LAURENSON, E.M. and MEIN, R.G. (1983). 'Hydraulic Design of Detention and Balancing Reservoirs.' *Proc.* Inter. Conf. on Hydraulic Aspects of Flood and Flood Control, London, 353-361, BHRA, Bedford, England.

LAURENSON, E.M. and MEIN, R.G. (1990) '*RORB Version 4 Runoff Routing Program: User Manual*.' Dept. of Civil Engineering, Monash University, 186pp.

LAURENSON, E.M. and MEIN, R.G. (1993). 'Unusual Applications of the RORB Program.' Engineering for Hydrology and Water Resources Conf., I.E. Aust., *Nat. Conf. Publ. 93/14*, 125-131.

MEIN, R.G. and APOSTOLIDIS, N. (1993). 'Application of a Simple Hydrologic Model for Sewer Inflow/infiltration.' *Proc.* 6th Inter. Conf. on Urban Storm Drainage. Niagara Falls, Ontario, Canada, 1994-1999.

MEIN, R.G. and LARSON, C.L. (1973) 'Modelling Infiltration during a Steady Rain'. *Water Resources Research*, **9**, 384-394. Also reply to Discussion, **9**, 1478.

MEIN, R.G. and LAURENSON, E.M. (1987). 'Adaptation of a Flood Estimation Model for Use on Urban Areas', in *Topics in Urban Drainage Hydraulics and Hydrology* (B.C. Yen, ed.), *Proc.* Tech. Session D, XXII IAHR Congress, Lausanne, 258-263.

MEIN, R.G., LAURENSON E.M. and McMAHON, T.A. (1974). 'Simple Nonlinear Model for Flood Estimation.' *J. Hyd. Div., ASCE*, **100**, HY11, 1507-1518.

MITCHELL, D.J. and LAURENSON, E.M. (1983). 'Flood Routing in a Complex Flood Plain.' Hydrology and Water Resources Symposium, I.E. Aust., *Nat. Conf. Publ. 83/13*, 315-318.

NASH, J.E. (1960). 'A Unit Hydrograph Study with Particular Reference to British Catchments'. Inst. Civ. Engrs. *Proc.*, **17**, 249-282.

THE INSTITUTION OF ENGINEERS, AUSTRALIA (1987). *Australian Rainfall and Runoff: A Guide to Flood Estimation*.

Chapter 6

TANK MODEL

M. Sugawara

6.1. WHAT IS THE TANK MODEL ?

6.1.1. The Tank Model for Humid Regions

The tank model is a very simple model, composed of four tanks laid vertically in series as shown in Fig. 6.1.

Precipitation is put into the top tank, and evaporation is subtracted from the top tank. If there is no water in the top tank, evaporation is subtracted from the second tank; if there is no water in both the top and the second tank, evaporation is subtracted from the third tank; and so on.

The outputs from the side outlets are the calculated runoffs. The output from the top tank is considered as surface runoff, output from the second tank as intermediate runoff, from the third tank as sub-base runoff and output from the fourth tank as baseflow. This may be considered to correspond to the zonal structure of underground water shown typically in Fig. 6.2.

In spite of its simple outlook, the behavior of the tank model is not so simple. If there is no precipitation for a long time, the top and the second tanks will empty and the tank model will look like Fig. 6.3a or 6.3b. Under such conditions, runoff is stable. In the case of Fig. 6.3a the discharge will decrease very slowly, and in the case of Fig. 6.3b the discharge will be nearly constant.

If there is a comparatively heavy rain of short duration under these conditions, the tank model will move to one of the states shown in Fig. 6.3c and 6.3d. In these cases, a high discharge of short duration will occur before the model returns to the stable state as before. In these cases, most of the discharge is surface runoff from the top tank and there is little or no runoff from the second tank.

If heavy precipitation occurs over a longer period then the tank model will take on the state shown in Fig. 6.3e. When the rain stops, the water in the top tank will run off quickly and the tank model will move to the state shown in Fig. 6.3f. Then, the output from the second tank will decrease slowly, forming the typical downward slope of the hydrograph following a large discharge.

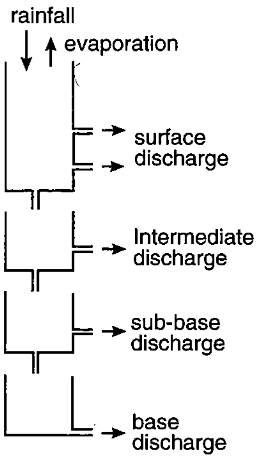


Fig. 6.1.

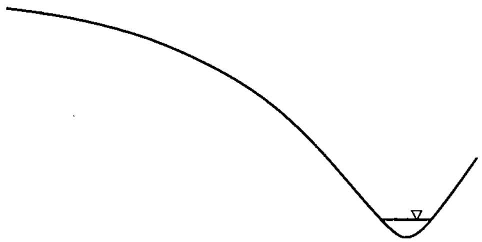


Fig. 6.2.

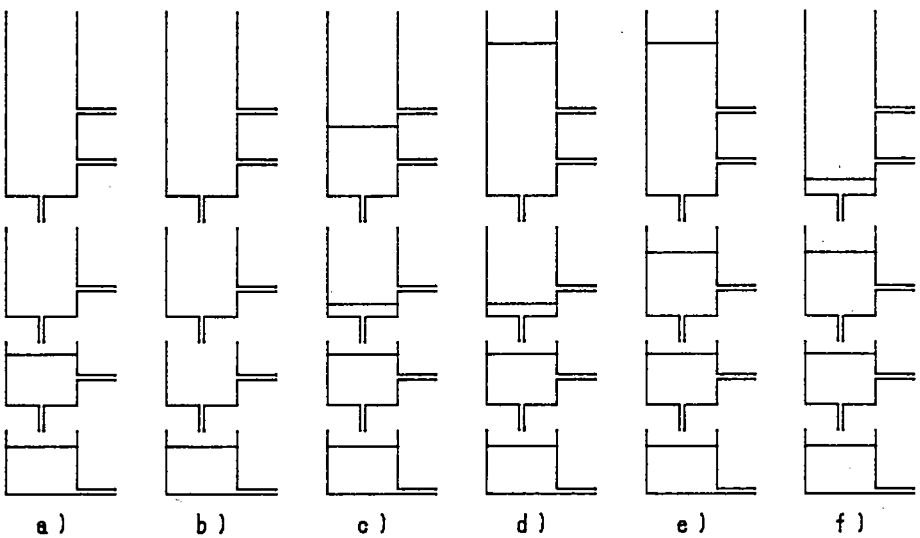


Fig. 6.3.

The tank model can represent such many types of hydrograph because of its non-linear structure caused by setting the side outlets somewhat above the bottom of each tank (except for the lowest tank).

The tank model described above is applied to analyse daily discharge from daily precipitation and evaporation inputs. The concept of initial loss

of precipitation is not necessary, because its effect is included in the non-linear structure of the tank model.

For flood analysis the tank model shown in Fig. 6.4 is applied, where the inputs are usually hourly precipitation and the outputs are hourly discharge. This model contains only two tanks; the third and the fourth tanks are replaced by a constant discharge because the flows from lower tanks form a negligible part of the large flood discharge.

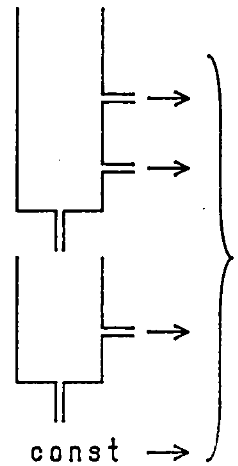


Fig. 6.4.

6.1.2. The Simple Linear Tank

If we move the side outlet(s) of each tank to the bottom of the tanks, we transform the model to one of the linear forms shown in Fig. 6.5a and Fig. 6.5b. Let us first consider a single linear tank of the form shown in Fig. 6.5b with input x(t) and output y(t) (Fig. 6.6). Then, if X(t) is the storage in the tank, the following equations hold:

$$\frac{d}{dt}X(t) = x(t) - y(t), \qquad y(t) = k\,X(t)$$

Therefore

$$\frac{d}{dt}X(t) + k\,X(t) = x(t),$$
$$\left(\frac{d}{dt} = k\right)X(t) = x(t), \quad \left(\frac{d}{dt} + k\right)y(t) = k\,x(t).$$

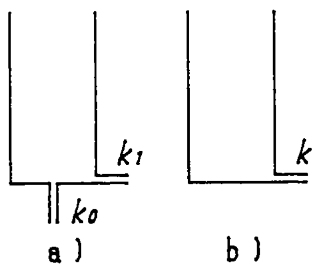


Fig. 6.5.

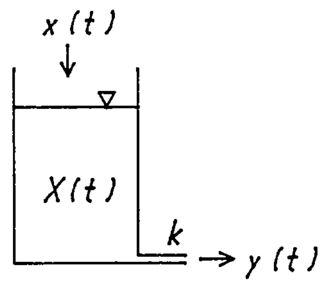


Fig. 6.6.

Or, using of differential operator D = d/dt

$$(D+K)X(t)=x(t), \quad (D+K)y(t)=k\,x(t)$$

$$X(t)=\frac{1}{D+k}x(t), \qquad y(t)=\frac{k}{D+k}x.$$

This shows, that the simple linear tank of Fig. 6.6 is the Heaviside operator k/(D+k), the first order lag system.

If the input at t = 0 to an empty linear tank is the δ function, i.e. x= $\delta(t)$, then the output will be the exponential function y(t) = k exp(-kt), as shown in Fig. 6.7. If we draw a tangent to this exponential curve at t = 0, then it will cut the horizontal axis(t-axis) at T = 1/k, which is the time constant of the operator k/(D+k) for a simple linear tank.

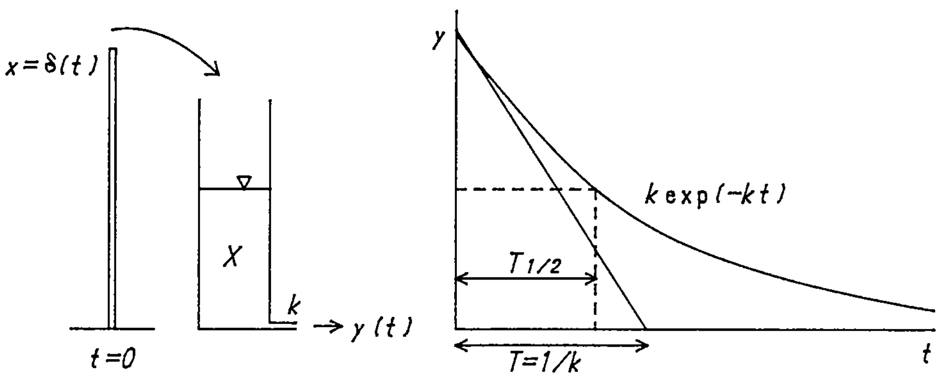


Fig. 6.7.

The meaning of the time constant T = 1/k is as follows: if the output y(t) is maintained at its initial value at t = 0, then the storage in the tank will

disappear after T = 1/k. In reality, the output decreases as the storage decreases and after time T = 1/k the storage will becomes $1/\ell \doteq 0.368$. In practice, the half period $T_{1/2}$ is a rather convenient number as this is the time at which both storage and discharge decrease to half their original values. This is given by:

$$T_{1/2}=T\ln 2=0.6931T=0.7T.$$

Consider now the non-linear tank shown in Fig. 6.8. When the water level is lower than the middle side outlet, discharge is controlled by the linear operator with a time constant T = $1/k_0$. When the water level is between the top and middle side outlets, discharge is controlled by the linear operator with time constant $1/(k_0+k_1)$, and when the water level is above the top outlet, the time constant is $1/(k_0+k_1+k_2)$. Such a tank structure has the property that the time constant becomes shorter as the storage becomes larger and should be a good representation of runoff phenomena.

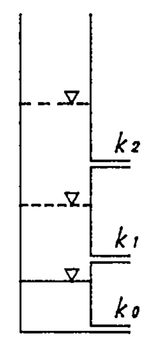


Fig. 6.8.

We can consider the time constant for each tank of the four-tank model by moving the side outlet or outlets to the bottom level. The top tank should have a time constant about a day or a few days, the second tank aweek or ten days or so, the third tank a few months or so, and the fourth tank should have a time constant of years. Roughly speaking, the time constants of the top, the second and the third tanks should have a ratio of $1:5:5^2$ or so. The time constant will also become longer as the catchment area becomes larger and will be proportional to the square root of the catchment area, i.e. to an equivalent diameter of the basin.

Regarding flood runoff, we have an approximate formula that the time constant T(hour) of the flood is given by

$$T=0.15\bullet A^{1/2},$$

where A(km^2) is the catchment area. We think that for flood analysis the time unit (T.U.) should be about one third of the time constant T, i.e.

$$T.U. = 0.15 \bullet A^{1/2}.$$

Using this formula, a basin of 10 km^2 would need precipitation and discharge data at a 10 minute interval, and a basin of 100 km^2 would need half-hourly data. This means that small experimental basins present difficult problems.

The formula $T = 0.15 \bullet A^{1/2}$ was obtained mostly from examples of Japanese basins where the catchment areas are small and the ground surface slope is rather high. Therefore, the coefficient will probably be larger than 0.15 for basins in large continents, perhaps 0.2 ~ 0.3, or so. In the case of the Chang Jiang (the Yangtze River) at Yichang, the catchment area (Sichuang basin) is about half million sq. kms. (5 $10^5 km^2$), and the time constant of the flood is about 10 days (240 hours). If we apply $T = C \bullet A^{1/2}$ to the Yangtze River at Yichang we get

$$C = 240/(5 \bullet 10^5)^{1/2} \doteq 0.34.$$

Even though the difference of the coefficient values between 0.15 and 0.34 is not small, when we consider the very large range of area from 10 to half a million km^2, it seems reasonable to say that the time constant of flood runoff is approximately proportional to the square root of the catchment area.

6.1.3. The Evidence for the Existence of Separated Storage of Groundwater

Despite giving good results for the calculation of river discharge from rainfall in many areas, the tank model has often been considered as only a black box model. However, there is important evidence that can give physical meaning to the tank model.

In Japan many stations measure crustal tilt for earthquake forecasting. The observations are often disturbed by noise, among which the largest is the effect of rain, as shown in Fig. 6.9 and Fig. 6.10. These data were observed in a tunnel at Nakaizu (position 138°59'48.4"E, 34°54'46.4"N; altitude 263m; lithology, tuffaceous sandstone). Figure 6.9 shows daily data and Fig. 6.10 shows hourly data. It is immediately apparent (particularly for the NS-component) that the curves of crustal tilt are very similar to hydrographs. By transforming the coordinate axis, the separation of the two types of tilt curve can be improved but this is not so important and so only original data are shown in the figures.

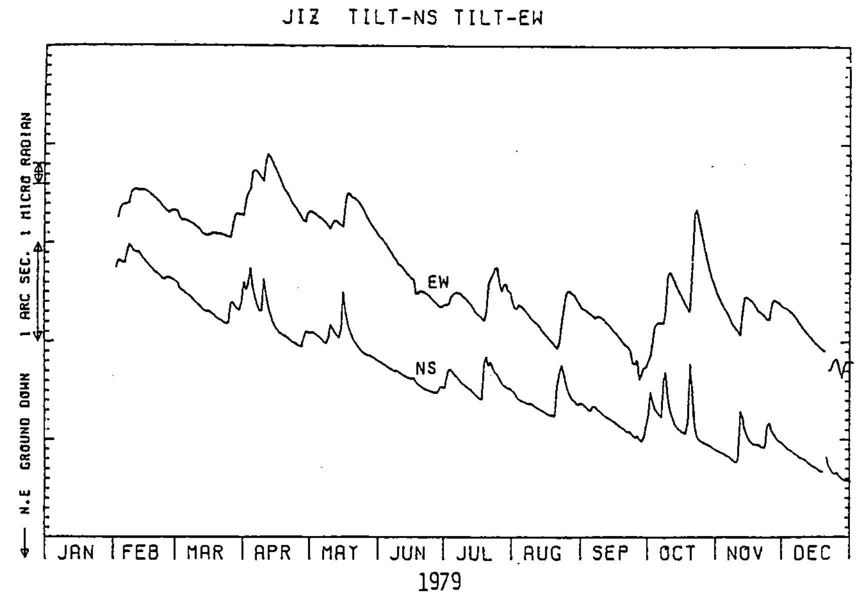
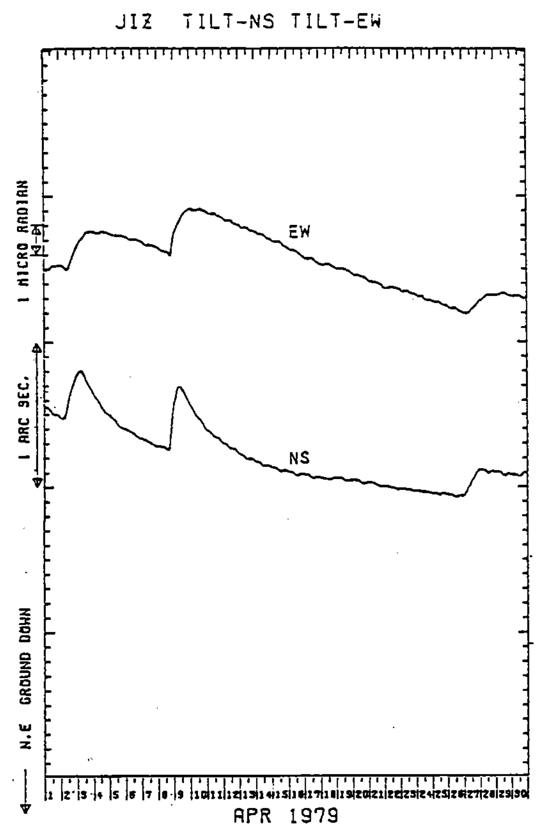


Fig. 6.9. Crustal tilt NS and EW-components at Nakaizu



Using precipitation input we began to use the tank model to simulate the NS-component of the tilt curve. After a while, however, we realized that, this would not work very well because the variation in the output of the tank model was much larger than that of the tilt curve; the discharge peaks were too high when compared with the peaks of the tilt curve. However, it seemed that the tilt curves could be approximated by the storage volumes of the tank model. We can approximate the NS-component by the sum of the storage in the first and the second tank, and the EW-component by the storage of the second tank; where NS means North-South and EW means East-West.

The explanation as to why the crustal tilt curve can be approximated by the storage of the tank model is that crustal tilt is effected by the weight of groundwater storage. In other words, the crust is some sort of spring balance which 'weighs' and so measures the ground water storage (see Fig. 6.11). The first attempt to simulate the crustal tilt curve by the output of the tank model can now be seen to be unreasonable because it assumed that the discharge or mass of water in the river channel effected the crustal tilt. However, the mass of water in the river channel is negligibe in comparison with the groundwater storage.

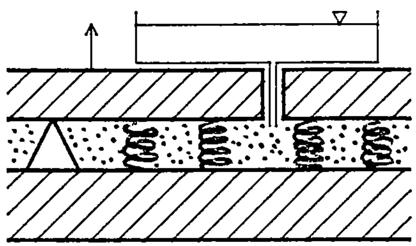


Fig. 6.11.

The main goal for seismologists is to eliminate the noise caused by the effect of rain. For hydrologists, however, the effect of rain on the crustal tilt curves is not noise but an important signal which clearly shows us the existence of two kinds of groundwater storage. There must be some unknown structure which 'weighs' two kinds of storage; storage in the first and the second tank for the NS tilt and storage in only the second tank for the EW tilt. Although the form of this structure is not clearly known, we can imagine such a structure, without any difficulty.

The most important point must be the separation of these two types of storage. For a long time the tank model was only an imaginary model in our minds, but now we can show that it corresponds to a real underground structure.

6.1.4. The Storage Type Model

It is reasonable to consider that runoff and infiltration from the ground surface layer are functions of the storage volume. Moreover, we can imagine that the relationship between runoff and storage must be of an accelerative type and that the relationship between infiltration and storage must be a saturation type, as shown in Fig. 6.12a. Such relationships can be simulated by a tank such as Fig. 6.12b. The two curves in Fig. 6.12a can be approximated by connected segments as shown in Fig. 6.13a and the corresponding functions can be simulated by a tank such as Fig. 6.13b. If we assume that infiltration is proportional to storage then the tank becomes as in Fig. 6.14.

This result lead the author to consider that all runoff phenomena could be simulated with the type of tank shown in Fig. 6.14 together with a simple linear tank for base discharge (see Fig. 6.15a). However, this was not

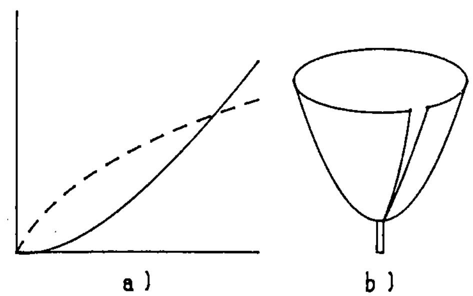


Fig. 6.12.

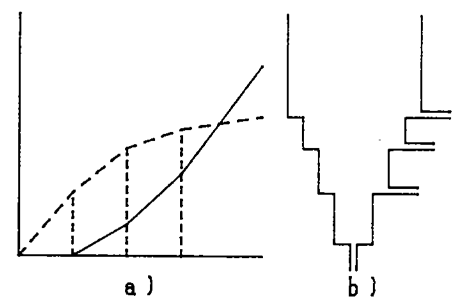


Fig. 6.13.

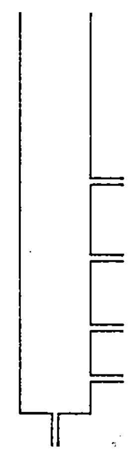


Fig. 6.14.

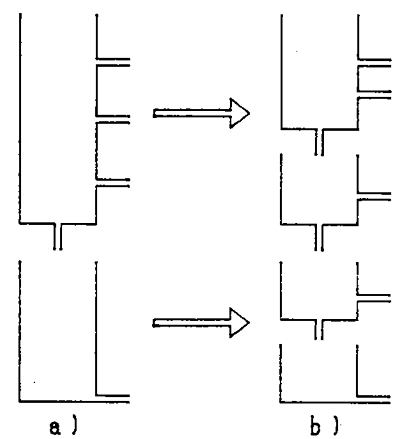


Fig. 6.15.

successful and to improve the performance we had to divide the upper storage type into two parts, one for surface runoff and another for intermediate runoff. Later, the lower tank was also divided into two tanks, one for sub-base runoff with a rather short time constant and one for stable base runoff with a long time constant (Fig. 6.15b).

In some cases, the tank with three side outlets is used as the top tank, and the tank with two side outlets is applied for the second tank. These are some sorts of variety of the usual tank model.

If we set many similar small-diameter evenly-spaced outlets on a tank, as shown in Fig. 6.16a, the relationship between runoff and storage is given by a parabola in the limiting case. In some cases, a tank of this type may be effective and the tank shown in Fig. 6.17a has been used as a top tank. In this case, when the storage is between H_1 and H_2, the relationship between runoff and storage is given by a parabola, and when the storage is larger than H_2, the relationship is given by the tangent of the parabola as shown in Fig. 6.17b.

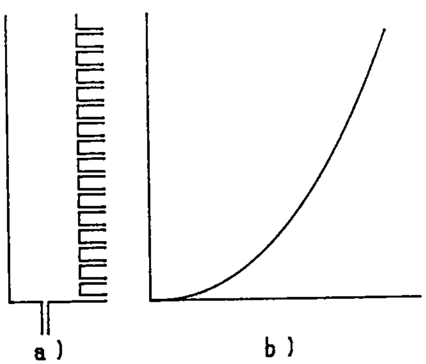


Fig. 6.16.

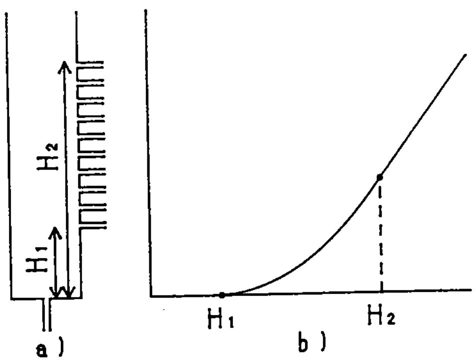


Fig. 6.17.

6.1.5. The Tank Model with Soil Moisture Structure

The ground surface layer is considered to have a soil moisture component. In humid regions without a dry season, such as Japan, the soil moisture is nearly always saturated and so the tank model of Fig. 6.1 can give good results without the need for a soil moisture component.

If, however we wish to consider the effect of soil moisture we must add a structure to the bottom of the top tank as shown in Fig. 6.18a. However, the effect of soil moisture is not so simple as shown in this model. If rain occurs on dry soil the moisture will first fill the space which is easy to occupy, i.e. free air space in the soil. Only then will water transfer gradually into the soil material itself.

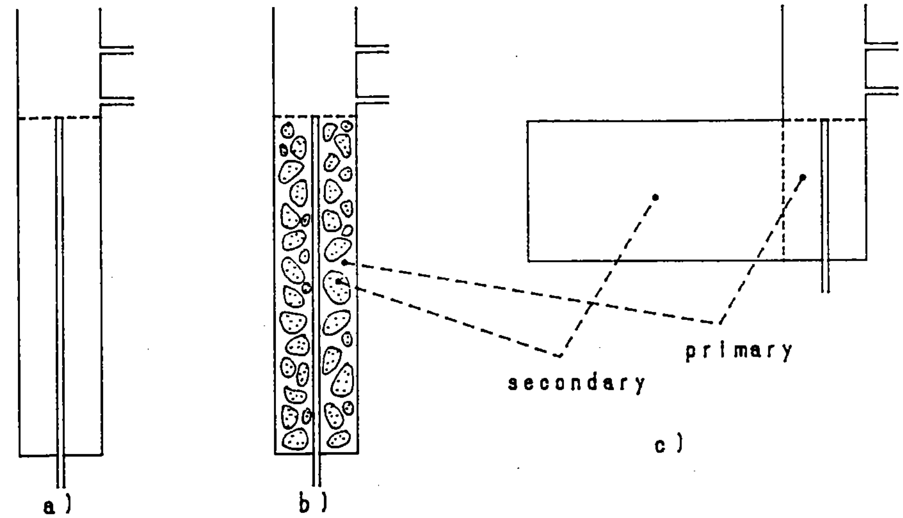


Fig. 6.18.

The first part is called the primary soil moisture and the latter the secondary soil moisture. They are shown symbolically in Fig. 6.18b. As this form of tank is somewhat inconvenient to use, we set the secondary soil moisture on the side of the tank as Fig. 6.18c shows.

The soil moisture structure attached to the tank model (see Fig. 6.19) is as follows :

1) Soil moisture storage has two components, the primary soil moisture storage X_P and the secondary soil moisture storage X_S, where the maximum capacity of each storage is S_1 and S_2, respectively.

2) The primary soil moisture storage and free water in the top tank together make a storage X_A in the top tank. Rainfall is added to X_A and evaporation is subtracted from X_A. When X_A is not greater than S_1, X_A is exclusively primary soil moisture storage and there is no free water, X_F, in the top tank; i.e.

$$X_P = X_A, \quad X_F = 0: \quad \text{when} \quad X_A \leqq S_1.$$

When X_A is larger than S_1, primary soil moisture is saturated and the excess part represents free water in the top tank; i.e.

$$X_P = S_1,$$

$$X_F = X_A - S_1:$$

when $\quad X_A > S_1.$

3) When the primary soil moisture is not saturated, and there is free water in the lower tanks, there is water supply T_1 to the primary soil moisture storage. For T_1 we can write

$$T_1 = K_1(1 - X_P/S_1).$$

4) There is water exchange between the primary and secondary soil moisture storages. This is regulated by the value T_2 where

$$T_2 = K_2(X_P/S_1 - X_S/S_2).$$

When T_2 is positive, it means water is transferred from primary to secondary storage, and vice versa; i.e. water transfers from the wet part to the dry part.

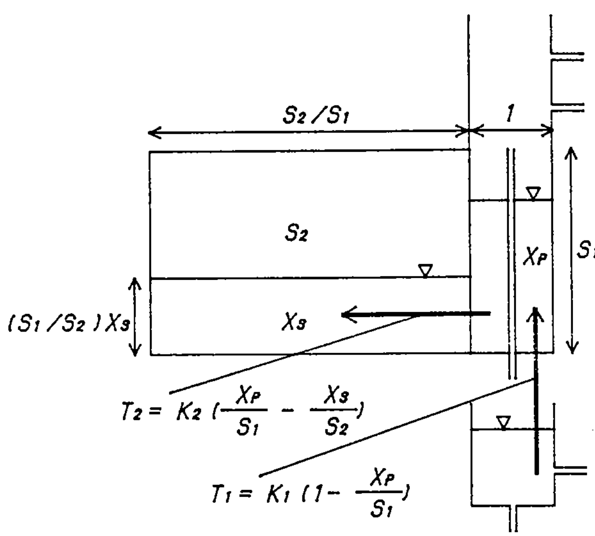


Fig. 6.19.

Now, we shall consider the behavior of the soil moisture storage under the condition that there is no rainfall but constant evaporation, E, there is no free water in the top tank but there is free water in the lower tanks.

Then, we can write the following equations:

$$\frac{dX_p}{dt} = -E + T_1 - T_2$$

$$= -E + K_1\left(1 - \frac{X_p}{S_1}\right) - K_2\left(\frac{X_p}{S_1} - \frac{X_s}{S_2}\right)$$

$$= -E + K_1 - \frac{K_1 + K_2}{S_1} X_p + \frac{K_2}{S_2} X_s,$$

$$\frac{dX_s}{dt} = T_2 = K_2\left(\frac{X_p}{S_1} - \frac{X_s}{S_2}\right)$$

$$= \frac{K_2}{S_1} X_p - \frac{K_2}{S_2} X_s.$$

To make these equations homogeneous, we put $X_P = x_P + c_P$, $X_S = x_S + c_S$.. Then, to eliminate the constant terms, we must satisfy the following equations:

$$-(K_1 + K_2)(c_P / S_1) + K_2 c_S / S_2 = E - K_1$$

$$K_2 c_P / S_1 = K_2 c_S / S_2.$$

By setting $c_P/S_1 = c_S/S_2 = c$, we obtain

$$-(K_1 + K_2)c + K_2 c = E - K_1$$

$$c = 1 - E / K_1.$$

Therefore,

$$c_P = S_1 c = S_1 - (E / K_1) S_1,$$

$$c_S = S_2 c = S_2 - (E / K_1) S_2.$$

The meaning of this result is simple and clear. If we determine new origins c_P and c_S, which are located $(E/K_1)S_1$ below the upper bound, as shown in Fig. 6.20, then the equations describing soil moisture storage can be rewritten in homogeneous form as:

$$\frac{dx_p}{dt} = -\frac{K_1 + K_2}{S_1} x_p + \frac{K_2}{S_2} x_s,$$

$$\frac{dx_s}{dt} = \frac{K_2}{S_1} x_p - \frac{K_2}{S_2} x_s,$$

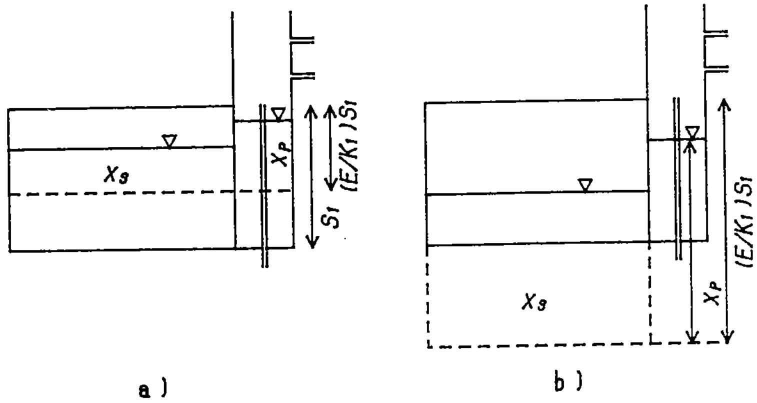


Fig. 6.20.

where x_P and x_S are primary and secondary soil moisture storage measured from the new origins c_S and c_P, respectively. When E is small ($E < K_1$) c_S and c_P lie relatively high, as shown in Fig. 6.20a, and when E is large ($E > K_1$) the new origins are low, as shown in Fig. 6.20b. Seasonal change in the level of these origins will play an important part in the behavior of soil moisture storage, as described later.

To solve the linear homogeneous equations, we put

$$x_p = A\exp(-kt), x_S = B\exp(-kt),$$

and obtain

$$-kA = -\frac{K_1 + K_2}{S_1}A + \frac{K_2}{S_2}B,$$

$$-kB = -\frac{K_2}{S_1}A - \frac{K_2}{S_2}B.$$

To have a meaningful solution other than A = B = 0, these equations must conform to the following characteristic equation:

$$\begin{vmatrix} -\frac{K_1 + K_2}{S_1}k & \frac{K_2}{S_2} \\ \frac{K_2}{S_1} & -\frac{K_2}{S_2} + k \end{vmatrix} = 0$$

$$\left(k - \frac{K_1 + K_2}{S_1}\right)\left(k - \frac{K_2}{S_2}\right) = \frac{K_2{}^2}{S_1 S_2},$$

$$k^2 - \left(\frac{K_1 + K_2}{S_1} + \frac{K_2}{S_2}\right)k + \frac{K_2{}^2}{S_1 S_2} = 0.$$

In our experience, S_2 is much larger than S_1, and K_2 is much larger than K_1. From Fig. 6.21, we can see, that the characteristic equation has two positive roots, k_1, k_2, where the large one satisfies the condition

$$k_1 > \frac{K_1 + K_2}{S_1} \text{ and the small satisfies } 0 < k_2 < \frac{K_2}{S_2}.$$

As k_1, k_2 satisfy equations

$$k_1 + k_2 = \frac{K_1 + K_2}{S_1} + \frac{K_2}{S_2}, \quad k_1 k_2 = \frac{K_1 K_2}{S_1 S_2},$$

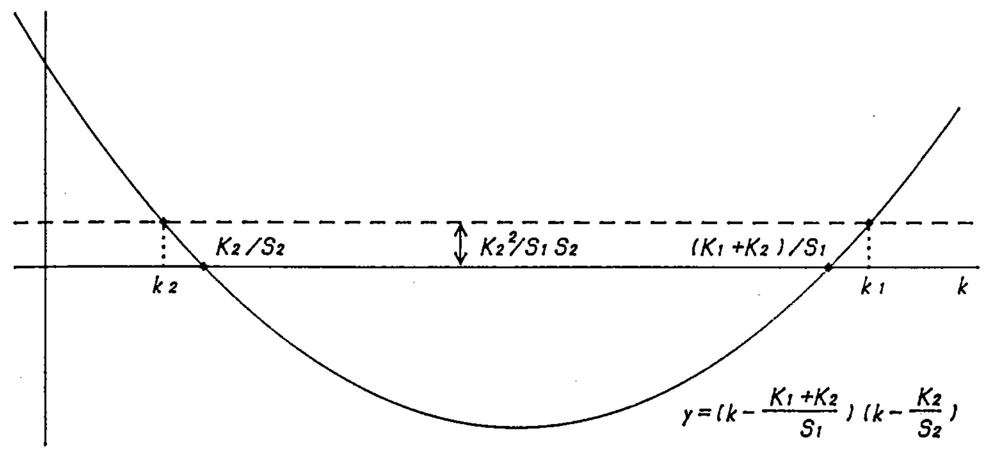


Fig. 6.21.

and k_2 is small, we can obtain the first approximation:

$$k_1{}' = \frac{K_1 + K_2}{S_1} + \frac{K_2}{S_2}, \quad k_2{}' = \frac{K_1 K_2}{S_1 S_2} \bullet \frac{1}{k_1{}'}.$$

Then, the second approximation is

$$k_1{}'' = \frac{K_1 + K_2}{S_1} + \frac{K_2}{S_2} - k_2{}', \quad k_2{}'' = \frac{K_1 K_2}{S_1 S_2} \bullet \frac{1}{k_1{}''}.$$

When $S_1 = 50$, $S_2 = 250$; $K_1 = 2$, $K_2 = 20$, the true values of k_1, k_2, and their first and second approximations are given as follows:

$$k_1 = 0.5137, \qquad k_1' = 0.52, \qquad k_1'' = 0.51385,$$
$$k_2 = 0.006228, \qquad k_2' = 0.00615, \qquad k_2'' = 0.006228.$$

To determine the vector (A,B), corresponding to the large characteristic value k_1, we use the first approximation

$$k_1' = \frac{K_1 + K_2}{S_1} + \frac{K_2}{S_2}$$

in the equation

$$-kA + -\frac{K_1 + K_2}{S_1} A + \frac{K_2}{S_2} B.$$

Then, we can obtain an approximate solution

$$A + B = 0.$$

By putting $A = C_1$ and $B = -C_1$ we can obtain the following approximate solution.

$$x_P = C_1 \exp(-k_1 t), \quad x_S = -C_1 \exp(-k_1 t).$$

The vector (A,B), corresponding to the small characteristic value k_2, which is very small, can be determined, by neglecting k in the equation

$$-kB = \frac{K_2}{S_1} A - \frac{K_2}{S_2} B.$$

Thus, we can obtain an approximate solution

$$\frac{A}{S_1} - \frac{B}{S_2} = 0.$$

If we set $A = S_1 C_2$ and $B = S_2 C_2$, we can obtain the following approximate solution;

$$x_P = C_2 S_1 \exp(-k_2 t), \quad x_S = C_2 S_2 \exp(-k_2 t).$$

Therefore, the general solution is of the form:

$$x_P = C_1 \exp(-k_1 t) + C_2 S_1 \exp(-k_2 t),$$
$$x_S = -C_1 \exp(-k_1 t) + C_2 S_2 \exp(-k_2 t).$$

The meaning of this general solution is simple and clear. If C1 = 0, we are left with the second component with a long time constant $T_2 = 1/k_2$ and both soil moisture storages are represented by

$$\frac{x_p}{S_1} = \frac{x_s}{S_2} = C_2 \exp(-k_2 t).$$

This means that both soil moisture storages have a common water content which decreases exponentially with the long time constant $T_2 = 1/k_2$. The first component is then a deviation from the common component and, itself, decreases exponentially with the short time constant $T_1 = 1/k_1$. We can also write this result in the following form:

$$\frac{x_p + x_s}{S_1 + S_2} = C_2 \exp(-k_2 t),$$

$$\frac{x_p}{S_1} - \frac{x_p + x_s}{S_1 + S_2} = \frac{C_1}{S_1} \exp(-k_1 t),$$

$$\frac{x_s}{S_2} - \frac{x_p + x_s}{S_1 + S_2} = -\frac{C_1}{S_2} \exp(-k_1 t).$$

These equations show that the relative wetness of the total soil moisture $(x_P+x_S)/(S_1+S_2)$ decreases exponentially with the long time constant $T_2 = 1/k_2$ and the deviation of the relative wetness of both primary and secondary soil moisture storages from the relative wetness of the total soil moisture decreases exponentially with the short time constant $T_1 = 1/k_1$.

If we assume that $S_1 = 50$, $S_2 = 250$, $K_1 = 2$ and $K_2 = 20$, both time constants can be estimated as

$$T_1 = 1/k_1 = 1.95(day), \quad T_2 = 1/k_2 = 160.6(day).$$

Now, we have to notice that the points of origin for x_P and x_S are c_P and c_S, respectively, as in Fig. 6.20. On the Pacific Ocean side of Japan, winter is a rather dry season, evaporation E is smaller than K_1, and c_P and c_S values are high. Accordingly, even if the weather is dry, the soil moisture is high because of the high values of c_P and c_S.

On the other hand, under hot and dry conditions, e.g. $S_1 = 50$, $S_2 = 250$, $K_1 = 2$, $K_2 = 20$ and $E = 6$ (mm/day), the soil moisture will vanish when the relative wetness $(x_P+x_S)/(S_1+S_2) = 2/3$, as shown in Fig. 6.22. And as ln 2/3 = -0.4055, even if the soil moisture starts from 100%, it will disappear after 160.6 ·0.4055 = 65.1 days. To take another example, if E = 8 the soil

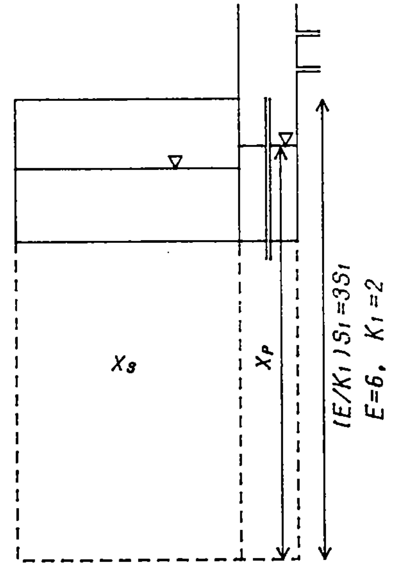


Fig. 6.22.

6.1.6 The Compound Tank Model for Arid Regions

moisture will disappear in 46 days from the saturation date. In conclusion, in spite of the long time constant T_2, the soil moisture will disappear completely in a relatively short period if the potential evaporation is high.

In arid regions where the annual potential evaporation is larger than the annual precipitation, mountainous areas become dry in the dry season, because groundwater moves downward by gravity. On the other hand, low areas along the river remain wet because they receive groundwater from higher areas. When the wet season returns and it rains, surface runoff occurs from the areas along the river, because they are still wet. In the dry mountain areas, however, rain water is absorbed as soil moisture and there is no surface runoff.

During the wet season the percentage of wet area increases with time and surface runoff increases. On the contrary, during the dry season the percentage of dry area increases with time and runoff decreases. As evaporation does not occur from dry areas the actual evaporation from the whole basin is smaller than the potential evaporation. Therefore, even if the annual potential evaporation is larger than the annual precipitation, there is the possibility of runoff.

Following these considerations we divide the basin into zones, for example, into the four zones shown in Fig. 6.23. Let the proportional areas of the zones be AR1, AR2, AR3 and AR4, where AR1+AR2+AR3+AR4 = 1 and apply a tank model of the same structure to each zone as in Fig. 6.24.

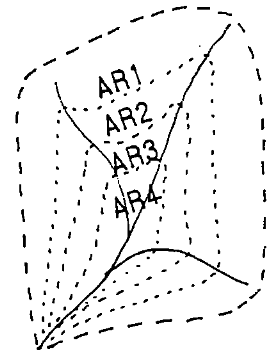


Fig. 6.23.

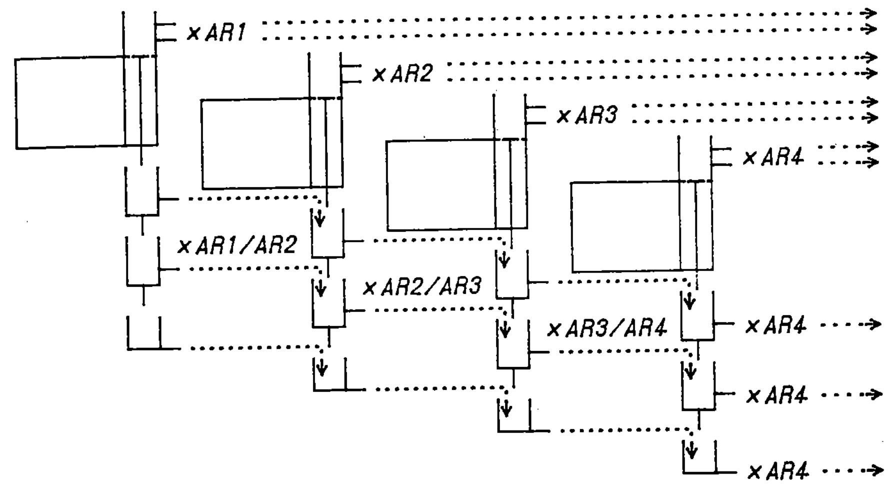


Fig. 6.24.

In this model we assume that, when surface runoff occurs in a zone, the lower zones are already wet and the surface runoff from the top tank of each zone will go directly to the river channel. For the second, third and fourth tanks, however, the output will go to the corresponding tank of the next lower zone, as Fig. 6.24 shows, because each tank corresponds to a groundwater layer (refer back to Fig. 6.2.

Consider the tank model for the first zone. The input to and output from the model (precipitation and evaporation) is measured and calculated in units of water depth, i.e. mm, but to transfer water to the next zone we have to convert to volume units by multiplying the outputs of each tank by AR1, the area of the first zone. The outputs from the tanks of the first zone are transferred as volumes into the corresponding tanks of the second zone and reconverted to depth units by dividing by AR2, the area of the second zone. Therefore, the outputs from the lower tanks of the first zone are multiplied by AR1/AR2 before they are put into the tanks of the second zone. The output from the top tank is multiplied by AR1 and goes directly to the river channel.

In the fourth zone, all outputs go directly to the river channel after being multiplied by AR4. The calculation system is shown in Fig. 6.24. The important factor in this type of tank model is the ratio of the zone areas, AR1 : AR2 : AR3 : AR4. If there are not enough data to determine this ratio, we have to determine it by trial and error method starting from a geometrical progression such as:

1	:	1	:	1	:		=	25%	:	25%	:	25%	:	25%
1.5^3	:	1.5^2	:	1.5	:	1	=	41.5%	:	27.7%	:	18.5%	:	12.3%
2^3	:	2^2	:	2	:	1	=	50%	:	25%	:	12.5%	:	6.25%
3^3	:	3^2	:	3	:	1	=	67.5%	:	22.5%	:	7.5%	:	2.5%
4^3	:	4^2	:	4	:	1	=	75.3%	:	18.8%	:	4.7%	:	1.2%

Roughly speaking, the drier the basin the larger ratio of the progression should be.

6.1.7. Effect of Irrigation Water Intake, Deformation of Hydrograph by Gorge, etc.

Effect of Irrigation Water Intake

In most Japanese river basins, there are wide paddy fields which use large volumes of irrigation water. In many cases, the irrigation water intake is nearly equal to the base discharge of the river because, over a long history, the paddy fields have been developed to the maximum possible extent considering the long-term river water supply. Therefore, if the effect of irrigation water is not considered, any runoff analysis would be meaningless. Fortunately, it is rather simple to calculate the effect of the irrigation water intake. Let the amount of irrigation water intake be Z; then Z is subtracted from the output Y of the tank model, and QE = Y-Z is the calculated discharge. However, irrigation water is not lost but infiltrates from paddy fields and adds to the groundwater, and so Z must be put back into the third tank as shown in Fig. 6.25. Using this simple method, good results have been obtained.

There is no need to subtract evaporation from irrigation water because evaporation has already been subtracted from the top tank. In most cases Z is determined by trial and error considering the local agricultural customs.

Deformation of the Hydrograph by Water Storage or Flooding Caused by Channel Restriction

If enough data are given about the relationship between the volume of stored or flooded water and the water level or discharge in a gorge or channel restriction then we can calculate the deformation of the hydrograph. If there are not enough data, we can estimate the hydrograph deformation using some sort of storage type tank such as shown in Fig. 6.26.

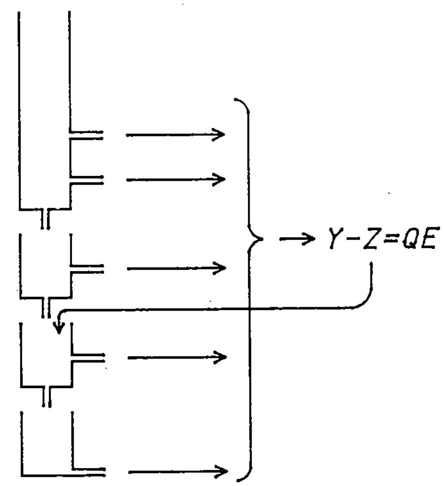


Fig. 6.25.

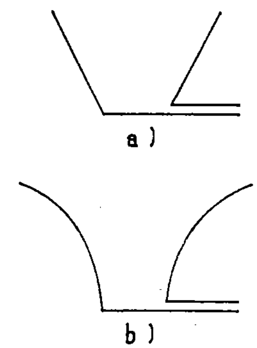


Fig. 6.26.

Water storage or inundation occurs above a restriction because of the limited channel capacity and this condition can be simulated with the storage type tank of Fig. 6.26. Using several tanks of such a type, we can attain a good result by trial and error.

Time Lag

Though the tank model itself can give some sort of time lag (the simple linear tank is equivalent to a kind of linear operator or first order lag) this time lag is often not enough. The output of the tank model Y(J), where J is the day number, may have a time lag when compared to Q(J), the observed discharge. To account for this it is necessary to introduce an artificial lag TL as follows:

$$QE(J) = (1 - D(TL)) \cdot Y(J + [TL]) + D(TL) \cdot Y(J + [TL] + 1),$$

where QE(J) is the calculated discharge, [TL] is the integer part of TL and D(TL) is the decimal part of TL.

As the discharge increases, the velocity also increases and, since time is inversely proportional to velocity, the time lag must decrease. However, since the change in velocity is usually small in comparison with the change in discharge, and the time lag itself is usually not large, we can often assume that the time lag is constant. However, for large basins with large time lags, we may have to consider the time lag as a function of discharge. In the case of the Yangtze River at Yichang the basin area is about half million sq. kms., the travel time from upstream points to Yichang is more than a week, and it is necessary to consider the time lag as the function of discharge.

6.2. HOW TO CALIBRATE THE TANK MODEL ?

6.2.1. Trial and Error Using Subjective Judgement

The tank model is non-linear and mathematics is nearly useless for non-linear problems. Therefore, the author could not use mathematics for the tank model calibration and the only solution was to use the trial and error method of numerical calculation. In 1951 when the author first applied the simple tank model for runoff analysis there were only a few computers in Japan and the author was not able to use one. Without mathematical solutions and with no computer, the numerical calculations necessitated long and hard labour.

However, the human mind is always curious and the labourious numerical calculations became not so boring but rather interesting as experience and judgements were built up in authors brain.

Gradually, calibration of the tank model became rather easy. Usually the first, second and even the third trials will not give good results and so we can make bold and large changes to the model parameters. However, after several trials the result should become fairly good and the fine adjustment of parameters can begin. After ten trials or so the result usually becomes very good.

However, another difficulty appeared. The author found difficulty in describing and explaining his tank model to others. Our language is based on common experiences; if two people have common experiences and knowledges about something, then they can talk and discuss the subject. However, if one person knows nothing about the subject, then it is difficult or impossible to discuss the subject because there is no common vocabulary.

Later, as the tank model became more well-known because of its good results, then the author had to find ways to describe and explain his method of working and his way of thinking. Some of these methods are:

The Initial Model

For the first trial, an initial tank model may be assumed; an example of such an initial model is shown in Fig. 6.27a. Or, we can derive an initial model to fit the basin by plotting the hydrograph in logarithmic scale, finding the peaks of the hydrograph and measuring the descending rates of discharge after the peaks. If the descending rate is r per day then the coefficients of the tank model shown in Fig. 6.27b may be determined as,

$$A0 = A1 = A2 = (1 - r)/3,$$

$$B0 = B1 = A0/5,$$

$$C0 = C1 = B0/5,$$

$$D1 = 0.001.$$

It is better to keep these coefficients as simple numbers, for example if A0 = 0.266 $\cdots$, it is better to make it 0.25 or 0.3.

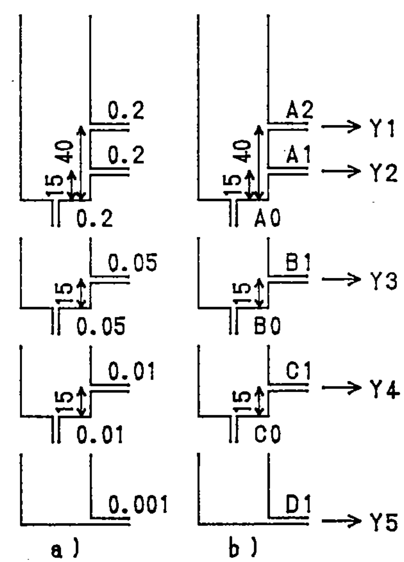


Fig. 6.27.

Find the Worst Point of the Running Model and Adjust the Parameter Values.

The entire hydrograph output of the running model must be plotted five times as Y5, Y4 + Y5, Y3 + Y4 + Y5, Y2 + Y3 + Y4 + Y5 and Y = Y1 + Y2 + Y3 + Y4 + Y5 (see Fig. 6.27b). Comparing the five calculated hydrograph components with the observed hydrograph, we can judge which component is the worst. In early trials, there may be many bad points and we have to select the worst one.

How to Make Some Component Larger or Smaller.

If the worst point is that the runoff from the top tank is too small, there may be two ways to correct for this. One way whould be to make A1 and A2 larger, and another way would be to make A0 smaller. However, the best way is to both make A1 and A2 larger, and to make A0 smaller, i.e. multiply A1 and A2 by k ($k > 1$) and divide A0 by k. If $0 < k < 1$, the output from the top tank will become smaller. The output from the second tank or the third tank can be adjusted in the same way.

In the case where judgement shows that the base discharge is too small, the method described above cannot work because the fourth tank has no bottom outlet. Therefore, we must increase the water supply to the fourth tank by making C0 larger. However this would decrease the runoff from the third tank and so it is better to supply more water to the third tank from the second tank. In such a case the adjustment to make the base runoff larger is made by increasing C0, B0 and A0 as:

$$C0' = k_1 C0, \quad B0' = k_2 B0, \quad A0' = k_3 A0,$$

where $k_1 > k_2 > k_3 > 1$, for example

$$k_1 = 1 + k, \quad k_2 = 1 + k/2, \quad k_3 = 1 + k/4,$$

$$\text{or} \quad k_1 = k, \quad k_2 = k^{1/2}, \quad k_3 = k^{1/4}.$$

How to Adjust the Shape of Hydrograph.

The adjustment described above will change the volume of each runoff component. But the shape of the hydrograph may also be a problem. For example, the peak of the calculated hydrograph may be too steep or too smooth when compared with the observed one. If the peak is too steep, we must make A0, A1, A2 smaller, by multiplying them some constant k ($0 < k < 1$). In such a way, we can adjust the shape of hydrograph of each runoff component.

How to Determine the Positions of Side Outlets.

In Japan, experience shows that if it rains less than 15mm after about 15 dry days then there will be no change in river discharge. The position of the lower side outlet of the top tank, HA1 = 15, is determined from this experience.

Also we know that when it rains more than 50mm or so the discharge will increase greatly. The position of upper side outlet of the top tank, HA2 = 40, is determined in this way considering also the water loss from the top tank by infiltration and runoff during the rainfall.

In the tank model of Fig. 6.27, the side outlets of the second and third tank are set at HB = HC = 15. These are determined to be similar to HA1 = 15, but without such good reasoning, because HB and HC are not as effective as HA1. The effect of HB or HC appears when the runoff component from the second or the third tank vanishes under dry condition. Examples are shown in Fig. 6.28 of hydrographs from tank models with HC = 15, HC = 30 and HC = 100. These hydrographs were obtained under the following conditions.

a) There is no precipitation but evaporation of 2mm/day.

b) Both the top and the second tanks are empty.

c) The runoff and infiltration coefficient of the third tank are given as C0 = C1 = 0.01. The initial storage of the third tank is 80mm, 100mm and 200mm, respectively, corresponding to the case of HC = 15, HC = 30 and HC = 100.

d) The runoff coefficient of the fourth tank is 0.001 and its initial storage is 1,000mm.

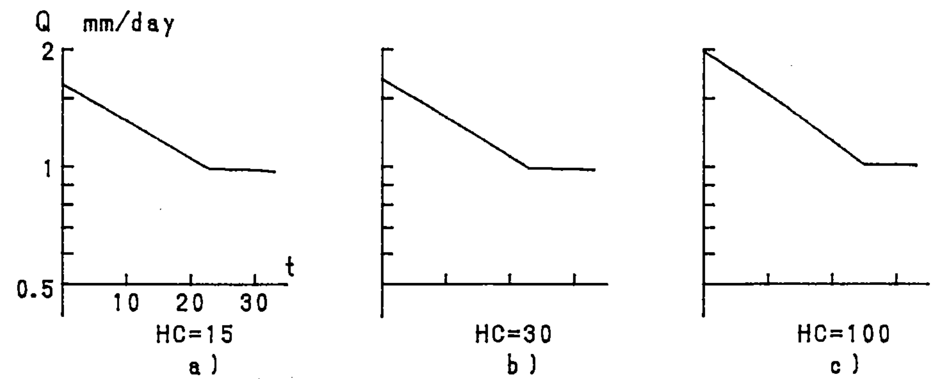


Fig. 6.28.

In some river basins where the observed hydrograph is similar to Fig. 6.28c, the side outlet of the third tank may be set at very high level.

In some cases the fourth tank may be replaced by the type shown in Fig. 6.29. In such a case, there is underground runoff from the basin and the baseflow disappears sometimes.

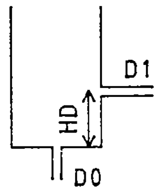


Fig. 6.29.

Volcanic Areas.

In regions where the land surface is covered by thick volcanic deposits the infiltration ability of the land surface is very high, the surface runoff is small and the base discharge is very large. The initial tank model of Fig. 6.27 is not suited for such regions and, instead, the tank models of Fig. 6.30 are recommended as the initial model.

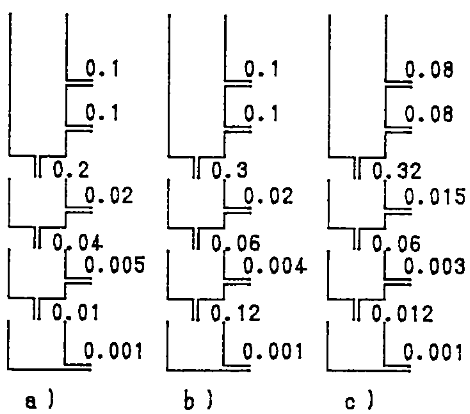


Fig. 6.30.

Some Sort of Balance and Harmony is Important for the Tank Model.

The coefficients of the initial tank models shown in Fig. 6.27 and Fig. 6.30, have the following relationships:

$$A0 : B0 : C0 = 1 : 1/5 : 1/5^2,$$

$$A1/A0 = B1/B0 = C1/C0 = k,$$

where k = 1 for the initial tank model of Fig. 6.27 and k = 1/2, k = 1/3 and k = 1/4 for the initial tank models of Fig. 6.30a, b, and c, respectively. These initial tank models are rather simple and they have some sort of balance and harmony within the parameters. We think that such balance and harmony are important for the initial tank model.

By measuring the slope of the observed hydrograph after a peak flow (plotted in logarithmic scale) we can estimate the value of A0+A1+A2. Then, by assuming a value for k, we can obtain an initial tank model of the type shown in Fig. 6.30.

Even after starting from an initial model that has balance and harmony, however, the balance and harmony may be lost during the course of iterated trials and the results cannot become good, in some cases. When the balance and harmony have been lost in this way it is usually impossible to restore them again. The author considers that this process is somewhat similar resembles to writing poetry or painting a picture painting. In these cases it may be better to start again from a different initial model, which should be made considering the past unsuccessful trials. If the new initial model is appropriate, the result will improve after a few trials.

6.2.2. Automatic Calibration Program (Hydrograph Comparison Method)

Why the author wished to develop the automatic calibration program ?

In 1975, the author retired as director of the institute and had more free time for research. However, the following year, the author was warned by medical authorities that he might soon die from disease. This opinion turned out to be wrong but the scare convinced the author that it was his duty to develop an automatic method of calibrating the tank model so that knowledge of calibration techniques should not be lost with his death.

Once the work had begun it turned out to be rather easy. It was the type of problem which simulated the author's way of thinking and, after a few months, a program had been written which ran successfully and actually gave better results than the manual calibration!

The program is a trial and error method carried out by computer. A rough outline of the program is as follows:

1) Start with an initial model.

2) Divide the whole period into five subperiods, each subperiod to correspond to one of the five runoff components.

3) Compare the calculated hydrograph with the observed hydrograph for each subperiod, and define the criteria RQ(I) and RD(I) (I = 1,2,⋯,5) where RQ(I) is the criterion for volume and RD(I) is the criterion for the hydrograph shape of the I-th subperiod.

4) Adjust the coefficients of the tank model according to the criteria RQ(I) and RD(I).

5) The hydrograph derived by the adjusted tank model is compared with the observed and an evaluation criterion CR is calculated.

6) The next trial is made, using the adjusted model.

7) The automatic calibration procedure will usually finish after several iteration of a few trials and the model of the least criterion CR is considered to be the best model.

Division of the hydrograph into five sub-periods so that in the I-th subperiod the I-th runoff component is the most important.

Evaluation of the effectiveness of the I-th runoff component must be made in the period in which the I-th runoff component is most important. To accomplish this, the entire hydrograph is divided into five subperiod as follows:

Subperiod 1: Days on which Y1, the output from the upper outlet of the top tank, is most important belongs to subperiod 1, i.e. each day that

$$Y1 > CY,$$

where C is a constant (usually set to 0.1) belongs to subperiod 1. Similarly the rules for the remaining four subperiods are:

Subperiod 2:	When Y1 < CY	and Y1+Y2 > CY.
Subperiod 3:	When Y1+Y2 < CY	and Y1+Y2+Y3 > CY.
Subperiod 4:	When Y1+Y2+Y3 < CY	and Y1+Y2+Y3+Y4 > CY.
Subperiod 5:	Otherwise.	

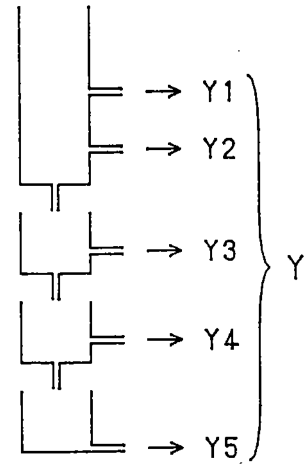


Fig. 6.31.

These rules were determined after some modification and improvements to simplify the programming.

Introduction of criteria RQ(I) and RD(I)

In each subperiod 1, 2, 3, 4 and 5, the discharge volume and the slope of the hydrograph (in logarithmic scale) of the calculated and observed discharges are compared using the following criteria:

$$RQ(I) = \sum_J QE(J) \,/\, \sum_J Q(J) \qquad (I = 1, 2, \ldots, 5)$$

$$RD(I) = \frac{\sum_J' (\log QE(J-1) - \log QE(J))}{\sum_J' (\log Q(J-1) - \log Q(J))} \qquad (I = 1, 2, \ldots, 5)$$

where Q is the observed discharge, QE is the calculated discharge, I is the index number of the subperiod, J is the day number, $\sum_J$ means the sum over the J days belonging to the subperiod I and $\sum_J'$ means the sum over the J days belonging to the subperiod I for which QE(J-1)-QE(J) is positive.

It should be noted that the conditions $\sum_J$ and $\sum_J'$ are determined only from the calculated discharge and not from the observed discharge. For example, in the definition of the criterion RD(I), the summation is carried out over the decreasing part of the calculated hydrograph. The observed hydrograph is not considered so that we avoid the random noise that will appear in it (Fig. 6.32).

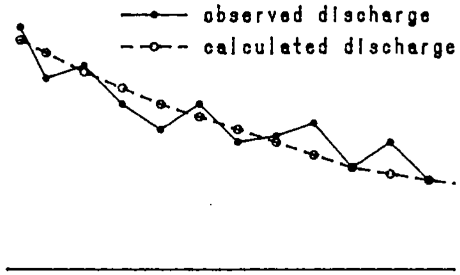


Fig. 6.32.

Later, an improvement to RD(1) and RD(2) was made based on the following reasoning: The top tank has two side outlets and one bottom outlet, and consequently, it has three coefficients A0, A1, and A2. However, there are four criteria RQ(1), RQ(2), RD(1) and RD(2) for the top tank, and so there is an 'excess' of criteria. To resolve this problem RD(1) is neglected (RD(1) = 1) and RD(2) is defined so that the summation Σ is made over the J days belonging to both subperiods 1 and 2 for which QE(J-1)-QE(J) is positive.

Feedback formulae

For the top tank, adjustments to the runoff and infiltration coefficients are made as follows:

If RD(2) > 1 (< 1) then this means that the slope of the runoff component from the top tank is too large (small) and so the sum of the coefficients should be adjusted to

$$(AM0+AM1+AM2)=(A0+A1+A2)/RD(2),$$

where AM0, AM1 and AM2 are adjusted coefficients.

If RQ(1) > 1 (< 1) then this means that the runoff component from the upper outlet is too large (small) and so the ratio of the coefficients A2/A0 should be adjusted to

$$(AM2/AM0)=(A2/A0)/RQ(1).$$

Similarly, A1/A0 is adjusted to

$$(AM1/AM0)=(A1/A0)/RQ(2).$$

By solving these equations we can derive the following feedback formulae (written in FORTRAN form):

$$A=(A2/A0)/RQ(1), \qquad B=(A1/A0)/RQ(2),$$

$$A0=(A0+A1+A2)/(RD(2)\bullet(1+A+B)$$

$$A1=B\bullet A0, \qquad A2=A\bullet A0.$$

Similarly, for the second tank, it is better to adjust the coefficients B0 and B1 by

$$(BM0+BM1)=(B0+B1)/RD(3),$$

$$(BM1/BM0)=(B1/B0)/RQ(3),$$

so that we can derive the following feedback formulae:

$$B=(B1/B0)/RQ(3),$$

$$B0=(B0+B1)/(RD(3)\bullet(1+B)$$

$$B1=B\bullet B0.$$

In the same way, we can derive feedback formulae for the third tank.

$$C=(C1/C0)/RQ(4),$$

$$C0=(C0+C1)/(RD(4)\bullet(1+C)$$

$$C1=C\bullet C0.$$

For the fourth tank, we cannot adjust the discharge amount by modifying the ratio of runoff and infiltration coefficients, because there is no bottom outlet. Therefore, we have to adjust it by modifying the water supply from the third tank, and, in consequence, from the second tank and the top tank, too. If RQ(5) > 1, the supply from the third tank should be made smaller by dividing C0 by RQ(5). If we decrease the value of C0, the runoff from the third tank will increase and we will have to reduce the supply from the second tank, i.e. we have to decrease the value of B0. In general this adjustment need not to be so large as the adjustment for the third tank. From these considerations, we can obtain the following adjustment formulae:

$$D1=D1/RD(5),$$

$$C0=C0/RQ(5),$$

$$B0=B0/(RQ(5))^{1/2},$$

$$A0=A0/(RQ(5))^{1/4}.$$

We would normally expect that, starting from an initial model and using the iterative feedback described above, the tank model would converge to a good fit. However, in order to make the feedback system converge, it may be necessary to include some additional modifications.

Some necessary modifications

RQ(I) and RD(I) should be limitted to the range (1/2, 2)

In some cases, some values of RQ(I) and RD(I) (especially RD(I)), may show values very different from 1. To avoid such extreme values we defined a maximum allowable range of 1/2 to 2, i.e. values of RQ(I) and RD(I) larger than 2 are replaced by 2, and values smaller than 1/2 are replaced by 1/2.

Effect of RD(I) must be halved

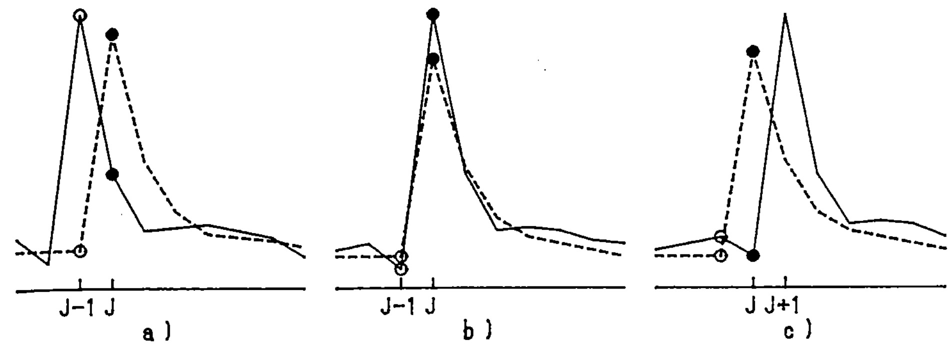
One of the reasons why the feedback system may not work well is the unreliability of RD(I) which is affected by large noise. We suppose that the subtraction included in the definition of RD(I) must be the main reason of its unreliability. To avoid this effect of noise, we halve the effect of RD(I) by replacing it by $(RD(I))^{1/2}$. At first, RD(I) was replaced by $(RD(I))^{1/4}$, but this was found to be too much.

RD(5) should be neglected

As the time constant of the fourth tank is very long, i.e. the slope of the calculated discharge in period 5 is very small, RD(5) is mainly governed by noise, and, accordingly, it is very unreliable. Therefore, the feedback of RD(5) is neglected by putting RD(5)=1.

The effect of random deviation of the time lag of peak discharge

RD(3) and RD(4) often show very unreliable values (i.e. negative values with large absolute values). This may occur when the peak calculated discharge has a one day lag when compared with the observed peak, as shown in Fig. 6.33a. One day difference between the peaks of observed and calculated hydrographs can easily occur because of the different measuring times of precipitation and discharge. For example, rain may occur just before or just after the normal measuring time of daily precipitation (usually 9 o'clock), and would then be recorded as either the precipitation of the foregoing day or of that specific day, only by chance. In the case of Fig. 6.33a, RQ(I) and RD(I) (especially RD(I)), are greatly affected when I = 3,4 or 5; and in the case of Fig. 6.33c, RD(I) for I = 1 or 2 is largely affected, too. To avoid cases in which day J belongs either to subperiod 1 or 2 and the previous day J-1 belongs to any of subperiods 3,4 or 5, the day J-1 is deleted from the definition of RQ(I) and RD(I) and the day J is deleted from the definition of RD(I).



○ : deleted in definition of RQ(I) and RD(I) where I=3 or 4 or 5

● : deleted in definition of RD(I) where I=1 or 2

Fig. 6.33.

Values of RQ(I) and RD(I) which are close to 1 should not be used for feedback

In the early stages of developing the automatic calibration program, feedback procedures using RQ(I) and RD(I) did not give good results. This was probably caused by feedback of RQ(I) and RD(I) for values near to 1. Such criteria which are close to 1 include little information but are mostly composed of noise, and although the feedback effects of each of these criteria is not large individually, their total effect is greater and may lead the calibration astray.

In calibrating the tank model by the trial and error method there is one important principle; we must always concentrate our attention on only one specific point. We have to improve the weak points of the model one by one and not try to adjust many bad points at the same time. This principle is probably true for the automatic calibration method, too. If we apply many feedback procedures at the same time, they may interfere with each other and cause bad result.

After some trials we finally developed a method by which we select two criteria within RQ(I) and RD(I) which are most distant from 1 and use only these two criteria for feedback; i.e. the other criteria are eliminated by putting them equal to 1.

Evaluation criterion

When the author was calibrating the tank model by his trial and error method of subjective judgement, the best model was also determined subjectively. However, for the trial and error procedure by computer, an evaluation criterion is necessary to decide on the best tank model. While the most usual criterion is the mean square error, it is probably better to divide this by the mean value to make the criterion dimensionless.

The criterion obtained is

$$MSEQ = \frac{\left(\sum_J (QE(J) - Q(J))^2 / \sum_J 1\right)^{1/2}}{\sum_J Q(J) / \sum_J 1}.$$

The weak of this criterion, however, is that it is weighted towards high discharges, i.e. it can be thought of as a criterion for floods.

For the overall evaluation of the model it may be better to use the natural logarithm of discharge, as follows:

$$MSELQ = \left(\sum_J \left(\ln QE(J) - \ln Q\left(J'\right)\right)^2 / \sum_J 1\right)^{1/2}.$$

If d is small compared to x, the following relationship holds:

$$\ln(x+d) - \ln x = \ln(1 + d/x) = d/x.$$

Therefore, in the case of the error criterion:

$$\ln QE(J) - \ln Q(J) \quad (QE(J) - Q(J))/Q(J).$$

This means that MSELQ is the mean square of the relative error and is probably better than MSEQ.

However, both MSEQ and MSELQ have a common weak point. In the case shown in Fig. 6.33a, QE shows a large error in days J-1 and J ; and in the case of Fig. 6.33c, QE shows large errors in days J and J+1. Since such cases occur frequently this means that both MSEQ and MSELQ are too high. Even if the difference between QE(J) and Q(J) is large in such cases, we can say that both hydrographs are similar except for the one day shift.

Compare QE(J) with Q(J-1), Q(J) and Q(J+1) and call the nearest one among these three Q(J'). Then a modified MSEQ may be defined as follows:

$$MSEQ = \frac{\left(\sum_J \left(QE(J) - Q\left(J'\right)\right)^2 / \sum_J 1\right)^{1/2}}{\sum_J Q(J) / \sum_J 1}.$$

In the same way MSELQ may be modified by searching ln Q(J-1), ln Q(J) and ln Q(J+1) for the nearest to ln QE(J), calling this ln Q(J'), and rewriting MSELQ as follows:

$$MSELQ = \left(\sum_J \left(\ln QE(J) - \ln Q\left(J'\right)\right)^2 / \sum_J 1\right)^{1/2}.$$

Finally, the evaluation criterion CR is defined by

$$CR = MSEQ + MSELQ.$$

Number of iteration of trials

In calibrating the tank model using the trial and error method of subjective judgement, if we make too many trials then the balance and harmony may be lost and the results may get worse. It is very interesting to note that the automatic calibration by computer also shows the same tendency. Since the computer program is a sort of simulation of the author's way of thinking, then it is also reasonable that the computer program should show this similar tendency.

There must be some reason why the number of iterations should be kept below a certain number. In the computer program, the tank model coefficients are adjusted by RQ(I)'s (I = 1,2,···,5) and RD(I)'s (I = 2,3,4). The two which are most distant from 1 within these eight RQ(I)'s and RD(I)'s, are selected, and these two are used to adjust the tank model coefficients. Therefore, after four iterations it is probable that all the effective RQ(I)'s and RD(I)'s will have been used for adjustment, the adjusted model will have become good, and all the RQ(I)'s and RD(I)'s will have become close to 1.

Generally speaking, each RQ(I) and RD(I) are composed of both signal and noise components. When they become close to 1 the signal component has mostly vanished and the remainder is noise. At this stage the RQ(I) and RD(I) are not usefull and may actually be harmful. Therefore, using such RQ(I)'s and RD(I)'s in the feedback will spoil the balance and harmony of the model. Once the balance is broken, the model can not become good again. It is some sort of divergence.

Uniqueness of the solution is probably nonsense

When the author started to develop the automatic calibration program he expected that there would be some final model, i.e. some definite solution, and that using the automatic calibration program we could approach this final best model. However, this was misplaced optimism. Starting from different initial models we will get different final models and it is very dificult to judge from the hydrographs of these different final models which one is the best.

The tank model is some sort of approximation of the runoff phenomena. There must therefore be many ways of approximating the result, i.e. there cannot be only one unique solution.

6.2.3. Automatic Calibration Program by Duration Curve Comparison Method

Why comparison of duration curve became necessary?

In 1975 and 1976, the author visited the Upper Nile Region around the Lake Victoria, and had the chance to analise the runoff of several basins. This region is located just under the equator and the rainfall is always local showers covering small areas. Therefore, even if there are several rainfall stations in the basin, this will not be enough to catch the rainfall over the basin. In some cases, nearly all stations may record heavy rain but the discharge from the basin is small. On the other hand, the raingauges may record very little but the discharge is large. Therefore, the calculated discharge from rainfall data may not show a good fit to the observed discharge.

Even under such bad conditions we were able to calibrate the tank model by comparing the shapes of the hydrographs, i.e. the slope of the discharge and the volumes of monthly discharge. Also, we were able to adjust the sub-base and the base runoff by comparing discharges in the dry season. However, it seemed to be very difficult to translate these ideas into a computer program.

Fortunately it occured to the author that he could use duration curves instead of hydrographs and it proved rather easy to develop an automatic calibration program using a duration curve comparison method.

The duration curve comparison method is much better than the hydrograph comparison method. It was first developed for basins in Upper Nile Region but it was applied to Japanese basins and gave better results than those obtained by the hydrograph comparison method. Now, the duration curve comparison method is used in every case.

Division of the duration curve into five sections corresponding to five subperiods

On the duration curve of calculated discharge derived from the working tank model, discharges for the days belonging to subperiod 1 are usually large and they gather in the left part of the curve and, conversely, discharges for the days belonging to subperiod 5 are usually small and they gather in the right part of the duration curve.

An example of such a distribution is shown in Fig. 6.34, where daily discharges for one year are ordered from the largest to the smallest, and each daily discharge is described by the number of the subperiod to which it belongs.

1111111111	1112122221	2222222212	2222222222	2222222222
2222222222	2222223223	2323322332	3333332222	2233222223
2333233333	3333333333	2333333333	2332222333	3322332332
3223333333	3333332333	3333333323	3323333333	3333343333
3323334432	3333433423	2433343234	4444343434	4434344333
4433344344	4444444444	4444434443	4344444444	4343443434
4444444444	4444444444	4444444444	4444444444	4444444444
4444444444	55555			

1111111111	2221121222	2222121222	2222222222	2222222222
2222222222	2222222232	2222233322	3323223223	3333333233
3223333333	3222332333	2333333233	3233323333	2333332223
3223333333	2333343333	2433333333	3332432322	2243332443
4433444224	4423344443	3434444344	4344444444	2444424434
4343444443	4343444434	3344444444	4444443443	3443444444
4444434443	4234434434	3434444344	4454554354	4554555555
5555555555	55555			

Fig. 6.34. Computed daily discharges of the River Sarugaishi at Taze (Japan) ordered and described by subperiod number (results for years 1969 and 1970)

After some consideration and hesitation we decided to use the following method: the duration curve is divided in the most simple way into five subsections, successively from the left, NQ(1) days, NQ(2) days,···, NQ(5) days, where NQ(I) is the number of days belonging to subperiod I.

Definition of RQ(I) and RD(I)

After the division of the duration curve into subsections the definition of the criteria RQ(I) and RD(I) is easy. Let Q(N) and QE(N) be the observed and calculated daily discharges in decreasing order, where N is the ordinal number of the day. Then, the five subsections are represented as follows, using the order number N:

subsection 1: $1 \leqq N \leqq NQ(1)$,

subsection 2: $NQ(1) + 1 \leqq N \leqq NQ(1) + NQ(2)$,

subsection 3: $\sum_{I=1}^{2} NQ(I) + 1 \leqq N \leqq \sum_{I=1}^{3} NQ(I)$,

subsection 4: $\sum_{I=1}^{3} NQ(I) + 1 \leqq N \leqq \sum_{I=1}^{4} NQ(I)$, subsection 5: $\sum_{I=1}^{4} NQ(I) + 1 \leqq N \leqq \sum_{I=1}^{5} NQ(I).$

We can define the criterion RQ(I) as the ratio of the sum of the calculated and observed discharges belonging to subsection I:

$$RQ(I) = \sum_{NY}\sum_{I} QE(N) / \sum_{NY}\sum_{I} Q(N) \quad (I = 1, 2, \ldots, 5),$$

where $\sum_{IL}$ is the sum over the subsection I, and $\sum_{NY}$ is the sum over years.

The criterion RD(I) can be defined as the ratio of inclination of calculated and observed duration curves in the subsection I, where the inclination is defined, for simplicity, by the difference of the mean value of duration curve in the left half and the right half of subsection I, respectively:

$$RD(I) = \frac{\sum_{NY}\left(\sum_{IL} QE(N) - \sum_{IR} QE(N)\right)}{\left(\sum_{NY}\left(\sum_{IL} Q(N) - \sum_{IR} Q(N)\right)\right)} \quad (I = 2, 3, 4),$$

where $\sum_{IL}$ and $\sum_{IR}$ mean the sum over the left half and the right half of the subsection I, respectively, and RD(2) is defined on the subsection, $1 \leqq N \leqq NQ(1) + NQ(2)$.

Feedback formulae

The feedback formulae are just the same as those used in the hydrograph comparison method. We knew that in the hydrograph comparison method we had to halve the effect of RD(I), because of its unreliability caused by the large noise effect. In the present case, RD(I) defined by a comparison of duration curves must be much more reliable because the duration curve is smooth and monotonously decreasing and, accordingly, we thought that they could be applied to the feedback formulae without dividing them by 2. Contrary to our expectation, we found that such a feedback formulae did not bring good results and we had to halve the effect of RD(I), just as we did before.

The feedback of RQ(I) probably corresponds to the displacement feedback in an automatic control system and the feedback of RD(I) probably corresponds to a velocity feedback. We also understand that, as a general principle, if we add a slight velocity feedback to a displacement feedback the total feedback system will usually work well. However, we can not understand the exact reason why the effect of RD(I) must be halved.

Evaluation criterion

For the duration curve comparison method we have to introduce new evaluation criteria as well as the ones previously used. Just as before, the mean square error of the duration curve of the calculated discharge or the natural logarithms of the calculated discharge are defined as follows:

$$MSEDC = \frac{\left(\sum_{NY}\sum_{N}(QE(N) - Q(N))^2 / \sum_{NY}\sum_{N} 1\right)^{1/2}}{\left(\sum_{NY}\sum_{N} Q(N) / \sum_{NY}\sum_{N} 1\right)}$$

$$MSEDC = \left(\sum_{NY}\sum_{N}(\ln QE(N) - \ln Q(N))^2 / \sum_{NY}\sum_{N} 1\right)^{1/2}.$$

And we can introduce the evaluation criterion CRDC for the duration curve as

$$CRDC = (MSEDC + MSELDC)/2,$$

in comparison to the earlier criterion, now renamed CRHY, which is given by

$$CRHY = (MSEQ + MSELQ)/2.$$

The final evaluation criterion used in the duration curve comparison method is given by

$$\begin{aligned} CR &= CRHY + CRDC \\ &= (MSEQ + MSELQ)/2 + (MSEDC + MSELDC)/2. \end{aligned}$$

This evaluation criterion seems to be fairly good because, by making the sum of four kinds of mean square error, the random deviation of all the components counteract each other and, consequently, the criterion becomes more reliable than each single criterion individually.

Some characteristics of the duration curve comparison method

The method is resistant to divergence

In this method RD(I)'s were defined under much better conditions than under the hydrograph comparison method and the result obtained was also better than that obtained by the hydrograph comparison method. Moreover, this method has the benefit that divergence hardly ever occurs.

The time lag problem is not important in this method

This method also has the benefit that we do not need to worry about the time lag. In the hydrograph comparison method it is important to determine an appropriate value of time lag and if the assumed time lag is inappropriate the criteria RD(I) become unreliable and, accordingly, the feedback procedures do not converge. Therefore, we have to determine the time lag first from the initial model which usually does not provide a good model for the object basin.

In the duration curve comparison method, such a troublesome problem does not exist at all. The criteria RQ(I) and RD(I) are defined by from the duration curves and so determination of the time lag is not necessary.

Goodness of fit of duration curves is probably more meaningful than goodness of fit of hydrographs

If we apply both methods to some basin and compare the resulting hydrographs, the hydrograph comparison method usually shows a slightly better result than the duration curve comparison method. This should be obvious because, in the hydrograph comparison method, the feedback procedures are made to give a good fit of the observed hydrograph, while, in the duration curve comparison method, the feedback procedures are made to give a good fit of the duration curves and usually, the evaluation criteria CRDC is smaller than CRHY, i.e. the duration curves fit better than the hydrographs.

We suppose that a good fit of duration curves is probably more meaningful than a good fit of hydrographs both from hydrological and practical points of view. It is clear, that for practical uses (say for hydroelectric power generation), a good fit of the duration curves is more meaningful than a good fit of the hydrographs. The hydrological point of view is not so obvious but can be developed as follows: There are many peaks in a hydrograph. In any particular storm rainfall data observed at some stations will show smaller values than the unknown real mean areal rainfall, and in some other storm the contrary case will occur. Therefore, in one case the observed peak discharge will be greater than the calculated one, and, in the other case, the observed peak will be smaller than the calculated one. Accordingly, to obtain a perfect fit between the observed and calculated hydrographs is impossible and to aim for a good fit of the duration curves seems to be more meaningful also from a hydrological point of view.

6.2.4. Calibration of other parameters than runoff-infiltration coefficients

Calibration of positions of side outlets

The automatic calibration program described above can calibrate runoff-infiltration coefficients A0, A1, A2, B0, B1, C0 and C1. However, the tank model has other coefficients than these: i.e. positions of side outlets, HA1, HA2, HB and HC; and parameters of soil moisture structure, S_1, S_2, K_1 and K_2.

The calibration of the positions of the side outlets is not so difficult. Parameters HA1 and HA2 relate to the behavior of the first and second runoff components Y1 and Y2, while HB relates to Y3 and HC relates to Y4. Therefore, within (HA1, HA2), HB and HC, there is not much interaction and we can calibrate them independently. For example, consider the case of HB and make several trials by changing its value. It is important to change the value of HB boldly with some large interval. For example, trials might be made, at first, setting HB as 10, 15 and 20, or 7.5, 15, 22.5. In each case, each trial is made by means of the automatic calibration program and the final model and criterion CR are obtained for each value of HB. In this way, three values of CR will be obtained for HB = h_1, h_2 and h_3 as shown in Fig. 6.35.

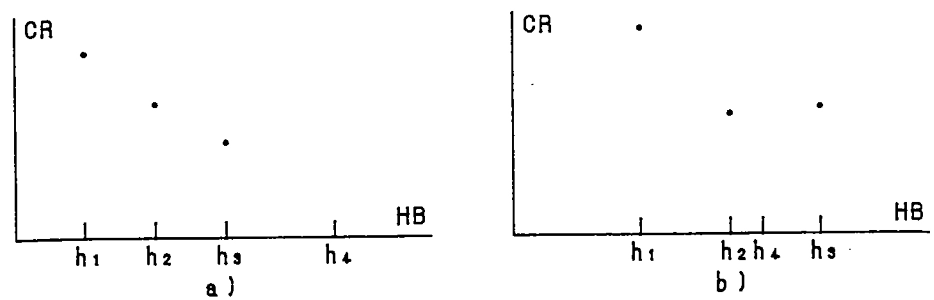


Fig. 6.35.

If the result obtained looks like Fig. 6.35a, we must make a fourth trial by putting HB = h_4, where the value of h_4 must be determined carefully and boldly. If the result obtained looks like Fig. 6.35b, we can determine h_4 without much difficulty, so that CR will attain a minimum at HB = h_4. In this way we can determine the value of HB after several trials. HC can be calibrated in the same way.

The top tank has two side outlets and so there are two parameters HA1 and HA2. We can consider that HA1 and HA2-HA1 will be nearly independent. In the initial model, HA1 = 15 and HA2 = 40 = HA1+25 and therefore we might make trials setting HA1 = 10, 15 and 20; HA2 = 10+25, 15+25 and 20+25. Next, the value of HA1 can be determined in the same way as the case of HB. After HA1 is determined, HA2-HA1 is changed to determine the optimum value and so HA1 and HA2 can be determined after ten trials or so.

Calibration of soil moisture structure

The soil moisture structure has four coefficients: S_1, S_2, K_1 and K_2 and therefore, if we wished to find the optimum values on a four dimensional mesh, the number of trials would become enormous. Therefore, we must find orthogonal factors and determine the optimum values one by one.

The soil moisture structure has two time constant T_1 and T_2. The short time constant T_1 shows how the difference between the relative wetness of primary and secondary soil moisture decreases with time as they approach the equibrium state. The long time constant T_2 shows how the total soil moisture decreases with time under drying condition. Therefore, T_1 and T_2 have entirely different hydrological effects, can be considered as independent factors and can be calibrated one by one.

Usually, S_2 is much greater than S_1, and K_2 is much greater than K_1 and so we can derive rough but simple approximate relations as follows:

$$T_1 = 1/k_1 \doteq S_1/K_2, \qquad T_2 = 1/k_2 \doteq S_2/K_1.$$

For simplicity, we put $T_1' = S_1/K_2$ and $T_2' = S_2/K_1$ so that T_1' and T_2' are values which nearly equal the short time constant T_1 and the long time constant T_2, respectively.

We can then calibrate S_1, S_2, K_1 and K_2 in the following four steps, starting from some initially assumed values of S_1, S_2, K_1 and K_2.

First step

To determine the value of $T_1' = S_1/K_2$, its value is changed, by changing the values of S_1 and K_2. For example, S_1 is changed to 0.95 S_1, S_1 and 1.05 S_1; and K_2 is changed to 1.05 K_2, K_2 and 0.95 K_2. This will change the value of T_1' about 10%. Trials are made, similar to the case of HB, and the value of T_1' is determined so as to give the minimum value of CR. As T_1' approximately equal to the short time constant T_1, which has important a hydrological meaning, it is important to determine the value of T_1' accurately.

Second step

Keeping the value of T_1' constant, the values of S_1 and K_2 are changed. For example, both are changed simultaneously ±10%. In this way, S_1 and K_2 are calibrated.

Third step

The value of T_2' is calibrated in the same way as in the first step.

Fourth step

The values of S_2 and K_1 are calibrated in the same way as in the second step.

6.3. SNOW MODEL AND SOME OTHER IMPORTANT FACTORS

6.3.1. Snow model

Division into zones

When winter comes, snow begins to deposit on high elevation area and then spreads to lower areas. When spring comes the snow deposit begins to melt first in low elevation areas and then moves up the elevation range of the basin. Therefore, it is necessary to divide the basin into elevation zones in order to calculate snow deposit and melt. The number of zones need not be too large; usually, it is sufficient to divide the basin into a few zones with equal elevation interval.

Temperature decrease with elevation

Snow deposit and melt are governed by air temperature, and so the rate of temperature decrease with elevation is one of the most important factors in the snow model. Usually, the temperature decrease per 1,000m in elevation is about 5°C~6°C, or so. In most cases, the precipitation and temperature station is located in the lowest zone and there are no stations in the higher zones. Therefore, we have to estimate the mean temperature of each zone from the value observed in the lowest zone. Under the assumption that the temperature decrease with elevation per 1,000m is 5.5°C, we can derive the mean temperature of the I-th zone T(I), from the observed temperature T at the station as,

$$T(I) = T + T0 - (I-1) \bullet TD,$$

where T0 is a correction term and TD is the temperature decrease per zone. These T0 and TD are initially given as,

$$T0 = (H_0 - H_1) \bullet 5.5 \bullet 10^{-3},$$

$$TD = H_d \bullet 5.5 \bullet 10^{-3},$$

where H_0 is the elevation of the station, H_1 is the middle elevation of the first zone and H_d is the elevation interval of zone.

Starting with initial values of T0 and TD we can calibrate them without much difficulty. In some basins, TD will show a seasonal change.

It is interesting to note that the calibrated value of T0 is usually much smaller than the initial value of T0 and is often negative. This means that the temperature at the recording station is usually higher than the temperature over most of the basin, after neglecting the effects of elevation. This must be due to the following two reasons: One, the station is, usually, set at a convenient place for observation, i.e. in or near a village or some place which is convenient to stay or visit. Such a place is usually warmer than the rest of the basin. The other reason might be that places where people live become warmer because of the effect of human activities.

Precipitation increase with elevation

Generally speaking, it rains more heavier in mountainous areas than on the plain. From a water balance based on a reasonable estimate of annual evaporation over the basin and the recorded annual runoff we can estimate the annual precipitation over the basin. Usually, it is about 120% ~ 130% or so of the annual precipitation observed at some place in the plain.

However, when the author tried to analyze a snowy basin in Japan for the first time, he was astonished by the fact that the discharge in the snowmelt season was very large compared with the total precipitation in winter as observed at several stations in the plain. This means that the precipitation increase with elevation must be very large.

Moreover, there was another curious fact, that the discharge in summer was not so large when compared with the rainfall observed in the plain. This means that the precipitation increase with elevation is not large in summer. So, there must be large seasonal change in the precipitation increase with elevation, i.e. it is large in winter and small in summer.

Such a large seasonal change can be considered reasonable under the following meteorological conditions. In Japan, the most snowy basins are located on the Japan Sea side. In winter, the seasonal north-west wind which blows over the warm Japan Sea brings heavy snowfall to the mountains. In such a case, the precipitation increase with elevation must be very large. In summer, the meteorological conditions are entirely different and rainfall occurs by typhoon, warm front, cold front, local shower, etc. In such cases, rainfall increase with elevation is small, because the wind is nearly independent of the mountain slope.

Assuming that the precipitation increase with elevation is linear and considering the seasonal change, the mean precipitation in the I-th zone is given by,

$$P(I) = (1 + P0 + (I-1) \bullet CP(M) \bullet PD) \bullet P,$$

where P is the observed precipitation, P0 is a correction term, PD is the precipitation increase per zone, CP(M) is the coefficient for the seasonal change and M is the month index.

In some case, it seems to be better to assume that the precipitation in the highest zone is the same as that in the next highest zone. For example, in the case of four zones the precipitation's are given as follows:

$$(1+P0) \bullet P, \qquad (1+P0+CP(M) \bullet PD) \bullet P,$$

$$(1+P0+2CP(M) \bullet PD) \bullet P, \qquad (1+P0+2CP(M) \bullet PD) \bullet P.$$

However, as the area of the fourth zone is not large, the effect of this assumption is not so large.

Snowmelt constant

Snowmelt is mostly governed by air temperature and we may assume that the amount of snowmelt in a day is proportional to the mean air temperature T of the day; the coefficient may be called the snowmelt constant SM. Moreover, if it rains, we can assume that the temperature of the rain water is equal to the air temperature. Then, the snowmelt by rainwater is PT/80, where P is the daily precipitation.

Therefore, the maximum amount of daily snowmelt is given by

$$SM \bullet T + PT/80.$$

When the author analyzed Japanese snow basins he obtained good results by putting SM = 6. However, when he analyzed the snowy foreign basins it was found that SM = 4 gave better results. Japan is very humid region and humid air has much more energy than dry air. This is probably the reason that SM = 6 gave good results for Japanese snow basins.

Avalanche effect

In some cases it is very cold in high zones and the temperature remains under 0 C most of the year. Then, the snow deposit in high zones grow larger and larger and, in the model, might tend to infinity. To avoid this tendency, we have to include the movement of snow by gravity. There are many kinds of a movement, e.g. avalanche, glacier, etc., but we will call such a movement the avalanche effect, for simplicity.

A simple way to consider the avalanche effect is to assume that some part of the snow deposit in the I-th zone moves down to the (I-1)-th zone everyday and that this volume is proportional to the total volume of snow deposit in the I-th zone. We will name this coefficient the avalanche constant AV. Let P be the annual precipitation in the I-th zone and we assume that the daily precipitation is constant, P/365. Then, the snow deposit in the I-th zone will be in an equibrium state when the daily avalanche volume is equal to the daily precipitation, i.e.,

$$AV \bullet SD = P / 365,$$

where SD is the equibrium snow deposit volume in the I-th zone. If AV is 1%, then SD = P/3.65 = 0.274 P; if AV = 0.5%, then SD = P/1.825 = 0.548 P and if AV = 0.1%, then SD = P/0.365 = 2.74 P.

Though the real value of snow deposit in the I-th zone is unknown we can estimate its approximate value, from which we can estimate the approximate value of AV as follows:

$$AV = (P / 365) / SD.$$

This equation is correct for the highest zone. In the next zone, the input is precipitation and the avalanche from the highest zone, which is equal to the precipitation on the highest zone in the equibrium state. Therefore, the input of the next zone per unit area is

$$(S(I) \bullet P(I) + S(I-1) \bullet P(I-1)) / S(I-1),$$

where P(I) and P(I-1) are the annual precipitation in the I-th and the (I-1)-th zone, respectively, and S(I) and S(I-1) are the area of the I-th and the (I-1)-th zone, respectively. Therefore, the avalanche constant in the (I-1)-th zone is given as follows:

$$AV(I-1) = (((P(I) \bullet S(I) + P(I-1) \bullet S(I-1)) / S(I-1) \bullet 365)) / SD(I-1).$$

In the same way, we can derive the formulae for AV(I-2).

The value of AV is not so important for the calculation of discharge; its importance is to modify the accumulation of snow in high zones.

Runoff calculation by the tank model

To calculate runoff from the tank model is rather easy. Though the basin is divided into zones, only one tank model is applied to the whole basin. The precipitation as rain and the snowmelt water are summed up in each zone, the sum is multiplied by the area of each zone, they are summed up and the total is divided by the area of the basin. This is then put into the tank model to be turned into runoff.

6.3.2. Precipitation Factor CP(M)

In snowy basins, precipitation increase with elevation often shows a large seasonal change. Correspondingly, in most non-snowy basins, the precipitation factor CP(M) also shows a seasonal change, where M is the month-index, and sometimes this may be very large. As the effect of CP(M) on the calculated discharge is direct and large, it is very important to determine the appropriate values of CP(M) in runoff analysis.

Fortunately, calibration of CP(M) is not so difficult. We start from some initial tank model and CP(M) = 1 (M = 1,2,···,12). Then, we calibrate CP(M) by comparing the calculated and observed monthly discharges in parallel with the calibration of the tank model. It is better to adjust or change the initial tank model, if necessary. We can also make an automatic calibration program to determine CP(M).

In some special case, CP(M) shows very large value in some specific month. It seems to be curious, but probably from some specific meteorological condition.

6.3.3. Weights of precipitation stations

The Thiessen polygon method is illogical

Even though the Thiessen polygon method has been widely used for a long time, it is illogical and, in reality, cannot give results.

We can clearly see the defect of the method by considering two stations close to the boundary of a basin, as in Fig. 6.36. Using the Thiessen polygon method, the weight of A is very small and that of B very large. However, if the two stations are located close to each other, their weights should be almost equal. If the station A is moved to a A', the weight of A' becomes very large and that of B becomes very small. If the Thiessen polygon method were logical, such an absurdity would not occur.

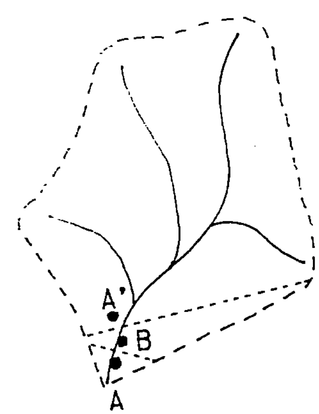


Fig. 6.36.

In principle, the weights of precipitation stations should be determined, not by their geometrical positions, but by the meteorological conditions. For example, in some cases, a station situated near the center of a basin may be almost useless for discharge forecasting, in spite of its large Thiessen weight. In such cases, due to micro-meteorological conditions, the rainfall in a small area around the station, may have a weak relationship with the areal precipitation over the whole basin.

The mean areal precipitation over the basin is unknown

When the basin is very small and there are no woods, no steep mountain, no lake, etc., that prevent the establishment of precipitation stations, then we can measure the approximate value of mean areal precipitation over the basin by setting many precipitation stations uniformly over the basin. It is some sort of sampling survey. However, in almost all cases, such a precipitation measurement is impossible, and so, to measure the areal precipitation over the basin is practically impossible.

As the mean areal precipitation over the basin is unknown, the problem of determining the weights of precipitation stations in order to get the best mean areal precipitation is meaningless.

When we have observed discharge

Even though the mean areal precipitation over the basin is unknown there is often observed discharge, which is some sort of areal integral of the precipitation over the basin. However, there is a large transformation between precipitation and discharge.

The problem then is to derive the mean areal precipitation over the basin from the observed discharge. To our regret, this problem seems to be impossible. The transformation from precipitation to discharge is an itegral-like operator. Most of the rain is stored in the ground before it is turned into discharge. For example, the tank model is some sort of a non-linear incomplete integral. Therefore, the inverse operator which will transform discharge into precipitation must be a differential-like operator, which means that it will amplify high frequency components and be practically useless. Even if we could get such an inverse operator, the areal precipitation over the basin derived from the observed discharge by this inverse operator would be full of noise.

The only remaining way is to transform the precipitation at each station into discharge time series. If these calculated discharge series corresponding to each precipitation station are compared with the observed discharge series, then weights may be determined such that the weighted sum of the calculated discharge series becomes the best approximation of the observed discharge series.

To derive the calculated discharge from precipitation some form of runoff model is necessary. So, we have to start the preliminary runoff analysis using some weights, usually, equal. If some precipitation station is found to be not representative in the course of runoff analysis, it is better to make the weight of this station small or neglect it. After the model has become good enough, the calibration of the weights using the derived runoff model can begin.

Determining the weights of precipitation stations by factor analysis

First, we make monthly or half-monthly sums of the observed discharge and the calculated discharge derived from the precipitation at the I-th station, where I = 1,2,...,N. Let the derived time series of monthly or half-monthly observed and calculated discharge be y(n), and $x_I(n)$ (I = 1,2,...,N), respectively.

Next, these time series are normalized by dividing them by the square root of the mean of square of each component: i.e. $(\sum_n y(n)^2/\sum_n 1)^{1/2}$ and $(\sum_n x_I(n)^2/\sum_n 1)^{1/2}$. Let the normalized time series be Y(n) and $X_I(n)$, respectively.

Then, the problem is to find the best approximation of Y(n) by the linear form $\{\sum_I w(I)X_I(n)\}$, where the coefficients w(I) is the weight of the precipitation station I. This problem can be solved easily by the method of least squares.

We put

$$D(n) = Y(n) - \sum_J^{NP} w(J)X_J(n) \quad (n = 1.2, ..., N).$$

where NP is the number of precipitation stations. Then, the problem is to determine w(I)'s which make $\sum_n D(n)^2$ minimum. To determine w(I), $\sum_n D(n)^2$ is differentiated partially with respect to w(I).

Putting $\frac{\partial}{\partial w(I)}\left(\sum_n D(n)^2\right) = 0$, we get the following equation:

$$\sum_J^{NP} A_{IJ} w(J) = B_I (I = 1, 2, ..., NP),$$

where

$$A_{IJ} = \sum_n X_I(n) X_J(n)/n, \quad B_I = \sum_n X_I(n) Y(n).$$

We had hoped that the solution of this equation would give good weights, but in most cases the solution was meaningless, showing positive and negative numbers with large absolute values. In most cases, the determinant of the matrix (A_{IJ}) is very close to zero, and the roots of the equation are given by the ratio of very small numbers, roughly speaking 0/0. Consequently, they are unreliable. The reason that the determinant $|A_{ij}|$ is very close to zero is that daily precipitation's at stations in or near the basin are usually similar to each other. In other words, they are highly correlated. Therefore, the daily runoffs derived from the precipitation at each station are also similar to each other, consequently, the time series of calculated monthly or half-monthly discharges. $x_I(n)$ (I = 1,2, ,NP) are similar to each other, as are the normalized time series $X_I(n)$. Hence, when we consider the time series $\{X_I(n)\}$ as a vector $\vec{X}_I$, the angle of intersection between any two of these vectors is small. Therefore, all A_{IJ} are close to 1, and, consequently, $\det(A_{IJ})$ shows very small value.

Moreover, if the working model has been calibrated well enough, the daily runoff derived from the precipitation at each station will be similar to the observed runoff. Accordingly, every vector $\vec{X}_I$, (I = 1,2, ,NP) is similar to $\vec{Y}$, i.e. the angle of intersection between $\vec{X}_I$, and $\vec{Y}$ is also small, and, consequently, B_I is close to 1.

To get a meaningful solution to such an equation the method of factor analysis was applied. The matrix (A_{IJ}) was transformed into diagonal form by orthogonal transformation. The derived diagonal elements are called the characteristic values of (A_{IJ}). By neglecting those characteristic values which are close to zero we can get a meaningful solution. Such a treatment is mathematics, not hydrology and so the author will stop the description here. It is not difficult to learn the method of factor analysis from some text book of statistical mathematics.

REFERENCES

Sugawara, M., Maruyama, F., 1952. Statistical method of predicting the runoff from rainfall. Proc. of the 2nd Jap. National Congress for App. Mech: 213-216.

Sugawara, M., 1961. On the analysis of runoff structure about several Japanese rivers. Jap. Jour. Geophys. 2(4):1-76.

Sugawara, M., 1967. The flood forecasting by a series storage type model. International Symp. on floods and their computation. Leningrad, USSR: 1-6.

Sugawara, M., et al., 1974. Tank model and its application to Bird Creek, Wollombi Brook, Bikin River, Kitsu River, Sanaga River and Nam Mune. Research note of the National Research Center for Disaster Prevention, No.11: 1-64.

Sugawara, M., 1979. Automatic calibration of the tank model. Hydrol. Sci. Bull. 24(3):375-388

Sugawara, M., et al., 1984. Tank model with snow component. Research note of the National Research Center for Disaster Prevention, No.65: 1-293.

Sugawara, M., 1993. On the weights of precipitation stations. Advances in theoretical hydrology, European Geophys. Soc. series on hydrological sciences, 1, Elsevier.

Chapter 7

THE XINANJIANG MODEL

R. J. Zhao and X. R. Liu

7.1. INTRODUCTION

The Xinanjiang model was developed in 1973 and published in 1980 (Zhao et al., 1980). Its main feature is the concept of runoff formation on repletion of storage, which means that runoff is not produced until the soil moisture content of the aeration zone reaches field capacity, and thereafter runoff equals the rainfall excess without further loss. This hypothesis was first proposed in China in the 1960s, and much subsequent experience supports its validity for humid and semi-humid regions. According to the original formulation, runoff so generated was separated into two components using Horton's concept of a final, constant, infiltration rate. Infiltrated water was assumed to go to the groundwater storage and the remainder to surface, or storm runoff. However, evidence of variability in the final infiltration rate, and in the unit hydrograph assumed to connect the storm runoff to the discharge from each sub-basin, suggested the necessity of a third component. Guided by the work of Kirkby (1978) an additional component, interflow, was provided in the model in 1980. The modified model is now successfully and widely used in China.

7.2. THE STRUCTURE OF THE XINANJIANG MODEL

7.2.1. The Concept of Runoff Formation on Repletion of Storage

Soil, or indeed any porous medium, possesses the ability of holding indefinitely against gravity a certain amount of water constituting a storage. This is sometimes called "field moisture capacity". By definition, water held in this storage cannot become runoff and the storage can be depleted only by evaporation or the transpiration of the vegetation. Hence evaporation becomes the controlling factor producing soil moisture deficiency.