

Ijking van grondwatermodellen met Monte Carlo en mathematische optimalisatie

Inleiding

Grondwatermodellen worden doorgaans met de hand geijkt. Als we de rapporteurs mogen geloven, dan schijnt dit meestal binnen enkele tientallen runs te lukken. Dit is verbazingwekkend, daar het via 'gissen en missen' gelijktijdig optimaliseren van meerdere parameters nagenoeg onmogelijk is. Het kan dan ook behoorlijk misgaan. Dit komt echter alleen aan het licht wanneer er van een gebied twee geijkte modellen bestaan en deze volslagen afwijkende voorspellingen



I.R. T. N. OLSTHOORN
Afdeling Procesontwikkeling,
Sector Hydrologie, GWA

opleveren [Olsthoorn, 1989]. Gezien de risico's is het van belang aandacht te besteden aan objectivering van het ijken van modellen.

Ijking is in dit artikel synoniem voor modelkalibratie. Dit is optimalisatie van modelparameters c.q. het 'draaien aan modelknoppen' totdat de relevant geachte berekende en de gemeten waarden nog slechts minimaal van elkaar verschillen. In dit artikel worden twee ijkmethoden gedemonstreerd: een Monte Carlo-methode en een directe mathematische optimalisatie van modelparameters. Beide methoden worden toegepast op het grondwatermodel van de Amsterdamse Waterleiding-duinen.

Voorwerk en nawerk bij modelijking

De te optimaliseren parameters representeren bij deterministische grondwatermodellen bepaalde fysische grootheden. Optimalisatie van deze parameters is echter pas zinvol nadat men zich zekerheid heeft verschaft over de mate waarin het gekozen model een geschikte representatie vormt van de werkelijkheid. Toetsing hiervan moet aan elke ijking voorafgaan.

Bij een gegeven model bestaat het ijken uit het gelijktijdig optimaliseren van een aantal parameters. Dit geschiedt door het minimaliseren van een nader te definiëren doelfunctie, waarvan de waarde geschikt moet zijn voor beoordeling van de juistheid van het model.

De meest gebruikelijke doelfunctie bij grondwatermodellen is de som van de kwadraten van het verschil tussen de gemeten en berekende grondwaterstijghoogten op alle bemeeten lokaties en tijdstippen. Per lokatie en tijdstip kan zonodig een weegfactor worden toegepast. Andere

Samenvatting

Twee ijkmethoden voor grondwatermodellen worden beschreven, een Monte Carlo-methode en een mathematische parameteroptimalisatie. Beide methoden worden toegepast op het grondwatermodel van de Amsterdamse Waterleiding-duinen.

De Monte Carlo-methode blijkt relatief onnauwkeurig en wordt onbruikbaar wanneer meer dan circa 5 parameters tegelijk moeten worden geoptimaliseerd. De toegepaste mathematische optimalisatie is efficiënter en nauwkeuriger. Zij heeft de voorkeur bij de kalibratie van modellen. Mathematische kalibratie is de enige keus bij gelijktijdige optimalisatie van een groter aantal parameters. Het gepresenteerde mathematische optimalisatie-algoritme is eenvoudiger en werkt beter dan bekende, meer geavanceerde methoden. Het algoritme is bovendien gemakkelijk te implementeren, zonder aanpassing van het (grondwater)model. De gegeven technieken lenen zich dan ook voor toepassing in praktijk. Zij zijn in dat opzicht als een doorbraak te beschouwen. Zij verdienen verre de voorkeur boven niet-reproduceerbare handmatige probeertechnieken.

meetgegevens en voorinformatie over de parameters kunnen in de doelfunctie worden opgenomen (zie hiertoe [Carrera en Neuman, 1986a]). In dit artikel beperken we ons tot de basale vorm:

$$F = \sum_{j=1}^m \omega_j (\hat{h}_j - h_j)^2 \quad (1)$$

Ter vergemakkelijking van de beoordeling van het ijkresultaat berekenen we tevens de stijghoogte deviatie [m]:

$$\text{dev} = \sqrt{\frac{\sum_{j=1}^m \omega_j (\hat{h}_j - h_j)^2}{\sum_{j=1}^m \omega_j}} \quad (2)$$

Hierin is de $\hat{h}_j$ berekende en h_j de gemeten stijghoogte op een bepaalde lokatie en eventueel tijdstip j . ω is een per lokatie en eventueel per tijdstip te kiezen gewichtsfactor. m is het aantal beschikbare metingen.

De volgende stap is de keuze van de te optimaliseren parameters. Hun aantal moet worden beperkt. Het moet in elk geval belangrijk kleiner zijn dan het aantal meetwaarden [Carrera en Neuman, 1986a]. Naarmate het aantal parameters stijgt, wordt de passing van model en metingen beter, maar neemt de betrouwbaarheid van de geoptimaliseerde parameters af. Het beste model levert een goede passing bij een zo klein mogelijk aantal parameters [Sun & Yeh, 1985; Yeh & Yoon, 1981].

Het reduceren van het aantal vrijheidsgraden of parameters wordt regularisatie genoemd. Een veel gebruikte methode is zonering van het modelgebied. Per zone moet de gemiddelde waarde van de

parameter worden bepaald.

Een andere mogelijkheid is de dikte van het pakket als 'hard' gegeven te hanteren met de gemiddelde doorlatendheid als te optimaliseren parameter [Olsthoorn, 1982; Olsthoorn en Kusse, 1986]. Hetzelfde kan men doen voor slecht doorlatende lagen, eventueel afhankelijk van hun samenstelling en voorspanning [Heij en Kester, 1977].

Het gebruik van specifieke doorlatendheden en laagweerstand leidt tot een minimaal aantal gelijktijdig te optimaliseren parameters. Mocht bijvoorbeeld uit boorgegevens blijken, dat de samenstelling van een formatie ruimtelijk belangrijk varieert, dan kunnen alsnog deelgebieden worden onderscheiden. In dit geval ontstaat een combinatie met voornoemde zonering.

Een laatste belangrijke vorm van regularisatie van de parameters is overschakeling op hun logaritmische waarde. Hiermee verdwijnen schaalverschillen en worden wijzigingen in feite relatief ten opzichte van hun grootte [Olsthoorn, 1982; Carrera en Neuman, 1986].

De mate waarin parameters optimaliseerbaar zijn, kan worden afgeleid uit de gevoeligheid van de doelfunctie voor variaties in hun waarde, wat weinig te maken heeft met de gevoeligheid van scenarioresultaten voor dezelfde parameters. Het laatste vergt altijd afzonderlijke aandacht in de vorm van een onzekerheidsanalyse achteraf.

Parameters waarvoor de doelfunctie gevoelig is, hoeven nog niet per sé onafhankelijk van elkaar optimaliseerbaar te zijn. Verschillende parametercombinaties kunnen dan een nagenoeg gelijk ijkresultaat opleveren. Het kan niet a priori worden uitgesloten dat geheel verschillende parametercombinaties tot een (lokaal) minimum van de doelfunctie

leiden. In beide gevallen is het ijkresultaat niet uniek.

Het al dan niet uniek zijn van de geoptimaliseerde parametercombinatie kan slechts achteraf worden onderzocht, in de regel door de ijking vanuit (extreem) verschillende parameterstartwaarden te herhalen. Dergelijk onderzoek is essentieel bij mathematische parameteroptimalisatie.

Monte Carlo-parameter-optimalisatie

De Monte Carlo-methode gaat uit van een groot aantal modelruns. Elke run wordt uitgevoerd met een bepaalde parametercombinatie. Deze combinaties worden verkregen door trekking uit de statistische verdeling die voor elke parameter afzonderlijk en a priori wordt gekozen. Na iedere modelrun wordt de waarde van de doelfunctie berekend. Uiteindelijk worden de parametercombinaties geselecteerd die de laagste waarde van de doelfunctie hebben opgeleverd. Hieronder volgt het recept voor de Monte Carlo-methode:

- Kies voor elk van de M te optimaliseren parameters een statistische verdeling die de verwachting weerspiegelt van de waarden die de parameter aan kan nemen (tabel I).

TABEL I – Gebruikte uitgangswaarden parametercombinaties.

Param. Dim	k1 [m/d]	c1 [d]	k2 [m/d]	c2 [d]
Min	5	5.000	2	1.000
Max	50	100.000	20	15.000

- Kies het gewenste aantal modelruns, bijv. N=1000.
 - Verdeel het traject van elk van de parameters in N stukjes van gelijke kans.
 - Trek binnen elk stukje van gelijke kans at random een waarde. Doe dit voor alle parameters.
 - Schud per parameter de getrokken trekkingen door elkaar en voeg de lijsten samen tot parametercombinaties. Op dit punt beschikt men over N parametercombinaties, samengesteld volgens de zogenoemde Latin Hypercube methode (zie ook [Peck *et al.*, 1988]).
 - Draai het grondwatermodel voor elk van de N parametercombinaties. Voor elk van de resultaten wordt de waarde van de doelfunctie berekend en genoteerd bij de parameterset die haar heeft voortgebracht (tabel II).
 - Sorteer de lijst (tabel III) en verricht nader onderzoek met de verkregen beste combinaties.
- De implementatie van de Monte Carlo-methode geschiedt buiten het model om. (De lijst parametercombinaties wordt tevoren aangemaakt.) Een BATCH-com-

TABEL II – Parametersets met berekende deviatie (vgl. 2).

SetNr []	k1 [m/d]	c1 [d]	k2 [m/d]	c2 [d]	dev [m]
1	9,92	12.243	5,39	6.896	0,36
2	10,48	19.229	3,31	5.437	0,52
3	20,07	21.491	9,42	5.727	0,52
4	6,26	4.668	4,51	2.656	0,95
5	9,20	9.408	8,44	11.501	0,35
6	24,37	41.037	6,60	2.971	0,70
7	51,51	66.715	7,50	2.386	1,04
8	13,63	21.725	11,12	6.651	0,44
9	14,10	21.077	6,08	3.839	0,51
10	30,38	48.595	18,39	5.485	0,73
635	24,81	32.084	7,27	5.249	0,56
636	14,82	19.199	7,23	4.568	0,48
637	28,59	19.478	5,78	2.251	0,93
638	33,09	46.827	4,65	4.680	0,64

TABEL III – Beste 10 en slechtste 4 parametersets met berekende deviatie (vgl. 2).

SetNr []	k1 [m/d]	c1 [d]	k2 [m/d]	c2 [d]	dev [m]
381	13,0	11.273	5,5	8.324	0,34
85	17,6	21.959	7,1	10.449	0,34
11	18,7	22.178	5,3	8.682	0,34
610	14,6	17.709	7,3	8.949	0,34
289	15,6	13.490	4,2	8.530	0,34
121	12,4	12.264	4,8	6.984	0,35
431	15,3	14.350	5,6	8.024	0,35
5	9,2	9.408	8,4	11.550	0,35
580	10,9	8.671	5,6	9.400	0,35
424	15,5	18.175	8,5	10.014	0,35
101	6,4	5.168	9,9	1.380	1,13
597	11,2	6.822	5,1	1.251	1,14
220	38,2	50.558	10,0	1.214	1,15
621	9,3	5.148	6,4	1.268	1,23

mando regelt de procedure als volgt. Het start een FORTRAN-programma dat steeds een volgende parameterset uit de lijst leest en de modelinvoer hieraan aanpast. Het start vervolgens het grondwatermodel. Hierna wordt een ander FORTRAN-programma opgestart. Dit vist de stijghoogten ter plaatse van meetlokaties uit de uitvoer, berekent de waarde van de doelfunctie en voegt deze met de desbetreffende parameterwaarden aan een bestand toe. Deze cyclus herhaalt zich automatisch tot alle parametercombinaties aan de beurt zijn geweest. Het resultaat is een lijst met parameterwaarden waaraan de waarde van de doelfunctie is toegevoegd (tabel II). Deze lijst wordt gesorteerd naar doelfunctiewaarde en gebruikt voor verdere analyse (tabel III).

Mathematische parameter-optimalisatie

De M parameters kunnen ook worden geoptimaliseerd door de doelfunctie mathematisch te minimaliseren. Deze doelfunctie stelt een (M+1)-dimensionaal vlak voor in een M-dimensionale parameteruimte. De procedure moet hiervan het laagste punt vinden. Over dergelijke optimalisaties zijn veel artikelen geschreven. Toepassing daarvan

wordt beperkt door hun complexiteit, zowel wiskundig als modelmatig. De hier te demonstreren methode is echter simpel en wordt, evenals de Monte Carlo-methode, geheel buiten het model om geïmplementeerd. Zoals aangetoond aan de hand van hun eigen voorbeeld is de ontwikkelde methode behalve eenvoudiger ook efficiënter dan de complexe aanpak van Carrera en Neuman [1986] (zie [Olsthoorn, 1994]).

De gewogen afwijking f_j tussen berekende $\hat{h}_j$ en gemeten stijghoogte h_j wordt m bij waarnemingen:

$$f_j = \sqrt{\omega_j}(\hat{h}_j - h_j) \quad j \in \{1..m\} \quad (3)$$

De doelfunctie F wordt dan in vectorvorm:

$$F = \underline{f}^T \underline{f} \quad (4)$$

De gradiënt van F binnen een gebiedje ter grootte $\Delta \underline{p}$ rond de huidige (log)-parameter vector $\underline{p}$ volgt uit een Taylorreeksontwikkeling:

$$\underline{g}(\underline{p} + \Delta \underline{p}) = \underline{g}(\underline{p}) + \underline{H}(\underline{p}) \cdot (\Delta \underline{p}) + O(\|\Delta \underline{p}\|) \quad (5)$$

Termen van hogere orde verwaarlozend en $\Delta \underline{p}$ zodanig gekozen dat deze gradiënt verdwijnt, levert:

$$\underline{H}(\underline{p}) \cdot (\Delta \underline{p}) = -\underline{g}(\underline{p}) \quad (6)$$

de standaard multi-dimensionale Newton-Raphson-relatie voor het vinden van het minimum van een functie. Hieruit volgt $\Delta \underline{p}$ mits $\underline{H}(\underline{p})$ en $\underline{g}(\underline{p})$ zijn gegeven en $\underline{H}(\underline{p})$ niet singulier is. De nieuwe parameter-vector wordt dan $(\underline{p} + \Delta \underline{p})$.

Daar de doelfunctie F de speciale vorm $\underline{f}^T \underline{f}$ heeft, kan hij worden gediifferentieerd om zo uitdrukkingen te krijgen voor de gradiënt en de Hessiaanse matrix $\underline{H}$ [Jennings & McKeown, 1992]:

$$\underline{g} = 2 \underline{f}^T \underline{f} \quad (7)$$

$$\underline{H} = 2(\underline{f}^T \underline{f} + \underline{B}) \quad (8)$$

Waarin $\underline{f}$ de (m x n)-matrix van Jacobi en $\underline{B}$, de som over alle waarnemingen $j \in \{1..m\}$ van de f_j -gewogen bijdrage van nietlineariteiten van $f_j(\underline{p})$ [Jennings & McKeown, 1992].

De Newton-Raphson-relatie wijzigt nu in:

$$(\underline{f}^T \underline{f} + \underline{B}) \cdot (\Delta \underline{p}) = -\underline{f}^T \underline{f} \quad (9)$$

De eerste matrix binnen $\underline{H}$ is positief-definiet. Dit hoeft niet te gelden voor de tweede matrix, $\underline{B}$. Deze wordt echter nul wanneer de berekende stijghoogten exact

gelijk zijn aan de meetwaarden, dus wanneer $f_j = 0$ voor alle j , zoals bij een goed ijkbaar model bij benadering het geval is in de buurt van het optimum. $\underline{B}$ wordt daarom normaliter verwaarloosd, wat de standaard Gauss-Newton-vergelijking oplevert.

$$\Delta p = -(\underline{J}^T \underline{J})^{-1} \underline{J}^T f \quad (10)$$

De Gauss-Newton-vergelijking vereist niet-singulariteit van de matrix $\underline{J}^T \underline{J}$. Dit kan niet worden gegarandeerd onder praktische omstandigheden. Een effectieve methode om dit soort problemen het hoofd te bieden danken we aan Levenberg en Marquardt (zie bijv. [Jennings & McKeown, 1992]). Hun aanpassing bestaat uit het optellen van een positief definitie diagonaalmatrix bij: $\underline{J}^T \underline{J}$.

$$\Delta p = -(\underline{J}^T \underline{J} + \mu \underline{I})^{-1} \underline{J}^T f \quad (11)$$

Zoals kan worden bewezen [Jennings & McKeown, 1992, p. 343] wordt niet- of bijna-singulariteit op deze wijze voorkomen, zelfs voor zeer kleine waarden van $\mu > 0$. Naarmate men μ groter kiest, neemt $\underline{J}^T \underline{J}$ af in vergelijking met $\mu \underline{I}$. De parameteraanpassing, Δp , wordt dan kleiner en wijst bovendien steeds meer in de richting van de steilste afdaling van de doelfunctie F .

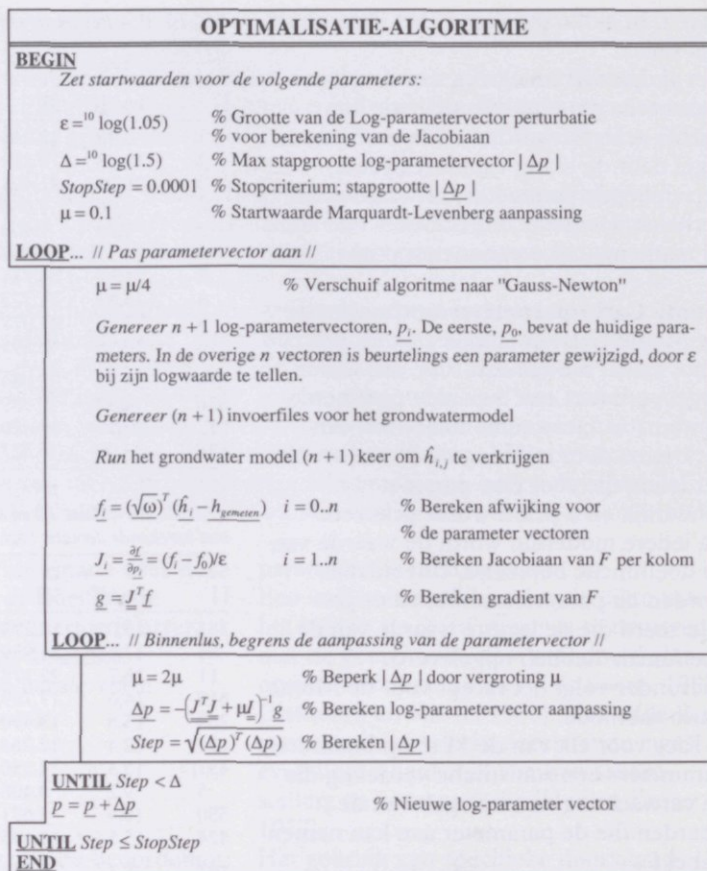
Zolang de parameterwaarden ver van hun optimum verwijderd zijn en de modelresultaten te veel afwijken van de meetwaarden, is $(\underline{J}^T \underline{J})$ slechts een grove benadering van de Hessiaanse matrix. Deze kan dan onjuiste parameteraanpassingen geven. Dit wordt voorkomen door de parameteraanpassing te begrenzen tot een 'vertrouwens'-gebiedje rond de huidige parametervector.

Deze begrenzing en daarmee een beheerste en robuuste minimalisatie wordt nu verkregen door aanpassing van de grootte van μ . Behalve de stapgrootte wordt daarmee automatisch ook de richting van Δp beheerst.

Het feit dat de Gauss-Newton-vergelijking per stap voor meerdere μ -waarden moet worden opgelost is niet van belang; bij enkele tientallen parameters is de hoeveelheid werk die hiermee gemoeid is verwaarloosbaar ten opzichte van een enkele modelrun. Bovendien varieert μ slechts geleidelijk van iteratie tot iteratie, waardoor in de praktijk gemiddeld per stap maar twee van dergelijke matrixinversies nodig zijn.

Bepalend voor de methode is het optimalisatie-algoritme, niet de berekening van de Jacobi-matrix $\underline{J}$. Vandaar dat hier voor de Jacobi-matrix de meest eenvoudige methode is toegepast. Deze bestaat uit het

Afb. 1 - Mathematisch kalibratie-algoritme.



afrekken van de stijghoogten berekend met de huidige parameterset van die waarbij steeds één van de logaritmische parameters met een kleine waarde ε is vergroot, en het verschil te delen door deze stapgrootte. Dit vergt $(M + 1)$ modelruns per iteratie, met M het aantal te optimaliseren parameters. (Zie voor meer geavanceerde methoden bij [Carrera & Neuman, 1986b; Sykes, 1985; Sun & Yeh, 1985]).

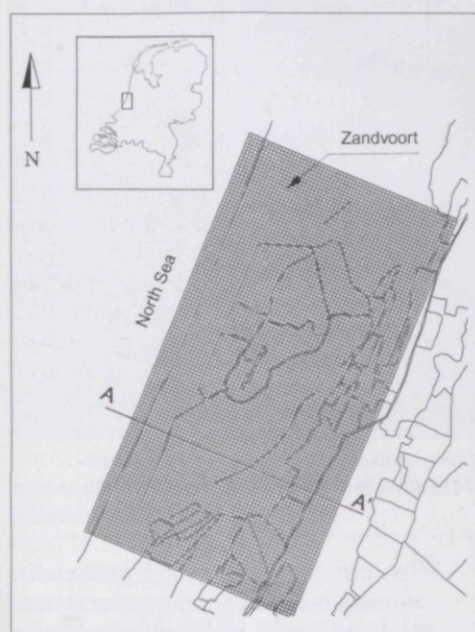
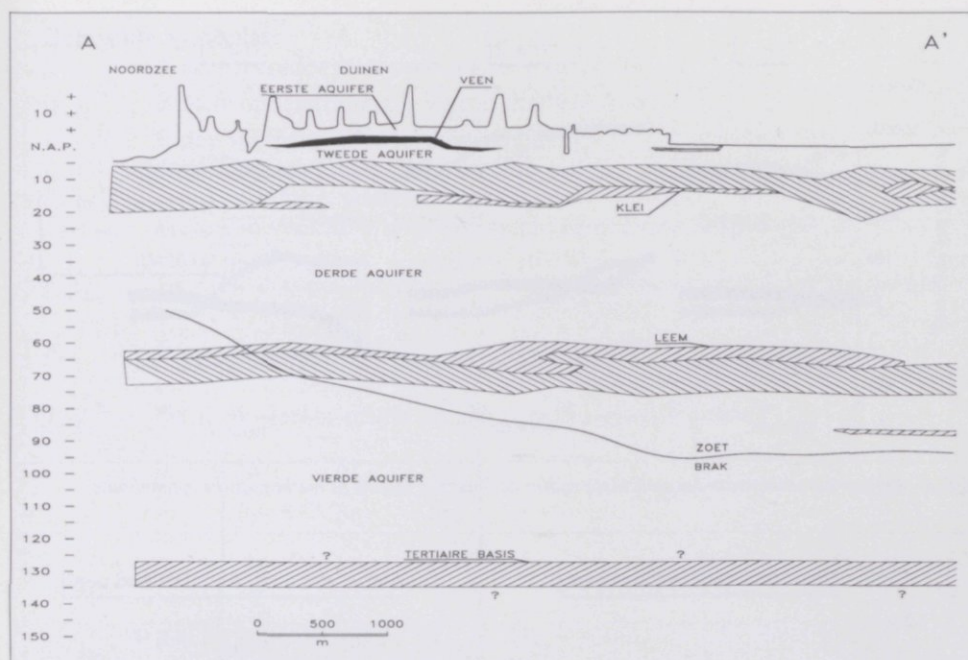
Het ontwikkelde algoritme is in afbeelding 1 uitgewerkt. Wij hebben het geïmplementeerd in het standaard matrixpakket MATLAB. De implementatie volgt strikt het gegeven algoritme en is even kort. MATLAB laat het grondwatermodel de runs uitvoeren die nodig zijn om de Jacobi-matrix te berekenen. Dat wil zeggen MATLAB berekent voor elke iteratie een nieuwe parametervector, schrijft deze naar een bestand en start een BATCH-commando. Dit start op zijn beurt een FORTRAN-programma dat de weggeschreven parameters leest en een gewijzigde invoer aanmaakt voor het grondwatermodel, dat vervolgens zelf wordt gestart. Hierna wordt een ander FORTRAN-programma gestart dat de berekende stijghoogten voor de meetlokaties uit de modeluitvoer vist en naar apart bestand wegschrijft. MATLAB neemt vervolgens het heft weer over en gebruikt de berekende stijghoogten voor

het samenstellen van de Jacobi-matrix en het afwikkelen van het optimalisatie-algoritme. De gevolgde aanpak maakt optimaal gebruik van het standaard grondwatermodel en het standaard matrixprogramma. Het model hoeft voor de optimalisatie op geen enkele manier te worden gewijzigd.

Het grondwatermodel van de Amsterdamse waterleidingduinen is beschreven door Olsthoorn *et al.* [1993]. Het wordt door-gerekend met het eindige differentie-model MODFLOW en bevat 4 aquifers (afb. 2). Het rechthoekige netwerk bestaat uit 64×124 vierkante cellen met ribben van 100 m (afb. 3).

De hier te beschrijven kalibratie van het model betreft eenvoudigheidshalve alleen de gemiddelde situatie tussen 1974 en 1990, een periode waarin de omstandigheden nagenoeg onveranderd zijn geweest.

Primair doel van het model was en is accurate modellering van het freatische grondwatervlak. De stijghoogte in aquifer 3 is voor dit onderzoek vastgehouden op de gemeten waarden, zoals geïnterpoleerd via kriging. Berekend is de stijghoogte in de bovenste twee aquifers. Gekalibreerd zijn de daarvoor relevante parameters. Dit zijn achtereenvolgens: de doorlatendheidscoëfficiënt k_1 van het zand boven het Hollandveen, de doorlatendheidscoëfficiënt k_2 van het zand daaronder doch boven de



Afb. 3 - Ligging model en netwerk.

kleilaag. De weerstandwaarde behorend bij de gemiddelde dikte van het Hollandveen op circa NAP en de weerstand behorende bij de gemiddelde dikte van de kleilaag tussen circa 10 en 20 m beneden NAP.

Het doorlaatvermogen T_{xy} op een willekeurige plek volgt dan uit:

$$T_{xy} = k_{gem} D_{xy} \quad (12)$$

en de weerstand op een willekeurige plek wordt:

$$c_{xy} = c_{gem} \frac{D_{xy}}{D_{gem}} \quad (13)$$

De ruimtelijk variërende dikte van elke laag D_{xy} is afkomstig van geologische kaarten [Stuyfzand, 1987]. Zij zijn verder als harde gegevens gehanteerd.

De kalibratie bestaat derhalve uit optimalisatie van vier parameters. Dit geldt zowel voor de Monte Carlo-methode als voor de mathematische optimalisatie. Met de mathematische optimalisatie kan echter zonder wezenlijke complicaties een groter aantal parameters worden geoptimaliseerd [Carrera & Neuman, 1986; Olsthoorn, 1994]. Met de Monte Carlo-methode is dit niet mogelijk. In de toekomst zullen onze modelkalibraties daarom met de mathematische methode worden uitgevoerd.

Voorafgaande aan de parameteroptimalisatie is het model getoetst door vergelijking van het resultaat van proefberekeningen met de door kriging uitgevoerde stijghoogte-interpolatie. Afwijkingen die niet konden worden toegeschreven aan krigingmethode en waarvan duidelijk was dat zij niet zouden kunnen worden weggelaten door parameteroptimalisatie zijn in het veld nader onderzocht met extra boringen en sonderingen. Dit leidde tot aanpassing van de geologische kaarten. Pas nadat een bevredigend model was verkregen is de parameteroptimalisatie uitgevoerd. De kalibratie heeft slechts betrekking op het zuidelijk deel van het duingebied. Hier waren stijghoogten beschikbaar op 154 lokaties, 42 daarvan boven het Hollandveen (aquifer 1) en de rest eronder en ernaast (aquifer 2). Elke lokatie is gewogen met de reciproke waarde van het aantal waarnemingslokalities binnen een straal van 100 m. Dit voorkomt op simpele wijze het overmatig

Afb. 2 - Doorsnede voor schematisatie grondwatermodel.

meetellen van dicht bij elkaar gelegen clusters waarnemingsputten.

Resultaten Monte Carlo-optimalisatie

Bij Monte Carlo-methode houdt toeval een aanzienlijk aandeel in het resultaat. Vooral de verkregen waarden van de minder gevoelige parameters kunnen sterk afwijken van het werkelijke optimum. Om deze redenen moet een aantal van de beste parametersets als plausibel worden beschouwd en verder in samenhang worden onderzocht. Bij de ijkking wordt steeds met de logaritme van de parameters gewerkt. Gekozen is voor een uniforme verdeling van de logaritme van de parameters, tussen een vooraf gekozen minimum en maximum waarde (tabel I). Elke parameter krijgt op deze wijze zijn eigen verdeling.

Tabel II geeft de eerste tien en de laatste vier van de 638 getrokken parametersets, met daarnaast de bijbehorende gewogen afwijking (vgl. 2) tussen gemeten en berekende stijghoogten.

Tabel III geeft na sortering van de lijst de beste tien en de slechtste vier van de 638 gebruikte parametercombinaties. Afhankelijkheid tussen geoptimaliseerde parameters kan worden onderzocht door opstellen van een correlatiematrix voor de beste parametersets (zie tabel IV).

TABEL IV - Correlatiematrix voor de 10 beste parametersets van tabel II.3.

r	k1	c1	k2	c2
k1	1	0.91	-0.17	-0.18
c1	0.91	1	0.17	0.07
k2	-0.17	0.17	1	0.80
c2	-0.18	0.07	0.80	1

De sterke onderlinge afhankelijkheid van de parameters k_1 en c_1 impliceert dat deze parameters moeilijk afzonderlijk kunnen worden geoptimaliseerd (althans met de Monte Carlo-methode). Er zitten dan ook tussen de 10 beste sets verschillende die nagenoeg hetzelfde ijkresultaat opleveren (zie tabel 3). De sterke correlatie tussen k_1 en c_1 is goed hydrologisch verklaarbaar. Een bepaalde waterstand in de 1e aquifer (boven de veenlaag, zie afb. 2) kan zowel worden verkregen met een kleine k_1 in combinatie met een kleine c_1 , als met een grote k_1 in combinatie met een grote c_1 . Het optimalisatieprobleem is derhalve niet optimaal geconditioneerd, wat ertoe leidt dat de Monte Carlo-methode geen scherp optimum vindt. Dit blijkt bijvoorbeeld ook

TABEL V – Gemiddelde en standaardafwijking over alle parametersets en over de 10 beste afzonderlijk.

Sets		k1 [m/d]	c1 [d]	k2 [m/d]	c2 [d]
Alle	Gem.	20,4	25.889	7,9	4.831
Alle	StdAfw	10,3	17.571	4,0	3.036
Beste 10	Gem.	14,3	14.948	6,2	9.086
Beste 10	StdAfw	2,8	4.612	1,4	1.236

uit tabel V, waarin te zien is dat de variantie van de parameters in de beste 10 sets nog aanzienlijk is.

Niet uniek zijn van een optimalisatieprobleem kan ook voorkomen wanneer er verschillende lokale optima bestaan. Bij de Monte Carlo-methode bestaat er in feite weinig kans dat men in een lokaal minimum terechtkomt en daarmee het absolute minimum mist.

De eenvoudigste tactiek bij voorspellingsberekeningen zou kunnen zijn deze uit te voeren met de plausibel geachte parametersets. Al doende wordt direct een beeld van de onzekerheid van de voorspelling verkregen. Het doen van harde uitspraken hierover is echter alleen mogelijk wanneer de acceptatie van de plausibele parametersets is geschied op basis van harde criteria. In dat geval wordt gesproken van acceptabele sets.

Kalibratie met Monte Carlo-technieken heeft belangrijke beperkingen. Zo is het resultaat onnauwkeurig, zeker binnen een als redelijk te hanteren aantal modelruns in de orde van 1000 (zie ook [Jakeman *et al.*, 1990]). Een belangrijker beperking is de onmogelijkheid om met Monte Carlo-methoden een groot aantal parameters gelijktijdig te optimaliseren:

Stel dat de Monte Carlo-zoekruimte wordt afgebakend door een hyperkubus met zijde D. Stel voorts dat het doel de hyper-bol is die juist in de hyperkubus past. Zoals blijkt uit onderstaande uit Scott [1992] overgenomen tabel, neemt deze kans sterk af met het aantal dimensies. Vandaar dat Monte Carlo-optimalisatie bij meer dan zeg 5 parameters een hopeloze zaak wordt.

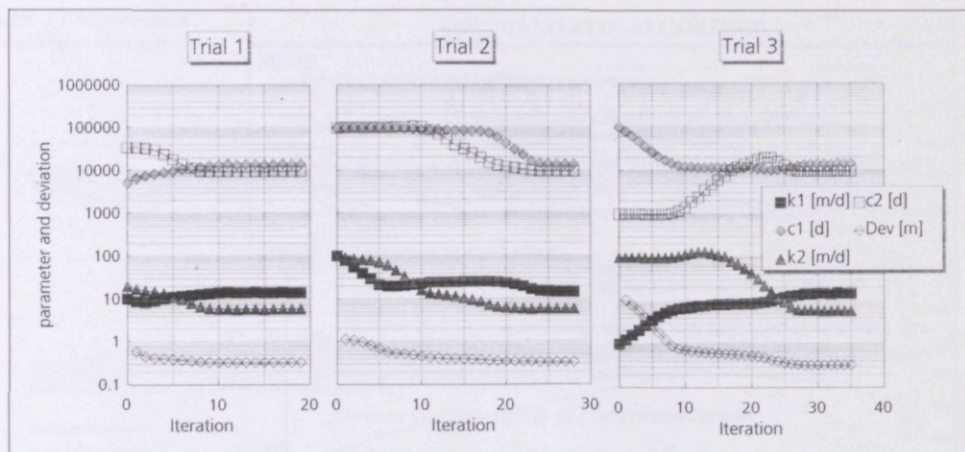
Resultaten mathematische optimalisatie

Het optimalisatie-algoritme blijkt uiterst effectief. Ondanks extreem uiteenlopende startwaarden voor de vier parameters, worden in de drie uitgevoerde kalibratieruns steeds dezelfde parameterwaarden gevonden (afb. 4).

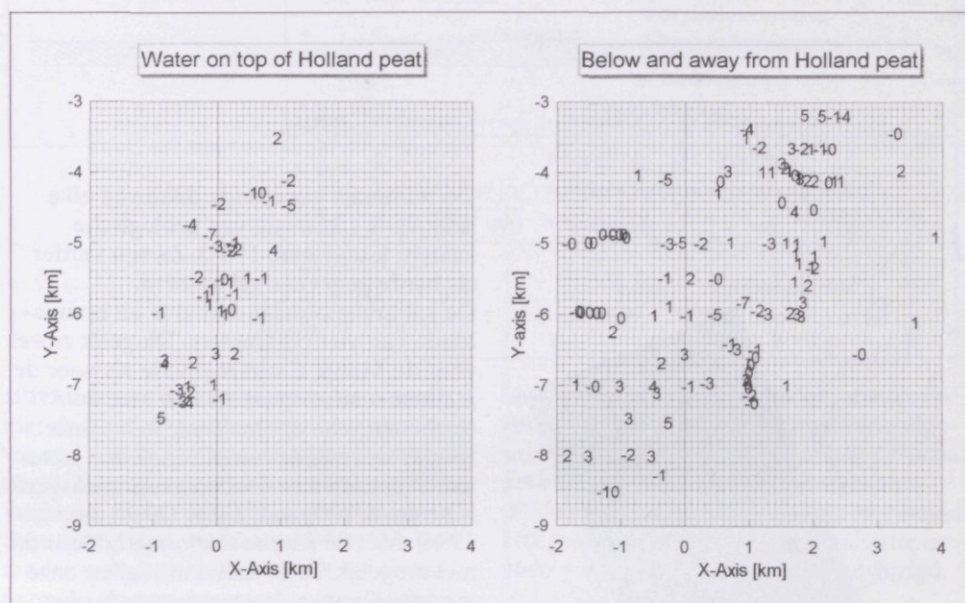
Afhankelijk van de startwaarden bedroeg het aantal parameteraanpassingen 19 tot

TABEL VI – Verhouding volume hyperbol en hyperkubus met dezelfde diameter.

Dimensie	1	2	3	4	5	6	7
Verhouding	1	0.785	0.524	0.308	0.164	0.081	0.037



Afb. 4 – Verloop parameterwaarden en deviatie tijdens de 3 ijkruns volgens de mathematische optimalisatie.



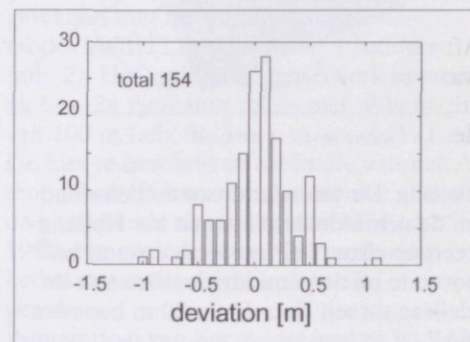
Afb. 5 – Deel van het modelnetwerk en de lokatie van de 154 waarnemingsputten met resterende stijghoogteafwijking in dm. (de linker benedenhoek van het modelnetwerk in afb. 3 heeft coördinaten $\{-2,4, -9,4\}$ [km]).

35 (tabel VII). De eindwaarden van de drie runs zijn gelijk. De laatste rij van tabel VII bevat de enigszins afwijkende eindwaarden verkregen met de Monte Carlo-methode.

Het aantal hiervoor benodigde modelruns bedraagt het vijfvoud (4 parameters + 1), dus 85 à 175. Dit is veel minder dan het aantal voor de Monte Carlo-methode

TABEL VII – Start- en eindwaarden van de parameters bij de 3 uitgevoerde mathematische optimalisaties en de Monte Carlo-methode.

	Trial	Iter.	k1 [m/d]	c1 [d]	k2 [m/d]	c2 [d]	Dev [d]
Start	1	0	10	5.000	20	34.000	0.587
Start	2	0	100	5.000	20	99.000	1.143
Start	3	0	1	5.000	20	1.000	10.314
Eind	1	19	14.2	5.000	20	9.455	0.3300
Eind	2	28	14.3	5.000	20	9.486	0.3300
Eind	3	35	14.3	5.000	20	9.467	0.3300
Monte Carlo:			14.3	14.948	6.2	9.086	0.34



Afb. 6 – Histogram van het verschil tussen de berekende en gemeten stijghoogten in 154 waarnemingsputten na ijk van het model.

gebruikte en benodigde runs. Behalve efficiënter is de methode stabiel (afb. 4) en nauwkeuriger dan de Monte Carlo-methode (dev=0,33 m in plaats van 0,34 m). In tegenstelling tot de Monte Carlo-methode is zij zonder wezenlijk verlies van efficiency toepasbaar bij een

Gebruikte symbolen

$\{\}^T$	T als superscript staat voor matrix transpositie
$\{\}, \{\}$	Enkele onderstreping is vector, dubbele matrix
i, j	Index van parameter resp. waarneming
n, m	Aantal parameters resp. waarnemingen
$\hat{h}_j, h_j$	Berekende resp. gemeten stijghoogte
f_j	Gewogen verschil tussen berekende en gemeten stijghoogte $f_j = \sqrt{\omega_j} (\hat{h}_j - h_j)$
$p, \Delta p$	$(n \times 1)$ --log-parametervector en zijn aanpassing van iteratie tot iteratie
F	Fout-criterium $f^T f = \sum_{j=1}^m f_j^2$
$\underline{J}, \underline{J}^T$	$(m \times n)$ --Jacobi-matrix en zijn $(n \times m)$ --getransporteerde
$\underline{J}$	$\begin{pmatrix} \frac{\partial f_1}{\partial p_1} & \dots & \frac{\partial f_1}{\partial p_i} & \dots & \frac{\partial f_1}{\partial p_n} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial f_i}{\partial p_1} & \dots & \frac{\partial f_i}{\partial p_i} & \dots & \frac{\partial f_i}{\partial p_n} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial p_1} & \dots & \frac{\partial f_n}{\partial p_i} & \dots & \frac{\partial f_n}{\partial p_n} \end{pmatrix}$
$\underline{g}$	$(n \times 1)$ --gradiënt van F , met elementen $\frac{\partial F}{\partial p_i}$
$\underline{H}$	$(n \times n)$ --Hessiaanse-matrix, met elementen $\frac{\partial^2 F}{\partial p_k \partial p_i}$
$\underline{I}$	$(n \times n)$ --identiteitsmatrix
μ	Levenberg-Marquardt-coëfficiënt
ε	Log-parameter perturbatie-stap om de Jacobi-matrix te berekenen
Δ	Maximum stapgrootte $ \Delta p $ per iteratie
N	Aantal parametercombinaties bij Monte Carlo-methode

groter aantal te optimaliseren parameters [Olsthoorn, 1994].

Afbeelding 5 geeft de lokatie van de waarnemingspunten en de resterende stijghoogte-afwijking. Afbeelding 6 vat de afwijkingen samen in een histogram. Verbetering is alleen mogelijk met nieuwe informatie over de opbouw van de ondergrond en met een groter aantal zinvol gekozen ijkparameters.

Nadat de parameters zijn geoptimaliseerd, biedt de Hessiaanse matrix de informatie die nodig is voor een onzekerheidsanalyse.

Conclusies

De Monte Carlo-methode blijkt relatief onnauwkeurig en wordt onbruikbaar wanneer meer dan zeg 5 onafhankelijke parameters tegelijk moeten worden geoptimaliseerd. (Monte Carlo-methoden zijn daarentegen uitstekend te gebruiken in onzekerheidsanalyses). Mathematische optimalisatie heeft de voorkeur bij de kalibratie van modellen.

Het is de enige keus bij gelijktijdige optimalisatie van een groter aantal parameters. Het gepresenteerde mathematische optimalisatie-algoritme werkt eenvoudiger en beter dan de geavanceerde methode van Carrera en Neuman [1986 a,b,c], die wereldwijd als meest vooruitstrevend wordt beschouwd [Olsthoorn, 1994]. Het algoritme is bovendien gemakkelijk te implementeren, zonder aanpassing van het (grondwater)model.

De gegeven technieken kunnen gemakkelijk in de praktijk worden toegepast. Zij verdienen verre de voorkeur boven niet-reproduceerbare handmatige probeertechnieken.

Verantwoording

Dit artikel is mogelijk gemaakt door ing. P. J. W. T. Kamps. Hij heeft het MODFLOW-model van de Amsterdamse Waterleidingduinen gemaakt, beide ijkmethoden geïmplementeerd, de benodigde hulpprogramma's geschreven en de berekeningen uitgevoerd.

Literatuur

- Carrera, J. and Neuman, S. P. (1986a, b, c). *Estimation of Aquifer Parameters Under Transient and Steady State Conditions*:
 1: *Maximum Likelihood Method Incorporating Prior Information*. Wat. Res. Res. Vol. 22, No. 2, 199-210.
 2: *Uniqueness, Stability and Solution Algorithms*. Wat. Res. Res. Vol. 22, No. 2, 211-227.
 3: *Application of Synthetic and Field Data*. Wat. Res. Res. Vol. 22, No. 2, 228-242.
 Carrera, J. (1988). *State of the Art of the Inverse Problem applied to the Flow and Solute Transport Equations*. In: Custodio E, e.a., (Eds). *Groundwater Flow and Quality Modelling*. D. Reidel Publ. Company. 549-583.
 Heij, G. J. and Kester, C. G. M. (1977). *De hydraulische weerstand van de Formatie van Sterksel/Kedichem in West-Utrecht*. RID-mededeling 1977-7, Rijksinstituut voor Drinkwatervoorziening, Den Haag, ca. 20 pp.
 Jakeman, A. J., Ghassemi, F. and Dietrich, C. R. (1990). *Calibration and reliability of an aquifer system model using generalized sensitivity analysis*. In: Kovar (Ed). *Calibration and Reliability in Groundwater Modelling*. ModelCare 90, The Hague 3-6 Sep. 1990, p 43-51, IAHS-publication No. 195.
 Jennings, A. and McKeown, J. J. (1992). *Matrix Computation*. John Wiley & Sons, Chichester etc. ISBN 0 471 93527 1.
 Olsthoorn, T. N. (1994). *Effective parameter optimization for groundwater model calibration*. To be published in Groundwater 1994.
 Olsthoorn, T. N., Kamps, P. T. W. J. and Driesen, W. J. (1993). *Groundwater modelling using GIS at the Amsterdam Water Supply*. In: Kovar e.a. (Eds). *Application of Geographic Information Systems in Hydrology and Water Resources management*. Vienna, April 19-22, 1993. IAHS publication No. 211. Wallingford. p. 665-674.
 Olsthoorn, T. N. (1989). *Grondwaterstandsverlagen ten gevolge van de Duitse bruinkoolwinning in de Roerdalslenk. Kwantitatieve analyse van de verschillen tussen de Duitse en de Nederlandse modelstudie*. RIVM, Bilthoven, Rapport Nr. 728610001, 28 pp.
 Olsthoorn, T. N. en Kusse, A. A. M. (1986). *Aanpak en gegevensverwerking bij grondwatermodellering, toegepast bij onderzoek naar de invloed van de Duitse bruinkoolwinning*. H₂O (19) 1986, Vol. 19, nr. 16, p. 367-373.
 Olsthoorn, T. N. (1982). *Praktisch ijken van grondwatermodellen*. Bouwvak, Misset, Doetinchem, Vol.1, No. 2.
 Peck, A. e.a. (1988). *Consequences of Spatial Variability in Aquifer Properties and Data Limitations for Groundwater Modelling Practice*. IAHS-publ.175, IAHS-press, Wallingford UK, ISBN 0-947571-61-2.
 Scott, D. W. (1992). *Multivariate Density Estimation*. Wiley & Sons, New York etc. ISBN 0-471-54770-0, 317 pp.
 Stuyfzand, P. J. (1987). *Hydrochemie en Hydrologie van duinen en aangrenzende polders tussen Noordwijk en Zandvoort aan Zee*. Kiwa, Nieuwegein, SWE-87.007, 337 pp.
 Sun N. Z. and Yeh, W. W. (1985). *Identification of Parameter Structure in Groundwater Inverse Problem*. Wat. Res. Res., Vol. 21, No. 6, 869-883.
 Sykes, J. F. (1985). *Sensitivity Analysis for Steady State Groundwater Flow Using Adjoint Operators*. Wat. Res. Res., Vol. 21, No. 3, 359-371.
 Yeh, W. W. and Yoon, Y. S. (1981). *Aquifer Parameter Identification with Optimum Dimension in Parameterisation*. Wat. Res. Res. Vol. 17, No. 3, 664-672.
 Yeh, W. W. (1986). *Review of parameter Identification Procedures in Groundwater Hydrology. The Inverse Problem*. Water Resources Research, 22, No. 2, 95-108.