PN-AAZ-782
5-174

TEL AVIV UNIVERSITY

FACULTY OF ENGINEERING
Department of Fluid Mechanics and Heat Transfer
P.O.Box 39040, Ramat Aviv, Tel Aviv 69978, ISRAEL
Telephone: (03)420930 Telex:342171 VERSY IL. Fax:972-3-419513

AID PROJECT DPE-G-SS-6006-00
"IDENTIFICATION OF AQUIFER HYDRAULIC PROPERTIES UNDER UNCERTAINTY"

THIRD PROGRESS REPORT
FOR THE PERIOD JULY 1 - DEC. 31, 1987

1. Activity of the Israeli team.

During this period the Israeli team has carried out the folowing tasks, along the project plan:

(i) A theoretical approach to identify the aquifer properties and the natural recharge under unsteady conditions has been developed. The method is based on a few simplifying assumptions and it is much simpler than existing stochastic approaches. The material forms part of an article we plan to publish in Water Resources Research. Enclosed please find the draft of this part of the projected article.

(ii) The method is being applied to a portion of the Israel Coastal Aquifer. The data have been collected in the preceding period and we are in the process of identifying parameters and recharge for this aquifer. This task should be completed within the next period.

2. Activity of the Portugaese team (reported by Drs. Lobo Ferreira and J. Rodriguez).

During the 6 month period starting August 1,1987 the Portuguese team worked in the following areas:

(i) Collection and analysis of piezometric and transmissivity data of the Rio Maior aquifer and the cretaceous and deep quaternary aquifers of Aveiro:

(a) Rio Maior aquifer: the latest collected 12 month sequential head measurement data of the Rio Maior aquifer have been analysed during this semester. This will allows for the future evaluation of the recharge of this aquifer, based on the transient analysis under development by the Israeli team.

(b) Cretaceous aquifer of Aveiro: additional transmissivity data has been analyzed during this period. The transmissivity data available for this aquifer are considered to be satisfactory. However, unlike what was expected during the first semester of 1987, it was not possible to be carry out the evaluation of the piezometry

Rec'd in SCI FEB 11 1988

of this aquifer . This is mainly due to the continuous extraction of groundwater by pumping wells owned by private industries and by the municipalities of the region. The temporary stopping of those extractions is necessary for the recovery of the head levels of the aquifer. Due to the lack of appropriate legislation it seems unprobable that this problem will be overcome in the near future. The project team will nevertheless establish contacts with the private industries and with the municipalities trying to make them aquainted with the importance fo the problem.

c) Deep quaternary aquifer of Aveiro: also for this aquifer additional transmissivity data have been analyzed during this semester. The transmissivity data available for this aquifer is considered also satisfactory. Additional pizometric measurements have been collected following the reclamation of several piezometers, suggested by the project team. Although the number of piezometric measurements is still not abundant, the available data allows for the application of the theory developed by the Israeli team to this aquifer.

(ii) Dr. Delgado Rodrigues from LNEC, in collaboration with Dr. Barradas and Dr. Marques da Silva from the University of Aveiro and Dr. Peixinho from the Directorate General for Natural Resources started the description of the geologic and hydrogeologic characteristics of the referred to aquifers.

3. Joint activity.

During this period Mr. Yoram Rubin stayed for a week in LNEC, working with Dr. Lobo Ferreira in the updating of the programs implemented in LNEC's computer system, and also in a program for the drawing of contours of numerical model outputs. During this visit data collected in Portugal and further implementation of the theoretical approach have been elaboarated.

Gedeon Dagan
Gedeon Dagan
Professor of Fluid Mechanics

STOCHASTIC IDENTIFICATION OF RECHARGE, TRANSMISSIVITY AND STORATIVITY IN AQUIFER UNSTEADY FLOW: 1. THEORY

by

GEDEON DAGAN and YORAM RUBIN

Dept. of Fluid Mechanics and Heat Transfer, Faculty of Engineering

Tel-Aviv University, Ramat-Aviv, Tel-Aviv, 69978, Israel

Abstract

In a sequence of two papers, a stochastic method to identify aquifer natural recharge, storativity and transmissivity under transient conditions is developed. Four main assumptions were adopted: Y, the logtransmissivity, is a normal random space function, the aquifer is unbounded, a first-order approximation of the flow equation is adopted, and the transients are slowly varying. Based on these assumptions, the expected value of Y and of the head H, as well as their covariances and crosscovariances, are expressed by analytical equations which depend on a parameters vector $\underline{\theta}$. A major part of the first paper is devoted to the development of these expressions, based on the two-dimensional flow equation. The proposed solution of the inverse problem is a double-stage procedure. First, $\underline{\theta}$ is identified stochastically, by a maximum-likelihood procedure applied to the measurements of Y and H. Then, $\underline{\theta}$ serves to estimate the spatial distributions of Y and H through their conditional mean and variances of estimation. The three main new features of the approach are: the possibility to identify the spatial distributions of Y and H through their first two statistical moments based on transient head data and in the presence of pumping/recharching wells; the identification of the

storativity and the stochastic identification of natural recharge. Since the proposed method make use of the analytic solution of the flow equation, it saves the need of laborious numerical schemes. Part 2 of this sequence (Rubin and Dagan, this issue) illustrates the application of the method to a real-life case.

1. INTRODUCTION.

The present study is concerned with the identification of aquifer flow parameters, known also as the inverse problem. The broad literature on the subject and the various methodologies employed in the past have been reviewed recently by Yeh (1986). The present work belongs to the recently developed stochastic approach, which regards the parameters as random variables or random space functions. These have to be identified in terms of their statistical moments, i.e. expected values, intervals of confidence etc. The determinstic approach to the inverse problem is believed to be plagued by non-uniqueness and stability problems (Yeh,1986, Yakowitz and Duckstein, 1980) and these difficulties may be alleviated in the stochastic framework (see, for instance, the discussions in Neuman et al, 1979, and Dagan, 1985 and Chavent, 1974).

So far, the stochastic identification approach has been applied mainly to steady aquifer flow (Clifton and Neuman, 1982, Kitanidis and Vomvoris, 1983, Townley and Wilson, 1983, Dagan, 1985, Rubin and Dagan, 1987, De Marsily, 1984), and only recently and at a lesser extent to unsteady flow (Carrera and Neuman, 1986, Loaiciga and Marino, 1987). In the present study, which is a direct continuation of our previous work (Rubin And Dagan, 1987), we address the problem of time-dependent flow, as encountered in many aquifers undergoing exploitation. A similar study about the subject is that of Carrera and Neuman (1986), who have cast the identification problem in a numerical frame, whereas we attack it in an analytical form, resulting in major computational simplifications. Besides, there are other significant differences between the various assumptions underlying the two studies, which will be recalled in Section 8, after rendering the readers familiar with the present stochastic

approach.

It is emphasized that the identification of the natural recharge, regarded as a random variable, is one of our main aims, as it was in Rubin and Dagan (1987). This emphasis is somewhat different from the exclusive concern of many previous studies with transmissivity identification.

2. MATHEMATICAL STATEMENT OF THE PROBLEM.

We consider here two-dimensional aquifer flow at the regional scale, i.e. over a planar horizontal extent much larger than the thickness. The well-known governing equation satisfied by the water head $H_t(\mathbf{x},t)$ is as follows

$$\frac{\partial}{\partial x}\left(T\,\frac{\partial H_t}{\partial x}\right) + \frac{\partial}{\partial y}\left(T\,\frac{\partial H_t}{\partial y}\right) = -R + S\,\frac{\partial H_t}{\partial t} + \sum_{i=1}^{N_w} Q_i(t)\,\delta(\mathbf{x}-\mathbf{x}_i) \qquad \underline{x} \in D \qquad (1)$$

where $\mathbf{x}(x,y)$ are Cartesian coordinates in the flow domain D, t is the time, $T(x,y)$ is the transmissivity, S is storativity, R is effective recharge, Q_i is the discharge of a well located at $\mathbf{x}=\mathbf{x}_i$ and δ stands for the Dirac distribution.

Eq. (1) is generally supplemented by boundary and initial conditions for H_t, e.g. H_t or its derivatives are given on the boundaries and H_t is given at t=0. The solution of (1) for H_t, for given T, S, R, Q_i, as well as given boundary and initial conditions, is known as the direct problem and it is well documented in the literature.

In the present study, dealing with the inverse problem, the given information is contained in : (i) values of measured transmissivities $T_j=T(\mathbf{x}_j)$

at $j=1,\ldots,M$ points within D ; (ii) records of the head $H_{t,lk}= H_t(\mathbf{x}_l,t_k)$ at $l=M+1,\ldots,N$ points and at various $k=1,\ldots,P$ times ; (iii) $Q_{ik}=Q_i(t_k)$, records of the pumping or recharge rates of the well located at $\mathbf{x}_i$ ($i=1,\ldots,N_w$) at time t_k and (iv) boundary and initial conditions for H_t.

The identification problem is defined then as determining $T(\mathbf{x})$, R and S based on the above information and with H satisfying (1).

While in cases of continuous and careful monitoring, the information of (ii) and (iii) above is available on a monthly basis, as it happens for the field example discussed in Part 2 of this sequence, sometimes the well discharge records are sketchy. A simplified approach to cope with such situations is presented in Appendix B. In a similar vein, sometimes the boundary and initial conditions are poorly defined. As we shall show, in our approach their knowledge is not required explicitely.

Our solution of the general identification problem stated above is carried out with the aid of a few simplifying assumptions, which are examined in the following two sections.

3. BASIC ASSUMPTIONS.

Our basic assumptions, along the procedure outlined by Kitanidis and Vomvoris (1983) and followed by Dagan (1985) and Rubin and Dagan (1987), are as follows :

(i) T is regarded as a random space function of lognormal distribution, i.e. the values of $Y_m = \ln T_m$ at an arbitrary set of points $\mathbf{x}_m$, form a multivariate normal vector. The stationary unconditional joint p.d.f. is characterized completely by the expected value $\langle Y\rangle=m_Y$=const and the auto-covariance $C_Y(\mathbf{x}_m,\mathbf{x}_n) = \langle Y'(\mathbf{x}_m)\, Y'(\mathbf{x}_n)\rangle = C_Y(r'_{mn})$, where $r'=|\mathbf{r}'|$ is the distance

between two points. In particular we shall adopt here also the exponential covariance

$$C_Y(r) = w\,[1-H(r')] + \bar{\sigma}_Y^2 \exp(-r'/I_Y) \tag{2}$$

where w represents an uncorrelated component of Y (a nugget), H(r') is here the Heaviside step function, $\bar{\sigma}_Y{}^2$ is the variance of spatially correlated Y values and I_Y is the correlation scale. Hence, Y is represented in terms of four parameters.

(ii) We regard R as a function of time only, whereas S is taken as constant. These assumptions are justified by the lesser space variability of R and S, generally observed in the field (cf. Freeze, 1975 for S), and by the limited availability of data, which precludes the simultaneous determination of T, R and S as general space functions.

(iii) A first-order approximation is adopted for the flow equation (1). Before expressing it in mathematical terms, we split the total head field into two components

$$H_t(\mathbf{x},t) = H_w(\mathbf{x},t) + H(\mathbf{x},t) \tag{3}$$

which satisfy the following two equations replacing (1)

$$\nabla.(T\,\nabla H_w) = S\,\frac{\partial H_w}{\partial t} + \sum_{i=1}^{N_w} Q_i(t)\delta(\mathbf{x}-\mathbf{x}_i) \tag{4}$$

$$\nabla.(T\,\nabla H) = -R + S\,\frac{\partial H}{\partial t} \tag{5}$$

Thus, H_w is a field related exclusively to the pumping wells, whose expression is given in Appendix A, and which is regarded as determinsitic and given. Furthermore, H_w is assumed to satisfy homogeneous boundary conditions and is singular at the wells locations. In contrast, H is related to the effect of recharge and of boundary conditions, and its spatial variability, regarded as random, is caused mainly by that of Y, and furthermore, H is a continous function of **x**.

With $h(\mathbf{x},t) = H - \langle H \rangle$ the residual, a first-order approximation is adopted for H, resulting in the following two equations which replace (5)

$$\nabla^2 \langle H \rangle = - \frac{R}{T_G} + \frac{S}{T_G} \frac{\partial \langle H \rangle}{\partial t} \tag{6}$$

$$\nabla^2 h = \frac{S}{T_G} \frac{\partial h}{\partial t} - \nabla.(Y' \nabla \langle H \rangle) \tag{7}$$

where $T_G = \exp(m_Y)$ is the geometric mean, and R and S are taken as time dependent and fixed, respectively.

The rational behind the first-order approximation has been discussed in the literature. It is recalled that: it simplifies considerably the original equation of flow with variable, random, coefficients (1), since (6) and (7) are linear; in principle it is valid only for small σ_Y^2, but it has been shown (Dagan, 1985b) that it leads to accurate results for σ_Y^2 as large as unity and, as demonstrated by the analysis of Hoeksema and Kitanidis (1984), the last condition is met by many aquifers.

The alternative formulation, able to handle in principle large variances σ_Y^2, is the numerical one. The numerical simulations of Smith and Freeze (1979) and

Desbarats (1987) have shown, however, that serious difficulties are encountered in the case of large σ_Y^2, and the subject can be considered as still opened.

Under these circumstances, the problem becomes the one of identifying the four parameters characterizing Y, as well as R(t) and S, for the given measurements of Y_j (j=1,...,M) and of $H_{t,lk}$ (l=M+1,...,N ; k=1,...,P), such that (6) and (7) are satisfied.

A numerical solution of the inverse problem may be envisaged now. By discretizing the space coordinates in D and the time, a numerical solution of <H> and h satisfying (6) and (7), respectively, may be obtained in terms of the boundary conditions and the unknown parameters. Then the problem becomes one of best fit between calculated and measured Y and H_t. Such an approach has been suggested for the steady-state flow by Hoeksema and Kitanidis (1984). Although the problem becomes considerably simpler than the original one expressed by (1) due to linearization, it is still a formidable one, involving large computational volumes and development of suitable algorithms. In line with our previous work on steady flow, we choose to further simplify the problem, by adopting additional assumptions.

4. ADDITIONAL SIMPLIFYING ASSUMPTIONS.

(iv) The head drop s caused by the pumping wells is represented as follows

$$s(\mathbf{x},t) = \sum_{j=1}^{N_W} \frac{S}{4\pi T_j}\int_0^t Q_j(t')\exp\left[\frac{-r'^2_j S}{4T_j(t-t')}\right]\frac{dt'}{t-t'}; \quad r'_j=|\mathbf{x}_j-\mathbf{x}| \tag{8}$$

i.e. the generalization of the Theis elementary solutions for each well and for

time depending Q_i (Carslaw and Jaeger, Chap.10). To simplify calculations, we separate the pumping wells into two classes. The first one contains M_1 wells for which the transmissivity T_j ($j=1,\ldots,M_1$) has been measured. For them, T_j has been taken in (8) as constant and given, since most of the drawdown is assumed to take place in an area of radius smaller than I_Y, in which the spatial variability of Y can be neglected. The second class contains the wells in areas of unknown transmissivity, in which we adopt the effective value T_G for T_j. This simplifies considerably the computations and can be justified by the fact that for a large number of wells the errors cancel each other. In any case, the basic requirement of water mass balance is obeyed by (8). Further transformations of (8), accounting for the discrete measurement of Q_j in time and for the fact that H_t is generally measured after a short recovery period, are given in Appendix A.

(v) The solution for h (7) is derived for an unbounded domain, disregarding the boundary conditions. This assumption can be justified by the existence of the forcing function related to spatial variability in the right-hand-side of (7), which dictates the behaviour of h. Boundaries might have an influence upon the statistical moments of h for points close to the boundary, but we shall apply the solution outside such regions. It is reminded that in many cases boundaries locations and boundary conditions are poorly defined.

(vi) The expected value $\langle H \rangle$ is regarded as having a slowly varying trend in space and time, and we adopt for it an expression similar to the quadratic trend taken by Rubin and Dagan (1987) for steady flow, i.e.

$$\langle H(\mathbf{x},t)\rangle \doteq H_0(t) - \mathbf{J}_0(t)\cdot\mathbf{x} - \frac{1}{2}\, c_{ij}(t)\, \mathbf{x}_i\, \mathbf{x}_j \qquad (i,j=1,2) \tag{9}$$

where H_0 is the head H at the origin, J_0 is its gradient there and c_{ij} are three independent coefficients ($c_{12}=c_{21}$). The summation convention for repeated indices is adopted in (9) and in the sequel.

Eq. (9) does not satisfy exactly Eq. (6), unless R, J_0 and c_{ij} are time independent and H_0 is a linear function of time. Then, substitution of (9) into (6) leads to the simple relationship

$$c_{11}+c_{22} = \frac{R - S\ dH_0/dt}{T_G} \tag{10}$$

generalizing the similar one for steady flow (Rubin and Dagan, 1987, eq. 14). The additional freedom offered by the remaining independent coefficients of the quadratic trend is presumably taking into account the departure of the average flow from a uniform one, due to boundary conditions.

We shall adopt (10) as a valid approximation even if H_0 is not linear and R is time depending. This is accurate if these functions are slowly varying in time, such that the time derivatives of J_0 and c_{ij}, appearing in (6) by substitution of (9), are negligible.

Hence, we regard the average head component $\langle H \rangle$, associated directly with the recharge R and indirectly with the boundary conditions, as a sequence of steady states. This assumption simplifies considerably the identification process and can be validated a-posteriori, as shown in Section 6.

(vii) The residual h satisfies the Poisson equation

$$\nabla^2 h = - \nabla.(Y' \nabla\langle H\rangle) \quad \text{i.e.} \quad \nabla^2 h = \frac{\partial}{\partial x_i}[(J_{0,i} - c_{ij}\ x_j)\ Y'] \tag{11}$$

replacing (7). The neglect of the time derivative of h in (7) is valid on the same grounds of slow variation in time. Indeed, by solving the quasi-steady version (10) of (7), the time shows up in the solution throughout $\mathbf{J}_0$ and c_{ij}, whose derivatives with respect to time are supposed to yield terms which are negligible compared to the others. Again, this assumption can be validated a-posteriori, as shown in Sect.6.

As a matter of fact, a time dependent uncorrelated noise can be added to h, without altering the results for the covariances of the Y,h field, which are shown in Section 5.

Summarizing the last two sections, the assumptions are identical to the ones adopted in our previous studies of steady flow (Dagan, 1985; Rubin and Dagan, 1987). The additional assumption is of quasi-steadiness, i.e. the expected value $\langle H \rangle$ and the head residual h are regarded as a sequence of steady states, such that at each instant t_k, they are identical to the ones corresponding to R, H_0, $\mathbf{J}_0$ and c_{ij} at the same instant, provided that R is supplemented by the term $S dH_0/dt$ (see eq.10). Hence, the unconditional statistical structure of H depends on these functions of time, besides the four parameters characterizing Y.

As mentioned above, the quasi-steady approximation is exact in the case in which $\langle H \rangle$ is a linear function of t, in which case t appears only in the term H_0 of (9). If H_0 changes nonlinearly with time, a time scale characterizing its departure from linearity may be defined, for instance by $\tau \approx |(dH_0/dt)/(d^2H_0/dt^2)|$. The validity of the quasi-steady approximation has been examined theoretically by Dagan (1982). It was shown there that in two-dimensional flow the head variogram adapts to its steady-state expression if the following criterion is satisfied

$$\frac{\tau\, T_G}{S\, I_Y^2} >> \frac{5\, r'/I_Y}{8\, \ln(r'/I_Y)} \tag{12}$$

where r' is the distance between the two points in the plane and $r'>>I_Y$. As we shall show in Part 2 of this sequence, this condition has been satisfied for the field case we have analyzed. The extension of our analysis to cases of relatively rapid time variations of the head is a subject for future investigations.

5. THE UNCONDITIONAL Y, H STATISTICAL MOMENTS.

In the stochastic approach, the aquifer is regarded as a realization belonging to an ensemble of formations with the same statistical structure of the transmissivity. The variables of interest, i.e. H_t and Y, are random space functions and are defined completely by their various statistical moments. Since Y is normal, the same is true in the first-order approximation for the residual h, which satisfies the linear equation in Y' (11). Hence, the entire statistical structure of the bivariate normal variables $Y(\mathbf{x})$ and $H(\mathbf{x}',t)$ is exhausted by the expected values $\langle Y\rangle$, $\langle H\rangle$ and the covariances $C_Y(\mathbf{x},\mathbf{x}') = \langle Y'(\mathbf{x})Y'(\mathbf{x}')\rangle$, $C_{YH}(\mathbf{x},\mathbf{x}',t) = \langle Y(\mathbf{x})h(\mathbf{x}',t)$ and $\Gamma_H(\mathbf{x},\mathbf{x}',t,t') = (1/2)\langle[h(\mathbf{x},t)-h(\mathbf{x}',t')]^2\rangle$. By unconditional moments we refer to the ensemble of formations in which the constraints imposed by measurements at the measurement points are not taken into account. Furthermore, all the above moments are for given values of S and R(t), which are regarded as fixed in space.

Under the assumptions of Sects. 3 and 4, we are able to write the above moments in terms of a series of parameters, by following closely the

procedure of Rubin and Dagan (1987). Thus, $\langle Y \rangle = m_Y$ and the given $C_Y(r')$ (2) are expressed in terms of the four parameters

$$\theta_1 = m_Y \; ; \; \theta_2 = \bar{\sigma}_Y^2 \; ; \; \theta_3 = w \; ; \; \theta_4 = I_Y \tag{13}$$

It is emphasized that other forms of covariances, e.g. semi-spherical, could be selected, with a somewhat different parametrization.

The expected value $\langle H_t \rangle = H_w + \langle H \rangle$ is given by (A.2) and (9), respectively, in terms of the additional parameters

$$\begin{gathered} \theta_5 = S \; ; \; : \theta_6 = J_{o,x}(t) \; ; \; \theta_7 = J_{o,y}(t) \; ; \\ \theta_8 = c_{11}(t) \; ; \; \theta_9 = 2c_{12}(t) \; ; \; \theta_{10} = c_{22}(t) \; ; \; \theta_{11} = H_0(t) \end{gathered} \tag{14}$$

The remaining two covariances C_{YH} and Γ_H are derived by using the equation of flow (1) for h. The solution is given by

$$h(\mathbf{x},t) = -\int_D \{\nabla''.[Y'(\mathbf{x}'')J(\mathbf{x}'',t)]\} G(\mathbf{x},\mathbf{x}'')d\mathbf{x}'' \tag{15}$$

where $G(\mathbf{x},\mathbf{x}'') = -(1/2\pi) \ln|\mathbf{x}-\mathbf{x}''|$ is the Green function, and ∇'' denotes differentiation with respect to $\mathbf{x}''$. It is immediately seen that the expressions of C_{YH} and Γ_H can be obtained from (15) by the same procedure which has been used in Rubin and Dagan (1987) for steady flow, the only difference being that θ_j (j=6,...,11) are now time-dependent. Hence, we can transfer directly the appropriate expressions as follows

$$C_{YH}(\mathbf{x},\mathbf{x}';t) = \bar{\sigma}^2_Y I_Y\left[c^{(1)}_{YH} + c^{(2)}_{YH}\right] \tag{16}$$

where

$$c^{(1)}_{YH}(\mathbf{x},\mathbf{x}',t) = \frac{J_{o,i}(t)r_i}{r^2}[e^{-r}(1+r)-1] \tag{17}$$

and

$$c^{(2)}_{YH}(\mathbf{r},t) = I_Y c_{ij}(t)\left[\delta_{ij}L + \frac{r_i r_j}{r}\frac{dL}{dr}\right] \quad ; \quad i,j=1,2 \tag{18}$$

where

$$L=e^{-r}(1+\frac{3}{r}+\frac{3}{r^2}) - \frac{3}{r^2} \quad \text{and} \quad r=\frac{r'}{I_Y}$$

The two parts comprising C_{YH} display different features: $c^{(1)}_{YH}$ was shown in the study of Rubin and Dagan (1987) to be highly anisotropic, being equal to zero in a direction perpendicular to the direction of the mean flow, and is also antisymmetrical. On the other hand, $c^{(2)}_{YH}$ is isotropic and symmetrical. The head residuals variogram is given by:

$$\bar{\Gamma}_H(\mathbf{x},t,\mathbf{x}',t') = \bar{\sigma}^2_Y I^2_Y[\bar{\gamma}^{(1)} + \bar{\gamma}^{(2)}] \tag{19}$$

The overbar for Γ_H denotes an isotropic average, i.e. $\bar{\Gamma}_H=\frac{1}{2\pi}\int_0^{2\pi}\Gamma_H d\theta$. $\bar{\Gamma}_H$ can be employed in lieu of Γ_H if the head measurement points are quite evenly distributed in the plane. In (19):

$$\bar{\gamma}^{(1)} = \frac{1}{2}\left[J_{o,i}(\bar{\mathbf{x}},t)J_{o,j}(\bar{\mathbf{x}},t')\right]\left[-E_i(-r)+\ln(r) + e^{-r}-1+E\right] \quad ;i,j=1,2 \tag{20}$$

where E=0.5772 is Euler's constant and $\bar{\mathbf{x}}=\frac{1}{2}(\mathbf{x}+\mathbf{x}')$. Also:

$$\bar{\gamma}^{(2)} = I_Y^2\{\alpha_0 - e^{-r}(\alpha_1 r^2+\alpha_2 r+\alpha_3) - \alpha_4 r^2 - \alpha_5[\ln(r) - E_i(-r)]\} \tag{21}$$

$$\alpha_0 = (\frac{25}{32}\frac{-3}{8}E)(c_{11}c_{11}'+c_{22}c_{22}') + (\frac{27}{8}\frac{-9}{4}E)c_{12}c_{12}' - (\frac{29}{32}\frac{-3}{4}E)(c_{11}c_{22}'+c_{11}'c_{22})$$

$$\alpha_1 = \frac{1}{8}c_{12}c_{21}' + \frac{1}{32}(c_{11}c_{22}'+c_{11}'c_{22}) + \frac{3}{32}(c_{11}c_{11}'+c_{22}c_{22}')$$

$$\alpha_2 = \frac{13}{32}(c_{11}c_{11}'+c_{22}c_{22}') + \frac{9}{8}c_{12}c_{12}' - \frac{5}{32}(c_{11}c_{22}'+c_{11}'c_{22})$$

$$\alpha_3 = \frac{25}{32}(c_{11}c_{11}'+c_{22}c_{22}') + \frac{27}{8}c_{12}c_{12}' - \frac{29}{32}(c_{11}c_{22}'+c_{11}'c_{22})$$

$$\alpha_4 = \frac{-7}{64}(c_{11}c_{11}'+c_{22}c_{22}') - \frac{3}{64}(c_{11}c_{22}'+c_{11} \qquad - \frac{1}{8}c_{12}c_{12}'$$

$$\alpha_5 = \frac{3}{8}c_{11}c_{11}'+c_{22}c_{22}') + \frac{9}{4}c_{12}c_{12}' - \frac{3}{4}(c_{11}c_{22}'+c_{22}c_{11}')$$

where E_i denotes the exponential integral. In the above expressions, c_{ij} stands for $c_{ij}(t)$, while $c_{ij}'=c_{ij}'(t')$. For $t=t'$, $c_{ij}=c_{ij}'$, and the above expression of $\bar{\Gamma}_H(\mathbf{x},t,\mathbf{x}',t')$ reduces to the one given in eq.(20) of Rubin and Dagan (1987a). An in-depth discussion of the structure and characteristics of the steady-state version of Γ_H is given in the study of Rubin and Dagan (1987). Herein a few points: $\gamma^{(1)}$ is the arithmetic mean of two head variograms - for $\mathbf{r}$ parallel and perpendicular to the mean flow direction, respectively. Its asymptotic behavior for large r is logarithmic; in contrast, $\gamma^{(2)}$ is isotropic for $c_{ij}=a\delta_{ij}$ and behaves like r^2 for large r.

Summarizing this section, we have arrived at the expressions of the unconditional moments of Y and H in terms of a vector of parameters $\underline{\theta}$ (13) and (14). If the time is discretized at values t_k ($k=1,\ldots,P$), then the parameters become a matrix θ_{jk} ($j=1,\ldots,11$; $k=1,\ldots,P$), with the dependence on k suppressed for the first five parameters, i.e. for $j=1,\ldots,5$.

6. OUTLINE OF PARAMETER IDENTIFICATION (UNCONDITIONAL) PROCEDURE.

The procedure follows closely the one developed for steady flow. In the present case the input, measured, variables are : the wells discharges Q_{ik} ($i=1,\ldots,N_w; k=1,\ldots,P$) where i designates the location of the well, while k stands for the time t_k; measured logtransmissivities $Y_j=Y(\mathbf{x}_j)$ ($j=1,\ldots,M$) and measured heads $H_{t,lk}=H_t(\mathbf{x}_l,t_k)$ ($l=M+1,\ldots,N; k=1,\ldots,P$).

In the stage of parameter identification, the objective is to identify the matrix of parameters θ_{jk} (13, 14). This is achieved by the procedure suggested by Kitanidis and Vomvoris(1983) and Hoeksema and Kitanidis(1984) and followed by Rubin and Dagan (1987), namely by a maximum likelihood procedure (MLP) applied simultaneously to the multivariate normal vector Y_j and H_{jk}, whose dimension is $M + (N-M)\cdot P$. The estimates $\hat{\theta}_{jk}$ are obtained from the minimization of the likelihood function, defined as usual by

$$F=(2\pi)^{\frac{-1}{2}[M+(N-M)P]}\,|\underline{\underline{Q}}|^{\frac{-1}{2}}\exp\left[-\frac{1}{2}(Z-\langle Z\rangle)^T\underline{\underline{Q}}^{-1}(Z-\langle Z\rangle)\right] \tag{22}$$

Here, Z is the vector of observations, angle brackets denotes the expectation operator and $\underline{\underline{Q}}$ is the covariance matrix defined by $\underline{\underline{Q}}=\langle(Z-\langle Z\rangle)(Z-\langle Z\rangle)^T\rangle$ where superscript T denotes transpose.

The output of the MLP is an estimator $\hat{\underline{\theta}}$ of $\underline{\theta}$, and the error covariance matrix Σ of $\hat{\underline{\theta}}$. The matrix Σ, the inverse of the Fisher information matrix, is sometimes referred to as the Cramer-Rao lower bound, and is found in many

cases (cf. Schweppe, 1973) to establish a tight asymptotic estimator for the parameters (co-)variances.

With given Σ, the confidence limits of the parameters can be evaluated. Approximate confidence limits are calculated from the standard error of the parameters and the appropriate value of Student's t-distribution

$$Pr\left[\ \theta_i - \frac{t(\Sigma_{ii}^{1/2})}{N^{1/2}} \leq \hat{\theta}_i \leq \theta_i + \frac{t(\Sigma_{ii}^{1/2})}{N^{1/2}}\right] = \rho \tag{23}$$

where Σ_{ii} is the parameter variance and $t \equiv t_{v\rho}$ is the value of the t-distribution for confidence level ρ and v degrees of freedom. A tight limit for the parameters is indicative of a correct model parametrization, and hence of the suitability of the suggested approach to the problem at hand.

At the completion of the stage of model validation, the recharge can be calculated, by employing eq.(10). First we define Δt as the (presumably equal) time period between t_k and t_{k+1}. Assuming that the estimates $\hat{c}_{11,k}$, $\hat{c}_{22,k}$, $\hat{S}$, $\hat{m}_Y$ and $\hat{H}_{o,k}$ are independent normal variables, we get

$$\hat{R}_k = (\hat{c}_{11,k} + \hat{c}_{22,k})\exp(\hat{m}_Y + \sigma^2_{\hat{m}}/2) - \frac{\hat{S}}{\Delta t}\langle \hat{H}_{o,k+1} - \hat{H}_{o,k} \rangle \quad k=1,..,P \tag{24}$$

The estimation variance of $\hat{R}_k$ is given by

$$\sigma^2(\hat{R}_k) = \left[\hat{c}^2_{11,k} + \hat{c}^2_{22,k}\right] \exp(2\hat{m}_Y)\exp(\sigma^2_{\hat{m}})\left[\exp(\sigma^2_{\hat{m}}) - 1\right] \tag{25}$$

$$+(\sigma^2_{\hat{c}_{11,k}} + \sigma^2_{\hat{c}_{22,k}})\exp(2\hat{m}_Y)\exp(2\sigma^2_{\hat{m}}) + \frac{1}{\Delta t^2}\sigma^2_{\hat{S}}\, V\langle \hat{H}_{o,k+1} - \hat{H}_{o,k} \rangle$$

where $\sigma^2_{\hat{c}_{11},k}$, $\sigma^2_{\hat{c}_{22},k}$, $\sigma^2_{\hat{m}}$ and $\sigma^2_{\hat{S}}$ are the variances of estimation of $\hat{c}_{11}$, $\hat{c}_{22}$, $\hat{m}_Y$ and $\hat{S}$, respectively.

The time sequence of $\hat{R}_k$ $(k=1,..,P)$ can supply useful information. For example, multi-seasonal water balances offer the possibility to estimate the relation between precipitation and natural recharge. A simple model for that would be $R_k=\alpha(P_k-\beta)$ where α and β are empirical coefficients, R_k and P_k are recharge and precipitation at time t_k, respectively. A quick and simple way to estimate this model quantitatively is by some best-fit technique applied to the series of R_k and P_k.

7. THE CONDITIONING STAGE

The unconditional moments derived previously define stochastically an ensemble of realizations. Only in a subset of them Y and H assume the measured values at the measurements points. The conditional p.d.f. of Y and H is defined as the one based on this subset, and the conditioned Y and H are multivariate normal as well.

The method for computing the conditional moments of $Y(\mathbf{x})$ is given in the study of Rubin and Dagan (Section 3, 1987). It is similar to kriging and it consists of a linear function of the head and transmissivity residuals (with respect to the unconditional mean) at the measurement points. The coefficient for each residual is a function of its relative position with regard to $\mathbf{x}_i$. The coefficients are computed by solving a linear system of equations, which require the knowledge of the unconditional second moments. In general, conditioning must be carried out simultaneously for $\mathbf{Z}$, the vector of all measurements at all times. When the parameters of the quadratic trend (9) change slowly in time, the unconditional second moments which are expressed by

these parameters, are practically constant in time. Consequently, the coefficients are also invariant in time, and they can be computed for any subset $\mathbf{Z}_k$ of $\mathbf{Z}$, which contain all the transmissivity measurements and the head measurements at time t_k, and conditioning is carried out on $\mathbf{Z}_k$ rather then $\mathbf{Z}$. The conditional mean of Y at different times t_k can then be averaged, since they might differ slightly.

Finally, the computation of the conditional variance of Y is carried out according to the method outlined in the previuos study of Rubin and Dagan (1987), and it includes the effect of parameters' variances of estimation.

The computation of the conditional moments of H is very similar to that of Y. The general relations between the conditional and unconditional moments are given in Dagan (1984) and they can be applied directly by using the expressions derived here for the unconditional moments.

8. CONCLUDING REMARKS

In this study, an analytic-stochastic method to solve the inverse problem under transient conditions was developed. This method provides a simple and inexpensive mean to identify the transmissivity field, the storativity and the recharge, based on a limited number of measurements. The method can also be employed to obtain a stochastic solution to the direct problem, i.e. the head field.

After rendering the reader familiar with the principles and methodology, we compare now briefly the present method with the work of Carrera and Neuman (1986), which is aimed at solving the inverse problem in a probabilistic frame. Their solution is based on a numerical solution of the flow equation (1), and thus it provides in principle more generality and flexibility, while the

analytical method described herein is limited to cases of $\sigma^2_Y \leq 1$ (see Dagan, 1985b). A recent survey by Hoeksema and Kitanidis (1985) showed, nevertheless, that this requirement is met in many aquifers.

Another point to be made when comparing numerically-based solutions with analytical ones is that of computational cost and complexity. In order to avoid artificial zonation and to capture accurately the aquifer heterogeneity, the numerical modeler has to resort to a fine mesh as computational grid. This may result in huge matrices that are costly to store and manipulate. Furthermore, it was pointed out by Sun and Yeh (1985) that in case of scarcity of data, fine-meshing often leads to unstable solution. The present method avoids these pittfalls, by providing an analytical solution to the flow equation, which is much simpler to apply. Depending on the amount of data at hand, it need not be applied to the whole aquifer, but only to a subsection of it. The problem of data scarcity is reduced, since there are fewer parameters to identify.

Besides these general observations, there are some specific differences between the present method and that of Carrera and Neuman (1986). The first refers to the definition of the parameters and their stochastic structure: Carrera and Neuman (1986) identify the logtransmissivity covariance and head variogram from measurements without employing the flow equation. Furthermore, they neglect the cross-covariance between heads and logtransmissivity. At the later stage of MLP applied to the measured values, the same covariance and variogram are utilized while assuming that they are known with certainty, except for a constant of proportionality. In the present case, the parameters are defined as structured space random functions. The covariances and cross-covariances of the parameters are obtained by a linearization of the flow equation and they are identified by a MLP procedure, as well as their variance of estimation.

In Part 2 of the present study (Dagan and Rubin, this issue), the application of the method to a subsection of Israel's coastal aquifer is presented in detail.

APPENDIX A

DERIVATION OF H_w

In the case of constant Q_i, eq.8 reduces to the sum of Theis' solution for each well, i.e.

$$H_w(\mathbf{x},t)=H_{in}+\sum_{i=1}^{N_w}\frac{Q_i}{4\pi T_i}E_i\left(\frac{-r'^2(\mathbf{x},\mathbf{x}_i)S}{4T_i t}\right) \tag{A.1}$$

where $H_{in}=H$ is the initial head distribution at t=0, and the pumping discharges Q_i are taken constant over time. It is more common for Q_i to change over time in a stepwise fashion. The common practice in Israel's coastal aquifer, for example, is to monitor heads and well discharges on a monthly basis. This periodic variation can be accomodated into eq.(A.1) as follows. We define first $\bar{Q}_i=\frac{1}{N}\sum_{k=1}^{P} Q_{i,k}$, where $Q_{i,k}$ is the average pumpage between t_{k-1} and t_k, and P is the number of (presumably equal) periods between t_0 and $t=t_P$, the last denoting the time of measuring $H_t(\underline{x},t_P)$. We further define the periodic fluctuations by $\Delta Q_{i,k}=Q_{i,k}-\bar{Q}_i$. Due to the linearity of the solution for (8), H_w can now be written as a function of the avergae $\bar{Q}_i$ and its fluctuations as follows

$$H_w(\mathbf{x},t_P)>=H_{in}(\mathbf{x},0)-\sum_{i=1}^{N_w}\frac{\bar{Q}_i}{4\pi\tau_i}\ln\left[\frac{2.25\tau_i t}{r'^2(\mathbf{x},\mathbf{x}_i)\hat{S}}\right]$$

$$+\sum_{i=1}^{N_w}\sum_{k=1}^{P}\frac{\Delta Q_{i,k}}{4\pi\tau_i}\left[E_i(-\bar{U}_{i,k})-E_i(-\bar{U}_{i,k+1})\right] \tag{A.2}$$

$$+\alpha(t,t_p)\alpha'(\mathbf{x},\mathbf{x}_i)\frac{\bar{Q}_i}{4\pi\tau_i}\ln\left[\frac{2.25\tau_i t}{r^2_{w_i}\hat{S}}\right] \quad (i=1,..,N_w)$$

where

$$\bar{U}_{i,k+1}=\frac{r^2(\mathbf{x},\mathbf{x}_i)\hat{S}}{4\tau_i(t-t_{k+1})}\ ;\ \bar{U}_{i,k}=\frac{r^2(\mathbf{x},\mathbf{x}_i)\hat{S}}{4\tau_i(t-t_k)}$$

and

$$\tau_i=\begin{cases} e^{\langle Y\rangle} & 1\le i\le M_1 \\ T_i & M_1+1\le i\le N_w \end{cases}$$

where $\alpha(t,t_p)$ is equal to unity when $t>t_p$ and to zero otherwise. Similarly, $\alpha'(\mathbf{x},\mathbf{x}_i)$ is equal to unity when $x=\mathbf{x}_i$ and zero otherwise. The last term in (A.2) represents the case in which a short period of recovery is allowed before heads are measured at the pumping wells, i.e. heads are not measured at t_p, but rather at $t>t_p$. Hence, $t-t_p$ is the duration of the recovery period and r_{w_i} is the radius of the well located at $\mathbf{x}_i$. This is the common practice, for example, followed by the authorities monitoring Israel's coastal aquifer, where a recovery period of 24-48 hours is usually allowed. It should be noted, however, that the recovery period affects only the close vicinity of the well.

Following our discussion in Section 4, we made here a distinction between M_1 wells of unknown transmissivities, and the rest M_2 wells with known transmissivities.

APPENDIX B

A SIMPLIFIED MODEL OF WELLS INFLUENCE

A more simple-minded approach to solve the problem defined in Section 2 is to define an effective constant recharge which combines the point source/sink terms of eq.1 together with the natural recharge R, i.e.,

$$R_{eff} = R - \frac{1}{A} \sum_{i=1}^{N_w} Q_i(t) \qquad (B.1)$$

where A is the area of the aquifer. The mean flow field is given now by eq.(9) solely, and there is no need to solve for H_w, which was given in eq.(A.2). Note, however, that the analysis directed toward the identification of the natural recharge of Section 6 is not applicable now, since the parameters representing the mean flow field reflect now the joint effect of the recharge and that of the flow towards the wells.

The stochastic analysis which is based on the substitution of R_{eff} into (1) is limited to the following cases:

(1) When point pumpage/discharge is negligible relative to the natural recharge.

(2) When the aquifer is uniformly pumped, i.e. a large number of wells are scattered quite evenly all over the aquifer, and it is known that they are being operated with similar pumping regimes. A sufficient recovery period before the periodical measurement of the heads enhances the applicability of

the simplified analysis, since the local effect of the particular well is attenuated.

(3) When no pumpage/recharge record is available, and it is assumed that the situation described in (2) prevails.

Two benefits are obtained in this manner. The first is the simplification of the computation, as discussed above. The second benefit is in the possibility to obtain estimates for the Y-covariance model, as well as to identify the areal distribution of the heads and logtransmissivities, even in the absence of pumpage/discharge record.

The main limitation of this approach resides in the inability to identify the storativity S and the natural rechargh R, since both cannot be separated from the sum in (B.1).

ACKNOWLEDGEMENT. This study is part of a thesis to be submitted by Y. Rubin to Tel-Aviv University in partial fulfilment of the requirements for a Ph.D degree. Partial support of Aid - Agency of International Development, is acknowledged with gratitude.

REFERENCES

Carrera, J., and S. P. Neuman, Estimation of aquifer parameters under transient and steady state conditions: maximum likelihood method incorporating prior information, Water Resour. Res., 22(2), 199-210, 1986.

Carslaw, H. S., and J. C. Jaeger, Conduction of heat in solids, 2nd edition, Oxford university press, 1959.

Chavent, G., Identification of functional parameters in partial differential equations, in Identification of parameters in distributed systems, edited by Goodson, R. E., and M. Polis, pp. 31-48, American Society of Mechanical Engineers, New York, 1974.

Clifton P. M., and S. P. Neuman, Effects of kriging and inverse modeling on conditional simulation of the Avra Valley aquifer in Southern Arizona, Water Resour. Res., 18(4), 1215-1234, 1982.

Dagan, G., Stochastic modeling of groundwater flow by unconditional and conditional probabilities: the inverse problem, Water Resour. Res., 21(1), 65-72, 1985.

Dagan, G., A note on the higher- order corrections of the head covariances in steady aquifer flow, Water Resour. Res., 21(4), 573-578, 1985b.

Dagan, G., Analysis of flow through heterogeneous random aquifers: 2. unsteady

flow in confined formations, Water Resour. Res., 18(5), 1571-1585, 1982.

Dagan, G., Solute transport in heterogeneous porous formations, J. Fluid Mech., 145, 151-177, 1984.

De Marsily, G., G. Lavedau, M. Boucher, and G. Fasanino, Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model, in Geostatistics for Natural Resources Characterization, edited by G. Verly, M. David, A. G. Journel, and A. Marechal, pp. 831-849, D. Reidel, Hingham, Mass., 1984.

Desbarates, A. J., Numerical estimation of effective permeability in sand-shale formations, Water Resour. Res., 23(2), 273-286, 1987.

Freeze, R.A., A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media, Water Resour. Res., 11(5), 725-741, 1975.

Hoeksema, R.J., and P.K. Kitanidis, An application of the geostatistical approach to the inverse problem in two dimensional groundwater modeling. Water Resour. Res., 20(7), 1009-1020, 1984.

Hoeksema R. J., and P.K. Kitanidis, Analysis of the spatial structure of properties of selected aquifers, Water Resour. Res., 21(4), 563-572, 1985.

Kitanidis P. K., and E. G. Vomvoris, A geostatistical approach to the inverse problem in groundwater modeling (steady - state) and one dimensional simulations, Water Resour. Res., 19(3), 677-690, 1983.

Loaiciga, H. A., and M. A. Marino, The inverse problem for confined aquifer flow: Identification and estimation with extensions, Water Resour. Res., 23(1), 92-104, 1987.

Neuman, S. P., and S. Yakowitz, A statistical approach to the inverse problem of aquifer hydrology, 1. Theory, Water Resour. Res., 15(4), 845-860, 1979.

Rubin Y., and G. Dagan, Stochastic identification of transmissivity and effective recharge in steady groundwater flow: 1. Theory, Water Resour. Res., In press, 1987.

Smith, L., and R. A. Freeze, Stochastic analysis of steady state groundwater flow in bounded domain, 2, Two-dimensional simulations, Water resour. Res., 15(6), 1543-1559,1979.

Schweppe, F. C., Uncertain dynamic systems, Prentice-Hall, New-Jersy, 1973.

Sun, N. Z., and W. W-G. Yeh, Identifiability of parameter structure in groundwater inverse problem, Water Resour. Res., 21(6), 869-883, 1985.

Townley, R. L., and J. L. Wilson, Conditional second moment analysis of groundwater flow: the cumulative effects of transmissivity and head measurements, Papers of the international conference on groundwater and man, Vol.1, pp.321-330, Australian Government Publishing Service, Canberra, 1983.

Yakowitz, S., and L. Duckstein, Instability in aquifer identification: Theory and case study, Water Resour. Res., 16(6), 1045-1064, 1980.

Yeh, W. W-G., Review of parameter identification procedures in groundwater hydrology: the inverse problem, Water Resour. Res., 22(2), 95-108, 1986.