RIVM report 703717013/2005

Reliability of travel times to groundwater abstraction wells: Application of the Netherlands Groundwater Model, LGM

K. Kovar[a], A. Leijnse[b], G.J.M. Uffink[a], M.J.H. Pastoors[a], J.H.C. Mülschlegel[a] and W.J. Zaadnoordijk[c]

Corresponding author: Karel Kovar
Agriculture and Rural Areas
karel.kovar@rivm.nl

a) RIVM, Bilthoven

b) TNO-NITG, Utrecht

c) Royal Haskoning, Rotterdam

Het rapport in het kort

Betrouwbaarheid van verblijftijden naar grondwaterwinningen; toepassing van Landelijk Grondwater Model – LGM.

Dit rapport beschrijft de ontwikkeling en een pilot-toepassing van een modelleringsmethode ten behoeve van de bepaling van betrouwbaarheid van verblijftijden van grondwater dat naar grondwateronttrekkingen stroomt. De methode, die van de eindige-elementechniek gebruik maakt, is opgenomen in het rekenmodule LGMLUC, een aanvullende module van het Landelijk Grondwatermodel LGM, van het RIVM. De betrouwbaarheid wordt voorgesteld als een band (zone) rondom de verwachtingswaarde van een verblijftijd-isochrone, bijvoorbeeld 25 jaar. De breedte van deze band voor een zekere waarschijnlijkheid van voorkomen (bijvoorbeeld tussen de 97,5 en 2,5 percentielwaarden) neemt toe met een toenemende onzekerheid van modelinvoer parameters. Gebruik is gemaakt van de First-Order Second-Moment (FOSM) methode voor de analyse van de voortplanting van fouten. De resultaten van de FOSM methode zijn vergeleken met die van de Monte Carlo aanpak voor een LGM-model en als een onafhankelijke test een TRIWACO-model. Uit deze vergelijking is geconcludeerd dat de FOSM-methode adekwaat en rekentechnisch effectief is voor het analyseren van de betrouwbaarheid van verblijftijden. Aangenomen is dat de kansdichtheidsverdeling van verblijftijden lognormaal verdeeld is. De methode houdt rekening met de onzekerheid in een aantal modelinvoer parameters, zijnde de factoren die de onzekerheid in verblijftijden tot gevolg hebben. De onzekerheid van de parameters is bepaald door middel van calibratie (invers model) en expert-judgement. De toepasbaarheid van de ontwikkelde methode is aan de hand van een pilot-studie getoond, gebruikmakend van het binnen het LGM bestaande deelmodel Utrecht. De methode kan bij verschillende dichtheid van eindige-elementengrid worden gebruikt, zowel voor problemen op lokale schaal (hoge griddichtheid) als op regionale schaal. De informatie over de betrouwbaarheid van verblijftijden kan worden benut voor beleidsmatige beslissingen, zoals bij onderzoek naar risico's binnen de bestaande grondwaterbeschermingsgebieden.

Trefwoorden: grondwater; verblijftijden; betrouwbaarheid; FOSM; Monte Carlo

Abstract

Reliability of travel times to groundwater abstraction wells: Application of the Netherlands Groundwater Model, LGM

A modelling approach was developed, incorporated in the finite-element method based program LGMLUC, making it possible to determine the reliability of travel times of groundwater flowing to groundwater abstraction sites. The reliability is seen here as a band (zone) around the expected travel-time isochrone, with the width of this probability occurrence zone (e.g. between 97.5 and 2.5 percentile values) increasing with increasing uncertainty in the model-input parameters. The modelling approach is based on the First-Order Second-Moment (FOSM) method for uncertainty propagation analysis. The FOSM results have been compared with a Monte Carlo analysis for an LGM model and another, LGM-independent model. From the match between FOSM and Monte Carlo results it was concluded that the FOSM approach is an adequate and computationally effective method to analyse the uncertainty of travel times. The travel time was assumed to be log-normally distributed. The uncertainty in several model-input parameters was accounted for as a factor influencing the reliability of travel times. These are aquifer transmissivity, hydraulic resistance of aquitards, drainage and infiltration resistance defining the head-flux relationship between the top aquifer and surface waters, effective porosity of aquifers and aquitards, and groundwater recharge rate. All these model-input parameters are assumed to be log-normally distributed. The uncertainty in these parameters was derived using inverse-method calibration and expert judgement. The applicability of the modelling approach was demonstrated for a pilot study area in the central part of the Netherlands. The approach can be used for problems with various spatial resolutions of the finite-element grid, problems ranging from local (high grid density) to regional. The information on the travel-time reliability can be used for policy-related decision-making, such as the examination of risks within the existing groundwater protection zones.

Keywords: groundwater; travel times; reliability; FOSM; Monte Carlo

Contents

Summary

In this study, aimed at setting up a methodology for determining travel time reliability, a comparison was made of two methods widely used for an uncertainty propagation analysis: (1) the Monte Carlo method (MC) and (2) the First-Order Second-Moment approach (FOSM). Monte Carlo is a direct, but computationally intensive method, while in the FOSM approach, the model calculates rapidly, but produces approximate results. The basic limitation of FOSM is that it only gives the mean and variance of the travel times, while MC produces the entire probability density function.

The FOSM approach was selected for subsequent development of a modelling tool to determine the travel time reliability resulting from the uncertainty in various model-input parameters. The choice for FOSM was justified in the comparison of the FOSM results with the Monte Carlo results, Monte Carlo representing the formally correct approach for uncertainty propagation analysis. A comparison has been made with a module of the Netherlands Groundwater model (LGM, "Landelijk Grondwater Model" in Dutch). In addition, an independent comparison of FOSM and MC results has been carried out. The independent results were produced with TrCalCon, the TRIWACO module for calibration and confidence analysis. The FOSM approach was found to produce results that agreed well with results of the Monte Carlo method. From the match between FOSM and Monte Carlo results, it can be concluded that the FOSM approach is a reliable and computationally effective method to analyse the uncertainty of travel times.

The programs, LGMLUC and LGMMUC, were developed as modules of the Netherlands Groundwater Model ("Landelijk Grondwater Model" in Dutch). These programs can be used to determine the uncertainty in groundwater travel times as described below:
− LGMLUC is based on the FOSM approach and assumes (as a user option) the travel time to be either normally or log-normally distributed. In this study we used only the log-normal option. The output consists of travel times for a lower percentile (e.g. 2.5%) and for an upper percentile (e.g. 97.5%). Both travel-time output sets are values at all nodes of a finite element grid and, thus, can be used for contouring isochrones. LGMLUC was used in the pilot-study application to determine the reliability of travel times for a number of groundwater abstraction sites in the Utrecht Model in the central part of the Netherlands (Chapter 3);
− LGMMUC uses Monte Carlo simulations to determine the distribution of travel time in a number of user-selected locations. The output is a table (file) with travel times for each selected location, one travel time for each Monte Carlo realization. Subsequently, the tables can be used to prepare a histogram and, based on the histogram, to show the travel times for any percentile (e.g. 2.5% and 97.5%). In view of the nature of the LGMMUC program itself, one does not have to make any assumption about the probability distribution of travel times in its use. This program was used for comparing the Monte Carlo and FOSM approaches (Chapter 2).

The uncertainty in the following parameters can be accounted for as a factor influencing the reliability of travel time: aquifer transmissivity, hydraulic resistance of aquitards, drainage and

infiltration resistance defining the head-flux relationship between the top aquifer and surface waters, effective porosity of aquifers, effective porosity of aquitards and groundwater recharge rate. LGMLUC and LGMMUC inputs are: (a) the "expected" parameter values and (b) the covariance matrix, or standard deviations and the correlation matrix. Depending on their origin, two groups of parameters can be distinguished:

− parameters determined by the inverse-method calibration using the LGM module LGMCAL. LGMCAL produces optimal ("expected") parameter values and their covariance matrix. The relevant parameters are aquifer transmissivity, hydraulic resistance of aquitards, and drainage and infiltration resistance defining the head-flux relationship between the top aquifer and surface waters;

− parameters not defined by means of calibration – the parameters whose variance was assumed by expert judgement. The parameters are: effective porosity of aquifers, effective porosity of aquitards and the groundwater recharge rate. In fact, effective porosity of aquifers and aquitards cannot even be calibrated for by using observed groundwater heads.

All uncertain parameters listed above are assumed to be log-normally distributed.

With the aim of demonstrating its applicability in practice, the FOSM approach, as implemented in the LGMLUC module, was applied to determine the reliability of travel times for 10 groundwater abstraction sites located in the pilot study Utrecht Model, in the central part of the Netherlands. In addition to the "expected" travel-time isochrone (e.g. 50 years), which could be generated by any deterministic approach, we produced two more travel-time isochrones, referred to as the "inner" and "outer" isochrones. These two isochrones are located on either side of the "expected" travel-time isochrone. To illustrate this principle, let us assume that the "inner" and "outer" isochrones refer to the probability of 0.975 and 0.025, respectively. With these two isochrones in the same plot, the total capture area is divided into three zones, listed in the order of increasing distance from the well location:

− A zone around the well up to the "inner" 50-years isochrone, representing the area where the probability that a water particle will reach the well within 50 years is higher than 0.975.

− A transition zone, where the probability of reaching the well within 50 years gradually decreases from 0.975 to 0.025, going from the "inner" to the "outer" isochrone.

− A zone outside the "outer" 50-years isochrone, in which a particle starting here reaches the well within 50 years with a probability of less than 0.025.

The width of the transition zone increases with increasing uncertainty in the model-input parameters, in other words, with variance in these parameters.

Summarizing, the modelling approach developed here on the basis of the FOSM method makes it possible to determine the reliability of travel times to groundwater abstractions. Since the procedure is computationally efficient, it can be used for problems at various spatial resolutions of the finite-element grid. These range from local problems with a high grid density to problems on a regional scale similar to the model area in this study.

Once the information on the travel-time reliability has become available it can be used for policy-related decision-making. One possible application is the examination of risks within the existing

groundwater protection zones. Knowing the travel-time reliability, one can better evaluate the vulnerable areas (potential sources of pollution), located both inside and outside the existing groundwater protection zones. Another way travel-time reliability could be applied is in decision-making with regard to space allocation; here claims are made by the various users (target sectors) on the same space. Groundwater-based drinking-water production is one of the users claiming space for groundwater protection zones.

1.   Introduction

Groundwater is an important source of public water supply. Sixty per cent of the drinking water in the Netherlands originates from abstracted groundwater, the total annual amount of abstracted groundwater being about 800 million m^3. Groundwater quality is endangered by various anthropogenic substances entering the subsurface. Groundwater has to be protected to prevent the pollution that, in turn, can cause public health risks. The pollution implies additional purification during drinking-water production or sometimes even the closure of groundwater abstraction sites.

The current groundwater protection policy is based on environmental considerations, such as the regulations for soil protection. In addition to the generally applicable environmental legislation, which also includes groundwater, specific protection is provided for groundwater abstractions. With regard to these abstractions, several types of spatial zones are distinguished in the Netherlands, as listed below in decreasing order of spatial extent:
–   the *capture zone* ("intrekgebied" in Dutch): the entire area from which the abstracted groundwater originates;
–   the *groundwater protection zone* ("grondwaterbeschermingsgebied" in Dutch): usually delineated using the 25-year travel time contour.;
–   the *microbiological protection zone* ("waterwingebied" in Dutch): delineated using the 60-day travel time contour, where it is assumed that a 60-day travel time is sufficient for die-off of pathogenic microorganisms in contaminated groundwater to the extent that health risks have been eliminated.
–   the first protective area around the wellhead (well-screen location), up to 30 metres from the wellhead.

The current groundwater protection policy has, in the past, yielded important results. However, this policy will have to be re-evaluated to achieve effective and long-term sustainable protection of groundwater as a source used in drinking-water production. To this end, policy measures, for example, might be considered to stimulate the decrease in load (leaching) of pollutants to the groundwater system or, in other words, to decrease the risk of groundwater pollution.

Another change in the groundwater protection policy will be introduced with the implementation of the Water Framework Directive (WFD), with specific reference to the delineation of bodies of groundwater. WFD requires each EU member state to evaluate the status of its groundwater quality and provides measures designed to maintain and, if needed, to improve the groundwater quality. The policy measures valid for the groundwater protection zones are important, especially since the EU also requires improvement in the quality of water at the source –i.e. input to groundwater at land surface. The intended gain is a decrease of effort needed for drinking-water purification.

The total area of the groundwater protection zones in the Netherlands measures about 1400 km^2, being about 4% of the entire surface area of the country. At the moment there are major claims being made on space in the Netherlands, with the consequence that conflicts will arise where claims are made on the same space by various users (target sectors). Indicating the spatial needs required

for drinking-water production in an early stage creates clarity during the space-allocation decision-making process. We distinguish here between the space that is directly needed for production facilities (buildings), and reservoirs and pipelines, and the space that is indirectly needed, such as the groundwater protection zones. One should realize that in a groundwater protection zone, limitations are imposed on certain activities, with the aim of avoiding or minimizing the risk of groundwater pollution. For example, specific agricultural practices have been prescribed to avoid pollution by pesticides and to minimize the nitrogen losses to groundwater and surface waters.

This study is carried out by order of the VROM Directorate General for Environmental Protection (Directorate for Soil, Water and Countryside) as part of project 703717. The title of the project is "Duurzaamheid bronnen drinkwatervoorziening" (Sustainability of water-sources for drinking-water supply). The information on the groundwater system provides the basis for developing the groundwater protection policy and the resulting legislation. The key building blocks for the design, implementation and enforcement of policy measures are (a) the location of the capture zone, (b) the location of the groundwater protection zone and (c) the spatial contour pattern of travel times. In dealing with risks we also need information on (d) the reliability of the capture-zone border and (e) the **reliability of the travel time contours**. Instead of using the term "reliability" here, we might also refer to "certainty", "uncertainty" or "probability of occurrence".

The reliability of travel times can be depicted as a band between two travel-time contours, each band indicating the probabilities of travel-time occurrence, for example, between the 97.5 and 2.5 percentiles. The contours and the reliability band can be used for risk-based decision-making. By way of an example, an activity within the reliability band, such as a production facility, could lead to leakage of a chemical substance to groundwater. Re-allocating this facility is possible but would involve considerable costs. Risked-based decision-making implies here that a decision on a re-allocation is based not only on the occurrence probability of spill-caused pollution but, also probabilistically, on the level of threat expressed in terms of travel time to the abstraction well.

The study presented in this report was aimed at:
−   setting up a methodology for determining the travel time reliability, and
−   testing the applicability of the methodology set up for a number of groundwater abstraction sites.
The modelling exercise was carried out for a pilot study area in the central part of the Netherlands.

The study is described in Chapters 2 – 4 as outlined below:
♦   Chapter 2 (uncertainty analysis of travel times) covers the selection of the methodology for calculating travel-time reliability. The choice for the FOSM approach (First-Order Second-Moment) is justified by comparing the FOSM results with the results obtained by a Monte Carlo approach, which is seen as the formally correct representation of reality.
♦   Chapter 3 (application of the FOSM approach in the pilot study) documents the application of the FOSM method to determine the reliability of travel times for a number of groundwater abstraction sites in the pilot study model Utrecht, located in the central part of the Netherlands.
♦   Chapter 4 lists the conclusions.

2.   Uncertainty analysis of travel times

2.1   Introduction

Traditionally, groundwater modelling exercises are classified in two categories: simulations of piezometric heads (flow models) and simulations of solute concentrations (solute transport models). The streamline and travel time analysis that is applied in this report falls somewhere in between these categories. Alongside with the word "streamline" we also use the term "pathline" in this report. Since we deal with groundwater systems in steady state, the meaning of both words is identical. Usually, streamlines and travel times are derived by a particle tracking technique under the assumption that the contaminant behaves conservatively and is transported only by advection. Despite such a simple concept, streamline and travel time analysis gives a good first impression of contaminant transport. A practice where travel time analysis is widely applied is the delineation of capture zones around groundwater pumping stations. Understanding the uncertainty of travel times is important for the design of protection zones around public drinking water supply wells.

The main objective of the present study is to set up a methodology to determine travel time uncertainty and test the proposed method for a number of pumping stations in the area Utrecht, in the Netherlands. In general, uncertainty may originate from different sources: aquifer heterogeneity, variations in groundwater recharge rates, dispersion, etc. Several of these types of uncertainty have been addressed to by other authors, e.g. Franzetti & Guadagnini (1996), Guadagnini & Franzetti (1999), Elfeki (1996), Uffink (1989), Neupauer & Wilson (2003), Hendricks Franssen *et al.* (2004), Varljen & Shafer (1991), Van Leeuwen (1997), and LaVenue *et al.* (1989). In the present report we focus on travel time uncertainty due to the parameter uncertainty resulting from calibration exercises. The following two methods for an uncertainty propagation analysis are both widely used: (1) the Monte Carlo method (MC) and (2) the First-Order Second-Moment approach (FOSM). Monte Carlo is a direct, but computationally intensive method, while the First-Order Second- Moment approach is fast but gives an approximation. The basic limitation of FOSM is that it only gives the mean and variance of the travel times, while MC gives the entire probability density function (*pdf*).

In chapter 3 of this report the FOSM method is applied to determine the reliability of travel time zones for a number of pumping stations in the Netherlands. The present chapter is dedicated to a justification of our choice for FOSM. This is achieved by comparing the FOSM results with results obtained by a Monte Carlo approach, first for a simple hydrological problem and secondly for an existing pumping station (Lochem) in the eastern part of the Netherlands. Prior to a discussion of the test case results we describe our methodology and briefly summarise the main features of the Monte Carlo and First-Order Second-Moment methods.

2.2   Methodology

Inverse model

Symbolically, a flow model may be written as:

$$\varphi = F(p_1, p_2, ..., p_m),$$ (2.1)

where φ or $\varphi(x, y, z)$ is the spatially distributed groundwater head and $p_1, p_2, ..., p_m$ denotes the set of model parameters. These parameters may be hydraulic conductivity's, storativity's, resistances of clay-layers, etc. The flow model (2.1) does not provide the travel times. For the determination of travel times (2.1) is only relevant as an inverse model, i.e. to obtain the set of model parameters. These model parameters are needed to calculate travel times. The inverse model gives information on the model parameters when a set of head measurements is available. In general the measured heads will differ from the (theoretical) model outcome. In inverse modelling we search the optimal values for the parameter set by minimising the object or penalty function U:

$$U = \sum_{i=1}^{k} (\varphi_i - \hat{\varphi}_i)^2$$

where $\hat{\varphi}_i$ denotes the measured head and φ_i the calculated head at location i. Usually, the parameters determined by the inverse model are a subset of the total set of parameters. For instance, some parameters may be difficult to obtain due to insensitivity to the measurements, while others may be known from other sources.

Inverse modelling is well documented (e.g. Sun, 1994). Symbolically, an inverse model may be expressed as:

$$\begin{pmatrix} p_1 \\ ... \\ p_n \end{pmatrix} = F^{-1}(\hat{\varphi}_1, \hat{\varphi}_2, ..., \hat{\varphi}_k)$$ (2.2)

The primary result is a set of optimal model parameters $p_1, p_2, ..., p_n$ or $\left[\mu_{p_1}, \mu_{p_1}, ..., \mu_{p_n} \right]$, where $n \leq m$. In addition the model gives information on the uncertainty of the parameters. The most complete characterization of parameter uncertainty is given by the (joint) probability density distribution (*pdf*). A less complete, but widely used uncertainty characterisation is given by the statistical moments of the *pdf*. Usually, only the first and second moments (mean and covariance) are used.

Streamline model

Determination of streamlines and travel times T towards a pumping well can be expressed as a model S:

$$T = S(p_1, p_2, ..., p_n) \tag{2.3}$$

Here, the travel time T is a spatially distributed variable depending, in principle, on the x, y and z coordinates of the starting point. It is tacitly understood that the streamline always starts at the phreatic surface. Since the elevation z of the phreatic surface is a function of x and y, we may leave out the z-coordinate and write $T(x,y)$. After the inverse model (2.2) is run, we may obtain travel times using (2.3). Since the (calibrated) model parameters are uncertain, model (2.2) represents a stochastic model and, therefore, produces a stochastic outcome variable T. Our purpose is to determine how the uncertainty in T is related to the uncertainty in the set of parameters.

In the following sections two methods for uncertainty propagation analysis are briefly discussed:

2.3 Monte Carlo

The Monte Carlo method is a direct and straightforward technique that has proved to be successful in various fields of science, especially since the introduction of fast computers. The first step consists of a sampling procedure in order to obtain sets of model parameters from a (multivariate) distribution. Secondly, a stream-line/travel-time model is run repeatedly for all sets parameters resulting in a large number of model outcomes (travel times). Finally, all outcomes are collected to obtain a frequency distribution of T. Of all methods for uncertainty propagation analysis, the Monte Carlo Method is potentially the most accurate. The major drawback is that the method is computationally intensive and time consuming. The accuracy depends on the number of runs. A reasonable accuracy often requires a high number of runs. The required number of runs depends mainly on the number of uncertain model parameters and the non-linearity of the system.

Percentiles

In the Monte Carlo approach the distribution of input and output variables can be of any form. Therefore, it is convenient to use percentiles to characterise the distribution in a concise way. Percentiles can be obtained from a cumulative frequency distribution of the output data. In the present report we frequently use the 2.5 and 97.5 percentiles for the extremes of the distribution. Further, we assume that the 50 percentile represents the optimal (or average) value. The percentiles are obtained as follows. In the horizontal plane a rectangular grid is defined that covers at least the expected capture zone of the well. The grid-points serve as starting points for a particle-tracking routine that yields the travel times towards the well. Each run gives travel times for every grid-point, as long as the grid-point is within the capture zone. After N runs the data are collected and a (cumulative) histogram is constructed and percentiles are obtained for each grid-point. In principle,

some points at the edge of the grid may fall outside the capture zone and a travel time can not be obtained. When no travel times are found in a substantial part of the Monte Carlo runs the point is disregarded.

Inner, outer and expected travel times and isochrones

With the percentiles for all grid-points a spatial distribution is obtained for $T_{2.5}(x,y), T_{50}(x,y)$ and $T_{97.5}(x,y)$. In this report we use the terms "outer", "expected" and "inner" travel-time to distinguish between these outcomes, so:

Outer: $T_{2.5}(x,y)$

Expected: $T_{50}(x,y)$ (2.4)

Inner: $T_{97.5}(x,y)$

Figures 2.1 and 2.2 illustrate the significance of the percentiles in the construction of probabilistic travel time zones (e.g. a 50 years-zone). Figure 2.1(upper) shows the "inner", "expected" and "outer" travel times as a function of the distance r to the well. The intersection of these graphs with the line T=50 marks the points r_a, r_b and r_c, which represent the positions of the "inner", "expected" and "outer" 50-years isochrone. Figure 2.1(lower) displays how the intervals between r_a, r_b and r_c may be interpreted in terms of probabilities to reach the well within 50 years. The interval (green) corresponds to positions from where a particle may reach the well within 50 years with a probability of 0.975 or higher. The yellow coloured interval, $r_a < r < \infty$, indicates points from where the well is reached in 50 years with a probability of 0.025 or less. The orange zone in between contains points from where the well is reached in 50 years with a probability that decreases gradually from 0.0975 to 0.025. From point r_b the well is reached in 50 years with a 0.5 probability. The orange zone marks the possible location of the 50-years capture zone with a reliability of 95%. Figure 2.2 gives a sketch of the situation in the horizontal plane. By contouring, the 50-years isochrone for "inner" travel time, $T_{97.5}(x,y)$, "expected" travel time $T_{50}(x,y)$, and "outer" travel time $T_{2.5}(x,y)$ can be plotted. Note that in this figure the boundary of the total capture zone is indicated although in principle this location is uncertain and varies when different combinations of parameters. The boundary in the picture may be interpreted as the "expected" capture zone. The "inner" and "outer" isochrones now divide the capture area into three zones:

1) A zone around the well bounded by the "inner" 50 year-isochrone (green). It represents the area where the probability that a water particle reaches the well within 50 years is higher than 0.975.

2) A zone outside of the "outer" 50-years isochrone (yellow). A particle starting in this zone reaches the well within 50 years with a probability less than 0.025. An alternative formulation is that with a probability of 0.975 or higher the particle does **not** reach the well within 50 years. Strictly speaking, the outer limit of the yellow zone is not the expected capture-zone boundary, but rather the outer occurrence probability contour (not shown in the picture) of the capture zone, which is bigger than the capture zone.

3) A transition zone (orange) where the probability to reach the well within 50 years is gradually decreasing from 0.975 to 0.025, going from "inner" to "outer" isochrone. Alternatively, one may say that this zone contains the "true" 50-years isochrone with a probability of 0.95. As the picture shows, the zone does contain the "expected" isochrone.

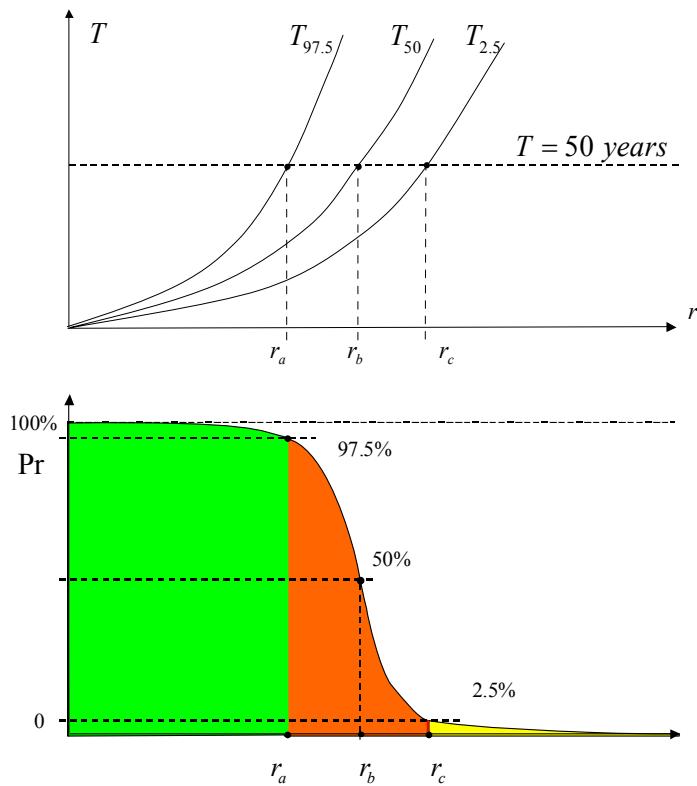*Figure 2.1 (upper) "Inner" ($T_{97.5}$), "outer" ($T_{2.5}$) and "expected" (T_{50}) travel times versus distance to well (r) and position of 50 years isochrones (r_a, r_b, r_c), (lower) Probability distribution to reach the well within 50 years.*
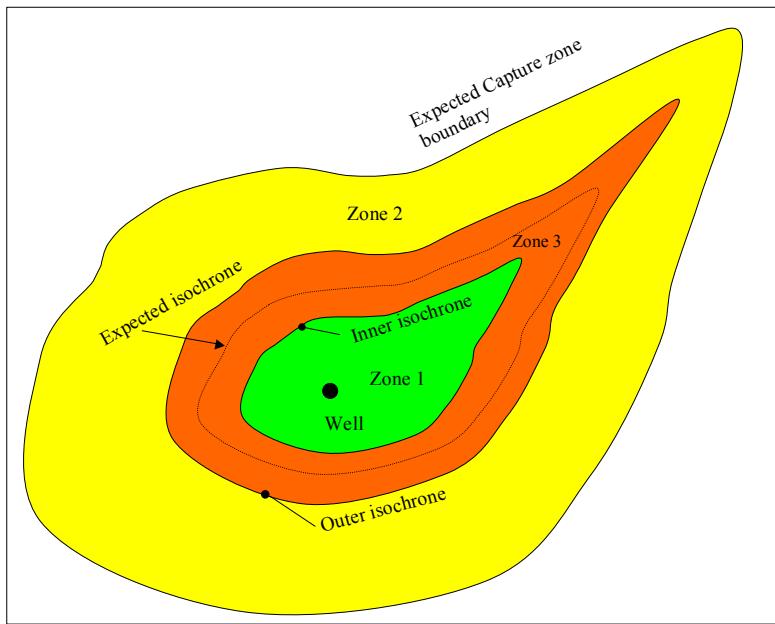
Figure 2.2 Example of capture area with "inner" ($T_{97.5}$),"outer" ($T_{2.5}$) and "expected" (T_{50}) isochrones.

2.4   First-Order Second-Moment

This method has been applied before in several groundwater studies. A general application is described by Dettinger & Wilson (1981). An application focusing on travel times is given by LaVenue *et al.* (1989). The first order method expresses the mean and variance of the model outcomes (travel times) in terms of means and variances of the model parameters. The mean travel time is obtained using the optimal values of the parameters ($\mu_{p_1},...,\mu_{p_m}$):

$$\mu_T(x,y) = S(\mu_{p_1}, \mu_{p_2},...,\mu_{p_n}) \tag{2.5}$$

This value can be obtained by one single run of the streamline model. An expression for variance of the travel times, σ^2_T , can be derived from a Taylor series expansion of S in terms of the model parameters. In a first order approximation the variance σ^2_T can be expressed in the (co)variances of the model parameter:

$$\sigma_T^{\ 2}(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(p_i, p_j)\frac{\partial S}{\partial p_i}\frac{\partial S}{\partial p_j} \tag{2.6}$$

The derivatives $\partial S/\partial p_i$ are called sensitivity derivatives (SD). In our study sensitivity derivatives have been evaluated by numerical differentiation:

$$\frac{\partial S}{\partial p_i} = \frac{S(p_1,...,p_i + \Delta p_i,...,p_n) - S(p_1,...,p_i,...,p_n)}{\Delta p_i}$$

This requires one additional model run for each parameter involved. The differentiation is performed using the mean parameter values $\mu_{p_1},...,\mu_{p_n}$, (except for the parameter we differentiate to). SD's are determined for all spatial points where a travel time is required, leading to a variance that varies in the horizontal plane.

As mentioned earlier, the FOSM method does not give the total *pdf* of the outcome variable and therefore we cannot find percentiles from a cumulative distribution. However, when a Gaussian distribution is assumed, the first and second moments are sufficient to construct the total *pdf* and accordingly to find values corresponding to the 2.5, 50 and 97.5 percentile. Two alternatives are discussed, namely (i) a normal and (ii) a log-normal Gaussian distribution. The choice between normal or log-normal distribution leads to different results.

2.4.1 Normal distribution

A normal distribution (*pdf*) is defined as:

$$p(T) = \frac{1}{\sigma_T \sqrt{2\pi}} \exp\left(-\frac{(T-\mu_T)^2}{2\sigma_T^{\,2}}\right)$$

From the formula it is clear that it is sufficient to know mean and variance. Instead of percentiles, we use a coefficient $\beta = (T-\mu_T)/\sigma_T$. The travel-time defined by a certain value for β will be denoted as T_β, so:

$$T_\beta = \mu_T(x,y) + \beta\sigma_T(x,y)$$

The cumulative distribution $P(\beta)$ as function of β becomes:

$$P(\beta) = \Pr\{T < T_\beta\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta} \exp\left(-\frac{u^2}{2}\right) du$$

The function $P(\beta)$ corresponds to the (pink) area under the curve (Figure 2.3). There is a direct relation between the exceeding probability for T_β, which is equal to $1-P(\beta)$ and the percentiles. For instance, $\beta=1$ defines the value $T_1 = \mu + \sigma$, which is exceeded with a 15.87% probability, $\beta=2$ gives $T_2 = \mu + 2\sigma$, with exceeding probability 2.27% etc. (see Table 2.1). A more extended table is found in Kreyzig (1970). For the test cases we are primarily interested in the following "statistics":

(outer): $T_{-2} = \mu - 2\sigma$

(expected) $T_0 = \mu$ (2.7)

(inner) $T_2 = \mu + 2\sigma$

As can be derived from Table 2.1, these values represent the 2.3, 50 and 97.7 percentiles. These values closely approach the percentiles used in the Monte Carlo method (2.5, 50 and 97.5). Therefore, in FOSM we shall use T_{-2}, T_0 and T_{-2} as estimates for "outer", "expected" and "inner" travel times.
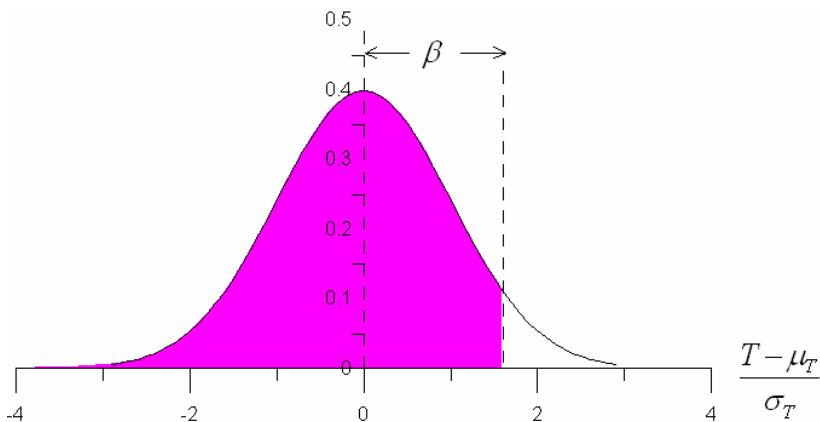
Figure 2.3 Normal distribution.

Table 2.1 Normal distribution, cumulative probabilities as a function of β.

β	$P(\beta)$
-4	0.00003
-3	0.0013
-2	0.0227
-1	0.1587
0	0.5
1	0.8413
2	0.9772
3	0.9986
4	0.999968

2.4.2 Log-normal distribution

Log-normal distributions are quite common in nature and arise in processes where the ratio between values is important instead of absolute differences. In case of log-normality the formulation needs to be slightly modified. When we introduce the mean m and variance s^2 of the log-transformed travel time $log(T)$, $(m = E\{\log(T)\}$ and $s^2 = E\{(\log(T) - m)^2\})$, we may express the log-normal *pdf* $\lambda(T)$ as follows:

$$\lambda(T) = \frac{1}{T\, s\sqrt{2\pi}} \exp\left(-\frac{(\log(T)-m)^2}{2s^2}\right)$$
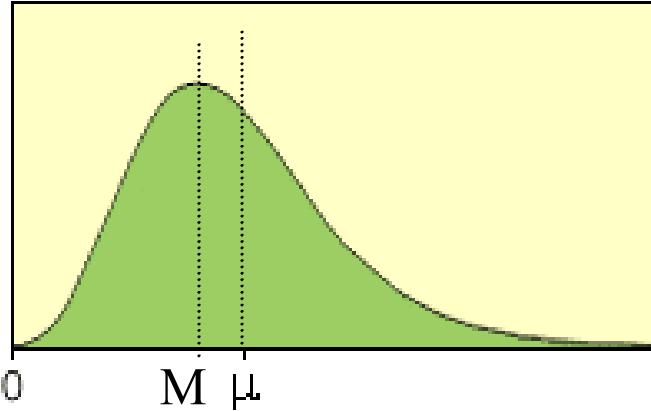
Figure 2.4 Log-normal distribution with median M and arithmetic mean μ.

In fact, we may consider *Y*, the log-transform of *T*, $Y = \log(T)$, as our model output, and write:

$$Y(x,y) = \overline{S}(p_1,...,p_n)$$

where $\overline{S} = \log(S)$ denotes the log-transformed streamline/traveltime model. Since *Y* is normally distributed, *m* and *s* (mean and standard deviation of *Y*) follow from (2.4) and (2.5), after replacing *S* by $\overline{S}$. Then, we obtain:

$$m(x,y) = \overline{S}(\mu_{p_1},\mu_{p_2},...,\mu_{p_n}) = \log\left(S(\mu_{p_1},\mu_{p_2},...,\mu_{p_n})\right) \qquad (2.8)$$

$$s^2(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(p_i,p_j)\frac{\partial\overline{S}}{\partial p_i}\frac{\partial\overline{S}}{\partial p_j} = \frac{1}{T^2}\sum_{i=1}^{n}\sum_{j=1}^{n} Cov(p_i,p_j)\frac{\partial S}{\partial p_i}\frac{\partial S}{\partial p_j} \qquad (2.9)$$

It is clear that with the same data we can compute either μ and σ from (2.5) and (2.6) or *m* and *s* from (2.8) and (2.9), depending on the choice between the normal or log-normal distribution. No additional model runs are required. Note that *T* in (2.9) corresponds to the value that produces the mean of the log-transformed *T*. This is the geometric mean of *T*, which when we use the [10]log can be written as: $\overline{T}_g = 10^m$.

With a similar coefficient β as above ($\beta = (\log T - m)/s$) a cumulative distribution $\Lambda(\beta)$ can be written:

$$\Lambda(\beta) = \Pr\{\log T < [m + \beta s]\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta} \exp\left(-\frac{u^2}{2}\right) du$$

The "statistics" of the log-transformed travel-time for $\beta = -2$, 0 and 2 are conform (2.7):

$$(\log T)_{-2} = m - 2s$$
$$(\log T)_0 = m$$
$$(\log T)_2 = m + 2s$$

or (using 10log) for the travel-time itself:

(outer) $\hat{T}_{-2} = 10^{m-2s}$
(expected) $\hat{T}_0 = 10^m$ (2.10)
(inner) $\hat{T}_{+2} = 10^{m+2s}$

where $\hat{T}_\beta$ is written to distinguish between the values of the normal and log-normal variant. One of the drawbacks of FOSM is that the method is not suited for (highly) non-linear models within the range of possible parameter values or for models with a non-linear correlation between the parameters. FOSM is most accurate around the mean, while in general one is interested in the tail of the curve. A second-order approach (e.g. SOSM, Zaadnoordijk, 2003) is more accurate, but involves evaluation of second-order sensitivity derivatives. Determination of the latter may be complicated and time consuming when it is done numerically. When FOSM is applied, the validity should be checked, for example, by comparing it to the MC method or to a higher order approach. Another drawback is that the method does not give complete information on the travel time distribution. Only its first and second moments are obtained. When a different types of *pdf*'s are assumed (e.g. normal or log-normal), the method yields different results.

Steps required concern the determination of:
- Covariances of the parameters. (In our study obtained by model calibration).
- Sensitivity coefficients. In general SD's need to be determined numerically, which may not be trivial. Additional runs are needed only to determine the sensitivity coefficients, one for each parameter.

2.5 Test cases

We compare the MC and FOSM approach for two cases that we shall refer to as Case 1 and Case 2.

2.5.1 Case 1 (TRIWACO)

This case consists of a well in a phreatic one-aquifer system, where the groundwater flow is assumed to be steady. The model covers a 5000 m wide and 10000 m long area (Figures 2.5 and 2.6) and is limited at the left and right side by a constant head boundary (left $\varphi = 0$, right $\varphi = -1$ m). The upper and lower boundaries (north and south) are impervious (no-flow boundaries). The well discharge $Q = 1000$ m^3 day^{-1}. The precipitation excess at the top has a rate of $N = 1$ mm day^{-1}. Figure 2.5 shows a cross-section, while Figure 2.6 gives a horizontal top-view. At the top of the aquifer a so-called top-system is present. The top-system simulates the interaction between the groundwater and the secondary drainage system of ditches and drains. The model calculates the total recharge to the aquifer system, q_{rch}, from the precipitation excess (or if you prefer that word, groundwater recharge) N and the discharge by drainage, q_{ts}:

$$q_{rch} = N + q_{ts}$$

The discharge by drainage depends on the phreatic head φ as follows:

$$q_{ts} = -\frac{\varphi - h_0}{c} \qquad (2.11)$$

where c is a resistance factor and h_0 is the drainage level. The drainage level is taken as a constant throughout the model area ($h_0 = -0.75$ m). The coefficient c is one of the parameters considered in the calibration exercise. For the southern part of the area the drainage relation (2.11) is slightly modified (Figure 2.6):

$$q_{ts} = \begin{cases} -\dfrac{\varphi - h_0}{c_{dra}}, & \varphi \geq h_0 \\[2ex] -\dfrac{\varphi - h_0}{c_{inf}}, & \varphi < h_0 \end{cases}$$

In the northern half of the model the resistance coefficient c is smaller when the system is draining than when it infiltrates ($c_{dra,n} < c_{inf,n}$). In the southern half the value for infiltration is very large, such that no infiltration from the surface water occurs and the groundwater recharge is less or equal to the precipitation recharge N. Note that close to the boundary at the right hand side, where $\varphi < h_0$, some infiltration occurs. In total we are interested in the following three values for the resistance coefficient: $c_{dra,n}$ (drainage, northern part), $c_{dra,s}$ (drainage, south), and $c_{inf,n}$ (infiltration northern part). All calculations are carried out with the simulation package TRIWACO (Royal Haskoning, 2002).

$$N = 1\,mm/d$$

$$\varphi = 0\,m$$

$$Q = 1000\,m^3/d$$

$$\varphi = -1\,m$$

$$\varphi_w > h_0$$

drainage level $h_0 = -0.75\,m$

$$kD = 1000\,m^2/d$$

5000 m
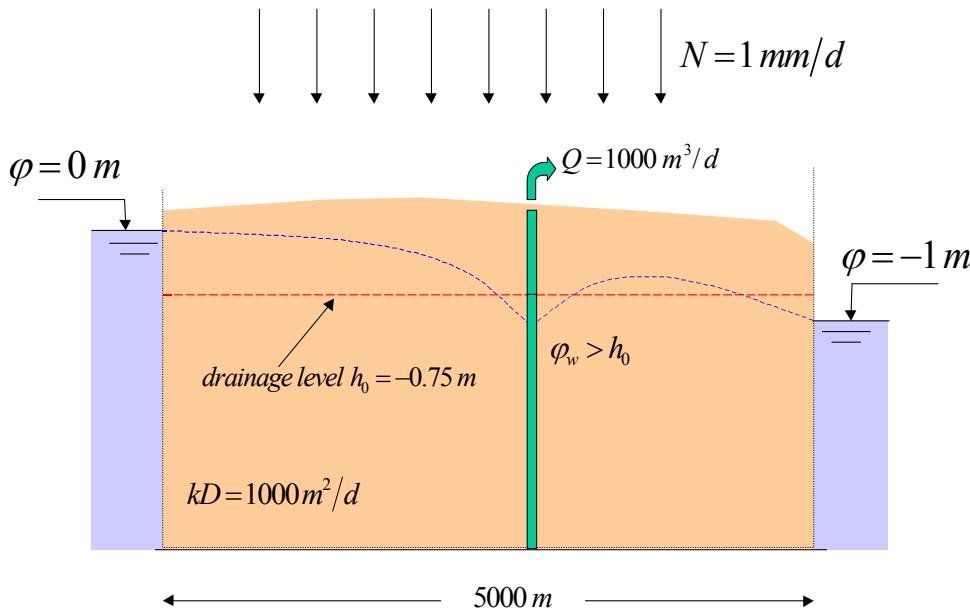
Figure 2.5 Cross-section Case 1.
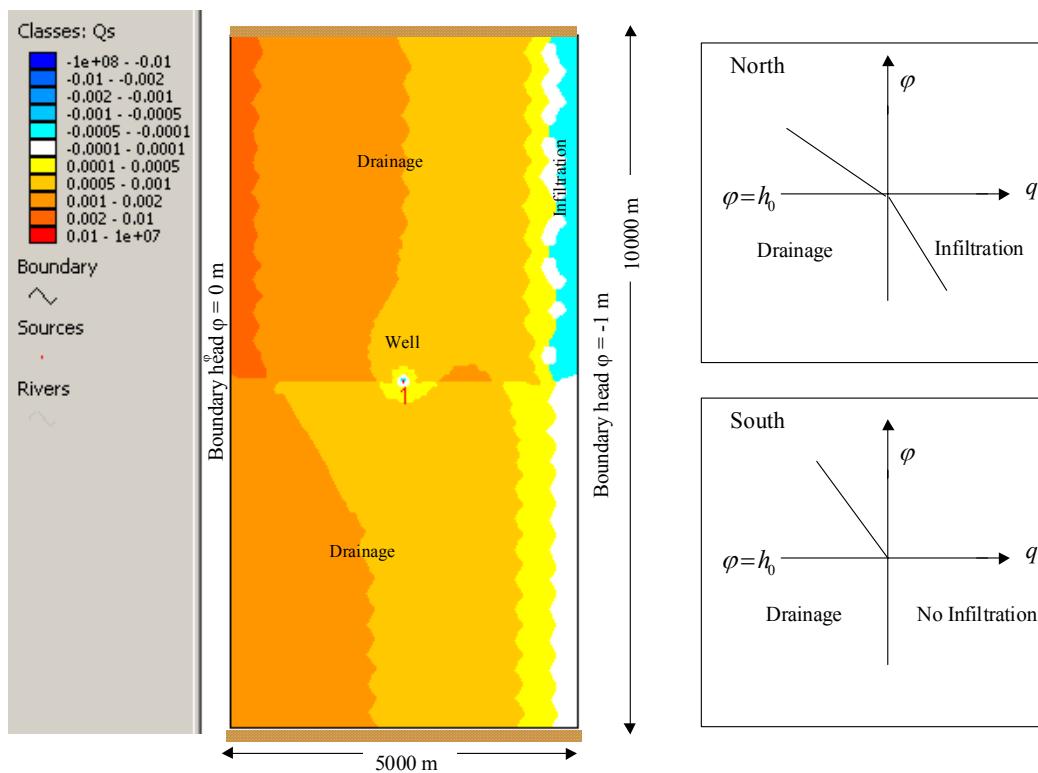
Figure 2.6 Plane view (left) and drainage relation (right) for Case 1 (TRIWACO).

First, a reference run is performed. In this run the parameters involved in the calibration are fixed (Table 2.2). The first parameter in Table 2.2, α, is a multiplication factor for the transmissivity of the aquifer (1000 m^2 day^{-1}). The remaining three parameters are the top system coefficients described above. The heads from the reference run are presented in Figure 2.7. Random components, sampled from a normal distribution with a standard deviation $\sigma_\varphi = 0.2$ m, have been

added to the heads to construct a set of 400 synthetic head observations $\hat{\varphi}_i$. These heads are located approximately within the rectangle shown on Figure 2.7. Secondly, TrCalCon, the TRIWACO module for calibration and confidence analysis was applied (Zaadnoordijk, 2001). The calibration routine is used with $\hat{\varphi}_i$ as "measured heads". The inverse model yields optimal parameter values that are identical to the ones used in the reference run (as expected), but it also produces the covariance $Cov(p_i, p_j)$.

Table 2.3 gives the optimal values and standard deviations σ_{p_i} as obtained by calibration. For the standard deviation of c_{inf} an extremely high value was found indicating that this parameter is not sensitive for head data. Therefore, this parameter is no longer included in the Monte Carlo exercise. Table 2.4 gives the correlation matrix ρ_{ij} for the remaining three parameters. The covariance follows from the correlation matrix and standard deviation by: $Cov(p_i, p_j) = \sigma_{p_i} \sigma_{p_j} \rho_{ij}$.

Table 2.2 Initial values of parameters as used in reference run. Case 1.

Parameter	Symbol	Value
P_1	α	1 (multiplication factor)
P_2	$C_{dra,n}$	250 days
P_3	$C_{dra,s}$	500 days
P_4	$c_{inf,n}$	500 days

Table 2.3 Calibrated values of parameters p_i. Case 1.

Parameter i	Optimal value μ_{p_i}	Standard deviation σ_i
1	1 (multiplication factor)	0.18
2	250 days	16 (days)
3	500 days	32 (days)
4	500 days	(extreme)

Table 2.4 Correlation matrix $\rho_{ij} = \dfrac{Cov(p_i, p_j)}{\sigma_{p_i}\sigma_{p_j}}$. Case 1.

$$\rho_{ij} = \begin{pmatrix} 1.00 & -0.214 & 0.367 \\ -0.214 & 1.00 & -0.296 \\ 0.367 & -0.296 & 1.00 \end{pmatrix}$$
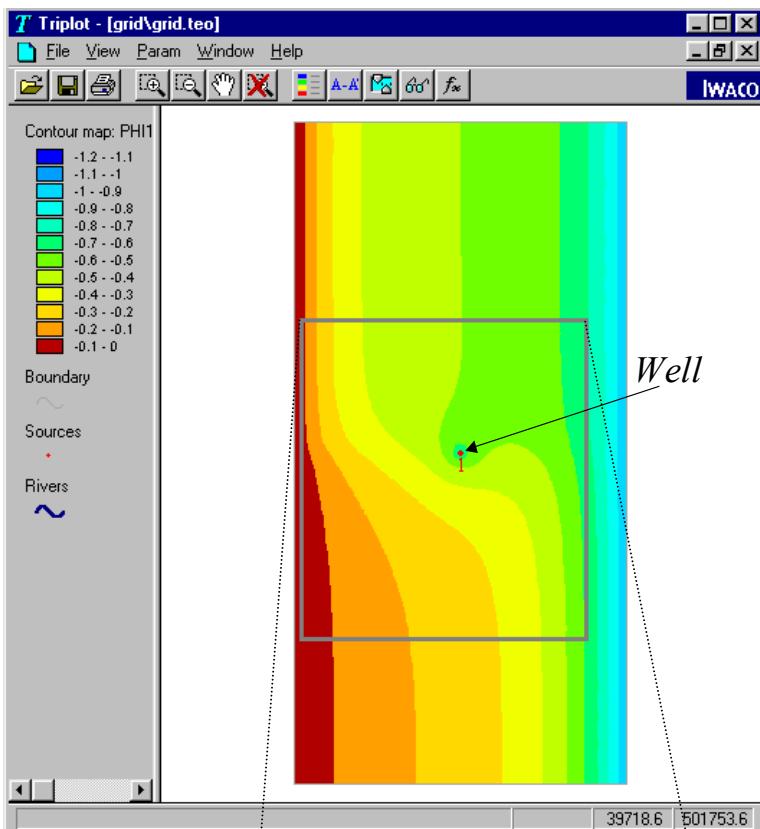
Figure 2.7 Head distribution for Case 1.
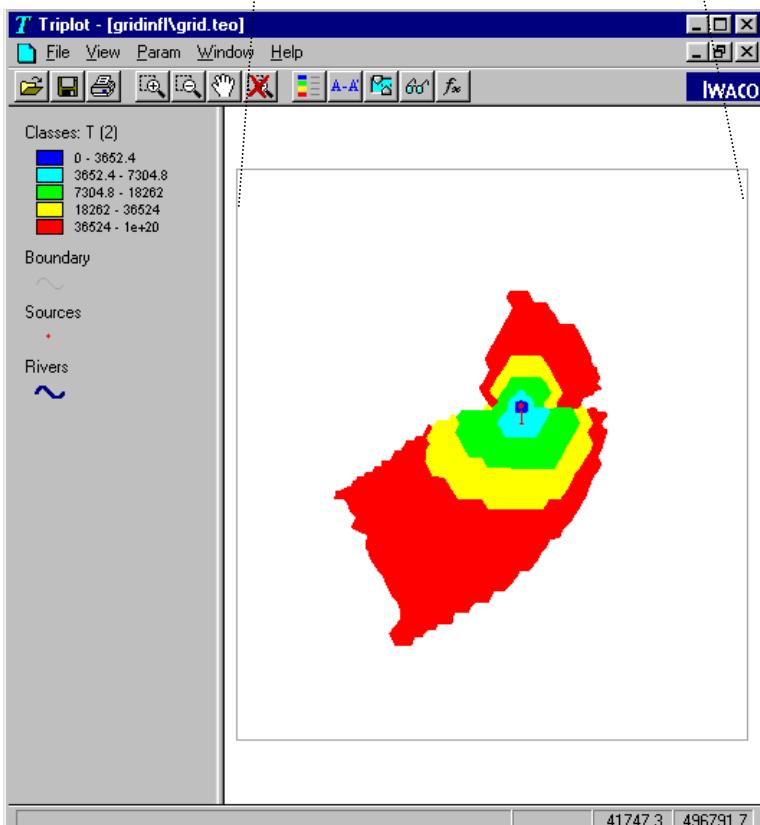
Figure 2.8 Map of expected travel times (T_{50}) for Case 1 (TRIWACO). Travel-time classes are 0-10, 10-20, 20-50, 50-100, and >50 year.

Travel times

With the optimal parameter values the "expected" travel times (T_{50}) have been obtained. Contours are given in Figure 2.8. The transitions of the colours mark the isochrones for 10, 20, 50 and 100 years. The (irregular) pattern of the travel time zones can be explained mainly by the drainage pattern via the top-system in combination with the heads (Figure 2.7). The picture in Figure 2.9 is a close-up from the area around the well. It contains, in addition to the contours of the "expected" travel time, the 10, 20, 50 and 100 years isochrones of "inner" and "outer" travel-time as obtained by MC (see (2.4)) and by FOSM (normal variant), see (2.6). Pink lines indicate "inner" and "outer" isochrones as obtained by the MC method, while the black lines give "inner" and "outer" isochrones found by FOSM. Note that the "expected" isochrones are not indicated explicitly, but coincide with the colour transitions.
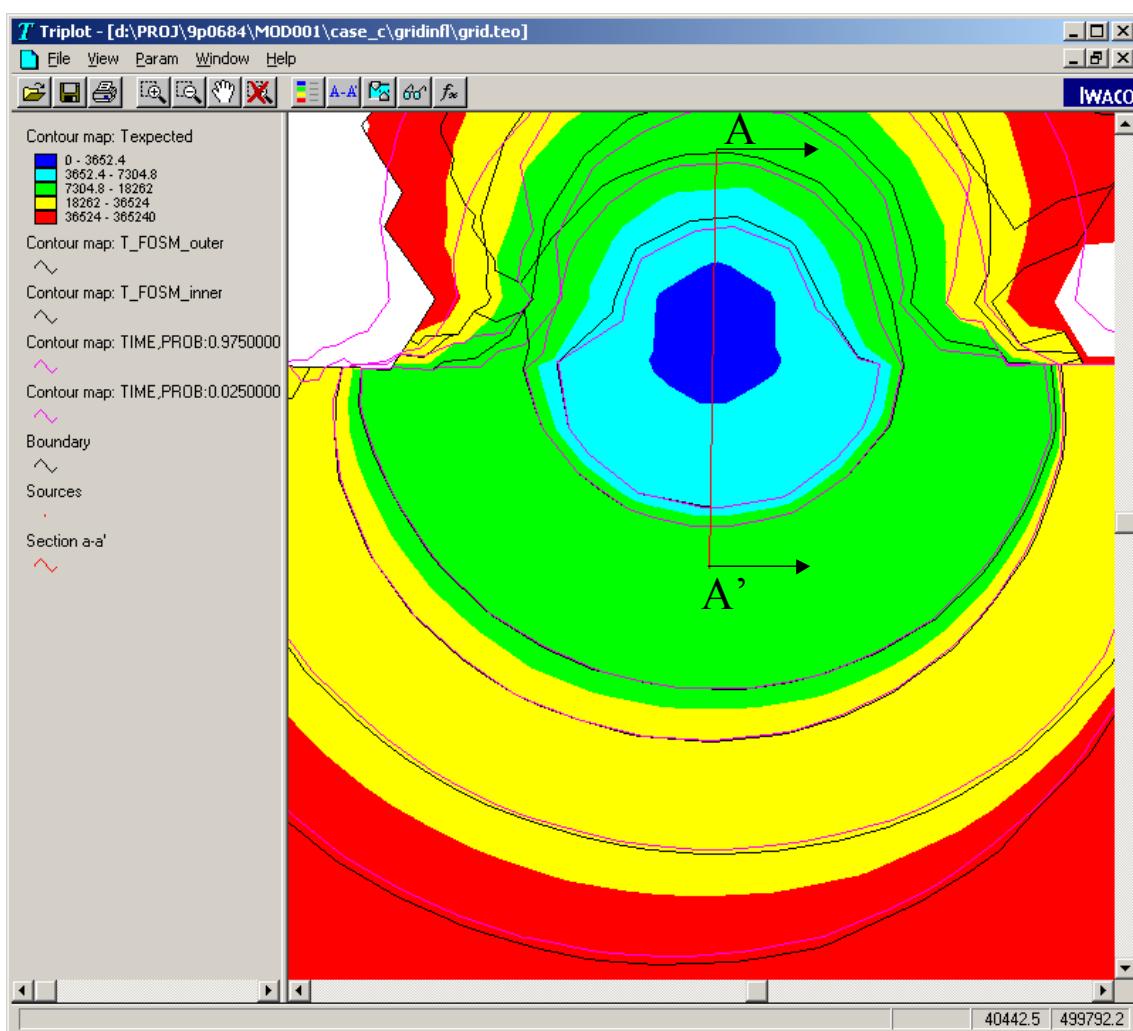
Figure 2.9 "Inner", "outer" and "expected" isochrones by MC and FOSM (10, 20, 50, and 100 years) for Case 1 (TRIWACO). For section A-A' see Figure 2.10. Travel-time classes are 0-10, 10-20, 20-50, 50-100, and >50 year.

Figure 2.9 shows that the overall agreement of the isochrones obtained by FOSM and MC is quite good. Further, it is noticed that when a difference occurs, it is a systematic one. The isochrones obtained from FOSM are always located closer to the well than the corresponding MC-isochrones. The distance between "inner" and "outer" isochrones appears to be the same for FOSM and MC. In an alternative presentation the situation is illustrated more clearly. Figure 2.10 shows the increase of travel time T along a cross-section A-A' as indicated on Figure 2.9. The travel times represent "inner" and "outer" times as obtained by FOSM (black) and MC (pink). The intersection of these lines with the line $T = 20$ years defines points on the isochrones. The order in which these points occur moving away from the well appears to be the same for the 10, 20, 50 and 100 year, as can be seen in Figure 2.9. We may conclude that the shape of the *pdf* is not reproduced optimally by FOSM but that the underestimation of the variance due to neglecting the higher order terms is very small.
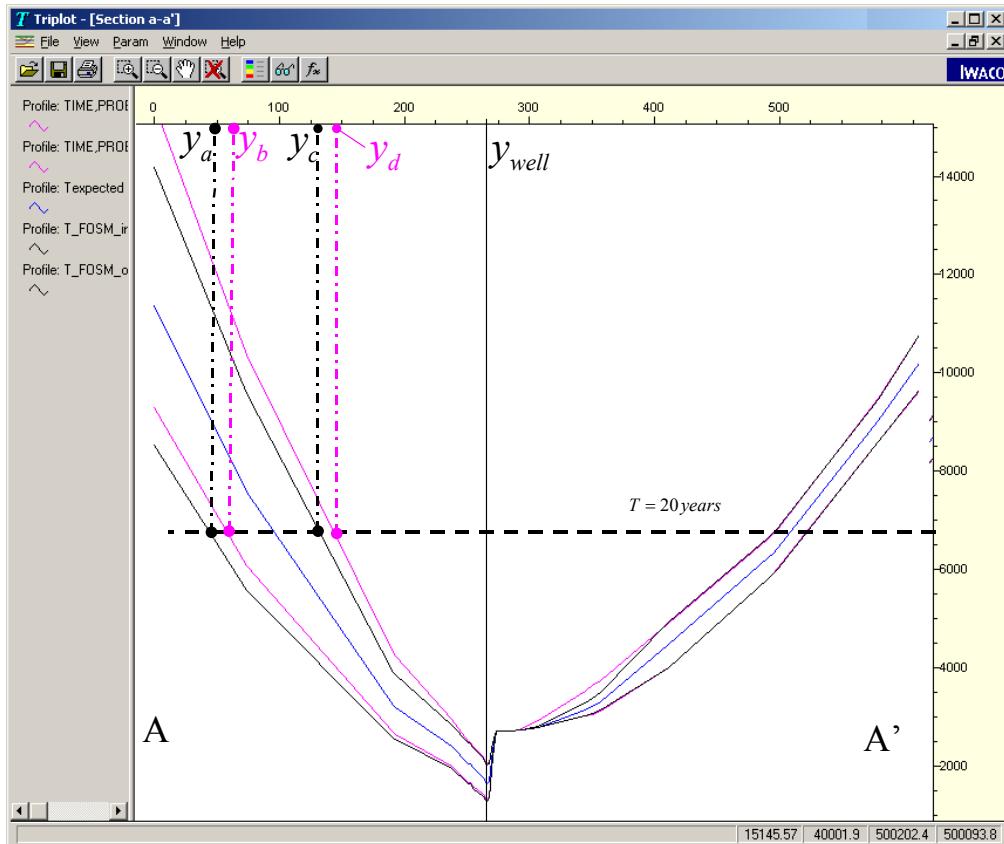
Figure 2.10 Case 1 (TRIWACO): travel-times along a section A-A' (see Figure 2.9) and position of "inner" and "outer" 20-years isochrones (r_a, r_b, r_c, r_d) with respect to the well: Inner-isochrone FOSM (r_a), Inner-isochrone MC (r_b), Outer-isochrone MC (r_c), Outer-isochrone FOSM (r_d).

2.5.2 Case 2 (Lochem)

The second case concerns the pumping station Lochem. The pumping station Lochem is located in the eastern part of the Netherlands (see Figure 2.11). The geohydrological system consists of an unconfined (phreatic) aquifer with an averaged thickness of 60 m. The impervious base of the
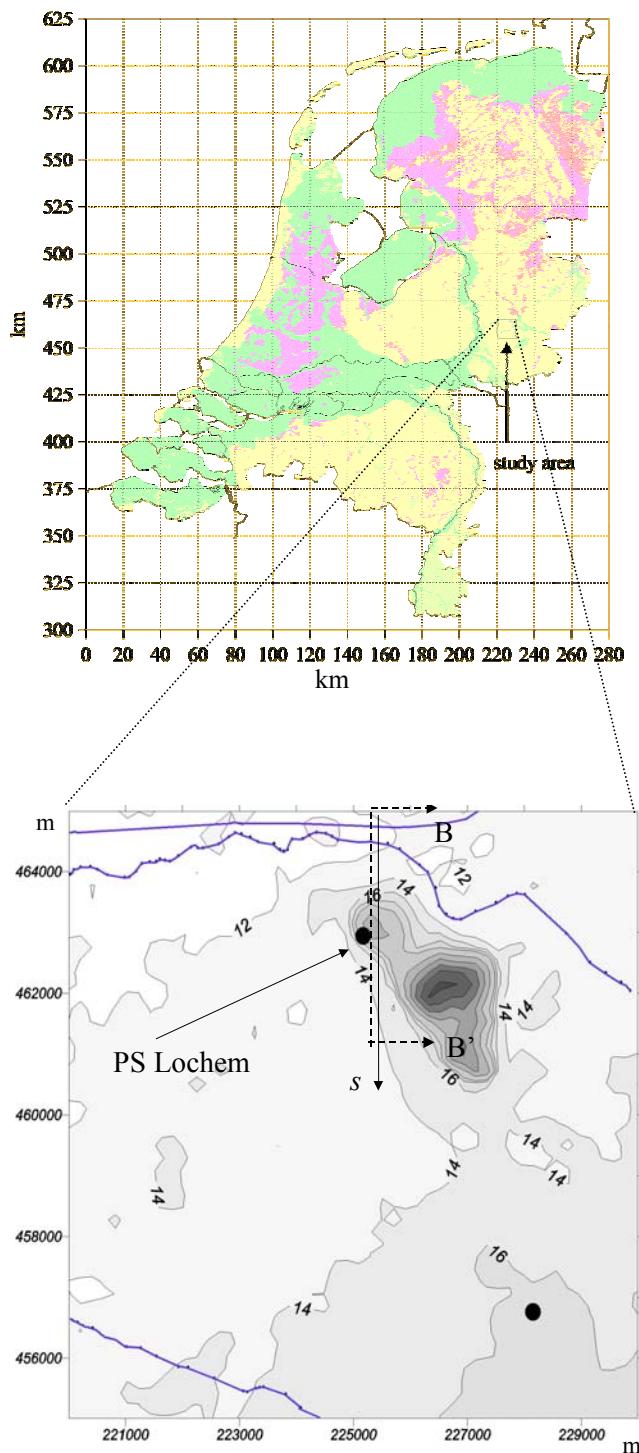
Figure 2.11 Case 2 (Lochem): location and ground-surface elevation around pumping station Lochem, and position of cross-section B-B'. Cross-section B-B' is located 30 m east of pumping station Lochem.

aquifer is found at a depth varying from 40 m below m.s.l. in the South to 60 m below m.s.l. in the North. The prevailing gradient of the groundwater flow is from South-East to North-West. The groundwater abstraction rate of the pumping station is 5133 m^3 day $^{-1}$. A typical feature in the landscape is the "Lochemer berg", an elevated region with a substantial infiltration of groundwater. Like Case 1, this case is also treated as a steady groundwater flow. More information on the geohydrological structure of the Lochem area can be found e.g. in Kovar *et al.* (1996), Uffink & Van der Linden (1998) and Uffink & Römkens (2001).

Calibration

With real measured groundwater heads as input data a model calibration has been performed (assuming steady-state) to obtain mean and (co)-variances for six model parameters. The calibration procedure itself is along the lines of the procedure described in detail by Leijnse *et al.* (2002a, 2002b). Among the calibrated parameters are the drainage coefficients in several sub-areas. These are areas belonging to four distinct classes of the Gt-value (i.e. Gt = 3, 5, 6, 7). The four drainage coefficients are W1(30), W1(50), W1(60) and W1(70). Two parameters (ALF11 and ALF21) represent a multiplication factor for the transmissivity in the first (topmost) and second aquifer, respectively. Results of the model calibration are presented as the first six items in Table 2.5.

The remaining three parameters in Table 2.5 involve:
–   multiplication factor F_PF1 for the effective porosity of the aquifers (PF1);
–   multiplication factor F_PA0 for the effective porosity of the aquitards (PA0);
–   multiplication factor F_QRE for the groundwater recharge rate (QRE).
These multiplication-factor parameters were not part of the calibration process, but are added to the list parameters to be varied during MC and FOSM. Mean value of each of these factors is 1. Mean values for the parameters themselves (PF1, PA0 and QRE) have been assumed: PF1 as spatially constant value of 0.3, PA0 as spatially constant value of 0.1, and QRE as spatially variable value, accounting for the spatial variability in precipitation and evapotranspiration. Also the variances of F_PF1, F_PA0 and F_QRE have been assumed. The variance $s^2 = 4.3 \times 10^{-4}$ has been established as follows. Let X denote any of these three parameters and let X_0 be the average value (geometric mean). Now, the variance is chosen such that 95% of the values is within in the interval:

$$\frac{X_0}{1.1} < X < 1.1 \, X_0$$

In a log-normal distribution (based on the 10-log) these limits can be expressed in (geometric) mean and standard deviation s of $^{10}\log X$ (also see (2.10)):

$$\frac{X_0}{10^{2s}} < X < X_0 \, 10^{2s}$$

By equating $10^{2s} = 1.1$ one obtains the variance $s^2 = 4.3 \times 10^{-4}$.

Travel times

As starting points for the calculation of travel-times for Case 2, we have chosen a line-segment B-B' of 5000 m length (see Figure 2.11). The *x*-coordinate of this line is 225200. It runs from the northern edge of the model-area (y = 465000) in southern direction. This position of the line segment in *x*-direction is such that it (almost) intersects the pumping station (coordinates pumping station *x* = 225170, *y* = 462945). Along the line B-B' travel-times ("inner", "outer" and "expected") have been determined with the program LGMLUC, which is based on the FOSM approach (see Appendix 1). A list of the uncertain parameters is given in Table 2.5. Mean ("expected") travel times have been calculated directly by particle tracking, as given by (2.8). "Inner" and "outer" travel times follow from (2.9) and (2.10). The results are displayed in Figure 2.12, along a distance *s* being the distance measured from the upper edge of the model. It is clearly seen that minimum travel times occur when the distance to the pumping well is smallest. The difference between "inner" or "outer" travel time compared to the expected value is in the order of 50% or more.

For 14 points along the line B-B' travel times are determined with the Monte Carlo method using the model LGMMUC (see Appendix 1). The number of MC runs is 2000. For the selection of parameter values LGMMUC uses a sampling procedure from Press *et al.* (1988). All parameters are assumed log-normally distributed with mean and covariances from Table 2.5 and 2.6. For several of the 14 chosen points not all 2000 runs produced a travel time. This is due to the fact that for some combinations of the model parameters the point under consideration falls outside the capture zone. For the points with a complete or almost complete table of travel times the "expected", "inner" and "outer" travel times have been determined and these values are plotted in Figure 2.12. It may be concluded from this figure that the FOSM results correspond well with the MC results. Especially for "outer" and "expected" travel times the match between FOSM and MC results is good. Some mismatch occurs between FOSM and MC with respect to the "inner" travel times.

Table 2.5 Mean and variances of uncertain parameters for test Case 2.

Code	Mean value	Variance s^2 (^{10}log)
W1(30)	122.7 [days]	0.482
W1(50)	338 [days]	0.447
W1(60)	291 [days]	0.463
W1(70)	263 [days]	0.475
ALF11	1.04 [-]	0.0913
ALF21	1.13 [-]	0.0926
F_PF1	1 [-]	4.3×10^{-4}
F_PA0	1 [-]	4.3×10^{-4}
F_QRE	1 [-]	4.3×10^{-4}

 (more information about the parameters is given in section 3.3)

Table 2.6 Covariances of uncertain parameters for test Case 2.

W1(30)	W1(50)	W1(60)	W1(70)	ALF11	ALF21	F_PF1	F_PA0	F_QRE
0.482	-0.024	-0.025	-0.017	0.0006	0.0017	0	0	0
-0.024	0.447	-0.05	-0.038	0.0016	0.0035	0	0	0
-0.025	-0.05	0.463	-0.033	0.0006	0.0014	0	0	0
-0.017	-0.038	-0.033	0.475	0.0006	0.0013	0	0	0
0.0006	0.0016	0.0006	0.0006	0.0913	0.0007	0	0	0
0.0017	0.0035	0.0014	0.0013	0.0007	0.0926	0	0	0
0	0	0	0	0	0	0.00043	0	0
0	0	0	0	0	0	0	0.00043	0
0	0	0	0	0	0	0	0	0.00043
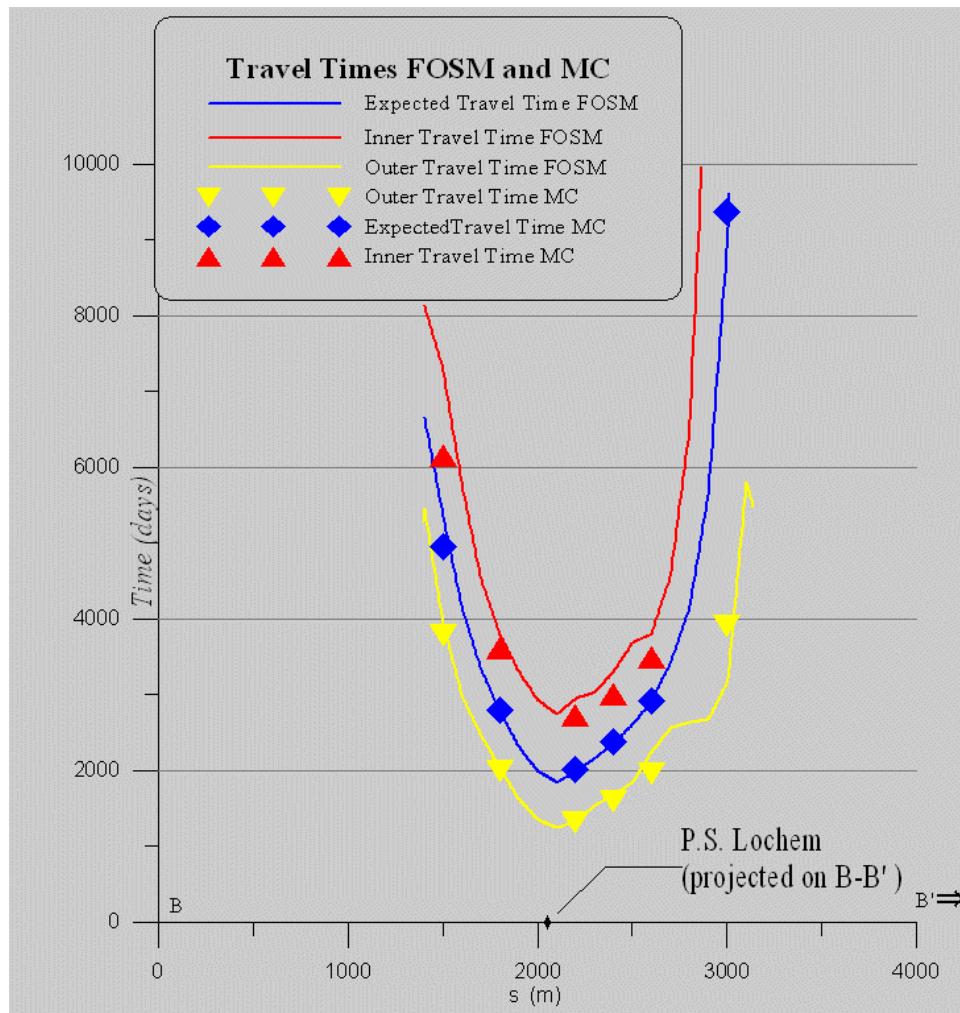
Figure 2.12 Case 2 (Lochem): "Inner", "outer" and "expected" travel times along the line B-B' (see Figure 2.11). FOSM results by LGMLUC, MC results by LGMMUC.

2.6   Discussion

The main points of interests with respect to the test case results are:

(i)   What type of *pdf*, normal or log-normal, matches best with the MC results and which of the two should be chosen in the FOSM approach?

(ii)   How do the results of the Monte Carlo approach compare to the results obtained by the first-order approach (FOSM) and is the first-order approach a suitable alternative for the costly Monte Carlo method?

For test Case 2 histograms are plotted for six points along the line B-B' in order to investigate the type of distribution of the travel times obtained by MC. Figure 2.13 shows the histograms for the travel time itself (*T*), while Figure 2.14 gives histograms for the log-transformed travel time ($^{10}\log(T)$). For five out of six points there is no decisive answer whether the normal or log-normal distribution fits better. Only for one point (Figure 2.13(f); Figure 2.14(f)) the distribution is clearly log-normal. In the literature several indications are found for log-normal distributed travel times (Andrews *et al.*, as reported by LaVenue *et al.*, 1989). Additional support is found in James & Chrysikopolous (2001) and Banton *et al.* (1997). The fact that in a normal distribution the variable, in principle, can be negative, is also a sign that a log-normal distribution is more appropriate for travel times. For these reasons the log-normal travel time distribution has been assumed for pilot study application reported in Chapter 3.

With respect to the match between FOSM and MC results we have found that test Case 1 (TRIWACO) and test Case 2 (Lochem) both show that first order results agree well with the results obtained by Monte Carlo method. A slight underestimation of the inner and outer travel times appears in the FOSM approach for the TRIWACO case. Outer travel times are also slightly underestimated by FOSM in the Lochem test case, but here the inner travel times agree very well with the MC results. The underestimation might be due to an incorrect assumption of log-normality. In addition a small underestimation of the bandwidth can be noticed in test Case 1 (Figure 2.10, Case 1, TRIWACO), which is due to neglecting the higher order terms.

Our overall conclusion from the match between FOSM and MC results is that the FOSM approach is a reliable and computationally effective method to analyse the uncertainty of travel time zones.
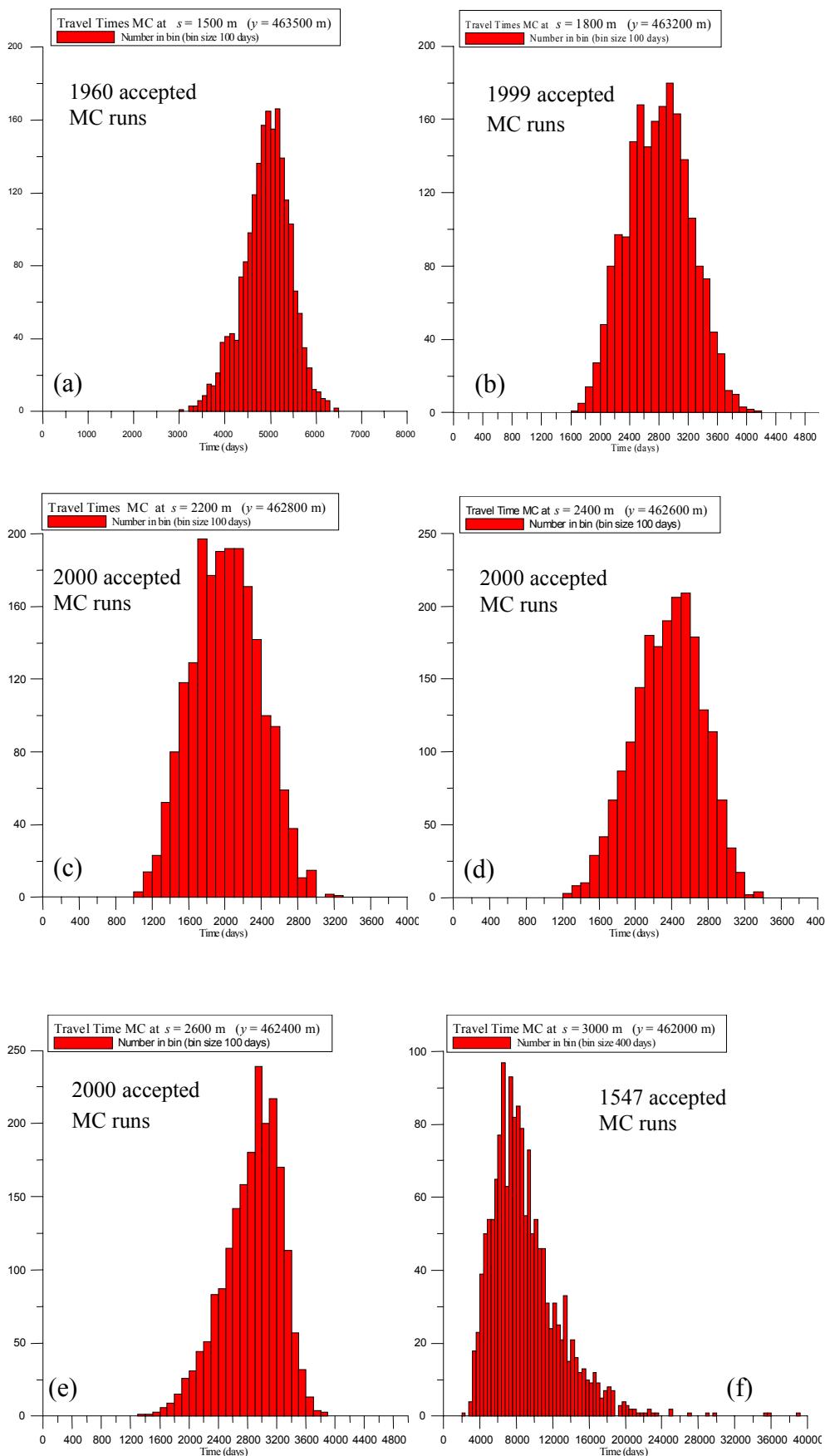
Figure 2.13 Histograms of T for six points along the line B-B'. Results for Case 2 (Lochem).
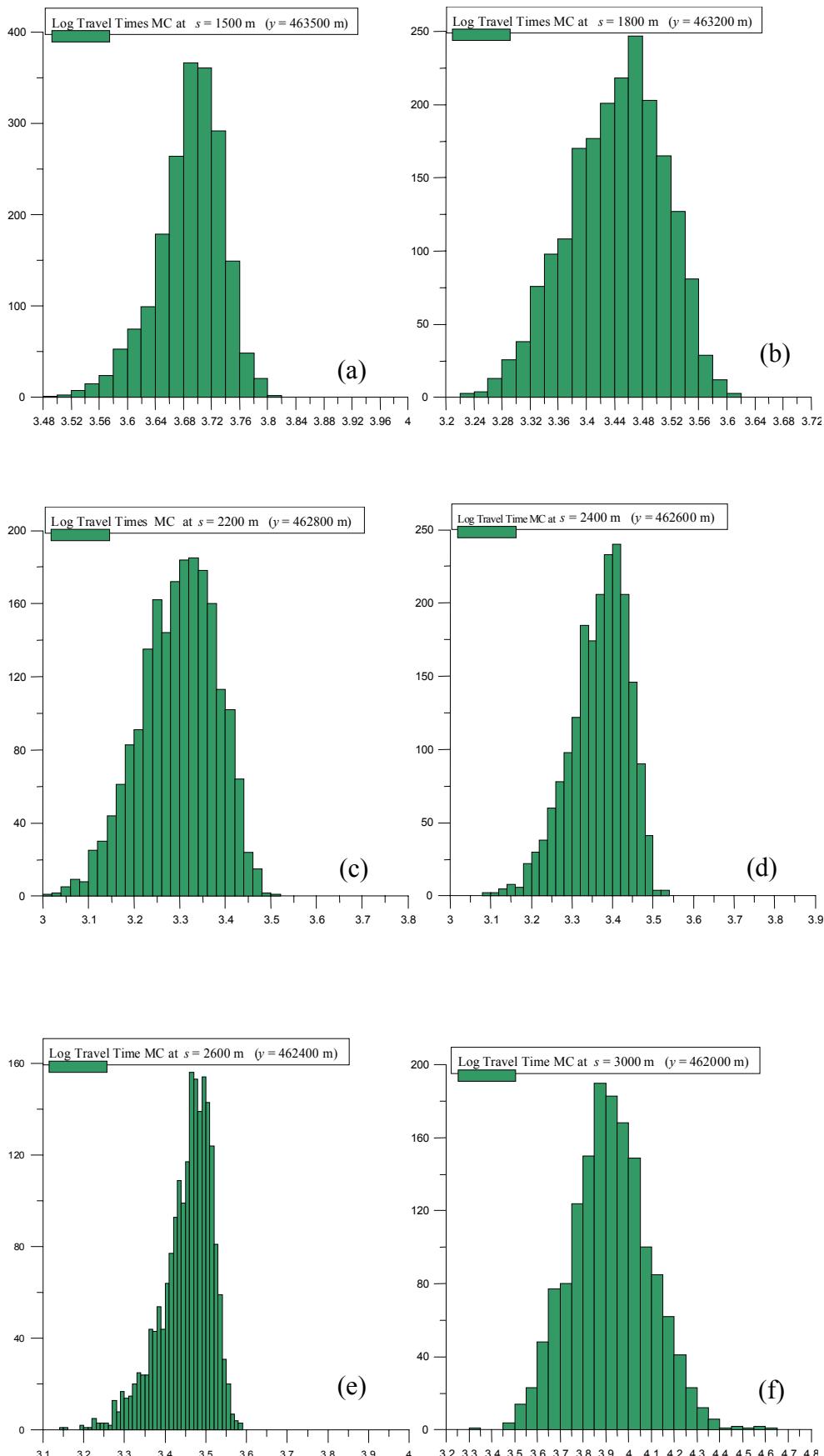
Figure 2.14 Histograms of $^{10}log(T)$ for six points along the line B-B'. Results for Case 2 (Lochem).

3. Pilot study application of FOSM approach

3.1 Introduction

For the pilot study reported here use is made of the existing model Utrecht. Figure 3.1 shows the location of the model in the Netherlands. This groundwater model (50 by 50 km) regards the saturated multi-aquifer geohydrological system consisting of four aquifers separated by aquitards. The steady-state flow in the system is assumed to be quasi-three-dimenensional.

The primary reason for selecting the Utrecht model was that, in a previous study, this model was calibrated by means of an inverse method (Leijnse *et al.*, 2002a, 2002b). The calibration output available at the onset of the current study were (a) the optimized parameter values of aquifer transmissivity (10 items), drainage resistance for the surface-water-groundwater interaction (8 items), and the hydraulic resistance of aquitards (4 items), and (b) the covariance matrix for those 22 parameter items. For details of the calibration refer to section 3.3.

Figure 3.1 Location of model area Utrecht (50 by 50 km) in the Netherlands. Model location is: xmin=123000, xmax=173000, ymin=425000, and ymax=475000 m.

The aim of this pilot study is to illustrate the applicability of the First-Order Second-Moment (FOSM) method for calculation of the reliability (or uncertainty, if one wishes to use that term) of travel times to groundwater abstraction sites. Specifically, it is the water-particle travel time that elapsed between the moment a water particle enters the system at the groundwater table and the moment when the particle is captured by a well screen.

The model area Utrecht comprises (a) groundwater abstractions for drinking-water supply and (b) abstractions for industrial purposes. Figure 3.2 depicts the location of all abstraction sites. However, for simplicity, we have selected only a few drinking-water abstractions for calculation of the travel-time reliability, i.e. the illustration of the applicability of the FOSM method. The latter abstractions are those in Figure 3.2 where the abstraction-site name (e.g. Zeist) is included at the

location marker. Those abstraction sites are also used for presentation of the travel-time reliability results in Figures 3.22 and 3.23.
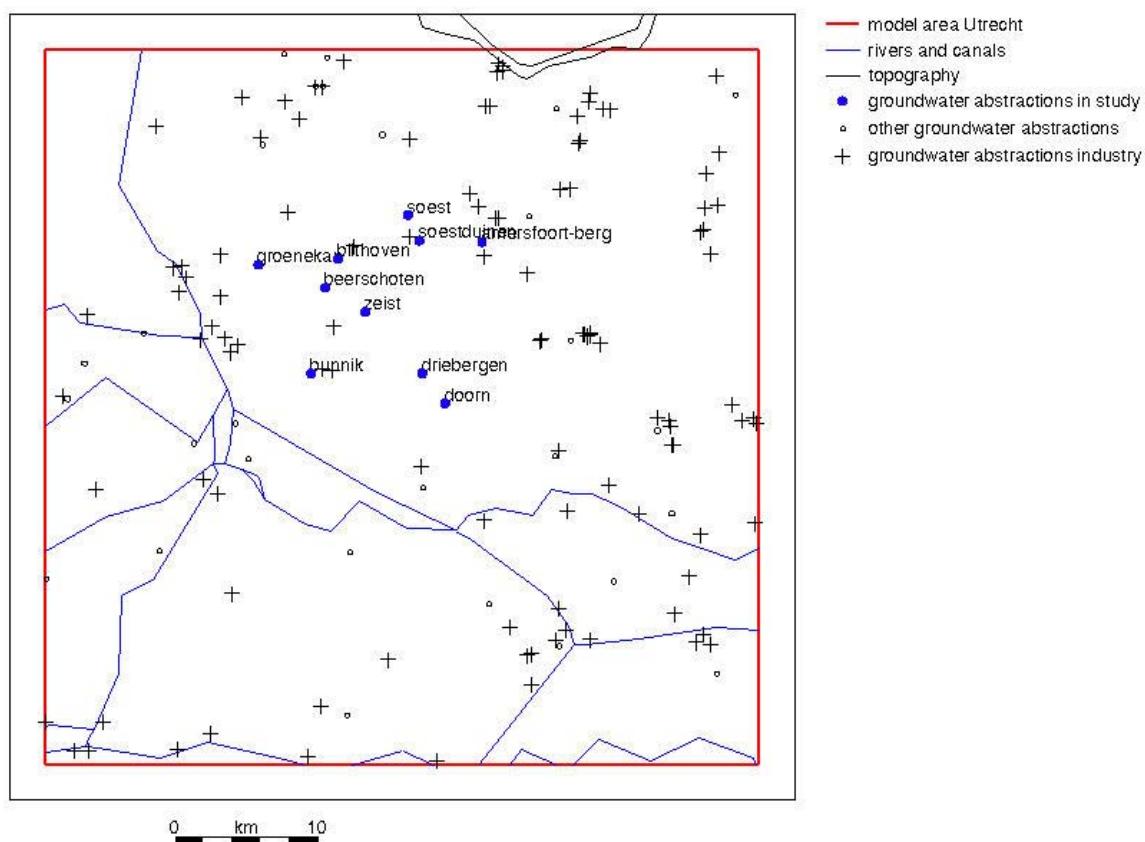
Figure 3.2 Location of groundwater abstractions for drinking-water supply and for industrial purposes. Ten drinking-water abstractions (blue marker with text label) are used to calculate reliability of travel times.

3.2   Modelling groundwater potential

3.2.1   Introduction

This section gives an overview of LGM and its application for modelling the groundwater potential problem, i.e. for the simulation of groundwater heads and fluxes. Specifically, it describes the concept for the modules LGMSAT and LGMCAL. The simulation results, the groundwater head and the flux, are presented in section 3.3.

LGM, the Netherlands Groundwater Model (in Dutch Landelijk Grondwater Model) is a model for the simulation of quantity and quality aspects of saturated groundwater systems. LGM was developed by the National Institute of Public Health and the Environment, RIVM (Pastoors, 1992; Kovar *et al.*, 1992) and applied in various studies (e.g., Kovar *et al.*, 1998; Kovar *et al.*, 2000; Stoppelenburg *et al.*, 2002). For this study the LGM version 3 was used.

LGM is based on the numerical technique of finite elements. The elements are quadrilaterals and triangles. LGM can be applied for any user-selected grid density, the grid density being dependent on the specific conditions of the problem to be solved. The grid can be locally refined, e.g. within a well capture area or in the vicinity of rivers. Figure 3.3 shows schematically the basic features of the LGM's finite element implementation, namely the grid nodes, the elements, and the node influence areas, A_{inf} [L^2]. The influence areas are highlighted at four nodes.
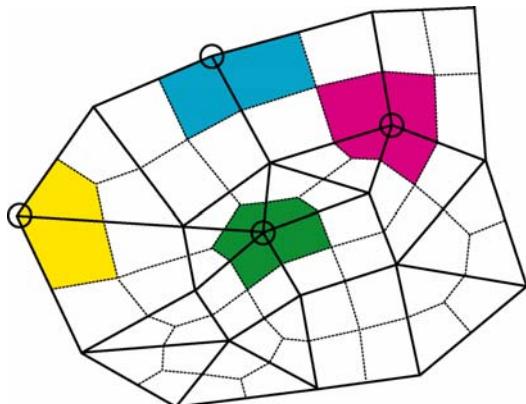
Figure 3.3 Example of finite element grid for LGM, containing nodes, elements and node influence areas (dashed-line polygons).

Figure 3.4 shows an overview of the functionality of the various computer programs (modules) within LGM, presented in the sequence in which the modules have to be used.

LGMGRID
| Generation of finite-element grid
|
Allocation of parameter values to grid nodes (Arc/INFO, LGM pre-processing modules)
|
LGMCAL (calls module LGMSAT to solve groundwater potential problem)
| Calibration of parameters in multi-aquifer system (yielding optimal values and covariance)
| (this step was not done here since we used results of existing study, see section 3.3)
|
LGMLUC and/or LGMMUC (calls modules LGMSAT and LGMFLOW)
| Calculation of reliability of travel times, by FOSM method or Monte Carlo approach
| (LGMLUC is used for pilot study reported in this chapter, section 3.5)

Figure 3.4 LGM modules and their interrelationships.

First the module LGMGRID is used to generate the finite element grid. Second, the spatially distributed data (stored in Arc/INFO) are discretized to the model input values. This data allocation step yields the parameter values at the nodes of the finite element grid.

Module LGMSAT can be used to carry out the calibration of selected parameters, using groundwater heads as observed variables. LGMSAT generates optimized parameter values and the

covariance (or correlation) matrix for those parameters. The inverse method used is described in Leijnse *et al.* (2002a, 2002b). The kernel of the calibration procedure is module LGMSAT, for calculation of groundwater heads in aquifers, the flux across aquitards, the flux between the top aquifer and rivers, and the flux between the top aquifer and the topsystem (small-scale surface waters, such as polder ditches).

Module LGMLUC can be used to calculate the reliability –or if you prefer that word, uncertainty– of the travel time from top of aquifer 1 to a well screen of a groundwater abstraction site. LGMLUC is based on the First-Order Second-Moment (FOSM) method. While running LGMLUC, the modules LGMSAT and LGMFLOW are pair-wise processed a number of times. LGMFLOW generates a pattern of pathlines in forward-tracking mode, the patlines being started at top of aquifer 1. Since the FOSM method –unlike the Monte Carlo approach– is computationally tractable even for grids consisting of large number of nodes, LGMLUC was used for the pilot study reported in this chapter. The methodology, and input and output for LGMLUC are described in Appendix 1.

Module LGMMUC also calculates the reliability of travel times. However, instead of using the approximate FOSM method of LGMLUC, it is based on the Monte Carlo approach. Repetitively, a series of input parameter values is prepared with the help of a random generator, taking into account the correlation between the parameters. For each input-parameter set realization, the modules LGMSAT and LGMFLOW are run. The outcome of LGMMUC is, for each of the given starting points, the table with generated travel times. This table can be used to calculate percentiles of travel time, e.g. 2.5, 5, 95, and 97.5. LGMMUC was not used for the production runs in this pilot study, it was used (see Chapter 2) to assess the applicability of the FOSM method used in LGMLUC. The methodology, and input and output for LGMMUC are described in Appendix 1.

3.2.2   Geohydrological system

Figure 3.5 depicts schematically the geohydrological system used in LGM, version 3. The groundwater system in the Netherlands can be described as a multi-aquifer system consisting of a sequence of aquifers and aquitards, where the groundwater head in the top aquifer (sometimes phreatic) is strongly influenced by the small-scale surface water system, and rivers and canals. For the description of the groundwater flow in this system, the Dupuit assumption can be adopted for the flow in the aquifers, while the flow in the aquitards is assumed to be vertical one-dimensional. This approach is often referred to as "quasi three-dimensional". Four aquifers are distinguished in LGM, version 3.
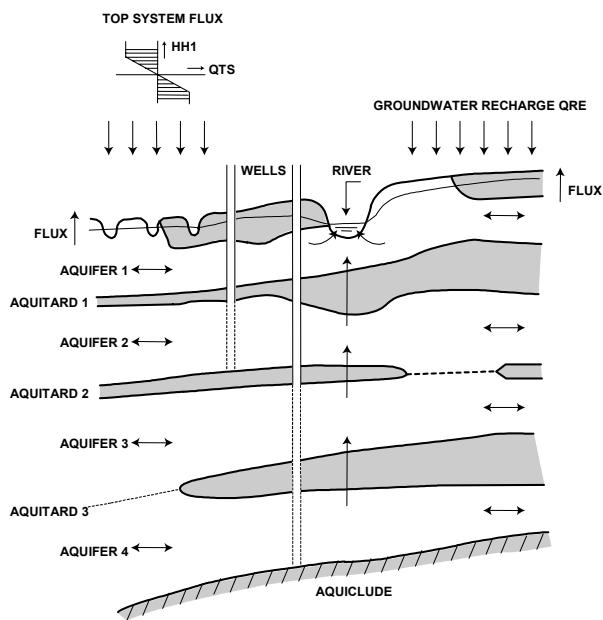
Figure 3.5 Geohydrological system for LGM, version 3.

The geohydrological system of aquifers 1 through 4, and the separating aquitards, is parameterized using the basic information listed in Pastoors (1992). Among other items, the database contains:
– the z-level, elevation (metres above m.s.l.) of the ground surface;
– the z-level, elevation (metres above m.s.l.) of the top and the bottom of each aquifer and aquitard;
– the hydraulic conductivity of aquifers;
– the relationship between the aquitard thickness and the aquitard depth, and the resulting hydraulic resistance of the aquitard. After all, the resistance does not have to be a linear function of the layer thickness. This information is spatially structured through the identification code of geological formations, i.e. each formation having its own unique characteristics.

The transmissivity of aquifers at each model node is calculated as the product of the hydraulic conductivity and the aquifer thickness, the latter being the difference of z-levels at aquifer tops and bottoms. Transmissivity is one of the uncertain parameters used for the assessment of travel-time reliability. We assumed that transmissivity of only aquifer 1 and 2 affects the travel-time reliability. The transmissivity, alongside to other parameters, was calibrated with observed groundwater heads as input data. The spatially variable transmissivity-multiplication factors for aquifer 1 and 2, **ALF1i** and **ALF2i**, i = 1 to 5, are listed in Table 3.2 in section 3.3.

The top aquifer, aquifer 1, is locally overlaid by a relatively poorly permeable layer composed of clay and peat. In the model, this layer is referred to as aquitard 0, characterized by a hydraulic resistance (in units of days) against the vertical flow. The layer is present in polder areas with a dense system of ditches with controlled surface-water level. The value of the hydraulic resistance at model nodes was also one of the uncertain parameters affecting the travel-time reliability. The calibration-derived multiplication factors **B0** for hydraulic resistance are listed in Table 3.2 in section 3.3.

The hydraulic resistance of deeper aquitards (numbered 1, 2, and 3) at model nodes is calculated as a function of the aquitard thickness and the depth of aquitard-top below ground level. At places where an aquitard is lacking, the value of hydraulic resistance is assumed very small (1 day). We assumed that only the hydraulic resistance of aquitard 1 is an uncertain parameter affecting the travel-time reliability. The calibration-derived multiplication factors **B1** for hydraulic resistance are listed in Table 3.2 in section 3.3.

The elevation of the ground surface is an important model input item. The elevation map for model Utrecht is given in Figure 3.6. A dominant local topographic feature is the ice-pushed ridge Utrechtse Heuvelrug, where the ground level reaches up to 50-60 m above m.s.l. The greatest depth of the groundwater table occurs also at the Utrechtse Heuvelrug, being about 40 m below ground level. This is however an exception, as in most parts of the model the groundwater depth is within a few metres below ground level. In the left-hand part of the model one can see the polder areas with ground-surface elevations up to 2.5 m below mean sea level.
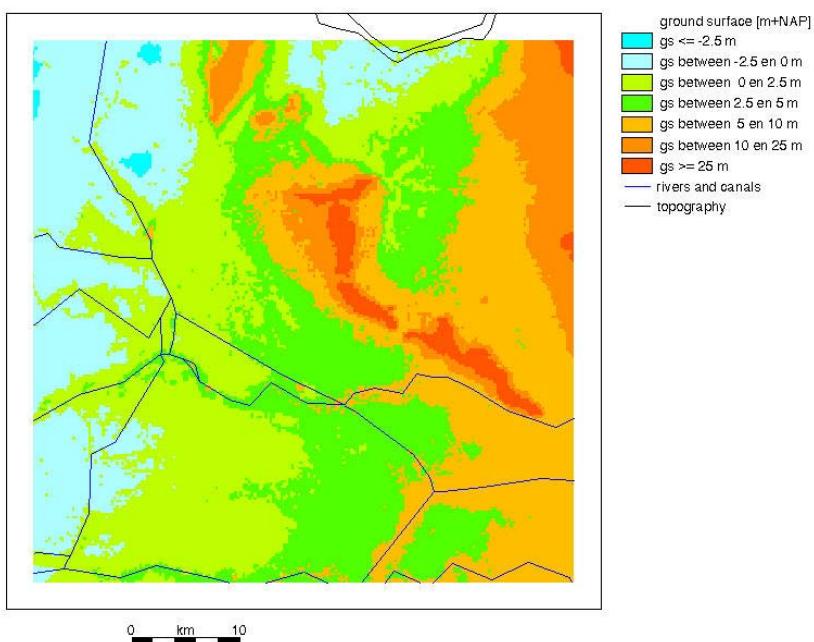
Figure 3.6 Ground surface level (metres above m.s.l.) in the model area Utrecht.

Finally, LGM requires as input the groundwater heads along the model periphery. The boundary-condition groundwater heads for each of the four aquifers were derived from TNO-NITG groundwater head observations for the year 1988 (Pastoors, 1992).

3.2.3   Groundwater-surface-water interactions

LGM features two options for the interaction with surface waters, namely (a) the so-called rivers and (b) the topsystem-flux relationship, to account for the small-scale surface waters.

Rivers

The rivers represent large (wide) river courses such as the Rhine River and its major branches, and the canals (e.g. the Amsterdam-Rhine Canal). Unlike the topsystem-flux relationship –which is spatially diffuse in nature– the rivers in the finite element grid follow their actual location in space. Figure 3.7 illustrates the location of a river, the centre line of the river being located along the element sides between five nodes. All river-related parameters can be variable along the water course. Those parameters are the river width, the hydraulic resistance of the river bottom, and the river-water level. The parameterization of rivers/canals of LGM, version 3, was based on the data from the previous version of LGM (Pastoors, 1992).
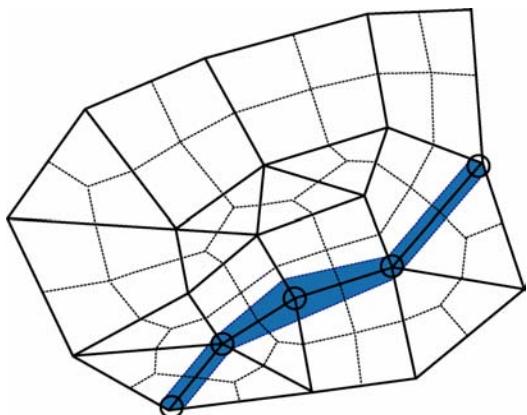
Figure 3.7 Rivers in LGM, coinciding with element sides. Dashed lines delineate the node influence areas.

In each river node, the flux between the top aquifer and the river, Q_{riv}, is defined by:

$$Q_{riv} = A_{riv} (h_{riv} - \varphi_1) / c_{riv} \tag{3.1}$$

where

Q_{riv} = flux $[L^3 T^{-1}]$ from the river to groundwater in the top aquifer
A_{riv} = river area within the influence area of the node $[L^2]$
h_{riv} = river-water level $[L]$
c_{riv} = hydraulic resistance of the river bottom $[T]$

Topsystem-flux relationship (small-scale surface waters)

The topsystem-flux relationship regards the interactions between the top aquifer and the small-scale surface waters (ditches, drains, brooks). The topsystem-flux q_{ts} $[L T^{-1}]$ is assumed constant within an influence area A_{inf} $[L^2]$ of a node, the total topsystem-flux Q_{ts} $[L^3 T^{-1}]$ at a node being expressed by:

$$Q_{ts} = A_{inf} q_{ts} \tag{3.2}$$

Figure 3.8 illustrates the shape of the topsystem-flux relationship for LGM, version 3. The relationship is composed of three topsystem points, each topsystem point being defined by a groundwater head h_{ts} and the flux q_{ts} occurring at this h_{ts}. The values $q_{ts}<0$ and $q_{ts}>0$ indicate the drainage and infiltration situation, respectively. The angle of a line in the relationshipship relates to the hydraulic resistance, a steeper the line implying higher value of hydraulic resistance. The multiple piece-wise relationships in Figure 3.8 can be seen as a combined effect of various types of flux-head relationships existing in the influence area of a given node, e.g. as a combination of the relationship for the primary system (wide ditches) and the relationship for the secondary system (narrow and shallow ditches). Note that the number of points in our topsystem relationship, i.e. three points, is identical at all grid nodes. Consequently, if no relation exists in a node or the relation is less complex than in other nodes, dummy topsystem points are introduced to preserve the same number of topsystem points for all nodes. The parameterization of the topsystem-flux relation (Pastoors, 1992) is based on the so-called Gt-value, the groundwater depth class number. The associated hydraulic resistances for drainage and infiltration stem from dewatering-design practice rules. An increasing value of the hydraulic resistance implies a steeper line in the topsystem-flux relationship in Figure 3.8.
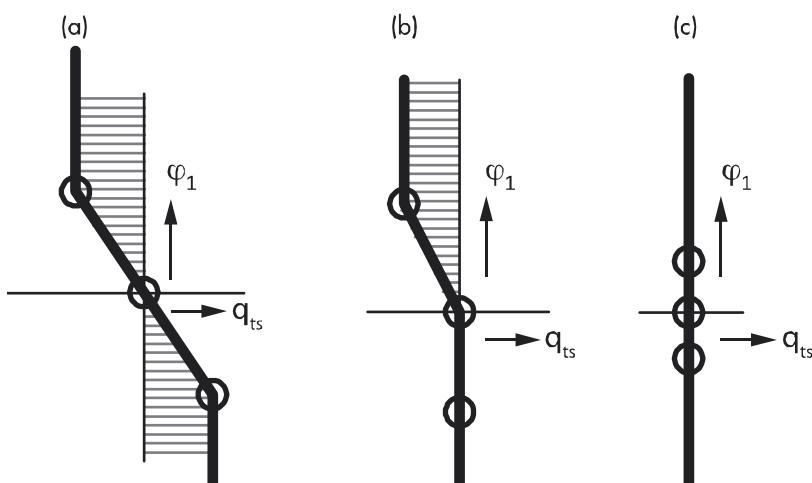
Figure 3.8 Examples of shape of the topsystem-flux relationship in LGM, version 3. The relation is defined by three topsystem points: (a) combined drainage/infiltration system, typically in polders with controlled ditch-water level, (b) drainage-only system (one dummy point below), typically on low reaches of sandy areas, (c) absence of any infiltration and drainage system (three dummy points, each with $q_{ts}=0$), typically at high elevated parts of sandy areas.

Since the topsystem-flux relationship strongly controls the behaviour of the geohydrological system, the hydraulic resistance for drainage, and for infiltration was assumed to be one of the uncertain parameters affecting the travel-time reliability. Two conditions were distinguished:
− the topsystem-flux relationship can only drain, as depicted in Figure 3.8(b). The calibration-derived drainage resistances **W1(30)** to **W1(70)** are listed in Table 3.2 in see section 3.3;
− the topsystem-flux relationship can both drain and infiltrate, with drainage and infiltration resistances being identical values, as depicted in Figure 3.8(a). The calibration-derived drainage resistances **W1(1030)** to **W1(1070)** are listed in Table 3.2 in see section 3.3.

3.2.4 Wells

The groundwater abstractions consist of (a) abstractions for drinking-water supply and (b) abstractions for industrial purposes. The LGM calculations were carried out for a single ideal well location for each abstraction site. In reality, the abstraction takes place in most cases by means of a series of wells spread over a certain area. A maximum distance between the wells within a well field of 500 m or more is not unusual. The ideal well location in LGM is supposed to be at the "centre of gravity" of the real well locations, the well rate in LGM being the total of all well rates.

Though not required, in our case the finite-element grid was generated without taking into account the well locations. The well location can be anywhere in the model, i.e. also inside a grid element. The program automatically allocates the groundwater abstraction rates to the grid nodes composing the element in which an abstraction site is located.

The parameterization of groundwater abstractions is described in Pastoors (1992). As already mentioned, this study is based on the existing model Utrecht. In the latter, the maximum-permitted abstraction rates were used, as in effect for the year 1988. The year 1988 was chosen because at the time when the model Utrecht was developed it was the year for which the most recent abstraction data were available for the entire country.

Table 3.1 gives an overview of the 10 groundwater abstraction sites selected for this study (cf. Figure 3.2). The so-called LAC-number is an identification code name (a text string) for an abstraction location. The LAC-numbers were assigned to abstraction locations in 1970s and 1980s by the RIVM Laboratory for Inorganic Chemistry (LAC), throughout the execution of chemical analyses of abstracted groundwater. The so-called LGM-number is the sequence number of an abstraction in the LGM database.

Table 3.1 Overview of groundwater abstraction sites considered in this study.

Abstraction site	Rate (m^3 $year^{-1}$)	LAC/LGM no.	IWL no.	Aquifer(s)	Phreatic/semi-confined
Soest	1.2×10^6	177 / 746	52,53	2,3	phreatic
Soestduinen	9.0×10^6	178 / 747	60,61	1,2	phreatic
Amersfoort-Berg	3.0×10^6	11 / 726	62	1	phreatic
Bilthoven	2.0×10^6	161 / 738	67,68	1,2	phreatic
Groenekan	10.0×10^6	165 / 745	69,70	2,3	semi-confined
Beerschoten	8.0×10^6	160 / 737	75	2	phreatic
Zeist	5.0×10^6	180 / 749	80	2	phreatic
Bunnik	3.5×10^6	162 / 739	108,109	2,3	semi-confined
Driebergen	0.9×10^6	164 / 743	110	2	mixed phreatic and semi-confined
Doorn	1.6×10^6	32 / 742	113	1	phreatic

As the geohydrological system of LGM, version 3, consists of four aquifers, the number of well screens at an abstraction site can vary between 1 and 4. This is because in most cases the actual well screen cuts across a number of aquifers. The IWL numbers are the internal LGM sequence numbers, ranging from 1 to *NWL*, where *NWL* is the total number of model-input well screens for

all abstraction sites and all aquifers. As an example, the abstraction LAC-no.=165, Groenekan, takes place in two aquifers, 2 and 3. The *NWL* model-input well screens are generated by the grid generator LGMGRID (also named EGGO), as a part of the data discretisation procedure.

Two types of groundwater abstractions can be distinguished, namely:

- **groundwater abstractions from *phreatic systems***, i.e. where the hydraulic resistance of semi-confining layers (aquitards) above the well screens is relatively low. Those abstractions are characterized by:
  (a) a more or less contiguous (non-fragmented) capure zone;
  (b) a capture zone that starts at the abstract location, and has a rather smooth (convex) shape (often unilaterally oriented or ellips-shaped) , and
  (c) a relatively high fraction of the area where travel time is less than 25 years.

- **groundwater abstractions from *semi-confined systems***, i.e. where the hydraulic resistance of semi-confining layers (aquitards) above the well screens is relatively high. Those abstractions are characterized by:
  (d) a discontinuous (fragmented) capture zone, in most cases, e.g. as patches of sub-capture zones;
  (e) a capture zone that sometimes is strongly dislocated with respect to the abstraction location itself, and has and has an uneven (concave, rugged) shape, and
  (f) a relatively high fraction of the area where travel time is greater than 50 years.

In our case, there are two distinct semi-confined abstractions, namely Groenekan and Bunnik, both abstracting from aquifer 2 and 3. An aquitard occurs both above aquifer 1 (see Figure 3.11) and between aquifer 1 and 2 (see Figure 3.13).

The groundwater abstraction Driebergen (well screen only in aquifer 2) is a mixed phreatic-semi-confined case. No aquitard occurs above aquifer 1 (see Figure 3.11). However, an aquitard can be distinguished (though not markedly developed in spatial extent) between between aquifer 1 and 2 (see Figure 3.13).

The remaining seven groundwater abstractions from Table 3.1 are typically phreatic, i.e. without any major semi-confining layer existent above the well screen.

3.2.5 Groundwater recharge

The travel times are affected not only by the amount of abstracted groundwater but also by the magnitude of groundwater recharge rate. The groundwater recharge is the net flux from the root zone into the saturated groundwater. Important are both the magnitude and the spatial distribution.

The procedure that was followed for the calculation of groundwater recharge rate for the model Utrecht is described in Pastoors (1992, section 4.3). The recharge rate is calculated as the difference

between precipitation and actual evapotranspiration, the latter being highly dependent on the land-use. The recharge rate assessment was based on the meteorological data for the year 1988. This year was used to maintain consistency with groundwater abstraction data, which are also specified for that year.

Since it was not considered essential, the spatial map of the groundwater recharge, input for LGMSAT, is not shown here. Referring to Kovar *et al.* (1998, section 2.4) the mean value at the model Utrecht is around 1 mm day^{-1}. Kovar *et al.* (1998 section 2.4, Table 2.2) also mention the statistics of the groundwater rechaage rate derived by a different approach, namely as the 1961-1990 average rate. The variance of this 1961-1990 average is about 0.17, implying the standard deviation of 0.41 mm day^{-1}.

As is described in section 3.5, the groundwater recharge rate, though it was not inverse-method calibrated, was included as one of the uncertain parameters affecting the travel times.

3.3   Calibration and parameter reliability

The reason for selecting the Utrecht model for the pilot study was that, in a previous study (Leijnse *et al.*, 2002b), this model was calibrated by means of an inverse method, implemented in the module LGMCAL. The development of the calibration procedure is described in Leijnse *et al.* (2002a, 2002b). Calibration is done on the basis of a large number of measured groundwater heads, using a weighted least-squares approximation, and including prior information on the hydraulic parameters to be estimated. The object function minimized was given by:

$$F = \left(h_m - h_c\right)W_h\left(h_m - h_c\right)^T + \left(P_p - P_c\right)W_P\left(P_p - P_c\right)^T$$

where h_m and h_c are, respectively, the measured and calculated groundwater heads, and P_p and P_c are the prior estimates and the optimized values of the hydraulic parameters. W_h and W_P are the weight matrices for the observed heads and the parameters, respectively.

The calibration study by Leijnse *et al.* (2002b) resulted in a model where 16 parameters were optimized. Subsequently, in a follow-up work, the model Utrecht was further elaborated upon (personal comm. A. Leijnse) by extending the number of calibration parameters and using a finer finite-element grid. The resulting model (not published), used for our study, consists of 22 optimized parameters (Table 3.2, data taken from LGMCAL output file *param.opt*).

Hence at the onset of the current study the following calibrated parameters were available:
– optimized drainage and infiltration resistance for the surface-water-groundwater interaction (8 items), parameters **W1(30)** to **W1(1070)**;
– optimized multiplication factor for hydraulic resistance of aquitard 0 (2 items), parameters **B0(-250)** and **B0(0)**;

– optimized multiplication factor hydraulic resistance of aquitard 1 (2 items), parameters **B1(12)** and **B1(18)**;
– optimized multiplication factor for transmissivity of aquifer 1 and 2 (10 items), parameters **ALF1*i*** and **ALF2*i***.

The calibration also produced the covariance for those 22 parameter items.

Table 3.2 Optimal parameter values and the standard deviation of their log values, for 22 calibrated parameters.

Code	Optimal value	St.Dev.	Description
W1(30) [a]	1.3467E+02	3.4415E-01	Drainage resistance of topsystem, in areas where Gt=3 (Fig. 3.9)
W1(40) [a]	1.9316E+02	5.3269E-01	Drainage resistance of topsystem, in areas where Gt=4 (Fig. 3.9)
W1(50) [a]	3.0969E+01	4.3005E-01	Drainage resistance of topsystem, in areas where Gt=5 (Fig. 3.9)
W1(70) [a]	1.8421E+03	7.0958E-01	Drainage resistance of topsystem, in areas where Gt=7 (Fig. 3.9)
W1(1030)	1.4426E+01	6.1059E-01	Drainage=infiltration resistance of topsystem, in areas where Gt=3 (Fig. 3.9)
W1(1040)	3.4491E+02	7.1760E-01	Drainage=infiltration resistance of topsystem, in areas where Gt=4 (Fig. 3.9)
W1(1050)	3.2450E+02	4.0919E-01	Drainage=infiltration resistance of topsystem, in areas where Gt=5 (Fig. 3.9)
W1(1070)	1.9801E+02	5.8501E-01	Drainage=infiltration resistance of topsystem, in areas where Gt=7 (Fig. 3.9)
B0(-250)	3.5991E+02	5.1016E-01	Multipl. factor for hydr. resistance of aquitard 0, in zone coded -250 (Fig. 3.10)
B0(0)	2.0916E+01	3.8847E-01	Multipl. factor for hydr. resistance of aquitard 0, in zone coded 0 (Fig. 3.10)
B1(12)	2.8366E+01	1.9006E-01	Multipl. factor for hydr. resistance of aquitard 1, in zone coded 12 (Fig. 3.12)
B1(18)	5.2960E+00	1.7133E-01	Multipl. factor for hydr. resistance of aquitard 1, in zone coded 18 (Fig. 3.12)
ALF11	4.0645E+00	2.9287E-01	Multiplication factor for transmissivity 1, at pilot point 1 (Fig. 3.14)
ALF12	1.5006E+00	3.0905E-01	Multiplication factor for transmissivity 1, at pilot point 2 (Fig. 3.14)
ALF13	6.0463E-01	3.1035E-01	Multiplication factor for transmissivity 1, at pilot point 3 (Fig. 3.14)
ALF14	5.9590E-01	3.1606E-01	Multiplication factor for transmissivity 1, at pilot point 4 (Fig. 3.14)
ALF15	7.6871E-01	3.0844E-01	Multiplication factor for transmissivity 1, at pilot point 5 (Fig. 3.14)
ALF21	9.4404E-01	3.0872E-01	Multiplication factor for transmissivity 2, at pilot point 1 (Fig. 3.14)
ALF22	2.5447E+00	2.6050E-01	Multiplication factor for transmissivity 2, at pilot point 2 (Fig. 3.14)
ALF23	9.0179E-01	3.2211E-01	Multiplication factor for transmissivity 2, at pilot point 3 (Fig. 3.14)
ALF24	7.5378E-01	3.1592E-01	Multiplication factor for transmissivity 2, at pilot point 4 (Fig. 3.14)
ALF25	3.7208E+00	9.6263E-02	Multiplication factor for transmissivity 2, at pilot point 5 (Fig. 3.14)

[a] no infiltration in this area, i.e. infiltration resistance not relevant

The correlation matrix of the 22 optimized parameters is included in Appendix 2. In the following, background information is given about the optimized parameters.

Drainage and infiltration resistance of topsystem-flux relationship

Among the calibrated parameters are the drainage and infiltration resistance **W1** [days] of the topsystem-flux relationship. The shape of the topsystem-flux relationship is shown in Figure 3.8. The resistance values are assumed constant in several sub-areas. These are areas belonging to four classes of Gt-value, Gt being the index number of the groundwater-depth class. In our case, we used Gt= 3, 4, 5 and 7. The zoning used for the calibration is depicted in Figure 3.9. Two conditions were distinguished:
– the topsystem-flux relationship can only drain (Figure 3.8(b)). The calibration-derived drainage resistances are named **W1(30)**, **W1(40)**, **W1(50)** and **W1(70)**;

– the topsystem-flux relationship can both drain and infiltrate, with drainage and infiltration resistances being identical values (Figure 3.8(a)). The calibration-derived resistances are named **W1(1030)**, **W1(1040)**, **W1(1050)** and **W1(1070)**.
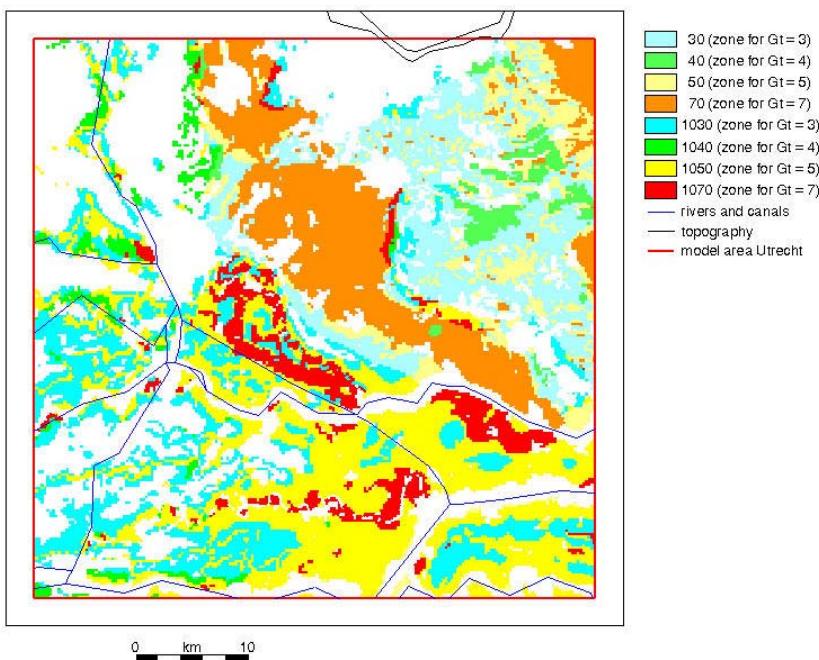
Figure 3.9 Zoning used for calibration of (a) drainage resistance W1 of topsystem (no infiltration in those areas), zone codes 30 (Gt=3), 40 (Gt=4), 50 (Gt=4), 70 (Gt=4), and (b) drainage and infiltration resistance W1 (one value, identical for drainage and infiltration), zone codes 1030 (Gt=3), 1040 (Gt=4), 1050 (Gt=5), 1070 (Gt=7)

Hydraulic resistance of aquitard 0

The next calibrated parameter was the hydraulic resistance of aquitard 0, i.e. the layer overlying the top aquifer. The spatially distributed value of the hydraulic resistance is based on an estimate of the thickness D of the aquitard. Based on available geological information the thickness D is assumed to be a known value. The hydraulic resistance RPL [days] of aquitard 0 then follows from:

$$RPL(x,y) = \mathbf{B0(x,y)} \times D(x,y)$$

where D [m] is the known layer thickness, and **B0** [day m^{-1}] is a multiplication factor that depends on the type of the geological formation. Note that **B0**, being the resistance per one metre thickness, is the inverse value of hydraulic conductivity, the smaller the hydraulic conductivity in vertical direction of the aquitard, the bigger the value of **B0**. The values of **B0**, constant within each of the **B0**-related sub-areas, have been estimated by means of calibration. The value of **B0** was assumed constant in two sub-areas, shown in Figure 3.10. The two areas are labelled "-250" and "0". Note in Figure 3.10 the white-coloured area, indicating that the layer is missing. In the area labelled "-250" the calibration started with the prior estimate **B0**=250, and resulted in the optimized value **B0**=359.91. In the area labelled "0" the calibration started with the prior estimate **B0**=100, and resulted in the optimized value **B0**=20.916.

The resulting spatial pattern of hydraulic resistance *RPL* is depicted in Figure 3.11. Note also the grey shaded area, labelled "no semipervious layer". At those parts of the model the aquitard is missing and a dummy small value [1 day] of hydraulic resistance *RPL* was used.
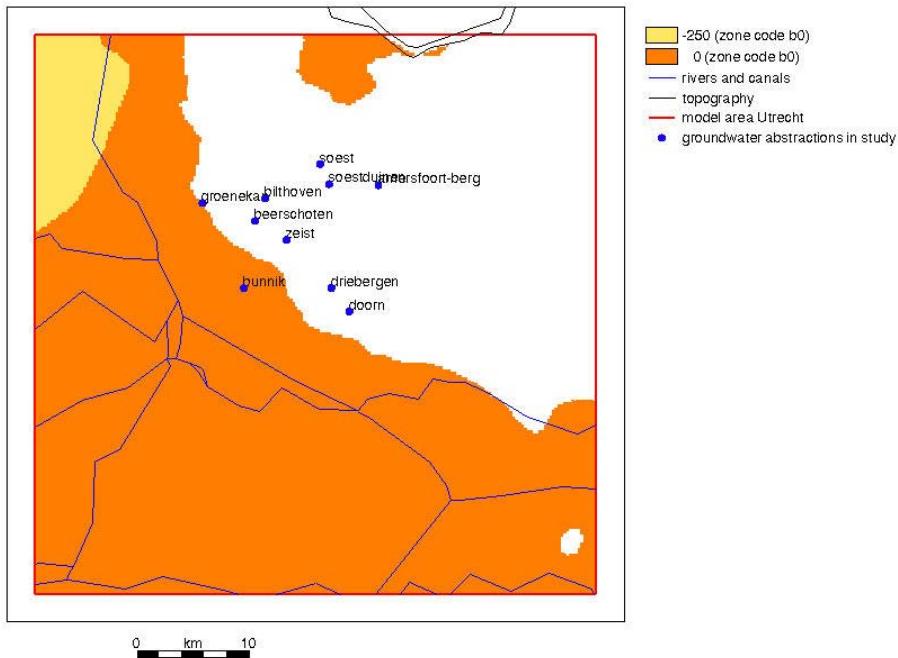
Figure 3.10 Zoning used for calibration of multiplication factor B0, for hydraulic resistance of aquitard 0.
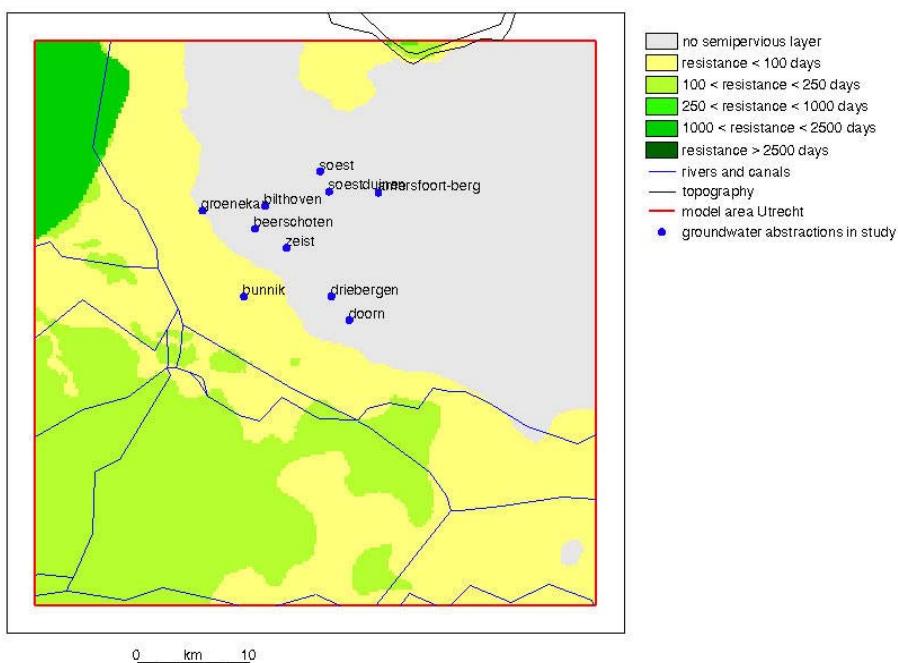
Figure 3.11 Hydraulic resistance RPL [days] of aquitard 0, i.e. layer above aquifer 1.

Hydraulic resistance of aquitard 1

The procedure for the calibration of hydraulic resistance of aquitard 1, i.e. the layer between aquifer 1 and 2, is similar to the procedure for hydraulic resistance of aquitard 0. The hydraulic resistance, denoted *RA1* [days], follows from:

$$RA1(x,y) = \mathbf{B1(x,y)} \times Depth(x,y) \times D(x,y)$$

where *Depth* [m] is the known depth of the layer below ground surface, *D* [m] is the known layer thickness, and **B1** [day m^{-2}] is a multiplication factor that depends on the type of the geological formation. Note that **B1** relates to the value of hydraulic conductivity, the smaller the hydraulic conductivity in vertical direction of the aquitard, the bigger the value of **B1**. The values of **B1**, constant within each of the **B1**-related sub-areas, have been estimated by means of calibration. The value of **B1** was assumed constant in two sub-areas, shown in Figure 3.12. The two areas are labelled "12" and "18". Note in Figure 3.12 the white-coloured area, indicating that the layer is missing. In the area labelled "12" the calibration started with the prior estimate **B1**=15, and resulted in the optimized value **B1**=28.366. In the area labelled "18" the calibration started with the prior estimate **B1**=2.5, and resulted in the optimized value **B1**=5.296.

The resulting spatial pattern of hydraulic resistance *RA1* is depicted in Figure 3.13. Note also the grey shaded area, labelled "no semipervious layer". At those parts of the model the aquitard is missing and a dummy small value [1 day] of hydraulic resistance *RA1* was used.
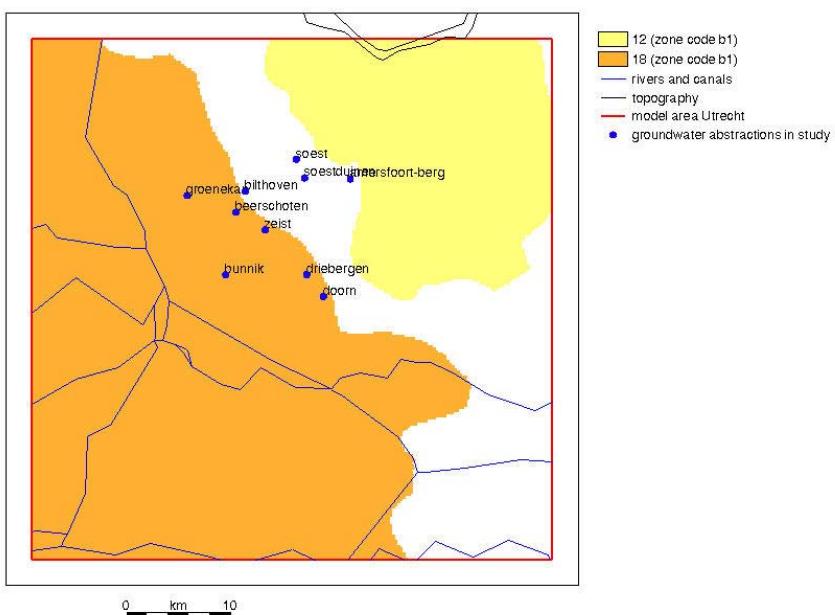
Figure 3.12 Zoning used for calibration of multiplication factor B1, for hydraulic resistance of aquitard 1.
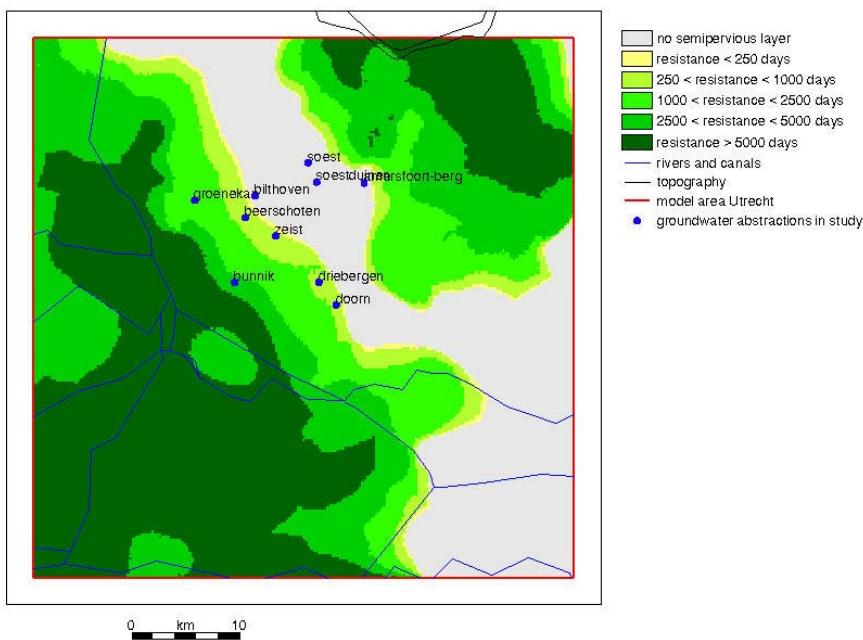
Figure 3.13 Hydraulic resistance RA1 [days] of aquitard 1, i.e. layer between aquifer 1 and 2.

Transmissivity of aquifer 1 and 2

Finally, the last set of calibrated parameters regards the aquifer transmissivities, for aquifer 1 and 2. The point of departure for the definition of the spatially distributed transmissivities of the aquifers are the prior transmissivity estimates, based on available geological information. These prior estimates are also spatially distributed. The transmissivity in each point of the model domain is then given by:

$$T(x,y) = ALF(x,y)\ T_p(x,y)$$

where T_p is the prior estimate of the transmissivity [m^2 day^{-1}] and *ALF* is a multiplication factor [-]. The spatially distributed multiplication factor *ALF* is defined on the basis of the four triangles formed by the diagonals in the rectangle (green line in Figure 3.14). Given the values of this multiplication factor in the five pilot points (the corner points and the midpoint of the rectangle, the value of ALF in each point of the model area is determined by a linear interpolation. The values of the multiplication factor in the pilot points, **ALF11** to **ALF15** for aquifer 1, and **ALF21** to **ALF25** for aquifer 2, are optimized by means of the calibration procedure.
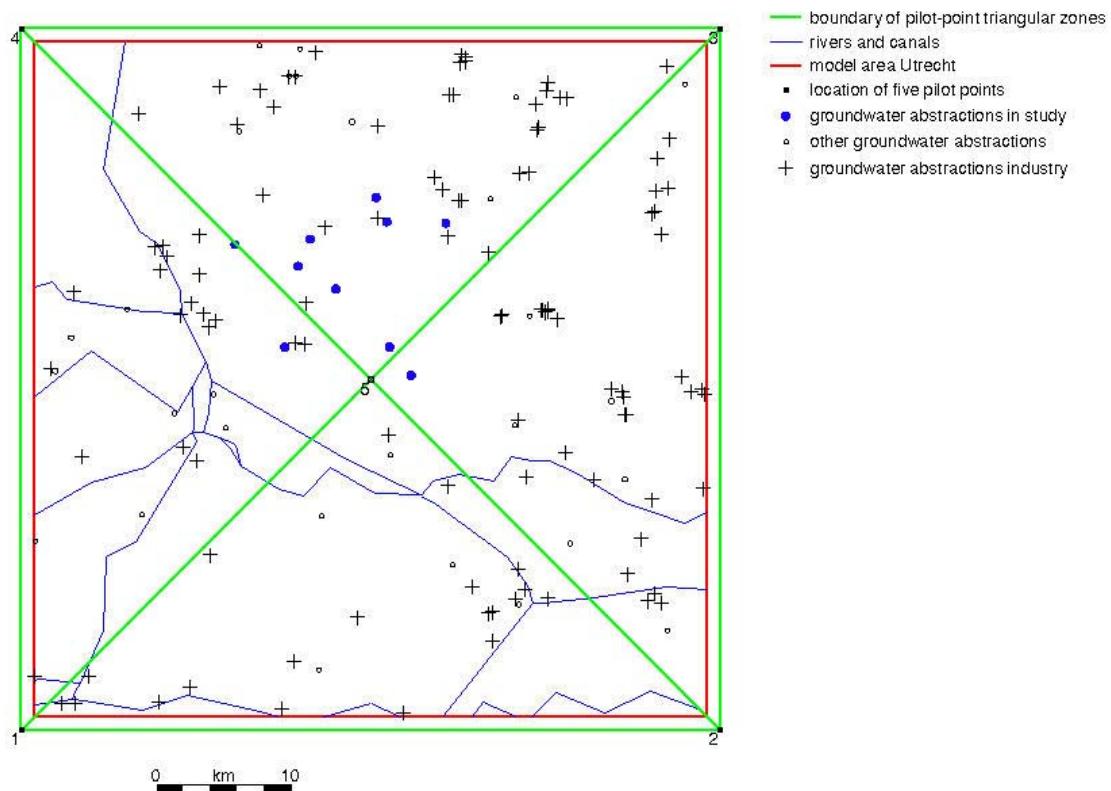
Figure 3.14 Location of five pilot points used for calibration of transmissivity multiplication factors: (a) factors ALF11 through ALF15 for transmissivity of aquifer 1, and (b) factors ALF21 through ALF25 for transmissivity of aquifer 2.

Solve groundwater potential problem for optimized model-input parameters

Once the calibration is completed (model parameters are optimized), the module LGMSAT can be applied for modelling the groundwater potential problem, i.e. for the simulation of groundwater heads and fluxes. Figure 3.15 depicts the simulated groundwater head in aquifer 1. Figure 3.16 shows the flux at the top of aquifer 1, this flux being the total of (a) user-specified groundwater recharge rate and (b) the calculated topsystem flux, the groundwater-head-dependent flux between groundwater in aquifer 1 and small-scale surface waters.
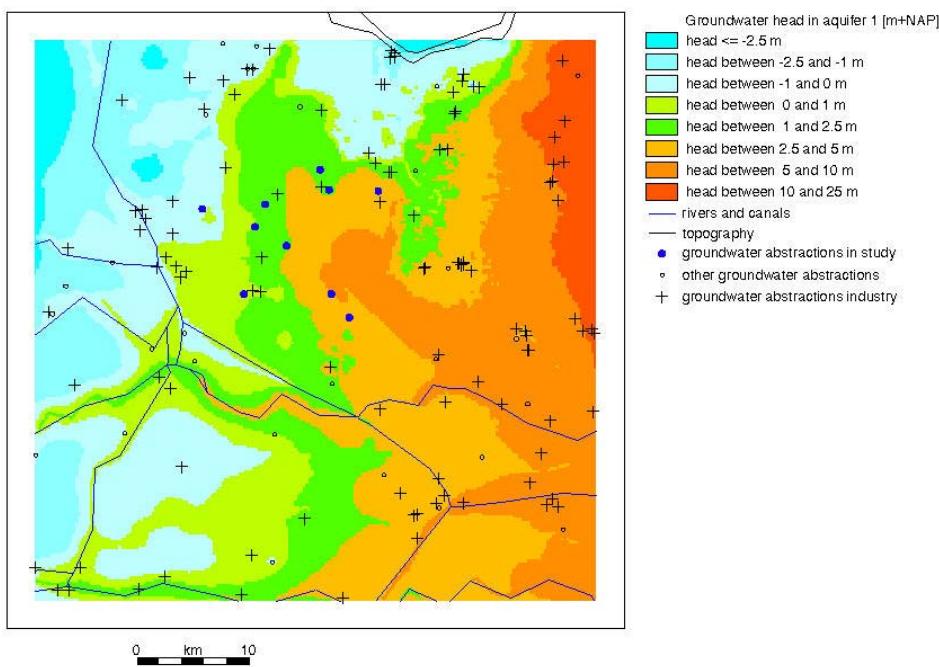
Figure 3.15 Groundwater head (m above m.s.l.) in aquifer 1, simulated with calibrated LGMSAT model.

In Figure 3.16 one can see the major recharge area (orange, flux between 0.5 and 2.5 mm day[-1]) at the elevated sandy hills of the Utrechtse Heuvelrug, and the seepage-flux areas (blue) in the polders in the west-southwest corner of the model.
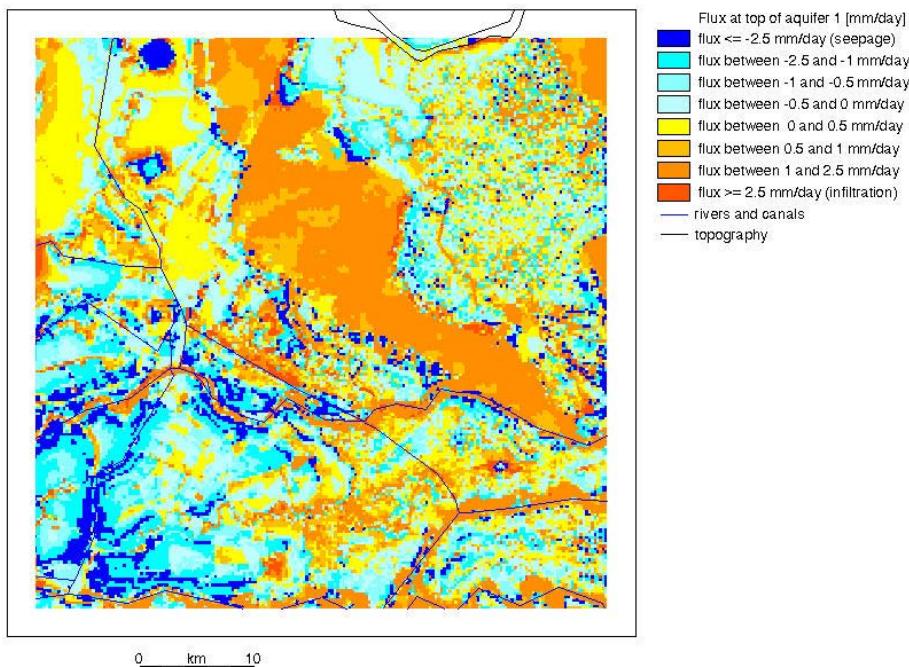
Figure 3.16 Flux (mm day[-1]) at top of aquifer 1, simulated with calibrated LGMSAT model.

The groundwater potential problem solved with the 22 optimal parameter values is the point of departure for the assessment of the travel-time reliability (section 3.5).

3.4   Modelling pathlines and travel times

The modelling method used for assessment of travel-time reliability, as implemented in the module LGMLUC, is based on a pair-wise repetitive solution of (a) the groundwater potential problem (module LGMSAT) and (b) the pathlines and travel times (module LGMFLOW).

The module LGMFLOW was used to calculate pathlines and travel times in the forward- tracking mode, i.e. when particles move in the same direction as groundwater flow. In our case, the term particle denotes a certain –very small– volume of groundwater. The pathlines start at the top of the system from a series of so-called *starting points*.

Two situations can be distinguished with regard to the z-level of the starting points:
− the z-level refers to the top of the system, i.e. the pathline tracking starts at ground level. This occurs mostly in polder areas with shallow water tables, where the groundwater table is within one metre below ground level;
− if groundwater head φ_1 in the top aquifer (aquifer 1) is below the top of this aquifer, the z-level of the starting point is set equal to φ_1. This usually occurs in phreatic water-table areas, where the top aquifer is not overlaid by a semi-permeable layer (in LGM terminology: aquitard 0).
In our study most starting points were located in the sandy-soil region of the Utrechtse Heuvelrug, where the starting z-level is equal to the phreatic water table.
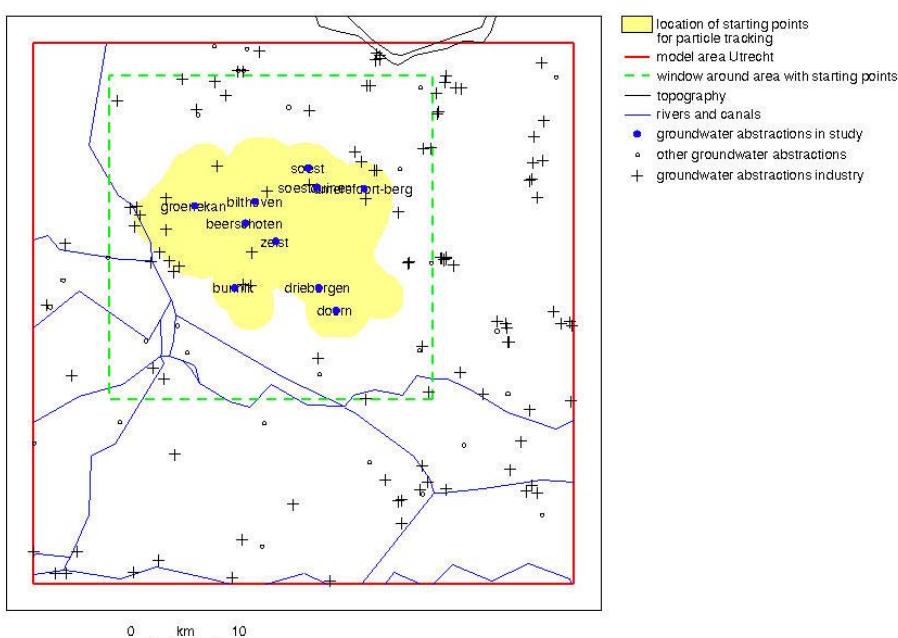
Figure 3.17 Location of starting points for forward particle tracking. Square-shaped frame (dashed line) is used for presentation in Figures 3.18, 3.19, 3.20, and 3.22.

Figure 3.17 depicts the area (yellow coloured) where the starting points are located. The boundary of this area is a 2-km hull around the capture zone for the 10 selected groundwater abstraction sites. Use is made of capture zones produced by Kovar *et al.* (1998). The starting points constitute a regular mesh, with point distances of 200 by 200 m, with the total number of starting points at 8020.

The travel time depends on the magnitude of the groundwater flux that is obviously variable in the three-dimensional space. However, it also depends on the effective porosity, the part of the grain pore space available for the passage (seepage) of groundwater. The input value of effective porosity was specified as:
- the effective porosity of aquifers, denoted by PF1 and a constant value (0.3) for all aquifers;
- the effective porosity of aquitards, denoted by PA0 and a constant value (0.1) for all aquitards;

Figure 3.18 presents the so-called destination of the pathlines' end-point, prepared using the LGMLUC output file *ms*fes_ict_ddg.unf.00. The suffix "00" indicates a file produced for the 22 optimal parameter values, as discussed in the previous section. Please allow us here to make the following notes:
- The well destination of pathlines (shown in orange) considers <u>all wells</u> acting in the system, i.e. (a) the 10 wells for drinking-water production used in this study, (b) all other wells for drinking-water production, and (c) all wells for industrial purposes.
- The destination of particles can be anywhere in the entire 50×50 km model, i.e. not only inside the area where the particles have originated.
- If a starting point is in a seepage area (dark green) then the particle tracking can obviously not be carried out. The seepage areas (upward flow, drainage) are shown in Figure 3.16.
- The maximum travel time for tracking was 5000 years. Most of the starting points where the tracking time exceeded this maximum are located along the edge of a capture zone, where stagnant zones occur (due to a negligible horizontal flow component).
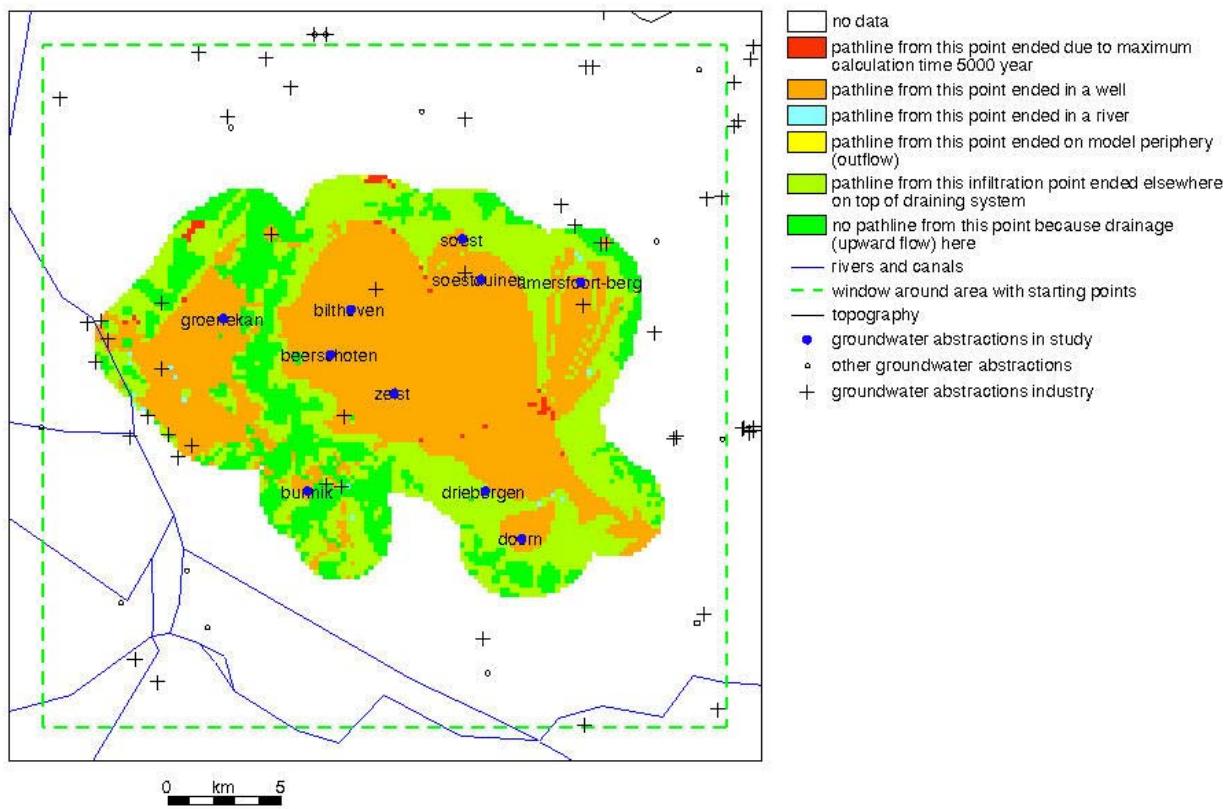
Figure 3.18 Destination of forward-tracking pathlines that started in area shown in Figure 3.17.

Figure 3.19 depicts the capture zone of the 10 wells considered for this pilot study. For detailed information about the 10 wells we refer to Table 3.1. The picture is prepared using the LGMLUC output file *ms*fes_iwl_ddg.unf.00, IWL being the sequence numbers of the well where a pathline has ended "her journey through the underground".

A capture zone of a well is the area where all abstracted groundwater originates. Usually, most of the area is replenished by groundwater recharge and infiltration from small-scale surface waters (ditches). However, though not relevant in our case, the capture zone can also include the bedding of rivers and canals, where river water seeps downwards through the river bottom into groundwater and subsequently flows towards the well.
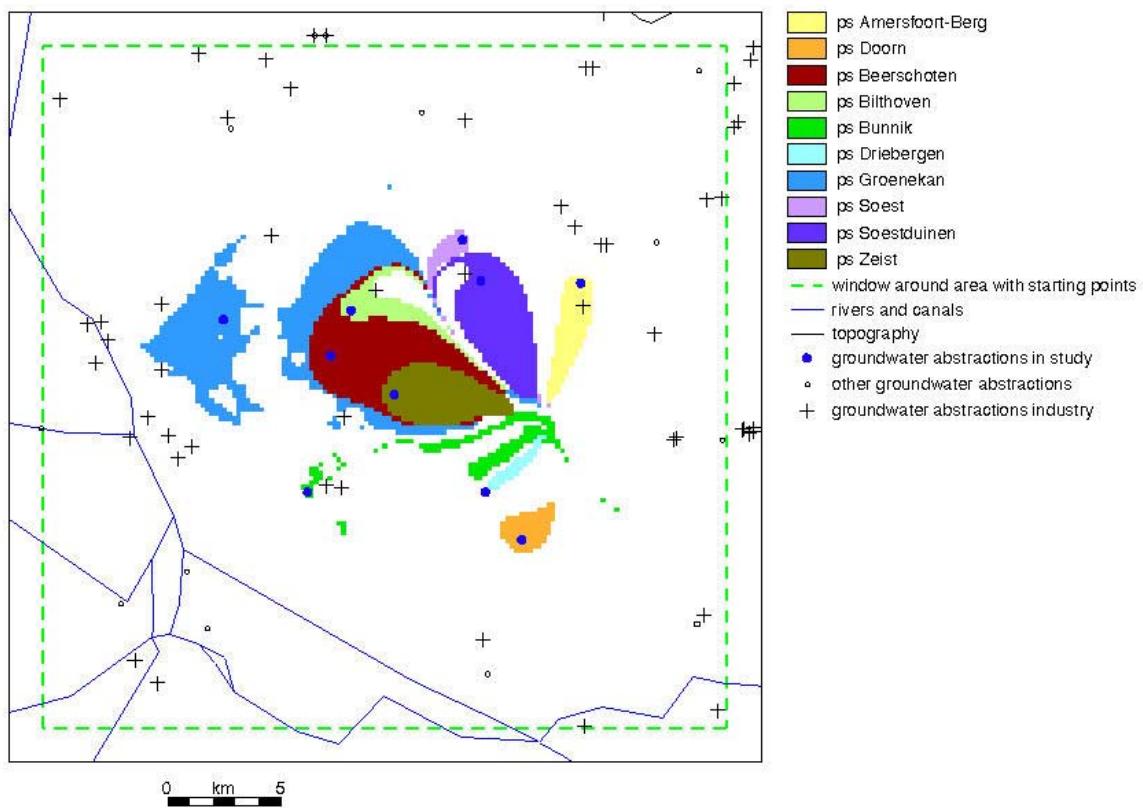
Figure 3.19 Capture zones for 10 groundwater abstractions considered in this study.

In our case, example of a phreatic groundwater abstraction site are Zeist and Beerschoten. The groundwater recharge takes place on the top of sandy-soil ridge Utrechtse Heuvelrug. The depth of the phreatic water table at the two locations varies between 3 and 5 m below ground surface. An example of the opposite is found in the semi-confined groundwater abstraction sites, Groenekan and Bunnik, where the top aquifer, aquifer 1, is overlaid by a peat-clay aquitard. In the latter case, the source of water for the abstraction is a dense system of ditches with controlled surface-water level. The groundwater in those polder areas is shallow, with a depth between 0.3 and 1 m.
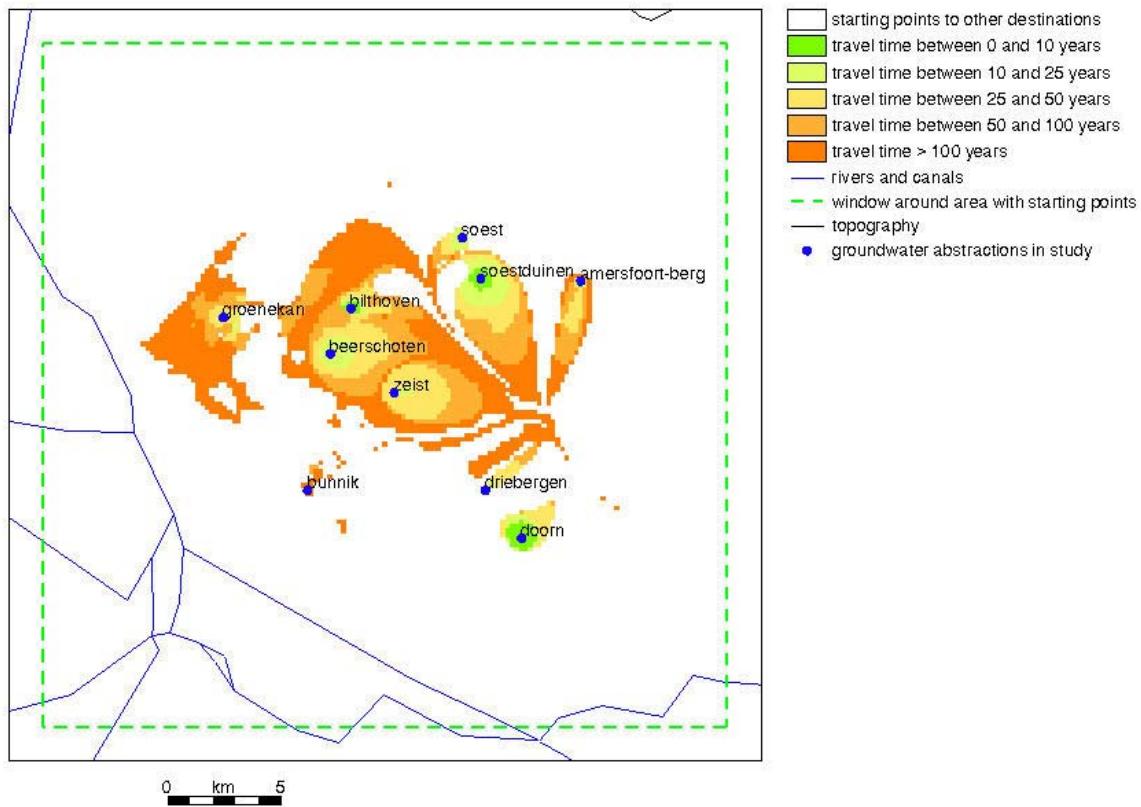
Figure 3.20 Travel time (years) for 10 groundwater abstractions considered in this study.

Figure 3.20 depicts the travel time to the 10 wells selected for this study. The figure was prepared using the LGMLUC output file *ms*fes_tti_ddg.unf.00. The travel time is the time between the moment of infiltration and the moment when the water particle is captured by a well screen. As previously noted, most starting points are located in the sandy-soil region of the Utrechtse Heuvelrug; therefore their starting z-level is the same as the phreatic water table.

Modelling pathlines and travel times forms part of the method for assessing the travel-time reliability as discussed in the next section.

3.5   Modelling reliability of travel times

The modelling approach used here to assess the reliability of travel times is based on the First-Order Second-Moment (FOSM) method for uncertainty propagation analysis. The method is incorporated in the program LGMLUC, one of the modules of LGM, the RIVM's Netherlands Groundwater Model. Instead of using the term "reliability" here, we might also refer to "certainty", "uncertainty" or "probability of occurrence".

While running LGMLUC, the modules LGMSAT and LGMFLOW are pair-wise processed *NP*+1 times, where *NP* is the number of uncertain parameters affecting the travel-time reliability. Obviously, the reliability of travel times decreases, or if you prefer, uncertainty increases with increasing uncertainty in the model-input parameters, in other words, with variance in these input parameters. The user's guide for LGMLUC, given in Appendix 1, describes the method, the input and the output.

A prerequisite for modelling the reliability of travel times is the availability of information on the uncertainty of the relevant model-input parameters. In the pilot-study application of LGMLUC reported here (Utrecht Model), the following uncertain parameters were assumed to affect the reliability of travel times:

(1) drainage resistance for the surface-water-groundwater flux relationship in areas where only drainage can take place. The shape of the relationship is illustrated in Figure 3.8(b);

(2) drainage and infiltration resistance for the surface-water-groundwater flux relationship in areas where both drainage and infiltration can take place. In our case we assumed a drainage resistance equal to infiltration resistance. The shape of relationship is illustrated in Figure 3.8(a);

(3) hydraulic resistance of aquitard 0 (the layer overlying aquifer 1);

(4) hydraulic resistance of aquitard 1 (the layer between aquifer 1 and 2);

(5) transmissivity of aquifer 1;

(6) transmissivity of aquifer 2;

(7) effective porosity of aquifers;

(8) effective porosity of aquitards;

(9) groundwater recharge rate.

Each of these hydrological parameters is assumed to be log-normally distributed.

Parameters (1) through (6) listed above were calibrated by means of an inverse method. The procedure and the results are discussed in section 3.3. The calibration resulted in 22 optimized parameter values. The optimal parameter values and their standard deviations are listed in Table 3.2. The correlation matrix for the 22 parameters is included in Appendix 2.

The remaining three parameters, (7), (8) and (9), involve multiplication factors:
– multiplication factor **F_PF1** for the effective porosity of the aquifers (PF1);
– multiplication factor **F_PA0** for the effective porosity of the aquitards (PA0);
– multiplication factor **F_QRE** for the groundwater recharge (QRE).

Each factor is used to multiply the entire spatially variable field of the relevant parameter (PF1, PA0 or QRE). These multiplication-factor parameters were not defined by means of calibration, but added to the list of parameters to be involved in the analysis. The mean value of the multiplication factor is, for all three parameters, equal to 1. The variance s^2 of the multiplication factors **F_PF1**, **F_PA0** and **F_QRE** was assumed by expert judgement, amounting to $s^2 = 4.3 \times 10^{-4}$, or S$tDev$ = 0.020736. In fact, effective porosity of aquifers and aquitards cannot even be calibrated by using observed groundwater heads. By the way, though theoretically possible, the current version of LGMCAL does not allow for calibration of groundwater recharge rate. The approach followed to

assess the variance $s^2 = 4.3 \times 10^{-4}$ is elaborated on in section 2.5.2. The variance is chosen so that 95% of the values of **F_PF1**, **F_PA0** and **F_QRE** fall within in the interval:

$$\frac{X_0}{1.1} < X < 1.1\,X_0$$

where X denotes any of these three multiplication-factor parameters and X_0, the average parameter value (geometric mean). Table 3.3 summarizes the information for the three uncertain model-input parameters added.

Table 3.3 Optimal parameter values and the standard deviation of their log values, for three expert-judgment parameters (PF1, PA0 and QRE) added.

Code	Optimal value	St.Dev.	Description
F_PF1	1.	0.020736	Multipliplication factor for effective porosity of aquifers (PF1)
F_PA0	1.	0.020736	Multiplication factor for effective porosity of aquitards (PA0)
F_QRE	1.	0.020736	Multiplication factor for groundwater recharge rate (QRE)

The three extra parameters were added to the 22 calibrated parameters. The resulting input data set, which, with its 25 uncertain parameters (*NP*=25), would be involved in the analysis, is given in Appendix 3. Since no information was available and no educated guesses could be made, we have assumed that no correlation exists between the three parameters added and the 22 parameters calibrated. Note again that all parameters are assumed to be log-normally distributed. Once the information about the parameter uncertainty is available, the LGMLUC program can be used to produce the travel-time reliability.
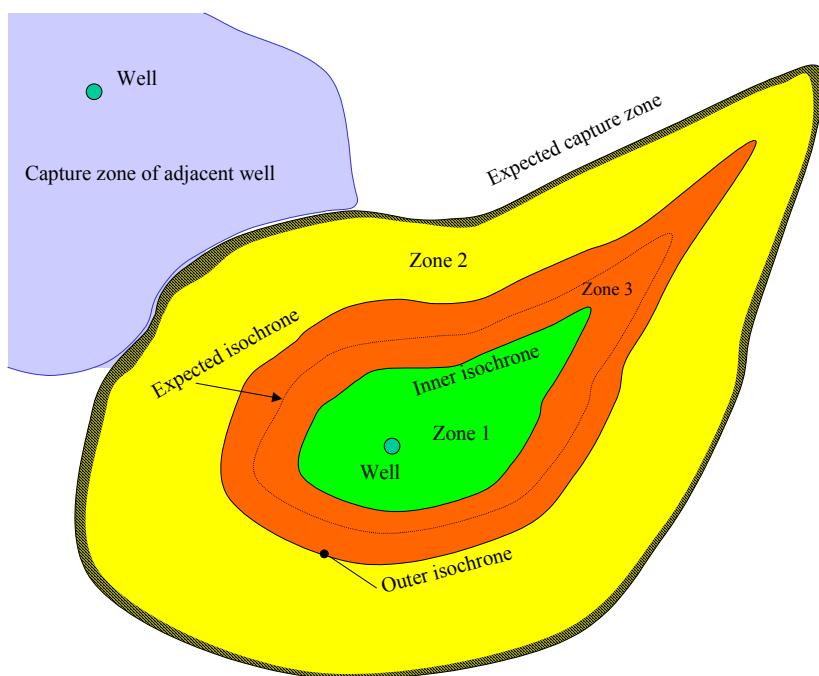
Figure 3.21 LGMLUC output: schematic representation of (1) expected capture zone, and (2) "inner", "outer" and "expected" isochrones. LGMLUC output is not possible in yellow hashed area on the inner side of the expected capture-zone boundary.

Figure 3.21 depicts the basic output features of LGMLUC. It gives a hypothetical example for a 50-year isochrone. Since it is a key factor for understanding the results of this study, we will reiterate part of the text taken from section 2.3. We have adopted the terms "inner", "outer" and "expected" isochrones to indicate the isochrone of $T_{97.5}(x,y)$, $T_{2.5}(x,y)$ and $T_{50}(x,y)$, respectively. With these two isochrones in the same plot, the total capture area is divided into three zones:

(1) A zone around the well bounded by the "inner" 50-year isochrone (green colour). This represents the area where the probability that a water particle will reach the well within 50 years is higher than, for example, 0.95.

(2) A zone outside the "outer" 50-year isochrone (yellow colour). A particle starting in this zone reaches the well within 50 years with a probability of less than, for example, 0.05. Re-wording this, we could say that, with a probability of 0.95 or higher, the particle will **not** reach the well within 50 years. Strictly speaking, the outer limit of the yellow zone is not the expected capture-zone boundary, but rather the outer occurrence probability contour (not shown in the picture) of the capture zone, which is larger than the capture zone.

(3) A transition zone (orange colour). Here, the probability to reach the well within 50 years gradually decreases from 0.95 to 0.05, going from "inner" to "outer" isochrone. Alternatively, it may be said that this zone contains the "true" 50-years isochrone with a probability of 0.90. As the picture shows, the zone contains the "expected" isochrone.

The algorithm used in LGMLUC is based on a linear approach. If we indicate the ^{10}log parameter values as P_i, $i=1,NP$, the variance σ_T^2 of the travel time is given by:

$$\sigma_T^2 = \sum_{i=1}^{NP}\left(\sum_{j=1}^{NP}\frac{\partial T}{\partial P_i}\frac{\partial T}{\partial P_j}Cov\left(P_i,P_j\right)\right) \tag{3.3}$$

Obviously, σ_T is the standard deviation of the travel time T. The above expression (3.3) is given for a normally distributed T. Nevertheless, a similar expression (not shown here) was used for this pilot study by considering the ^{10}log value of the travel time and replacing T with ^{10}log(T). Note that LGMLUC can handle both normally distributed T and ^{10}log(T) as an input option for the user.

LGMLUC calculates in each of the pathline starting points: (a) the "expected" isochrone (i.e. 50 percentile travel time), and (b) the "inner" and "outer" isochrones. The "expected" travel times are already shown in Figure 3.20. The "inner" and "outer" isochrones are calculated by using the travel time variance σ_T^2. However, due to the numerical (discrete) nature of LGMLUC, it is not possible to calculate σ_T^2 –and consequently, neither the "inner" or "outer" isochrone – in starting points in a narrow-band area along the inner side of the capture zone boundary. The relevant area is indicated in hashed yellow in Figure 3.21. The reason for not being able to calculate σ_T^2 is that, in a given pathline starting point, neither of the sensitivity derivatives, $\partial T/\partial P_i$ or $\partial T/\partial P_j$, can be calculated for each of the NP parameters (NP=25). The latter occurs when the pathline from a

starting point within the capture zone, due to a derivative increment ΔP, does not end in "its own well screen" but in the well screen of a neighouring capture zone. The width of the no-results-band depends primarily on the density of the finite-element grid, and on the value of the derivative increment ΔP. In our case, the increment $\Delta P=10\%$ of the ^{10}log parameter values P_i was used.

Figure 3.22 depicts the resulting reliability of travel times for the 10 groundwater abstraction sites considered in this study. The picture is prepared using the LGMLUC output files with the following travel times:
- *ms*fes_ttiexp_ddg.unf (expectation value),
- *ms*fes_ttiinner_ddg.unf ("inner" travel time contour is closer to the well than the expected travel time contour);
- *ms*fes_ttiouter_ddg.unf ("outer" travel time contour is further from the well than the expected travel time contour).

The figure shows the probability that a water particle reaches a well within 50 years, for the probability range of 5-95%. The dark-blue bands represent the zone where the probability to reach the well within 50 years gradually decreases from 0.95 to 0.05, going from "inner" to "outer" isochrone. One may also say that this zone contains the "true" 50-year isochrone with a probability of 1.0–0.05–0.05, i.e. 0.90.
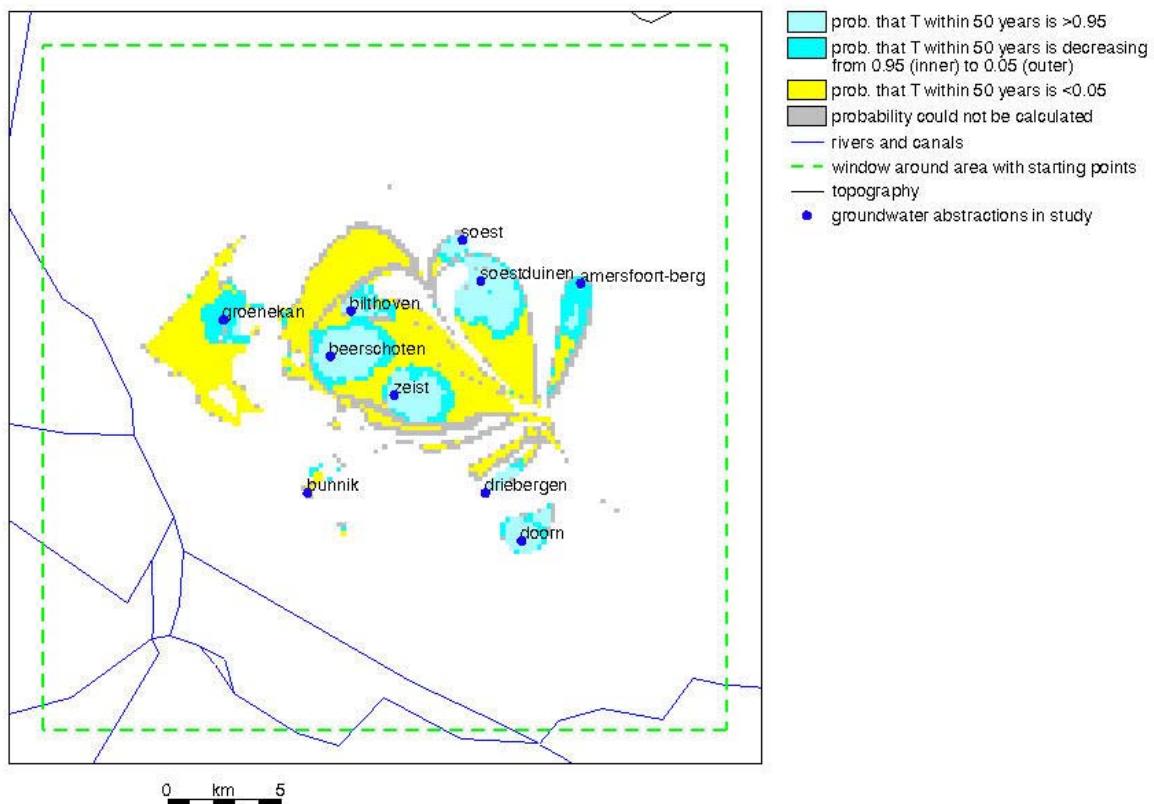
Figure 3.22 Probability that a water particle will reach the well within 50 years. Probability range of 5-95%, for 10 groundwater abstractions considered in this study. The inner or outer boundary of a probability band is an isochrone located either closer to (inner) or further from the well (outer).

In Figure 3.22 note too the grey-shaded spots, a narrow-band along the inner side of the capture zone boundaries. Here the probability could not be calculated. These grey-shaded plots are identical to the hashed yellow-coloured zone in Figure 3.21.

The travel-time probability can also be presented in one figure for a single groundwater abstraction site applying to a series of travel times. An example of such way of presenting the results is depicted in Figure 3.23.
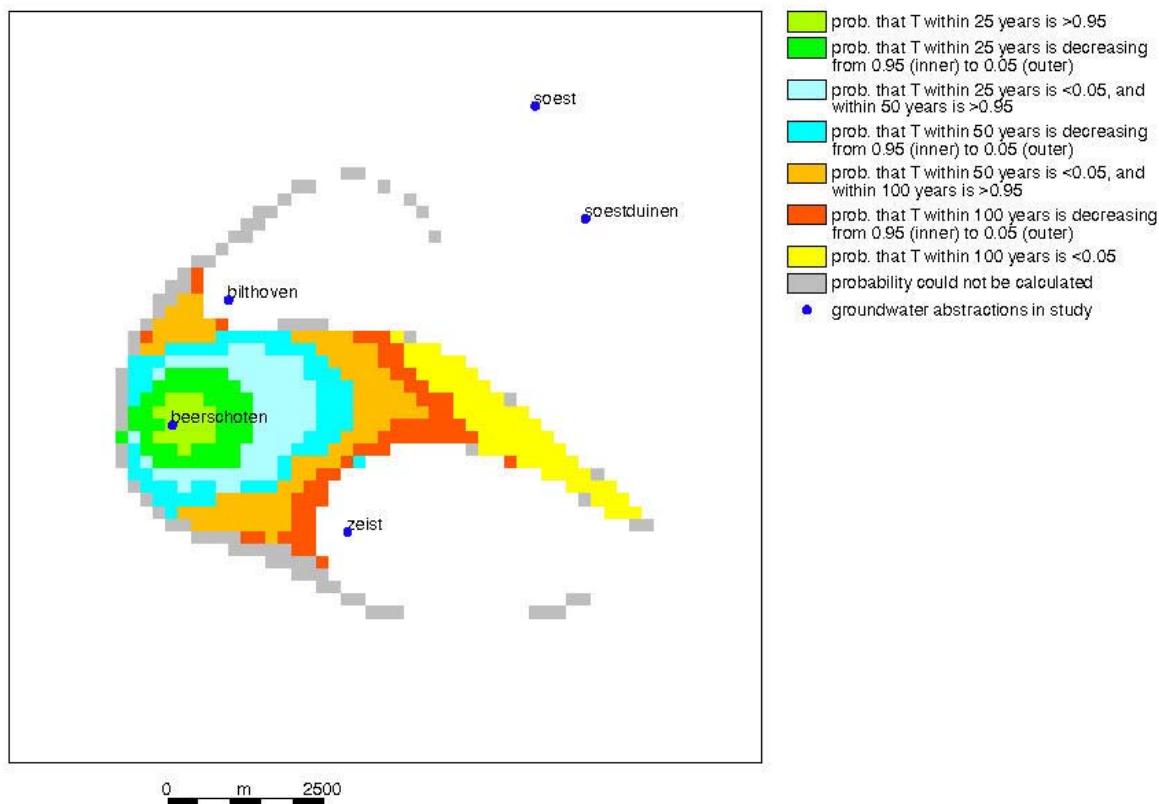
Figure 3.23 Probability that a water particle will reach the well within 25, 50 and 100 years. Probability range is 5-95% for groundwater abstraction Beerschoten The inner or outer boundary of a probability band is an isochrone located either closer to (inner) or further from the well (outer).

Finally, a note about the limitation of this method. One of the additional output items is the "expected" capture-zone boundary. This zone is derived by means of a single deterministic run for the "expected" values of model input parameters. However, since the model-input parameters are associated with uncertainty, the location of the capture-zone boundary is also uncertain and varies with the different combinations of parameters. It should be noted that although the method presented here is suitable for determining travel-time reliability, it cannot be used to determine the uncertainty in the location of the capture-zone boundary. The uncertainty of the capture-zone boundary can be assessed by means of a Monte Carlo based method (not used in this study). The FOSM method might be used for this purpose, but with a different numerical procedure for determining the travel-time derivatives. This could be investigated in a follow-up study.

4.   Conclusions

The modelling approach developed here on the basis of the First-Order Second-Moment (FOSM) method for uncertainty propagation analysis, incorporated in the program LGMLUC, makes it possible to determine the reliability of travel times of water particles to well screens abstracting groundwater. The reliability is presented as a band (zone) around the "expected" travel time isochrone (Figures 3.22 and 3.23). The band is located between the so-called "inner" and "outer" isochrones. The width of this probability occurrence zone, e.g. between 97.5 and 2.5 percentile values, increases with increasing uncertainty in the model-input parameters, in other words, with variance in these parameters.

Since the procedure is computationally efficient, it can be used for problems at various spatial resolutions of the finite-element grid. These range from local problems with a high grid density to problems on a regional scale similar to the model area in this study. The accuracy of the FOSM method has been checked by means of comparison with Monte Carlo analyses. One comparison has been made for an LGM model and another, independent, comparison for a TRIWACO model.

Because of the numerical procedure adopted for the implementation of the FOSM method in LGMLUC, LGMLUC is not suitable for calculating reliability of travel times in a narrow-band area along the inner side of the "expected" capture-zone boundary (see discussion of Figure 3.21 in Chapter 3). The reason for not being able to calculate the reliability in that area is that the travel-time sensitivity derivatives cannot be calculated for either of the involved uncertain model input parameters.

One of the output items of the LGMLUC program is the "expected" capture zone boundary. This zone is derived by means of a single deterministic run for the "expected" values of model input parameters. However, since the model-input parameters are associated with uncertainty, the location of the capture-zone boundary is also uncertain and varies with different combinations of parameters. It should be noted that although the method presented here is suitable for determining travel-time reliability, it cannot be used to determine the uncertainty in the location of the capture-zone boundary. The uncertainty of the capture-zone boundary can be assessed by means of a Monte Carlo based method (not used in this study). Possibly, also the FOSM method could be used for this purpose, but with a different numerical procedure for determining the numerical derivatives (see also previous section).

The method incorporated in the program, LGMLUC, can be used for simulation of travel-time reliability on the regional scale in the Netherlands. A prerequisite for its application is that information is available about the uncertainty of those model-input parameters that influence the reliability of travel times. The relevant parameters are: (a) aquifer transmissivity, (b) hydraulic resistance of aquitards, (c) drainage and infiltration resistance defining the head-flux relation between the top aquifer and surface waters, (d) effective porosity of aquifers, (e) effective porosity of aquitards and (f) groundwater recharge rate. The uncertainty in the model input parameters (a)

through (c) can be derived using inverse-method calibration, for example, as applied in the LGMCAL module (developed in a previous study). However, the uncertainty in the model input parameters (d) through (f) will have to be assessed separately by expert judgement, similarly to the approach followed in this study (section 2.5.2).

The information on the reliability of travel time – generated in the methodology presented here– provides policy-makers and administrators with additional insight into the location of vulnerable areas located both inside and outside the existing groundwater protection zones. A vulnerable area could be one that contains a potentially polluting object or activity located close to the fringe of the protection zone.

The information on the reliability of travel time can also be used in decision-making with regard to space allocation where claims are made on the same space by various users (target sectors). Groundwater-based drinking-water production is one of the users claiming space for groundwater protection zones. Here, the reliability of travel times can serve as input for developing policy measures for long-term sustainable protection of groundwater. Specifically, reliability information can be applied in two ways:

− The outer boundary of the 97.5-2.5 percentile reliability band for the 25-year travel time – the 2.5 percentile travel-time contour – can be interpreted as delineation of the groundwater protection zone by means of which the maximum protection can be achieved. This zone, larger than the one based on contour for the expected travel time of 25 years, would ensure that the groundwater polluted is minimalized .

− The area within a reliability band, for example, between the 97.5 and 2.5 percentiles, can be used for risk-based decision-making. An example of this application would be the decision about re-allocating a potentially polluting activity, taking into account the occurrence probability of pollution and, also probabilistically, the level of threat expressed in terms of travel time to the abstraction well.

References

Aitchison, J., Brown, J.A.C. (1957) The Log-Normal Distribution. Cambridge University Press, New York.

Banton, O., Porel, G., Delay, F. (1997) Coupling of the time domain random walk method with the finite fragment method to simulate flow and transport in 1-D heterogeneous media. J. Hydrol. 201, 49–61.

Dettinger, M.D., Wilson, J.L. (1981) First order analysis of uncertainty in numerical models of groundwater flow. Part. 1. Mathematical development. Wat. Resour. Res. 17(1), 149–161.

Elfeki, A.M.M. (1996) Stochastic Characterization of Geological Heterogeneity and its Impact on Groundwater Contaminant Transport. Ph. D. thesis. Delft University of Technology, the Netherlands.

Franzetti, S., Guadagnini, A. (1996) Probabilistic estimation of well catchments in heterogeneous aquifers. J. Hydrol. 174, 149–171.

Guadagnini, A., Franzetti, S. (1999) Time-related capture zones for contaminants in randomly heterogeneous formations. Groundwater 37(2), 53-260.

Hendricks Franssen, H.J., Stauffer, F., Kinzelbach, W. (2004) Joint estimation of transmissivities and recharges. Application: stochastic characterization of well capture zones. J. Hydrol. 294, 87–102.

James, S.C., Chrysikopoulos, C.V. (2001) An efficient particle tracking equation with specified spatial step for the solution of the diffusion equation. Chemical Engineering Science 56(23), 6535–6543.

Kovar, K., Leijnse, A., Gan, J.B.S. (1992) Groundwater Model for Netherlands; Mathematical Model Development and User's Guide. RIVM report no. 7143050002, Bilthoven, the Netherlands.

Kovar, K., Pastoors, M.J.H., Tiktak, A., van Gaalen, F.W. (1998) Application of the Netherlands Groundwater Model, LGM, for calculating concentration of nitrate and pesticides at abstraction wells in sandy soil areas of the Netherlands. RIVM report no. 703717002, Bilthoven, the Netherlands.

Kovar, K., Pastoors, M.J.H., Mülschlegel, J.H.C., van Ooijen, S.P.J. (2000) Toepassing van het Landelijk Grondwater Model, LGM, voor berekening van ruwwaterconcentraties op een aantal locaties van oevergrondwaterwinningen in Nederland (Application of the Netherlands Groundwater Model, LGM, for calculating solute concentration at a number of groundwater abstraction sites in vicinity of large rivers in the Netherlands). RIVM report no. 703717008, Bilthoven, the Netherlands.

Kovar, K., Uffink, G.J.M., Pastoors, M.J.H. (1996) Evaluation of the LGM Groundwater Model for the Netherlands for the calculation of pathlines, travel times and concentrations at abstraction wells. RIVM report no. 703717001, Bilthoven, the Netherlands.

Kreyzig, E. (1970) Introductory Mathematical Statistics. Principle and Methods. Wiley and Sons, New York.

LaVenue, M., Andrews, R.W., RamaoRao, B.S. (1989) Groundwater travel time uncertainty analysis using sensitivity derivatives. Wat. Resour. Res. 25(7), 1551–1566.

Leijnse, A., Pastoors, M.J.H., Roelofsen, F.J. (2002a) The effect of weighting the observations on parameter estimates in groundwater modelling. In: Calibration and Reliability in Groundwater Modelling (ed. by K. Kovar & Z. Hrkal), Proc. Int. Conf. ModelCARE'2002, June 2002, Prague, Czech Republic, IAHS Publ. No. 277, 204–211.

Leijnse, A., Pastoors, M.J.H., Tiktak, A. (2002b) Linear first order estimate of the spatially distributed prediction accuracy of groundwater models: A practical example. In: Accuracy'2002 (ed. by Gary Hunter & Kim Lowel), Proc. Accuracy'2002, 5th Int. Symp. on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, July 2002, Melbourne, 75–82.

Neupauer, R.M., Wilson, J.L. (2003) Backward location and travel time probabilities for a decaying contaminant in an aquifer. J. Contam. Hydrol. 66, 39–58.

Pastoors, M.J.H. (1992) Landelijk Grondwater Model; conceptuele modelbeschrijving (National Groundwater Model; description of model concept). RIVM report no. 7143050004, Bilthoven, the Netherlands.

Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T. (1988) Numerical Recipes in C, The Art of Scientific Computing. Cambridge University Press.

Royal Haskoning (2002) TRIWACO a simulation package for groundwater, version 3.0. Haskoning Nederland B.V., Rotterdam, the Netherlands.

Stoppelenburg, F.J., Kovar, K., Pastoors, M.J.H., Tiktak, A., Leijnse, A. (2002) Two-way coupling of 1-D unsaturated-saturated flow model SWAP with 3-D saturated regional groundwater flow model LGM: Time-average coupling approach. Acta Universitatis Carolinae, Geologica, 46(3-4), 575–580.

Sun, N.Z. (1994) Inverse Problem Theory in Groundwater Modelling. Kluwer Academic Publishers, the Netherlands.

Uffink, G.J.M. (1989) Application of Kolmogorov's backward equation in random walk simulations of groundwater contaminant transport. In: Contaminant Transport in Groundwater (ed. by H.E. Kobus & W. Kinzelbach), Proc. Int. Symp., April 1989, Stuttgart, Germany. Balkema, Rotterdam, 283–289.

Uffink, G.J.M., van der Linden, A.M.A. (1998) Dilution of pesticides in groundwater during advective dispersive transport. RIVM report no. 716601002, Bilthoven, the Netherlands.

Uffink, G.J.M., Römkens, P.F.A.M. (2001) Nitrate transport modeling in deep aquifers. Comparison between model results and data from the groundwater monitoring network. RIVM report no. 711401010, Bilthoven, the Netherlands.

Van Leeuwen, M. (1997) Stochastic Determination of Capture Zones of Groundwater Wells. Ph. D. thesis. Delft University of Technology, the Netherlands.

Varljen, M.D., Shafer, J.M. (1991) Assessment of uncertainty in time-related capture zones using conditional simulation of hydraulic conductivity. Groundwater 29, 737–748.

Zaadnoordijk, W.J. (2001) Analytical relations between model input statistics and output reliability for verification of a numerical groundwater model. In: Impact of Human Activity on Groundwater Dynamics (ed. by Hans Gehrels, Norman E. Peters, Eduard Hoehn, Karsten Jensen, Chris Leibundgut, Jasper Griffioen, Bruce Webb & Willem Jan Zaadnoordijk), Proc. IAHS Assembly, July 2001, Maastricht, the Netherlands, IAHS Publ. No. 269, 263–269.

Zaadnoordijk, W.J. (2003) Analytic expressions for output reliability: Verifying automated calibration and confidence calculations of a groundwater flow simulator. Wat. Resour. Res. 39(12), 1351.

Appendix 1: User's guide for LGMLUC and LGMMUC

The programs LGMLUC and LGMMUC are used to determine the uncertainty in groundwater travel times due to uncertainty in the values of input parameters, such as transmissivities, hydraulic resistances, effective porosities and groundwater recharge rate:
- LGMLUC uses a linear approach and assumes the travel time to be either normally or log-normally distributed. The output are travel times for a lower percentile (e.g. 2.5%) and travel times for an upper percentile (e.g. 97.5%). Both travel-time output sets are values at all nodes of a finite element grid and, thus, can be used for contouring of isochrones;
- LGMMUC uses Monte Carlo simulations to determine the distribution of the travel time in a number of user-selected locations. The output is a table of travel times, for each selected location, one travel time for each Monte Carlo realization. Subsequently, the tables can be used to prepare a histogram and, based on the histogram, the travel times for any percentile. For LGMMUC, due to its nature, one does not have to make any assumption about the probability distribution of travel times.

For both LGMLUC and LGMMUC, the uncertain input parameters are assumed to be log-normally distributed. Those input parameters consist of the output from the calibration module LGMCAL and the user-added data for (a) effective porosities of aquifers and aqutards and (b) the groundwater recharge rate.

LGMLUC has the user-input option to use either a normally distributed or a log-normally distributed travel time.

The programs LGMLUC and LGMMUC require the presence of control input files for LGM-programmes LGMSAT (also named EPGO) and LGMFLOW (also named EFGO), and all files referred to in these control input files. Those control input files are:
- file **msinpepg.asc**, for LGMSAT (EPGO);
- file **msinpefg.asc**, for LGMFLOW (EFGO).

In the following, only the terms "EPGO" and "EFGO" will be used instead of LGMSAT and LGMFLOW, respectively.

The two-character text string "*ms*" at the beginning of file names represents the so-called scenario identification text.

Uncertain parameters affecting reliability of travel times
The following input (base) files, representing uncertain hydraulic parameters, are used to simulate the uncertainty in travel times. Two groups of parameters are distinguished, namely (a) parameters calibrated (optimized) by LGMCAL, and (b) parameters added later, with an assumed variance:
(a) Parameters calibrated by LGMCAL:
- **msc0_tab.asc**, for hydraulic resistance of aquitard 0 (above aquifer 1) (calibrated for the factor in the formula to calculate RPL)
- **mscc_tab.asc**, for hydraulic resistance of aquitards j = 1 to NAQF-1 (calibrated for the factor in the formula to calculate RAj)
- **mstr_tab.asc**, for transmissivity TRi of aquifers i = 1 to NAQF (calibrated for the multiplication factor, in 1 or 5 pilot points)
- **msts_tab.asc**, for shape of the topsystem relation (calibrated for the drainage/infiltration resistance)
(b) Parameters added later, with assumed variance:
- **mspajddg.unf**, this is *ddgop.unf* file for porosity of aquitards (e.g. pa1, for aquitard 1, spatially variable) (only one uncertainty for an entire aquitard can be taken into account)
- **mspfiddg.unf**, this is *ddgop.unf* file for porosity of aquifers (e.g. pf1, for aquifer 1, spatially variable)

(only one uncertainty for an entire aquifer can be taken into account)
-       **msqreddg.unf,** this is *ddgop.unf* file for groundwater recharge rate
        (spatially variable)
        (only one uncertainty for the entire model can be taken into account).
The latter three parameters (pa*j*, pf*i*, qre) were not part of the calibration
process, but are added later to the LGMCAL-produced optimization results. Their
variance has been assumed. When X denotes any of these three parameters and X_0
is the average value (geometric mean), then the variance is chosen such that 95%
(between 2.5 and 97.5 percentiles) of the distribution is in the interval:

$$\frac{X_0}{1.1} < X < 1.1\, X_0 \qquad\qquad (1)$$

In a 10-log normal distribution, **which is the case for pa*j*, pf*i* and qre**, these
limits are

$$\frac{X_0}{10^{2s}} < X < X_0\, 10^{2s} \qquad\qquad (2)$$

leading to a variance of 10-log-assumed pa*j*, pf*i* and qre: $s^2 = 4.3 \times 10^{-4}$. In the
log-normal case, X_0 represents the median (50 percentile).

Files required for running LGMLUC and LGMMUC
The files applicable for the uncertainty assessment are to be selected below.
The total number of parameter items associated with uncertainty is denoted by
NP, *kk* = 1 through NP.

The above mentioned files will be adapted by both LGMLUC and LGMMUC. LGMLUC
adapts the files in order to calculate the derivative of travel time with
respect to a parameter item *kk*, while LGMMUC adapts the files in order to
generate values for the hydraulic parameters for the Monte Carlo simulations.
Because of these adaptions, it is strongly advised to store the input (base)
files in backup files before starting the calculation with either LGMLUC or
LGMMUC. It is suggested to use the following notation:
-       store the original *msc*0_tab.asc as **msc0_tab.asc_opt**
-       store the original *ms*cc_tab.asc as **mscc_tab.asc_opt**
-       store the original *ms*tr_tab.asc as **mstr_tab.asc_opt**
-       store the original *ms*ts_tab.asc as **msts_tab.asc_opt**
-       store the original *ms*pa*j*ddg.unf as **mspajddg.unf_opt**
-       store the original *ms*pf*i*ddg.unf as **mspfiddg.unf_opt**
-       store the original *ms*qreddg.unf as **msqreddg.unf_opt**
The input base files will be restored if LGMLUC or LGMMUC finishes properly.
However, if this is not the case, you will need these *_opt files to restore the
input (base) files. Restore by:
-       copy **msc0_tab.asc_opt** *msc*0_tab.asc
-       copy **mscc_tab.asc_opt** *ms*cc_tab.asc
-       copy **mstr_tab.asc_opt** *ms*tr_tab.asc
-       copy **msts_tab.asc_opt** *ms*ts_tab.asc
-       copy **mspajddg.unf_opt** *ms*pa*j*ddg.unf
-       copy **mspfiddg.unf_opt** *ms*pf*i*ddg.unf
-       copy **msqreddg.unf_opt** *ms*qreddg.unf
For easiness we recommend to include the above restore-copy commands in a
uniquely named "run file" (referred to as script file in UNIX environment), such
as *runcopy_opt_files*.

How to run LGMLUC and LGMMUC
Once the control input file *ms*inpluc.asc (for LGMLUC) or *ms*inpmuc.asc (for
LGMMUC) is prepared, start the execution by specifying:

        lgmluc *ms*inpluc.asc {enter}
        or
        lgmluc *ms*inpmuc.asc {enter}

Calculation process in LGMLUC
The algorithm used in LGMLUC is based on a linear approach. If we indicate the ^{10}log parameter values as P_i, $i=1,NP$, the variance of the travel time is given by:

$$\sigma_T^2 = \sum_{i=1}^{NP} \left(\sum_{j=1}^{NP} \frac{\partial T}{\partial P_i} \frac{\partial T}{\partial P_j} Cov(P_i, P_j) \right) \qquad (3)$$

where σ is the standard deviation of the travel time T. A similar expression is used if the ^{10}log value of the travel time is considered, replacing T by ^{10}log(T).

The contribution of the different parameters to the variance of the travel time is estimated as:

$$\sigma_{T|P_i}^2 = \left(\frac{\partial T}{\partial P_i} \right)^2 \sigma_{P_i}^2 \qquad (4)$$

where σ_P is the standard deviation of the parameter P. This estimate (note that covariances are not taken into account) is based on the assumption that the correlations between the different parameters are small.

The calculation process in LGMLUC now consists of five major steps:
a)   calculate the base situation, with the optimized (calibrated) values of the parameters; this involves running EPGO and EFGO in sequence. The following files are generated:
-       **msfes_ict_ddg.unf.00** containing the exit codes for the starting particles. The following codes ICT are distinguished:
-          ICT=0: particle tracking stopped because maximum tracking time was exceeded;
-          ICT=1: pathline ended in a well (well IWL);
-          ICT=2: pathline ended in a river;
-          ICT=3: pathline ended on model periphery (outflow from an aquifer);
-          ICT=4: pathline ended on top of system: (a) if the starting point is in an infiltration area then pathline ends in a draining area or (b) if the starting point is in a drainage area then particle tracking cannot occur;
-       **msfes_iwl_ddg.unf.00** containing the wells by which the particles are captured. The IWL numbers are internal LGM sequence numbers, ranging from 1 to *NWL*, where *NWL* is the total number of model-input well screens in the model. If a well intersects various aquifers, this will result in two or more IWL wells. NWL is generated by the grid generator EGGO, as part of the data discretisation procedure;
-       **msfes_tti_ddg.unf.00** (mean, deterministic, travel times, **in seconds**) This file contains the expected values of the travel time, later stored in the output file *ms*fes_ttiexp_ddg.unf. However, in contrast to *ms*fes_tti_ddg.unf.00, the file *ms*fes_ttiexp_ddg.unf will contain the value −9999 in those starting points where the variance (see equ. 3) could not be calculated;
b)   store the initial (base) files in temporary files:
-          store the original *ms*c0_tab.asc as **msc0_tab.asc.bas**
-          store the original *ms*cc_tab.asc as **mscc_tab.asc.bas**
-          store the original *ms*tr_tab.asc as **mstr_tab.asc.bas**
-          store the original *ms*ts_tab.asc as **msts_tab.asc.bas**
-          store the original *ms*pajddg.unf as **mspajddg.unf.bas**
-          store the original *ms*pfiddg.unf as **mspfiddg.unf.bas**
-          store the original *ms*qreddg.unf as **msqreddg.unf.bas**
      The *.bas files are identical to the *_opt files mentioned above.

c)   calculate the derivatives of travel times with respect to the ^{10}log values of the NP parameter items, for each derivative re-running EPGO and EFGO. The derivatives are calculated numerically by a finite difference approach. One parameter value is generated that deviates a certain percentage (cf records 19-20 of the input file *ms*inpluc.asc/*ms*inpmuc.asc) from the original (optimized) value while keeping all other parameter values the same as for the base run; the travel times for this parameter combination is calculated, and the derivatives are calculated as the difference in travel times in this run and in the base run, divided by the change in ^{10}log(parameter value).
For the parameters *kk* = 1 through NP, the following files are generated:
-    **msfes_ict_ddg.unf.*kk***, containing the exit codes for the particles;
-    **msfes_iwl_ddg.unf.*kk***, containing the wells by which the particles are caught;
-    **msdtdp_tgp.*kk***, containing the derivative of the (^{10}log) travel time with respect to the ^{10}log of the parameter values.
d)   assemble the NP derivatives and the parameter covariance matrix to the variance of the (^{10}log) travel time (cf equ. 3). The variance of the travel time cannot be calculated if either of the NP+1 pathlines have not ended in the same well. This will be indicated by the value –9999 in the LGMLUC output files (*ddgop.unf*) listed in the next section.
e)   restore the original base input files and delete the temporary files *.bas

Description of output files from LGMLUC
The following output files are generated by LGMLUC:
-    **mssigmaddg.unf.*kk*** , *kk* = 1 through NP
     These files contain an estimate of the contribution of the parameters *kk* to the variance of the (^{10}log) travel time (equ. 4). This is a *ddgop.unf* file with a parameter SIGMA.
-    **msfes_ttiexp_ddg.unf**
     This file contains the expected (deterministic) value of the travel times (**in days!**), i.e. the travel times in the base run. This is a *ddgop.unf* file with a parameter TTIEXP. If either of the NP+1 pathlines have not ended in the same well, then TTIEXP value is set –9999. The travel times in this file are identical to those in **msfes_tti_ddg.unf.00**, with the exception of starting where the variance of the travel time could not be calculated. In the latter nodes **msfes_ttiexp_ddg.unf** contains –9999, while *ms*fes_tti_ddg.unf.00 contains the base travel-time value
-    **msfes_ttiinner_ddg.unf**
     This file contains he travel times (**in days!**) for the upper value of the confidence limits (records 36-37 of the input file *ms*inpluc.asc/ *ms*inpmuc.asc)(e.g. expectation+2sigma, for the 97.5 percentile). This is a *ddgop.unf* file with a parameter TTIINNER. If either of the NP+1 pathlines have not ended in the same well, then TTIINNER value is set –9999.
-    **msfes_ttiouter_ddg.unf**
     This file contains the travel times (**in days!**) for the lower value of the confidence limits (records 36-37 of the input file *ms*inpluc.asc/ *ms*inpmuc.asc) (e.g. expectation-2sigma, for the 2.5 percentile). This is a *ddgop.unf* file with a parameter TTIOUTER. If either of the NP+1 pathlines have not ended in the same well, then TTIOUTER value is set –9999.
-    **msfes_ttiiwlall_ddg.unf**
     This file contains the internal LGM sequence number of the destination well, reached for all NP+1 simulations. This is a *ddgop.unf* file with a parameter TTIIWLALL. If either of the NP+1 pathlines have not ended in the same well, then TTIIWLALL value is set –9999.
-    **msoutluc.frm**
     This formatted output file contains information about the parameter values and gives a table of the expected travel times, the variances in the (^{10}log) travel times and an estimate of the contributions of the different parameters to the variance in the travel times. This tabel contains only

those starting points for which the particle has ended in the same well for all NP+1 simulations.

Improper ending of LGMLUC

For some reason, LGMLUC might end improper, e.g. because of a system crash, network problems, disk full, etc. In that case, it is possible to restart LGMLUC by defining a starting parameter (record 1 of the input file *ms*inpluc.asc/ *ms*inpmuc.asc). The parameter to start with can be determined on the basis of the presence of the intermediate files *ms*fes_ict_ddg.unf.*jj*, *ms*fes_iwl_ddg.unf.*jj* and *ms*dtdptgp.*jj*. All parameters *kk* =< *jj* have already been treated, and, hence, the starting parameter in the follow-up run should be *jj*+1. However, make sure that all intermediate files with extension *kk* =< *jj* are present, since these are required for the calculation of the variance in the travel time (step d).

If the LGMLUC run starts with a parameter item *kk*>1, then the input (base) files have to be restored before starting LGMLUC. Though both the *_opt files and the *.bas files can be used, we suggest to use the *_opt files:
- copy *ms*c0_tab.asc_opt to ***msc0_tab.asc***
- copy *ms*cc_tab.asc_opt to ***mscc_tab.asc***
- copy *ms*tr_tab.asc_opt to ***mstr_tab.asc***
- copy *ms*ts_tab.asc_opt to ***msts_tab.asc***
- copy *ms*pajddg.unf_opt to ***mspajddg.unf***
- copy *ms*pfiddg.unf_opt to ***mspfiddg.unf***
- copy *ms*qreddg.unf_opt to ***msqreddg.unf***

In order to avoid typos, it is advisable to store the above set of copy instructions in a unix script file, e.g. named *runcopy_opt_files*

Calculation process in LGMMUC

Like in LGMLUC, the algorithm in LGMMUC is based on the assumption that the parameters that contribute to the uncertainty in the travel times are log-normally distributed. A large number of parameter values will be generated according to their optimal (calibrated) values and the known covariance matrix. For each of the parameter combinations EPGO and EFGO are run to determine the traveltimes from a limited number of starting points towards a predetermined well. If the particle is caught by that particular well, the calculated travel time is added to the list of travel times. For a large enough number of simulations, a histogram of the travel times gives a good indication of the probability distribution of the calculated travel times. Contrary to simulations with LGMLUC, no prior assumptions regarding the form of this distribution has to be made.

During the Monte Carlo simulations, after each simulation a check is performed on the value of the mean and the standard deviation of the travel times. If these do not change by more than 10^{-5} times the present value for all starting points, convergence is assumed, and the simulations are stopped. Otherwise, simulations are carried out until the maximum number (see record 39 of the *ms*inpluc.asc/*ms*inpmuc.asc input file) is reached.

Description of output files from LGMMUC

LGMMUC generates the following output files:
- NPMC table output files *ms*tti_muc_iiiiii.tab.
  NPMC is the number of starting points (from record 40).
  The text "iiiiii" is the starting point number (from record 41),
  e.g. l0tti_muc_000655.tab.
  The contents of a *ms*tti_muc_iiiiii.tab contains:
  - information about the starting node number, the well that has to capture the particle and the accepted number of Monte Carlo simulations.
  - a table of the calculated travel times in increasing order. This table allows for an easy construction of the histogram, and an easy

calculation of the percentiles in the distribution of the travel times. Basically, the percentiles follow from counting the number of hits. If e.g the number of accepted simulations is 900, the 10 percentile of the travel time distribution is given by the 90th number in the tabel, the 20 percentile by the 180th number, etc. In order to construct a histogram, classes of the travel time have to be defined either on a linear scale or on a log scale, and the number of calculated travel times in each of these classes can be plotted.

- An ASCII output file called *ms*outmuc.frm that shows the progress of the Monte Carlo simulations. For each simulation, the present value of the mean and standard deviation of the travel time for each starting point is given.

Improper ending of LGMMUC:

For some reason, LGMMUC might end improper, e.g. because of a system crash, network problems, disk full, etc. There is no provision for a restart of LGMLUC, and the complete calculation has to be repeated.

Before a start of LGMMUC after an improper ending, the input (base) files have to be restored before starting LGMMUC. See the section on improper ending of LGMLUC for further information.

Contents of the control input file *ms*inpluc.asc / *ms*inpmuc.asc

Record 1   *START_WITH_PARAMETER_10*
                  START_WITH_PARAMETER_ is literal text.
                  Follows the starting sequence number *kk* of the parameter item
                  to calculate the derivative files *ms*dtdp.tgp.*kk*

                  Allowed to be 1=<*kk*.
                  kk = 1 indicates an initial run, where the base case is also
                  simulated.
                  For *kk*>1, the base case will not be carried out. LGMLUC
                  starts with the calculation of the derivative for the parameter
                  item *kk* given in record 1. This option, *kk*>1 is of practical
                  value when continuing the calculation process after abnormal
                  termination of the run. Note that in that case the base input
                  files need to be restored.

                  If *kk* is given a value > the number of parameters NP, none of the
                  derivatives is calculated, but only the assembly to determine the
                  variances and the percentiles is repeated. This will allow for
                  re-specification of the lower/upper confidence percentiles (see
                  records 36 and 37). The result is the new set of files
                  msfes_ttiexp_ddg.unf (unchanged), **msfes_ttiiiner_ddg.unf** and
                  msfes_ttiouter_ddg.unf.

                  Record 1 needs to be given for both LGMLUC and for LGMMUC.
                  However, the value of *kk* is irreklevant for LGMMUC.

Record 2   *LOG*
                  Literal text. Should be LOG or LINEAR,
                  for LOG(Travel time) or Travel time respectively
                  Record 2 must be present for both LGMLUC and LGMMUC. However, for
                  LGMMUC this record does not play a role.

Record 3   *10*
                  Scenario identification text *ms*. Two characters long.

                  The scenario identifier *ms* is, among others, used for naming
                  the LGMMUC output files, namely *ms*tti_muc_iiiiii.tab.

Record 4   ***10inpluc_covp.dat***
              File name for (ASCII) parameter covariance matrix.
              Suggested file name is *ms*inpluc_covp.dat, or *ms*inpmuc_covp.dat.
              An example of this file is given at the bottom of this document.

Records 5 through 8 have to be specified for both LGMLUC and LGMMUC. However,
LGMMUC does not use these file names.

Record 5   ***10fes_ttiexp_ddg.unf***
              Output file (*ddgop.unf*) name for expected travel times (days).
              Value −9999 is used if either of NP+1 pathlines has not ended
              in the same well.

Record 6   ***10fes_ttiouter_ddg.unf***
              Output file (*ddgop.unf*) name for lower percentile travel times
              (days), e.g. for 2.5% (specified in records 36 and 37)
              Value −9999 is used if either of NP+1 pathlines has not ended
              in the same well.

Record 7   ***10fes_ttiinner_ddg.unf***
              Output file (*ddgop.unf*) name for upper percentile travel times
              (days), e.g. for 97.5% (specified in records 36 and 37)
              Value −9999 is used if either of NP+1 pathlines has not ended
              in the same well.

Record 8   ***10fes_ttiiwlall_ddg.unf***
              Output file (*ddgop.unf*) name for well idents of starting points
              Based on all NP+1 runs.

Record 9   ***10outluc.frm***
              Formatted output file name for LGMLUC (*ms*outluc.frm)
              or LGMMUC (*ms*outmuc.frm)

Record 10   ***10fes_feg_egg.unf***
              Grid input file containing possible starting points.
              Should be identical to the one defined in EFGO input file
              (in record 46 in file *ms*inpefg.asc)

Record 11   ***10fes_ist_ddg.unf***
              Input file to identify starting points in grid file
              Should be identical to the one defined in EFGO input file
              (in record 47 in file *ms*inpefg.asc)

Record 12   ***10inpepg.asc***
              ASCII input file for EPGO (*ms*inpepg.asc)

Record 13   ***10outepg***
              Base of EPGO output files (text string *ms*outepg)

Record 14   ***10inpefg.asc***
              ASCII input file for EFGO (*ms*inpefg.asc)

Record 15   ***10outefg***
              Base text for EFGO output files (text string *ms*outefg)

Record 16   ***10fes_tti_ddg.unf***
              Output file (*ddgop.unf*) name for travel times per EFGO-run
              Should be identical to the one defined in EFGO input file
              (in record 48 in file *ms*inpefg.asc)

Record 17   **10fes_ict_ddg.unf**
                     Output file (*ddgop.unf*) name for exit codes per EFGO run
                     Should be identical to the one defined in EFGO input file
                     (in record 49 in file *ms*inpefg.asc)

Record 18   **10fes_iwl_ddg.unf**
                     Output file (*ddgop.unf*) name for well numbers per EFGO run
                     Should be identical to the one defined in EFGO input file
                     (in record 50 in file *ms*inpefg.asc)

Record 19   **Percentage DPAR**
                     Literal text
Record 20   **10**
                     Percentage change in parameter values to calculate derivatives.
                     DPAR applies to parameter values themselves, i.e. not log-values.
                     For example, using DPAR=5 for the parameter base value 1000
                     will result in the parameter increment of 50.
                     Default value DPAR=10

Record 21   **Gt**
                     Literal text
Record 22   **30**
                     Gt code number for which resistance was determined in calibration
                     Should be available in msTS_TAB.ASC file
                     Repeat as many times as necessary to define all Gt code numbers
                     optimized

Record 23   **C0**
                     Literal text
Record 24   **30**
                     C0 code number for which resistance was determined in calibration
                     Should be available in msC0_TAB.ASC file
                     Repeat as many times as necessary to define all C0 code numbers
                     optimized

Record 25   **C1**
                     Literal text
Record 26   **30**
                     C1 code number for which resistance was determined in calibration
                     Should be available in msCC_TAB.ASC file
                     Repeat as many times as necessary to define all CC code numbers
                     optimized

Record 27   **kD**
                     Literal text
Record 28   **1**
                     Number of values per aquifer for multiplication factor (1 or 5)
Record 29   **1**
                     Aquifer number for which multiplication factor(s) was(were)
                     determined in calibration
                     Repeat as many times as necessary to define all aquifers for
                     which multiplication factor(s) was(were) optimized

Record 30   **Aquifer Porosity**
                     Literal text
Record 31   **1 '10pf1ddg.unf' 'pf1'**
                     Aquifer number (i>=1), input file (*ddgop.unf*), and parameter name
                     for aquifer porosity.
                     Should be the same as defined in EFGO input file *ms*inpefg.asc.
                     Note: file name and parameter name in quotes
                     Repeat as many times as necessary to define all aquifers for

which variation in porosity should be taken into account
All numbers outside the range 1 through the number of aquifers in
the system will be neglected.

Record 32   ***Aquitard Porosity***
                 Literal text
Record 33   ***1 '10pa0ddg.unf' 'pa0'***
                 Aquitard number (j>=0), input file (*ddgop.unf*), and parameter
                 name for aquitard porosity.
                 Should be the same as defined in EFGO input file *ms*inpefg.asc.
                 Note: file name and parameter name in quotes
                 Repeat as many times as necessary to define all aquitards for
                 which variation in porosity should be taken into account
                 All numbers outside the range 1 through the number of aquitards
                 in the system will be neglected.

Record 34   ***Recharge***
                 Literal text
Record 35   ***'10qreddg.unf' 'qre'***
                 Input file (*ddgop.unf*) and parameter name for recharge.
                 Should be the same as defined in EPGO input file *ms*inpefg.asc.
                 Note: file name and parameter name in quotes

Record 36   ***Confidence Percentiles***
                 Literal text
Record 37   ***20   80***
                 Minimum and maximum percentile of travel time distribution to be
                 stored in *ddgop.unf* output files defined in records 6 and 7.
                 Default values 2.5 and 97.5
                 The minimum and maximum percentile are used to generate the files
                 *ms*fes_ttiouter_ddg.unf and *ms*ttiinner_ddg.unf, respectively.
                 The **minimum** confidence percentile is / implies…
                 The **maximum** confidence percentile is / implies…

                 The confidence percentiles given through records 36 and 37 are
                 relevant for program LGMLUC only.
                 Once LGMLUC has completed the calculation for a given set of
                 minimum/maximum percentiles, the results are stored in output
                 files *ms*fes_ttiouter_ddg.unf (see record 6, lower percentile)
                 and *ms*fes_ttiiiner_ddg.unf (see record 7, upper percentile).
                 If you wish to calculate a different set of lower/upper
                 travel-time uncertainties, then:
                 - specify "***START_WITH_PARAMETER_XX***" at record 1, where
                   XX is a value > #parameter items NP;
                 - specify the file name *ms*fes_ttiouter_ddg.unf (see record 6,
                   for the lower percentile);
                 - specify the file name *ms*fes_ttiinner_ddg.unf (see record 7,
                   for the upper percentile);
                 - rerun LGMLUC.

Record 38   ***Monte Carlo simulation***
                 Literal text
Record 39   ***1000***
                 Maximum number of Monte Carlo simulations
Record 40   ***4***
                 Number of starting points for Monte Carlo simulations
Record 41   ***655 849 851 1153***
                 Starting points numbers selected from those defined in
                 the file *ms*fes_ist_ddg.unf given in record 11

Record 42   *4*   *4*   *4*   *4*
                     LGM-internal sequence number IWL of the well expected to catch
                     the pathlines (out of the NMC realizations) from the given
                     starting point. Note that an IWL value refers to a well-screen
                     in one aquifer. Hence, if well abstraction at one x/y location
                     takes place from two or more aquifers, then two or more IWL
                     values identify that abstraction location. Currently, the IWL
                     values can be visually defined via Arc/Info. The IWL values can
                     also be taken from the formatted output file (FRM) of the grid
                     generator EGGO, if running the well discretization option.

Records *1* through *18* should be given in this predetermined order. They will be
                     read by both LGMLUC and LGMMUC, even though some of these records
                     are immaterial for LGMMUC
The groups: *19-20*, *21-22*, *23-24*, *25-26*, *27-29*, *30-31*, *32-33*, *34-35*, *36-37* and
                     38-42 can, following record *18* be given in any order. Each of
                     these groups can be omitted, with the exception of *38-41* in case
                     LGMMUC is run.
Groups *19-20* and *36-37* are immaterial for runs with LGMMUC. They may however
                     be present.
Group *38-42* is immaterial for runs with LGMLUC. It may however be present.

===

Structure of the parameter covariance matrix file
***ms*inpluc_covp.dat or *ms*inpmuc_covp.dat**
(see record 4 in control input file *ms*inpluc.asc / *ms*inpmuc.asc):

Record I   **NP**
                     Number of active parameters that determine the uncertainty in the
                     travel times. This number defines the number of entries in the
                     covariance matrix (NP^2), and should be the same as the number of
                     active parameters that follow from the definition in the input
                     file *ms*inpluc.asc / *ms*inpmuc.asc, records 21 through 35.

Record II   **COVARIANCE_MATRIX**   or
                     STDEV_AND_CORRELATION_MATRIX
                     Literal text.
                     This records indicates the way the covariance matrix will be
                     given.

Record III is given if the literal text in record II is **COVARIANCE_MATRIX**
Record III   **COV(I,J),J=1,NP**
                     COV are the elements of the full covariance matrix.
                     This record is given NP times in order to define the complete
                     covariance matrix.

Records IV-V are given if the literal text in record II is
STDEV_AND_CORRELATION_MATRIX
Record IV   **STDEV(I),I=1,NP**
                     STDEV are the standard deviations of the parameter values as
                     obtained from the calibration.
Record V   **COR(I,J),J=1,I**
                     COR are the elements of the correlation matrix as obtained from
                     the calibration. This record is given NP times in order to define
                     the complete symmetric correlation matrix. Note that only the
                     lower part of the correlation matrix needs to be given.

Appendix 2: Correlation matrix for 22 LGMCAL-optimized parameters

The correlation matrix for the 22 parameter values, optimized by module LGMCAL for the model area Utrecht:

	W1(30)	W1(40)	W1(50)	W1(70)	W1(1030)	W1(1040)	W1(1050)	W1(1070)	B0(-250)	B0(0)	B1(12)	B1(18)	ALF11	ALF12	ALF13	ALF14	ALF15	ALF21	ALF22	ALF23	ALF24	ALF25
W1(30)	1.000																					
W1(40)	-0.117	1.000																				
W1(50)	-0.377	-0.011	1.000																			
W1(70)	-0.090	0.014	-0.056	1.000																		
W1(1030)	-0.005	-0.007	0.029	0.096	1.000																	
W1(1040)	-0.013	0.001	0.011	0.006	-0.010	1.000																
W1(1050)	-0.013	0.005	-0.067	0.003	-0.015	0.003	1.000															
W1(1070)	0.019	-0.004	0.040	-0.027	-0.029	0.019	-0.059	1.000														
B0(-250)	-0.001	0.000	0.000	0.009	0.000	-0.001	0.002	0.000	1.000													
B0(0)	0.004	0.003	-0.061	-0.017	-0.076	-0.021	-0.400	-0.082	0.009	1.000												
B1(12)	-0.261	-0.051	-0.041	-0.077	0.048	-0.021	-0.004	0.110	0.000	-0.010	1.000											
B1(18)	0.017	0.001	-0.009	-0.007	0.007	0.000	-0.041	-0.022	-0.003	0.126	0.011	1.000										
ALF11	-0.003	-0.001	0.015	-0.006	-0.022	-0.019	-0.003	0.026	-0.047	-0.359	-0.006	0.044	1.000									
ALF12	0.014	-0.005	0.038	-0.001	-0.003	0.000	-0.038	-0.014	-0.001	-0.018	0.010	-0.002	-0.005	1.000								
ALF13	0.014	0.003	0.018	0.011	-0.009	-0.004	-0.003	0.011	-0.004	0.001	0.134	-0.004	-0.001	0.003	1.000							
ALF14	0.021	-0.002	0.003	0.017	-0.031	-0.007	0.007	0.005	-0.001	-0.009	0.025	0.003	-0.032	-0.007	0.006	1.000						
ALF15	0.070	0.000	0.056	0.045	-0.011	-0.015	-0.032	0.013	-0.003	-0.036	0.067	0.003	-0.055	-0.011	-0.015	-0.013	1.000					
ALF21	0.003	0.001	-0.004	0.002	-0.001	0.000	0.001	0.003	0.000	-0.004	0.005	0.036	-0.002	-0.001	-0.001	-0.001	0.004	1.000				
ALF22	-0.082	-0.024	0.103	-0.016	-0.034	-0.006	-0.025	-0.051	0.001	-0.018	-0.105	-0.023	0.012	-0.053	-0.017	0.007	0.028	-0.007	1.000			
ALF23	-0.039	-0.011	-0.006	0.017	-0.056	-0.001	0.011	0.001	-0.005	0.009	-0.111	-0.053	0.003	-0.021	-0.003	0.001	-0.006	-0.009	-0.035	1.000		
ALF24	0.033	-0.003	-0.032	0.021	-0.056	0.001	0.005	0.008	-0.002	0.005	0.025	-0.024	-0.003	0.002	0.002	-0.010	-0.014	0.023	0.038	0.004	1.000	
ALF25	0.029	0.001	0.148	0.110	-0.007	0.029	0.005	-0.122	0.006	0.017	0.010	0.221	0.015	-0.019	-0.010	-0.025	-0.234	-0.014	-0.270	-0.238	-0.167	1.000

value correlation is between 0.1 and 0.2, or between –0.1 and –0.2

value correlation is > 0.2 or < -0.2

Appendix 3: Standard deviation and correlation matrix for 25 LGMLUC-input parameters

File *u1inpluc_covp.dat*, an input file for LGMLUC / LGMMUC, specified via control input file *u1inpluc.asc*:

```
25   (number of hydrological parameters affecting reliability of travel times)
STDEV_AND_CORRELATION_MATRIX
3.4415E-01 5.3269E-01 4.3005E-01 7.0958E-01 6.1059E-01
7.1760E-01 4.0919E-01 5.8501E-01 5.1016E-01 3.8847E-01
1.9006E-01 1.7133E-01 2.9287E-01 3.0905E-01 3.1035E-01
3.1606E-01 3.0844E-01 3.0872E-01 2.6050E-01 3.2211E-01
3.1592E-01 9.6263E-02 0.020736   0.020736   0.020736

 1.000
-0.117  1.000
-0.377 -0.011  1.000
-0.090  0.014 -0.056  1.000
-0.005 -0.007  0.029  0.096  1.000
-0.013  0.001  0.011  0.006 -0.010  1.000
-0.013  0.005 -0.067  0.003 -0.015  0.003  1.000
 0.019 -0.004  0.040 -0.027 -0.029  0.019 -0.059  1.000
-0.001  0.000  0.000  0.009  0.000 -0.001  0.002  0.002  1.000
 0.004  0.003 -0.061 -0.017 -0.076 -0.021 -0.400 -0.082  0.009  1.000
-0.261 -0.051 -0.041 -0.077  0.048 -0.021 -0.004  0.110  0.000 -0.010  1.000
 0.017  0.001 -0.009 -0.007  0.007  0.000 -0.041 -0.022 -0.003  0.126  0.011  1.000
-0.003 -0.001  0.015 -0.006 -0.022 -0.019 -0.003  0.026 -0.047 -0.359 -0.006  0.044  1.000
 0.014 -0.005  0.038 -0.001 -0.003  0.000 -0.038 -0.001 -0.018 -0.006  0.010 -0.002 -0.005  1.000
 0.014  0.003  0.018  0.011 -0.009 -0.004 -0.003 -0.004  0.001  0.010 -0.004 -0.001 -0.001  0.003  1.000
 0.021 -0.002  0.003  0.017 -0.031 -0.007  0.007  0.011  0.001  0.134 -0.004  0.003 -0.001  0.006  0.006  1.000
 0.070  0.000  0.056  0.045 -0.011 -0.032 -0.015  0.005  0.013  0.025 -0.009 -0.032 -0.055 -0.011 -0.015 -0.013  1.000
 0.003  0.001 -0.004  0.002  0.000  0.001  0.000  0.013  0.003  0.067 -0.036 -0.011 -0.001 -0.001 -0.001  0.004  1.000
-0.082 -0.024 -0.103 -0.016 -0.006 -0.025 -0.034  0.001  0.000  0.005 -0.004  0.036  0.012  0.007  0.028 -0.007  1.000
-0.039 -0.011 -0.006 -0.056 -0.001  0.011 -0.034 -0.018 -0.051  0.000  0.005 -0.023 -0.003  0.001 -0.006 -0.009 -0.035  1.000
 0.033 -0.003  0.021 -0.056  0.001  0.005  0.008  0.025 -0.111 -0.001 -0.105  0.009 -0.053 -0.003  0.001 -0.009  0.023  0.038  0.004  1.000
 0.029  0.001  0.148  0.110 -0.007  0.029  0.005 -0.122  0.025 -0.002 -0.053  0.005 -0.021  0.002 -0.010 -0.014 -0.270 -0.238 -0.167  1.000
 0.001  0.017  0.010  0.010  0.006  0.221  0.015 -0.019 -0.003  0.002  0.010 -0.025 -0.234
 0.       1.
 0.       0.       1.
 0.       0.       0.       1.
```