality. Van Nostrand Reinhold, New York.

APPELO, C.A.J., G.J.W. KRAJENBRINK, C.C.D.F. VAN REE & L. VASAK (1982). Controls on groundwater quality in the NW Veluwe catchment. Soil Protection Series 11, Staatsuitgeverij, Den Haag. (in Dutch)

APPELO, C. A. J & D. POSTMA (1993). Geochemistry, Groundwater and Pollution. Balkema, Rotterdam, The Netherlands.

BAGGELAAR, P.K. 1992). Towards a monitoring system groundwater quality at drinking water well fields. Kiwa-Report 117, Nieuwegein, The Netherlands.

BAGGELAAR, P.K. (1996). An overall strategy for groundwater quality monitoring by drinking water supply companies. Proceedings Monitoring Tailor Made II: Information Strategies in Water Management, pp. 75-85. Published by RIZA, Lelystad, The Netherlands.

BAGGELAAR, P. K. & C.G.E.M. VAN BEEK (1997). Suggestions for the optimization of large scale groundwater quality monitoring networks. Report KOA 95.107, Kiwa, Nieuwegein. (in Dutch)

BALL, J.W. & D.K. NORDSTROM (1991). User's manual for WATEQ4F. U.S. Geol. Survey, Open file Report 91-183.

BARBER, L.B., E.M. THURMAN & D.D. RUNNELS (1992). Geochemical heterogeneity in a sand and gravel aquifer: effects of sediment mineralogy and particle size on the sorption of chlorobenzenes. Journal of Contaminant Hydrology 9, pp. 35-54.

BARLOW, P.M. (1994). Two-and three-dimensional pathline analysis of contributing areas to public-supply wells of Cape Cod, Massachusetts. Groundwater 32 (3), pp. 399-410.

BEAR, J. (1972). Dynamics of fluids in porous media. Elsevier, New York.

BEAR, J. & J. VERRRUIJT (1987). Modeling groundwater flow and pollution. Reidel, Dordrecht.

BEEKMAN, H.E. (1991). Ion chromatography of fresh and salt water intrusion. Ph.D. Thesis, Free University, Amsterdam.

BEEKMAN, W. (1998). Manual SPREAD: prediction of nitrate, hardness, chloride and sulphate in shallow groundwater. Kiwa report nr. SWE 98.012, Nieuwegein, The Netherlands. (in Dutch)

BELTMAN, W.H.J., J.J.T.I. BOESTEN & S.E.A.T.M. VAN DER ZEE (1995). Analytical modelling of pesticide transport from the soil surface to a drinking water well. Journal of Hydrology 169, pp. 209-228.

BELTMAN, W.H.J., J.J.T.I. BOESTEN, S.E.A.T.M. VAN DER ZEE. & J.J. QUIST (1996). Analytical modelling of effects of application frequency on pesticide concentrations in wells. Groundwater 34 (3), pp. 470-479.

BEUGELINK, G.P. & J.H.C. MÜHLSCHEGEL (1989). Prognoses of the evolution of the groundwater quality at phreatic well abstraction in The Netherlands. National institute of public health and environmental protection (RIVM), Report no728803001, Bilthoven, The Netherlands. (in Dutch)

BIERKENS, M.F.P. (1994) Complex confining layers: a stochastic analysis of hydraulic properties at various scales. KNAG/Ruimtelijke Wetenschappen Universiteit Utrecht. Netherlands Geographical Studies 184, 272 pp.

BLEEKER, A. & J.W. ERISMAN (1996). Deposition of acidifying components in The Netherlands during 1980-1995. Report National institute of public health and environment, RIVM Report 722108018, Bilthoven, The Netherlands. (in Dutch)

BLYTH, C.R. & H.A. STILL (1983). Binomial confidence intervals. American Statistical Association Journal 78, pp. 108-116.

BÖTTCHER, J., O. STREBEL & H.M. DUYNISVELD (1985). Vertikale Stoffkonzentrationsprofile im Grundwasser eines Lockersteins-Aquifers und deren Interpretation. Zeitschrift Deutscher Geologischer Gesellschaft 136, pp. 543-552.

BOUMANS, L.J.M. & D. FRATERS (1995). The quality of the upper groundwater. In: H.F.M. Aarts (ed): Weide- en voederbouw op de marke; op zoek naar de balans tussen productie en emissie. (in Dutch)
BROERS, H.P (1988). Leaching experiments with clay samples of 4 geological Formations in western Noord-Brabant. Kiwa, Report SWI 88.114, Nieuwegein, The Netherlands. (in Dutch)
BROERS, H.P.(1990). Design of a monitoring network for groundwater quality for the province of Noord-Brabant. Netherlands Institute of Applied Geoscience TNO, Report OS 90-15-A. (in Dutch)
BROERS, H.P. (1993). Interpretation of tritium measurements in groundwater quality monitoring wells. Case study: the provincial monitoring network of Noord-Brabant. Netherlands Institute of Applied Geoscience TNO, Report OS 93-93-A. (in Dutch)
BROERS, H.P. (1996a). Evaluation of the groundwater quality monitoring network of Drenthe. Netherlands Institute of Applied Geoscience TNO, Report GG R-96-78A. (in Dutch)
BROERS, H.P. (1996b). The groundwater quality of Drenthe. Report on the monitoring year 1995 of the provincial network. Netherlands Institute of Applied Geoscience TNO, Report GG R-96-78-A. (in Dutch)
BROERS, H.P. (1998). Release of Ni, Co and As from aquifer sediments at the Oostrum well field. Proceedings of the conference on Groundwater Quality, Remediation and Protection, Tuebingen, Germany. Ed. M. Herbert & K.Kovar. Tuebinger Geowissenschaftiche Arbeiten, C36, pp.106-108.
BROERS, H.P. (1999). A strategy for sampling reactive aquifer sediments at drinking water well fields. Netherlands Institute of Applied Geoscience TNO, Report no. 99-291-B, Delft, The Netherlands. (in Dutch)
BROERS H.P., S.W.M. PETERS & A. BIESHEUVEL (1990). Design of a groundwater quality monitoring network with GIS and remote sensing. Proceedings of the first European Conference on Geographic Information Systems, EGIS'90, pp. 95-106.
BROERS, H.P., A.LOURENS, & F.C. VAN GEER (1991). Optimization of groundwater head monitoring networks at the drinking water well fields in Friesland. Netherlands Institute of Applied Geoscience TNO, Report OS 91-22-A. (in Dutch)
BROERS, H.P. & G. BERGSMA (1992). Installation of the groundwater quality monitoring network of Noord-Brabant: monitoring locations and well completion. Netherlands Institute of Applied Geoscience TNO, Report. OS 92-04-A. (in Dutch)
BROERS, H.P. & J. GRIFFIOEN (1992). The groundwater quality network of the province of Noord-Brabant: design and first results. H2O 25, pp. 728-735. (in Dutch)
BROERS, H.P. & J. GRIFFIOEN (1994). Reactivity of natural sediments in relation to groundwater quality in several hydrogeochemical environments. In: International Symposium on Transport & Reactive Processes in Aquifers, ed. Dracos, Th. & Stauffer, F. Zurich, Switzerland, April 11-15, 1994. Balkema, Rotterdam, The Netherlands, pp.269-275.
BROERS, H.P., J. GRIFFIOEN, & E.A. BUIJS (1994). Relations between groundwater quality and sediment composition for 5 well locations of the provincial monitoring network of Noord-Brabant. Netherlands Institute of Applied Geoscience TNO, Report. OS 94-38-A. (in Dutch)
BROERS H.P. & E.A. BUIJS (1996) Temporal trends in the groundwater quality of Drenthe. Netherlands Institute of Applied Geoscience TNO, report GG R-96-80-A. (in Dutch)
BROERS, H.P. & E.A. BUIJS (1997). Origin of trace metals and arsenic at the Oostrum well field. Netherlands. Institute of Applied Geoscience TNO, Report NITG 97-189-A. (in Dutch)
BROERS, H.P. & J. PEETERS (2000). Evaluation of provincial groundwater monitoring networks; methodology for groundwater quality monitoring networks. Netherlands Institute of Applied Geoscience TNO, Report NITG 00-247-B, 65 pp. (in Dutch)
BRONSWIJK, J.J.B. & H. F. PRINS (2001). Nitrogen inputs and nitrate concentrations in the deeper groundwater of The Netherlands. H2O 25/26, pp. 27-29.BROWN, C.J., J.D. COATES & M.A.A. SCHOONEN (1999). Local sulfate-reducing zones in a coastal plain aquifer Groundwater 37 (4), pp. 505-516.
BRUN, A. et al. (1998). Water quality modelling at the Langerak deep-well recharge site. Proceedings of the third international symposium on artificial recharge of groundwater TISAR 98. Balkema, Rotterdam, pp. 305-310.
BRUS, D.J. & J.J. DE GRUIJTER (1993). Design-based versus model-based estimates of spatial means: theory and application in environmental soil science. Environmetrics (4), no. 2, pp. 123-152.
BURN, D.H. & M.A. HAG ELNUR (2002). Detection of hydrologic trends and variability. Journal of Hydrology 255, pp. 107-122
CAIN, D., D.R. HELSEL & S.E. RAGONE (1989). Preliminary evaluation of regional groundwater quality in relation to land use. Groundwater 27 (2), pp. 230-244.

CBS (1998). Transport and use of animal manure and minerals. (Transport en gebruik van mest en mineralen 1994-1998). Report Netherlands Central Bureau of Statistics (CBS). (in Dutch)
CBS (1999). Production of animal manure and available amount of minerals. Report Netherlands Central Bureau of Statistics : Milieustatistieken in de Landbouw 1970-1990.
CHILTON, P.J. & C.J. MILNE (1994). Groundwater quality assessment: a national strategy for the NRA. British Geological Survey, Technical Report WD/94/40C.
CHILTON, P.J. & S.S.D. FOSTER (1996). Monitoring for groundwater quality assessment: current constraints and future strategies. Proceedings Monitoring Tailor-Made II: Information strategies in water management. September 1996, Nunspeet, The Netherlands, pp. 53-64.
CHRISTIANSEN, J.S., P. ENGESGAARD & P.L. BJERG (1998). A physically and chemically heterogeneous aquifer: field study and reactive transport modelling. Proceedings Groundwater Quality, Remediation and Protection GQ'98. IAHS Publication no. 250, pp. 329-336.
CLEVELAND , W.S.& S.J. DEVLIN (1988). Locally weigthed regression; an approach to regression analysis by local fitting. Journal American Statistical Association 83, pp. 596-610.
COCHRAN, W.G. (1977). Sampling Techniques. John Wiley & Sons, New York.
COPLEN, T.B. (1993). Uses of environmental isotopes. In: Alley, W.M. ed. Regional groundwater quality. New York, van Nostrand Reinhold, pp. 227-294.
DAHMKE, A., G. MATTHESS, A. PEKDEGER, D. SCHENK & H.D. SCHULZ (1986). Near-surface geochemical processes in Quaternary sediments. Journal Geological Society 143, pp. 667-672.
DAVIS, J.A., C.C. FULLER, C.A. COSTON, K.M. HESS & E. DIXON (1993). Spatial heterogeneity of geochemical and hydrologic parameters affecting metal transport in groundwater. US-EPA Environmental Research Brief EPA/600/s-93/006.
DILLON, P.J. (1989). An analytical model of contaminant transport from diffuse sources in saturated porous media. Water Resources Research 25, pp. 1028-1218.
DIXON, W. & B. CHISWELL (1996). Review of aquatic monitoring program design. Water Resources 13 (9), pp. 1935-1948.
DRIA, M.A., S.L. BRYANT, R.S. SCHECHTER & L.W. LAKE (1987). Interacting precipitation/ dissolution waves: the movement of inorganic contaminants in groundwater. Water Resources Research 23, pp. 2076-2090.
DUFFY. C.J & D.H. LEE (1992). Base flow response from non-point source contamination; simulated spatial variability in source, structure, and initial condition. Water Resources Research 28, pp. 905-914.
DZOMBAK, D.A. & F.F.M. MOREL (1990). Surface complexation modeling: hydrous ferric oxide. Wiley & Sons, New York.
EKWURZEL, B., P. SCHLOSSER, W.M. SMETHIE, L.N. PLUMMER, E. BUSENBERG, R.L. MICHEL, R. WEPPERNIG & M. STUTE (1994). Dating of shallow groundwater: comparison of the transient tracers 3H/3He, chlorofluorcarbons and 85Kr. Water Resources Research 30 (6), pp. 1693-1708.
ELDOR, M. & G. DAGAN (1972). Solutions of hydrodynamic dispersion in porous media. Water Resources Research 8, pp. 1316-1331.
ENGELEN, G.B. (1981). A systems approach to groundwater quality - methodological aspects. The Science of the Total Environment 21, pp. 1-15.
ENGELEN, G.B. (1986). Water systems in the South-eastern part of The Netherlands. In Engelen, G.B. & Jones, G.P. Developments in the analysis of groundwater flow systems. IAHS Publication 163, pp. 205-224.
ENGESGAARD, P., K.H. JENSEN, J. MOLSON, E. FRIND & H. OLSEN (1996). Large-scale dispersion in a sandy aquifer: simulation of subsurface transport of environmental tritium. Water Resources Research 32 (11), pp. 3253-3266.
ERISMAN, J.W. & R. BOBBINK (1997). Scientific background of the ammonia problem. Landschap 14 (2), pp. 87-104.
ERNST, L.F. (1973). Determination of travel times of groundwater by saturated groundwater flow. ICW Nota 755, Wageningen, 42 pp. (in Dutch)
EU (2000). Directive 2000/60/EU of the European Parliament and of the Council of 23 October 2000 Establishing a framework for Community Action in the field of water policy. Official Journal of the European Communities L327.
FOGG, G.E. AND R.K. SENGER (1985). Automatic generation of flow nets with conventional groundwater modeling algorithms. Groundwater 23 (3), pp. 336-344.
FOPPEN, J.W. & T. KREMERS (1997). Well field Noordbargeres. Design of the groundwater quality monitoring network. Netherlands Institute of Applied Geoscience TNO, Report 97-125-B. (in Dutch)

FOSTER, S.S.D. (1987). Fundamental concepts in aquifer vulnerability, pollution risk and protection strategy. In: Vulnerability of soil and groundwater to pollutants. Proceedings and Information/TNO Committee on Hydrologic Research; No. 38, pp. 69-85.
FRAPPORTI, G., S.P. VRIEND & P.F.M. VAN GAANS (1993). Hydrochemistry of the shallow Dutch ground water; interpretation of the National Ground Water Monitoring Network. Water Resources Research 29, pp. 2992-3004.
FRAPPORTI, G. (1994). Geochemical and statistical interpretation of the Dutch national groundwater quality monitoring network. Ph.D thesis, Geologica Ultraiectina no. 115, University of Utrecht, The Netherlands.
FRAPPORTI, G., J.H. HOOGENDOORN & S.P. VRIEND (1995). Detailed hydrochemical studies as a useful extension of national ground water monitoring networks. Ground Water 33 (5), pp. 817-828.
FRATERS, D., L.J.M. BOUMANS, G. VAN DRECHT, T. DE HAAN & W.D. DE HOOP (1998). Nitrogen monitoring in groundwater in the sandy regions of The Netherlands. Environmental Pollution 102, pp. 479-485.
FRITZ, P. & J.C. FONTES (1980). Handbook of environmental isotope chemistry. Vol. 1 The terrestrial environment. Elsevier, Amsterdam.
FULLER, C.C., J.A. DAVIS, J.A. COSTON & E. DIXON (1996). Characterization of metal adsorption variability in a sand and gravel aquifer, Cape Cod, Massachusetts, U.S.A. Journal of Contaminant Hydrology 22, pp. 165-187.
GEHRELS, J.C. (1999). Groundwater level fluctuations. Separation of natural from anthropogenic influences and determination of groundwater recharge in the Veluwe area, The Netherlands. Ph. D. Thesis, Vrije Universiteit, Amsterdam.
GEOCHEM (1992). Analysis of a selection of sediment samples from the drilling programme of the regional groundwater quality monitoring network of Noord-Brabant. Report GEOCHEM Utrecht. (in Dutch).
GILBERT, R.O. (1987). Statistical methods for environmental pollution monitoring Van Nostrand Reinhold, New York.
GLYNN, P. & J. BROWN (1996). Reactive transport modeling of acidic metal-contaminated groundwater at a site with sparse spatial information. In: 'Reactive transport in porous media', eds. P.C. Lichtner, C.I. Steeferl & E.H. Oelkers. Reviews in Mineralogy 34. Mineralogical Society of America, Washington D.C., pp.377-438.
GRAF-PANNATIER, E., H.P. BROERS, P. VENEMA & G. VAN BEUSEKOM, G. (2000). A new process-based hydro-geochemical classification of groundwater. Application to The Netherlands national monitoring system. Netherlands Institute of Applied Geoscience TNO, Report NITG 00-143-B.
GRIFFIOEN, J. (1992). Cation-exchange and carbonate chemistry in aquifers following groundwater flow. Ph.D. Thesis, Free University Amsterdam.
GRIFFIOEN, J. (2001). Potassium Adsorption ratios as indicator for the fate of agricultural potassium in groundwater. Journal of Hydrology. 254, pp. 244-254.
GRIFFIOEN, J. & H.P. BROERS (1993). Characterization of sediment reactivity: the feasibility of sequential extraction techniques. Netherlands Institute of Applied Geoscience TNO, Report OS-93-65-A, Delft, The Netherlands.
GRIFFIOEN, J. & J.H. HOOGENDOORN (1993). Infiltration depths of diffuse agricultural pollutants in the groundwater of Salland: controls by hydrogeochemistry and hydrogeology. H2O 27, pp. 138-144. (in Dutch)
GRIFFIOEN, J., E.A. BUIJS & H.P. BROERS (1994). Occurrence and behaviour of some trace metals in shallow groundwater-sediment systems. A case study on 5 wells in the province of Noord-Brabant. Netherlands Institute of Applied Geoscience TNO, Report no. OS 94-39-A. Delft, The Netherlands.
GRIFFIOEN, J., Y. HOUTHUESEN & B. VINK (1997). A geochemical, process-based interpretation of the Dutch national monitoring network. H2O 19, pp. 579-582. (in Dutch)
GRIFFIOEN, J. et al. (1998). An integrated groundwater quality model based on hydrogeochemical environments. In: Proceedings Water Rock Interaction 9 (Taupo, New Zealand). (ed. by G. B. Arehart & J. R. Hulston), Balkema, Rotterdam, The Netherlands, pp.153–156.
GRIFFIOEN, J. & H.P. BROERS (1999). Manure in groundwater: the fate of nutrients in the subsurface. Landschap 16 (3), pp. 169-178. (in Dutch with English summary)
GRIFFIOEN, J. & T. KEIJZER (2001). Biogeochemical cycles of chloride, nitrogen, sulphate and iron in a phreatic aquifer system in The Netherlands. In: R. Cidu (ed.), Proceeding of the tenth international symposium on Water-Rock Interaction, Vol 1, pp. 521-524. Balkema, Lisse, The Netherlands.
HALE, W.E. (1972). Sample size determination for the log-normal distribution. Atmospheric Environment 6, pp. 419-422

HALLBERG G.R. & D.R. KEENEY (1993). Nitrate. Chapter 12 in: Alley (1993). Regional Groundwater Quality. Chapter 12, pp. 297-317.

HARTOG, N., J. GRIFFIOEN, P.F. VAN BERGEN & C.H. VAN DER WEIJDEN (2001). Determining the reactivity of reduced components in Dutch aquifer sediments. In: Impact of Human Activity on Groundwater Dynamics, IAHS Publication 269, pp. 221-229.

HARTOG, N., J. GRIFFIOEN & C.H. VAN DER WEIJDEN (in press). Distribution and Reactivity of O_2-Reducing Components in Sediments from a Layered Aquifer. Environmental Science and Technology.

HELSEL D.R. & R.M. HIRSCH (1992): Statistical methods in water resources. Studies in Environmental Science 49. Elsevier, Amsterdam.

HELSEL, D.R. (1993). In: Alley, W.M. ed. (1993). Regional Ground-Water Quality. Van Nostrand Reinhold, New York, pp.12.

HELSEL, D.R. (1997). The value of water quality information for national environmental policy. In: Proceedings of the international workshop Monitoring Tailor Made II, Nunspeet, The Netherlands.

HELVOORT, P.J. VAN, H.P. BROERS, P. SCHIPPER & C.A.J. APPELO (2000). Zware metalen in het grondwater: pyrietoxidatie en desorptie. 1. Veld- en laboratoriumonderzoek Oostrum. H_2O 24, pp. 15-18.

HERON G. & T.H. CHRISTENSEN (1994). The role of aquifer sediment in controlling redox conditions in polluted groundwater. In: Transport and Reactive Processes in Aquifers. Dracos & Stauffer (eds), Balkema, Rotterdam, pp. 73-79.

HETTERSCHIJT, R. & J.W. FOPPEN (1998). Well field Noordbargeres: design of a monitoring network for groundwater quality: MT3D scenario calculations. Netherlands Institute of Applied Geoscience TNO, Report 98-35-B. (in Dutch)

HOOGENDOORN, J.H. (1983). Hydrochemistry Eastern Netherlands. Netherlands Institute of Applied Geoscience TNO. Report no. OS 83-38. (in Dutch)

HUDAK, P.F. & H.A. LOAIGICA (1993). An optimization method for monitoring network design in multilayered groundwater flow systems. Water Resources Research 29, pp. 2835-2845.

HUISMAN, D. J. (1998). Geochemical characterization of subsurface sediments in The Netherlands. PhD. Thesis, Agricultural University Wageningen, The Netherlands.

ISAAKS, E.H. & R.M. SRIVASTAVA (1989). Applied Geostatistics. Oxford University Press, Oxford.

JACKS, G. (1993). Acid precipitation. Chapter 16 in: Alley, W. Regional Groundwater Quality, New York, van Nostrand Reinhold, pp. 405-421.

JACKS, G., G. KNUTSON, L. MAXE & A. FYLKNER (1984). Effects of acid rain on soil and groundwater in Sweden. In: Pollutants in porous media, ed. Yaron, G., Dagan, G. & Goldschmid, J. Springer-Verlag, Berlin, pp. 95-114.

JEDLITSCHKA, J. (1996). Groundwater monitoring in Germany. Proceedings Monitoring Tailor-Made II: Information strategies in water management. September 1996, Nunspeet, The Netherlands, pp. 65-73.

JOHNSTON, C.T., P.G. COOK, S.K. FRAPE, L.N. PLUMMER, E. BUSENBERG & R.J. BLACKPORT (1998). Groundwater age and nitrate distribution within a glacial aquifer beneath a thick unsaturated zone. Groundwater 36, pp. 171-180.

JONG, C. de (1986). Pollution loading in the 25 year protection zone of the Vierlingsbeek well field. Report Kiwa SWE 86.112. (in Dutch)

JOUSMA, G., J.C. GEHRELS, M.P. LAEVEN & J.G.F. VAN DAEL (1997). Evaluation of groundwater monitoring networks; methodology and examples. Netherlands Institute of Applied Geoscience TNO, Report NITG 97-256-B. (in Dutch)

KASSE, K. (1988). Early-Pleistocene tidal and fluviatile environments in the southern Netherlands and northern Belgium. Ph.D-thesis, Free University, Amsterdam.

KOLPIN, D.W., J.E. BARBASH & R.J. GILLIOM (1998). Occurrence of pesticides and shallow groundwater of the United States: initial results from the National Water Quality Assessment Program. Environmental Science and Technology 32, pp. 558-566.

KOROM, S. (1992). Natural denitrification in the saturated zone: a review. Water Resources Research 28, pp. 1657-1668.

LAEVEN, M. (1997). Nitrate at the Noorbargeres well field. Kiwa Report SWE 97.005, Nieuwegein, The Netherlands. (in Dutch)

LEAHY, P.P., B.J. RYAN & A.I. JOHNSON, A. I. (1993). An introduction to the U.S. Geological Survey's National Water-Quality Assessment Program. Water Resources Bulletin 29, pp. 529-532.

LOAIGICA, H.A. (1989). An optimization approach for groundwater quality monitoring network design. et Water Resources Research 25, pp. 1771-1782.

LOAIGICA, H.A. al. (1992). Review of groundwater quality monitoring network design. Journal of Hydraulic Engineering 118, pp. 11-37

LOFTIS, J.C. (1996) Trends in groundwater quality. Hydrological Processes 10, pp. 335-355
LOFTIS, J.C., G. B. McBRIDE & J.C. ELLIS (1991) Considerations of scale in water quality monitoring and data analysis. Water Resources Bulletin 27, pp. 255-264
MACDONNELL, L.J. & D.J. GUY (1991). Approaches to Groundwater Protection in the Western United States. Water Resources Research 27, pp. 259-265.
MALOSZEWSKI, P. AND A. ZUBER (1982). Determining the turnover time of groundwater systems with the aid of environmental tracers. 1. Models and their applicability. Journal of Hydrology 57, pp. 207-231.
MATSUNAGA, T., G. KARAMETAXAS, H.R.VON GUNTEN & P.C. LICHTNER (1993). Redox chemistry of iron and manganese minerals in river-recharged aquifers: a model interpretation of a column experiment. Geochimica et Cosmochimica Acta 57, pp. 1691-1704
MCDONALD, M.G. & A.W. HARBAUGH (1986). A modular three-dimensional finite-difference groundwater flow model. U.S. Geological Survey, Reston, Virginia, USA.
MEINARDI, C.R. (1994). Groundwater recharge and travel times in the sandy regions of The Netherlands. Ph.D. Thesis, Free University of Amsterdam.
MENKE, M.A. (1992). Loads of manure and fertilizer on Dutch soils. Report Waterloopkundig Laboratorium, Delft, The Netherlands T794. (in Dutch)
MEYER, P.D. & E.D. BRILL (1988). A method for locating wells in a groundwater monitoring network under conditions of uncertainty. Water Resources Research 24, pp. 1277-1282.
MODICA, E., T.E. REILLY & D.W. POLLOCK (1997). Patterns and age distribution of groundwater flow to streams. Groundwater 35, pp. 523-537.
MOURA, M.L. & S.B. KROONENBERG (1990). Geochemistry of Quaternary fluvial and eolian sediments in the southeastern Netherlands. Geologie & Mijnbouw 69, pp. 359-373.
MUELLER, D.K., P.A. HAMILTON, D.R. HELSEL, K.J. HITT & B.C. RUDDY (1995). Nutrients in groundwater and surface water of the United States - An analysis of data through 1992. U.S. Geological Survey: Reston, VA, Water resources Investigations Report 95-4031.
MYLONA, S. (1996). Sulphur dioxide emissions in Europe 1880-1991 and their effect on sulphur concentrations and depositions. Tellus 48B, pp. 662-689.
NOLAN, B.T., B. RUDDY, K.J. HITT & D.R. HELSEL. (1997). Risk of nitrate in groundwaters of the United States-a national perspective. Environmental Science and Technology 31, pp. 2229-2236.
PARKHURST, D.L. & C.A.J. APPELO (1999). User's guide to PHREEQC (version 2),- a computer program for speciation, batch-reaction, one-dimensional transport and inverse geochemical calculations. USGS Water-Resources Investigation Report 99-4259.
PEBESMA, E.J. (1996). Mapping Groundwater Quality in The Netherlands. Nederlandse Geografische Studies 199. Faculty of Geographical Sciences, Utrecht / Amsterdam. (Ph.D. thesis)
PEBESMA E.J. & J.W. DE KWAADSTENIET (1995). Changes in The Netherlands' groundwater quality between 5 and 17 m depth. National Institute of Public Health and the environment, RIVM Report 714810015. (in Dutch)
PEBESMA, E.J. & J.W. DE KWAADSTENIET (1997). Mapping groundwater quality in The Netherlands. Journal of Hydrology 200, pp. 364-386.
PEETERS, S.W.M. (1990). Land use classification of Noord-Brabant using LANDSAT images. Netherlands Institute of Applied Geoscience TNO. Report no. OS 90-18-A. (in Dutch)
PETERS, J.H. ed. (1998). Artificial recharge of groundwater. Proceedings of the third international symposium on artificial recharge of groundwater, TISAR 98. Balkema, Rotterdam.
PLUMMER, L.N., R.L. MICHEL, E.M. THURMAN & P.D. GLYNN (1993). Environmental tracers for age dating young groundwater. In: Alley, W.M. ed. Regional groundwater quality. New York, van Nostrand Reinhold, pp. 255-294.
POLLOCK, D.W. (1994). User's guide for MODPATH: A particle tracking post-processing package for MODFLOW. U.S. Geological Survey Open File Report 94-464.
POSTMA, D., C. BOESEN, H. KRISTIANSEN & F. LARSEN (1991). Nitrate reduction in an unconfined sandy aquifer: water chemistry, reduction processes, and geochemical modeling. Water Resources Research 27, pp. 2027-2045.
PRONK, A. (1994).Yearly statistics of fertilisers 1993-1994 (Jaarstatistiek van de Kunstmeststoffen 1993/'94), Periodieke rapportage 66-93/94. LEI-DLO, Den Haag. 44p. (in Dutch)
PROVINCIE NOORD-BRABANT (2000). Provincial Environment Programme 4. ISSN 0920-1408, 's-Hertogenbosch.
RAATS, P.A.C.(1978). Convective transport of solutes by steady flows II: specific flow problems. Agricultural Water Management 1, pp. 219-232.

RAATS, P.A.C. (1981). Residence times of water and solutes within and below the root zone. Agricultural Water Management 4, 63–82.
REIJNDERS, H.F.R., G. VAN DRECHT, H.F. PRINS & L.J.M. BOUMANS (1998): The quality of groundwater in The Netherlands. Journal of Hydrology 207, pp. 179-188.
RIVM (2001). Environment Balance ISSN 1383-4959, Bilthoven. (in Dutch)
ROBERTSON W.D. & J.A. CHERRY (1989). Large-scale dispersion in a sandy aquifer: simulation of subsurface transport of environmental tritium. Water Resources Research 32, pp. 3253-3266.
ROBIN, M.J.L., E.A. SUDICKY, R.W. GILLHAM & R.G. KACHANOSKI (1991). Spatial variability of strontium distribution coefficients and their correlation with hydraulic conductivity in the Canadian Forces base Borden Aquifer. Water Resources Research 27, pp. 2619-2632.
ROUHANI, S. & T.J. HALL (1988). Geostatistical schemes for groundwater sampling Journal of Hydrology 103, pp. 95-102.
SAALTINK, M.W., C. AYORA, P.J. STUYFZAND & H. TIMMER (1998). Modelling the effects of deep artificial recharge on groundwater quality. Proceedings of the third international symposium on artificial recharge of groundwater, TISAR 98. Balkema, Rotterdam, pp. 423-425.
SCHIPPER, P., P.J. VAN HELVOORT, C.A.J. APPELO & H.P. BROERS (2000). Zware metalen in het grondwater: pyrietoxidatie en desorptie. 2. Geochemisch modelonderzoek Oostrum. H_2O 24, pp. 19-22. (in Dutch)
SNELTING, H. ed. (1990). Installation and exploitation of the national monitoring network for groundwater quality. Report National institute of public health and environmental protection RIVM, no. 728517061. (in Dutch)
SOLOMON, D.K., R.J. POREDA, S.L.K. SCHIFF & J.A. CHERRY (1992). Tritium and Helium 3 as groundwater Age tracers in the Borden aquifer. Water Resources Research 28, pp. 741-755.
SPRUILL, T. B & L.C. CANDELA (1990). Two approaches to design of monitoring networks. Groundwater 28, pp. 430-442.
STATLINE (1999). - Website Statline of Netherlands Central Bureau of Statistics (www.statline.cbs.nl).
STEENVOORDEN, J.H.A.M. (1987). The fight against nitrate leaching from agricultural soils. In: Vulnerability of soil and groundwater to pollutants. Proceedings and Information No. 38 of the International Conference held in The Netherlands, TNO Committee on Hydrologic Research, Delft, The Netherlands , pp. 679-690.
STEENVOORDEN J.H.A.M., H. FONCK & H.P. OOSTEROM (1986). Losses of nitrogen from intensive grassland systems by leaching and surface runoff. In: Van der Meer HG, Ryden JC & Ennik GC (eds.) Nitrogen fluxes in intensive grassland systems. Martinus Nijhoff Publishers, Dordrecht, pp. 85-97.
STIBOKA. Various sheets of the 1:50,000 soil map series of The Netherlands.
STUURMAN, R.J., J.L. VAN DER MEIJ & A. BIESHEUVEL (1989). Hydrological systems analysis of western Noord-Brabant. Netherlands Institute of Applied Geoscience TNO, Report OS 89-59-A, Delft, The Netherlands. (in Dutch)
STUURMAN, R.J., J.L.VAN DER MEIJ, G.B. ENGELEN, A. BIESHEUVEL & F.J. VAN ZADELHOFF (1990). Hydrological systems analysis of western Noord-Brabant. Netherlands Institute of Applied Geoscience TNO, Report OS 90-25-A. (in Dutch)
STUYFZAND, P.J. (1991). The composition of rain water along the Dutch coast. Report Kiwa SWE 91.010. (in Dutch)
STUYFZAND, P.J. (1993). Hydrochemistry and hydrology of the coastal dune area of the western Netherlands. Ph.D. Thesis Free University of Amsterdam, Kiwa, Nieuwegein.
STUYFZAND, P.J. (1996). The geohydrochemical approach in optimizing groundwater quality monitoring. Proceedongs Monitoring Tailor Made II: Information Strategies in Water Managament, pp. 415-419. Published by RIZA, Lelystad, The Netherlands.
STUYFZAND, P.J. (1998a). Quality changes upon injection into anoxic aquifers in The Netherlands. Evaluation of 11 experiments. Proceedings of the third international symposium on artificial recharge of groundwater, TISAR 98. Balkema, Rotterdam, pp. 283-291.
STUYFZAND, P.J. (1998b). Simple models for reactive transport of pollutants and main constituents during artificial recharge and bank infiltration. Proceedings of the third international symposium on artificial recharge of groundwater, TISAR 98. Balkema, Rotterdam, pp. 427-434.
STUYFZAND, P.J. & J.A. MEIMA (eds.) (2000). Manual for geochemical research for aquifers. Report Kiwa, KOA 00.021, Nieuwegein, The Netherlands.
SVERDRUP, H.U. & P. WARFVINGE (1988). Weathering of primary silicate minerals in the natural soil environment in relation to a chemical weathering model. Water, Air & Soil Pollution 38, pp. 387-408.

TCB (1989). Advise on additional groundwater protection against diffuse sources of soil contamination. Report Netherlands Technical Committee on Soil Protection (TCB), Leidschendam. (in Dutch)
UFFINK, G.J.M. & P.F.A.M. ROMKENS (2001). Nitrate transport modelling in deep aquifers. Comparison between model results and data from the groundwater monitoring network. National institute of public health and environmental protection (RIVM), Report no711401010, Bilthoven, The Netherlands.
UFFINK, G.J.M. (2001). Modelling nitrate transport in deep aquifers: comparison of model results and field measurements. In: Impact of Human Activity on Groundwater, IAHS Publication no. 269, 2001, pp. 339-343.
UIL, H & J.E. VLOT (1991). Design of the primary groundwater quality monitoring network of the province of Drenthe. Netherlands Institute of Applied Geoscience TNO, Report nr. OS 91-26-B. (in Dutch)
VAN BEEK, C.G.E.M., H. BOUKES, D. VAN RIJSBERGEN & R. STRAATMAN (1988). The threat to The Netherlands waterworks by nitrate in the abstracted groundwater as demonstrated on the well field of Vierlingsbeek, The Netherlands. Water Supply 6, pp. 313-318.
VAN BEEK, C.G.E.M., F.A.M. HETTINGA & R. STRAATMAN (1989). Release of heavy metals in groundwater due to manure spreading. Proceedings 6th international symposium on Water-Rock Interaction, Malvern, U.K. 3-8 August 1989. Balkema, Rotterdam, Netherlands.
VAN BREEMEN, N., P.A. BURROUGH, E.J. VELTHORST, H. VAN DOBBEN, T. DE WIT, T.B. RIDDER & H.F. REIJNDERS (1982). Soil acidification from atmospheric ammonium sulphate in forest canopy throughfall. Nature 299, pp. 548-550.
VAN BREUKELEN, B.M., C.A.J. APPELO & T.N. OLSTHOORN (1998). Hydrogeochemical transport modelling of 24 years of Rhine water infiltration in the dunes of the Amsterdam water supply. Journal of Hydrology 209, pp. 281-296.
VAN BRUSSEL, J.F.M. (1990). Nitrate and sulfate predictions for the well fields Vessem, Gilze and Nuland. Kiwa, Report SWO 89.255, Nieuwegein, The Netherlands. (in Dutch)
VAN DER EEM, J.P. (1991). Pesticides and phreatic groundwater abstraction. Kiwa-Report SWO 90.312, Kiwa, Nieuwegein, The Netherlands, 39 pp. (in Dutch)
VAN DER GRIFT, B. & C.G.E.M VAN BEEK (1996). Hardness of abstracted groundwater: indicative predictions. Report Kiwa Onderzoek & Advies, Nieuwegein, The Netherlands. (in Dutch)
VAN DER SLUIS, P. & J.J. DE GRUIJTER (1985). Water table classes: a method to describe seasonal fluctuation and duration of water tables on Dutch soil maps. Agricultural Water Management 10, pp. 109-125.
VAN DRECHT. G. (1993). Modeling of regional scale leaching from agricultural soils, The Netherlands. Applied Geochemistry, Suppl. 2, pp. 175-178.
VAN DRECHT, G., F.R. GOOSSENSEN, M.J.D. HACK-TEN BROEKE, E.J. JANSEN AND J.H.A.M. STEENVOORDEN (1991). Calculation of nitrate-leaching to groundwater using simple models. Report Staring Centrum (Alterra), no. 163, Wageningen.
VAN DRECHT, G., H.F.R. REIJNDERS, L.J.M. BOUMANS. & W.VAN DUIJVENBOODEN (1996). The quality of the groundwater between 5 and 30 m depth in The Netherlands in the year 1992 and changes in the period 1984-1993. National institute of public health and environment, RIVM Report 714801005, Bilthoven, The Netherlands. (in Dutch)
VAN DUIJVENBOODEN, W & H.G. VAN WAEGENINGH, H.G. (1987). Vulnerability of soil and groundwater to pollutants. Proceedings and Information No. 38 of the International Conference. TNO Committee on Hydrologic Research, Delft, The Netherlands,
VAN DUIJVENBOODEN, W (1987). Groundwater quality monitoring networks: design and results. In: Vulnerability of soil and groundwater to pollutants. Proceedings and Information No. 38 of the International Conference, TNO Committee on Hydrologic Research, Delft, The Netherlands pp. 179-192.
VAN DUIJVENBOODEN, W. (1993). Groundwater quality monitoring in The Netherlands. In: Alley, W.M. ed. Regional groundwater quality. New York, van Nostrand Reinhold, pp. 515-535.
VAN DUIJVENBOODEN, W., J.TAAT & L.F.L. GAST (1985). The Netherlands monitoring network for groundwater quality: Report on the installation phase. National institute of public health and environmental protection Report no. 840382001, Leidschendam, The Netherlands. (in Dutch)
VAN DUIJVENBOODEN, W. & A. BREEUWSMA (eds.) (1987). Vulnerability of the groundwater Mapping properties of the Dutch soil in relation to vulnerability of groundwater for contaminants. National institute of public health and environment, RIVM Report no. 840387003, Bilthoven, The Netherlands. (in Dutch)

VAN ELBURG, H., G.B. ENGELEN, & C.J. HEMKER (1993). FLOWNET, Manual, version 5.4. Free University of Amsterdam.
VAN GEER, F.C. (1987). Design of groundwater head monitoring networks at drinking water well fields Stage 1: evaluation of existing networks. Netherlands Institute of Applied Geoscience TNO, Report OS 87-39. (in Dutch)
VAN GEER, F.C. (1988) Determination of groundwater head drawdown at the well field Spannenburg using transfer-noise modelling. H2O 21. pp. 451-454. (in Dutch)
VAN OMMEN, H.C.(1986). Influence of diffuse sources of contamination at the quality of outflow in groundwater including non-equilibrium sorption and decomposition. Journal of Hydrology 88, pp. 79-95.
VAN OMMEN, H.C., H.C. VAN GENUCHTEN, W.H. VAN DER MOLEN, R. DIJKSMA, & J. HULSHOF (1989). Experimental and theoretical analysis of solute transport from a diffuse source of pollution. Journal of Hydrology 105, pp. 225-251.
VAN VLIET, M. (2000). Interpretation, evaluation and optimization of the groundwater quality monitoring networks of the province of Noord-Brabant: Reports 3 and 4. Netherlands Institute of Applied Geoscience TNO, Reports 00-234-B and 00-235-B. (in Dutch).
VAN VUGHT A. & B. VAN DER EIJNDEN (1999). Transport modelling of the Noordbargeres well field. A study on the feasibility of monitoring network optimization using a fine-discretisized transport model. M.Sc. Thesis, University of Utrecht, Netherlands Institute of Applied Geoscience TNO, Report NITG-99-118-B. (in Dutch)
VELDKAMP, A. & S.B. KROONENBERG (1993). Application of bulk sand geochemistry in mineral exploration and Quaternary research: a methodological study of the Allier and Dore terrace sands, Limagne rif valley, France. Applied Geochemistry 8, pp. 177-187.
VENEMA, P., H.P. BROERS, & J. PEETERS (2000). The quality of the groundwater in Noord-Brabant in 1998., Netherlands Institute of Applied Geoscience TNO, Delft, The Netherlands Report NITG 00-232-B. (in Dutch)
VON GUNTEN U., & J. ZOBRIST (1993). Biogeochemical changes in groundwater-infiltration systems: column studies. Geochimica et Cosmochimica Acta 57, pp. 3895-3906.
VRIES, F. DE & J. DENNEKOM (1992). The digital soil map of The Netherlands. Technical document no.1. Staring Centre DLO (Alterra), Wageningen, The Netherlands.
VRIES. J.J. DE (1977). The stream network in The Netherlands as a groundwater discharge phenomenon Geologie & Mijnbouw 5, pp. 103-122.
VRIES, J.J. DE (1994). Dynamics of the interface between streams and groundwater systems in lowland areas, with reference to streamnet evolution. Journal of Hydrology 155, pp. 39-56.
VRIES, J.J. DE (1995). Seasonal expansion and contraction of stream networks in shallow groundwater systems. Journal of Hydrology 170, pp. 15-26.
VRIES, W. DE (1988). Critical deposition levels for nitrogen and sulphur on Dutch forest ecosystems. Water, Air, and Soil Pollution 42, pp. 221-239.
VRIES W, DE (1993). Average critical loads for nitrogen and sulfur and its use in acidification abatement policy in The Netherlands. Water, Air, and Soil Pollution 68, pp. 399-434.
VRIES, W. DE & A. BREEUWSMA (1987). The relation between soil acidification and element cycling. Water, Air, and Soil Pollution 35, pp. 293-310.
WARD, R.C., J.C. LOFTIS & G.B. MCBRIDE (1986). The 'data-rich but information-poor' syndrome in water quality monitoring. Environmental Management 10, pp. 291
WEERTS, H.J.T. (1996). Complex confining layers; architecture and hydraulic properties of Holocene and Late Weichselian deposits in the fluvial Rhine-Meuse delta, The Netherlands. Ph.D.Thesis Utrecht University, Netherlands Geographical Studies 213.
WOLDT, W. & I. BOGARDI (1992). Groundwater monitoring network design using multiple criteria decision-making and geostatistics. Water Resources Bulletin 28, pp. 45-62.
WOLTER, R., S. ROSENBAUM & S. HANNAPPEL (2001). The German groundwater monitoring network. Proceedings of Monitoring Tailor Made (III), September 2000, Nunspeet, The Netherlands.
YUE, S., P. PILON & G. CAVADIAS (2002). Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. Journal of Hydrology 259, pp. 254-271.
ZHANG K. (1996). Modelling tools for groundwater monitoring network design. M.Sc. Thesis H.H. 274, International Institute for Infrastructural, hydraulic and environmental engineering HIE, Delft, The Netherlands.
ZOELLMANN, K., W. KINZELBACH, & C. FULDA (2001). Environmental tracer transport (^3H and SF6) in the saturated and unsaturated zones an its use in nitrate pollution management. Journal of Hydrology 240, pp. 187-205.

Summary

Diffuse contamination of groundwater is monitored at the national scale, at the regional scale of provinces and at the local scale of phreatic well fields. Monitoring studies often address one of the individual fields of monitoring statistics, hydrology or hydrogeochemistry. The aim of the present study was to integrate statistical, hydrological and hydrogeochemical methods and information in the design, the data analysis, the evaluation and the optimization of regional and local scale monitoring networks. The central hypothesis is that more effective monitoring is achieved when using hydrological and hydrogeochemical information plus concepts of advective and reactive transport to steer the monitoring design and data analysis. The specific research issues and aims of this study were:

1. to investigate the influence of a drainage network and aquifer heterogeneity on the groundwater age distribution and to test whether regional mapping can be used to predict the age distribution in a regional groundwater quality monitoring network
2. to integrate information on groundwater age and hydrogeochemical processes in the design and data analysis of regional monitoring networks, and to investigate if such an approach yields extra value in the identification of groundwater quality patterns
3. to design a framework for the evaluation and optimization of regional groundwater quality monitoring networks, using area-specific information goals for areas with low, moderate and high risks for the contamination of deeper groundwater
4. to improve the detection and interpretation of groundwater quality changes with time, combining time-series information, concentration-depth profiles, age dating and concentration-depth prognosis based on data on the historical input of solutes
5. to judge the effectiveness of monitoring configurations for phreatic well fields, using a three-dimensional travel time approach and scenarios with advective and simple reactive transport
6. to formulate sampling objectives and initiatory data analysis protocols for the reactive properties of sediments at drinking water well fields in order to predict the evolution of the quality of the extracted groundwater.

Figure S.1 shows how the six research issues fit in a general scheme for groundwater quality monitoring. Together, the research issues 1 to 4 address the 8 stages of regional groundwater quality monitoring, using the regional networks of Noord-Brabant and Drenthe as case studies. Conceptual groundwater modelling, hydrogeochemical calculations, age dating and non-parametric statistical methods were used for designing and evaluating two regional networks and for identifying and interpreting spatial and temporal contamination patterns in the monitored regions. Research issues 5 and 6 relate to the monitoring of drinking water well fields, and emphasis is on the information analysis, and design and installation stages of the network operation. Strategies for groundwater quality monitoring and the sampling of reactive sediments were evaluated using simple models of advective and simple reactive transport.

Research issue 1 - Groundwater age distribution in regional monitoring
Groundwater age distribution is a key factor determining the distribution of dissolved contaminants in the subsurface when contamination loadings have increased in time. The effects of superficial drainage and aquifer heterogeneity on the groundwater age distribution in unconsolidated aquifers in flat areas were investigated, and consequences were presented for the monitoring of contaminants from diffuse sources. First, the effects were assessed using model simulations. Second, the groundwater age distribution was evaluated in the two regional monitoring networks of Noord-Brabant and Drenthe using tritium measurements.

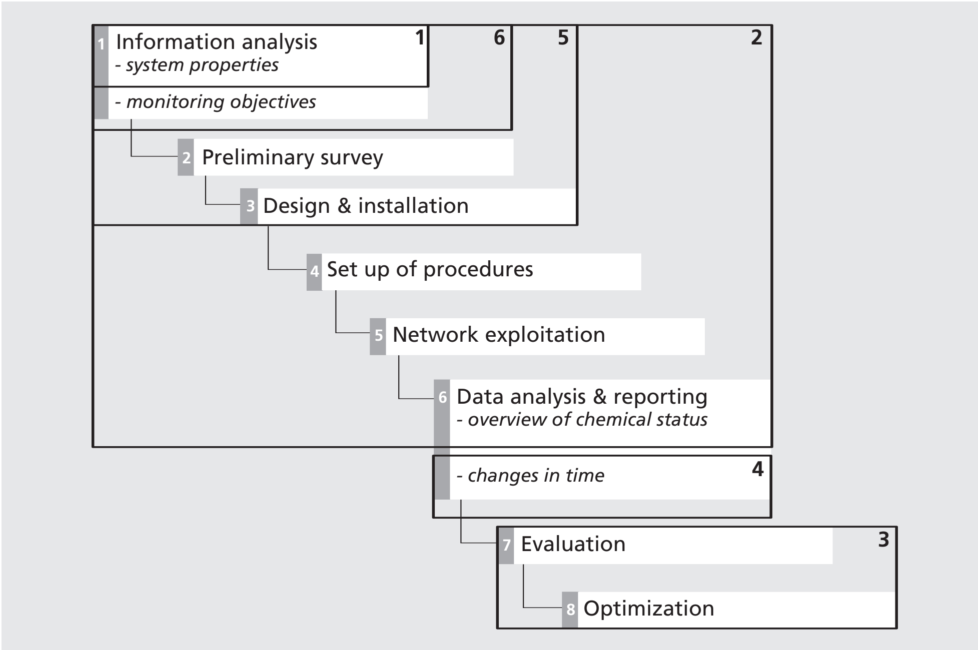


Figure S.1 - Aspects of groundwater quality monitoring that are addressed in the research issues 1 to 6 (bold numbers)

Theoretically, a simple groundwater age distribution is present in homogeneous aquifers with natural groundwater recharge, characterized by a horizontal pattern of residence time isochrones and a gradual increase of groundwater age with depth. The model simulations show that the isochrone pattern becomes distorted in areas with a superficial drainage network, resulting in relatively old groundwater at shallow depth and larger variation in groundwater age at a specific depth. This drainage effect on the groundwater age distribution is relatively large compared with effects of aquifer heterogeneity or spatially varying groundwater recharge.

The effects of superficial drainage on the groundwater age distribution were confirmed by tritium measurements made in the regional monitoring networks of two provinces in the Netherlands. At about 22 m depth, the proportion of post-1950 groundwater in drained areas was significantly less and the groundwater age variation was larger than in recharge areas that lack a drainage network. The age of the groundwater appeared to be related to the drainage network density and the water table class. A preliminary survey showed that contamination patterns in the two networks agree well with the proportion of post-1950 groundwater.

Research issue 2 - Integrating groundwater age and reactive processes in the design and data analysis of regional monitoring

The design and data-analysis of two regional networks in the Netherlands was based on sampling from land-use/soil-type/geohydrology strata. These strata were called *homogeneous areas* because the variation between those areas was expected to be larger than the variation within them. Two main hypotheses were examined: (1) extra geohydrological stratification helps to reduce the variation in the data, and (2) extra indicators based on geochemical knowledge provide a better identification and understanding of contamination patterns than the sole statistical analysis of concentrations of targeted contaminants. The study concentrated on the vulnerable sandy areas in the two provinces. Non-parametric statistical methods were applied to estimate typical concentrations and proportions of contaminated groundwater in the homogeneous areas.

The results indicate that the extra geohydrological stratification, based on hydro-geomorphological characteristics of areas, helps detect general contamination patterns in the provinces. Often the effects of land use on groundwater quality would have been overlooked when using only land-use and soil-type as stratification factors, because a large part of the variation is explained by differences in groundwater age in recharge, intermediate and discharge areas. Hydrogeochemical processes explain another part of the variation between and within the homogeneous areas. Extra indicators that were based on geochemical knowledge, including the oxidation capacity and the hardness/alkalinity ratio, helped detect the impact of these processes and the identification and understanding of contamination patterns. For example, oxidation capacity was used as indicator for agricultural pollution, because it behaves conservatively under the condition that pyrite oxidation is the only reaction that controls the transport of nitrate.

Overall, the geohydrological stratification and the use of conditionally conservative chemical indicators in the data analysis yielded extra value because they both lead to more homogeneity and smaller skewness of the concentrations in the homogeneous areas. Distinctive groundwater quality patterns between similar homogeneous areas in the provinces Noord-Brabant and Drenthe would have remained unnoticed without the extra stratification and the additional geochemical indicators.

Research issue 3 - Framework for evaluation and optimization in regional monitoring
In the design stage of the Dutch national and regional monitoring networks neither the monitoring information goals were clearly defined, nor were the corresponding statistical information goals specified. The study aimed to design a framework for the evaluation and optimization of the Dutch regional monitoring networks which is based on the vulnerability and the pollution loading of the individual homogeneous areas. The underlying design strategy of these networks was to have higher ambitions for areas with high risks for contamination of deep groundwater than for areas with moderate or low risks. The presented framework assigns area-specific statistical information goals to high-risk, moderate-risk and low-risk areas. The approach uses the collected data and non-parametric methods to evaluate the sample size and the monitoring frequencies and depths and to optimize the network step-by-step. The advantages of the methodology are: (1) a stronger focus on the high-risk areas, which results in less uncertainty in the estimates of typical concentrations and the proportion of contaminated groundwater, and (2) feasibility of cost-reduction in low-risk and moderate-risk areas, while maintaining the opportunities for the detection of groundwater contamination and for the assessment of regional groundwater quality patterns.

The methodology was demonstrated for two regional monitoring networks. The results indicated that the sample size of the networks is sufficient to assess general groundwater quality patterns and to obtain a regional overview of groundwater chemical status. However, judging the monitoring effectiveness in individual homogeneous areas, especially the sample size in many high-risk areas appeared insufficient to monitor typical concentrations of targeted contaminants or the proportion of contaminated groundwater. Net cost reduction was achieved by decreasing monitoring frequency in low-risk and moderate-risk areas, while increasing sample size in high-risk areas.

Research issue 4 - Detection and understanding of temporal changes
Changes in agricultural practices are expected to affect groundwater quality by changing the

loads of nutrients and salts in recharging groundwater, but regional monitoring networks installed to register the changes often fail to detect them and interpretation of trend analysis results is difficult. Normally, the trend detection is limited to the analysis of time series of individual wells or groups of wells. Trends in groundwater quality can also be detected using concentration-depth information, because depth and groundwater age are interrelated, which is most pronounced in recharge areas. The study aimed at improving the detection and the understanding of groundwater quality changes with time. For *trend detection*, a combination was used of trend analysis on time series at specific depths and time-averaged concentration-depth profiles. To reveal trends that have become obscured by chemical reactions, additional *conditionally conservative* indicators were introduced that are insensitive to those reactions under specific conditions. Significant trends were matched with prognoses of conservative and reactive transport to aid the *understanding* of trends. The prognoses are based on groundwater age determinations using tritium and historical time series of the inputs of components in recharging groundwater A hydrogeochemical model was used for reactive transport prognoses. Data of the regional network in 2 homogeneous areas in the Dutch province of Noord-Brabant were used to illustrate the approach.

The results from the monitoring network of Noord-Brabant indicated that many targeted contaminants have become retarded or delayed and that quality changes were hard to detect for many reactive solutes, including the nutrients nitrate and potassium. As a result, pollution fronts of these targeted chemical components are limited to the first 15 m of the subsoil. At deeper level, about 20-25 m, the effects of agricultural pollution and acidification were indicated only by chemical indicators that have not been considered by others: oxidation capacity, the sum of cations and chloride. The downward movement of the agricultural pollution fronts was demonstrated for the 2 homogeneous areas in Noord-Brabant with intensive livestock farming. Increasing trends of the conditionally conservative indicators 'oxidation capacity' and 'sum of cations' were found at larger depth (18-25 m below surface). Increasing trends for potassium were found at smaller depth (7-13 m), which is explained by retardation of potassium due to cation-exchange with calcium and magnesium. The modelled cation-exchange explained the shape of the concentration-depth profile and the increasing trends at shallow level in the aquifer. No significant quality changes could be demonstrated for nitrate, due the disappearance of nitrate through oxidation by pyrite and organic matter.

Overall, the combined use of time series information, the evaluation of concentration-depth profiles, age dating and conservative and reactive prognoses has large advantages in the detection and interpretation of temporal trends.

Research issue 5 - Monitoring configurations at phreatic well fields

Observation networks around phreatic well fields are installed with three main monitoring objectives: (1) signaling of unexpected threats to the quality of extracted groundwater (*early warning*), (2) supporting operational decisions by the prediction of future quality changes (*prediction*), and (3) evaluating protection measures in the protection zone (*protection*). Monitoring configurations of the well fields are often based on a horizontal two-dimensional concept of groundwater flow, focusing on early warning using wells at 10 or 15 years residual transit time to the pumping well. Interestingly, the design of the Dutch regional monitoring networks (research issues 1 to 4) emphasized the vertical flow component and the residence time from the earth surface to the observation screens. Groundwater modelling was used to assess the 3D travel time distribution at well fields in order to judge different monitoring configurations for the three objectives mentioned for advective and simple reactive transport.

The model calculations showed the need to evaluate the travel time distribution in three dimensions for effective monitoring. The model results indicated that the location and especially the depth of the observation wells should be carefully chosen, in accordance with the residence time from the surface to the observation well, the residual transit times to the extraction well, and the transformation and retardation rates. The larger the degradation rates or retardation, the shallower should the monitoring be for effective early warning and prediction of future groundwater quality. Shallow monitoring was most functional for a variety of objectives and conditions.

Shallow monitoring at the boundary of the 10 year protection zone is found most suitable for the early warning objective and also appropriate for use in prediction models. Deeper monitoring at this boundary, which is frequent in existing monitoring networks in the Netherlands, appears ineffective because of the long residence times to the screens and the late arrival of pollution fronts relative to the arrival in the pumping well.

Research issue 6 - Sampling reactivity

Sediment reactivity is one of the largest unknowns in the interpretation of groundwater quality data and the largest impediment for prediction of groundwater quality changes. For prognoses of the evolution of groundwater quality at the scale of well fields, there is an increasing need for specific information goals and standardised procedures for the sampling of reactive sediments. The aim of the study was to formulate sampling objectives and initiatory data analysis protocols for the reactive properties at drinking water well fields, in order to produce input for transport models that are used to predict the evolution of the groundwater quality.

Information analysis was done for the acquisition of reactivity data using conceptual models for transport in geochemically and hydraulically layered porous media. The results show that different information goals are needed for the sampling at phreatic well fields and the sampling at deep-well recharge systems, because of different directions of the propagation of the reaction fronts relative to the main direction of geochemical variation. Reaction fronts at phreatic well fields move vertically through predominantly horizontally layered aquifers. For this situation, average contents of the reactive properties are sufficient for transport modelling. Reaction fronts at deep-well recharge sites move horizontally through the horizontally layered sediments. Information about the vertical variation of sediment reactivity is needed to predict solute breakthrough. This information is obtained by estimating percentiles of the frequency distribution of the reactive property.

The different information goals for the two situations result in different sampling designs. A case study of the Oostrum aquifer demonstrated that the uncertainty of the estimated reactive properties directly affects the uncertainty of the calculated breakthrough. Hence, a proper design of the sampling strategy for aquifer reactivity helps to reduce the uncertainty of the transport model predictions.

General conclusions

The study has shown that significant improvement of monitoring effectiveness is obtained by integrating concepts of advective and reactive transport and using hydrological and hydrogeochemical information in the design, the data analysis and the evaluation and optimization of groundwater quality monitoring networks. An effective monitoring strategy needs well-focused monitoring objectives and statistical information goals and substantial understanding of the hydrological and hydrogeochemical system properties. The most important hydrological and hydrogeochemical system properties for the monitoring of diffuse contaminants are: (1) the 3D travel time distribution in the aquifer, and (2) the reactivity of the introduced solutes and the subsurface sediments. Evaluation of residence times and hydrogeochemical calculations should preferably be integrated in the regular, statistical methods that are used to match the monitoring objectives and specific information goals.

Monitoring efficiency is best when the monitoring information goals and statistical information goals are defined before the actual installation and operation of the monitoring networks, and data analysis protocols are established during the design stage.

Samenvatting

Het doel van onderhavige studie is om hydrogeochemische, hydrologische en meetnet-statistische kennis en methoden te integreren in het ontwerp, de data-analyse, de evaluatie en de optimalisatie van grondwaterkwaliteitsmeetnetten De studie richt zich op de monitoring van grondwaterkwaliteit in relatie tot diffuse verontreinigingen. De centrale hypothese is dat een effectiever meetnet ontstaat als hydrologische en hydrogeochemische kennis en informatie wordt gecombineerd via concepten van advectief en reactief transport. De specifieke doelstellingen van de studie zijn:

1. Nagaan hoe de leeftijdsopbouw van het grondwater wordt beïnvloed door oppervlakkige afwatering en een heterogene opbouw van de ondergrond, en welke consequenties dat heeft voor grondwaterkwaliteitsmonitoring
2. Integreren van informatie over verblijftijden en hydrogeochemische processen in het ontwerp en de data-analyse van regionale meetnetten en nagaan wat de meerwaarde daarvan is voor het herkennen van ruimtelijke grondwaterkwaliteitspatronen
3. Ontwerpen van een methode voor de evaluatie en optimalisatie van regionale meetnetten met specifieke meetdoelstellingen voor gebieden met lage, middelbare en hoge risico's voor de verontreiniging van het diepe grondwater
4. Verbeteren van de detectie en interpretatie van grondwaterkwaliteitsveranderingen in de tijd
5. Beoordelen van de effectiviteit van meetnetconfiguraties bij freatische waterwingebieden
6. Formuleren van meetdoelen en data-analyse protocollen voor de bemonstering van reactieve eigenschappen van sedimenten in waterwingebieden.

Figuur S.1 (zie Summary) toont hoe de zes onderwerpen passen in een generiek schema voor grondwaterkwaliteitsmonitoring. De doelstellingen 1 tot met 4 worden behandeld aan de hand van de regionale meetnetten van Drenthe en Noord-Brabant. Bij het ontwerp en de evaluatie van de twee meetnetten zijn conceptuele grondwatermodellering, hydrogeochemische berekeningen, grondwaterdatering en non-parametrische statistische methoden gebruikt om ruimtelijke en temporele verontreinigingspatronen te identificeren en te interpreteren. De studieonderwerpen 5 en 6 hebben betrekking op het monitoren van waterwingebieden. De nadruk ligt op de fasen van 'informatie-analyse' en 'ontwerp en installatie' van de meetnetten. Om meetstrategieën voor grondwaterkwaliteit en sedimentreactiviteit te evalueren zijn eenvoudige modellen van advectief en simpel reactief transport gebruikt.

1 - Leeftijdsopbouw van het grondwater op regionale schaal

De leeftijdsopbouw van het grondwater is een cruciale factor voor de ruimtelijke patronen van diffuse grondwaterverontreiniging. Daarom zijn de effecten van een oppervlakkig afwateringsstelsel en een heterogene opbouw van de ondergrond op de leeftijdsopbouw onderzocht met behulp van modelsimulaties en tritiummetingen in twee regionale meetnetten.

In homogene aquifers met natuurlijke grondwateraanvulling neemt de ouderdom van het grondwater normaalgesproken met de diepte toe en is een horizontaal patroon van isochronen, lijnen van gelijke verblijftijd, aanwezig. Dit isochronenpatroon wordt volgens de modelsimulaties verstoord in gebieden met een oppervlakkig afwateringsstelsel, hetgeen resulteert in relatief oud grondwater op geringe diepte en in een grote variatie in grondwaterleeftijd op een bepaalde diepte. Het effect van oppervlakkige drainage is relatief groot ten opzichte van de effecten van een heterogene opbouw van de ondergrond en van een ruimtelijk variabele grondwateraanvulling. De resultaten van de modelsimulaties worden bevestigd door tritiummetingen in de meetnetten van de provincies Drenthe en Noord-Brabant. In gebieden met een oppervlakkig afwateringsstelsel is het percentage water dat na 1950 is geïnfiltreerd (post-1950) kleiner dan in gebieden zonder zo'n afwateringsstelsel. Ook is de variatie in leeftijden groter, met name op de reguliere meetdiepte van 20-25 meter onder maaiveld. Het percentage post-1950 grondwater kan worden gerelateerd aan de slootdichtheid en de grondwatertrap. De verontreinigingspatronen in de twee regionale meetnetten blijken duidelijk gerelateerd aan het percentage post-1950 grondwater.

2 - Ontwerp en data-analyse van twee regionale meetnetten
Het ontwerp en de data-analyse van de provinciale meetnetten van Drenthe en Noord-Brabant zijn gebaseerd op een gestratificeerde steekproef uit zogenaamde homogene gebiedstypen (landgebruik/bodemtype/geohydrologie strata). De a priori aanname was dat de variatie tussen de homogene gebiedstypen groter is dan de variatie binnen de gebiedstypen zelf. In de studie zijn twee basishypothesen getoetst: (1) extra geohydrologische stratificatie reduceert de variatie in de gebiedstypen, en (2) extra analyse van hydrogeochemische indicatoren leidt tot een scherpere identificatie en een beter begrip van verontreinigingspatronen ten opzichte van het uitsluitend gebruik van verontreinigingsindicatoren waarvoor een wettelijke normstelling aanwezig is. De gepresenteerde studie beperkt zich tot de Pleistocene zandgebieden in de twee provincies.

De extra geohydrologische stratificatie, die is gebaseerd op hydrogeomorfologische gebiedseigenschappen, verbetert de detectie van verontreinigingspatronen. Het verschil in grondwaterleeftijdsopbouw in infiltratiegebieden, intermediaire gebieden en kwelgebieden verklaart een groot deel van de variatie in verontreinigingsconcentraties binnen een bepaalde landgebruik-bodem categorie. De stratificatie leidt aldus tot een scherpere afbakening van verontreinigde gebieden. De hydrogeochemische processen verklaren een ander deel van de variatie tussen en binnen de homogene gebiedstypen. De extra hydrogeochemische indicatoren, waaronder het oxidatievermogen en de hardheid/alkaliniteit-ratio, helpen om de impact van deze processen op te sporen en verontreinigingspatronen te identificeren en te begrijpen. Het oxidatievermogen blijkt bijvoorbeeld een goede indicator voor 'vermesting', omdat de indicator zich conservatief gedraagt onder de voorwaarde dat pyrietoxidatie het enige proces is dat transport van nitraat controleert. Het gebruik van dergelijke 'voorwaardelijk conservatieve' indicatoren en de extra geohydrologische stratificatie leiden beide tot een kleinere spreiding van concentraties en een minder scheve frequentieverdeling in de homogene gebiedstypen en leveren zo een meerwaarde in de data-analyse. Zonder de extra stratificatie en de hydrogeochemische indicatoren zouden duidelijke verschillen tussen gelijksoortige gebieden in Drenthe en Noord-Brabant niet zijn opgemerkt. Het diepe grondwater in de Drentse infiltratiegebieden blijkt bijvoorbeeld kwetsbaarder voor nitraatuitspoeling dan het diepe grondwater in Noord-Brabant.

3 - Evaluatie en optimalisatie van regionale meetnetten
In het ontwerpstadium van de Nederlandse landelijke en provinciale grondwaterkwaliteitsmeetnetten zijn weliswaar beleidsmatige meetdoelen opgesteld, maar die zijn niet expliciet uitgewerkt in specifieke (statistische) meetdoelen. Daarom is een evaluatie- en optimalisatiemethode uitgewerkt voor de provinciale meetnetten die uitgaat van de kwetsbaarheid en de verontreinigingsbelasting in de homogene gebiedstypen. De methode definieert gedifferentieerde ambitieniveaus voor gebieden met verschillend risico voor de verontreiniging van het diepe grondwater. De ambitieniveaus zijn vertaald in specifieke statistische informatiedoelen voor gebieden met een laag, middelbaar en hoog risico. De aanpak maakt gebruik van de verzamelde meetgegevens en non-parametrische methoden om het aantal meetpunten (sample size) en de meetfrequenties en meetdiepten te evalueren en stap-voor-stap te optimaliseren. De

voordelen van de methode zijn: (1) nadruk op hoog-risico gebiedstypen, hetgeen leidt tot kleinere onzekerheid in de bepaling van typische concentraties in het percentage verontreinigd grondwater, (2) kostenreductie in laag- en middel-risico gebiedstypen, met handhaving van de mogelijkheden om achtergrondwaarden te bepalen en regionale grondwaterkwaliteitspatronen te identificeren. De methode is uitgewerkt voor de provinciale meetnetten van Drenthe en Noord-Brabant. Uit de evaluatie blijkt dat het aantal meetpunten voldoende is om een regionaal overzicht van de chemische toestand van het grondwater te verkrijgen. De meetnet-effectiviteit in hoog-risico homogene gebiedstypen is daarentegen onvoldoende om typische concentraties of het percentage verontreinigd grondwater met voldoende precisie vast te stellen. De optimalisatie van de meetnetten, en netto verlaging van de exploitatiekosten, kan worden bereikt met een gerichte uitbreiding van het aantal meetpunten in hoog-risico gebiedstypen en reductie van meetfrequentie in laag- en middel-risico gebiedstypen.

4 - Detectie en interpretatie van temporele trends

Het vaststellen van ontwikkelingen in de grondwaterkwaliteit in de tijd is één van de belangrijkste doelstellingen van grondwaterkwaliteitsmonitoring. Ondanks het feit dat in het Nederlandse grondwater een stijgende trend in concentraties mag worden verwacht op basis van de historische ontwikkeling in de belasting met de nutriënten en zouten in infiltrerend grondwater, bleek het in het verleden lastig deze trends daadwerkelijk aan te tonen. Ook is de interpretatie van de gevonden trends moeizaam. Trendanalyse beperkt zich meestal tot analyse van tijdreeksen in individuele putten of groepen van putten. In principe kunnen trends in de grondwaterkwaliteit ook worden gedetecteerd met behulp van concentratie-diepte informatie, omdat de grondwaterleeftijd en de diepte gerelateerd zijn, met name in infiltratiegebieden. De doelstelling van de studie was om de trenddetectie en de interpretatie te verbeteren door informatie uit tijdreeksen van een bepaalde diepte te combineren met tijd-gemiddelde concentratie-diepte profielen. Om trends te detecteren die gemaskeerd worden door geochemische processen, werd gebruik gemaakt van de onder punt 3 geïntroduceerde 'voorwaardelijk conservatieve' indicatoren. Om de gevonden significante trends te interpreteren zijn de trendresultaten vergeleken met prognoses van conservatief en reactief transport. De prognoses zijn gebaseerd op grondwaterdatering met tritium en tijdreeksen van de historische belasting met stoffen in infiltrerend grondwater. Voor de prognoses met reactief transport is een hydrogeochemische modelcode gebruikt.

De gegevens van twee homogene gebiedstypen met intensieve veehouderij in de provincie Noord-Brabant zijn gebruikt om de methode te illustreren. De verontreinigingsfronten van milieuindicatoren zoals nitraat, aluminium en kalium zijn doorgedrongen tot de eerste 15 meter van de ondergrond. De invloed van 'vermesting' is echter op grotere diepte aantoonbaar (20-25 m) hetgeen blijkt uit de 'voorwaardelijk conservatieve' indicatoren oxidatievermogen, som kationen en chloride. Het vermestingsfront in de twee gebiedstypen verplaatst zich in de diepte; stijgende trends voor oxidatievermogen en som kationen zijn aangetoond tussen 18 en 25 meter diepte. Deze verplaatsing is in overeenstemming met de prognose voor conservatief transport. Stijgende trends voor kalium werden op kleinere diepte gevonden (7-13 m) hetgeen wordt verklaard door retardatie van kalium. De vorm van het gemeten concentratie-diepte profiel en de stijgende trends tussen 7 en 13 meter zijn in overeenstemming met de reactieve prognose voor kaliumtransport met kationuitwisseling tegen calcium en magnesium.

Concluderend biedt het gecombineerde gebruik van tijdreeksinformatie, concentratie-diepte profielen, grondwaterdatering en prognoses voor conservatief en reactief transport duidelijke voordelen bij de detectie en de interpretatie van temporele trends.

5 - Monitoringconfiguraties in freatische waterwingebieden
Doelstellingen van meetnetten van freatische drinkwaterwinningen zijn: (1) signaleren van onverwachte bedreigingen voor de kwaliteit van het onttrokken grondwater (early warning), (2) voorspellen van toekomstige kwaliteitsveranderingen, en (3) evalueren van beschermingsmaatregelen binnen het waterwingebied. Bestaande meetnetontwerpen gaan vaak uit van een horizontaal 2-dimensionaal stromingsconcept en zijn vooral gericht op het tijdig signaleren met behulp van waarnemingsputten op tien tot vijftien jaar residuele reistijd naar de pompputten. In het ontwerp van regionale meetnetten daarentegen, gaat de aandacht vooral uit naar de verblijftijden tussen aardoppervlak en waarnemingspunt. In de studie is de 3-dimensionale reistijdverdeling gesimuleerd om de geschiktheid te beoordelen van verschillende meetnetconfiguraties voor de drie genoemde doelstellingen en scenario's met conservatief en eenvoudig reactief transport. Voor effectieve monitoring blijkt het nodig om de reistijdverdeling in 3 dimensies te kennen. De lokaties en vooral de diepte van de waarnemingsfilters dienen zorgvuldig afgestemd te worden op de verblijftijden tussen aardoppervlak en waarnemingspunt, de residuele reistijd van het waarnemingpunt tot de pompput, en op de eventuele omzetting- en retardatieprocessen. Hoe groter de omzettingssnelheid of retardatie, hoe ondieper de monitoring moet zijn om effectief te kunnen signaleren en voorspellen. Voor uiteenlopende doelstellingen en omstandigheden is ondiepe monitoring (tussen maaiveld en 15 m diepte) het meest functioneel. Vooral het ondiep monitoren op de grens van de 10 jaar beschermingszone is geschikt voor het tijdig signaleren en levert bruikbare gegevens voor voorspellingsmodellen. Dieper monitoren op de grens van de beschermingszone, zoals gebruikelijk in veel Nederlandse meetnetten, is ineffectief vanwege de lange verblijftijden. De verontreiniging bereikt eerder de pompput dan de waarnemingsfilters.

6 - Bemonsteren van reactieve sedimenten
Sedimentreactiviteit is één van de grootste onbekenden bij de interpretatie van grondwaterkwaliteitsgegevens en bij de voorspelling van grondwaterkwaliteitsveranderingen bij drinkwaterwinningen. Er is daarom een groeiende behoefte aan specifieke informatiedoelen en richtlijnen voor bemonstering van reactieve sedimenten. Doel van de studie was om meetdoelen te bepalen voor bemonstering van reactieve sedimenten en een aanzet te geven voor data-analyse protocollen voor het gebruik in hydrogeochemische transportmodellen. Daartoe is een informatie-analyse uitgevoerd met behulp van conceptuele modellen voor transport in geochemisch en hydraulisch gelaagde poreuze media. De resultaten tonen aan dat verschillende informatiedoelen nodig zijn voor de bemonstering in freatische waterwingebieden en in diepinfiltratie-systemen, vanwege de verschillende verplaatsingsrichting van de reactiefronten ten opzichte van de richting van geochemische variatie. Reactiefronten in freatische waterwingebieden verplaatsen zich verticaal door de overheersende horizontaal gelaagde pakketten. In die situatie voldoet de bepaling van het gemiddelde gehalte van de reactieve component in een reactieve laag. Reactiefronten in diepinfiltratie-systemen verplaatsen zich horizontaal door de horizontaal gelaagde pakketten. In dat geval is informatie over de verticale variaties in sedimentreactiviteit nodig voor de voorspellingsberekeningen. Deze informatie wordt verkregen door percentielen van de frequentieverdeling te bepalen. De verschillende informatiedoelen leiden tot een verschillende bemonsteringsstrategie voor de twee situaties. Uit een case study blijkt dat de onzekerheid van de geschatte reactieve eigenschappen direct doorwerkt in de onzekerheid van de berekende doorbraak van stoffen in de pompput. Een goed ontworpen bemonsteringsstrategie voor sedimentreactiviteit, waarin expliciet rekening wordt gehouden met informatiedoelen en reactieve transportprocessen, helpt dus om de onzekerheid van voorspellingsberekeningen te verminderen.

Algemene conclusies
De studie heeft aangetoond dat een aanzienlijke verbetering van de meetneteffectiviteit kan worden bereikt door hydrologische en hydrogeochemische informatie en concepten van advectief en reactief transport te integreren in het ontwerp, de data-analyse en de evaluatie en optimalisatie van grondwaterkwaliteitsmeetnetten. Een effectieve meetstrategie vraagt om scherp geformuleerde (statistische) meetdoelen en begrip van de hydrologische en hydrogeochemische systeemeigenschappen. De belangrijkste systeemeigenschappen voor de monitoring van diffuse verontreinigingen zijn: (1) de 3D reistijdenverdeling in de aquifer, en (2) de reactiviteit van de geïntroduceerde opgeloste stoffen en de doorstroomde sedimenten. De evaluatie van reistijden en hydrogeochemische berekeningen wordt bij voorkeur geïntegreerd in reguliere, statistische methoden die worden toegepast om de beleidsmatige meetdoelen en specifieke informatiedoelen te realiseren. De meetnetten winnen aan effectiviteit wanneer specifieke meetdoelen en data-analyse protocollen voorafgaand aan installatie en exploitatie van de meetnetten worden vastgesteld.

Dankwoord

De totstandkoming van dit proefschrift was niet mogelijk geweest zonder de ondersteuning en vriendschap van veel vakgenoten, collega's, vrienden en familie. Een aantal van hen wil ik graag persoonlijk bedanken.

Mijn promotoren, professor Peter Burrough en professor Chris Spiers wil ik bedanken voor het geschonken vertrouwen en de deskundige wijze waarop zij me hebben begeleid. Het kostte hen enige overtuigingskracht om het aantal geplande hoofdstukken telkens met één terug te brengen. Chris' typisch Engelse humor was een voortdurende steun in moeilijke tijden, getuige zijn opbeurende woorden: 'If you've seen one monitoring network, you've seen them all'. Dit enigszins tegen het zere been van Frans van Geer, mijn co-promotor en gewaardeerde collega bij NITG-TNO. Frans' analytische geest en statistische kennis en intuïtie hebben mij gescherpt in de meetnetologie. Daarnaast delen we de liefde voor het Wad en het zeilen, al bleek ik te ongedurig voor het admiraalzeilen (zeilen met de handrem erop). Jasper Griffioen is mijn tweede co-promotor. Zonder Jasper's aanmoedigingen was dit proefschrift er wellicht nooit gekomen, en was ik teveel onder de indruk gebleven van Dick Lyklema's woorden 'op meetnetten kun je niet promoveren'. Jasper overtuigde me van het tegendeel en met zijn opbouwend bedoelde kritiek en de discussies over advectief en reactief transport heeft hij erg geholpen. Jasper en Frans, dank voor het vele lees- en discussieerwerk.

De leden van de leescommissie, te weten Prof. van den Akker, Prof. van der Zee, Prof. van Cappellen, dr. Bronswijk en dr. van Gaans, ben ik zeer erkentelijk voor de aandacht die zijn hebben besteed aan het manuscript.

Hoewel onbewust, werd de kiem voor dit proefschrift al in 1989 gelegd bij het schrijven van een projectvoorstel voor het ontwerp van het meetnet grondwaterkwaliteit voor de provincie Noord-Brabant. De eerste bijeenkomst van de begeleidingscommissie zal ik niet snel vergeten vanwege de levendige discussie met Wil van Duijvenbooden, de grondlegger van het onvolprezen, en in internationaal licht bezien unieke, Nederlandse meetnet voor de grondwaterkwaliteit. Sindsdien is monitoring voor mij met tussenpozen steeds een onderwerp van studie gebleven. Het proefschrift bouwt voort op studies die zijn uitgevoerd voor provincies en waterleidingbedrijven. Veel mensen hebben daar direct en indirect aan bijgedragen. Vooral bij de provincie Noord-Brabant heb ik ondertussen al heel wat ambitieuze meetnetbeheerders 'versleten'. Zo hebben bijvoorbeeld Henk van Zoelen, Ingrid van Tiel, Rob Ruytenberg en Corinne Geujen elk hun enthousiaste bijdrage geleverd aan een goed functionerend meetnet. Ik waardeer ook de bijdragen van Wil van Duijvenbooden en Hans Bronswijk van het RIVM aan het ontwerp, de data analyse en de evaluatie van het Brabantse meetnet. Cees Meinardi wil ik bedanken voor het gebruik van de tritiumgegevens. In de provincie Drenthe was het een genot samen te werken met Thea Harmelink, Anton Dries, Erik Blom en Nico van der Moot. Paul Torfs en Hans Bronswijk wil ik bedanken voor hun opbouwende commentaar op de meetnet-evaluatiemethodiek. Het hydrogeochemisch onderzoek bij de waterwinplaats Oostrum (Noord-Limburg) zou zonder de enthousiaste inbreng van Chris Janssen en Joyce Wakker van de WML niet hebben kunnen plaatsvinden.

Het proefschrift verraadt ongetwijfeld de invloeden van de Amsterdamse school in de hydrologie, die met veel bezieling is opgezet door mijn afstudeerbegeleider prof. dr. G.B. Engelen. Het integreren van informatie en het streven naar 'convergentie van bewijs' heb ik vooral van hem geleerd. Paul Baggelaar zette me op het spoor van de non-parametrische statistiek. Paul, bedankt voor de open en hartelijke gedachtenwisselingen. Edzer Pebesma, Giuseppe Frapporti, Paulien van Gaans en Simon Vriend wil ik bedanken voor de open discussies over meetdoelen en data-analysemethoden in de regionale meetnetten. Door jullie heb ik ingezien dat verschillen in benadering uiteindelijk meestal zijn terug te voeren op (vaak impliciete) verschillen in meetdoelstellingen. Pieter Stuyfzand en Tony Appelo wil ik bedanken voor discussies over hydrogeochemie in de breedste zin. 'Many joyful insights'; zo'n nieuwjaarswens tekent Tony.

Mijn ICHU-collega's Hanneke Verweij, Jasper Griffioen, Gualbert Oude Essink, Thom Bogaard, Derk-Jan Karssenberg en Martin Hendriks wil ik bedanken voor de ongelooflijk goede teamspirit bij de opzet van het onderwijsprogramma hydrologie. Aan de traditie van ICHU-afscheidsetentjes komt nu een roemloos einde. Hanneke was voor 4 jaar mijn gewaardeerde kamergenoot en de gezamenlijke opzet van het Ouvèze-veldwerk was één van de leukste werkzaamheden van de afgelopen jaren. Gu, bedankt voor je collegialiteit, de modeldiscussies en de prettige koffietijd. Martin, ik zal met weemoed terugdenken aan de terrasjes in Vaison La Romaine en de wijnhellingen onder de Mont Ventoux. Bedankt ook voor je commentaar op eerdere versies van het manuscript. Thom, ik heb genoten van de veldpractica rond Ossendrecht. Op één of andere manier lukte het ons om ook de studenten ervan te overtuigen dat hydrologie een leuk vak is. Je steun in promotieland heeft zeker 3 maanden tijdwinst opgeleverd. Margriet Ganzeveld en Margot Stoete van Kartlab verdienen alle lof voor de fraaie opmaak van het boekje. Let wel, de kleur oranje hebben zij niet verzonnen.

Veel collega's bij TNO-NITG hebben op directe of indirecte wijze bijgedragen aan onderdelen van dit proefschrift. In enigszins chronologische volgorde gaat mijn dank uit naar Wil Ewalts, Wiel Senden, Aris Lourens, Roelof Stuurman, John Lambert, Koos Uil, Judith Vlot, Alice Buijs, Jan Willem Foppen, Gerrit Jousma, Rolf Hetterschijt, Chris te Stroet, Judith Peeters, Marielle van Vliet, Peter Venema, Bas van der Grift, Monique van der Aa en Ruben Busink. Marielle, Judith, Peter en Ruben hebben veel bijgedragen aan de data-analyse en evaluatie van het meetnet in Noord-Brabant. Alice heeft veel van het veldwerk georganiseerd en een fikse bijdrage geleverd aan het uitwerken van de gegevens. Bas' inbreng was onontbeerlijk bij de concentratie-diepte prognoses. Roelof was altijd goed voor (on)gegrond commentaar en suggesties voor nieuwe invalshoeken. Arco van Vught en Bart van der Eijnden droegen als stagiairs bij met hun stoftransportmodel rond de winning Noordbargeres.

Mijn ouders wil ik bedanken voor de ruimte die ze me hebben gegeven om te spelen en te leren. Stenen zoeken op de Bussumsche hei, zeilen op Loosdrecht en Bach en Händel in het Concertgebouw waren bouwstenen, waarvan de afgeleide activiteiten nog altijd anker- en rustpunten zijn. Marieke, bedankt voor je steun en liefde, ook als het soms even moeilijk was. Je bent nu wel vaak genoeg alleen heen en weer gehopt tussen Stellendam en Lowestoft en andere mij onbekende oost-Engelse oorden. Ik moest ook maar eens mee oversteken.

Een leuke en leerzame periode is afgesloten.
And now for something completely different.

Curriculum Vitae

Hans Peter Broers is op 16 juni 1963 geboren in Naarden. In 1981 deed hij eindexamen VWO aan het Chr. College Stad en Lande te Huizen. In 1984 behaalde hij zijn kandidaatsexamen Fysische Geografie aan de Faculteit Aardwetenschappen van de Vrije Universiteit in Amsterdam. In 1988 studeerde hij af in de richting Geografische Hydrologie en Hydrogeologie, met een bijvak Hydrogeochemie aan de Faculteit Aardwetenschappen van de Universiteit Utrecht.

Zijn hobby's zijn (wedstrijd)zeilen en bergwandelen. Hij is gediplomeerd zeilwedstrijdtrainer. In 1988 werd hij samen met zijn broer Arjan Nederlands kampioen in de internationale Vaurienklasse.

Sinds 1988 werkt hij als geohydroloog/hydrogeochemicus bij TNO. Daar verricht hij onderzoeks- en advieswerk op het raakvlak van grondwaterstroming en grondwaterkwaliteit. Hij hield zich ondermeer bezig met grondwatermonitoring, verontreinigingshydrologie, hydrogeochemisch veld- en laboratoriumonderzoek en geo-informatiesystemen. Hij maakte deel uit van de redactie van het hydrologisch vakblad Stromingen en was lid van het kennisintegratieteam van NOBIS (het Nederlands Onderzoeksprogramma Biotechnologische In-Situ Sanering).

Tussen 1997 en 2002 was hij tevens in deeltijd werkzaam als docent/onderzoeker Hydrogeologie bij het Interfacultair Centrum Hydrologie Utrecht (ICHU), een samenwerkingsverband tussen de Faculteiten Ruimtelijke Wetenschappen en Aardwetenschappen van de Universiteit Utrecht.