Table of Contents

Session plan

		Time (min.)
1.	General Overhead - Highlighted text (2)	5
2.	Completion from another record at the same station	5
3.	Interpolating discharge gaps of short duration Overhead - Highlighted text	5
4.	Interpolating gaps during recessions Overhead - Fig. 1 - Semi logarithmic plot of hydrograph	5
5.	Interpolation using regression Overheads - Highlighted text and bullet points (3) Overhead - Fig. 2 - example of fitted relation between observed and estimated flow	5
6.	Water balance and flow routing methods Overhead - Highlighted text 6.1 Muskingum method of flow routing Overhead - highlighted text and equations Overhead - Table 1 - Muskingum routing	10
7.	Using rainfall runoff simulation Overhead - highlighted text	5
	Total time	**40 min.**

1. General

- **Correction and completion will be carried out at the State Data Processing Centre.**

- **After validation a number of values will be flagged as incorrect or doubtful. Some records may be missing due to non-observation or loss on recording or transmission**

- **Incorrect and missing values will be replaced where possible by estimated values based on interpolation or other observations at the same station or neighbouring stations. The process of filling in missing values is generally referred to as 'completion'.**

- **It must be recognised that values estimated from other gauges are inherently less reliable than values properly measured. Doubtful original values will therefore be generally given the benefit of the doubt and will be retained in the record with a flag. Where no suitable neighbouring observations or stations are available, missing values will be left as 'missing' and incorrect values will be set to 'missing'**

- **Procedures for correction and completion depend on the type of error, its duration, and the availability of suitable source records with which to estimate.**

2. Completion from another record at the same station

All streamflow stations equipped with autographic or digital recorders have a back-up of level observations made by the observer. Where there is an equipment failure the observer's record is used to complete the instrument record. However this is normally done as level rather than as discharge.

3. Interpolating discharge gaps of short duration

Unlike rainfall, streamflow shows strong serial correlation; the value on one day is closely related to the value on the previous and following days especially during periods of low flow or recession.

Where gaps in the record are short, during periods of low flow (say, less than 2 days), it may be acceptably accurate to use linear interpolation between the last value before the gap and the first value after it. To confirm that this is acceptable, a graphical display of the hydrograph at the station and one or more neighbouring stations is inspected to ensure that that there are no discontinuities in the flow sequence over the gap.

4. Interpolating gaps during recessions

During periods of recession when the flow is dependent on surface and sub-surface storage rather than rainfall, the flow exhibits a pattern of exponential decay giving a curved trace on a simple plot of Discharge versus Time, but a straight line on a logarithmic plot (Fig.1). **During long recession periods, interpolation between the logarithmically transformed points before and after the gap will result in a more realistic recession than simple linear interpolation.** It is possible to make this interpolation as stage rather than as discharge but, as the principle is based on depletion of a storage volume, it is conceptually simpler to apply the interpolation to discharge rather than to stage.

The slope of the logarithmically transformed flow recession α (also called a reaction factor), from the last value before the gap $Q_{t\,0}$ at time t_0 to the first value after the gap $Q_{t\,1}$ at time t_1 is:

$$a = \frac{\ln Q_{t_0} - \ln Q_{t_1}}{t_1 - t_0} \qquad (1)$$

and k, a reservoir coefficient is:

$$k = \frac{1}{a} \qquad (2)$$

Hence at time t within the gap, Q_t is:

$$Q_t = Q_{t_0} \exp\left(-\frac{t - t_0}{k}\right) \qquad (3)$$

The gap is filled incrementally with no discontinuity at the beginning and end of the gap. In suitable conditions periods of a month or more may be interpolated in this way.

5. Interpolation using regression

During periods of variable flow or in longer gaps, simple interpolation should not be used and relation/regression equations may be applied to fill in missing data, provided there are suitable stations on the same river or on a neighbouring catchment. Regression relations may be obtained for annual, monthly or daily series.

Since relations between stations may change seasonally, HYMOS offers the option of fitting and applying a relation only to a limited period in the year, e.g. daily or monthly values of one to a few months for sequential years. Where two or more such relations are applied to fill a single gap, the resulting interpolated hydrograph is inspected to ensure that there is no serious discontinuity at the junction between the periods of application.

A time shift may be applied to a series to allow for the average lag time between a station represented by the dependent series and the independent station series where they are on the same river - or the difference in rainfall response time if on neighbouring catchments. The time shift may be computed by the user based on the physical properties of the channel or an optimal time shift may be computed using HYMOS.

For a user computed shift (t_1) between an upstream station X and a downstream one Y, spaced at a distance of *s* km, the following may be applied. Given the average flow velocity of *u* m/sec (possibly available from gauging), the propagation velocity or celerity of the flood wave *c* is approximately *1.7u* and the time shift to be applied between X and Y is:

$t_1 = -1000\ s / (1.7\ u)$ (secs), or
$t_1 = -0.00772\ s / u$ (days)

When the river is in flood and especially when the river extends over the floodplain the celerity may differ from 1.7u and a different time shift may be necessary.

HYMOS also provides an analytical procedure to compute time shift on the basis of cross correlation analysis. Correlation coefficients are derived for sequential time shifts. A parabolic relationship is fitted between shift time and correlation coefficient and the maximum of this parabola is displayed as the optimum time shift.

As described in the validation by regression module, HYMOS provides a range of regression relations as follows:

- Single independent variable
  - simple linear
  - polynomial
  - logarithmic
  - power
  - exponential
  - hyperbolic
- Multiple independent variables
  - linear multiple
  - linear stepwise

To establish the spread of the relationship between X and Y and to view its functional form, a scatter graph should first be plotted. In general a simple linear relationship should be tried first, then a polynomial - but rarely exceeding order 3.

After the time period of analysis is specified, the parameters of the relation curve of selected functional form are computed and the graph displayed of the fitted relationship. An example is shown of Fig. 2. Once an acceptable relationship has been established, the parameters may be stored for further use in computing missing values of Y from X

Multiple regression may be applied, for example, between:

- downstream station and two or more upstream tributaries
- downstream station and upstream station and rainfall
- upstream station and downstream station and intervening tributary

Irrespective of the scatter, regression analysis will produce a functional relationship, but if the relationship is poor it should not be used to in-fill missing values. What criteria then, should be used for acceptance? It is suggested that a correlation coefficient r of less than 0.90 be generally considered the lower limit for acceptance but reference should also be made to the standard error of estimate. Where no acceptable relationship is found, the missing values should be left 'missing' or an alternative method of in-filling used.

Application of regression analysis may also produce a discontinuity between observed flows before and after the gap and the in-filled values due to error in the relationship. The in-filled hydrograph plot should be inspected for such discontinuities and suitable adjustment applied.

Moreover, water balance approach for in-filling/correcting, which is explained in the following section, would be preferred as it will have better physical basis.

6. Water balance and flow routing methods

Regression analysis may be used to estimate long periods of missing values or to extend a record. However, in-filling missing values by regression does not ensure water balance between neighbouring stations. Thus application may give significantly less volume at a downstream station than upstream (taking outflows into consideration). To check the water balance, HYMOS provides an option to compare upstream and downstream stations or multiple stations around a junction and to display the balance graphically. **If the balance conflicts with common sense, the functional relationship should be reviewed and if necessary rejected.**

Alternatively, to achieve a satisfactory balance between stations flow routing methods may be applied. The mass-balance equation for a system states that the difference between the input and output is equal to the rate of change in storage.. In flow routing, routing two parameters K and X are determined from measured hydrographs at upstream and downstream stations and applied to route the flow from upstream to a missing downstream station. Inflows and abstractions from the intervening reach can be incorporated to achieve a water balance. Flow routing is usually applied to floods but can be extended for use in low flows. An example of the classical Muskingum method of flow routing follows. More sophisticated alternatives are available using a third parameter.

6.1 Muskingum method of flow routing

The Muskingum method of flow routing is based on the continuity equation:

$$I - Q = \frac{dS}{dt} \tag{4}$$

and

$$S = K[XI + (1 - X)Q] \tag{5}$$

where I = inflow, Q = outflow, S = storage, X = weighting factor, K = storage coefficient, and t = time.

The values of storage at time t and t+1 can be written, respectively, as

$$S_t = K[XI_t + (1 - X)Q_t] \tag{6}$$

and

$$S_{t+1} = K[XI_{t+1} + (1 - X)Q_{t+1}] \tag{7}$$

Using eqs. (6) and (7), the change in the storage over time interval Δt is

$$S_{t+1} - S_t = K[XI_{t+1} + (1 - X)Q_{t+1}] \quad - \quad K[XI_t + (1 - X)Q_t] \tag{8}$$

Considering the variation of inflow and outflow over the interval is approximately linear, the change in storage can also be expressed using eq. (4) as:

$$S_{t+1} - S_t = \frac{(I_t + I_{t+1})}{2} \Delta t \quad - \quad \frac{(Q_t + Q_{t+1})}{2} \Delta t \tag{9}$$

Coupling of Eqs. (8) and (9) in finite difference forms leads to

$$Q_{t+1} = C_0 I_{t+1} + C_1 I_t + C_2 Q_t \tag{10}$$

where C_0, C_1, and C_2 are routing coefficients in terms of Δt, K, and X as follows:

$$C_o = \frac{(\Delta t / K) - 2X}{2(1 - X) + (\Delta t / K)} \tag{11a}$$

$$C_1 = \frac{(\Delta t / K) + 2X}{2(1 - X) + (\Delta t / K)} \tag{11b}$$

$$C_2 = \frac{2(1-X)-(\Delta t/K)}{2(1-X)+(\Delta t/K)} \qquad (11c)$$

Since $(C_0+C_1+C_2)=1$, the routing coefficients can be interpreted as weighting coefficients.

If observed inflow and outflow hydrographs are available for a river reach, the values of K and X can be determined. Assuming various values of X and using known values of the inflow and outflow, successive values of the numerator and denominator of the following expression for K, derived from eq. (8) and (9).

$$K = \frac{0.5\,\Delta t\,[(I_{t+1}+I_t)-(Q_{j+1}+Q_j)]}{X\,(I_{j+1}-I_j)\;+(1-X)\,(Q_{j+1}-Q_j)} \qquad (12)$$

The computed values of the numerator and denominator are plotted for each time interval. With numerator on the vertical axis and the denominator on the horizontal axis. This usually produces a graph in the form of a loop. The value of X that produces a loop closest to a single line is taken to be the correct value for the reach, and K, according to eq. (12), is equal to the slope of the line. Since K is the time required for the incremental flood wave to traverse the reach, its value may also be estimated as the observed time of travel of the peak flow through the reach.

The application of Muskingum method is demonstrated with the help of the following example.

An inflow hydrograph to a channel is shown in Col. 2 of Table 1. Using Muskingum method, route this hydrograph with K=2 days and X=0.1 to calculate an outflow hydrograph. Assume baseflow is 352 m^3/s.

Table 1 shows the computation of outflow using Muskingum method. For given values of K=2 days and X=0.1 (derived from the historical record of inflow and outflow sets), C_0=0.1304 (Eq. 7a), C_1=0.3044 (Eq. 7b), and C_2=0.5652 (Eq. 7c). It is noted that the sum of these routing coefficients is equal to 1.0. Taking Q_1 = 352 m^3/s which is the baseflow, partial flows shown in columns 3 through 5 of Table 1 can be computed and summed up to obtain total outflow given in Col. 6.

The principal difficulty in applying routing methods to infilling missing values is in the assessment of ungauged lateral inflows and outflows, and the method should not be used where these are large and variable.

7. Using rainfall runoff simulation

Regression analysis and flow routing imply that neighbouring flow records are available for infilling. Where no such records are available but the catchment contains sufficient raingauges to calculate areal rainfall, simulation modelling may be used to infill missing records or to extend the record.

The Sacramento model which is incorporated in HYMOS has been outlined in the previous module. With respect to application to infilling missing records, it is assumed that the model has already been calibrated and the parameters selected and optimised and that rainfall data series are available for the period of the missing flow record. The model is again run with the given rainfall (and evapotranspiration) input and parameters to give the outflow record at the specified time step for the missing record period.

Like regression analysis the quality of the generated record cannot be guaranteed. It may be limited by the reliability of the rainfall and flow record during the calibration period. Generally a correlation coefficient of less than 0.70 between observed and simulated flow for the calibration period should not be applied for infilling. Such records should be left missing.

Table 1. Muskingum routing

Time (day)	Inflow (m^3/s)	Partial flows (m^3/s)			Total Outflow (m^3/s)
		C_0I_2	C_1I_1	C_2Q_1	
1	2	3	4	5	6
0	352.0	-	-	-	352
1	587.0	76.5	107.1	199.0	382.6
2	1353.0	176.4	178.7	216.3	571.4
3	2725.0	355.3	411.9	322.9	1090.1
4	4408.5	574.9	829.5	616.1	2020.5
5	5987.0	780.7	1341.9	1142.0	3264.6
6	6704.0	874.2	1822.4	1845.2	4541.8
7	6951.0	906.4	2040.7	2567.0	5514.1
8	6839.0	891.8	2115.9	3116.6	6124.3
9	6207.0	809.4	2081.8	3461.4	6352.6
10	5346.0	697.1	1889.4	3590.5	6177.0
11	4560.0	594.6	1627.3	3491.3	5713.2
12	3861.5	503.5	1388.1	3229.1	5120.7
13	3007.0	392.1	1175.4	2894.2	4461.8
14	2357.5	307.4	915.3	2521.8	3744.5
15	1779.0	232.0	717.6	2116.4	3066.0
16	1405.0	183.2	541.5	1732.9	2457.7
17	1123.0	146.4	427.7	1389.1	1963.2
18	952.5	124.2	341.8	1109.6	1575.6
19	730.0	95.2	289.9	890.6	1275.7
20	605.0	78.9	222.2	721.0	1022.1

21	514.0	67.0	184.2	577.7	828.9
22	422.0	55.0	156.5	468.5	680.0
23	352.0	45.9	128.5	384.3	558.7
24	352.0	45.9	107.1	315.8	468.8

Fig. 1 Semi-logarithmic plot of a hydrograph during recession

Fig. 2 **Example of a fitted relationship between two series**