

Stochastic determination of capture zones for the well field 'Het Rot' (Belgium)

L. Feyen¹, M. Dessalegn², F. De Smedt¹, O. Batelaan¹, S. Gebremeskel¹

1. Department of Hydrology and Hydraulic Engineering, Free University Brussels, Belgium, lfeyen@vub.ac.be

2. Inter University Program in Water Resources, Free University Brussels – Catholic University Leuven, Belgium

ABSTRACT

This paper presents a stochastic approach to delineate the capture zones for the well field 'Het Rot' at Nieuwrode (Belgium). The conductivity of the production aquifer is modelled as a random space function (RSF). For the conductivity of the overlying semi-pervious unit and the top phreatic aquifer homogeneous parameters are used. The parameters of the stochastic model and the homogeneous parameters are treated as random variables to account for the fact that they are unknown. Conductivity measurements for the production aquifer update the prior distributions for the structural parameters of this layer using Bayes' theorem, yielding posterior parameter distributions. Conditional realisations of the production aquifer conductivity field are generated using parameter sets sampled by Monte Carlo from the posterior distributions of the structural parameters. The realisations are combined with values for the other unknown parameters, sampled by Monte Carlo from the corresponding prior distributions. In a second application of Bayes' theorem, head observations are used to assign probability-based weights to the parameter realisations. Using particle tracking the 25-year capture zone and the well catchment are determined for each realisation, which are weighted by the associated probability of the parameter realisation. Statistical analysis of the set of weighted capture zones results in a capture zone probability distribution.

KEYWORDS

Bayesian inference, capture zone, groundwater, spatial stochastic approach

Introduction

The awareness of the importance to maintain the quality of groundwater in the vicinity of supply wells has resulted both in the U.S. and Europe in regulations aiming to protect groundwater supplies from accidental contamination. In 1986, a wellhead protection program was established by the U.S. Environmental Protection Agency (U.S. EPA) to help municipalities protect public groundwater water supplies from contamination. One of the regulations is the definition of the Wellhead Protection Area (WHPA), defined as the area around a public water supply well that contributes water to the well (U.S. EPA, 1987), also referred to as the well catchment. In Belgium, current regulations differentiate

between four types of protection zones, for each of which a list of prohibited activities are defined. Zone III or the chemical zone corresponds to the well catchment, limited to a maximum distance of 2 km to the well field (Vlarem II 1995). In practice it turns out that the other zones (Zones I and II) are very limited in extent. Therefore, we will delineate the total catchment for the well field 'Het Rot' at Nieuwrode (Belgium). In addition the 25-year capture zone is determined.

The shape of capture zones depends on the composite effects of many interacting processes, and on the geometry and hydrogeological properties of the system. In general, most numerical modelling approaches provide a deterministic estimate of the capture zones. However, due to the different sources of error inherent in any modelling exercise, the problem should be approached from a probabilistic point of view. A quantitative measure of the uncertainty associated with the model predictions is needed, which allows the regulatory instances to implement different degrees of protection for areas with different degrees of uncertainty. In this paper we apply the numerical stochastic approach presented in Feyen et al. (2001, 2002) to predict the capture zones for the well field 'Het Rot'.

Description of the study area: 'Het Rot'

The well field 'Het Rot' is located at Nieuwrode, northeast of Brussels. The well field consists of 8 abstraction wells with an average total discharge of approximately $5000 \text{ m}^3 \text{ d}^{-1}$. The area is hilly with the topography varying between 10 to 80 m above sea level. The battery of wells is located on a plateau, which forms a topographic water divide. South of the divide water is drained in the Winge-Molenbeek valley at approximately 2 km south of the well field, whereas water to the north of the divide is drained in the Demer river located approximately 3 km north of the site. The average precipitation in the area is 780 mm a year.

Water is abstracted from the Brusseliaan aquifer, a permeable formation consisting of coarse to fine sands. In the lower parts of this aquifer high permeable zones of very coarse sand occur, which have been deposited in broad southwest-northeast channels. The well field is installed in one such coarse sand channel. The Brusseliaan aquifer is bounded at the bottom by an impervious clay layer (Ieperiaan) and is overlain by a semi-pervious unit of fine sands with variable clay content (Tongeriaan). The top of the system consists of a phreatic aquifer of coarse to fine glauconitic sands (Diestiaan).

Deterministic hydrological model for the area

A hydrological model comprising 3 layers is constructed for the area using MODFLOW. The model boundaries and the location of the abstraction and observation wells are depicted in Figure 1. The horizontal discretisation comprises 202 columns and 160 rows, with a constant grid spacing of 50 m, covering an area of 80.8 km² and resulting in a total of 96960 grid cells in the model. A no flow boundary is defined at the bottom of the model to represent the impervious clay layer. The model is bounded in the north by the Demer river, in the east by the Motte river and in the south by the Winge-Molenbeek river system. For the west boundary of the model a constant head boundary is specified based on regional observations of the groundwater levels. The river system is modelled in layer 1 using the MODFLOW river package. Other small canals and ditches present in the area are simulated as drains.

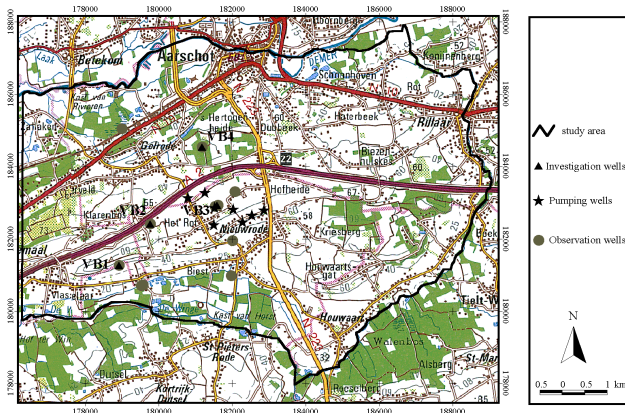


Fig. 1. Model boundaries and location of wells

Within the area of interest, the production aquifer is characterized by an average thickness of 30 m, showing an inclination of 0.6% in the north direction. The thickness of the different model layers varies throughout the area and is obtained by interpolating borehole data and geological profiles. The topography of the area is obtained from a DEM of scale 1:10.000. In a site investigation conducted before the actual well implementation, hydraulic conductivity values for the Brusselianaan have been determined using residual drawdown at four locations. The effective porosity of the model layers is assumed homogeneous. As no direct measurements at the site are available, porosity values are chosen based on lithological descriptions and the literature. The capture zones are determined with the particle-tracking code of MODPATH. Forward particle tracking is applied to ideal tracer particles uniformly distributed over the area. Well catchments are not static as their shape and extension depend on time-variable inputs such as recharge and pumping rates. However, generally, the temporal variations in these inputs are on a time scale that is considerable smaller than the typical travel time of a particle to the well. Such changes with a higher frequency are averaged out by the inertia of the aquifer. Therefore, the delineation of capture zones is usually based on long-term average conditions using a steady-state model, which is the approach followed here. All time-variable inputs and outputs are averaged over the period 1998-2000. The steady-state recharge to the groundwater table is obtained with WetSpass

(Batelaan and De Smedt, 2001), a physically-based water balance model for calculating the quasi-steady state spatially-variable groundwater recharge.

In the calibration process, the deterministic model is adjusted to match the observed heads at 13 observation locations, some of which have filters at different depths. Results show that the flow in the upper layers is mainly vertical and that the upper phreatic and lower confined aquifer are hydraulically connected, which is also indicated by the hydraulic heads observed in the two layers. The calibrated values for K of each layer are given in Table 1.

Main sources of uncertainty

A sensitivity analysis is conducted to evaluate the influence of the parameters on the flow model output and the capture zones. The main sources of uncertainty are the spatial distribution of the hydraulic conductivity in the Brusselianaan layer and the homogeneous conductivity values of the other model layers. The deposit history of the Brusselianaan aquifer, resulting in gully structures of high permeable zones (Houthuys, 1989), and the variability in the values of the four measurements, indicate that a spatially uniform value does not allow to characterize the true heterogeneity of the conductivity in the production aquifer. Therefore, the theory of RSF is used to represent the spatial variability of the conductivity in the production aquifer. The conductivity values of the other layers are assumed homogeneous. This is justified because the flow in the first two layers is mainly vertical, which implies that the influence of the spatial variability of the conductivity in the top two layers of the system on the location of the capture zones is limited compared to the influence of the spatial variability of the conductivity in the production aquifer.

The rate of natural recharge entering the system is an important factor for the location of the capture zones. However, given the detailed method to determine the average steady-state recharge, at present we do not consider uncertainty in the recharge rate. Pumping rates vary in time and do effect the well capture zones but the rates are generally well known and, as mentioned earlier, effects of time-varying pumping rates are filtered out on the longer time scale at which the capture zones are determined.

Uncertainty in the effective porosity is not accounted for, as no data are available that allow the effective porosity to be estimated in the inverse analysis. We would like to emphasise that the effective porosity only influences the time-related capture zones. The total well catchment does not depend on the aquifer effective porosity, as the latter does not influence the head distribution but only the velocity field, and as a result the time scale of the capture zones.

Bayesian inference of well catchments

The stochastic method starts with specifying distributions for the uncertain inputs of the groundwater flow model. For the homogeneous conductivity in the upper two layers lognormal distributions have been specified, with the mean and variance based on borehole descriptions and values found in the

literature (Bronders and De Smedt, 1991). We denote the unknown parameters of the Diestiaan and Tongeriaan layers by $\boldsymbol{\theta}_{D,T} = (K_D, K_T)$. The prior distributions for the parameters are described in Table 1.

Table 1. Calibrated values for the conductivities of the deterministic model, and prior parameter distributions (with values for the mean and variance) for the conductivities of the different layers in the stochastic model

Geological unit	Calibrated values K (m d ⁻¹)	Prior distributions (log K)
Diestiaan	7.5	$N(1, 0.5)$
Tongeriaan	0.12	$N(-1, 0.5)$
Brusseliaan	8	spat. stoch. app.

The log conductivity $Y = \log K$ of the Brusseliaan aquifer is modelled as a realisation of an isotropic, multi-Gaussian random space function with mean μ , variance σ^2 and correlation function $\rho(u) = \exp[-(u/\varphi)]$, where u is the Euclidian distance between two points, and φ is the integral scale. We denote the 3 structural parameters for the log K field of the Brusseliaan by $\boldsymbol{\theta}_B = (\mu, \sigma^2, \varphi)$ and their joint prior distribution by $[\boldsymbol{\theta}_B]$. The following functional forms have been used for the priors of the parameters in $\boldsymbol{\theta}_B$: (i) for the mean a normal distribution, $\mu | \sigma^2 \sim N(\mu_0, \sigma^2/\kappa_0)$, with hyperparameters $\boldsymbol{\theta}_\mu = (\mu_0, \kappa_0) = (0.778, 1)$, κ_0 being equivalent to the number of prior measurements; (ii) for the variance a scaled-inverse χ^2 -distribution, $\sigma^2 \sim \text{Sc-Inv-}\chi^2(\nu_0, \sigma_0^2)$, with hyperparameters $\boldsymbol{\theta}_{\sigma^2} = (\nu_0, \sigma_0^2) = (1, 1)$, ν_0 being the degrees of freedom and σ_0^2 being the scale; and (iii) for the integral scale a locally uniform distribution, $\varphi \sim U(\varphi_{ll}, \varphi_{ul})$, with hyperparameters $\boldsymbol{\theta}_\varphi = (\varphi_{ll}, \varphi_{ul}) = (100, 10000)$, φ_{ll} and φ_{ul} being the lower and upper limit, respectively. The values specified for the hyperparameters are based on values found in the literature about the Brusseliaan (Bronders and De Smedt, 1991).

Next, we update the prior distribution $[\boldsymbol{\theta}]$ with the information in the measurements $\mathbf{y} = (y_1, y_2, y_3, y_4)^T$, expressed by the likelihood function $L[\boldsymbol{\theta}|\mathbf{y}] \equiv [\mathbf{y}|\boldsymbol{\theta}]$, using Bayes' rule ($[\boldsymbol{\theta}|\mathbf{y}] \propto [\boldsymbol{\theta}]L[\boldsymbol{\theta}|\mathbf{y}]$), yielding the posterior distribution $[\boldsymbol{\theta}|\mathbf{y}]$ for the parameters. The likelihood $L[\boldsymbol{\theta}|\mathbf{y}]$ is a function of the parameter vector $\boldsymbol{\theta}$ and, for the spatial model assumed here for $Y = \log K$, has an expression given by the equation of a multivariate normal distribution. The posterior distribution $[\boldsymbol{\theta}|\mathbf{y}]$ reflects the uncertainty about the structural parameters after the log K data have been incorporated. The posterior $[\boldsymbol{\theta}|\mathbf{y}]$ does not define a standard probability distribution and inference by numerical simulation using a Monte Carlo sampling algorithm is adopted. The effect of updating the prior parameter distributions $[\boldsymbol{\theta}]$ with the log K data is shown in Figure 2, where histograms for μ , σ^2 and φ , obtained by sampling from the prior and posterior distributions, are presented. The histograms show that the 4 measurements considerably reduce the initial uncertainty about the mean and variance of the field, whereas the uncertainty about the integral scale remains fairly large.

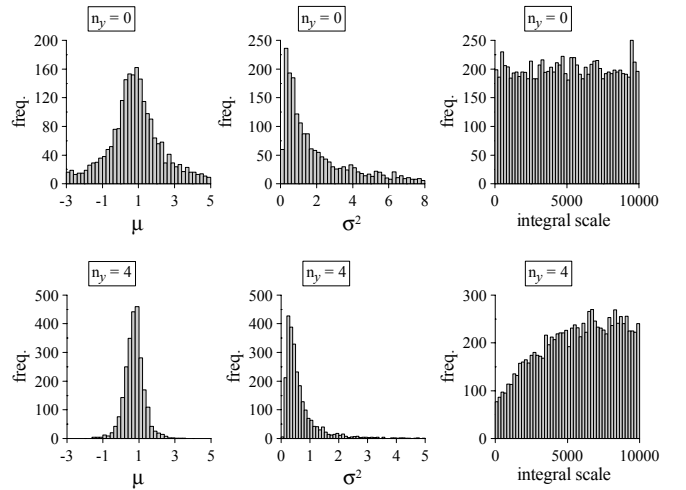


Fig. 2. Histograms of sampled parameter sets from parameter distributions $[\boldsymbol{\theta}|\mathbf{y}]$, with $n_y = 0$ and $n_y = 4$.

For each parameter set $\boldsymbol{\theta}_{B,i}$ sampled from $[\boldsymbol{\theta}_B|\mathbf{y}]$, equiprobable realisations of the log K field, honouring the 4 log K values at the measurement locations, are generated. We use $Y_{\boldsymbol{\theta}_{B,i},j}$ to denote the j th realisation obtained with the i th-sampled parameter set $\boldsymbol{\theta}_{B,i}$. All log K realisations generated with $\boldsymbol{\theta}_{B,i}$ are combined with the i th-sampled set of values for the parameters in $\boldsymbol{\theta}_{D,T}$, resulting in parameter realisations denoted by $\boldsymbol{\theta}_{ij} = (Y_{\boldsymbol{\theta}_{B,i},j}, \boldsymbol{\theta}_{D,B,i})$. The parameter realisations are used as input in the flow model, resulting in a corresponding head field for each realisation $\boldsymbol{\theta}_{ij}$. Observations of system state variables may suggest that some parameter realisations are more likely than others to represent reality. In this study the observations of state variables are limited to head observations, which are used to obtain probability-based weights for each parameter realisation. To apply Bayes' theorem, a probability density function needs to be specified for the head residuals, which is consistent with the information available about the errors. The joint conditional distribution $[\boldsymbol{\varepsilon}_h|\boldsymbol{\theta}, \boldsymbol{\psi}]$ describes the distribution of the residuals, given $\boldsymbol{\theta}$ and the parameters $\boldsymbol{\psi}$ of the error model. This expression, seen as a function of $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$, is called the likelihood function of the unknown quantities $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$, and expresses the probability of observing the residuals, given the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$. Since the structure of the flow model is known, $L[\boldsymbol{\theta}, \boldsymbol{\psi}|\mathbf{h}]$ is actually proportional to the probability distribution of the observations \mathbf{h} . We assume that the combination of forward modelling and measurement errors is unbiased and Gaussian, in which case the likelihood is given by

$$L[\boldsymbol{\theta}, \boldsymbol{\psi}] = (2\pi)^{-n_h/2} (\det \mathbf{V}_{\boldsymbol{\varepsilon}_h})^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{\varepsilon}_h(\boldsymbol{\theta}))^T \mathbf{V}_{\boldsymbol{\varepsilon}_h}^{-1}(\boldsymbol{\varepsilon}_h(\boldsymbol{\theta}))\right)$$

where $\mathbf{V}_{\boldsymbol{\varepsilon}_h}$ is the error covariance matrix of dimension n_h , and $\boldsymbol{\psi} = \mathbf{V}_{\boldsymbol{\varepsilon}_h}$. Assuming that the head residuals are uncorrelated between the observations locations, $\mathbf{V}_{\boldsymbol{\varepsilon}_h}$ is diagonal, with the terms in the diagonal given by the respective variances $\sigma_{\boldsymbol{\varepsilon}_h,i}^2$ at each location. We treat the error variances $\sigma_{\boldsymbol{\varepsilon}_h,i}^2$ explicitly as unknown and integrate out their effect. The prior distribution for the noise values $\sigma_{\boldsymbol{\varepsilon}_h,i}^2$ is given by a scaled-inverse χ^2 -distribution, $\sigma_{\boldsymbol{\varepsilon}_h,i}^2 \sim \text{Sc-Inv-}\chi^2(\nu_{\boldsymbol{\varepsilon}_h,0}, \sigma_{\boldsymbol{\varepsilon}_h,0}^2)$, with the degrees of freedom $\nu_{\boldsymbol{\varepsilon}_h,0}$ equal to one and the scale $\sigma_{\boldsymbol{\varepsilon}_h,0}^2$ of each error

component equal to the variance of the time-variable observations at each observation location. Evaluating the likelihood function $L[\theta, \psi]$ for each realisation θ_{ij} , integrating out the error variances and normalizing the probabilities, results in the conditional distribution $[\theta | y, h]$, which reflects the probability $p(\theta_{ij} | y, h)$ associated with each realisation θ_{ij} to represent reality after incorporating the conductivity and head data.

In prediction, each capture zone is weighted by the conditional probability $p(\theta_{ij} | y, h)$ associated with the corresponding parameter realisation θ_{ij} . Statistical processing of the set of weighted capture zones results in the predictive capture zone distribution (capd), defined by

$$[CAP(x, t) | y, h] = \frac{\sum_{i=1}^m \sum_{j=1}^n p(\theta_{ij} | y, h) (I(x, t) | y, h)_{ij}}{\sum_{i=1}^m \sum_{j=1}^n p(\theta_{ij} | y, h)}$$

where m is the number of parameter sets sampled from $[\theta_B | y]$ and $[\theta_{D,T}]$, n is the number of generated realisations for each sampled parameter set $\theta_{B,i}$, and $p(\theta_{ij} | y, h)$ is the conditional or posterior probability attached to parameter realisation θ_{ij} . Parameter realisations that better reproduce the head observations will have higher posterior probabilities and therefore have more weight in the predictions. The term $(I(x, t) | y, h)_{ij}$ is the probability of intake I by the well and equals one if the particle released in the point $x = x_1$ is captured by the well within the time span $t = t_1$, and zero otherwise. The capd $[CAP(x, t) | y, h]$ defines at a point $x = x_1$ and for a time $t = t_1$ the probability $p(CAP(x, t) | y, h)$ that an inert particle released at this point reaches the well within the specified time span t_1 .

Figures 3 and 4 show the simulated probability distributions for the 25-year capture zone and for the catchment of the well field, respectively. For a travel time of 25 years, the capture zones of the individual wells can still be distinguished, whereas the total catchments of the individual wells overlap, forming a united catchment for the well field. The capture zones exhibit most uncertainty southeast of the well field. This is because this is the main contributing area of the well field, which results in longer particle pathlengths toward the wells and implies that more heterogeneity is encountered by the particles.

The capd provides a quantitative measure of the uncertainty associated with the model predictions. Deterministic approaches neglect the uncertainty that is inherent in any modelling practice and underestimate the uncertainty associated with the model predictions. If, for example, the deterministic model parameters result in smaller estimates of the capture zone, e.g., in the case of a smaller estimated mean conductivity or when neglecting spatial correlation, this could lead to under-protection of the well field. Even though the predictions of the stochastic approach can be more conservative, they allow the regulatory instances to implement different degrees of protection for areas with different degrees of uncertainty.

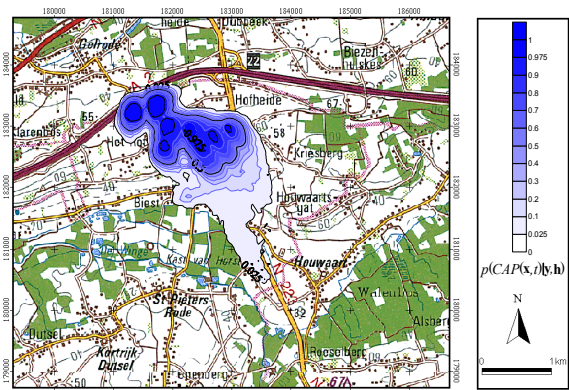


Fig. 3. Stochastic 25-year capture zone for the well field

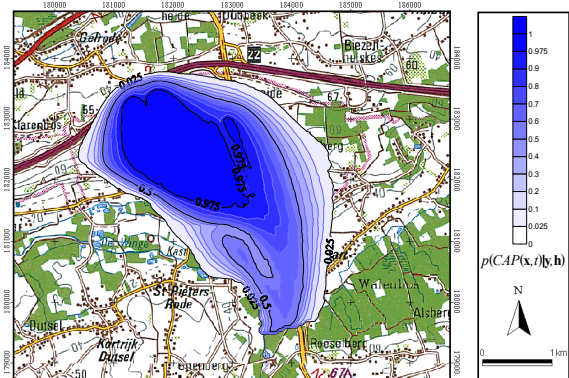


Fig. 4. Stochastic catchment for the well field

Acknowledgements

The authors wish to thank K. Beven and J. Freer for the access to the parallel system at the IENS of Lancaster University. The first author wishes to acknowledge the Fund for Scientific Research – Flanders for providing a Research Assistant Scholarship. We wish to acknowledge the VMW for providing the data about the site.

REFERENCES

- BATELAAN, O. & DE SMEDT, F., 2001. WetSpa: a flexible, GIS based, distributed recharge methodology for regional groundwater modelling. In: H. Gehrels *et al.* (Eds), *Impact of Human Activity on Groundwater Dynamics*. IAHS **269**, 11-17.
- BRONDERS, J., & DE SMEDT, F., 1991. Geostatistical analysis of the hydraulic conductivity of aquifers in the centre of Belgium (in Dutch), *Water*, **59**, 127-132.
- FEYEN, L., K.J. BEVEN, F. DE SMEDT, & J. FREER, 2001. Stochastic capture zone delineation within the GLUE-methodology: Conditioning on head observations, *Water Resources Research*, **37**(3), 625-638.
- FEYEN, L., P. J. RIBEIRO JR., F. DE SMEDT, & P. J. DIGGLE, 2002. Bayesian methodology to stochastic capture zone determination: conditioning on transmissivity measurements. *Water Resources Research* (in press).
- HOUTHUYS R., 1989. Vergelijkende studie van de afzettingsstructuur van getijdenzanden uit het Eoceen en de huidige Vlaamse banken, *Aardkundige mededelingen* **5**.
- U.S. ENVIRONMENTAL PROTECTION AGENCY 1987. *Guidelines for delineation of wellhead protection areas*. Washington, D.C.: U.S. EPA, Office of Ground-Water Protection.
- VLAREM II 1995. Besluit van de Vlaamse Regering van 1 juni 1995 houdende algemene en sectorale bepalingen inzake milieuhygiëne. Belgisch Staatsblad **31VII**.