

UC Berkeley

Technical Completion Reports

Title

Optimal Determination of Stratified Groundwater Basin Characteristics

Permalink

<https://escholarship.org/uc/item/97v8p6dv>

Authors

Shephard, Ronald W
Todd, David K.

Publication Date

1970-07-01

G402
XU2-7

no. 264

TECHNICAL COMPLETION REPORT

on

Project W 264

OPTIMAL DETERMINATION OF STRATIFIED GROUNDWATER
BASIN CHARACTERISTICS

WATER RESOURCES CENTER
University of California

(Optimal Identification of Groundwater Reservoir Parameters
by Nonlinear Decomposition - John Labadie)

Submitted by

Ronald W. Shephard and David K. Todd
University of California

JULY 1970

WATER RESOURCES CENTER ARCHIVES

~~6/13~~

I. THE PROBLEM

In order to predict future water level response to well pumpage in large groundwater basins, aquifer characteristics must be known to a reasonable degree of accuracy. Optimal pumping policies can then be developed on the basis of the following criteria:

1. Meet all future water demand.
2. Minimize pumping costs (which are directly related to depth of water level or drawdown).
3. Guard against harmful "mining" of the groundwater resource, with land subsidence and irreversible loss of storage potential as inevitable consequences.

A partial differential equation has been derived which describes the unsteady response of water level or head (if groundwater is under pressure) to well pumpage (Davis and DeWiest [2]):

$$(1) \quad \frac{\partial^2 h}{\partial r^2} + \frac{1}{r} \frac{\partial h}{\partial r} = \frac{S}{T} \frac{\partial h}{\partial t}$$

$$\lim_{t \rightarrow 0} h(t, r) = \lim_{r \rightarrow \infty} h(t, r) = h_0$$

$$\lim_{r \rightarrow 0} r \left(\frac{\partial h}{\partial r} \right) = \frac{Q}{4\pi T}$$

where

$h(t, r)$ = water level or head (L)

t = time (T)

r = radial distance from the pumping well (L)

$Q(t)$ = pumping rate (L^3/T)

S = coefficient of storage (related to porosity)

T = transmissivity (related to permeability) (L^2/T) .

A good aquifer must have relatively high values of both S and T . For example, clay can have a higher porosity than sand, but is a poorer aquifer since it allows very little seepage. Considerable water may be in storage, but it cannot be pumped out.

The solution to Equation 1 is

$$(2) \quad h(t,r) = h_0 - \frac{Q}{4\pi T} \int_{\frac{Sr^2}{4Tt}}^{\infty} \frac{e^{-u}}{u} du$$

the integral can be expanded into a convergent series.

$$(3) \quad h(t,r) = h_0 - \frac{Q}{4\pi T} \left[-0.5772 - \ln u + u - \frac{u^2}{2 \cdot 2!} + \frac{u^3}{3 \cdot 3!} - \frac{u^4}{4 \cdot 4!} \dots \right]$$

where drawdown $d(t,r) = h_0 - h(t,r)$, and

$$(4) \quad u = \frac{Sr^2}{4Tt}$$

There are essentially two kinds of aquifers: unconfined and confined (where water is under pressure):

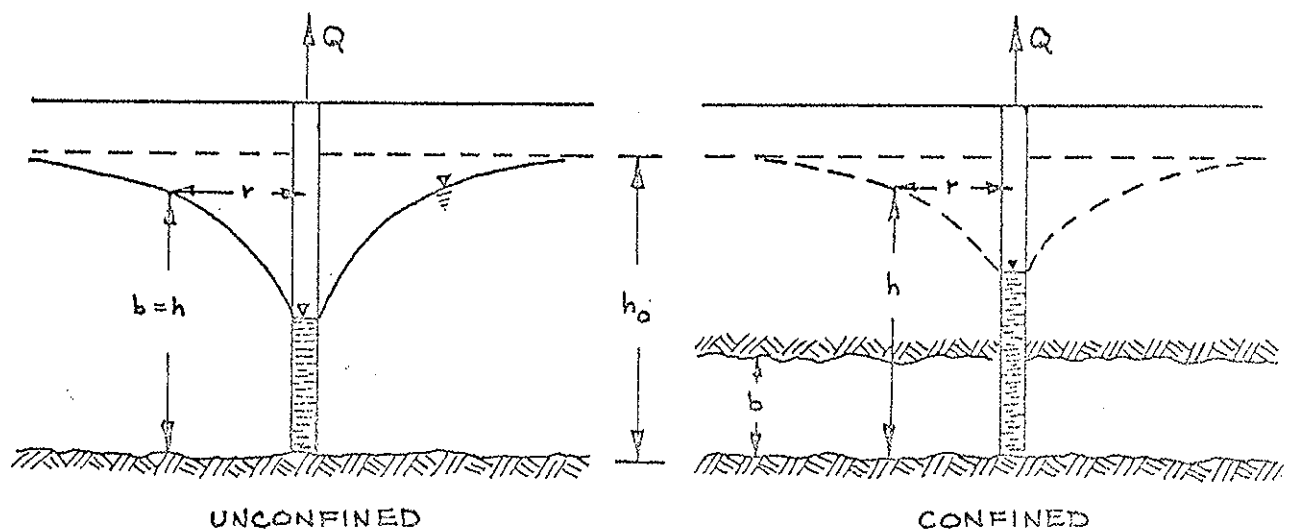


FIGURE 1: TYPES OF AQUIFERS

For confined aquifers, transmissivity T is relatively constant, whereas $T = Kh$ for unconfined aquifers. If the amount of drawdown is small, a reasonable approximation for T would be

$$(5) \quad \bar{T} = K\bar{h}$$

where

K = the constant coefficient of permeability (L/T)

\bar{h} = the average drawdown over a period of time (L) .

The important *assumptions* associated with equations 2 and 3 include (Todd [8]):

1. The aquifer is of infinite extent.
2. The aquifer is homogeneous and isotropic.
3. The well completely penetrates the aquifer.
4. Water is removed instantaneously from storage.

The usual method of determining or "identifying" aquifer parameters S and T is to perform a "pumping test." A constant flow of water is discharged during a short period of time, and the change in water level at a nearby observation well noted continuously or at discrete intervals. The parameters in Equation 3 are then varied until the closest possible fit is obtained between the computed head $h(t,r)$ (where r is the radial distance from the pumping well to the observation well) and the actual observed drop in head $\hat{h}(t,r)$.

The standard pumping test will generally achieve a much smaller drawdown than that observed during actual operation of the well. In many cases, the result is an inadequate representation of parameter values. The difficulty in using historical records compiled during actual operation of a well can be seen in Figure 2. Large drawdown will result in considerable interaction between adjacent pumping wells. The resulting water table is obtained by superposition of the uninfluenced "cones of depression" for each well.

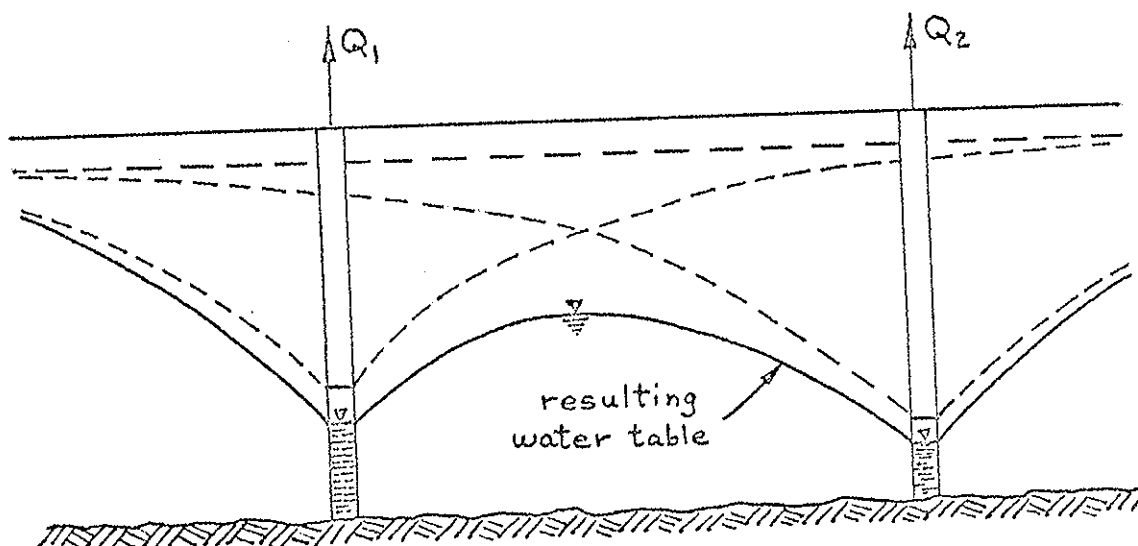


FIGURE 2: INTERACTION BETWEEN TWO PUMPING WELLS

When a large number of wells (say N) are involved, the overall identification problem can become unwieldy. Using the standard least-square criterion for closest fit, the problem is posed as follows (assuming $M + 1$ discrete observations):

$$(6) \quad \text{minimize } E = \min_{\substack{T_1, S_1 \\ \vdots \\ T_N, S_N}} \sum_{i=1}^N \sum_{k=0}^M \left[\sum_{j=1}^N [d_j(t_k, r_{ij})] - \hat{d}_i(t_k, r_{ii}) \right]^2$$

where $\hat{d}_i(t_k, r_{ii}) = h_0 - \hat{h}_i(t_k, r_{ii})$

where r_{ij} is the distance from well i to well j and r_{ii} is the radius of well i .

The parameters T_i and S_i now take on a slightly different interpretation. Assumption 2 stated that the aquifer must be homogeneous and isotropic. This assumption is obviously violated since there is little chance that all of the parameters computed from the minimization problem of Equation 6 will turn out to be

the same for a realistic problem. With violation of assumption 2, T_i and S_i lose some of their physical significance and are interpreted more as "black box" parameters whose primary function is the prediction of output (water level response) from input or excitation (well pumping). On the other hand, the result is an increase in realism since groundwater basins are generally nonhomogeneous and nonisotropic.

The parameters T_i and S_i are now interpreted as those values found from observations on well i which are uninfluenced by adjacent pumping. The fact that the actual observations $\hat{h}_i(t_k; r_{ii})$, $i = 1, \dots, N$, $k = 0, \dots, M$, are indeed influenced by adjacent pumping gives rise to the complication and high dimensionality of this minimization problem.

If N is very large, an algorithm for solving Equation 6 should take advantage of the fact that a particular well will only be directly influenced by the few wells surrounding it. As shown in Figure 3, historical records can verify that, with a high probability, beyond a certain range the uninfluenced cones of depression of adjacent wells will never reach well i .

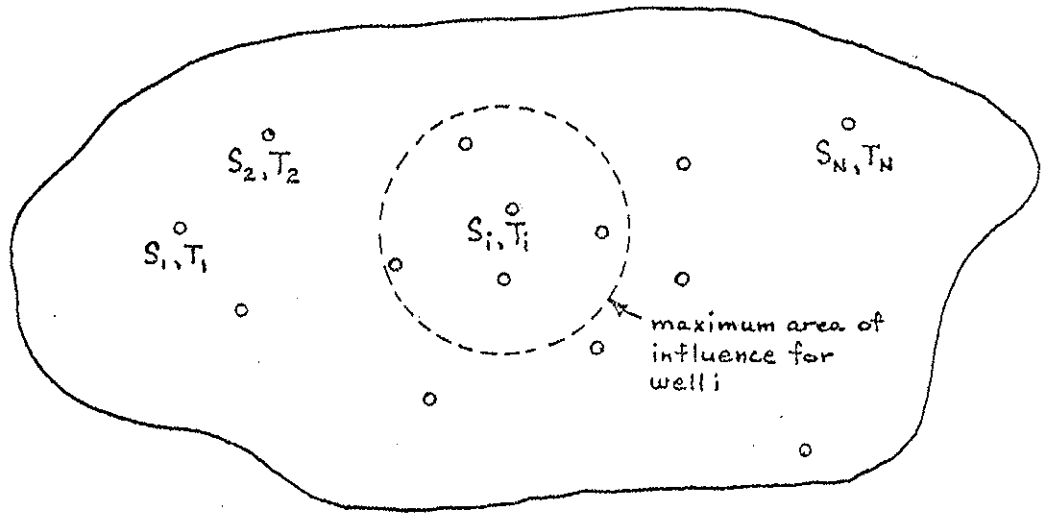


FIGURE 3: LARGE GROUNDWATER BASIN

Haines, *et. al.* [10], have used a decomposition approach to identifying aquifer parameters by considering that a hypothetical "no-flow" line can be mathematically constructed between wells. Each well is then contained in a wedge, within which aquifer characteristics are assumed to be homogeneous. The parameters, as well as the positions of the "no-flow" lines, are optimally determined from well records.

A groundwater basin can be considered as a discrete distributed parameter system, where wells are analogous to spatially distributed nodes. Nonuniqueness of parametric estimation is always a possibility when attempts are made to surmise continuous properties from spatially discrete data. The approach presented here makes no such attempt. In the following sections, a promising decomposition algorithm for large nonlinear problems, developed mainly by Lasdon [6], is discussed and applied to an example problem.

As mentioned previously, the parameter T for unconfined aquifers is actually a variable, which renders impossible an explicit solution of Equation 1. Future reports will deal with this problem, but at present, the emphasis is on confined aquifers. In addition there will be consideration of stratified groundwater basins where parameters vary with depth as well as spatially.

II. SEPARABILITY OF LARGE PROBLEMS

The essential idea behind nonlinear decomposition is to write a Lagrangian for the large nonlinear programming problem and then decompose the Lagrangian into smaller problems which can be solved independently. It is essential that the objective function and constraints are additively separable. It will be subsequently shown, however, that this requirement presents no real difficulty if some non-separable coupling does occur in the objective function and/or constraints.

Consider the following large nonlinear problem.

- (1) Minimize $f(x)$
- (2) Subject to $g_i(x) \leq 0 \quad i = 1, \dots, m$
- (3) $x \in S \quad S \subset R^n$.

The Lagrangian can be defined as

$$(4) \quad L(x, u) = f(x) + u g(x)$$

where

$$u = (u_1, \dots, u_m)$$
$$g(x) = (g_1(x), \dots, g_m(x))^T.$$

Suppose that the above problem is to be decomposed into p subproblems, where $p < n$. Separability requires that

$$(5) \quad f(x) = \sum_{i=1}^p f_i(x_i)$$

$$(6) \quad g_i(x) = \sum_{k=1}^p g_{ik}(x_k)$$

where

$$x = (x_1, x_2, \dots, x_p)$$

$$x_i \in S_i$$

$$S = S_1 \times S_2 \times \dots \times S_p .$$

The Lagrangian can now be separated into p independent subproblems

$$(7) \quad \min_{x_k \in S_k} \left\{ f_k(x_k) + \sum_{i=1}^m u_i g_{ik}(x_k) \right\}, \quad k = 1, \dots, p .$$

If u_i^* , $i = 1, \dots, m$, can be found such that Equation 2 is satisfied, then the original problem has been solved (Equations 1-3).

Suppose that the objective function $f(x)$ is not completely separable, or that some variables must be shared between subproblems. It will now be shown that complete separability can be attained at the expense of an increase in dimensionality of the subproblems. This procedure can also be applied to problems with nonseparable constraints. However, for the purposes of illustration, the constraints are assumed to be properly separable and so will be temporarily ignored (without loss of generality).

Examination of $f(x)$ indicates that the vector x can be partitioned into two vectors

$$(8) \quad x = \left[\begin{array}{c|c} y & z \end{array} \right]$$

where

y is in R^{n_1} and each component is associated with only one subproblem.
 z is in R^{n_2} and each component is associated with two or more subproblems.

$$(9) \quad n_1 + n_2 = n .$$

The vector z is the troublesome coupling vector. If $n_2 = 0$, then $f(x)$ is completely separable.

Define the following:

p = the total number of subproblems.

$C = \|\|c_{ij}\|\|$ is a $p \times p$ matrix of 0's and 1's which specifies the coupling between subproblems i and j ; with $c_{ii} = 0, i = 1, \dots, p$.

$s_i = \sum_{\substack{j=1 \\ i \neq j}}^p c_{ij} z_j$ are the vectors directed to subproblem i from other

subproblems $(s_i \in T_i, T \subset R^n, T = T_1 \times \dots \times T_p)$.

$f_i(y_i, z_i, s_i)$ = the objective function associated with subproblem i .

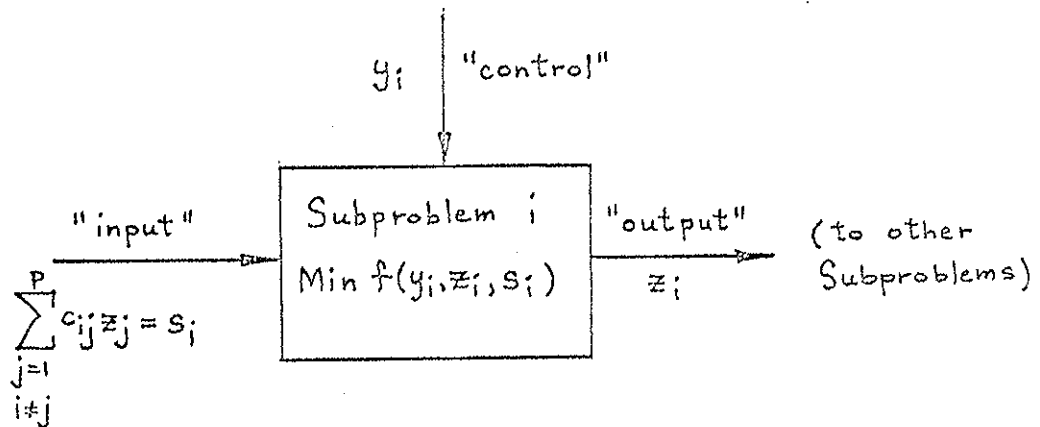


FIGURE 4: COUPLED SUBPROBLEM

The original problem is still intact, since coupling has not been broken. Decomposition is carried out by designating the vectors $s_i, i = 1, \dots, p$, as *pseudo-variables* which are independent of the vectors incoming from other subproblems (Baumann [1] and Wismer [9]).

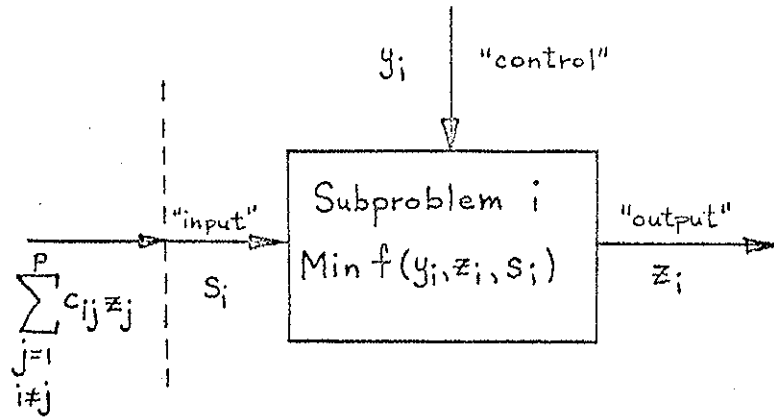


FIGURE 5: UNCOUPLED SUBPROBLEM

The following new problem is now defined:

$$(10) \quad \min_{y, z, s} \sum_{i=1}^P f_i(y_i, z_i, s_i)$$

subject to

$$(11) \quad \sum_{j=1}^P c_{ij} z_j - s_i = 0 \quad i = 1, \dots, p.$$

And,

$$(12) \quad \min_{y, z, s} \sum_{i=1}^P f_i(y_i, z_i, s_i) = \sum_{i=1}^P \min_{y_i, z_i, s_i} f_i(y_i, z_i, s_i).$$

The addition of Equation 11 establishes the equivalence between this problem and the original coupled problem. Define

$$(13) \quad L = \sum_{i=1}^p f_i(y_i, z_i, s_i) + \sum_{i=1}^p \lambda_i \left[\sum_{j=1}^p c_{ij} z_j - s_i \right]$$

where λ_i^T , $i = 1, \dots, p$, are Lagrange multipliers of the same dimension as s_i , $i = 1, \dots, p$.

For fixed values of λ_i , L is to be minimized. If there exist λ_i^* , $i = 1, \dots, p$, such that Equation 11 is satisfied, then the original problem is solved (such λ_i^* may not exist, however).

The problem at hand is to decompose Equation 13 into p separate subproblems (Lasdon and Schoeffler [7])

$$(14) \quad \sum_{i=1}^p \lambda_i \left[\sum_{j=1}^p c_{ij} z_j - s_i \right] = \sum_{i=1}^p \sum_{j=1}^p \lambda_i [c_{ij} z_j] - \sum_{i=1}^p \lambda_i s_i$$

$$(15) \quad \sum_{i=1}^p \sum_{j=1}^p \lambda_i [c_{ij} z_j] = \sum_{j=1}^p \sum_{i=1}^p \lambda_i [c_{ij} z_j] = \sum_{j=1}^p \left[\sum_{i=1}^p \lambda_i c_{ij} \right] z_j .$$

Switching i and j in Equation 15 gives

$$\sum_{i=1}^p \left[\sum_{j=1}^p \lambda_j c_{ji} \right] z_i .$$

Define

$$(16) \quad \gamma_i = \sum_{j=1}^p \lambda_j c_{ji} .$$

The Lagrangian is

$$(17) \quad L = \sum_{i=1}^p [f_i + \gamma_i z_i - \lambda_i s_i]$$

$$(18) \quad L_i = f_i + \gamma_i z_i - \lambda_i s_i \quad i = 1, \dots, p.$$

Separability has been attained through addition of pseudo-variables s_i , $i = 1, \dots, p$. The constraints (Equation 2) can now be incorporated into the following new problem.

$$(19) \quad \min L_k = \min_{\substack{x_k \in S_k \\ s_k \in T_k}} \left\{ f_k(x_k, s_k) + \gamma_k z_k - \lambda_k s_k + \sum_{i=1}^m u_i g_{ik}(x_k) \right\} \quad k = 1, \dots, p$$

where $x_k = \begin{bmatrix} y_k \\ z_k \end{bmatrix}$. For the more general case where nonseparable coupling occurs in the constraints also, the same procedure of adding appropriate pseudo-variables applies.

Figure 6 illustrates a two-level scheme for adjusting λ_i and u_i , $i = 1, \dots, m$, until the coupling constraints (Equation 11) and problem constraints (Equation 2) are eventually satisfied. The 2nd level controller passes down initial guesses $\lambda^{(0)}$, $u^{(0)}$ to the subproblems. The subproblems (Equation 19) are then solved and the results passed back to the 2nd level. The error across the cuts and the values of $g(x)$ are computed and a new set of prices generated such that the coupling error is reduced and $g(x)$ moves closer to being nonpositive. This process continues until all constraints are satisfied.

Decomposition Principle:

If those λ^* and u^* are found such that each subproblem attains a global optimum and all problem constraints and coupling constraints are satisfied, then a global optimum has been found for the overall problem. (Note: if one or more of the subproblems attain only local optima, then the overall optimum is also local.)

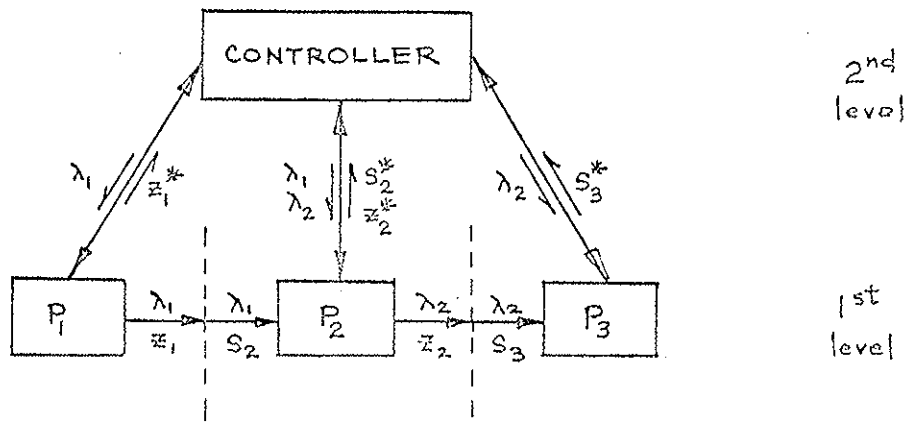


FIGURE 6: DECOMPOSITION

If the number of Lagrange multipliers becomes excessive for large problems, a hierarchy of control can be developed. Each controller can then be concerned with adjusting a limited number of multipliers.

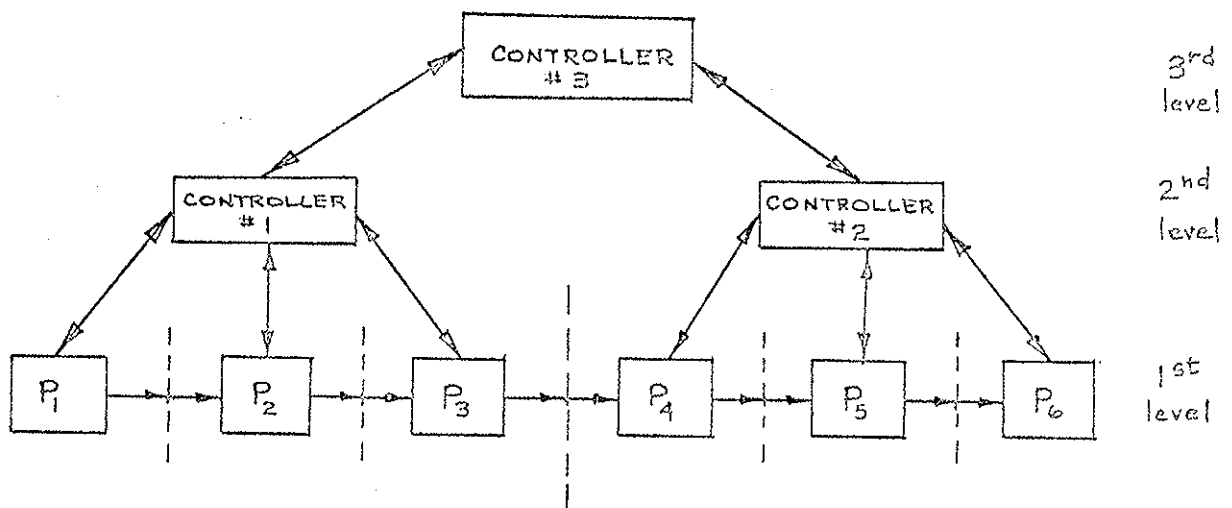


FIGURE 7: MULTI-LEVEL DECOMPOSITION

In the following section conditions are specified which guarantee convergence of the second-level controller to λ^* and u^* . As pointed out by Everett [3], however, it is usually valuable to attempt this method even though these conditions cannot be properly met. In many cases, upper and lower bounds on the overall solution can be determined, if not the solution itself.

III. CONVERGENCE OF THE 2ND LEVEL PROBLEM

In order to simplify the following treatment, consider the following *primal problem* (note: min or max will always be assumed to exist):

- (1) Minimize $f(x)$
(2) Subject to $g_i(x) \leq 0$, $i = 1, \dots, m$
(3) $x \in S$ $S \subset \mathbb{R}^n$.

It is easy to show that the following results are equally valid if equality constraints are added to the primal problem.

Define the *dual function* as

$$(4) \quad h(u) = \min_{x \in S} L(x, u)$$

where

$$(5) \quad L(x, u) = f(x) + u g(x)$$

$$u = (u_1, \dots, u_m)$$

$$u \in D$$

$$D = \left\{ u \mid u \geq 0, \min_{x \in S} L(x, u) \text{ exists} \right\}$$

$$g(x) = (g_1(x), \dots, g_m(x))^T.$$

Theorem 1:

$h(u) \leq f(x)$ for all x satisfying Equations 2 and 3, for all $u \in D$.

Proof:

$$\begin{aligned}h(u) &= \min_{x \in S} (f(x) + ug(x)) \\ &\leq f(x) + ug(x) \quad x \in S, u \in D.\end{aligned}$$

For all $x \in S$ satisfying $g(x) \leq 0$,

$$(6) \quad h(u) \leq f(x) \quad ||$$

A *duality gap* exists if $h(u) < f(x)$. Therefore,

$$\min_{x \in S} \max_{u \geq 0} L(x, u) = \max_{u \geq 0} \min_{x \in S} L(x, u)$$

if and only if there exists a *saddle point* (Lasdon [6] and Karlin [5]).

Theorem 2:(Lasdon [6])

The dual function $h(u)$ is concave over any convex subset of its domain D .

Proof:

Let D be a convex subset of \hat{D} and let $u_1, u_2 \in \hat{D}$, $\alpha \in [0, 1]$. Then

$$h(\alpha u_1 + (1 - \alpha)u_2) = \min_{x \in S} L(x, \alpha u_1 + (1 - \alpha)u_2)$$

$$\begin{aligned}L(x, \alpha u_1 + (1 - \alpha)u_2) &= f(x) + (\alpha u_1 + (1 - \alpha)u_2)g(x) \\ &= \alpha f(x) + \alpha u_1 g(x) + (1 - \alpha)f(x) + (1 - \alpha)g(x)\end{aligned}$$

therefore,

$$\begin{aligned}
h(\alpha u_1 + (1 - \alpha)u_2) &= \min_{x \in S} [\alpha L(x, u_1) + (1 - \alpha)L(x, u_2)] \\
&\geq \alpha \min_{x \in S} L(x, u_1) + (1 - \alpha) \min_{x \in S} L(x, u_2) \\
&= \alpha h(u_1) + (1 - \alpha)h(u_2) . \quad ||
\end{aligned}$$

Having established the concavity of the dual function (even though f and g may be nonconvex), it is clear that maximization of $h(u)$ over \hat{D} will yield a global answer. Equation 6 says that this answer will be at least a lower bound for $f(x)$. It remains to specify the conditions under which a saddle point is assumed to exist.

Define the nonincreasing function $w(y)$ as

$$(7) \quad w(y) = \min \{f(x) \mid g(x) \leq y, x \in S\} .$$

The domain of $w(y)$ is

$$(8) \quad F = \{y \mid \exists x \in S \ni g(x) \leq y\}$$

Theorem 3:

If f , g , and S are convex, then $w(y)$ is convex over F .

Proof: $((\cdot))^*$ refers to the global optimum)

Define

$$(9) \quad w(y^1) = \min_{x \in S} \{f(x) \mid g(x) \leq y^1\}$$

where $x^* = x^1$, $y^1 \in F$

$$(10) \quad w(y^2) = \min_{x \in S} \{f(x) \mid g(x) \leq y^2\}$$

where $x^* = x^2$, $y^2 \in F$. For any $\alpha \in [0,1]$,

$$(11) \quad w(\alpha y^1 + (1 - \alpha)y^2) = \min_{x \in S} \{f(x) \mid g(x) \leq \alpha y^1 + (1 - \alpha)y^2\}$$

where $x^* = \bar{x}$. Since

$$(12) \quad \begin{aligned} g(\alpha x^1 + (1 - \alpha)x^2) &\leq \alpha g(x^1) + (1 - \alpha)g(x^2) \\ &\leq \alpha y^1 + (1 - \alpha)y^2 \end{aligned}$$

then $\alpha x^1 + (1 - \alpha)x^2$ satisfies the constraints in Equation 11. Hence

$$(13) \quad f(\bar{x}) \leq f(\alpha x^1 + (1 - \alpha)x^2)$$

$$(14) \quad f(\alpha x^1 + (1 - \alpha)x^2) \leq \alpha f(x^1) + (1 - \alpha)f(x^2)$$

therefore,

$$(15) \quad w(\alpha y^1 + (1 - \alpha)y^2) \leq \alpha w(y^1) + (1 - \alpha)w(y^2) . \quad ||$$

Consider the set of points R in R^{m+1} on and above the graph of $w(y)$:

$$(16) \quad R = \{(y_0, y) \mid y \in F, y_0 \geq w(y)\}$$

Theorem 4: (Lasdon [6])

If x^0 solves the primal problem, then there exist $u^0 \geq 0$ such that x^0 minimizes $L(x, u^0)$ over S if and only if the set R (Equation 10) has a supporting hyperplane at point $(f(x^0), g(x^0))$.

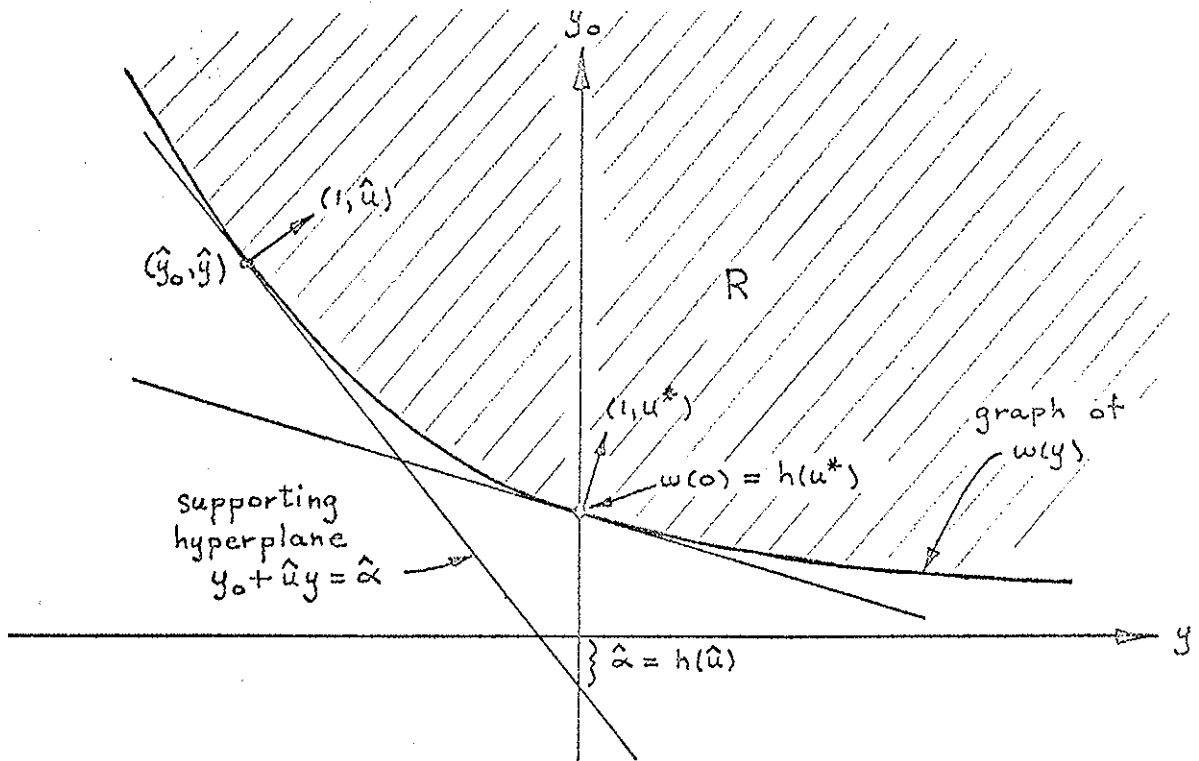


FIGURE 8: SUPPORTING HYPERPLANE FOR SET R

Figure 6 illustrates Theorem 4 for the case where $w(y)$ is strictly convex. The primal problem is to minimize the y_0 intercept of the points in R , while the dual problem is to maximize the y_0 intercept of the supporting hyperplanes for R . The Lagrange multipliers u are the slopes of the hyperplanes, and are always nonnegative since $w(y)$ is nonincreasing.

The dual problem corresponds to the problem of the second-level controller. Concavity of the dual function assures that the second-level controller will converge to a global answer. Convexity of $w(y)$ guarantees that this answer is a saddle point. Some nonconvex problems can be solved through use of the dual function as long as a saddle point exists (i.e., a supporting hyperplane for R exists at the optimal primal solution). Otherwise, a duality gap exists, which corresponds to the inability of the second-level controller to equalize flows across all the "cuts" in the large problem.

Since the dual function is to be maximized with respect to multipliers u , conditions under which the gradient exists must be specified.

Theorem 5: (Lasdon [6])

If (i) S is closed and bounded and (ii) $f(x)$ and $g(x)$ are continuous on S , then the dual function $h(u)$ is differentiable at point \bar{u} if and only if each $g_i(x)$ is constant over $X(\bar{u})$, where

$$(17) \quad X(\bar{u}) = \{x \mid x \text{ minimizes } L(x,u) \text{ over } S\}.$$

In this case

$$(18) \quad \left. \frac{\partial h}{\partial u_i} \right|_{u=\bar{u}} = g_i(x) \quad \forall x \in X(\bar{u})$$

The gradient of h with respect to u_i is simply $g_i(x)$. Continuous 1st partial derivatives are not required. Under the given assumptions, any number of algorithms utilizing this gradient will converge to the optimal dual solution and hence, if a saddle point exists, the optimal primal solution.

An Alternative 2nd-level Controller

An alternative 2nd-level controller to the previous one proposed by Lasdon [6] offers considerable reduction in dimensionality of the individual subproblems. The basic difference is that the pseudo-variables s utilized for uncoupling subproblems are treated as parameters rather than variables. Equation II-19 from the previous section is written as

$$(19) \quad \min L_k = \min_{x_k \in S_k} \left\{ f_k(x_k; s_k) + \gamma_k z_k + \sum_{i=1}^m u_i g_{ik}(x_k) \right\} \quad k = 1, \dots, p$$

where $x_k = \begin{bmatrix} y_k \\ -z_k \end{bmatrix}$. Note that the 3rd term in Equation II-19 has been canceled

since both λ_k and s_k are parameters. In addition to the assumptions listed in Theorem 5, this algorithm requires that f belong to class C^1 . If pseudo-variables are also required to uncouple the constraints g , then g must also possess continuous 1st partial derivatives. Assume for now that g is separable.

Procedure:

1. The 2nd level controller sends initial guesses $s_k^{(0)}, \lambda_k^{(0)}, u_i^{(0)}$; $k = 1, \dots, p$, $i = 1, \dots, m$, to the subproblems. Set $\ell = 1$.

2. Each subproblem computes x_k^* using Equation 19 and $\left. \frac{\partial s_k}{\partial \lambda^{(\ell)}} \right|_{x_k = x_k^*}$ using Equation II-19 from the previous section ($k = 1, \dots, p$) where

$$(20) \quad \frac{\partial s_k}{\partial \lambda^{(\ell)}} = \left[\frac{\partial s_k}{\partial \lambda_1^{(\ell)}}, \frac{\partial s_k}{\partial \lambda_2^{(\ell)}}, \dots, \frac{\partial s_k}{\partial \lambda_p^{(\ell)}} \right]$$

and

$$(21) \quad \frac{\partial s_k}{\partial \lambda_i^{(\ell)}} = \left[\frac{\partial L_k}{\partial s_k} \right]^{-1} \cdot \frac{\partial L_k}{\partial \lambda_i^{(\ell)}} \quad i = 1, \dots, p.$$

3. The second-level utilizes any efficient algorithm for maximizing the Lagrangian function - preferably in the direction of the gradient specified by Theorem 5. Appropriate changes in the multipliers, $\Delta \lambda_k^*$ and Δu_i^* ; $k = 1, \dots, p$, $i = 1, \dots, m$, result.

$$\lambda_k^{(\ell+1)} = \lambda_k^{(\ell)} + \Delta \lambda_k^* \quad k = 1, \dots, p$$

$$u_i^{(\ell+1)} = u_i^{(\ell)} + \Delta u_i^* \quad i = 1, \dots, m$$

(22)

$$s_k^{(\ell+1)} = s_k^{(\ell)} + \theta_k \cdot \Delta\lambda^* \cdot \left[\frac{\partial s_k}{\partial \lambda^{(\ell)}} \right]_{x_k = x_k^*}^T$$

$$\theta_k > 0, \Delta\lambda^* = [\Delta\lambda_1^*, \dots, \Delta\lambda_p^*]$$

$$k = 1, \dots, p.$$

4. Is the Lagrangian maximized with respect to u and λ ?

Yes: Stop.

No: Continue.

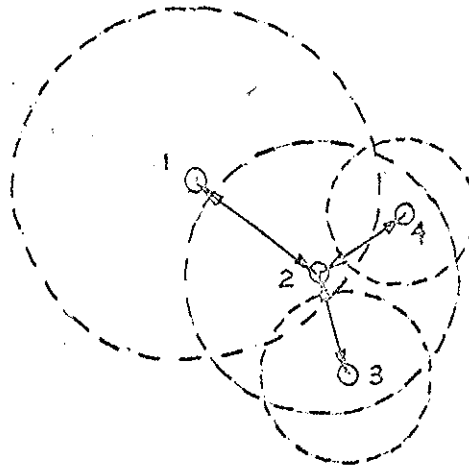
5. $\ell \rightarrow \ell + 1$

Go to 2.

It is conceivable that with careful selection of step sizes θ_k , $k = 1, \dots, p$, this algorithm can maximize the dual as rapidly as Lasdon's method, and yet require much less computation at the 1st level (subproblem level) due to decreased dimensionality. Since, however, the changes in pseudo-variables s_k are linear approximations to the actual changes which would occur if s_k were treated as subproblem variables, $\Delta\lambda_k^*$ must be small enough to avoid large inaccuracies.

IV. APPLICATION TO THE GROUNDWATER PROBLEM

Consider a system of four wells, as shown in Figure 9.



$$r_{12} = 5000 \text{ ft.}$$

$$r_{23} = 3500 \text{ ft.}$$

$$r_{24} = 2500 \text{ ft.}$$

FIGURE 9: 4-WELL SYSTEM

The circle surrounding well 1 depicts the expected maximum extent of the cone of depression for well 1, $i = 1, \dots, 4$. Notice that well 2 is affected by well 1, but not *vice versa*. Also, well 2 affects wells 3 and 4.

For purposes of illustration, assume that parameters T_i , $i = 1, \dots, 4$, are accurately known. The primal problem is

$$(1) \quad \min E = \min_{S_1, \dots, S_4} \sum_{i=1}^4 \sum_{k=0}^M \left[\sum_{j=1}^4 [d_j(t_k, r_{ij})] - d_i(t_k, r_{ii}) \right]^2$$

Decomposition is accomplished by introducing pseudo-variables σ_i , $i = 1, \dots, 3$, where

$$\begin{aligned}
 (2) \quad & s_1 - \sigma_1 = 0 \\
 & s_2 - \sigma_2 = 0 \\
 & s_2 - \sigma_3 = 0
 \end{aligned}$$

and separating the Lagrangian into four independent problems

$$L = \sum_{i=1}^4 L_i .$$

From Equation II-18 in Section II, the subLagrangians are

$$(3) \quad L_1(s_1; \lambda_1) = \sum_{k=0}^M \left[d_1(t_k, r_{11}; s_1) - \hat{d}_1(t_k, r_{11}) \right]^2 + \lambda_1 s_1$$

$$\begin{aligned}
 (4) \quad L_2(s_2, \sigma_1; \lambda_1, \lambda_2, \lambda_3) &= \sum_{k=0}^M \left[d_1(t_k, r_{21}; \sigma_1) + d_2(t_k, r_{22}; s_2) - \hat{d}_2(t_k, r_{22}) \right]^2 + \\
 &+ (\lambda_2 + \lambda_3) s_2 - \lambda_1 \sigma_1
 \end{aligned}$$

$$\begin{aligned}
 (5) \quad L_3(s_3, \sigma_2; \lambda_2) &= \sum_{k=0}^M \left[d_2(t_k, r_{32}; \sigma_2) + d_3(t_k, r_{33}; s_3) - \hat{d}_3(t_k, r_{33}) \right]^2 - \\
 &- \lambda_2 \sigma_2
 \end{aligned}$$

$$\begin{aligned}
 (6) \quad L_4(s_4, \sigma_3; \lambda_3) &= \sum_{k=0}^M \left[d_2(t_k, r_{42}; \sigma_3) + d_4(t_k, r_{44}; s_4) - \hat{d}_4(t_k, r_{44}) \right]^2 - \\
 &- \lambda_3 \sigma_3 .
 \end{aligned}$$

The original 4-dimensional problem has been decomposed into three 2-dimensional problems and one 1-dimensional problem, plus a 3-dimensional concave dual problem.

Lasdon's gradient controller was utilized for the 2nd level or dual problem of adjusting λ_i , $i = 1, \dots, 3$ so as to maximize L with respect to λ or minimize the total error across the cuts.

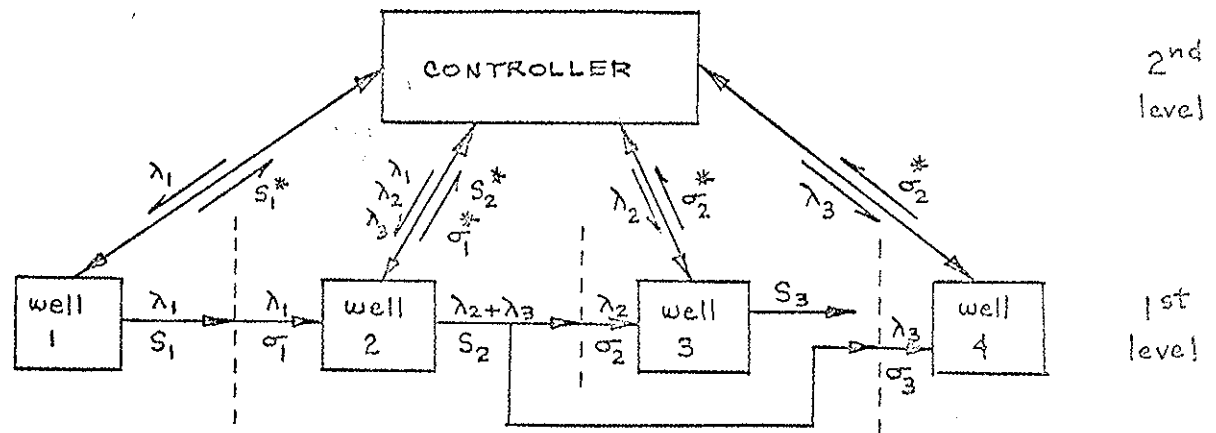


FIGURE 10: DECOMPOSED GROUNDWATER PROBLEM

$$(7) \quad \lambda_1^{(n+1)} = \lambda_1^{(n)} + \theta_1 [S_1^* - \sigma_1^*]$$

$$(8) \quad \lambda_2^{(n+1)} = \lambda_2^{(n)} + \theta_2 [S_2^* - \sigma_2^*]$$

$$(9) \quad \lambda_3^{(n+1)} = \lambda_3^{(n)} + \theta_3 [S_3^* - \sigma_3^*]$$

where $\theta_i > 0$, $i = 1, \dots, 3$.

The subproblem minimizations (Equations 3 to 6) were carried out by a 2nd-order gradient search procedure.[†]

Hypothetical field data for the system of wells of Figure 9 are given in Tables 1 and 2. Initial guesses for parameters $S_i^{(0)}$ are given, as well as the known values for parameters T_i^* . Drawdown observations $(h_o - \hat{h})$ are used instead of water head \hat{h} .

[†]Subroutine HO CAL NMIN, Computer Center Library, University of California, Berkeley.

Computation results are illustrated in Figures 10 and 11. Figure 10 depicts the maximization of the dual or 2nd-level problem. Convergence is not monotonic until after the 3rd iteration, due to rather large step sizes initially.

Figure 11 indicates that a saddle point was found since the total absolute error across the cuts

$$|e| = |s_1^* - \sigma_1^*| + |s_2^* - \sigma_2^*| + |s_2^* - \sigma_3^*|$$

converged to zero.

The initial set of multipliers were set equal to zero - $\lambda_i^{(0)} = 0$, $i = 1, \dots, 3$. The final optimal values were

$$\lambda_1^* = 7.368 \times 10^4$$

$$\lambda_2^* = 1.0028 \times 10^3$$

$$\lambda_3^* = 1.0028 \times 10^3 .$$

Table 3 lists the optimal parameter solutions s_i^* , $i = 1, \dots, 4$, and final drawdown values. The final squared-error between computed and observed drawdown was

$$E^* = 8.5106 \text{ ft}^2$$

TABLE 1				
GROUNDWATER BASIN DATA				
	WELL			
	i = 1	i = 2	i = 3	i = 4
$h_o(\text{ft})$ $b(\text{ft})$	1000 50	1000 50	1000 50	1000 50
$r_{ii}(\text{ft})$	1	1	1	1
$Q_i(\text{gal/min})$	1000	600	450	350
$s_i^{(o)}$.00042	.00048	.0072	.0072
$T_i^* \left(\frac{\text{gal/day}}{\text{ft}} \right)$	9250	6500	4850	4025

TABLE 2				
DRAWDOWN OBSERVATIONS ($h_o - \hat{h}$) (ft.)				
Time t (days)	WELL			
	i = 1	i = 2	i = 3	i = 4
30	236	194	145	150
45	241	203	152	157
60	245	210	157	162
90	250	219	164	170
180	258	234	177	183
360	267	250	191	197
720	275	266	206	211

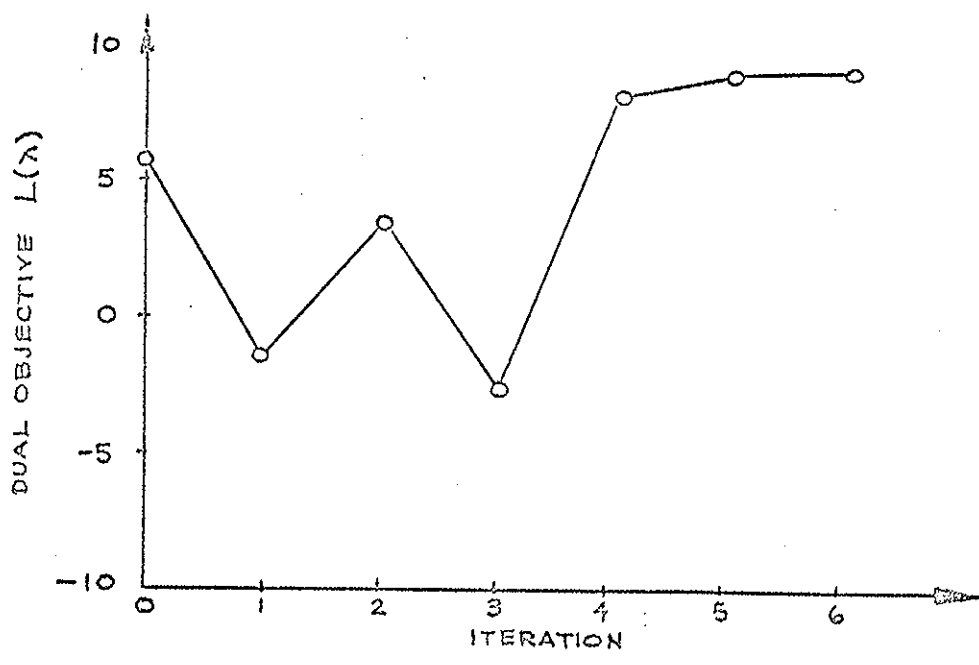


FIGURE 10: MAXIMIZATION OF THE 2ND LEVEL PROBLEM

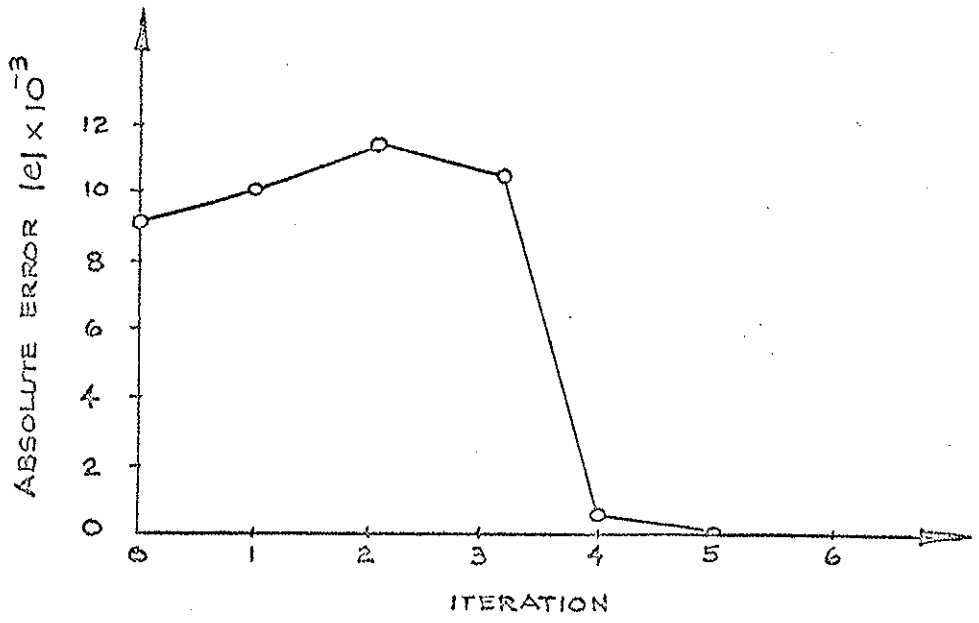


FIGURE 11: MINIMIZATION OF TOTAL CUT ERROR

TABLE 3

OPTIMAL SOLUTIONS

Time t (days)	WELL			
	i = 1	i = 2	i = 3	i = 4
30	235.9	194.1	145.9	150.5
45	240.9	203.1	152.3	157.6
60	244.4	209.5	157.3	162.8
90	249.5	218.7	164.6	170.4
180	258.1	234.5	178.0	184.0
360	266.6	250.3	192.0	197.8
720	275.2	266.2	206.3	211.9
S_i^*	.000449	.00718	.064	.020

V. DISCUSSION

A generalized decomposition approach to the solution of large nonlinear programming problems has been presented and applied to the optimal identification problem in groundwater hydrology. In addition, a new 2nd level algorithm has been proposed which offers further reduction of dimensionality in a decomposition framework.

Computational experience has revealed the following:

1. After the first iteration of the second-level controller, subproblem solutions converge more rapidly. Solutions from previous iterations are utilized as accurate initial approximations for first level solutions in succeeding iterations.
2. For intermediate iterations of the second level, it seems unnecessary that subproblem solutions converge to highly accurate values. Total computation time can be significantly lessened by allowing subproblem solutions to fall within some ϵ neighborhood of the true optimum. Then as the second level problem begins to converge, let $\epsilon \rightarrow 0$. This idea raises some interesting questions of stability and trade off between decreased subproblem computational times and increased inaccuracies.

REFERENCES

- [1] Baumann, Edward James, "Multilevel Optimization Techniques with Application to Trajectory Decomposition," Ph.D. Dissertation, Department of Engineering, University of California, Los Angeles, (1966).
- [2] Davis, S. N. and R. J. M. DeWiest, HYDROGEOLOGY, John Wiley and Sons, Inc., New York, (1966).
- [3] Everett, H., "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources," Operations Research, Vol. 11, pp. 399-417, (1963).
- [4] Geoffrion, A. M., "Elements of Large-Scale Mathematical Programming," Report #R-481-PR, RAND Corporation, Santa Monica, California, (November 1969).
- [5] Karlin, S., MATHEMATICAL METHODS AND THEORY IN GAMES, PROGRAMMING AND ECONOMICS, Vol. 1, Addison-Wesley, Reading, Massachusetts, (1959).
- [6] Lasdon, L. S., "Duality and Decomposition in Mathematical Programming," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-4, No. 2, pp. 86-100, (July 1968).
- [7] Lasdon, L. S. and J. D. Schoeffler, "Decentralized Plant Control," ISA Transactions, Vol. 5, pp. 175-183, (April 1966).
- [8] Todd, D. K., GROUND WATER HYDROLOGY, John Wiley and Sons, Inc., New York, (1959).
- [9] Wismer, D. A. (Ed.), OPTIMIZATION METHODS FOR LARGE-SCALE SYSTEMS, forthcoming.
- [10] Haimes, Y. Y., R. L. Perrine and D. A. Wismer, "Identification of Aquifer Parameters by Decomposition and Multilevel Optimization," Presented at the British-Israeli Colloquium on Operational Research, Haiffa, Israel, (December 1967).