

Kapitel 14 INVERS MODELLERING

Torben Obel Sonnenborg
Geologisk Institut, Københavns Universitet

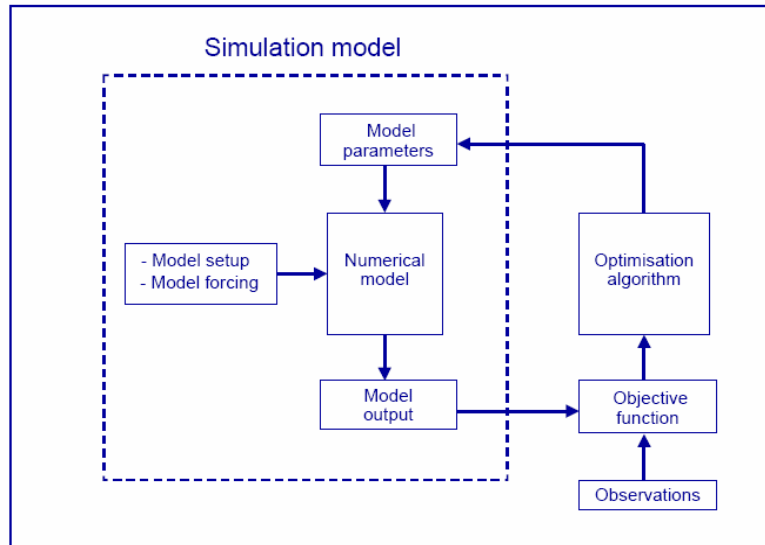
Anker Lajer Højberg
Hydrologisk Afdeling, GEUS

Nøglebegreber: Objektivfunktion, vægtning af residualer, optimeringsalgoritmer, parameterusikkerhed og –korrelation, vurdering af kalibreringsresultat.

ABSTRACT: Invers modellering, ofte også benævnt automatisk parameterestimation eller invers modellering, involverer flere kritiske valg og kan ved uhensigtsmæssig brug resultere i dårlige resultater. Opstilling af objektiv funktionen, der er den størrelse, som den inverse model vil forsøge at minimere ved at justere parameterværdierne, er et af de kritiske skridt i automatisk parameterestimation. Det skal specificeres, hvordan observationer af samme type og observationer af forskellige type skal vægtes indbyrdes og i forhold til hinanden. Optimeringsalgoritmen, der skal minimere den definerede objektiv funktion, er et andet vigtigt valg i parameterestimationen. I dette kapitel præsenteres en global metode og en lokal metode, hvor vægten lægges på sidstnævnte. Problemer med lokale minima og manglende entydighed beskrives, og der angives retningslinier for, hvordan problemerne kan minimeres. Desuden beskrives problemer med estimation af urealistiske parameterværdier. Et eksempel illustrerer, hvordan valg af observationsdata, vægtning af disse data, og parametrisering påvirker resultaterne fra den inverse optimering. Endelig gives nogle retningslinier for, hvilke spørgsmål der bør stilles, når resultaterne fra den inverse model skal evalueres.

14.1 INDLEDNING

Poeter & Hill (1997) opfordrede i artiklen ”Inverse models: A necessary next step in ground-water modeling” til øget brug af automatisk parameteroptimering i forbindelse med grundvandsmodellering. Som argumenter for at bruge invers modellering anfører Poeter & Hill, at metoden har evnen til at (1) automatisk at finde de parametre, der giver den bedste overensstemmelse mellem observeret og simuleret hydraulisk trykniveau og vandflukse, (2) kvantificere kvaliteten af kalibreringen, (3) identificere mangler og behov mht. observationsdata, (4) beregne konfidensintervaller på de estimerede parametre, (5) identificere parameterkorrelation og sensitivitet af parametre, og (6) tilvejebringe et objektivt grundlag for sammenligning af forskellige konceptuelle modeller. Siden Poeter & Hills artikel udkom, er automatiske optimeringsalgoritmer blevet indbygget i de fleste modelsystemer på markedet (f.eks. Groundwater Vistas, GMS, MIKE SHE), hvilket gør det relativt let at foretage invers optimering. Sideløbende hermed stilles der i større og større grad krav om anvendelse af automatiske kalibreringsmetoder fra vandressourceforvalternes side. Dette har gjort, at inverse metoder i dag anvendes ekstensivt f.eks. i forbindelse med modellering af indsatsområder. Det er imidlertid ikke altid, at de inverse metoder anvendes lige hensigtsmæssigt, og oftest udnyttes kun en lille del af den information, en invers model potentielt kan give brugeren. Det er f.eks. sjældent, at informationen i de seks punkter anført ovenfor bliver udnyttet samtidigt i de opgaver der udføres i Danmark.



Figur 14 Ramme for automatisk kalibrering, hvor sammenhængen mellem simuleringssmodellen og optimeringsalgoritmen ses (fra Madsen & Jacobsen, 2002).

Automatisk (eller invers) parameteroptimering kan illustreres som vist på figur 14.1. Den inverse model består af følgende elementer:

1. en numerisk kode (f.eks. MODFLOW)
2. den hydrogeologiske tolkningsmodel ("model setup" eller konceptuel model – indgående strømningselementer, randbetingelser, etc.) samt påvirkninger ("model forcing" – nedbørsinput, oppumpning, vandudledning, etc.)
3. parametrisering og valg af kalibreringsparametre
4. observationsdata
5. kalibreringskriterium (objektiv funktion)
6. optimeringsalgoritme

I dette kapitel vil der blive fokuseret på punkterne 5 og 6, idet de øvrige punkter er behandlet andre steder i håndbogen. Retningslinier for modelkalibrering vha. invers modellering er desuden beskrevet udførligt i Hill (1998), hvor der bl.a. er anført en checkliste på 16 punkter, som beskriver god kalibreringspraksis. I forhold til sidstnævnte reference vil der her blive fokuseret på nogle af problemerne ved anvendelse af automatisk parameterestimation.

14.2 KALIBRERINGSKRITERIUM (OBJEKTIV FUNKTION)

I automatisk parameterestimation skal der defineres et enkelt udtryk, som kan minimeres vha. en matematisk/numerisk optimeringsalgoritme. Dette udtryk kaldes også for objektiv-funktionen, G , som er et mål for afvigelsen mellem en observeret og simuleret variabel fra det modellerede system. Normalt defineres G som en vægtet sum af kvadratet på residualerne $r_i = \psi_{\text{obs},i} - \psi_{\text{sim},i}$, hvor $\psi_{\text{sim},i}$ er den simulerede værdi af $\psi_{\text{obs},i}$ og ψ er en variabel fra grundvandssystemet

$$G(\underline{b}) = \sum_{i=1}^n w_i r_i^2 \tag{14.1}$$

\underline{b} er vektoren af kalibreringsparametre, som i forbindelse med grundvandsmodellering primært vil være sammensat af hydrauliske ledningsevner for forskellige lag og zoner, vandløbskonduktivitet (lækagekoefficient), drænkonduktivitet (drænkoefficient), samt eventuelt randbetingelser såsom

værdien af trykniveauet på en rand med fastholdt tryk eller grundvandsdannelsen til modellen. Endelig er w_i vægtningen af det enkelte residual. Residualerne kan være baseret på forskellige datatyper, f.eks. hydraulisk trykniveau eller vandløbstilstrømning, og det er derfor mest hensigtsmæssigt at differentiere mellem residualer af samme type og residualer af forskellige type, da vægtningen vil være forskellig alt efter om der er tale om vægtning indenfor datatype eller mellem datatype. I de to efterfølgende afsnit behandles hhv. vægtning indenfor og mellem datatyper.

14.2.1 *Vægtning indenfor datatyper*

I dette afsnit antages, at der kun indgår én type observation i G , f.eks. data for hydraulisk trykniveau, og det er derfor kun vægtningen af de enkelte residualer, w_i , der skal kvantificeres.

Vægtningen kan foretages på mange forskellige måder. Den ene ekstrem består i subjektivt at tildele observationer, hvor man ønsker en nøjagtig simulering, høje vægte, mens observationer, hvor der ikke er behov for en nøjagtig reproduktion af de observerede værdier, tildeles lave vægte. Som modsætning til denne fremgangsmåde kan det forsøges at estimere usikkerheden på de enkelte residualer (summen af observationsfejl og modelleringsfejl), hvilket anbefales stærkt f.eks. i Hill (1998) og Hill & Tiedeman (2002). Teoretisk set er det fornuftigt at tildele en lille vægt til et residual, hvis der er stor usikkerhed om f.eks. den observerede værdi, da det ikke er rimeligt at tilstræbe en nøjagtig simulering af en observationsværdi, der er stor usikkerhed på. Når der udføres manual kalibrering, kan den slags overvejelser foretages implicit, idet hydrogeologen, der udfører kalibreringen, bevidst eller ubevidst vil bruge mindre energi på at tilpasse modellen i områder, hvor f.eks. målingerne af hydraulisk trykniveau er af ældre dato, eller hvor der ikke er information tilgængelig om filtersætning. Med den automatiske optimeringsalgoritme er det ligeledes muligt at foretage en lignende afvejning af usikkerhederne, ved at kvantificere disse og give dem som input til modellen i form af vægte på de enkelte residualer. En ofte anvendt metode er at tildele residualerne en vægt, der er omvendt proportional med variansen på fejlen på residualerne, σ^2

$$w_i = \frac{1}{\sigma_i^2} \quad (14.2)$$

Problemet ved at tildele vægte på baggrund af usikkerheder er, at det kan være vanskeligt at kvantificere hhv. observationsfejl og modelfejl. Observationsfejlen består af faktorer som måleusikkerhed, barometereffekt, og kote fejl (se kapitel 15, Kalibrering af strømningsmodel), og det vil i de fleste tilfælde være muligt at komme med et kvalificeret bud på disse typer af fejl. Modelfejlen består af elementer som diskretiseringsfejl, manglende modellering af geologisk heterogenitet (udtrykt ved den rumlige variation i hydraulisk ledningsevne), samt manglende viden om den arealdistribuerede grundvandsdannelse. Effekten af geologisk heterogenitet på usikkerheden i modelleret trykniveau er en af de vigtigste modelusikkerheder og er beskrevet i kapitel 11 (Skalaforhold og heterogenitet). Denne usikkerhed kan kvantificeres f.eks. vha. udtryk (11.16), hvis der er en tilstrækkelig datamængde til rådighed om hydraulisk ledningsevne til at de geostatistiske egenskaber for K kan bestemmes.

Usikkerheden på grundvandsdannelsen er den størrelse, som det er sværest at håndtere. Dels er det vanskeligt at kvantificere usikkerheden på grundvandsdannelsen, dels er det vanskeligt at kvantificere, hvordan usikkerheden på grundvandsdannelsen forplanter sig til en usikkerhed på hydraulisk trykniveau. Gelhar (1993) behandler emnet, men kommer ikke frem til generelle udtryk for usikkerheden på hydraulisk trykniveau, som det er tilfældet for effekten af heterogenitet i hydraulisk ledningsevne. Gelhars analyse giver imidlertid en indikation af, at grundvandsdannelsen er en vigtig faktor for usikkerheden på hydraulisk trykniveau. Det må forventes, at grundvandsmagasiner tæt på jordoverfladen (typisk frie magasiner) vil være mere påvirket af denne effekt end dybereliggende magasiner. Det må ligeledes forventes, at grundvandsdannelsen til magasiner, der er overlejret af komplekse geologiske formationer som f.eks. moræneforekomster, og hvor terrænet er stærk kuperet, vil være behæftet med betydelig større usikkerhed end grundvandsdannelsen til magasiner med relativ homogen geologi og fladt terræn.

De fleste datasæt for hydraulisk trykniveau vil have en uensformig rumlig fordeling. Typisk finder der mange målinger i områder med store vandindvindingsinteresser, og der kan derfor være en

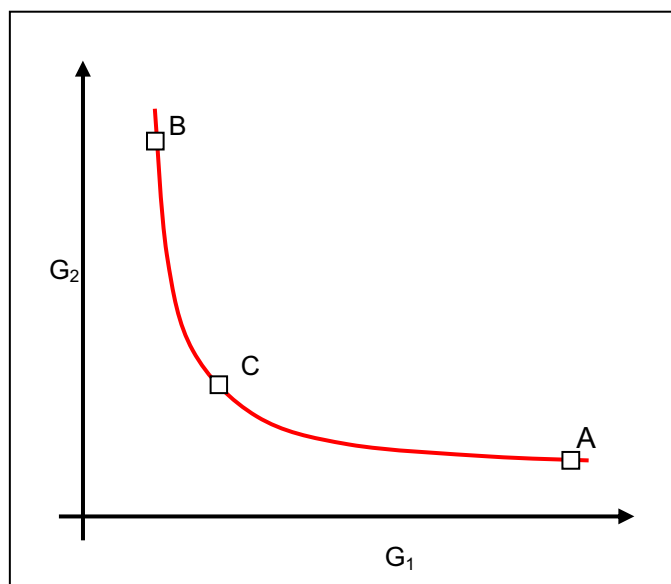
tendens til, at disse målinger kommer til at styre optimeringen i højere grad end informationsværdien i målingerne berettiger til. Det skyldes, at når målingerne er klumpet sammen ("clustering"), må det forventes, at de er positivt korrelerede, hvilket betyder, at hvis den ene måling stiger vil en nærliggende måling ligeledes stige. Derfor kan det være berettiget og hensigtsmæssigt for optimeringen at tildele målinger i områder med mange observationer en mindre vægt end en uafhængig analyse af fejlene på målingen og modelresultaterne på målelokaliteten vil resultere i. Da der ikke på forhånd er viden om korrelationen mellem målingerne, må vægtningen foretages ad hoc, og må eventuelt foretages iterativt idet der gennemføres flere inverse optimeringer med forskellige vægtninger.

Som den ovenstående gennemgang viser, vil en kvantificering af vægtene på baggrund af usikkerheden på residualerne ikke kunne gennemføres uden en vis brug af mere eller mindre subjektive overvejelser. Alligevel kan det anbefales, at usikkerheden estimeres, da det dels tvinger hydrogeologen til at tage stilling til kvaliteten af de enkelte observationer, og da det dels giver en plausibel vægtning af de enkelte residualer. Eventuelt kan vægtning på basis af usikkerheder kombineres med en subjektiv vægtning, hvor residualer beliggende i områder med stor interesse vægtes højere end områder med lille fokus. Dette svarer til, at man accepterer en større modelfejl i områderne med lille fokus, da hverken grundvandsdannelsen eller variationen i hydraulisk ledningsevne her beskrives i detalje. Eksempelvis vil man forvente en relativ upræcis reproduktion af det hydrauliske trykniveau i områder, hvor heterogene grundvandsmagasiner beskrives vha. store homogene enheder. Det kan derfor forventes, at den resulterende RMS-værdi bliver relativt høj, da effekten af småskalheterogeniteter ikke fanges af modellen. Der skal imidlertid stadig stilles krav til en fornuftig middelfejl (f.eks. ME-værdi), da det kan være vigtigt for det tilgrænsende fokusområde, at området fungerer som en fornuftig randbetingelse.

14.2.2 Vægtning mellem forskellige datatyper

Ofte er det både muligt og hensigtsmæssigt at kalibrere mod mere end én type observationsdata. F.eks. kan der ud over trykniveaudata være vandløbsafstrømning, Q , stofkoncentration, C , eller dateringer af grundvandets alder, T , til rådighed. Hvis de enkelte residualer vægtes med usikkerheden på observationen, vil de vægtede residualer være dimensionsløse, og det vil derfor være uproblematisk at beregne summen af de kvadrerede residualer og minimere denne størrelse. Normalt vil der imidlertid være flere målinger til stede af den ene type observation end den anden, og der vil derfor være en måske utilsigtet tendens til, at den type data, som der er flest observationer af, vil dominere optimeringen.

Problemet kan illustreres vha. figur 14.2, hvor G_1 og G_2 kan være objektiv-funktioner for hhv. trykniveau og vandløbsafstrømning



Figur 14.2 Illustration af Pareto-fronten i et tilfælde med to objektivfunktioner G_1 og G_2 . Punktet A angiver en optimering, hvor kun type 2 observationer er anvendt, B angiver et tilfælde hvor kun type 1 observationer er anvendt, men C angiver et balanceret optimum baseret på begge typer informationer.

$$G_1 = \sum_{i=1}^n w_i (h_{obs,i} - h_{sim,i})^2$$

$$G_2 = \sum_{i=1}^m w_i (Q_{obs,i} - Q_{sim,i})^2$$
(14.3)

Kurven på figur 14.2 angiver minimum af funktionen

$$G = v_1 G_1 + v_2 G_2$$
(14.4)

hvor $0 \leq v_i \leq 1$ og $\sum v_i = 1$. Punktet A er G 's minimum, hvis der kun inddrages observationer af vandløbsafstrømningen i optimeringen ($v_1 = 0$; $v_2 = 1$), mens punktet B svarer til en optimering, hvor der kun bruges observationer af trykniveau i optimeringen ($v_1 = 1$; $v_2 = 0$). Alle andre punkter på kurven, der forbinder A og B, repræsenterer optimum af G for forskellige valg af v_1 og v_2 , og betegnes Pareto-fronten. Hvor på Pareto-fronten man ønsker at befinde sig er et subjektivt valg og afhænger af den givne problemstilling. Hvis begge observationstyper tillægges lige meget vægt, vil det være hensigtsmæssigt at kunne ramme punktet C, som kan betegnes som et balanceret optimum. Det er imidlertid ikke på forhånd muligt at bestemme værdierne af v_1 og v_2 , der sikrer, at punkt C rammes. Teoretisk set skal hele Pareto-fronten estimeres i hver konkret problemstilling, før det balancerede optimum kan bestemmes. I praksis kan det imidlertid resultere i et uoverkommeligt antal modelsimuleringer, og denne fremgangsmåde vil derfor ikke være mulig generelt.

En anden og meget pragmatisk måde at finde en passende vægtning af forskellige datatyper på, består i først at bestemme punkt A og B på figur 14.2. Det kræver, at der gennemføres to inverse kørsler, hvor der hhv. kun optimeres mod trykniveaudata og vandløbsafstrømning. Her ud fra kan følgende vægte bestemmes

$$v_i = 1 - \frac{\sum_{j=1}^n G_{i,j}}{\sum_{i=1}^n \sum_{j=1}^n G_{i,j}}$$
(14.5)

hvor $G_{i,j}$ er den optimale værdi af objektivfunktionen for datatype i optimeret udelukkende på basis af datatype j.

Hopmans et al. (2002) har foreslået, at der vægtes med antallet af observationer, som er til rådighed indenfor de forskellige datatyper

$$v_i = 1 - \frac{N_i}{N}$$
(14.6)

hvor N_i er antallet af observationer af datatype i og N er det totale antal observationer, $N = \sum N_i$. Denne vægtning forudsætter dels, at de indgående residualer har samme dimension (evt. er dimensionsløse), dels at alle observationer indenfor de enkelte datatyper indeholder lige meget information. Sidstnævnte antagelse vil oftest ikke være opfyldt, da observationer foretaget på visse lokaliteter vil have større værdi end andre steder. Eksempelvis vil en trykniveauobservation fra en boring placeret tæt på en fastholdt trykniveau-rand være af mindre værdi for optimeringen end en observation placeret i midten af modelområdet. Hvis der er relativt mange observationer i et område, vil værdien af hver enkelt observation sandsynligvis være mindre end en observation fra et datafattigt område. Ligning (14.6) vil dog kunne give et kvalificeret estimat af, hvordan de enkelte datatyper skal vægtes.

Hvis G_1 og G_2 (givet ved udtryk (14.3), hvor $w_i = 1/\sigma_i^2$) er bestemt til hhv. 45 og 8 i en situation, hvor der kun er optimeret mod 110 trykniveauobservationer, og G_1 og G_2 er bestemt til hhv. 67 og 4 i en optimering, hvor der kun er anvendt 3 afstrømningsobservationer, fås vha. (14.5) $v_1 = 0.1$ og

$v_2 = 0.9$. Havde (14.6) været anvendt ville de tilsvarende vægte være blevet estimeret til $v_1 = 0.03$ og $v_2 = 0.97$.

I praksis vil det ofte være nødvendigt at anvende trial-and-error metoden for at bestemme en hensigtsmæssig vægtning mellem forskellige datatyper, hvor forskellige vægte specificeres, og resultaterne vurderes successivt. I dette tilfælde er det ikke nødvendigt, at de enkelte residualer er dimensionsløse, da forskellen mellem den numeriske værdi af residualer af forskellig type kan vurderes manuelt. Anvendelse af udtryk (14.5) kræver ligeledes ikke, at der arbejdes med dimensionsløse residualer (eller residualer med samme dimension).

Nogle modelsystemer giver kun mulighed for at angive én vægt på hver observation, og i dette tilfælde er det nødvendigt at multiplicere vægtene v og w med hinanden og indsætte resultatet som vægtningen af observationspunkterne.

14.2.3 Brug af a priori information

Observationer af parametre, der skal estimeres, kan også inddrages i optimeringen. Normalt defineres en ny objektiv funktion af følgende form

$$G_b = \sum w_i (b_{obs,i} - b_i)^2 \quad (14.7)$$

hvor $b_{obs,i}$ er den observerede værdi af parameteren b_i . G_b omtales ofte som en "penalty function", da optimeringen bliver "straffet", hvis parameteren b_i bevæger sig for langt væk fra den observerede værdi. En penalty funktion kan forhindre estimation af urealistiske parametre og afhjælpe problemer med manglende entydighed (Carrera & Neuman, 1986). Til gengæld skal man være påpasselig, når der skal specificeres observationsværdier for en parameter. Som beskrevet i kapitel 4, er målinger af hydrauliske ledningsevne skalaafhængige og det vil derfor være fejlagtigt at anvende f.eks. en K-værdi baseret på slugtest som apriori parameter i en model, der arbejder på oplandsskala. Hvis de målte data svarer til modelleringsskalaen er der stadig risiko for, at de tilgængelige punktmålinger ikke er repræsentative for de parameterzoner, der er defineret under parametriseringen af modellen. Samlet må det anbefales at anvende målinger af parametre i optimeringen af modellen med stor forsigtighed. Der er i litteraturen stor uenighed om, hvorvidt a priori information skal medtages i objektiv funktionen. Uenigheden strækker sig fra helt at undgå det (Ginn & Cushman, 1990) til anbefalinger om, at det kan styrke optimeringen betragteligt (Carrera & Neuman, 1986). Alternativt kan observationer af parametre anvendes til at identificere realistiske intervaller for den pågældende parameter, indenfor hvilket det må forventes at den estimerede parameter skal ligge.

14.3 OPTIMERINGSALGORITMER

Objektivfunktionen G givet ved (14.1) kan i matrix-notation skrives som

$$G(\underline{b}) = \underline{r}^T \underline{w} \underline{r} \quad (14.8)$$

hvor $\underline{\bullet}$ angiver en vektor og $\underline{\underline{\bullet}}$ angiver en matrice. For et simpelt eksempel med blot 5 trykni-
veauobservationer, hvor de 3 første observationer har en standardafvigelse på 2 m og de to sidste har en standardafvigelse på 3 m fås (vægtene antages omvendt proportionale med variansen)

$$\underline{r} = \begin{bmatrix} h_{obs,1} - h_{sim,1} \\ h_{obs,2} - h_{sim,2} \\ h_{obs,3} - h_{sim,3} \\ h_{obs,4} - h_{sim,4} \\ h_{obs,5} - h_{sim,5} \end{bmatrix} \quad \underline{\underline{w}} = \begin{bmatrix} (1/2)^2 & 0 & 0 & 0 & 0 \\ 0 & (1/2)^2 & 0 & 0 & 0 \\ 0 & 0 & (1/2)^2 & 0 & 0 \\ 0 & 0 & 0 & (1/3)^2 & 0 \\ 0 & 0 & 0 & 0 & (1/3)^2 \end{bmatrix} \quad (14.9)$$

Idet vægtningsmatricen kun består af diagonalelementer antages det implicit, at trykniveauobservationerne er indbyrdes uafhængige.

Matematisk formuleret skal optimeringsalgoritmen finde den kombination af parameterværdier, der minimerer G

$$\hat{\underline{b}} = \text{Min}\{G(\underline{b})\} \quad (14.10)$$

hvor $\hat{\underline{b}}$ er vektoren af parameterestimer, der resulterer i, at G bliver mindst muligt. Der findes forskellige typer af algoritmer, der er udviklet til dette formål. Overordnet kan de inddrages i globale og lokale metoder (Sooroshian & Gupta, 1995). De lokale metoder er designet til at finde minimum for objektivfunktioner, hvis værdi kontinuert aftager mod et enkelt minimum. De globale metoder er designet til at finde minimum for objektivfunktioner, der har mere end et hul, og hvor der derfor er en risiko for at havne i et lokalt minimum. I det følgende vil kun de mest almindelige algoritmer blive kort beskrevet. Mht. til den globale metode drejer det sig om den såkaldte Suffle complex evolution algoritme (modul i MIKE SHE), mens der for den lokale metode er tale om de gradient-baserede metoder, som kendes fra f.eks. MODFLOW2000, PEST og UCODE.

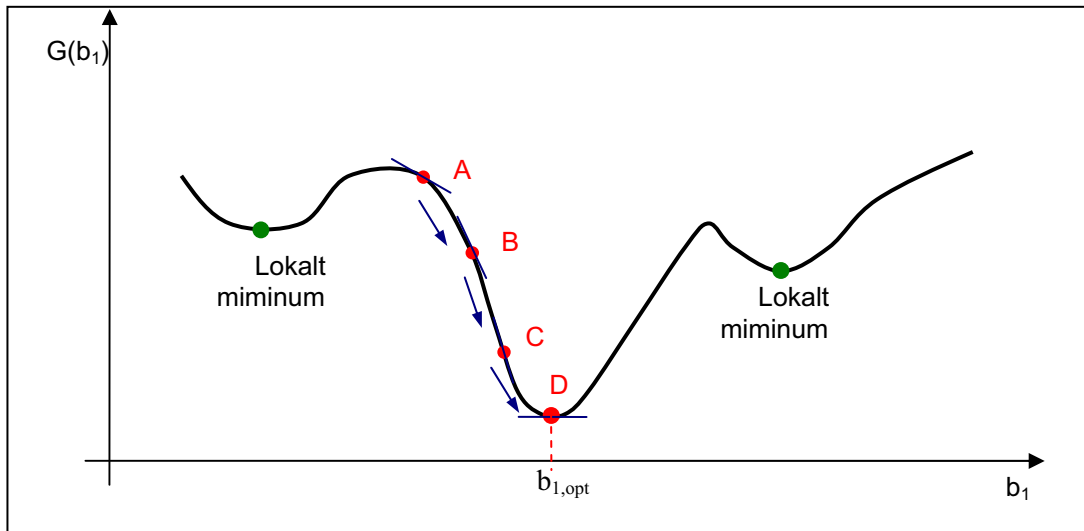
14.3.1 Suffle complex evolution method (SCE-metoden)

Indenfor grundvandsmodellering benyttes der primært lokale, gradientbaserede optimeringsalgoritmer til automatisk modeloptimering. Optimeringsalgoritmen i MIKE SHE er dog en undtagelse fra denne tendens, idet der her anvendes en global optimeringsalgoritme ved navn suffle complex evolution method (Duan et al., 1992; Madsen, 2000). Denne metode benytter sig af en i princippet relativt simpel strategi. Initielt samples der et antal (N) tilfældige parameterkombinationer i et på forhånd defineret parameterum (f.eks. givet ved minimum og maksimum for hver kalibreringsparameter), og modellen eksekveres for alle parameterudfaldene. I næste iteration udvælges igen N parameterudfald, men denne gang samples der mere målrettet i omegnen af det parameterudfald, der i første iteration gav den bedste overensstemmelse mellem observerede og simulerede værdier. Således fortsætter optimeringen indtil optimeringen konvergerer. SCE metoden er lidt mere avanceret end beskrevet ovenfor, idet der kan defineres et antal "complexes", dvs. grupper, således at optimeringen efter første iteration forløber i flere grupper, der kan køre i separate retninger, hvorved der er mindre risiko for at havne i et lokalt minimum.

SCE-metoden er designet til at optimere meget ikke-lineære problemstillinger, hvor det kan forventes, at objektivfunktionen er plaget af et eller flere lokale minima. Eksempelvis kan det forventes, at modeller der indeholder den umættede zone vil være stærkt ikke-lineære. Ligeledes vil fuldt integrerede hydrologiske modeller (med eller uden en fysisk beskrivelse af den umættede zone) have relativt stor chance for at være stærkt ikke-lineære. I forhold til de nedenfor beskrevne gradientmetoder, er der større chance for at optimeringen finder et globalt minimum, men til gengæld skal der også benyttes flere modelsimuleringer, når SCE-metoden benyttes.

14.3.2 Gradientbaserede metoder

I de gradientbaserede metoder minimeres G ved funktionsanalyse, idet nulpunktet for den afledede af G søges iterativt. Figur 14.3 illustrerer i det simple tilfælde med kun en parameter, hvordan



Figur 14.3 Illustration af princippet i de gradientbaserede metoders optimering. De korte liniestykker, der tangerer kurven, repræsenterer gradienten af objektivfunktionen G.

optimeringen principielt forløber. Optimeringen påbegyndes f.eks. i punktet A, og for at finde en lavere værdi af objektivfunktionen G, bestemmes hældningen af G i punktet. Herved findes retningen i hvilken modelparameteren skal justeres. Herefter flyttes parameteren til punkt B, en ny hældning findes, og optimeringen fortsættes på denne iterative måde, indtil hældningen af G er nul, og det derfor må antages, at objektivfunktionens minimum er fundet.

Til bestemmelse af, hvordan modelparametrene skal ændres i hver iteration beregnes den såkaldte Jacobi-matrice, J, som indeholder sensitivitetskoefficienterne

$$\frac{\partial \psi_{sim,i}}{\partial b_i} \tag{14.11}$$

De enkelte sensitivitetskoefficienter angiver, hvor meget simuleringsværdien $\psi_{sim,i}$ ændrer sig, når parameteren b_i ændres. En stor sensitivitetskoefficient er udtryk for, at den simulerede værdi ændrer sig meget når den pågældende parameter ændre størrelse. Parameteren har med andre ord stor indflydelse på simuleringen af f.eks. hydraulisk trykniveau i det pågældende punkt.

I eksemplet med 5 observationer af hydraulisk ledningsevne og 3 parametre fås følgende Jacobi-matrice

$$J = \begin{bmatrix} \frac{\partial h_{sim,1}}{\partial b_1} & \frac{\partial h_{sim,1}}{\partial b_2} & \frac{\partial h_{sim,1}}{\partial b_3} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \frac{\partial h_{sim,5}}{\partial b_1} & \frac{\partial h_{sim,5}}{\partial b_2} & \frac{\partial h_{sim,5}}{\partial b_3} \end{bmatrix} \tag{14.12}$$

Præcisionen af beregningen af sensitivitetskoefficienterne i Jacobi-matricen er af afgørende betydning for en gradientbaseret metode. UCODE og PEST anvender en forward eller central approksimation til beregning af sensitivitetskoefficienterne givet ved

$$\frac{\partial \psi_i}{\partial b_j} \cong \frac{\psi_i(b_j + \Delta b_j) - \psi_i(b_j)}{\Delta b_j} \quad \text{forward}$$

$$\frac{\partial \psi_i}{\partial b_j} \cong \frac{\psi_i(b_j + \Delta b_j) - \psi_i(b_j - \Delta b_j)}{2 \Delta b_j} \quad \text{central}$$
(14.13)

Central-approksimationen har større præcision end forward approksimationen, men til gengæld kræver det dobbelt så mange eksekveringer af grundvandsmodellen at beregne central-approksimationen. Den bedste præcision opnås ved brug af den såkaldte "sensitivity-equation" metode, som er indbygget i MODFLOW2000 (Hill et al., 2000).

Med en gradientbaseret metode opnås de bedste resultater, hvis G er en kontinuert funktion, der varierer gradvist med parametrene. Dette vil være tilfældet for artesisk grundvandsstrømning, men der findes en lang række tilfælde, hvor G er diskontinuert. F.eks. vil inaktivering af celler i et frit magasin, der pga. små ændringer i parameterværdierne kommer til at ligge over grundvandsspejlet, resultere i en markant ændring af trykniveauresponset, som følgende kan have en u hensigtsmæssig effekt på beregningen af sensitivitetskoefficienterne. Aktivering af drænflow, når trykniveauet overstiger en forudbestemt værdi, eller aktivering af overfladeafstrømning, når trykniveauet overstiger terrænkoten er andre mere fysiske eksempler på diskontinuerte reaktioner i hydrologiske modeller. I sådanne tilfælde vil det være svært for en gradientbaseret metode at bestemme dels retningen og størrelsen af ændringen i parametervektoren \underline{b} .

Som det ses af figur 14.3 afhænger optimeringen af den skitserede objektivfunktion med en gradientbaseret metode af værdien af det initiale parameterestimat. Hvis optimeringen startes i punkt A, vil det globale minimum blive fundet, men hvis optimeringen påbegyndes enten til venstre eller højre for dalen, hvor det globale minimum befinder sig, vil metoden resultere i et lokalt minimum. De gradientbaserede metoder er derfor bedst egnede til forholdsvis lineære problemstillinger.

UCODE og PEST, som er de mest anvendte generelle optimeringsmodeller indenfor grundvandsmodellering, er begge gradientbaserede metoder, og er mht. de basale antagelse og algoritmer meget ens. I PEST foretages der en mere raffineret bestemmelse af den skridtlængde de enkelte parametre tager i hver iteration, hvilket kan resultere i en hurtigere konvergens end det er tilfældet for UCODE. Desuden er det i PEST muligt at angive grænser for det interval, de enkelte parametre må bevæge sig indenfor under optimeringen. Herved kan urealistiske parameterestimer undgås, og denne facilitet gør, at PEST bliver foretrukket frem for UCODE af nogle brugere.

14.4 ENTYDIGHEDSPROBLEMER

Manglende entydighed ("non-uniqueness") optræder, når ét sæt observationsdata, f.eks. hydrauliske trykniveaudata, fra det betragtede system kan resultere i mere end en invers løsning mht. parametrene \underline{b} . Dette kan forekomme i principielt set to forskellige situationer. I den ene situation er det muligt at simulere det samme respons ved anvendelse af forskellige parametersæt. Dette vil eksempelvis være tilfældet, hvis to parametre er perfekt korrelerede, og der derfor findes et uendeligt antal kombinationer af de to parametre, der resulterer i nøjagtigt samme trykniveaufordeling og dermed også i samme værdi af objektivfunktionen G . Der er med andre ord mulighed for at generere det samme trykniveaubillede og derfor også samme strømningssituation med forskellige kombinationer af modelparametrene. Problemet er imidlertid, at modeller parametriseret med de forskellige parametersæt vil reagere forskelligt på f.eks. nye randbetingelser og påvirkninger af systemet (oppumpning, dræning, etc.), og modellen er derfor uegnet til at lave prediktioner med.

Den anden type af manglende entydighed optræder, hvis flere forskellige parameterkombinationer giver samme G -værdi, men ikke nødvendigvis samme trykniveaufordeling. Dette svarer til, at der er flere stort set lige dybe huller i objektivfunktionen, hvor overensstemmelsen mellem observeret og simuleret respons er lige god, men hvor residualerne er fordelt forskelligt. Denne situation

benævnes også "equi-finality", og kan opstå i selv velformulerede problemstillinger, hvor der ikke er problemer med manglende parameter-sensitivitet eller høj parameterkorrelation. De forskellige løsninger vil generere forskellige strømningssystemer, som vil resultere i forskellige resultater mht. variable, som der ikke er blevet kalibreret imod. Hvis der f.eks. kun kalibreres mod observationer af hydraulisk trykniveau, vil de alternative parametersæt resultere i simulering af forskellig tilstrømning til vandløb, forskellige opholdstider i de enkelte magasiner, etc. Derfor vil inddragelse af andre typer af observationsdata i kalibreringsprocessen ofte kunne afhjælpe problemet med manglende entydighed, da der pga. de ekstra data lægges flere bånd på optimeringen, og antallet af mulige løsninger indsnævres – i bedste fald til kun en enkelt.

Risikoen for manglende entydighed øges, når antallet af kalibreringsparametre stiger. Derved forøges sandsynligheden for, at optimeringen resulterer i et sæt parameterestimer, der ikke er i overensstemmelse med virkeligheden. Hvis en model parametriseret med et sådant parametersæt udsættes for en valideringstest, er der stor sandsynlighed for at der opnås et dårligt valideringsresultat. Et eksempel på dette problem ses i figur 15.1 i kapitlet om validering, hvor der genereres utilfredsstillende valideringsresultater, når antallet af parametre bliver for højt.

Manglende entydighed vil med sikkerhed optræde, hvis der estimeres flere parametre, end der er observationsdata til rådighed. Det er imidlertid ikke tilstrækkeligt blot at kræve, at antallet af observationsdata er større end antallet af kalibreringsparametre for at opnå entydighed. Værdien af den enkelte observation vil bl.a. afhænge af placering og nøjagtighed, og der kan derfor ikke opstilles generelle retningslinier for sammenhængen mellem antallet af observationsdata, der er til rådighed for optimeringen, og antallet af parametre der entydigt kan estimeres.

Der kan imidlertid gives nogle anbefalinger for, hvordan entydighed bedst muligt tilstræbes (bl.a. Carrera & Neuman, 1986; Hill, 1998):

- Minimering af fejlen (usikkerheden) på observationsdata
- Anvendelse af flere forskellige typer data og dermed objektivfunktioner
- Basere parametriseringen (f.eks. afgrænsning af zoner, indenfor hvilke parametrene er homogene) på geologiske og andre fysiske tolkninger, der retfærdiggør den valgte opdeling af systemet i parameterzoner. Herved minimeres risikoen for at parameterestimationen reduceres til en ren datafitningsøvelse.
- Anvendelse af logaritmisk transformeret hydraulisk ledningsevne, da det minimerer graden af ikke-linearitet mellem modelparametre og trykniveau (Carrera & Neuman, 1986)
- Tilstræbe højst mulig sensitivitet af modelparametrene. Dette er der flere muligheder for:
 - Det er vigtigt at foretage et hensigtsmæssigt valg af kalibreringsparametre, der ikke resulterer i en eller flere usensitive parametre. Dette vil ofte betyde, at antallet af kalibreringsparametre skal holdes relativt lavt.
 - Det kan ligeledes anbefales at anvende specificeret fluks (eller gradient) frem for specificeret trykniveau som modelrandbetingelse, da det resulterer i højere parametersensitiviteter.

14.5 PARAMETERUSIKKERHED

Fordelen ved at anvende en statistisk baseret invers model er, udover at opnå forhåbentligt optimale parameterestimer, at der bestemmes en række statistiske størrelser for de estimerede parametre. Afhængig af, hvor mange antagelser der er opfyldt, kan der udtrages mere eller mindre information om parametrenes statistiske egenskaber.

Kovariansmatricen på de estimerede parametre $\hat{\underline{b}}$ kan beregnes som (Hill, 1998)

$$Cov(\hat{\underline{b}}) = s^2 \left[\underline{\underline{J}}^T \underline{\underline{W}} \underline{\underline{J}} \right]^{-1} \quad (14.14)$$

hvor s^2 er den såkaldte estimations-varians ("calculated error variance" eller "goodness of fit")

$$s^2 = \frac{1}{N_{obs} - N_b} G(\hat{b}) = \frac{1}{N_{obs} - N_b} \left[\underline{r}^T \underline{w} \underline{r} \right] \quad (14.15)$$

Udtryk (14.14) er udledt under antagelse om linearitet mellem modelparametre og modelrespons (f.eks. hydraulisk trykniveau) omkring de optimale parametre \hat{b} , hvilket ikke er opfyldt mht. grundvandsstrømning. Desuden antages det, at residualerne r_i er tilfældigt fordelt og ukorrelerede, hvilket også kan være svært at opfylde i praksis. Udtryk (14.14) er derfor en approksimation til den sande kovariansmatrice, og det kan forventes, at udtrykket generelt vil underestimere den sande usikkerhed på parametrene.

Indsættes (14.15) i (14.14) fås

$$\begin{aligned} Cov(\hat{b}) &= \frac{1}{N_{obs} - N_b} \left[\underline{r}^T \underline{w} \underline{r} \right] \left[\underline{J}^T \underline{w} \underline{J} \right]^{-1} \\ &= \frac{1}{N_{obs} - N_b} \left[\underline{r}^T \underline{r} \right] \left[\underline{J}^T \underline{J} \right]^{-1} \end{aligned} \quad (14.16)$$

Det ses af (14.16), at den estimerede kovariansmatrice udelukkende afhænger af residualerne $r(\hat{b})$ og Jacobi-matricen \underline{J} , som indeholder sensitivitetskoefficienterne. Intuitivt ville man forvente, at parameterusikkerheden afhænger af den usikkerhed, som de benyttede observationsdata påhæftes. Dette er imidlertid ikke tilfældet, da observationsusikkerheden alene udnyttes i estimationen af parametre via deres plads i vægtningsmatricen \underline{w} . Alt andet lige vil parameterusikkerheden øges, når residualerne r_i vokser, mens den vil blive reduceret, når parametersensitiviteten øges. Som det ses af udtryk (14.16) er der omvendt proportionalitet mellem parameterusikkerheden og parametersensitiviteten, hvilket betyder, at parametre, der har stor indflydelse på simuleringen af observationsdataene (høj sensitivitet), er godt bestemt og derfor tillægges en lille varians.

Det ses også, at præcisionen af sensitivitetskoefficienterne i Jacobimatricen \underline{J} er specielt vigtig i denne forbindelse, idet de multipliceres parvis og summeres for at beregne kovariansmatricen.

For tilfældet med 3 parametre fås en kovariansmatrice af formen

$$Cov(\hat{b}) = \begin{bmatrix} Var(b_1) & Cov(b_1, b_2) & Cov(b_1, b_3) \\ Cov(b_2, b_1) & Var(b_2) & Cov(b_2, b_3) \\ Cov(b_3, b_1) & Cov(b_3, b_2) & Var(b_3) \end{bmatrix} \quad (14.17)$$

hvor $Var(b_i)$ er variansen på parameter nr. i og $Cov(b_i, b_j)$ er kovariansen mellem parameter nr. i og j . Kovariansmatricen er symmetrisk, dvs. $Cov(b_i, b_j) = Cov(b_j, b_i)$. Korrelationsmatricen fås tilsvarende til

$$Cor(\hat{b}) = \begin{bmatrix} 1 & \frac{Cov(b_1, b_2)}{(Var(b_1)Var(b_2))^{0.5}} & \frac{Cov(b_1, b_3)}{(Var(b_1)Var(b_3))^{0.5}} \\ \frac{Cov(b_2, b_1)}{(Var(b_2)Var(b_1))^{0.5}} & 1 & \frac{Cov(b_2, b_3)}{(Var(b_2)Var(b_3))^{0.5}} \\ \frac{Cov(b_3, b_1)}{(Var(b_3)Var(b_1))^{0.5}} & \frac{Cov(b_3, b_2)}{(Var(b_3)Var(b_2))^{0.5}} & 1 \end{bmatrix} \quad (14.18)$$

Korrelationen mellem de enkelte parametre kan ligge i intervallet $[-1; 1]$. Hvis korrelationen er tæt på 1 (numeriske værdi), er de to parametre stærkt korreleret. Det er i dette tilfælde tvivlsomt, om der kan opnås entydige estimater af de to parametre, da det kun er forholdet mellem dem, som har

betydning for resultatet. Ifølge Hill et al. (1998) kan korrelationskoefficienter på over 0.95 give problemer med entydigheden.

Konfidensintervaller på de estimerede parametre kan under forudsætning om normalfordelte residualer beregnes vha. (Hill, 1998)

$$b_i \pm t\left(n, 1.0 - \frac{\alpha}{2}\right) s_{b_i} \quad (14.19)$$

ved et signifikansniveau på α , hvor $t(\cdot)$ er t-fordelingen, og s_{b_i} er standard afvigelsen for parameter nr. i . For et signifikansniveau på 5%, svarende til et 95% konfidensinterval, fås for $n \rightarrow \infty$, at $t(\infty, 0.975) = 1.960$, mens den tilsvarende værdi på 32% niveau, svarende til et 68% konfidensinterval, er på 0.995. Hvis der er mange observationsdata til rådighed i forhold til antallet af parametre (30 eller derover) vil man (approsimativt) med 68% sikkerhed kunne sige, at parameter b_i vil ligge i intervallet $[b_i - s_{b_i}; b_i + s_{b_i}]$, mens der er 95% sikkerhed for, at parameteren vil ligge i intervallet $[b_i - 2s_{b_i}; b_i + 2s_{b_i}]$.

Variationskoefficienten, C_v , givet ved

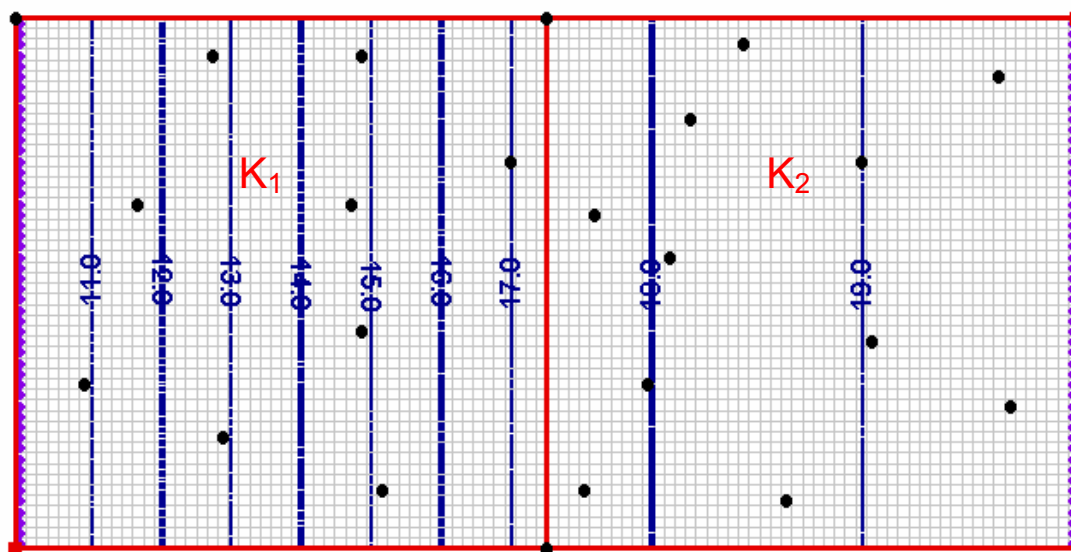
$$C_v = \frac{\sqrt{\text{Var}(\hat{b}_i)}}{\hat{b}_i} = \frac{s_{b_i}}{\hat{b}_i} \quad (14.20)$$

er et udtryk for den relative usikkerhed på et parameterestimat. En variationskoefficient $C_v \leq 1.0$ vil normalt betyde, at parameteren er godt bestemt. Hvis $C_v > 10$ vil det i nogle tilfælde være mere hensigtsmæssigt at bestemme parameteren ud fra forhåndsviden frem for at estimere den vha. optimeringsalgoritmen. Et 95% konfidensinterval for en parameter med $C_v = 10$ er approsimativt givet ved $[b_i - 20 b_i; b_i + 20 b_i]$, hvilket indikerer, at parameteren er relativt dårligt bestemt. I nogle tilfælde vil det være muligt at angive et mindre bredt interval for den sande værdi af parameteren ud fra erfaringsværdier og foreliggende undersøgelser. I så fald er der grund til at gentage optimeringen, hvor kun parametre, der kan estimeres præcist, medtages i estimationen.

14.6 EKSEMPEL

I figur 14.4 er et to-dimensionalt strømningssystem vist. Randbetingelserne består af fastholdt trykniveau på 10 m på den vestlige rand og fastholdt trykniveau på 20 m ved den østlige rand. Mod nord og syd er der impermeable grænser. Der er ingen grundvandsdannelse i området. Aquiferen simuleres som et artesisk magasin med en ensformig lagtykkelse på 10 m. Diskretiseringen i horisontal retning er på 200 m, mens der i vertikal retning benyttes et modellag. Fordelingen af hydraulisk ledningsevne er givet ved to lige store områder, hvor den vestlige ende har en ledningsevne på $K_1 = 8.64 \text{ m/d}$ ($1 \cdot 10^{-4} \text{ m/s}$), og den østlige ende har en ledningsevne, der er tre gange så stor, $K_2 = 26 \text{ m/d}$ ($3 \cdot 10^{-4} \text{ m/s}$). Der er defineret 20 observationspunkter jævnt fordelt over området, og i første omgang vil det ud fra disse 20 observationer af hydraulisk trykniveau (med en observationsusikkerhed svarende til en standardafvigelse på 0.5 m) blive forsøgt at estimere den hydrauliske ledningsevne i området vha. optimeringsprogrammet PEST.

I første omgang antages K at være homogen og horisontal isotrop. Initielt sættes $K = 1 \text{ m/d}$. I optimering nr. 1 kalibreres der kun mod de 20 trykniveau-observationer. Optimeringen resulterer i et estimat af K på 25 m/d med et 95% konfidensinterval på $[2.5 \cdot 10^{-299}; 1 \cdot 10^{300} \text{ m/d}]$, se tabel 14.1, dvs. K -værdien i systemet er praktisk talt ubestemmelig. Optimeringen konvergerer efter tre iterationer, hvilket er det antal iterationer, som PEST er blevet bedt om at foretage uden forbedringer i objektivværdien G . Der opnås nemlig ingen ændring i G , når K ændres, da trykniveaufordelingen er totalt uafhængig af den specificerede K -værdi. Trykniveaufordelingen i et system med fastholdte randbetingelser i opstrøms og nedstrøms ende og kun én K -værdi, er uafhængig af den valgte K -værdi, og ved nærmere eftersyn findes der meget små sensitiviteter og composite parameter sensitiviteter (se afsnit 14.8). Estimatet af K i denne opsætning er helt afhængig af hvilken K -værdi, der bruges som startværdi for optimeringen.



Figur 14.4 Illustration af setup. De røde linier angiver zonerings af hydraulisk ledningsevne. De sorte punkter angiver lokaliteter, hvor der er observationer af hydraulisk trykniveau. De blå linier er konturer af fordelingen af hydraulisk ledningsevne.

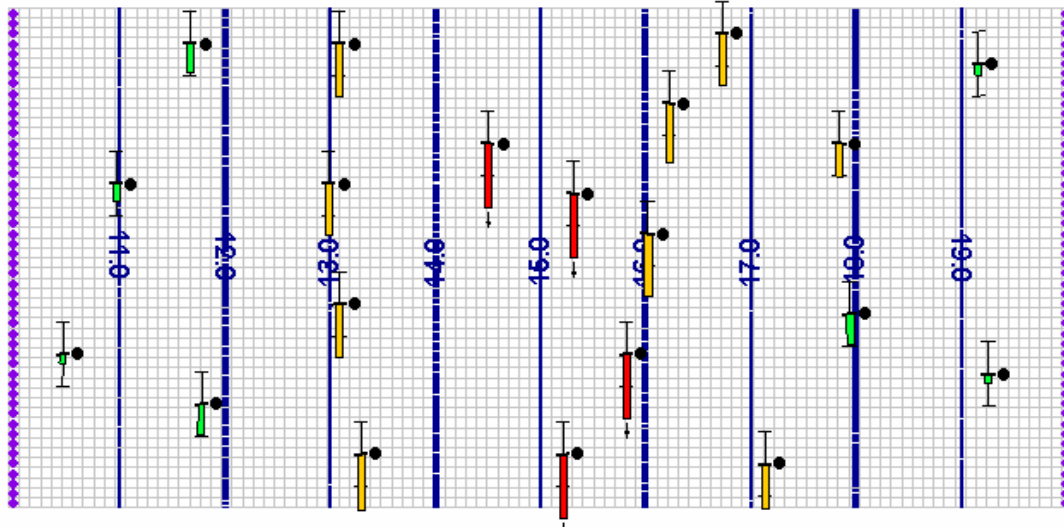
Det er derfor nødvendigt at bibringe optimeringen mere information, hvis der skal kunne estimeres entydige parametre. Ud fra den sande løsning kendes fluksen til nedstrøms rand i den vestlige ende af systemet, $Q = 655 \text{ m}^3/\text{d}$. Variationskoefficienten antages at være ca. 10% af afstrømningsmålingen, hvilket resulterer i en usikkerhed på afstrømningsmålingen, der svarer til en standardafvigelse på ca. $65 \text{ m}^3/\text{d}$. Inkluderes denne observation i optimeringen findes resultatet angivet under optimering nr. 2 i tabel 14.1. Der estimeres en hydraulisk ledningsevne, som ligger mellem de to sande værdier på 8.64 og 26 m/d, hvor de 13 m/d svarer til det harmoniske gennemsnit af de to værdier (strømning gennem serieforbindelse), hvorved den korrekte fluks til nedstrøms rand beregnes. Da den homogene K-værdi stadig ikke har nogen indflydelse på simuleringen af det hydrauliske trykniveau, findes den samme overensstemmelse mellem simuleret og observeret trykniveau, se figur 14.5.

Tabel 14.1 Resultater fra optimering af en enkelt homogen K-zone. SD angiver standardafvigelsen på observationerne, som benyttes i vægtningen af data.

Optimering nr.	Obs. Usikkerhed	K-estimat (m/d)	Lav 95% konf. Int. (m/d)	Høj. 95% konf. Int. (m/d)	RMS _h (m)
1	SD _h = 0.5 m	25.0	$2.5 \cdot 10^{-299}$	$1 \cdot 10^{+300}$	1.52
2	SD _h = 0.5 m; SD _Q = 65 m ³ /d	13.0	7.1	23.6	1.52
3	SD _h = 0.1 m; SD _Q = 65 m ³ /d	13.0	0.66	255.9	1.52
4	SD _h = 0.1 m; SD _Q = 13 m ³ /d	13.0	7.1	23.6	1.52

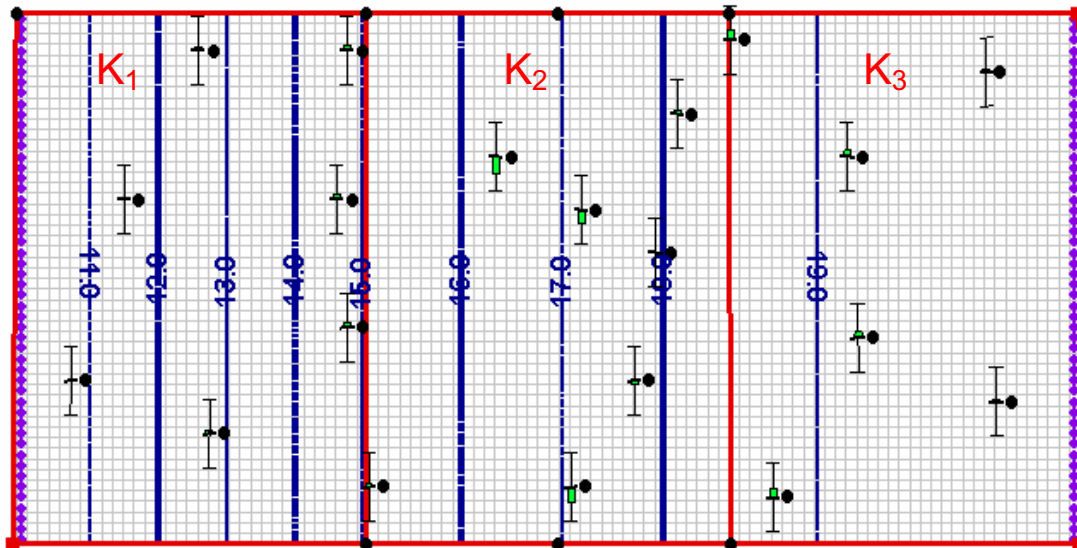
For at undersøge indflydelsen af vægtningen mellem trykniveauobservationerne og fluksobservationen, sættes standardafvigelsen på de observerede trykniveauer ned med en faktor 5 fra 0.5 m til 0.1 m. Optimeringen gentages og resultatet ses under optimering nr. 3 i tabel 14.4. Estimatet af K er stadig på 13 m/d som i optimering nr. 2, men 95% konfidensintervallet er nu betydeligt bredere, til trods for at usikkerheden på de observerede trykniveauer er reduceret. Dette tilsyneladende paradoksale resultat skyldes, at fluksobservationen nu tillægges relativt mindre vægt, og da det er

fluksobservationen, der giver hele informationen til bestemmelsen af K, bliver konfidensintervallet bredere. Hvis standardafvigelsen på fluksobservationen også reduceres med en faktor 5 (fra 65 til 13 m³/d), findes nøjagtigt samme resultat som angivet under optimering nr. 2. Det absolutte niveau af vægtningen af data har altså ingen betydning for estimationen af usikkerheder på kalibreringsparametrene, det er kun fordelingen af vægtene, der er afgørende.



Figur 14.5 Overensstemmelse mellem “observeret” og simuleret trykniveau for optimeringer nr. 1-4 (se tabel 14.1).

For at reproducere de observerede trykniveauobservationer bedre, besluttes det at definere tre zoner for hydraulisk ledningsevne, se figur 14.6. Både fluks ved nedstrøms rand og de 20 trykobservationer benyttes som kalibreringsgrundlag (standardafvigelse på hhv. 65 m³/d og 0.5 m). Resultatet af optimeringen fremgår af tabel 14.2.



Figur 14.6 Illustration af setup for de inverse kørsler nr. 4 – 6, se tabel 14.1. De røde linier angiver zoner af K. De blå linier er konturer af beregnet hydraulisk ledningsevne.

Tabel 14.2 Resultat af estimering af de tre K-zoner illustreret i figur 14.6. $RMS_h = 0.21$ m.

	K-Estimat (m/d)	95% konf. lav (m/d)	95% konf. høj (m/d)
K_1	8.4	7.6	9.2
K_2	12.4	11.0	14.1
K_3	31.7	26.6	37.7

Som det ses i tabel 14.2 resulterer optimering i fornuftige estimater af de hydrauliske ledningsevner. Den dårlige parametrisering, hvor K-zonerne ikke er i overensstemmelse med de sande K-zoner, resulterer imidlertid i, at estimatet af K_3 , som ligger helt inde i zonen med en sand hydraulisk ledningsevne på 26 m/d, er overestimeret, og selv 95% konfidensintervallet indeholder ikke den sande værdi.

Ofte vil der i praktiske situationer være betydelig usikkerhed om grundvandsdannelsen til modellen. Det forsøges derfor at estimere grundvandsdannelsen sammen med de tre K-parametre fra det ovenstående eksempel. Det skal nævnes, at der i dette tilfælde både introduceres en parametriseringsfejl (forkert zonerings af K) og en konceptuel fejl (forkert øvre randbetingelse). Eksemplet er imidlertid ikke urealistisk, da begge fejl ofte vil blive begået i virkelige tilfælde, uden at man er klar over det. Startværdien for grundvandsdannelsen sættes til $1 \cdot 10^{-4}$ m/d og resultatet af optimeringen ses i tabel 14.3.

Tabel 14.3 Resultat af estimering af de tre K-zoner illustreret i figur 14.6, samt en homogen grundvandsdannelse. $RMS_h = 0.20$ m.

	Estimat (m/d)	95% konf. lav (m/d)	95% konf. høj (m/d)
K_1	7.6	6.6	8.8
K_2	8.1	4.8	13.5
K_3	12.4	2.6	58.6
Grundv.dan.	$9.3 \cdot 10^{-6}$	$3 \cdot 10^{-8}$	$1.9 \cdot 10^{-5}$

Grundvandsdannelsen estimeres relativt godt, idet den nærmer sig den sande værdi på nul. Til gengæld har inkluderingen af grundvandsdannelsen i optimeringen stor indflydelse på optimeringen af de hydrauliske ledningsevner. Specielt for K_3 opnås en værdi, der ligger langt fra den sande værdi på 26 m/d, og samtidig ses 95% konfidensintervallet at være betydeligt bredere end i det foregående eksempel. Årsagen til det relativt dårlige resultat kan findes ved at undersøge parameterkorrelationerne, se tabel 14.4.

Tabel 14.4 Korrelationsmatrice for kalibreringsparametrene anført i tabel 14.3.

	K_1	K_2	K_3	Grundv.dan.
K_1	1.0	0.79	0.78	-0.69
K_2	0.79	1.0	0.97	-0.95
K_3	0.78	0.97	1.0	-0.99
Grundv.dan.	-0.69	-0.95	-0.99	1.0

Der ses at være stor korrelation mellem grundvandsdannelsen og de hydrauliske ledningsevner, og korrelationen er specielt høj mht. K_3 , som både var svær at estimere og havde det brede konfidensinterval. Eksemplet illustrerer, hvordan høj parameterkorrelation resulterer i problemer med at bestemme entydige estimater af kalibreringsparametrene.

Det skal bemærkes, at der i de ovenstående eksempler kun er anvendt støjfri observationsdata, dvs. modellen er blevet kalibreret mod de sande værdier af hydraulisk trykniveau og fluks. I praksis vil disse data altid være fejlbehæftede i større eller mindre grad, hvilket vil forstærke problemerne med at estimere troværdige parameterverdier. Endelig skal det nævnes, at når den rigtige zonerings af K, dvs. to lige store zoner som vist i figur 14.2, gives som input til den inverse model, finder den automatiske kalibrering uden problemer frem til de sande værdier af K.

14.7 ANALYSE AF OUTPUT FRA INVERS MODEL

Et af de meget brugbare resultater, som en invers model kan tilvejebringe, er sensitiviteterne, som er de afledte af modelresultaterne i observationspunkter mht. kalibreringsparametrene. Sensitiviteterne i Jacobi-matricen (14.12) kan udnyttes på to måder. Hvis der foretages en summation af elementerne i en søjle i J findes en slags akkumuleret (eller "composite") parameter-sensitivitet, der angiver den samlede sensitivitet af den pågældende parameter. En parameter med stor sensitivitet vil have stor indflydelse på resultatet af modellens prediktioner i de punkter, hvor der er observationsdata til rådighed. F.eks. vil man oftest finde, at for højpermeable enheder vil den vertikale hydrauliske ledningsevne være insensitiv, mens den horisontale hydrauliske ledningsevne vil være sensitiv. Omvendt forholder det sig normalt med lav-permeable enheder. Hvis sensitiviteterne i Jacobi-matricen summeres rækkevis, vil resultatet være en slags akkumuleret observations-sensitivitet, der angiver informationsværdien af den enkelte observation. En lille akkumuleret observations-sensitivitet betyder, at den pågældende måling ikke har ret stor betydning for bestemmelsen af kalibreringsparametrene. Omvendt betyder en stor akkumuleret observations-sensitivitet, at målingen er vigtig for estimation af parametrene.

UCODE/MODEFLOW2000 og PEST foretager beregningerne af de akkumulerede sensitiviteter på to forskellige måder. I UCODE/MODEFLOW2000 findes de såkaldte dimensionsløse skalerede sensitiviteter ("dimensionless scaled sensitivity"), ss_{ij}

$$ss_{ij} = \frac{\partial \psi_{sim,j}}{\partial b_i} \frac{b_i}{s_{jj}} \quad (14.21)$$

hvor $s_{jj} = 1/(w_{jj})^{0.5}$. Hvis vægtene w_{jj} specificeres som anført i (14.2), vil s_{jj} være lig standardafvigelsen på observation nr. j. En dimensionsløs skaleret sensitivitet fortæller om den enkelte observations informationsindhold mht. en enkelt parameter. Man kan også vælge at tolke ss som den enkelte parameters indflydelse på en enkelt observation.

Beregningen af "composite scaled parameter sensitivity" foretages som (idet vægtningsmatricen w antages at være en diagonal-matrice)

$$css_i = \sqrt{\frac{1}{N_{obs}} \sum_{j=1}^{N_{obs}} \left(\frac{\partial \psi_{sim,j}}{\partial b_i} \frac{b_i}{s_{jj}} \right)^2} \quad (14.22)$$

hvor N_{obs} er antallet af observationsdata. Der foretages, som beskrevet ovenfor, en søjlevis summation af (de kvadrerede) sensitiviteter i Jacobimatricen (14.12). I PEST beregnes tilsvarende der en "composite parameter sensitivity"

$$cps_i = \frac{1}{N_{obs}} \sqrt{\sum_{j=1}^{N_{obs}} \left(\frac{\partial \psi_{sim,j}}{\partial b_i} \frac{1}{s_{jj}} \right)^2} \quad (14.23)$$

der svarer til (14.22) ovenfor. Det ses, at cps ikke er dimensionsløs, da sensitiviteterne ikke multipliceres med parameterværdien, som det er tilfældet i (14.22).

Til forskel fra UCODE/MODEFLOW2000 bestemmes der i PEST også en "composite observation sensitivity", cos , som er givet ved

$$cos_j = \frac{1}{N_b} \sqrt{\sum_{i=1}^{N_b} \left(\frac{\partial \psi_{sim,j}}{\partial b_i} \frac{1}{s_{jj}} \right)^2} \quad (14.24)$$

hvor N_b er antallet af kalibreringsparametre. Udtryk (14.24) kvantificerer informationsværdien af observation nr. j. Som anført i Doherty (2000) skal man være opmærksom på, at en høj værdi af cos ikke nødvendigvis betyder, at observationen er meget vigtig for optimeringen. Hvis der f.eks. findes to observationer lige ved siden af hinanden, vil begge få beregnet en høj cos , men pga.

indbyrdes korrelation, vil informationsværdien af de to målinger tilsammen ikke svare til summen af deres akkumulerede (composite) observations-sensitivitet.

Det er vigtigt at pointere, at udpegelsen af vigtige parametre på baggrund af sensitivitetsanalyser, der udføres i kalibreringssituationen, ikke nødvendigvis kan overføres til prediktionssituationen. Hvis det er formålet at bestemme sænkningen som følge af installation af en ny kildeplads, kan det være, at simuleringen af sænkingskeglen omkring kildepladsen er sensitiv overfor nogle parametre, der ikke blev fundet sensitive i kalibreringssituationen. I dette eksempel skabes der i prediktionssituationen et andet strømningsfelt end det var tilfældet i kalibreringssituationen, og der er derfor en vis sandsynlighed for, at det er nogle andre parametre, der er vigtige i den nye situation. Ligeledes kan det være, at prediktionen af indvindingsoplande eller transporttiden til en eksisterende kildeplads er sensitiv overfor andre parametre end det blev konstateret under kalibreringen. I dette eksempel er strømningsfeltet i kalibrerings- og prediktions-situationen den samme, men det er nogle andre variable, der er fokus på, f.eks. er der en vis sandsynlighed for, at transporttiden til en kildeplads er sensitiv overfor nogle parametre, der ikke havde stor indflydelse på simuleringen af det hydrauliske trykniveau.

Det er derfor vigtigt at afrapportere resultatet af en sensitivitetsanalyse foretaget med de optimale parameterestimer, så det er klart, hvilke parametre der er godt bestemt, og hvilke parametre, der er mindre godt bestemt. Dette vil for kalibreringsparametrene også fremgå af den estimerede parameterusikkerhed, f.eks. i form af 95% konfidensintervaller eller angivelse af den estimerede standardafvigelse. Desuden er det vigtigt, at der også gennemføres en sensitivitetsanalyse eller følsomhedsanalyse i prediktionssituationen, for at finde frem til, hvilke parametre der er vigtige for simulering af de variable, der fokuseres på i denne situation. Såfremt der ikke er overensstemmelse mellem parametrene, der blev fundet sensitive i hhv. kalibreringssituationen og prediktionssituationen, er der grund til at være påpasselig, for i den situation skal der genereres simuleringresultater med en model, hvor en eller flere af de vigtige parametre er dårligt bestemt. Dette er en vigtig information at viderebringe, så der kan blive ageret i henhold dertil. Det kan eventuelt blive besluttet på den foreliggende information at søge at tilvejebringe information om de parametre, som der er problemer med, hvis dette er fysisk og økonomisk muligt. Alternativt skal de producerede resultater tolkes med varsomhed, idet de sandsynligvis vil være behæftet med stor usikkerhed.

14.7.1 Urealistiske parameterestimer

Anvendelse af automatisk parameteroptimering resulterer til tider i estimation af urealistiske parameterverdier, hvor f.eks. den hydrauliske ledningsevne af et lerlag er urealistisk høj, den hydrauliske ledningsevne for et sandlag er lavere end for et siltlag eller vandløbskonduktiviteten i et sandet område er meget lille. Den slags problemer opstår normalt ikke, når der udføres manuel kalibrering, da hydrogeologen ikke vil afprøve urealistiske parameterkombinationer under kalibreringen.

Der kan være forskellige forklaringer på, at der estimeres urealistiske parametre. Dels kan optimeringen være blevet fanget i et lokalt minimum, og i det tilfælde vil overensstemmelsen mellem observerede og simulerede data sandsynligvis også være dårlig. En anden forklaring kan være, at optimeringen pga. entydighedsproblemer har konvergeret mod et minimum, der giver god overensstemmelse med observationsdata, men som resulterer i parameterverdier, der ikke er i overensstemmelse med virkeligheden. Dette problem svarer til, hvad der af Beven (1993) benævnes "equifinality", hvilket beskriver, at forskellige modeller kan resultere i repræsentationer af virkeligheden, som ikke kan afvises som utroværdige. I afsnit 14.3 er forskellige årsager til entydighedsproblemer beskrevet mere udførligt.

En tredje forklaring på urealistiske parameterestimer er, at der er fejl i den opstillede model (Poter & Hill, 1996). Der kan f.eks. være tale om, at den konceptuelle model (hydrogeologisk tolkningsmodel) er fejlbehæftet, eller at der er foretaget en forkert parametrisering. Den inverse model vil forsøge at kompensere for disse fejl ved at tillægge kalibreringsparametrene værdier, der ikke er i overensstemmelse med det virkelige system. Urealistiske parameterestimer kan i dette tilfælde ikke forkastes under påskud af, at den inverse model ikke "virker". I stedet bør et sådant resultat anvendes konstruktivt, hvor resultaterne fra den inverse model analyseres (f.eks. fordeling af resi-

dualer i systemet og usikkerhedsestimater på parametrene) sammen med f.eks. den geologiske model eller den anvendte parametrisering i et forsøg på at finde evt. fejl i modelopsætningen.

Alternativt kunne flere forskellige modeller opstilles og sammenlignes på baggrund af primært (a) hvor realistiske de estimerede parametre anses for at være, (b) hvor god overensstemmelse der er med observationsdata (f.eks. RMS-værdi), samt (c) hvor tilfældigt residualerne er fordelt i modellen (Hill, 1998). På baggrund af de tre kriterier kan de alternative modeller rangordnes og nogle kan eventuel helt forkastes ud fra nogle på forhånd opstillede kriterier.

14.8 EVALUERING AF KALIBRERINGSRESULTAT

Reilly & Harbaugh (2004) anbefaler, at der tages stilling til følgende spørgsmål ved evaluering af, om en model er kalibreret tilstrækkeligt godt:

1. En den konceptuelle model (hydrogeologiske tolkningsmodel, parametrisering, etc.) for det betragtede system rimelig?
2. Er den matematiske repræsentation af randbetingelserne rimelige for formålet med opgaven?
3. Reproducerer det simulerede trykniveau og strømningsfordeling de vigtige aspekter af strømningssystemet, såsom størrelse og retning af trykniveaunkonturerne?
4. Er der en rimelig overensstemmelse mellem observeret og simuleret trykniveau og afstrømning (udtrykt ved kvantitative mål som f.eks. RMS-værdien) i forhold til formålet med undersøgelserne?
5. Fordeler residualerne sig ligeligt over modelområdet, dvs. er der en tilfældig fordeling af hhv. over- og under-prediktering af de observerede trykniveauer? Hvis det ikke er tilfældet, er der så en hydrogeologisk begrundelse for at ændre modellen og derved opnå en mere tilfældig rumlig fordeling af residualerne?

De to første spørgsmål er primært relateret til den konceptuelle forståelse af det betragtede system samt modelopsætningen for systemet. Disse spørgsmål er der normalt taget stilling til inden kalibreringen påbegyndes, men hvis det i løbet af kalibreringsprocessen viser sig, at det ikke er muligt at opnå en tilfredsstillende overensstemmelse med de observerede størrelser, vil det være naturligt at stille spørgsmålene igen, og eventuelt forsøge at korrigere den hydrogeologiske tolkningsmodel inkl. geologisk model og randbetingelser).

Specifikt for resultater af en invers kalibrering vil det være hensigtsmæssigt yderligere at stille følgende spørgsmål:

- A. Har de estimerede parametre rimelige størrelser, der ligger indenfor de forventede intervaller?
- B. Er de estimerede parameterusikkerheder og –korrelationer rimelige?
- C. Er der opnået et globalt minimum i den inverse kalibrering?

Det skal sandsynliggøres, at svarene på ovenstående spørgsmål er bekræftende. Hvis der ikke kan gives positive svar, skal det forklares, hvad der er årsag til de fundne problemer.

14.9 REFERENCER

Beven, K.J. (1993) Phophesy, reality and uncertainty in distributed hydrological modelling, *Adv. Water Resources*, 16, 41-51.

Carrera, J., and Neuman, S.P. (1986) Estimation of aquifer parameters under transient and steady state conditions: 2. Uniqueness, stability, and solution algorithms, *Water Resources Research*, 22(2), 211-227.

Doherty, J. (2000) PEST, Model independent parameter estimation. *Watermark Numerical Computing*, Australia, 249 p.

Duan, Q., Sorooshian, S., Gupta, V. (1992) Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resources Research*, 28(4), 1015-1031.

- Ginn, T.R., Cushman, J.H. (1990) Inverse methods for subsurface flow: A critical review of stochastic techniques, *Stochastic Hydrol. Hydraul.*, 4, 1-26.
- Hill, M.C. (1998) Methods and guidelines for effective model calibration, U.S. Geological Survey, Water-Resources investigations Report 98-4005, 90 p.
- Hill, M.C., Cooley, R.L., Pollock, D.W. (1998) A controlled experiment in groundwater flow model calibration, *Ground Water*, 36(3), 520-535.
- Hill, M.C., Banta, E.R., Harbaugh, A.W., Anderman, E.R. (2000), MODFLOW-2000, The U.S. Geological Survey modular ground-water model – User guide to the Observation, Sensitivity, and Parameter-estimation processes and three post-processing programs, U.S. Geological Survey, Open-file Report 00-184, 209 p.
- Hill, M.C., Tiedeman, C.R. (2002) Weighting observations in the context of calibrating groundwater models, In: K. Kovar & Z. Hrkal (eds.), *Calibration and reliability in groundwater modeling: A few steps closer to reality*, ModelCARE 2002, IAHS Publication 277, 196-203.
- Hopmans, J.W., Šimunek, J., Romano, N., Durner, W. (2002) Inverse methods, In: *Methods of soil analysis, Part 4, Physical methods*, Eds.: J.H. Dane & G.C. Topp, Soil Science Society of America, Inc., Madison, Wisconsin, USA, 963-1008.
- Madsen, H. (2000) Automatic calibration of a conceptual rainfall-runoff model using multiple objectives, *J. Hydrology*, 235, 276-288.
- Madsen, H. (2003) Parameter estimation in distributed hydrological catchment modeling using automatic calibration with multiple objectives, *Advances in Water Resources*, 26, 205-216.
- Madsen, H., Kristensen, M. (2002) A multi-objective calibration framework for parameter estimation in the MIKE SHE integrated hydrological modeling system, In: *ModelCARE 2002, Proceeding of the 4th international conference on Calibration and Reliability in Groundwater Modelling*, 6 p.
- Madsen, H., Jacobsen, T. (2001) Automatic calibration of the MIKE SHE integrated hydrological modeling system, 4th DHI softwater conference, 20 p.
- Poeter, P.P., and Hill, M.C. (1996) Unrealistic parameter estimates in inverse modeling: a problem or a benefit for model calibration? In: *Calibration and Reliability in Groundwater Modelling, Proceeding of the Model CARE 96 Conference*, IAHS Publ. no. 237, 277-285.
- Poeter, E.P., Hill, M.C. (1997) Inverse models: A necessary next step in ground-water modeling, *Ground Water*, 35(2), 250-260.
- Poeter, E. P., Hill, M.C. (1998) Documentation of UCODE, A computer code for universal inverse modeling: U.S. Geological Survey Water-Resources Investigations Report 98-4080, 116 p.
- Poeter, E.P., Hill, M.C. (1999) UCODE, a computer code for universal inverse modeling, *Computers & Geosciences*, 25, 457-462.
- Reilly, T.E., and Harbaugh, A.W. (2004) Guidelines for evaluating ground-water flow models, U.S. Geological Survey Scientific Investigations Report 2004-5038, 30 p.

Sonnenborg, T.O. (2001) Kalibrering af strømningsmodel, Ståbi i grundvandsmodellering, Ed. Henriksen et al., 22. p.

Sorooshian, S., Gupta, V.K. (1995) Model calibration, In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology, Water Resources Publications, Colorado, 23-68.