MATHEMATICAL MODEL OF A PROTOTYPE
WATER SYSTEM

M. Fiering

June 1974 WP-74-19

Note: This document will be amended by computations and other sections will be added by Messrs. Wood and Ostrum.

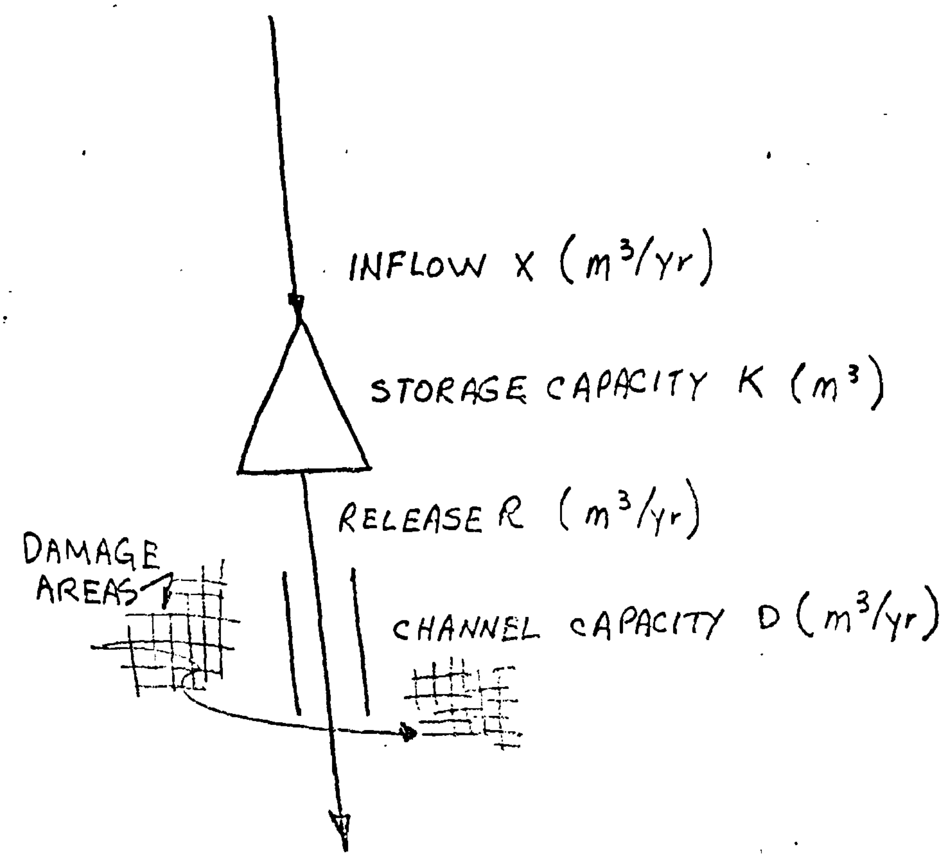
by M. Fiering

June 5, 1974

MATHEMATICAL MODEL OF A PROTOTYPE WATER SYSTEM

Consider the model shown in Figure 1. There is a single stream which yields an annual inflow X_i, where the index i represents time. The flow enters a reservoir of capacity K, from which an annual release of R_i is made. The units are compatible with respect to annual volumes so that X_i is measured in volume/year, K in volume and R in volume/year. It is understood that the total amount of water available at the beginning of any year is the storage at the end of the previous year plus the annual inflow, $S_{i-1} + X_i$. In other words, it is assumed that the annual inflow is known on the first day of the current year and that the characteristic time interval of the model is one year so that the inflow and release values, which are really rates, can be thought of as volumes for a single year.

The reservoir services some upstream demand in the vicinity of the dam; typically this might be hydro-electric power. After leaving the reservoir the channel leads through an area subject to flood damage. As shown in the figure, this area is protected by dykes along the channel. Enough is known about the hydraulic configuration of the system to assert that an annual release from the reservoir is associated with a particular flood surge which, in turn, is attenuated in some prescribed fashion between the reservoir and the protected area. Thus in this simplictic model we do not deal with the realities of flood routing, determination of peak flows, or other complications. Everything is expressed in terms of annual flow, and it is assumed that the model is sufficiently regular in its hydrologic aspects to enable us



X	0	1	2	3	4	5	6	7
p(X)	.05	.12	.15	.20	.20	.10	.10	.08

K	0	1	2	3	4	5	6
$C_1(K)$	2	4	9	25	40	57	62

D	5	6	7
$C_2(D)$	7	15	25

FIGURE 1 - MODEL & BASIC DATA

to deduce the downstream consequences associated with reservoir releases. We express all this by noting that the channel capacity at the point of potential overflow is given the symbol D (volume/year). For example, if the release R is 5 and the channel capacity D is 5, it implies that there is associated with the annual release some surge or peak flow which, when attenuated through the system, produces at the point of damage a peak flow which can just be contained within the channel. This does not mean that the channel capacity is itself 5 units, but rather that it is convenient to express the channel capacity in terms of an equivalent upstream release which, when routed through the system, would be just contained within the banks of the channel.

The inflow vector X represents a random process without serial correlation; the probability density of any particular flow in a given year is given in Figure 1. The capital cost of reservoir construction is given by the function $C_1(K)$, and a geologic investigation of the area shows that it is infeasible to construct a storage capacity in excess of 6 units. It is possible, of course, to build no reservoir at all; but even this action is associated with some cost for investigation, planning, data collection and decision-making. The storage capacity K is one of the design variables in the system.

The channel capacity of the unimproved system, measured at the point of potential overflow, is given as 4 units. This is not really the capacity of the channel because it will be recalled that the capacity is given in terms of an-equivalent annual release at the reservoir. Therefore the value "4" is a surrogate for the actual channel capacity, but for purposes of this model it will be sufficient to refer only to the annual release from the reservoir when dealing with flows through the damage area. Dykes can be built to increase the channel capacity, and Figure 1 shows the cost, $C_2(D)$, for D = 5, 6 and 7. It can be seen from the figure that the inflow X is divided into 8 discrete values ranging from zero to 7, so that under ordinary circumstances it would be quite unusual for the release from the reservoir ever to exceed 7 units; that is, if the reservoir is full and the worst possible flow is received, it will simply pass that flow without any storage. Therefore the maximal discharge passing the damage area is that associated with a reservoir release of seven units per year. Channel improvements, or increases in carrying capacity, are represented by the second design parameter for our system, the quantity D.

We now consider some of the economic characteristics which govern system operation. Figure 2 A gives the system operating policy; it is the standard or Z-shaped policy characterized by the storage capacity K and the target release, T. If the total amount of water available is less than the target, all of it is released and the reservoir remains empty. If more than the target is

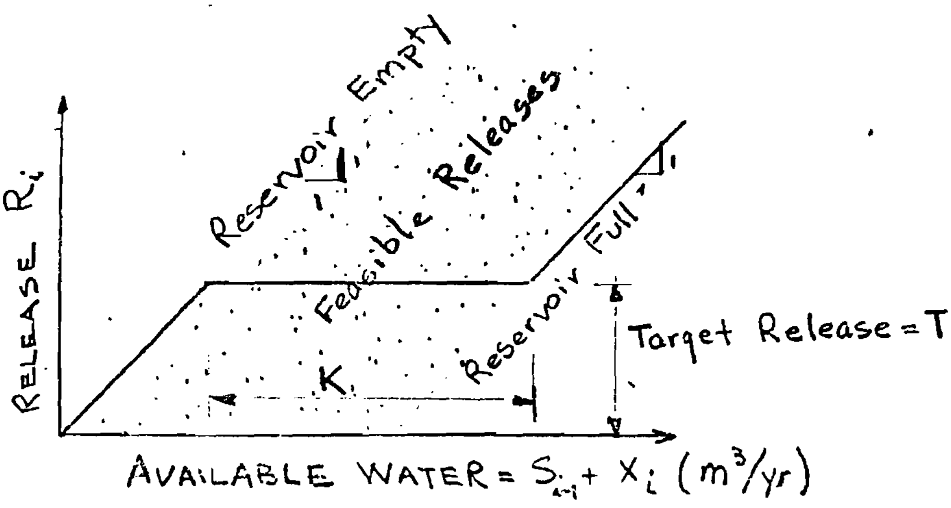


FIGURE 2A- OPERATING POLICY

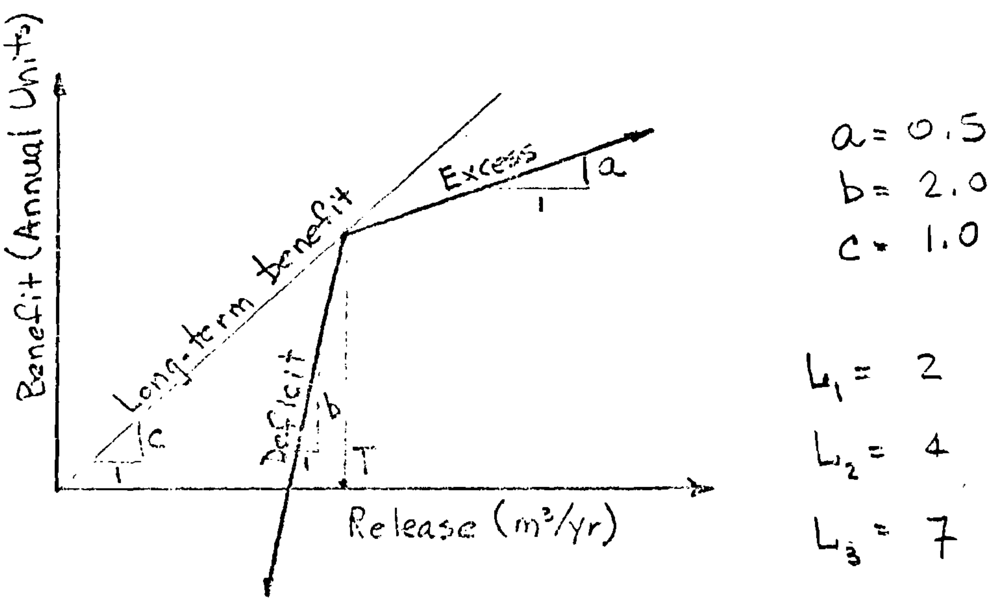


FIGURE 2B- BENEFIT FUNCTIONS

available, the release is set equal to the target until such a point is reached where additional releases must be made because the reservoir cannot store the remaining water. These two constraints, reservoir empty and reservoir full, define the band within which all feasible releases necessarily lie. The horizontal distance across the band width is precisely K, and any point which lies outside the shaded band cannot be attained by the reservoir system. The slope of the band is unity.

Figure 2B gives the benefit function for the upstream (hydroelectric or other) release. The benefit function is a three-part linear function characterized by a long-term component and two short-term componenets. The long-term benefit is a single (in this case, linear) relationship between annual benefits and target release. It represents the fact that increasing the capital investment in turbines, generators, and other facilities, necessarily implies an increased commitment to deliver water and, moreover, that the increased physical output can be sold at a constant marginal price of c per unit of output. In the case of hydroelectric energy, the output is given in megawatt hours but for purposes of this model all electrical units are converted to equivalent flows of water required to service these facilities at their design or nominal operating levels. Having decided upon the long-term or capital investment which specifies T, the operation in any year can produce precisely T units, an excess, or a defecit. If there is a large flow so that some excess energy is developed, it can be sold at a marginal rate of a, but as shown in the figure a is smaller than c to accommodate the fact that dump energy is less valuable on the market than firm energy. Similarly, should there be a water def cit, other sources of energy will have to be made available at a greater price; this implies a serious drop in the economic benefits, as reflected by the slope b being much greater than c and a. In other words, there is an economic penalty associated with failing to meet the target (or commitment), and the magnitude of the penalty is greater than the magnitude of the bonus associated with generating excessive levels of system output.

In addition to benefits at the reservoir, the system can provide flood control benefits by reducing the probability and severity of extreme flows. It will be recalled that flows of 5 6 or 7 are associated with peak discharges which produce damage in the unimproved reach of the system. The probabilities of these flood events are p_5, p_6 and p_7. If a system of reservoir and dykes is built and operated reasonably, we would expect that these three probabilities should be reduced. For example, if we specify the design D = 5, the probability of overflow in the area of potential damage is changed as follows: there can be no overflow if the release is 5 because the entire discharge can be contained in the channel, the probability of the first level of overflow is then given by the probability that the release is 6, and the probability of the second level of overflow is then given by the probability that the release is 7. The probability of the most serious overflow, that which would have occurred without improvement if the flow had been 7, goes automatically to zero. We postulate in this simple model that the flood benefits are equal to the average annual damages averted, taken over the three potential flood levels. These damages are defined as L_1, L_2 and L_3. Numerical values for these parameters are shown in Figure 2 B.

It is tempting to claim that the objective function for system design is the maximization of some combination of benefits and costs. Typically this might be the ratio, the difference, or some other function which takes account of various budgetary constraints and physical requirements. In the ordinary calculus of such a system, it is traditional to specify a discounting factor which trades on the availability and price of money required for the capital investment, and which is used to discount to present value the time stream of annual economic benefits. There are some difficulties with this notion when dealing with different economic systems, and in our model we show the effect of the rate of interest by accumulating the present value of benefits for a few sample interest rates, among which we include zero to represent the condition of no discounting.

Moreover, it is clear that there might be different social and political weights assigned to the benefits perceived by the upstream and downstream users of the the system. These weights might be widely different so that the optimal design for the system could vary enormously as a function of whose weight dominates the benefit calculation. The system design consists of three numbers; we have already identified the storage capacity (K) and the dyke level (D) as design decisions, and to these we now add the target release (T). These three parameters define a series of points in a response surface, and the usual workings of a design procedure require that this multi-dimensional space be examined in the hope of identifying the optimal response (however that might be measured). But if the response is perceived to attain different values for each of the institutions represented in the decision-making process, it is clear that the sum of benefits is not necessarily the best metric for system evaluation.

Therefore, before moving on to discussing the analysis of the model, it should be clear that we do not purport to develop an optimal solution because we recognize that optimality implies some judgements concerning the way in which benefits should be measured, discounted and combined. We will show only how to calculate some of the physical responses, how to convert these to benefits at their points of origin and how to tabulate these in such a way that additional methological tools (for example, Paretian analysis) can be employed to identify negotiation frontiers, side payments and other cost-and revenue-sharing devices for reaching a harmonious design under competition. In so doing we anticipate that the essential economic parameters, those which must be refined before agreement can be reached, will be identified; this will lead, in our judgement, to a program of inquiry which can fruitfully be pursued in order to identify optimal data collection techniques, methodological issues and, ultimately, an acceptable design program.

Analytical Procedures

We first specify a design vector which consists of numerical assignments for the storage capacity K, the dyke level D and the target draft T. Thereafter the analysis is directed at identifying the probability distribution of releases R from the reservoir. Associated with each release is some economic benefit which can be read directly from the benefit function for hydroelectric energy (Figure 2B) and augmented by the amount of flood damage alleviation associated with the level assigned to D.

These economic values are then weighted by their respective probabilities and summed in accordance with the schemes (and with attention to the warnings) described above, whereupon the trial design vector is then available for ranking and negotiation as part of the more comprehensive planning process.

In order to calculate the draft probabilities it is necessary first to have the steady state reservoir probabilities; these are identified by the symbol P_i. We tabulate first all the possible values associated with the trial system. The maximal flow is 7, and we assume an initial design vector (for this example only) of K = 4, D = 6 and T = 2. This means that there can be at most 4 units of water available in storage so that the total amount of water available at any time cannot exceed 11 or 7 + 4. The operating policy specifies that a target release of 2 will be attempted; as shown in the second column of the attached table, two

$T=2$, $K=4$, $D=6$

vailable $S_{i-1}+X_i$	Release R_i	Remaining S_i	Probability
0	0	0	$p_0 P_0$
1	1	0	$p_1 P_0 + p_0 P_1$
2	2	0	$p_2 P_0 + p_1 P_1 + p_0 P_2$
3	2	1	$p_3 P_0 + p_2 P_1 + p_1 P_2 + p_0 P_3$
4	2	2	$p_4 P_0 + p_3 P_1 + p_2 P_2 + p_1 P_3 + p_0 P_4$
5	2	3	$p_5 P_0 + p_4 P_1 + p_3 P_2 + p_2 P_3 + p_1 P_4$
6	2	4	$p_6 P_0 + p_5 P_1 + p_4 P_2 + p_3 P_3 + p_2 P_4$
7	3	4	$p_7 P_0 + p_6 P_1 + p_5 P_2 + p_4 P_3 + p_3 P_4$
8	4	4	$p_7 P_1 + p_6 P_2 + p_5 P_3 + p_4 P_4$
9	5	4	$p_7 P_2 + p_6 P_3 + p_5 P_4$
10	6	4	$p_7 P_3 + p_6 P_4$
11	7	4	$p_7 P_4$

$$\begin{cases} P_0 = (p_0 P_0) + (p_1 P_0 + p_0 P_1) + (p_2 P_0 + p_1 P_1 + p_0 P_2) \\ P_1 = p_3 P_0 + p_2 P_1 + p_1 P_2 + p_0 P_3 \\ P_2 = p_4 P_0 + p_3 P_1 + p_2 P_2 + p_1 P_3 + p_0 P_4 \\ P_3 = p_5 P_0 + p_4 P_1 + p_3 P_2 + p_2 P_3 + p_1 P_4 \\ P_4 = (p_6 P_0 + p_5 P_1 + p_4 P_2 + p_3 P_3 + p_2 P_4) + (p_7 P_0 + p_6 P_1 + p_5 P_2 \\ \quad + p_4 P_3 + p_3 P_4) + (p_7 P_1 + p_6 P_2 + p_5 P_3 + p_4 P_4) \\ \quad + (p_7 P_2 + p_6 P_3 + p_5 P_4) + (p_7 P_3 + p_6 P_4) + (p_7 P_4) \end{cases}$$

$$P_0 + P_1 + P_2 + P_3 + P_4 = 1$$

units are released unless there is not enough water or there is so much water that the capacity is inadequate to store that which remains. The table shows all the possible combinations of available flow and release, from which the remaining storage is deduced by subtraction.

The only way in which there can be no water available is if there is no storage and no inflow; this is given by the product p_0P_0. One unit is available under two possible combinations: one unit in storage coupled with no inflow and nothing in storage coupled with one unit of inflow. These combinations are independent so the sum of their joint probabilities is the probability of the compound event, as represented by $p_1P_0 + p_0P_1$. The argument continues all the way through the table, but it will be noted that there is no entry beyond P_4 because the design vector specifies that the reservoir cannot contain more than 4 units. Similarly, there is no inflow probability beyond p_7 because 7 units is the maximal annual flow. This suggests that the compound events are represented by sums of increasing number of terms until some maximum is reached, after which the number of terms decreases until the last event, a total availability of 11 units, is reached if and only if there are 4 units in storage and the inflow is seven.

We then seek to solve for the steady state probabilities P_i, and note that that the only way in which the reservoir can terminate in an empty condition is if the available flow is zero, 1 or 2 ; this corresponds to the fact that the remaining storage for those 3 events, shown in the third column of the table, is zero. Now because all of these events are independent, the probability that the system shall end in a state with zero storage is the sum of the probabilities derived from the last column, or the cumulative probability identified as line 1 in the set of equations which follow the table. Similar equations can be derived for all reservoir storage states, resulting in equations 1 through 5, giving the steady state probabilities for each of the five possible reservoir conditions.

These conditions, however, are not independent and an additional condition is required; this is the requirement that the sum of all steady state probabilities be precisely unity because the reservoir must be in one state or another at any time, and this condition is represented by the 6th equation. The solution procedure would then be to select any 4 of the first 5 equations and the 6th, noting that all of these are linear in P_i, and then to solve directly for the set of P_i. It would seem to be most sensible to eliminate equation 5 because it is the most cumbersome, but this is a matter of individual preference.

For example, the set of 6 equations is shown in the attached table, along with the solution for the steady state probabilities P_i. Clearly this vector depends on two of the three design variables: K and T. The third variable, the channel capacity or dyke level D, does not enter explicitly in determining the reservoir probabilities; under more complicated operating policies, in which the release itself is a function of D, it would enter, but this is not the case in this simple example.

It is a little more compact to generate the solution for the reservoir's steady state probability vector using matrix notation. As shown in the attached tables we write first the 5 X 12 matrix for the probability of the total water available, given the initial contents. This is essentially the information inherent in the probability density of inflows to the reservoir. We write also the 12 X 5 matrix of final (or remaining) contents, given the water available. This matrix contains zeros and ones because the operating rule, which is contained in this matrix, is deterministic, so all the probability elements are 0 or 1. The dimensions of this matrix correspond to the maximal available flow of 11 units and the maximal storage of 4 units. If we multiply the first matrix by the second, the product has dimension 5 X 5 and is the probability of final storage conditioned on the initial storage. This result is a Markov matrix whose elements are the transfer probabilities between reservoir states in time period i and those in time period (i - 1). From this Markov matrix it is a trivial matter to write the simultaneous linear equations (including the condition that the sum of all steady state probabilities must be in unity) required to solve for the steady state probabilities P_i.

Available Water, $X_i + S_{i-1}$

S_{i-1}	0	1	2	3	4	5	6	7	8	9	10	11
0	.05	.12	.15	.20	.20	.10	.10	.08	0	0	0	0
1	0	.05	.12	,15	.20	.20	.10	.10	.08	0	0	0
2	0	0	.05	.12	.15	.20	.20	.10	.10	.08	0	0
3	0	0	0	.05	.12	.15	.20	.20	.10	.10	.08	0
4	0	0	0	0	.05	.12	.15	.20	.20	.10	.10	.08

Final Storage, S_i

X_i+S_{i-1}	0	1	2	3	4
0	1	0	0	0	0
1	1	0	0	0	0
2	1	0	0	0	0
3	0	1	0	0	0
4	0	0	1	0	0
5	0	0	0	1	0
6	0	0	0	0	1
7	0	0	0	0	1
8	0	0	0	0	1
9	0	0	0	0	1
10	0	0	0	0	1
11	0	0	0	0	1

MARKOV MATRIX

S_{i-1}	S_i 0	1	2	3	4
0	0.32	0.20	0.20	0.10	0.18
1	0.17	0.15	0.20	0.20	0.28
2	0.05	0.12	0.15	0.20	0.48
3	0	0.05	0.12	0.15	0.68
4	0	0	0.05	0.12	0.83

$p(S) = \{0.01, 0.02, 0.07, 0.13, 0.77\}$

$p(R) = \{.050, .120, .150, .200, .200, .100, .100, .080\}$ (original)

$p(R) = \{0.001, .002, .388, .190, .175, .096, .087, .061\}$ (storage)

R	Ben(R)	p(R)Ben(R)	Δp
0	-2	-.002	
1	0	0	
2	2	.776	
3	2.5	.475	
4	3.0	.525	
5	3.5	.336	0.004
6	4.0	.348	0.013
7	4.5	.275	0.019
		2.733	

FLOODS

Flow	Level	L	p(level)	Expected Damage	Modified p (level)	Modified L	Expected Damage
5	1	2	0.100	0.20	0	-	-
6	2	4	0.100	0.40	0.087	2	0.174
7	3	7	0.080	0.56	0.061	4	0.244
				1.16			0.418

Average Annual Flood Benefit = 1.16 - 0.42 = 0.74

Cost of Dam = 40

Cost of Dike = 7

Upstream Gross Avg. Ann. Benefit = 2.73

Time Horizon = 25 years

Discount Rate = 0, 4%

P.V. of Benefits @ 0% = 68.25 - 40 = 28.25 PROFIT
@ 4% = 42.64 - 40 = 2.64 PROFIT

Downstream Gross Avg. Ann. Benefit = 0.74

P.V. of Benefits @ 0% = 18.50 - 7 = 11.50 PROFIT
@ 4% = 11.56 - 7 = 4.56 PROFIT

NB: Present Value of 1 unit = $\frac{1}{r}\left[(1 + r)^{n} - 1\right](1 + r)^{-n}$

where r = annual rate and n = economic time horizon.

For n = 25 and r = 0.04, PV = 15.62

On the assumption that the inputs and reservoir states are independent, it is now an easy maater to identify the several combinations which give rise to the complete range of releases R_i, to assign specific probabilities to these, and then to preceed with the economic analysis. In our example, the table shows the probability associated with each release for all twelve lines. These are summed according to their arguments, and the unconditional or marginal release probabilities are written directly.

The expected gross annual benefit from upstream utilization is tabulated as shown, and it remains only to calculate downstream benefits due to flood control. If there were no dyke, there would still be some reduction in flood probability as shown in the tables. But because the dykes can contain the peak flow associated with an annual release of 5, we assign to a damaging overflow of magnitude 5 the probability zero. The damaging overflow occasioned by a release of 6 is assigned a unit damage level associated with that of 5 and the damage level for a release of 7 is assigned the damage level previously associated with a release of 6. Thus the effect of the dyke is to change the unit damage function while the effect of the reservoir is to reduce the probability of flood events. As shown in the calculations, this combined effect produces a benefit for the downstream users.

These costs and benefits are now combined over a range of interest rates to produce the numerical basis for decision-making under competition. It is clear that this simplistic model contains many assumptions which are not tenable in models of real situations. These are centered around the independence of the inputs, the highly simplified operating policy, the use of annual events rather than seasonal or instantaneous peaks and complete avoidance of the details of flood routing and other dynamic events associated with time-varying flow and with releases from the reservoir. But the point here is to suggest that these several technological difficulties, and an equivalent number of economic ones, can be the subject of intensive investigation by the IIASA Water Resources Project; what is of interest is a model framework within which the Tisza, Vistula and other basins might be structured.

For example, a groundwater resource might be included and its source and sink effects easily modelled within the framework of this analysis. Stochastic operating rules, serial correlation amongst the inputs, and other advanced control phenomena might be incorporated. These details are for the moment not important except to note that they do not perturb the basic structure of the model and that the essential conflicts between users, between uses and between difference of geographic locations in the basin can still be highlighted by the formalisms, even though they become extremely complex and rigorous.

Dear Mike:

I read with great interest and profit your mathematical model of a prototype water system. As a pedagogic device it is superb. Four suggestions follow:

(1) In calculating benefits, if I am not mistaken, the entire flood control benefit arising both due to storage <u>and</u> due to ~~the~~ dyke of 'capacity' 5 is compared against the cost of dyke. It seems to me that it will be more instructive to compare the benefits (of flood control) of alternative dyke capacities ~~under~~ separately for the cases of no storage and storage of size 4. My calculations lead me to:

	Flow	D = 4: L	Prob.	Exp. Benefit	D = 5: L	Prob.	Exp. Benefit	D = 6: L	Prob.	Exp. Benefit
	5	2	0.100	0.200	0			0		
Storage	6	4	0.100	0.400	2	0.100	0.200	0		
	7	7	0.080	0.560	4	0.080	0.320	2	0.08	0.160
		Total		1.160			0.5200			0.160
	5	2	0.096	0.192	0			0		
	6	4	0.087	0.348	2	0.087	0.174	0		
Storage 4	7	7	0.061	0.427	4	0.061	0.244	2	0.061	0.122
				0.967			0.418			0.122

Assuming that initially there is no storage and a dyke capacity of 4 exists, we get the following benefit situation:

Expected Benefits (Flood Control)

Additional Dyke Capacity	1	2	3
With Storage	0.549	0.845	0.967
Without Storage	0.640	1.000	1.160

Thus if storage was being built anyway, the downstream decision makers will ~~have~~ compare the cost of additional dyke capacities against the additional benefits in the row with storage. If no storage was to be built then the second row above will be the relevant row. In neither case, will one compare the cost of additional dyke with the ~~added~~ benefits accruing both due to storage and dyke over ~~a~~ a situation in which no storage exists and ~~=~~ the dyke capacity is at the initial level of 4.

~~No~~ One can also ask the question: what is the maximum contribution that the downstream decision makers will be willing to pay for the construction of storage capacity 4 upstream? To answer these, we note first that at the existing dyke capacity, storage contributes 0.193 (= 1.160 − 0.967) in terms of annual benefits or ~~...~~ 3.015 in terms of present value at 4% discount rate. Let c denote the cost of additional dyke capacity that will yield the same annual benefit of 0.193 in the absence of storage. Then the answer to the question posed is Minimum of the numbers 3.015 and c, assuming that the relevant interest rate is 4%.

(2) To the extent upstream benefits are due to production of electricity, they depend not only on the release but also on the "head" which in turn depends on the reservoir level. Thus a release of 2 when the reservoir is full will produce more power than when it is empty.

T. N.