RR-74-17

SOME SYSTEM APPROACHES TO WATER RESOURCES PROBLEMS
I. OPERATION UNDER WATER SHORTAGE

Yu. A. Rosanov

October 1974

Some System Approaches to Water Resources Problems

I. Operation under Water Shortage

Yu. A. Rozanov

1. Nowadays we encounter a shortage of water resources and are obliged to deal with the problem of optimal water distribution.

When we treat a large river basin as a Large Scale System, we have to identify corresponding inputs with available water resources which appear as inflows, water reservoirs, etc.

Let us begin with a look at the river basin as a system comprising the proper components S_i; $i = 1,...,m$ which could be water users or river basin regions with a number of water users.

Suppose we are interested in distributing our water resources amongst these components S_i and that each region S_i is situated upstream with respect to S_{i-1}. The water requires some time T_i to pass through S_i to S_{i+1}.

Let us assume that the water Δw_o, which comes to the first users S_1 as the main flow during a time period $t_1, t_1 + \Delta$, will pass through S_i during the corresponding time period $[t_i, t_i + \Delta]$ where

$$t_i = t_1 + \sum_{k=1}^{i-1} T_k \quad .$$

Furthermore, if we assume that w_i is the total amount of water available for S_i during $[t_i, t_i + \Delta]$, and that x_i is the water consumption in the region S_i during this period of time, then the rest

$$\Delta w_i = w_i - x_i \tag{1}$$

is available for all users situated downstream. According to these assumptions, the water available for S_i during $[t_i, t_i + \Delta]$ is

$$w_i = w_{i-1} + \xi_i \quad , \tag{2}$$

where ξ_i denotes an additional "inflow" including water in dams, rainfall, etc. In the case of S_i being a dam (or a system of dams), the corresponding consumption is assumed to be

$$x_i = w_i - \Delta w_i \quad ,$$

where w_i signifies the water released from S_i for S_{i+1} during the interval $[t_i, t_i + \Delta]$.

In some way we will find a reasonable or even optimal water distribution $x = (x_1, \ldots, x_m)$ under conditions of water resources shortage.

2. Suppose S_i is a region with a number of users $j = 1, \ldots, n$.

The water amount x_i has to be distributed amongst them. Let x_{ij} be the water amounts supplying users $j = 1, \ldots, n$ in response to their demands $\tilde{x}_{ij}$. Say, $\tilde{x}_{ij}$ means the water amount for a normal operation of the corresponding user, and we are given a proper estimate of a loss

$$f_{ij}[x_{ij},\tilde{x}_{ij}]$$

which takes place if we, during the period $[t_i,t_i + \Delta]$, supply $x_{ij} < \tilde{x}_{ij}$ instead of $\tilde{x}_{ij}$ (which actually happens under resources shortage).

There are constraints of the following type:

$$x_{ij} \geq a_{ij} \quad , \quad \sum_{j=1}^{n} x_{ij} = x_i \quad , \qquad (3)$$

where a_{ij} are minimal demands and x_i is the total available water resources. It seems reasonable to distribute the water amount x_i in such a way to minimize the total loss

$$\sum_{j=1}^{n} f_{ij}[x_{ij},\tilde{x}_{ij}]$$

under the constraints (3). If we solve this traditional dynamic programming problem for a number of different values x_i, we could estimate the minimal loss

$$f_i[x_i,\tilde{x}_i] = \min \sum_{j=1}^{n} f_{ij}[x_{ij},\tilde{x}_{ij}]$$

as a function of the total demand

$$\tilde{x}_i = \sum_{j=1}^{n} \tilde{x}_{ij} \quad ,$$

and the actual water supply

$$x_i = \sum_{j=1}^{n} x_{ij} \quad .$$

3. Our first goal is to minimize the total loss

$$f(x) = \sum_{i=1}^{n} f_i(x_i, \tilde{x}_i)$$

as a function of $x = \{x_i\}$ under the constraints

$$a_i \leq x_i \leq \min(\tilde{x}_i, w_i) \tag{4}$$

which are assumed to be feasible.

We have also assumed the variables $w_o, \xi_1, \ldots, \xi_n$ are known. We verify $x_i = w_i$ later in a case where S_i is a dam and Δw_i is the corresponding water release for S_{i+1}.

Let us assume that

$$f_i(x_i, \tilde{x}_i) = f_i(\tilde{x}_i - x_i)$$

are convex functions of the corresponding water deficit $y = \tilde{x}_i - x_i$ which are given with a proper "broken line" approximation:

$$f_i(y) = \sum_{j=1}^{k(y)} \lambda_{ij} \tilde{x}_{ij} + \lambda_{ij(y)} \left[y - \sum_{j=1}^{k(y)} \tilde{x}_{ij} \right] \tag{5}$$

where

$$k(y) = \max k \; : \; \sum_{j=1}^{k} \tilde{x}_{ij} \leq y \quad ,$$

and λ_{ij}, x_{ij} are some non-negative constants (here

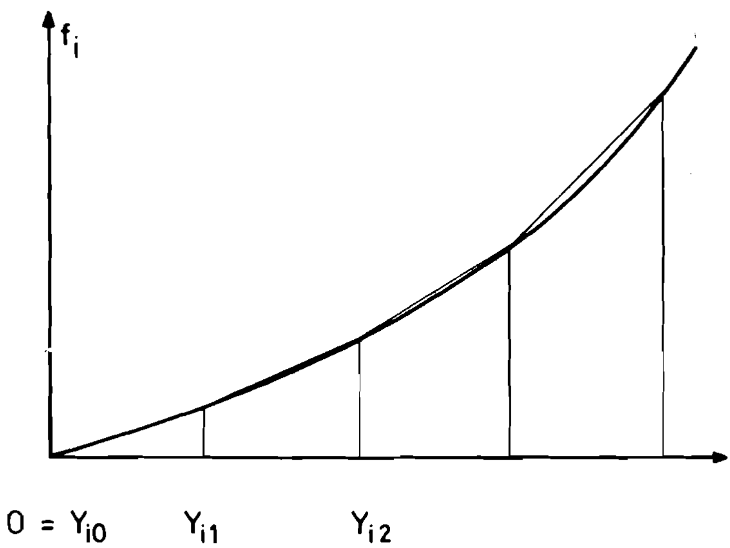

FIGURE 1

$x_{ij} = Y_{ij} - Y_{ij-1}$ where Y_{ij} means the points where linear approximations are broken (see Figure 1) although it may have no connection with water demands in the region S_i which was considered above).

Let us make the substitution

$$x_i - a_i \to x_i, \tilde{x}_i - a_i \to \tilde{x}_i, w_i - a_i \to w_i \quad .$$

Obviously the problem is one concerned with minimization of

$$\sum_{i=1}^{m} f_i(\tilde{x}_i - x_i)$$

as a function of $\{x_i\}$.

$$0 \leq x_i \leq \min\ (\tilde{x}_i, w_i)$$

is equivalent to the minimization of a function

$$\sum_{i=1}^{m} \sum_{j=1}^{k_i} \lambda_{ij}(\tilde{x}_{ij} - x_{ij})$$

where $k_i = k(\tilde{x}_i)$ and

$$\sum_{j=1}^{k_i} x_{ij} = x_i (x_{ij} \geq 0)$$

because under these last constraints,

$$\min \sum_{j=1}^{k_i} \lambda_{ij}(\tilde{x}_{ij} - x_{ij}) = f_i(\tilde{x}_i - x_i)$$

The corresponding minimum points x_{ij}^{o} give us the optimal water distribution as

$$x_i^o = \sum_{j=1}^{k_i} x_{ij}^o \quad . \tag{6}$$

Formally it is possible to consider additional water ξ_{ij} with respect to all demands $\tilde{x}_{ij}$.

In our particular case, it is convenient to order these demands in such a way that x_{ij} precedes x_{pq} if $i < p$ or (in the case $i = p$) if $j > q$ -- remember that because of loss functions convexity we have $\lambda_{ij} > \lambda_{il}$ if $j > l$ and in this sense, it is more important to meet the demand $\tilde{x}_{ij}$ than $\tilde{x}_{il}$. So we set

$$\xi_{ij} = \begin{cases} \xi_i & \text{if } j = k_i \\ 0 & \text{if } j < k_i \end{cases}$$

and reorder all demands with one index as it was explained above.

4. After that our scheme is the following: to minimize

$$f(x,\xi) = \sum \lambda_i (\tilde{x}_i - x_i)$$

as a function of $x = \{x_i\}$,

$$0 \le x_i \le \min\,(\tilde{x}_i, w_i)$$

where w_i is connected with x_i by relations (1) and (2). Here we consider newly ordered variables $\tilde{x}_i$ and x_i which appeared earlier as $\tilde{x}_{ij}$ and x_{ij}.

Suppose x_i, $i < k$, have already been determined; then the problem is to determine the next water supply x_k.

Let us set

$$X_{kj} = \sum_{k<i\leq j} \xi_i \quad , \quad \tilde{X}_{kj} = \sum_{\substack{k<i\leq j, \\ \lambda_i > \lambda_k}} \tilde{x}_i$$

$$\Delta_{kj} = \tilde{X}_{kj} - X_{kj} \tag{7}$$

$$\Delta_k = \max \{0 \text{ and } \Delta_{kj} \ , \ j > k\} \quad .$$

Theorem.

<u>The optimal water supply x_k^o for the corresponding k^{th} user under the given parameters w_k and ξ_j, $j > k$, is</u>

$$x_k^o = \min \{\tilde{x}_k, \max (0, w_k - \Delta_k) \quad . \tag{8}$$

Proof.

Let λ_{i_p} be the p^{th} number amongst $\lambda_i > \lambda_k$ $(i > k)$.

We have to meet the demand x_{i_1} with expenses of all previous consumers, and in the case of the deficit $\Delta_{ki_1} > 0$, the remaining $\Delta w_k = w_k - x_k^o$ must be not less than $\min w_k, \Delta_{ki_1}$. Particularly if $\Delta_{ki_1} \geq w_k$, then $\Delta_{wk} = w_k$ and $x_k^o = 0$. Suppose that $0 < \Delta_{ki_1} < w_k$; then we can cover a deficit Δ_{ki_1} with the expenses of the k^{th} consumer. We have

$$0 \leq x_k^o \leq w_k - \Delta_{ki_1} \quad .$$

In case the next consumer is of lesser importance than i_1, its deficit is

$$\tilde{x}_{i_2} - \sum_{i_1<i\leq i_2} \xi_i = \tilde{x}_{i_1} + \tilde{x}_{i_2}$$

$$- \sum_{k<i\leq i_2} \xi_i - \Delta_{ki_1}$$

$$= \Delta_{ki_2} - \Delta_{ki_1} \quad ,$$

(remember $\Delta_{ki_1} > 0$) and if

$$\Delta_{ki_2} - \Delta_{ki_1} > 0 \quad ,$$

we must satisfy i_2 with the expenses of the k^{th} user in such a way as to supply $\Delta_{ki_2} - \Delta_{ki_1}$ from the rest $w_k - \Delta_{ki_1}$ so that

$$0 \leq x_k^o \leq \max \left\{0, w_k - \Delta_{ki_1} - \left(\Delta_{ki_2} - \Delta_{ki_1}\right)\right\}$$

$$= \max \left\{0, w_k - \Delta_{ki_2}\right\} \quad .$$

Of course, if $\Delta_{ki_1} < 0$, then the corresponding deficit for i_2 has to be defined as

$$\tilde{x}_2 - \sum_{i_1<i\leq i_2} \xi_i + \Delta_{ki_1} = \Delta_{ki_2} \quad ,$$

and this amount of water has to be subtracted from w_k.

Let us consider a case when i_2 is more important than i_1. If

$$\tilde{x}_{i_2} - \sum_{k<i\leq i_2} \xi_i > 0 \quad ,$$

then the total deficit for both users i_1 and i_2 is

$$\tilde{x}_{i_2} - \sum_{k<i\leq i_2} \xi_i + \tilde{x}_{i_1} = \Delta_{ki_2} \quad ,$$

and it should be covered with the expenses of the k^{th} user so that

$$0 \le x_k^o \le \max\left(0, w_k - \Delta_{ki_2}\right) \quad .$$

If

$$\tilde{x}_{i_2} - \sum_{k<i\le i_2} \xi_i < 0 \quad , \quad \tilde{x}_{i_2} - \sum_{i_1<i\le i_2} \xi_i > 0 \quad ,$$

then the deficit

$$\tilde{x}_{i_2} - \sum_{i_1<i\le i_2} \xi_i$$

should be covered for expenses of i_1; thus, we need an amount of water

$$\Delta_{ki} + \left(\tilde{x}_{i_2} - \sum_{i_1<i\le i_2} \xi_i\right)$$
$$= \left(\tilde{x}_i - \sum_{k<i\le i} \xi_i\right) + \left(\tilde{x}_{i_2} - \sum_{i_1<i\le i_2} \xi_i\right)$$
$$= \Delta_{ki_2} \quad .$$

We see that in any case a total deficit for both users i_1 and i_2 which takes place under the additional resources $\sum_{k<i\le i_2} \xi_i$ is $\max\left(0, \Delta_{ki_1}, \Delta_{ki_2}\right)$ so that

$$0 \le x_k^o \le \max\left\{0, w_k - \max\left(0, \Delta_{ki_1}, \Delta_{ki_2}\right)\right\} \quad .$$

If we assume it holds true for any p, $p \le m - 1$, then in a similar manner we can obtain the corresponding deficit for m users $i_1, \ldots, i_m$ which is

$$\max\left(0,\Delta_{ki_1},\dots,\Delta_{ki_m}\right) \quad .$$

Specifically let us set $\Delta_{kk} = 0$ and suppose that

$$\max\left(\Delta_{kk},\Delta_{ki_1},\dots,\Delta_{ki_{m-1}}\right) = \Delta_{ki_p} \quad .$$

Then

$$\Delta_{ki_{m-1}} = \Delta_{ki_p} + \sum_{\substack{i_p<i\leq i_{m-1}\\ \lambda_i>\lambda_k}} \tilde{x}_i - \sum_{i_p<i\leq i_{m-1}} \xi_i \leq \Delta_k \quad ,$$

and in any case

$$\sum_{i_p<i\leq i_{m-1}} \xi_i - \sum_{\substack{i_p<i\leq i_{m-1}\\ \lambda_i>\lambda_k}} \tilde{x} > 0$$

is the available water for i_m; so if

$$\tilde{x}_{i_m} - \xi_{i_m} - \sum_{i_p<i<i_{m-1}} \xi_i + \sum_{\substack{i_p<i\leq i_{m-1}\\ \lambda_i>\lambda_k}} \tilde{x}_i \leq 0 \quad ,$$

then, as before, the total deficit is Δ_{ki_p}.

If

$$\tilde{x}_{i_m} - \xi_{i_m} - \sum_{i_p<i\leq i_{m-1}} \xi_i + \sum_{\substack{i_p<i\leq i_{m-1}\\ \lambda_i>\lambda_k}} \tilde{x}_i > 0 \quad ,$$

Then the total deficit is

$$\Delta_{kp} + \sum_{\substack{p<i\le i_m \\ \lambda_i>\lambda_k}} x_i - \sum_{p<i\le i_m} \xi_i = \Delta_{ki_m} \quad .$$

Thus

$$0 \le x_k^o \le \max \left\{0, w_k - \max \left(0, \Delta_{ki_1}, \ldots, \Delta_{ki_m}\right)\right\}$$

so that Equation (8) holds true.

5. This water distribution $\{x_k^o\}$ is optimal for general schemes with ordered demands $\tilde{x}_i$. It is not necessary to know your loss exactly; the only thing you need to know is the demands $\tilde{x}_i$ that you have to meet after $\tilde{x}_k$, which are more important than $\tilde{x}_k$ itself.

Note that if the scheme is chosen in such a way that demands $\tilde{x}_k$ are approximately of the same size as the consumed water unit, then we can simplify the water supply and set

$$x_k^o = \begin{cases} \tilde{x}_k, w_k - \Delta_k \ge \tilde{x}_k \\ 0, \; w_k - \Delta_k < \tilde{x}_k \end{cases} \quad . \tag{9}$$

6. One can say that we considered above only one arc of the whole river basin net which has a tree structure and may be quite complicated (see Figure 2).

Let us divide this tree and examine it from the very top (a) to the corresponding ends (d_1, d_4, e_2; f_1, f_2; f_3, f_4, f_5, f_6).

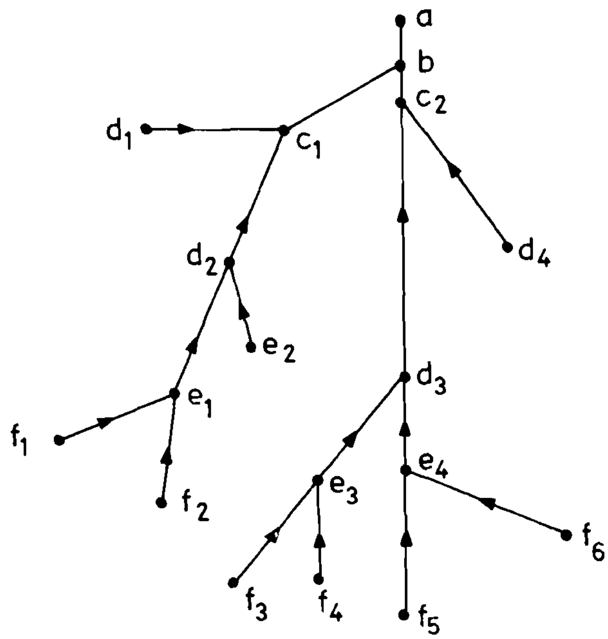

FIGURE 2

Suppose at the end (.) of each branch we are given a total inflow w(.) and the problem is to determine a water supply x(.) for the corresponding user situated at point (.).

In order to avoid some difficulties with notation let us consider the river-basin net which is represented in Figure 2 and the corresponding supply $x(f_6)$ for the point (f_6).

Let us consider that (f_6,e_4) are open interval, $[f_5,e_4)$, $[f_4,e_3),\ldots,$ are semi-open and $[b,a]$ are closed.

Let us treat an arbitrary interval Γ with demands $\tilde{x}_i$ and additional water resources ξ_i; $i = 1,\ldots,n$, in the same way as earlier. Namely, let us set

$$X_k(\Gamma) = \sum_{i \le k} \xi_i, \tilde{X}_k(\Gamma) = \sum_{\substack{i \le k \\ \lambda_i > \lambda}} \tilde{x}_i$$

$$\Delta_k(\Gamma) = \tilde{X}_k(\Gamma) - X_k(\Gamma) \qquad (10)$$

$$\Delta(\Gamma) = \max \{0 \text{ and all } \Delta_k(\Gamma)\} \quad ,$$

and also let us determine

$$\xi(\Gamma) = \sum_{k_m < j \le n} \xi_j - \sum_{k_m < j \le n} \tilde{x}_j \qquad (11)$$

where λ is the loss coefficient for the user considered (f_6) with the demand $\tilde{x} = \tilde{x}(f_6)$ and k_m being the maximum point: $\Delta(\Gamma) = \Delta_{k_m}(\Gamma)$.

As was shown before, we have, first of all, to meet the demands $\tilde{x}_i$ which are more important than $\tilde{x}$. In particular, we must supply nothing for the user (f_6) if

$$\Delta(f_6,e_4) \ge w(f_6) \quad .$$

Thus, in this case $x^o(f_6) = 0$.

Remember that $\Delta(\Gamma)$ means the water deficit along the corresponding arc Γ and in the case $\Delta(\Gamma) = \Delta_{k_m}(\Gamma)$, there is no deficit after the point k_m. More precisely, all demands $\tilde{x}_j$, $k_m < j \le n$, are covered with the corresponding additional water resources ξ_j, $k_m < j \le n$, and the variable $\xi(\Gamma)$, determined by the Equation (11), means the rest of the total resources $\sum_{k_m < j \le n} \xi_j$ which come at the end of the considered arc Γ. Thus for the arc $\Gamma = (f_5,e_4)$ the variable

$$\xi(e_4) = \max\{0, w(f_5) - \Delta[f_5,e_4)\} + [\xi\ f_5,e_4) \tag{12}$$

means the water amount which comes to the knot (e_4) through the arc $[f_5,e_4)$, if we attempt to meet only those demands $\tilde{x}_j$ at $[f_5,e_4)$ that were more important than $\tilde{x} = \tilde{x}(f_6)$.

In a similar way let us set

$$\xi(d_3) = \max\{0, w(e_3) - \Delta[e_3,d_3)\} + \xi[e_3,d_3) \tag{13}$$

where $w(e_3)$ is defined as

$$w(e_3) = \max\{0, w(f_3) - \Delta[f_3,e_3)\} + \xi[f_3,e_3)$$

$$+ \max\{0, w(f_4) - \Delta[f_4,e_3)\} + \xi[f_4,e_3) \quad ,$$

and also

$$\xi(c_2) = \max\{0, w(d_4) - \Delta[d_4,c_2)\} + \xi[d_4,c_2) \quad , \tag{14}$$

$$\xi(b) = \max\{0, w(c_1) - \Delta[c_1,b)\} + \xi[c_1,b) \tag{15}$$

where the variable $w(c_1)$ is defined as

$$w(c_1) = \max\{0, w(d_1) - \Delta[d_1,c_1)\} + \xi[d_1,c_1)$$

$$+ \max\{0, w(d_2) - \Delta[d_2,c_1)\} + \xi[d_2,c_1) \quad ,$$

and in its turn

$$w(d_2) = \max\{0, w(e_1) - \Delta[e_1,d_2)\} + \xi[e_1,d_2)$$

$$+ \max\{0, w(e_2) - \Delta[e_2,d_2)\} + \xi[e_2,d_2) \quad ,$$

$$w(e_1) = \max\{0, w(f_1) - \Delta[f_1, e_1)\} + \xi[f_1, e_1)$$

$$+ \max\{0, w(f_2) - \Delta[f_2, e_1)\} + \xi[f_2, e_1)$$

The variables (12) through (15) mean the additional water resources at the corresponding knots (e_4), (d_3), (c_2) and (b), which are available if we attempt to meet only those demands that were more important than $\tilde{x}(f_6)$. Thus the actual water deficit along the branch

$$(f_6) \rightarrow (e_4) \rightarrow (d_3) \rightarrow (c_2) \rightarrow (b) \rightarrow (a)$$

with respect to the demands that are more important than $x(f_6)$, can be determined as was done previously -- see (7) and (10):

$$\Delta(f_6, a) = \max\{0 \text{ and all } \Delta_k(f_6, a)\} \tag{16}$$

where

$$\Delta_k(f_6, a) = \sum_{\substack{i \leq k \\ \lambda_i > \lambda}} \tilde{x}_i - \sum_{i \leq k} \eta_i$$

$$\eta_i = \begin{cases} \xi_i \;, & i \neq (.) \\ \xi_i + \xi(.) \;, & i = (.) \end{cases}$$

$$(.) = (e_4), (d_3), (c_2), (b) \;.$$

According to the general Equation (8), <u>the optimal water supply $x^o = x^o(f_6)$ at the point (f_6) is</u>

$$x^o = \min\left[\tilde{x}, \max\{w(f_6) - \Delta(f_6, a)\}\right] \;. \tag{17}$$

7. There is difficulty in using our (or any other) programming model for water supply optimization. Namely, future demands and available water resources are not known exactly in our scheme, so we have at each step of our decision making to estimate corresponding parameters $\xi_i, \tilde{x}_i$.

We suggest treating the proper parameters ξ_i, $i = 1,2,\ldots,n$ as random variables with a corresponding conditional probability distribution $P(./W)$ which depends on the current information $\{W\}$ concerning the additional water resources $\{\xi_i\}$.

We also suggest considering the demands $x_i,\ldots,\tilde{x}_n$ during the fairly short time cycle Δ as known constants.

Remember that in our scheme (see Figure 2) the problem is to choose the current water supply $x = x(f_6)$.

Let $x = \{x_i\}$ be our decision concerning components x_i which are functions of the current information about available water resources.

Let $f(x,\xi)$ be the actual loss due to the water shortage which takes place under the corresponding parameters $\xi = \{\xi_i\}$ with respect to our choice of $x = \{x_i\}$.

We suggest considering only such types of decision making for which the loss $f(x,\xi)$ becomes less if the additional water resources are increased; more precisely, the loss $f(x,\xi)$ is a monotone decreasing function of each component ξ_i. Note that the operation we suggest below has exactly this property.

Suppose we are given some reliable lower estimates $\underline{\xi}_i$ of the unknown (random) variables ξ_i such that

$$P\ \{\xi_i \geq \underline{\xi}_i \quad ; \quad 1 \leq i < n\} > 1 - \varepsilon$$

where ε is small enough.[1]

Remember that $P(./W)$ is the conditional probability distribution with respect to the current information $\{W\}$ about $\xi = \{\xi_i\}$ so the estimates $\underline{\xi}_i$; $i = 1,\ldots,n$ are functions of the information $\{W\}$.

If we agree to take the risk with the probability $\sum$ then the following criterion for the choice of x looks quite reliable: the optimal water supply $x = x^o$ is the minimum point of the maximum loss

$$g(x) = \max_{\xi_i > \underline{\xi}_i} f(\xi,x) \quad ,$$

namely

$$g(x^o) = \min g(x) \quad .$$

Obviously in our case because of the assumption concerning $f(\xi,x)$

$$g(x) = f(\underline{\xi},x) \quad , \quad \underline{\xi} = \{\underline{\xi}_i\} \quad ,$$

i.e. the vector function $x^o = x_i^o$ means water supply which gives us the minimal loss $f(\underline{\xi},x)$ under the assumption that $\underline{\xi} = \underline{\xi}_i$ are the actual water resources.

With these additional water resources $\underline{\xi}_i$ the

---

[1]In order to simplify the choice of $\underline{\xi}_i$ one can determine $\underline{\xi}_i$ from the relation

$$\sum_{i=1}^{n} P\{\xi_i < \underline{\xi}_i\} \leq \varepsilon \ .$$

corresponding minimum point (under the best operation) according to equation (17) is

$$x_i^o = \min \{\tilde{x}_i, \max (w_1 - \bar{\Delta})\} \quad , \tag{18}$$

where $\bar{\Delta}$ can be obtained with the substitution of $\xi_i \to \underline{\xi}_i$ from Equation (16).

The similar procedure is suggested at the next step. One has only to use the proper conditional probability distribution for $\xi_i; i = 2,\ldots,n$.

With this procedure the current water supply is determined depending on the corresponding deficit (see Equation (17)). So it is not necessary to know the probability distribution of $\{\xi_i\}$ in order to determine the upper estimate $\bar{\Delta}$ of the actual deficit Δ. For instance, if we know the probability distribution P of Δ then we can determine $\bar{\Delta}$ from the relation

$$P \{\Delta \le \bar{\Delta}\} > 1 - \varepsilon \quad .$$

8. The suggested procedure of decision making implies computation of **the conditional distributions** $P_\xi(\cdot/w)$ of the water resources $\xi = \{\xi_i\}$, $i = 1,\ldots,n$ during the future under the corresponding water data $w = \{w_j\}$. Remember it was suggested that the reliable estimates $\underline{\xi} = \{\underline{\xi}_i\}$ are such that

$$P\{\xi_i > \xi_i \ ; \quad i = 1,\ldots,n/w\} > 1 - \varepsilon \quad .$$

A collection of water data $w(t) = \{w_1(t),\ldots,w_n(t)\}$ for a water basin may be considered as a multivariate random process. The usual tool of its analysis is based on the corresponding mean value vector function

$$A(t) = Ew(t) = \{A_1(t),\ldots,A_n(t)\} \quad ,$$

and correlation matrix function

$$B(s,t) = D^{\frac{1}{2}}(s)\ R(s,t)\ D^{\frac{1}{2}}(t)^{*} \quad ,$$

where

$$D^{\frac{1}{2}}(t) = \{D^{\frac{1}{2}}(t),\ldots,D_n^{\frac{1}{2}}(t)\}$$

is a vector of standard deviations

$$D_k^{\frac{1}{2}}(t) = \mathrm{Var}\ W_k(t)^{\frac{1}{2}} \quad , \quad k = 1,\ldots,n \quad ,$$

and

$$R(s,t) = R_{kj}(s,t)$$

is a matrix of correlation coefficients

$$R_{kj}\ (s,t) = \frac{E[w_k(s) - A_k(t)]\ [w_j(t) - A_j(t)]}{D_k^{\frac{1}{2}}(s)\ D_j^{\frac{1}{2}}(t)}$$

Concerning various inputs $w_1(t),\ldots,w_n(t)$ of WR systems which are water streams, levels of water reservoirs, etc., it seems reasonable to assume that these components are <u>positively</u> correlated:

$$R_{kj}(s,t) \geq 0$$

because their increasing (or decreasing) usually occurs for the same reason -- snow melting, rain, drought, etc. Increasing (or decreasing) of some components occurs with the same phenomena as for other components, and a similar connection takes place in time.

We are not going to discuss in detail a structure of functions $A(t)$, $D(t)$ or $R(s,t)$, but note that usually $A(t)$, $D(t)$ are assumed to be periodic functions (with seasonal periods), and

$$R(s,t) = R(t - s)$$

is assumed to be a correlation function of the multivariate stationary Markov type random process (multidimensional auto-regression model).

Concerning probability distributions for $w_1(t),\ldots,w_n(t)$, one usually assumes that each component has a proper <u>gamma distribution</u>.

Now the following problem arises: what type of <u>multi-dimensional distribution</u> for the vector input

$$w(t) = \{w_1(t),\ldots,w_n(t)\}$$

is consistent with all the properties mentioned above, namely with the given positive correlation coefficients $R_{kj}(s,t)$ and marginal gamma distributions?

We suggest considering some kind of multidimensional gamma distribution which is completely determined with the corresponding parameters $A(t)$, $D(t)$ and $R(s,t)$.

We prefer to describe this multivariate distribution in a way that is convenient for actual modelling (for synthetic hydrology).

Let

$$\{\zeta_{ik}(t) \quad , \quad i = 1,\dots,m_k \quad , \quad k = 1,\dots,n\}$$

be a series of independent standard Gaussian variables. With well known linear methods we can obtain identically distributed Gaussian processes

$$\eta_{ik}(t) \quad , \quad i = 1,\dots,m_k \quad ,$$

(independent for different $i = 1,\dots,m_k$) such that

$$E\eta_{ik}(t) = 0 \quad , \quad \mathrm{Var}\ \eta_{ik}(t) = \sigma_k(t)^2$$

and

$$E\eta_{ik}(s)\ \eta_{ij}(t) = r_{kj}(s,t) \geq 0 \quad .$$

Let us consider

$$w_k(t) = \sum_{i=1}^{m_k} \eta_{ik}^2(t) \quad , \quad k = 1,\dots,n \quad . \tag{18}$$

We have

$$Ew_k(t) = m_k E\eta_{1k}^2(t) = m_k \sigma_k(t)^2 \quad ,$$

$$\mathrm{Var}\ w_k(t) = m_k\ \mathrm{Var}\ \eta_{1k}^2(t) = 2m_k \sigma_k(t)^4 \quad ,$$

and

$$E[w_k(s) - Ew_k(s)]\ [w_j(t) - Ew_j(t)]$$

$$= \min\ (m_k, m_j)\ E[\eta_{1k}(s)^2 - \sigma_k(s)^2]\ [\eta_{1j}(t)^2 - \sigma_j(t)^2]$$

$$= 2 \min\ (m_k, m_j)\ r_{kj}(s,t)^2 \quad .$$

Thus if we set

$$2m_k \sigma_k(t)^4 = D_k(t) \quad ,$$

$$\frac{\min\ (m_k, m_j)}{\sqrt{2m_k m_j\ \sigma_k(s)^2\ \sigma_j(t)^2}}\ r_{kj}(s,t)^2 = R_{kj}(s,t) \quad ,$$

then the multivariate random process with components

$$w_k(t) + A_k(t) - m_k \sigma_k(t)^2 \quad , \quad k = 1,\ldots,n \quad ,$$

has the given parameters

$$A_k(t) \quad , \quad D_k(t) \quad \text{and} \quad R_{kj}(s,t) \quad , \quad j = 1,\ldots,n \quad .$$

All marginal distributions are gamma distributions;[2] namely, the probability density of the variable $w_k(t)$ is

$$f(x) = \frac{x^{\beta-1} e^{-x/\alpha}}{\alpha^{\beta}\Gamma(\beta)} \quad , \quad x > 0 \quad , \qquad (19)$$

---

[2] More general multidimensional gamma-type distributions were suggested by D.R. Krishnaiah and H.M. Rao: "Remarks on a Multivariate Gamma-distribution," Amer. Math. Monthly, Vol. 68, No. 4, 1961; see also U.U. Siddiqui: "Some Properties of Empirical Distribution Function of a Random Process," J. of Research of National Bureau of Standards, Vol. 65B, No. 2, 1961.

where

$$\alpha = 2\,\sigma_{kk}(t)^2 \quad , \qquad \beta = m_k/2 \quad .$$

9. Let us return to the question concerning the conditional distributions $P_\xi(\cdot/w)$.

Suppose all components ξ_i are the corresponding values of independent random processes, namely

$$\xi_i = \xi_i(t_i) \quad , \qquad i = 1,\ldots,n \quad ,$$

where $\xi_i(t)$ means water resources at the point i during the time interval $[t, t + \Delta]$.

Suppose the corresponding water data collection $w = \{w_i\}$ consists of the values

$$w_i = \xi_i(t) \quad , \qquad i = 1,\ldots,n$$

at the current time -- moment t, $t < t_i$.

Because of our assumption, we have

$$P_\xi(\cdot/w) = \prod_{i=1}^{n} P_{\xi_i}(\cdot/w_i)$$

and the problem of finding the conditional distributions $P_{\xi_i}(\cdot/w_i)$ of separate variables $\xi_i = \xi_i(t_i)$ under the given $w_i = \xi_i(t)$.

Let us consider the i^{th} component as the random process with the distribution of Γ-type described above and assume that the variables $\xi = \xi_i$, $w = w_i$ have the following structure:

$$\xi = \sum_{j=1}^{m} \eta_{1k}^2$$

and

$$W = \sum_{k=1}^{m} \eta_{2k}^2$$

where $\{\eta_{11},\ldots,\eta_{1m}\}$ and $\{\eta_{21},\ldots,\eta_{2m}\}$ are Gaussian vectors with identically distributed independent components with zero mean values.

Suppose we are given $\{\eta_{21},\ldots,\eta_{2m}\}$. The corresponding conditional distribution $P_\xi(\cdot/\eta_{21},\ldots,\eta_{2m}\}$ is similar to the Γ-type. More precisely, under the condition of $\{\eta_{21},\ldots,\eta_{2m}\}$ the vector $\{\eta_{11},\ldots,\eta_{2m}\}$ is distributed as when the components η_{1k}, $k = 1,\ldots,m$ are independent Gaussian variables with the same variants

$$\delta^2 = \text{Var}\ \eta_{1k} = E(\eta_{1k} = a_k)^2 \quad ,$$

and mean values

$$a_k = E\{\eta_{1k}/\eta_{2k}\} = r\eta_{2k}$$

where

$$r = \frac{E\eta_{1k}\eta_{2k}}{6^2} \quad , \quad 6^2 = E\eta_{2k}^2 \quad .$$

Let $P_\xi(\cdot,a)$ be the corresponding (non-central) Γ-distribution of

$$P_\xi(\cdot,a) = P_\xi(\cdot/_{\eta_{21},\ldots,\eta_{2m}}) \quad .$$

(Here $a = \{a_k\}$ is the vector with components $a_k = r\eta_{2k}$).
Obviously,

$$P_\xi(\cdot/w) = E\{P_\xi(\cdot/\eta_{21},\ldots,\eta_{2m})/w\}$$

$$= \int_{S_w} P_\xi(\cdot,a)\, da / \int_{S_w} da$$

where

$$S_w = \left\{ a : \sum_{k=1}^{m} a_k^2 = r^2 w \right\}$$

is the sphere of the radius $r\sqrt{w}$; remember that the probability density of $\eta_{21},\ldots,\eta_{2m}$ is

$$\frac{1}{(2\pi)^{m/2}6^m} \exp\left\{-\frac{1}{26^2}\sum_{k=1}^{m}\eta_{2k}^2\right\} ,$$

so the variables $\eta_{21},\ldots,\eta_{2m}$ are distributed under the condition

$$\sum_{k=1}^{m} \eta_{2k}^2 = w$$

uniformly.

The conditional distribution $P_\xi(\cdot/w)$ is continuous and the reliable estimate $\underline{\xi}$ can be obtained from the relation

$$F(\underline{\xi}/w) = \varepsilon ,$$

where $F(x/w)$ is the corresponding distribution function.

It seems interesting to note that

$$E\{\xi/w\} = r^2 w + m\delta^2 \quad .$$

Of course, there are similar, more sophisticated methods of stochastic approaches. Namely, one can consider general (non-central) Γ-distributed multidimensional random process $\xi(t) = \{\xi_i(t)\}$ with correlated components $\xi_i(t)$; $i = 1,\ldots,n$.

10. We have considered above the actual water consumers S_i during the corresponding time cycles $[t_i, t_i + \Delta]$. As promised, we will now turn to the dams S_i for which one needs to determine the optimal water release Δw_i (see Section 1). Obviously we have to take into account the water demands x_i during not only the time cycle $[t_i, t_i + \Delta]$ but also for the future time, say during sequential cycles.

Let us consider the scheme of Section 6 (see Figure 2) and suppose there are dams at the knots (c_1), (d_3), and (e_1).

By using Equations (10) through (16) with

$$\lambda = 0 \quad ; \quad w(c_1) = w(d_3) = w(e_1) = 0 \quad ,$$

we can determine the proper water deficits during different time cycles $t = 1,2,\ldots,n$. Say they are

$$\Delta_t(c_1,b) \quad , \quad \Delta_t(d_3,b) \quad , \quad \Delta_t(e_1,c_1) \quad , \quad \Delta_t[b,a]$$

for the corresponding arcs (c_1,b) (d_3,b), (e_1,c_1) and $[b,a]$ These deficits have to be covered during the time with water from the reservoirs (c_1), (d_3), and (e_1) according to the demands

$$\tilde{x}_t(c_1) \quad , \quad \tilde{x}_t(d_3) \quad , \quad \tilde{x}_t(e_1)$$

which may be defined as

$$\tilde{x}_t(c_1) = \lambda_t(c_1)\ \Delta_t(c_1,b) + \sigma_t(c_1)\ \Delta_t\ [b,a] \quad ,$$

$$\tilde{x}_t(d_3) = \Delta_t(d_3,b) + \sigma_t(d_3)\ \Delta_t[b,a] \quad ,$$

$$\tilde{x}_t(e_1) = \Delta_t(e_1,c_1) + \lambda_t(e_1)\ \Delta_t(c_1,b) + \sigma_t(e_1)\Delta_t[b,a]$$

where the weight coefficients

$$\lambda_t(c_1) + \lambda_t(e_1) = 1 \quad ,$$

$$\sigma_t(c_1) + \sigma_t(d_3) + \sigma_t(e_1) = 1$$

are chosen with respect to water amounts $w_t(e_1)$, $w_t(d_3)$, and $w_t(e_1)$ in the dams (c_1), (d_3), and (e_1), and the capacities of the corresponding channels, etc.

The problem concerning the corresponding <u>water releases</u>

$$x_t(c_1) \quad , \quad x_t(d_3) \quad , \quad x_t(e_1)$$

is the following.

Let $w_o(\cdot)$ be an initial water amount in a dam $(\cdot)$ and $\xi_t(\cdot)$ denote an additional inflow during the corresponding time cycle t. One has to determine an optimal water release

$x_t(\cdot)$ in such a way as to minimize a loss

$$\sum_t f_t(\tilde{x}_t - x_t) \quad .$$

Here $\tilde{x}_t - x_t$ means a lack of water during the time cycle t, and $f_t(\tilde{x}_t - x_t)$ is the corresponding loss as a function of $\tilde{x}_t - x_t$ which may be determined similarly to (5).

Obviously this scheme is of the same type as was considered earlier, except that here we are dealing with water distribution in time rather than space.