An approach to quantum-computational hydrologic inverse analysis

Daniel O'Malley

EES-16, Los Alamos National Laboratory

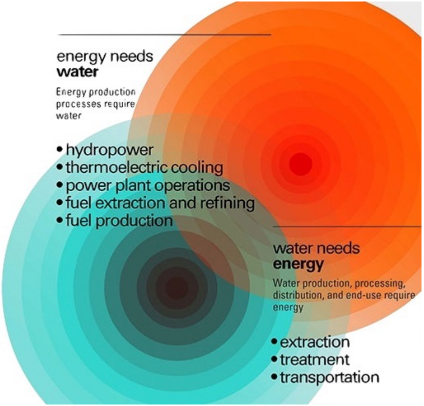
September 25, 2018

LA-UR-17-23780

Water is important

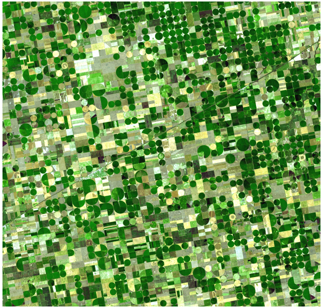
We are water

Energy



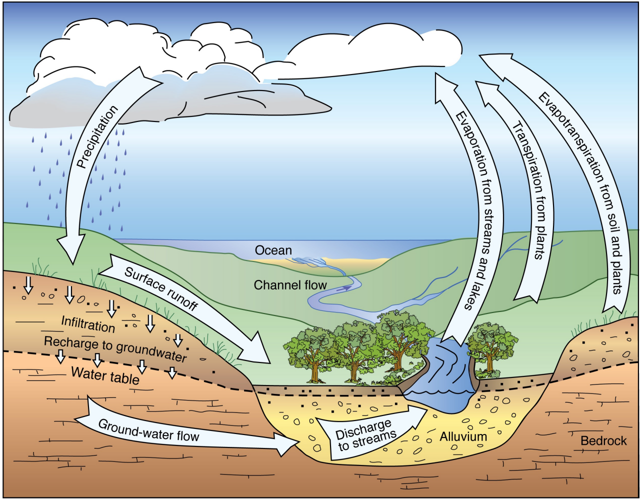
www.worldbank.org/thirstyenergy

Food



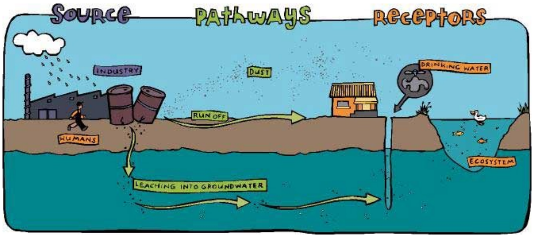
NASA Earth Observatory

What is hydrology?



THE HYDROLOGIC CYCLE

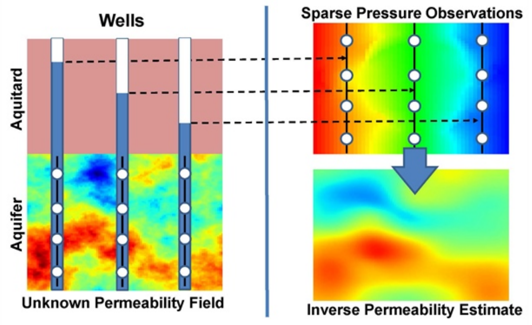
Contaminated groundwater



- The US has more than 100,000 contaminated groundwater sites and the total cost to remediate them will exceed $100,000,000,000[1]
- In order to design cost-effective remedial measures, the properties of the subsurface must be understood
- Hydrologic inverse analysis helps us understand the properties of the subsurface

[1] National Research Council. Alternatives for managing the nation's complex contaminated groundwater sites. National Academies Press, 2013.

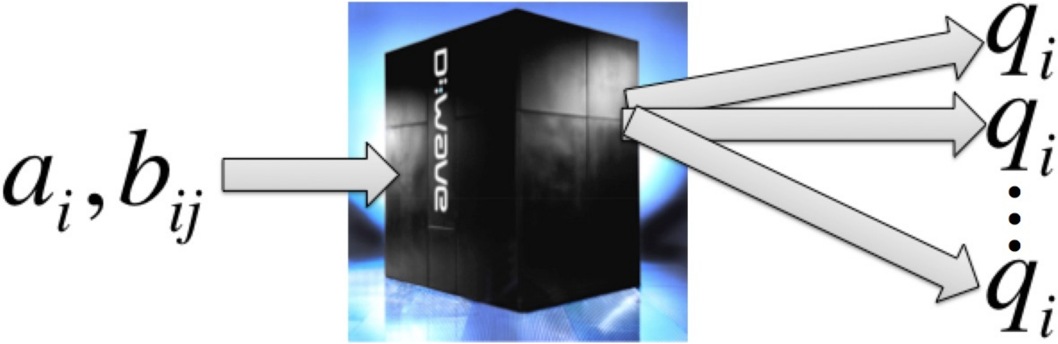
Hydrologic inverse analysis



$$\nabla \cdot (k \nabla h) = 0$$

Going from k to h is easy
Going from h to k is hard

D-Wave: What does it do?



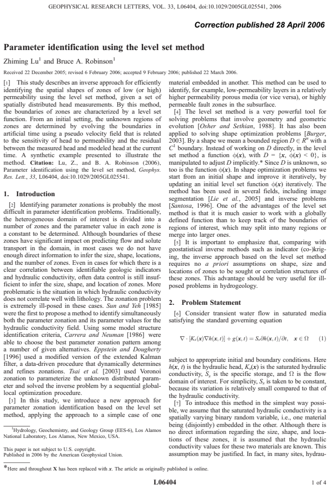
0^{th} order approximation

$$\hat{\mathbf{q}} = \min_{\mathbf{q} \in \{0,1\}^n} \sum_{i=1}^{n} a_i q_i + \sum_{i=1}^{n} \sum_{j=1}^{i-1} b_{ij} q_i q_j$$

1^{st} order approximation

$$\mathbf{q} \sim e^{-\beta\left(\sum_{i=1}^{n} a_i q_i + \sum_{i=1}^{n} \sum_{j=1}^{i-1} b_{ij} q_i q_j\right)}$$

Binary variables in hydrology? Indeed.



GEOPHYSICAL RESEARCH LETTERS, VOL. 33, L06404, doi:10.1029/2005GL025541, 2006

Correction published 28 April 2006

Parameter identification using the level set method

Zhiming Lu[1] and Bruce A. Robinson[1]

Received 22 December 2005; revised 6 February 2006; accepted 9 February 2006; published 22 March 2006.

[1] This study describes an inverse approach for efficiently identifying the spatial shapes of zones of low (or high) permeability using the level set method, given a set of spatially distributed head measurements. By this method, the boundaries of zones are characterized by a level set function. From an initial setting, the unknown regions of zones are determined by evolving the boundaries in artificial time using a pseudo velocity field that is related to the sensitivity of head to permeability and the residual between the measured head and modeled head at the current time. A synthetic example presented to illustrate the method. **Citation:** Lu, Z., and B. A. Robinson (2006), Parameter identification using the level set method, *Geophys. Res. Lett.*, *33*, L06404, doi:10.1029/2005GL025541.

1. Introduction

[2] Identifying parameter zonations is probably the most difficult in parameter identification problems. Traditionally, the heterogeneous domain of interest is divided into a number of zones and the parameter value in each zone is a constant to be determined. Although boundaries of these zones have significant impact on predicting flow and solute transport in the domain, in most cases we do not have enough direct information to infer the size, shape, locations, and the number of zones. Even in cases for which there is a clear correlation between identifiable geologic indicators and hydraulic conductivity, often data control is still insufficient to infer the size, shape, and location of zones. More problematic is the situation in which hydraulic conductivity does not correlate well with lithology. The zonation problem is extremely ill-posed in these cases. *Sun and Yeh* [1985] were the first to propose a method to identify simultaneously both the parameter zonation and its parameter values for the hydraulic conductivity field. Using some model structure identification criteria, *Carrera and Neuman* [1986] were able to choose the best parameter zonation pattern among a number of given alternatives. *Eppstein and Dougherty* [1996] used a modified version of the extended Kalman filter, a data-driven procedure that dynamically determines and refines zonations. *Tsai et al.* [2003] used Voronoi zonation to parameterize the unknown distributed parameter and solved the inverse problem by a sequential global-local optimization procedure.

[3] In this study, we introduce a new approach for parameter zonation identification based on the level set method, applying the approach to a simple case of one material embedded in another. This method can be used to identify, for example, low-permeability layers in a relatively higher permeability porous media (or vice versa), or highly permeable fault zones in the subsurface.

[4] The level set method is a very powerful tool for solving problems that involve geometry and geometric evolution [*Osher and Sethian*, 1988]. It has also been applied to solving shape optimization problems [*Burger*, 2003]. By a shape we mean a bounded region $D \in R^n$ with a C^1 boundary. Instead of working on D directly, in the level set method a function $\phi(\boldsymbol{x})$, with $D = \{\boldsymbol{x}, \phi(\boldsymbol{x}) < 0\}$, is manipulated to adjust D implicitly.* Since D is unknown, so too is the function $\phi(\boldsymbol{x})$. In shape optimization problems we start from an initial shape and improve it iteratively, by updating an initial level set function $\phi(\boldsymbol{x})$ iteratively. The method has been used in several fields, including image segmentation [*Lie et al.*, 2005] and inverse problems [*Santosa*, 1996]. One of the advantages of the level set method is that it is much easier to work with a globally defined function than to keep track of the boundaries of regions of interest, which may split into many regions or merge into larger ones.

[5] It is important to emphasize that, comparing with geostatistical inverse methods such as indicator (co-)kriging, the inverse approach based on the level set method requires no *a priori* assumptions on shape, size and locations of zones to be sought or correlation structures of these zones. This advantage should be very useful for ill-posed problems in hydrogeology.

2. Problem Statement

[6] Consider transient water flow in saturated media satisfying the standard governing equation

$$\nabla \cdot [K_s(\boldsymbol{x})\nabla h(\boldsymbol{x},t)] + g(\boldsymbol{x},t) = S_s \partial h(\boldsymbol{x},t)/\partial t, \quad \boldsymbol{x} \in \Omega \qquad (1)$$

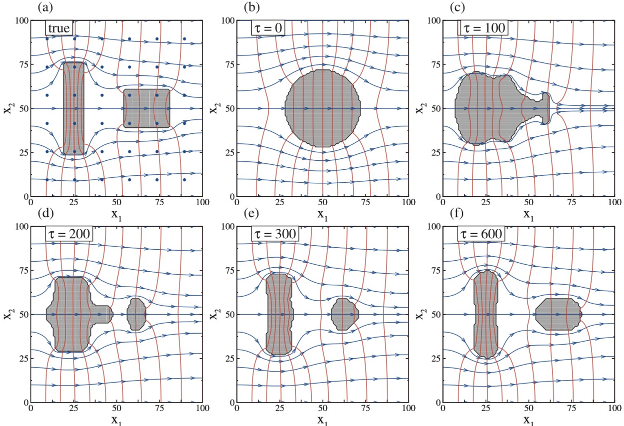
subject to appropriate initial and boundary conditions. Here $h(\boldsymbol{x}, t)$ is the hydraulic head, $K_s(\boldsymbol{x})$ is the saturated hydraulic conductivity, S_s is the specific storage, and Ω is the flow domain of interest. For simplicity, S_s is taken to be constant, because its variation is relatively small compared to that of the hydraulic conductivity.

[7] To introduce this method in the simplest way possible, we assume that the saturated hydraulic conductivity is a spatially varying binary random variable, i.e., one material being (disjointly) embedded in the other. Although there is no direct information regarding the size, shape, and locations of these zones, it is assumed that the hydraulic conductivity values for these two materials are known. This assumption may be justified. In fact, in many sites, hydrau-

[1] Hydrology, Geochemistry, and Geology Group (EES-6), Los Alamos National Laboratory, Los Alamos, New Mexico, USA.

*Here and throughout **X** has been replaced with *x*. The article as originally published is online.

"This study describes an inverse approach for efficiently identifying the spatial shapes of zones of low (or high) permeability using the level set method, given a set of spatially distributed head measurements."



Computational hydrology: a historical perspective

ground water

Historical Note/ David Deming/ History Editor

Jacob Bear: An Autobiography

by Jacob Bear[1]

It is not easy to write about myself, as requested by the history editor of *Ground Water*, especially when the request includes the explanation: "A hundred years from now, people will want to know what sort of person you were: This is your opportunity to tell them." I told him that, so far, whenever I wanted to say something to the scientific/professional community, I did so through my papers and books. Nevertheless, I agreed to try.

I was born in Haifa, Israel (then Palestine), in 1929. My parents immigrated from Russia and Poland to Israel in the early 1920s, as Zionists driven by the urge to participate in the establishment of a homeland for the Jews in their ancestral land so that Jews will have at least one place in the world to where they can flee whenever persecuted. I married my wife, Siona, in 1951. We have three children and six grandchildren.

Being a member of the underground HAGANA (defense) units since the age of 16, it was natural that when I graduated from high school, I joined the PALMACH (the Israeli underground fighting units) and fought in the War of Independence (1947–1948). In 1949, following a short period in a kibbutz, I started my studies of civil engineering at the Technion-Israel Institute of Technology, Haifa. In my senior year, I elected to graduate in the option of water resources, realizing that water—in fact, the scarcity of water—would, no doubt, be a central issue and a limiting factor in Israel's development. I liked the challenge. Also, by watching my professors, J. Breuer (water works) and S. Irmay (hydraulics) and their activities, I was convinced that this was an exciting field in which I'd be able to combine theory and practice. I liked that, too.

I received the B.Sc. (summa cum laude) in 1953, and the Dipl. Engenieur (equivalent to a P.E.) degree in 1954, both in civil engineering. I started to work as an engineer in the Planning Division of TAHAL-Water Planning for Israel Ltd., the government (now private) company in charge of water resources planning and development in Israel.

A couple years later, I was awarded a scholarship by the Dutch government and spent a year with the Government Institute for Water Supply in Scheveningen, Netherlands, studying ground water hydraulics, seawater intrusion, and the use of laboratory models as tools for solving ground water problems. My supervisor was Professor Krul of TUDelft. I learned a lot also from Mr. Santing, an excellent engineer and wonderful person.

One of the main models that I learned to use was the Hele-Shaw (parallel-plate) model (Figure 1). During the 1960s, I used this model extensively to investigate seawater intrusion (in the vertical cross section) into layered coastal aquifers, under the sharp interface approximation. To remind you, this was before the era of computers, when no other tool was available for solving regional ground water problems with a phreatic surface and/or a (assumed) sharp freshwater/seawater interface (Stefan problems). Within the framework of cooperation between the Delft Technological University and the Technion, I developed the Horizontal Hele-Shaw model that enabled investigations of seawater encroachment into (horizontal) regional coastal aquifers. For this research, I received my M.Sc. in civil engineering from the Technion in 1957.

Returning to Israel from the Netherlands, I started to work in ground water hydrology. In Israel, ground water, primarily from the coastal (sandstone) aquifer and the (limestone) mountain aquifer, constitutes the major source of water. However, the total available annual sustainable water yield is rather small, to the extent that it seriously

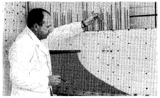


Figure 1. Investigations of seawater intrusion using a Hele-Shaw model, 1963.

[1]Jacob Bear, Professor Emeritus, Department of Civil and Environmental Engineering, Technion-Israel Institute of Technology, Haifa, 32000 Israel; 972-4-8292290; fax 972-4-8228898; cvrbear@tx.technion.ac.il; www.technion.ac.il/~cvrbear/; www.heath-hydrology.com/cmdl/

"In 1979, *Hydraulics of Groundwater* was published, in which I tried to bring the comprehensive approach and mathematical modeling of flow and contaminant transport to the field of ground water hydrology."

"Nowadays, models are accepted as fundamental tools in practice, but not long ago the question of whether models should legitimately be used as a prediction tool was still being debated."

Computational hydrology: a historical perspective

VOL. 11, NO. 6 WATER RESOURCES RESEARCH DECEMBER 1975

Identifying the Parameters of an Aquifer Cell Model

E. HEFEZ,[1] U. SHAMIR, AND J. BEAR

Department of Civil Engineering, Technion-Israel Institute of Technology, Haifa, Israel

Cell models are commonly used for forecasting water levels in aquifers. Calibration of such models is achieved through identification of their parameter values, the transmissivities and storativities of all cells, using historical data. Several methods of formulating the identification as a linear or quadratic programing problem are presented. Examples are given, results of the various methods compared, and the sensitivity of these results to errors in the data is discussed. Inflows or outflows during historical periods may also be determined by the same methods; an example for a real aquifer is presented.

INTRODUCTION

The problem of identifying the parameters of an aquifer model is known also as the inverse problem. Here identification of parameters means the determination of the distributions of storativity $S = S(x, y)$ and transmissivity $T = T(x, y)$ of an aquifer in which we assume two-dimensional flow in the horizontal plane. The complete distributions are required if one wishes to forecast the future regime in the aquifer in response to various imposed activities of pumping and recharge.

An indirect way of obtaining the sought distributions is by some trial and error technique of adjusting the various parameters until an acceptable agreement is reached between the response of the model and that of the aquifer itself under some specified operation regime. A prerequisite for the application of this approach is the availability of hydrological data (e.g., water levels and rates of pumping, discharge of springs, artificial recharge, and natural replenishment) for some period in the past called the identification or calibration period. Sometimes, certain parts of the historical hydrological data mentioned above are not known. In this case we may regard the missing information items also as unknown parameters, the values of which have to be determined during the identification procedure. The natural replenishment of an aquifer is often considered such an unknown parameter. Unless otherwise specified, we shall consider henceforth the problem of identifying the distributions of S and T only.

The main disadvantage of a trial and error technique is that it does not involve an algorithm for seeking the solution systematically. In recent years, advanced mathematical methods have been developed and implemented for determining unknown model parameters. Among these one may mention the works of *Deininger* [1969], *Vemuri and Karplus* [1969], *Korganoff* [1970], *Coats et al.* [1970], *Emsellem and de Marsily* [1971], *Kleinecke* [1971], and *Neuman* [1973]. The present work is another attempt in this direction.

In this work the aquifer is represented by a finite difference model, and linear and quadratic programing procedures are employed as tools for identifying its parameters. The proposed methods have been tested on synthetic models, the parameters of which were a priori known. This technique made it possible to check results and compare the different methods. The application of the proposed techniques to cases of practical interest is now under way and will be reported separately.

[1] Now with the Department of Computer Sciences, Haifa University, Haifa, Israel.

FINITE DIFFERENCE MODEL

Various numerical models of groundwater systems may be employed. Essentially, the models considered in this paper deal with flow in a confined aquifer or in a phreatic one in which the spatial variations of the water table are small with respect to the thickness of the aquifer. It is assumed that the aquifer is isotropic and that the flow in it is essentially two-dimensional in the horizontal (x, y) plane.

The continuity equation for this model is

$$\frac{\partial}{\partial x}\left(T\frac{\partial \phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(T\frac{\partial \phi}{\partial y}\right) + r - p = S\frac{\partial \phi}{\partial t} \qquad (1)$$

where $T = T(x, y)$ is the transmissivity of the aquifer, $S = S(x, y)$ is the storativity of the aquifer, $\phi = \phi(x, y, t)$ indicates the elevation of the water table or the piezometric surface, $r = r(x, y, t)$ represents inflow per unit (horizontal) area (e.g., by artificial and natural replenishment), and $p = p(x, y, t)$ represents withdrawal from the aquifer per unit (horizontal) area (e.g., by pumping).

The flow domain is subdivided into a network of rectangular cells (Figure 1) which serve in constructing the finite difference representation of (1).

In the cell model the continuous variables transform into discrete ones in the following way:

$$x \to x_i \quad i = 1, 2, \cdots, I \qquad (2)$$

$$y \to y_j \quad j = 1, 2, \cdots, J \qquad (3)$$

$$t \to t^n \quad n = 0, 1, 2, \cdots, N \qquad (4)$$

$$\phi(x, y, t) \to \phi_{i,j}{}^n \qquad (5)$$

$$p(x, y, t) \to p_{i,j}{}^{n+1/2} = (p_{i,j}{}^{n+1} + p_{i,j}{}^n)/2 \qquad (6)$$

$$r(x, y, t) \to r_{i,j}{}^{n+1/2} = (r_{i,j}{}^{n+1} + r_{i,j}{}^n)/2 \qquad (7)$$

$$T(x, y) \to T_{i,j} \qquad (8)$$

$$S(x, y) \to S_{i,j} \qquad (9)$$

In Figure 1 the centers of the cells are indicated in order to emphasize the fact that any property of the cell is represented by a single value which is assigned to its center. Note that the value of the transmissivity is also assigned to the centers of the cells and not to the border lines between adjacent cells, as is sometimes done. The number of transmissivity values, which are to be identified, is equal to the number of cells, whereas assignment of values to cell boundaries would result in roughly twice this number.

Two finite difference schemes, commonly used for

"The identification problem as stated in the present work is solved as a linear or a **quadratic programing problem**. The solution in the latter case is much more complicated, whereas the solution of the linear programing problem is based on readily available computer programs."

"Examination shows that the best results were obtained when [a **quadratic programming problem**] was used."

The D-Wave "solves" binary **quadratic programming problems**.

Identifying the parameters of an aquifer cell model with D-Wave

1D groundwater flow equation

Finite difference equation: $0 = \nabla \cdot (k \nabla h)$

$$0 = k_1(h_1 - h_2) + k_2(h_3 - h_2)$$

Reformulate as a least squares problem

$$0 \approx [k_1(h_1 - h_2) + k_2(h_3 - h_2)]^2$$

Fill in, say, $h_1 = 1$, $h_2 = \frac{1}{3}$, $h_3 = 0$

$$0 \approx \left(\frac{2k_1}{3} - \frac{k_2}{3}\right)^2$$

Discretize $k_i = 1 + q_i$, $q_i \in \{0, 1\}$

$$0 \approx \left(\frac{2 + 2q_1}{3} - \frac{1 + q_2}{3}\right)^2 = \frac{8}{9}q_1 - \frac{1}{9}q_2 - \frac{4}{9}q_1 q_2 + \frac{1}{9}$$

h_1 k_1 h_2 k_2 h_3

Identifying the parameters of an aquifer cell model with D-Wave

1D groundwater flow equation

Finite difference equation: $0 = \nabla \cdot (k \nabla h)$

$$0 = k_1(h_1 - h_2) + k_2(h_3 - h_2)$$

Reformulate as a least squares problem

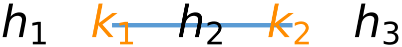
$$0 \approx [k_1(h_1 - h_2) + k_2(h_3 - h_2)]^2$$

Fill in, say, $h_1 = 1$, $h_2 = \frac{1}{3}$, $h_3 = 0$

$$0 \approx \left(\frac{2k_1}{3} - \frac{k_2}{3}\right)^2$$

Discretize $k_i = 1 + q_i$, $q_i \in \{0, 1\}$

$$0 \approx \left(\frac{2 + 2q_1}{3} - \frac{1 + q_2}{3}\right)^2 = \frac{8}{9}q_1 - \frac{1}{9}q_2 - \frac{4}{9}q_1 q_2 + \frac{1}{9}$$



Identifying the parameters of an aquifer cell model with D-Wave

1D groundwater flow equation

Finite difference equation: $0 = \nabla \cdot (k \nabla h)$

$$0 = k_1(h_1 - h_2) + k_2(h_3 - h_2)$$

Reformulate as a least squares problem

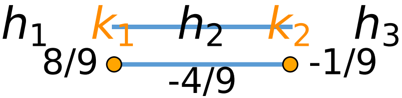
$$0 \approx [k_1(h_1 - h_2) + k_2(h_3 - h_2)]^2$$

Fill in, say, $h_1 = 1$, $h_2 = \frac{1}{3}$, $h_3 = 0$

$$0 \approx \left(\frac{2k_1}{3} - \frac{k_2}{3}\right)^2$$

Discretize $k_i = 1 + q_i$, $q_i \in \{0, 1\}$

$$0 \approx \left(\frac{2 + 2q_1}{3} - \frac{1 + q_2}{3}\right)^2 = \frac{8}{9}q_1 - \frac{1}{9}q_2 - \frac{4}{9}q_1 q_2 + \frac{1}{9}$$

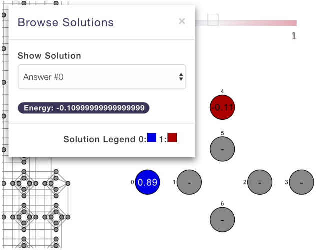


What did D-Wave have to say?

1,000 "solutions," but only 2 are distinct

Solution	Energy	Occurrences
[0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...] Graph Show	−0.11	869
[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...] Graph Show	0	131

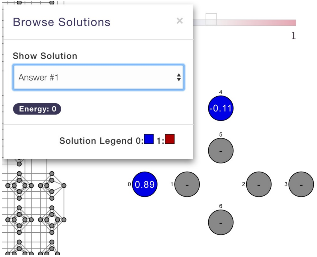
Best solution



$$q_1 = 0 \implies k_1 = 1$$
$$q_2 = 1 \implies k_2 = 2$$

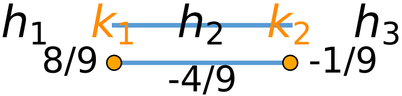
Correct 869 times out of 1,000

2nd best solution



$$q_1 = 0 \implies k_1 = 1$$
$$q_2 = 0 \implies k_2 = 1$$

Wrong 131 times out of 1,000



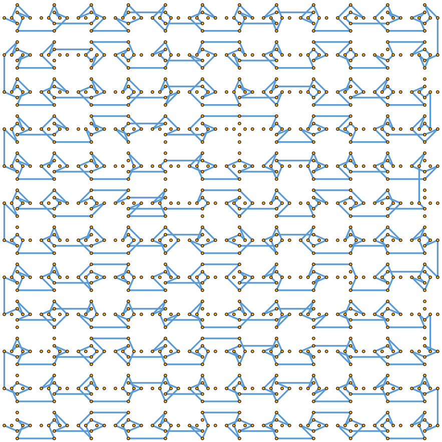
Why these solutions? Why these frequencies?

$$
\begin{aligned}
f(q_1, q_2) &= \frac{8}{9}q_1 - \frac{1}{9}q_2 - \frac{4}{9}q_1 q_2 + \frac{1}{9} \\
P(Q_1 = q_1, Q_2 = q_2) &\propto \exp\left[-\beta f(q_1, q_2)\right] \\
\beta &\approx 16.6
\end{aligned}
$$

q_1	q_2	$f(q_1, q_2)$	$P(Q_1 = q_1, Q_2 = q_2)$	D-Wave probabilities (10^6 samples)
0	0	$\frac{1}{9}$	0.136	0.136
1	0	1	5×10^{-8}	10^{-6}
0	1	0	0.863	0.863
1	1	$\frac{4}{9}$	0.0005	0.00018

Can we go bigger? Yes, much bigger, but not *that* big.

The graph for a bigger problem is like a long string



How to go pretty big

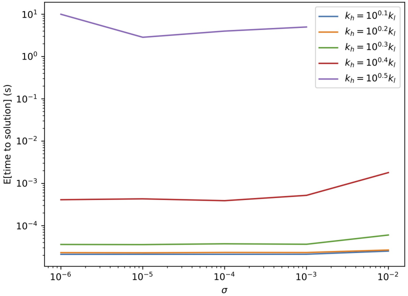
- To go big, we have to find a long, non-intersecting path through the D-Wave's graph
- Finding the longest such path is NP-hard
- Tools like KaLP[2] are designed to solve this problem
- KaLP chokes on the D-Wave 2X's graph
- Exploit what structure is in the hardware graph, use KaLP on smaller graphs
  - Decompose the hardware graph into subgraphs (several neighboring unit cells) in a snake-like pattern
  - Use KaLP to find the longest path through these subgraphs
  - Connect the paths through the subgraphs

[2] Balyo, Tomas, Kai Fieger, and Christian Schulz. "Optimal Longest Paths by Dynamic Programming." arXiv preprint arXiv:1702.04170 (2017).

What happens when we solve a "big" problem?

What is the expected amount of time to get all 972 k_i's correct?

$$\begin{aligned} h_i^{obs} &= h_i + \sigma Z \\ k_i &= k_l + q_i(k_h - k_l) \\ k_l &= 1 \\ Z &\sim N(0,1) \end{aligned}$$

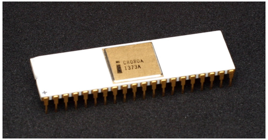


Not bad, considering the number of possible answers is
39,916,806,190,694,396,233,127,454,260,736,771,321,349,025,208,709,150,830,050,944,848,744,237,837,
379,281,315,699,159,309,852,714,021,786,848,936,883,849,904,879,448,759,767,871,873,214,783,435,965,
696,628,406,400,113,459,021,713,530,350,754,428,887,259,743,653,067,134,890,878,479,866,616,209,102,
417,407,777,777,105,368,960,883,150,142,418,137,515,120,832,847,169,904,606,880,198,557,696

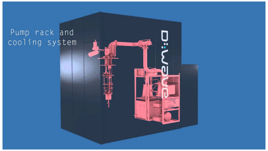
Struggles with large contrasts though (a significant practical limitation)

Is this problem big?

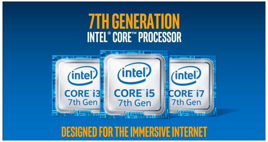
Intel 8080



D-Wave 2X



Modern CPUs



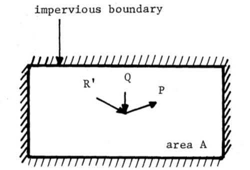
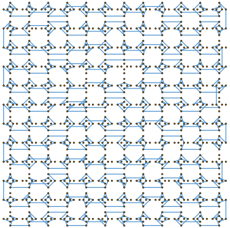
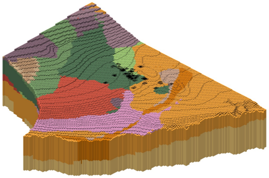


Fig. 4. A single-cell model of the Yarkon-Taninim aquifer.





1 node
2 parameters

973 nodes
972 parameters

766,283 nodes
252 parameters

Hefez, Shamir, and Bear also solved a 2D problem

VOL. 11, NO. 6 WATER RESOURCES RESEARCH DECEMBER 1975

Identifying the Parameters of an Aquifer Cell Model

E. HEFEZ,[1] U. SHAMIR, AND J. BEAR

Department of Civil Engineering, Technion-Israel Institute of Technology, Haifa, Israel

Cell models are commonly used for forecasting water levels in aquifers. Calibration of such models is achieved through identification of their parameter values, the transmissivities and storativities of all cells, using historical data. Several methods of formulating the identification as a linear or quadratic programing problem are presented. Examples are given, results of the various methods compared, and the sensitivity of these results to errors in the data is discussed. Inflows or outflows during historical periods may also be determined by the same methods; an example for a real aquifer is presented.

INTRODUCTION

The problem of identifying the parameters of an aquifer model is known also as the inverse problem. Here identification of parameters means the determination of the distributions of storativity $S = S(x, y)$ and transmissivity $T = T(x, y)$ of an aquifer in which we assume two-dimensional flow in the horizontal plane. The complete distributions are required if one wishes to forecast the future regime in the aquifer in response to various imposed activities of pumping and recharge.

An indirect way of obtaining the sought distributions is by some trial and error technique of adjusting the various parameters until an acceptable agreement is reached between the response of the model and that of the aquifer itself under some specified operation regime. A prerequisite for the application of this approach is the availability of hydrological data (e.g., water levels and rates of pumping, discharge of springs, artificial recharge, and natural replenishment) for some period in the past called the identification or calibration period. Sometimes, certain parts of the historical hydrological data mentioned above are not known. In this case we may regard the missing information items also as unknown parameters, the values of which have to be determined during the identification procedure. The natural replenishment of an aquifer is often considered such an unknown parameter. Unless otherwise specified, we shall consider henceforth the problem of identifying the distributions of S and T only.

The main disadvantage of a trial and error technique is that it does not involve an algorithm for seeking the solution systematically. In recent years, advanced mathematical methods have been developed and implemented for determining unknown model parameters. Among these one may mention the works of *Deininger* [1969], *Vemuri and Karplus* [1969], *Korganoff* [1970], *Coats et al.* [1970], *Emsellem and de Marsily* [1971], *Kleinecke* [1971], and *Neuman* [1973]. The present work is another attempt in this direction.

In this work the aquifer is represented by a finite difference model, and linear and quadratic programing procedures are employed as tools for identifying its parameters. The proposed methods have been tested on synthetic models, the parameters of which were a priori known. This technique made it possible to check results and compare the different methods. The application of the proposed techniques to cases of practical interest is now under way and will be reported separately.

[1] Now with the Department of Computer Sciences, Haifa University, Haifa, Israel.

FINITE DIFFERENCE MODEL

Various numerical models of groundwater systems may be employed. Essentially, the models considered in this paper deal with flow in a confined aquifer or in a phreatic one in which the spatial variations of the water table are small with respect to the thickness of the aquifer. It is assumed that the aquifer is isotropic and that the flow in it is essentially two-dimensional in the horizontal (x, y) plane.

The continuity equation for this model is

$$\frac{\partial}{\partial x}\left(T\frac{\partial \phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(T\frac{\partial \phi}{\partial y}\right) + r - p = S\frac{\partial \phi}{\partial t} \quad (1)$$

where $T = T(x, y)$ is the transmissivity of the aquifer, $S = S(x, y)$ is the storativity of the aquifer, $\phi = \phi(x, y, t)$ indicates the elevation of the water table or the piezometric surface, $r = r(x, y, t)$ represents inflow per unit (horizontal) area (e.g., by artificial and natural replenishment), and $p = p(x, y, t)$ represents withdrawal from the aquifer per unit (horizontal) area (e.g., by pumping).

The flow domain is subdivided into a network of rectangular cells (Figure 1) which serve in constructing the finite difference representation of (1).

In the cell model the continuous variables transform into discrete ones in the following way:

$$x \to x_i \quad i = 1, 2, \cdots, I \quad (2)$$

$$y \to y_j \quad j = 1, 2, \cdots, J \quad (3)$$

$$t \to t^n \quad n = 0, 1, 2, \cdots, N \quad (4)$$

$$\phi(x, y, t) \to \phi_{i,j}^n \quad (5)$$

$$p(x, y, t) \to p_{i,j}^{n+1/2} = (p_{i,j}^{n+1} + p_{i,j}^n)/2 \quad (6)$$

$$r(x, y, t) \to r_{i,j}^{n+1/2} = (r_{i,j}^{n+1} + r_{i,j}^n)/2 \quad (7)$$

$$T(x, y) \to T_{i,j} \quad (8)$$

$$S(x, y) \to S_{i,j} \quad (9)$$

In Figure 1 the centers of the cells are indicated in order to emphasize the fact that any property of the cell is represented by a single value which is assigned to its center. Note that the value of the transmissivity is also assigned to the centers of the cells and not to the border lines between adjacent cells, as is sometimes done. The number of transmissivity values, which are to be identified, is equal to the number of cells, whereas assignment of values to cell boundaries would result in roughly twice this number.

Two finite difference schemes, commonly used for

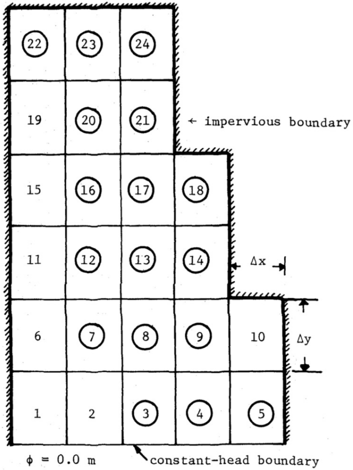
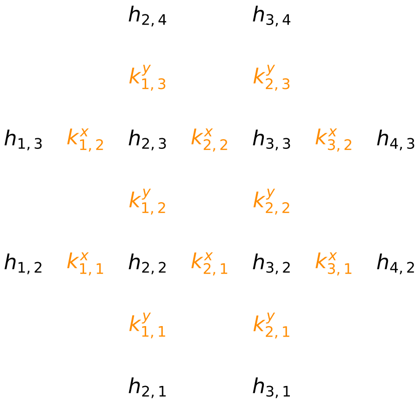


Fig. 3. The model studied in example 1.

Can we solve a 2D problem on the D-Wave? Yes.

2D groundwater flow equation



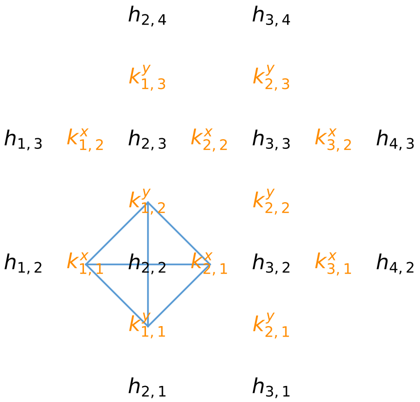
2D finite difference equation

$$\begin{aligned} 0 = & k^y_{1,1}(h_{2,1} - h_{2,2}) + k^y_{1,2}(h_{2,3} - h_{2,2}) \\ & + k^x_{1,1}(h_{1,2} - h_{2,2}) + k^x_{2,1}(h_{3,2} - h_{2,2}) \end{aligned}$$

Reformulate as a least squares problem

$$\begin{aligned} & [k^y_{1,1}(h_{2,1} - h_{2,2}) + k^y_{1,2}(h_{2,3} - h_{2,2}) \\ & + k^x_{1,1}(h_{1,2} - h_{2,2}) + k^x_{2,1}(h_{3,2} - h_{2,2})]^2 \end{aligned}$$

2D groundwater flow equation



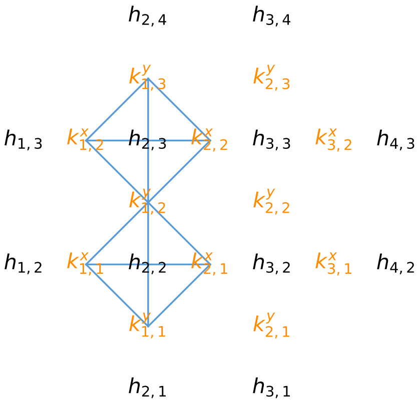
2D finite difference equation

$$\begin{aligned} 0 = &k^y_{1,1}(h_{2,1} - h_{2,2}) + k^y_{1,2}(h_{2,3} - h_{2,2}) \\ &+ k^x_{1,1}(h_{1,2} - h_{2,2}) + k^x_{2,1}(h_{3,2} - h_{2,2}) \end{aligned}$$

Reformulate as a least squares problem

$$\begin{aligned} &[k^y_{1,1}(h_{2,1} - h_{2,2}) + k^y_{1,2}(h_{2,3} - h_{2,2}) \\ &+ k^x_{1,1}(h_{1,2} - h_{2,2}) + k^x_{2,1}(h_{3,2} - h_{2,2})]^2 \end{aligned}$$

2D groundwater flow equation



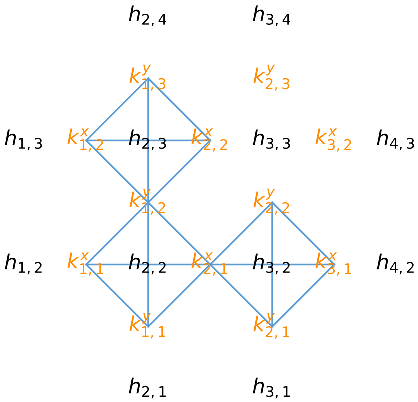
2D finite difference equation

$$\begin{aligned} 0 =& k^y_{1,1}(h_{2,1} - h_{2,2}) + k^y_{1,2}(h_{2,3} - h_{2,2}) \\ &+ k^x_{1,1}(h_{1,2} - h_{2,2}) + k^x_{2,1}(h_{3,2} - h_{2,2}) \end{aligned}$$

Reformulate as a least squares problem

$$\begin{aligned} &[k^y_{1,1}(h_{2,1} - h_{2,2}) + k^y_{1,2}(h_{2,3} - h_{2,2}) \\ &+ k^x_{1,1}(h_{1,2} - h_{2,2}) + k^x_{2,1}(h_{3,2} - h_{2,2})]^2 \end{aligned}$$

2D groundwater flow equation



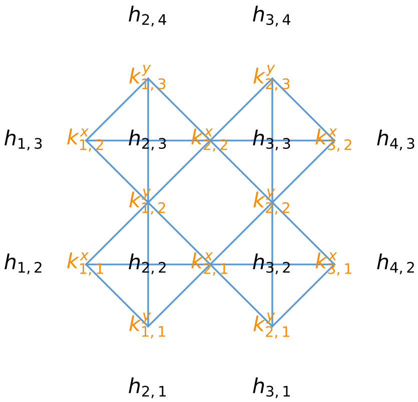
2D finite difference equation

$$\begin{aligned} 0 = & k^y_{1,1}(h_{2,1} - h_{2,2}) + k^y_{1,2}(h_{2,3} - h_{2,2}) \\ & + k^x_{1,1}(h_{1,2} - h_{2,2}) + k^x_{2,1}(h_{3,2} - h_{2,2}) \end{aligned}$$

Reformulate as a least squares problem

$$\begin{aligned} & [k^y_{1,1}(h_{2,1} - h_{2,2}) + k^y_{1,2}(h_{2,3} - h_{2,2}) \\ & + k^x_{1,1}(h_{1,2} - h_{2,2}) + k^x_{2,1}(h_{3,2} - h_{2,2})]^2 \end{aligned}$$

2D groundwater flow equation



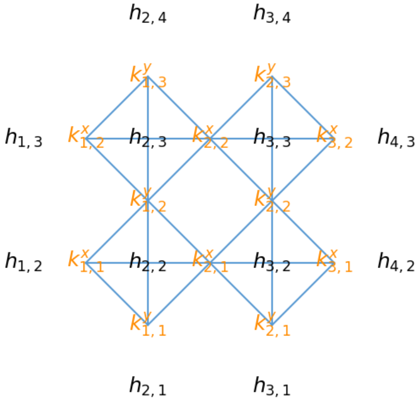
2D finite difference equation

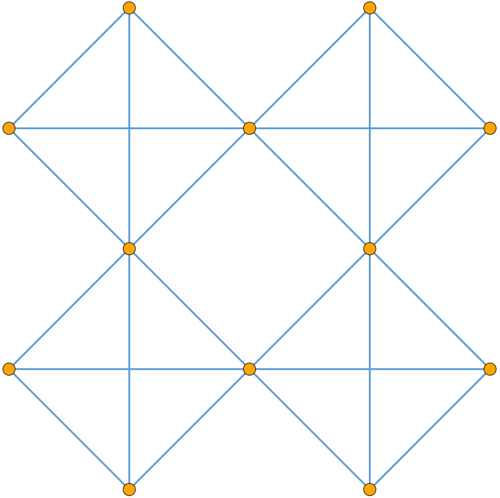
$$\begin{aligned} 0 =& k^y_{1,1}(h_{2,1} - h_{2,2}) + k^y_{1,2}(h_{2,3} - h_{2,2}) \\ &+ k^x_{1,1}(h_{1,2} - h_{2,2}) + k^x_{2,1}(h_{3,2} - h_{2,2}) \end{aligned}$$

Reformulate as a least squares problem

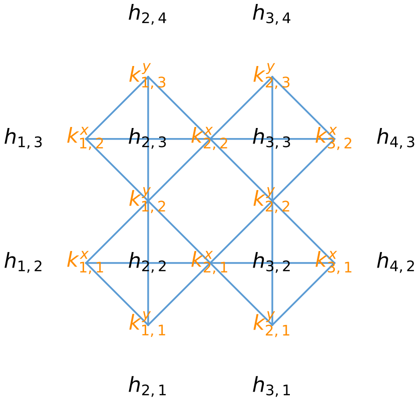
$$\begin{aligned} &[k^y_{1,1}(h_{2,1} - h_{2,2}) + k^y_{1,2}(h_{2,3} - h_{2,2}) \\ &+ k^x_{1,1}(h_{1,2} - h_{2,2}) + k^x_{2,1}(h_{3,2} - h_{2,2})]^2 \end{aligned}$$

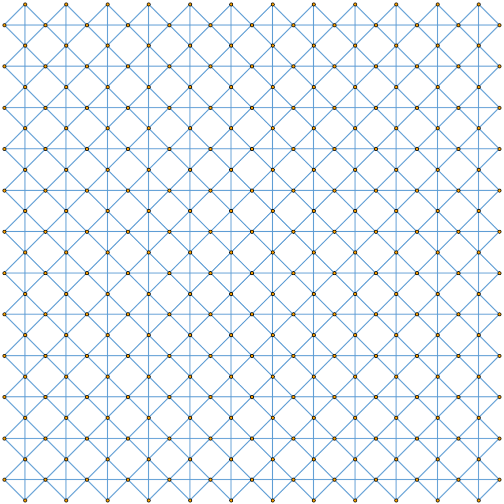
2D finite difference grid and graph



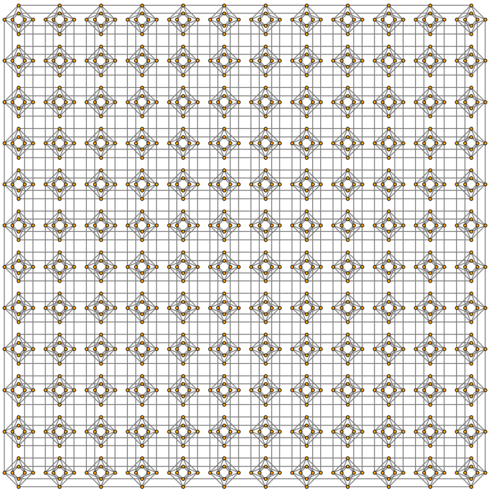


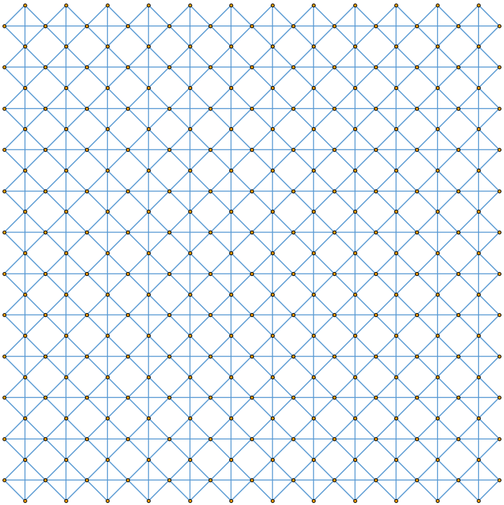
A bigger 2D finite difference graph



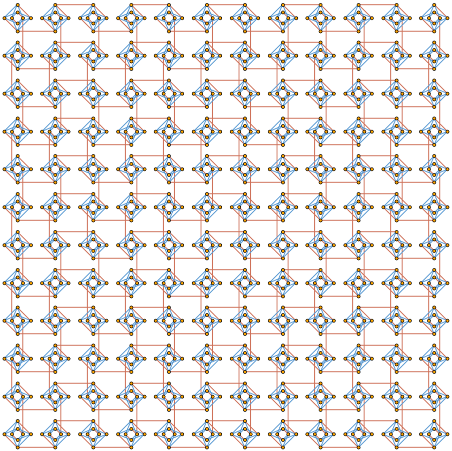


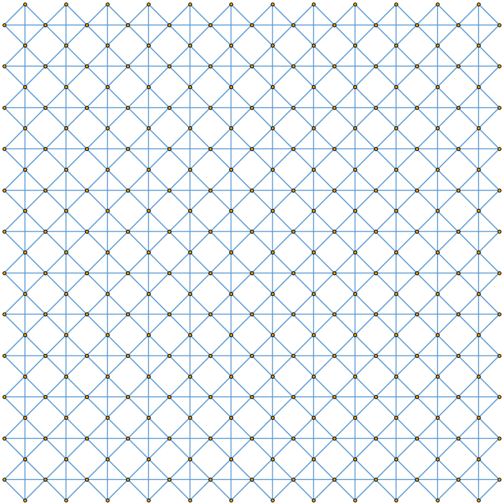
D-Wave graph vs. 2D finite difference graph



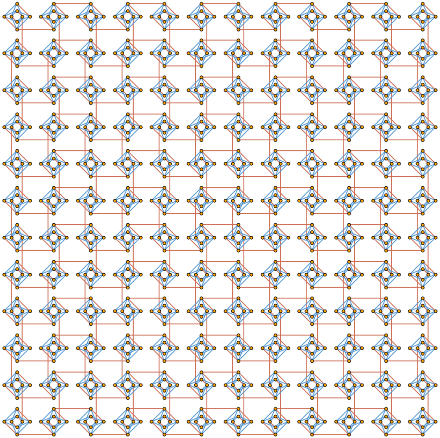


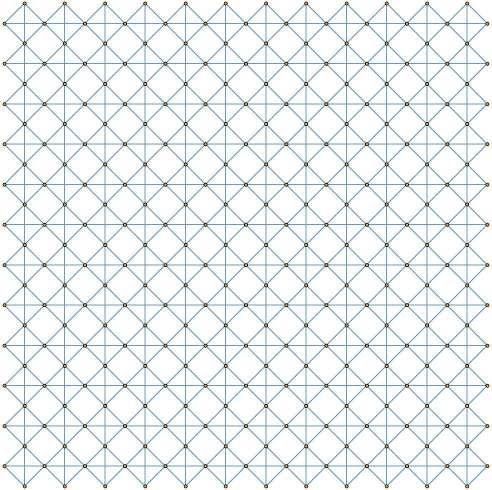
D-Wave graph→2D finite difference graph (embedding)





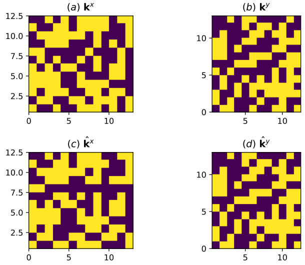
D-Wave graph→2D finite difference graph (embedding)

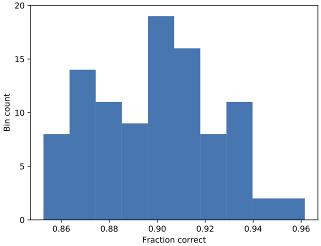




What does D-Wave have to say?

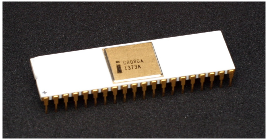
- If you take 10,000 samples from the virtual full yield solver
  - D-Wave gets $\mathbf{k}^y$ correct everywhere
  - D-Wave gets $\mathbf{k}^x$ correct in $\sim$90% of the locations
- Why is it better at $\mathbf{k}^y$ than $\mathbf{k}^x$?
  - $\mathbf{k}^y$ is aligned with the large-scale pressure gradient, and $\mathbf{k}^x$ is perpendicular to it
  - QUBO coefficients associated with $\mathbf{k}^y$ tend to be larger than those associated with $\mathbf{k}^x$





Is this problem big?

Intel 8080



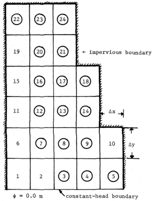
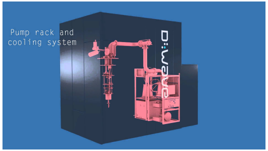
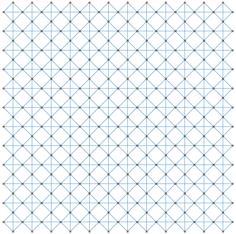


Fig. 3. The model studied in example 1.

24 nodes
48 parameters

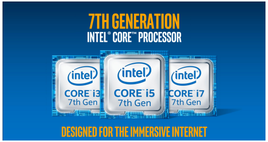
D-Wave 2X

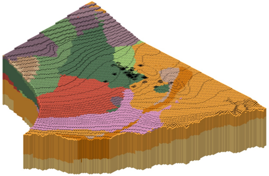




196 nodes
312 parameters

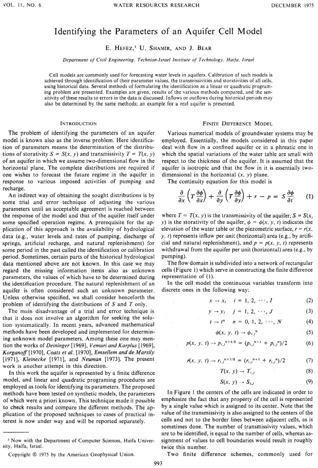
Modern CPUs





766,283 nodes
252 parameters

Computational hydrology: a historical perspective



VOL. 11, NO. 6 WATER RESOURCES RESEARCH DECEMBER 1975

Identifying the Parameters of an Aquifer Cell Model

E. HEFEZ,[1] U. SHAMIR, AND J. BEAR

Department of Civil Engineering, Technion-Israel Institute of Technology, Haifa, Israel

Cell models are commonly used for forecasting water levels in aquifers. Calibration of such models is achieved through identification of their parameter values, the transmissivities and storativities of all cells, using historical data. Several methods of formulating the identification as a linear or quadratic programing problem are presented. Examples are given, results of the various methods compared, and the sensitivity of these results to errors in the data is discussed. Inflows or outflows during historical periods may also be determined by the same methods; an example for a real aquifer is presented.

INTRODUCTION

The problem of identifying the parameters of an aquifer model is known also as the inverse problem. Here identification of parameters means the determination of the distributions of storativity $S = S(x, y)$ and transmissivity $T = T(x, y)$ of an aquifer in which we assume two-dimensional flow in the horizontal plane. The complete distributions are required if one wishes to forecast the future regime in the aquifer in response to various imposed activities of pumping and recharge.

An indirect way of obtaining the sought distributions is by some trial and error technique of adjusting the various parameters until an acceptable agreement is reached between the response of the model and that of the aquifer itself under some specified operation regime. A prerequisite for the application of this approach is the availability of hydrological data (e.g., water levels and rates of pumping, discharge of springs, artificial recharge, and natural replenishment) for some period in the past called the identification or calibration period. Sometimes, certain parts of the historical hydrological data mentioned above are not known. In this case we may regard the missing information items also as unknown parameters, the values of which have to be determined during the identification procedure. The natural replenishment of an aquifer is often considered such an unknown parameter. Unless otherwise specified, we shall consider henceforth the problem of identifying the distributions of S and T only.

The main disadvantage of a trial and error technique is that it does not involve an algorithm for seeking the solution systematically. In recent years, advanced mathematical methods have been developed and implemented for determining unknown model parameters. Among these one may mention the works of *Deininger* [1969], *Vemuri and Karplus* [1969], *Korganoff* [1970], *Coats et al.* [1970], *Emsellem and de Marsily* [1971], *Kleinecke* [1971], and *Neuman* [1973]. The present work is another attempt in this direction.

In this work the aquifer is represented by a finite difference model, and linear and quadratic programing procedures are employed as tools for identifying its parameters. The proposed methods have been tested on synthetic models, the parameters of which were a priori known. This technique made it possible to check results and compare the different methods. The application of the proposed techniques to cases of practical interest is now under way and will be reported separately.

[1] Now with the Department of Computer Sciences, Haifa University, Haifa, Israel.

FINITE DIFFERENCE MODEL

Various numerical models of groundwater systems may be employed. Essentially, the models considered in this paper deal with flow in a confined aquifer or in a phreatic one in which the spatial variations of the water table are small with respect to the thickness of the aquifer. It is assumed that the aquifer is isotropic and that the flow in it is essentially two-dimensional in the horizontal (x, y) plane.

The continuity equation for this model is

$$\frac{\partial}{\partial x}\left(T\frac{\partial \phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(T\frac{\partial \phi}{\partial y}\right) + r - p = S\frac{\partial \phi}{\partial t} \qquad (1)$$

where $T = T(x, y)$ is the transmissivity of the aquifer, $S = S(x, y)$ is the storativity of the aquifer, $\phi = \phi(x, y, t)$ indicates the elevation of the water table or the piezometric surface, $r = r(x, y, t)$ represents inflow per unit (horizontal) area (e.g., by artificial and natural replenishment), and $p = p(x, y, t)$ represents withdrawal from the aquifer per unit (horizontal) area (e.g., by pumping).

The flow domain is subdivided into a network of rectangular cells (Figure 1) which serve in constructing the finite difference representation of (1).

In the cell model the continuous variables transform into discrete ones in the following way:

$$x \to x_i \quad i = 1, 2, \cdots, I \qquad (2)$$

$$y \to y_j \quad j = 1, 2, \cdots, J \qquad (3)$$

$$t \to t^n \quad n = 0, 1, 2, \cdots, N \qquad (4)$$

$$\phi(x, y, t) \to \phi_{i,j}^n \qquad (5)$$

$$p(x, y, t) \to p_{i,j}^{n+1/2} = (p_{i,j}^{n+1} + p_{i,j}^n)/2 \qquad (6)$$

$$r(x, y, t) \to r_{i,j}^{n+1/2} = (r_{i,j}^{n+1} + r_{i,j}^n)/2 \qquad (7)$$

$$T(x, y) \to T_{i,j} \qquad (8)$$

$$S(x, y) \to S_{i,j} \qquad (9)$$

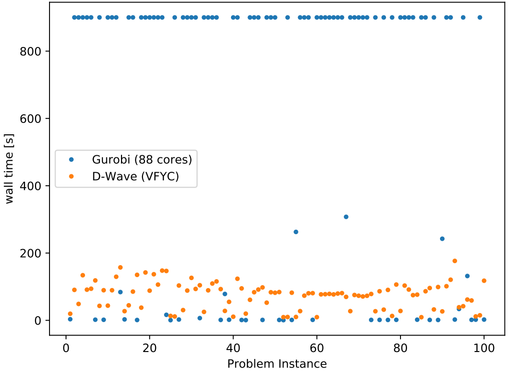
In Figure 1 the centers of the cells are indicated in order to emphasize the fact that any property of the cell is represented by a single value which is assigned to its center. Note that the value of the transmissivity is also assigned to the centers of the cells and not to the border lines between adjacent cells, as is sometimes done. The number of transmissivity values, which are to be identified, is equal to the number of cells, whereas assignment of values to cell boundaries would result in roughly twice this number.

Two finite difference schemes, commonly used for

"The identification problem as stated in the present work is solved as a linear or a quadratic programing problem. **The solution in the latter case is much more complicated, whereas the solution of the linear programing problem is based on readily available computer programs.**"

GUROBI OPTIMIZATION

D-Wave vs. Gurobi in a time-to-target benchmark



- D-Wave sets a solution quality target. How long does it take Gurobi to match or beat it?
- Gurobi is a state-of-the-art mathematical programming solver that can solve BQPs/QUBOs
- Gurobi lost the race in 69/100 cases and hit the 15 minute time limit in 64/100 cases

In the instance shown previously, we relaxed the 15 minute time limit. Gurobi exhausted the 256GB of memory on the machine after $\sim$ 4 hours without matching the target. We reran Gurobi in a mode that uses less memory for 24 hours and it failed to match the target set by the D-Wave.

Conclusions

- The D-Wave can be used to solve hydrologic inverse problems
- We solved problems with D-Wave's 3^{rd} generation chip that are large compared to what Hefez et al solved with Intel's 3^{rd} generation chip
- In many instances of the 2D problem we solved, the D-Wave outperformed a state-of-the-art classical tool whose use is consistent with the motivations of Hefez et al
- There is still a ways to go before practical applications to hydrology can be made
  - Both in terms of methods and hardware improvements
- O'Malley, D. (2018). An approach to quantum-computational hydrologic inverse analysis. *Scientific Reports*, 8(1), 6919.