Technical Assistance
HYDROLOGY PROJECT

World Bank & Government of The Netherlands funded

Training module # 5

Understanding the Concept of Optimal Monitoring frequency of DWLR

New Delhi, March 2000

CSMRS Building, 4th Floor, Olof Palme Marg, Hauz Khas,
New Delhi – 11 00 16 India
Tel: 68 61 681 / 84 Fax: (+ 91 11) 68 61 685
E-Mail: dhvdelft@del2.vsnl.net.in

DHV Consultants BV & DELFT HYDRAULICS

with
HALCROW, TAHAL, CES, ORG & JPS

Table of contents

Page

1. Module context

While designing a training course, the relationship between this module and the others, would be maintained by keeping them close together in the syllabus and place them in a logical sequence. The actual selection of the topics and the depth of training would, of course, depend on the training needs of the participants, i.e. their knowledge level and skills performance upon the start of the course. This module is related to module 2, 3 & 4 and should be referred during the discussions.

2. Module profile

Title	:	Understanding the Concept of Optimal Monitoring frequency of DWLR
Target group	:	Hydrogeologists, Asst-Hydrogeologists, Senior Technical Assistant
Duration	:	One Session of 45 minutes
Objectives	:	After the training the participants will be able to: • Arrive at optimal monitoring frequency for any given intended use of the DWLR data.
Key concepts	:	• Credibility of derived attribute(s) • Preserving the hydrograph shape • Identification of Cycles • Correlation
Training methods	:	Lecture
Training tools required	:	OHS
Handouts	:	As provided in this module
Further reading and references	:	

3. Session plan

No	Activities	Time	Tools
1	• Describe the strategies to be adopted for arriving at optimal monitoring frequency, ensuring the credibility of the attributes, preserving the shape of hydrograph.	30 min	
2	• Illustration	5 min	OHS
3	***Feedback***	10 min	

4. Main text

Contents

Understanding the Concept of Optimal Monitoring frequency of DWLR

1. Objective based Optimal Frequency

The high frequency data emanating from the DWLRs shall assist the hydrogeologists in performing their professional activities more objectively and hence more credibly. The requirement of the frequency can apparently not be uniquely defined and shall depend upon the intended use of the high frequency data as well as upon the local hydrogeological and hydrological characteristics. For example, an aquifer having low specific yield may display faster water level variations and hence may need a higher frequency of monitoring. Similarly, if the objective is to estimate only the peaks and troughs, a higher frequency may be adopted around the beginning and the end of the rainy seasons and a lower frequency may be adopted at other times. On the other hand if the objective is to arrive at a true hydrograph, a uniformly high frequency may have to be adopted. A few strategies for arriving at the optimal monitoring frequency are described in the following paragraphs:

2. Optimizing Criteria

It follows from the preceding paragraph that there does not exist any unique optimal monitoring frequency. The optimal monitoring frequency would depend upon the expectation from (or intended use/uses of) the hydrograph to be monitored. Some criteria could be as follows:

2.1 Credibility of Derived Attribute(s)

There may normally be a few well-defined objectives of monitoring the water table/piezometric head. The objectives may relate to deriving one or more of the following attributes from the observed hydrograph.

- Peak of the hydrograph
- Trough of the hydrograph
- Range of water level fluctuation
- Time of shallow water level, i.e., time during which the water level rises above a stipulated shallow critical level
- Time of deep water level, i.e., time during which the water level falls below a stipulated deep critical level

Thus, the criteria would be to arrive at such monitoring frequency that the desired attribute(s) as derived from the observed hydrograph is (are) close enough to the true values (i.e., the values derived from the true hydrograph).

2.2 Preserving the Hydrograph Shape

This implies that the selected monitoring interval should be such that the monitored hydrograph resembles *closely* with the *true hydrograph.* This is indeed the most stringent and all-encompassing expectation requiring uniformly small monitoring intervals. The intervals would depend upon the degree of the desired resemblance. As discussed subsequently, *correlation* is an index of similarity between the shapes of two time series. It shall be an index of resemblance if the two time series display the variation of the same variable. Thus, the criteria could be to arrive at such an interval which ensures a *high* enough correlation between the true and the monitored hydrographs.

2.3 Identification of Cycles

As already stated, the cycle of smallest time period that can be identified (or separated) by the Spectral analysis is $4\Delta t$, where Δt is the interval between two successive water level data. Thus, the monitoring interval has to be at lone fourth of the time period of the smallest cycle intended to be identified.

3. Optimizing Strategy

It follows from the above discussion that for optimising the monitoring frequency, we shall require the true hydrograph. The hydrograph of smallest *feasible interval (*say hourly) could be *deemed* as the *true* hydrograph. Therefore, it is necessary to first procure a time series of water level with small enough monitoring interval. Subsequently *under-sampled* hydrographs of increasing monitoring intervals, are simulated as follows:

- Knock off the intermittent data from the *true* hydrograph to simulate the under-sampled series of the chosen larger interval.

- Simulate the under-sampled hydrograph by estimating the knocked off data by linear interpolation.

The simulated under-sampled hydrographs may be analysed to arrive at the optimal monitoring frequency as follows:

3.1 Credible Attribute Estimation

Compute the desired attribute from the *true* hydrograph and from each of the simulated under-sampled hydrographs. Terming the attribute value as computed from the *true* hydrograph as X and the values computed from i^{th} simulated under-sampled hydrograph as A_i, compute the *loss array* [$L_i = ABS(X - A_i)/S$]. Here, S is a relevant quantity for normalizing the error. Thus the array represents the *loss of information* (expressed as a fraction of the chosen S) on account of increasing the time interval from the minimum feasible to the one incorporated in i^{th} simulated under-sampled hydrograph. This array could be interpreted for specific attributes as follows:

Peak, Trough and Range: (Over-estimation of depth to peak, under-estimation of depth to trough, and consequent under-estimation of the range) Selecting S as the *true* range (that is, vertical distance from trough to peak in the *true* hydrograph), the array shall represent the loss of information expressed as a fraction of the *true* range.

Time of shallow/deep water level: Selecting S as *true* period, the array shall represent the loss expressed as a fraction of the *true* time.

Plot the computed array [L_i] versus the corresponding intervals of the simulated under-sampled hydrographs. Assigning an *acceptable* level of the loss, pick up the optimal interval of monitoring.

3.2 Preservation of Hydrograph Shape

The simulated under-sampled hydrographs of increasing time interval may be successively compared visually with the *true* hydrograph. This may lead to a threshold interval beyond which the two series may cease to resemble each other. Alternatively, this could be done more objectively in the following steps:

- Compute the *correlation* (described in the following section) between the *true* hydrograph and each of the simulated under-sampled hydrographs.

- Plot the computed correlations against the time interval.
- Pick up the optimal interval corresponding to the minimum desired similarity between the two hydrographs.
-

3.3 Identification of Cycles

The longest permissible interval for identifying a cycle of any time period, by Harmonic analysis is half the time period. For a more reliable identification of the cycle, it may be desirable to further restrict the longest permissible interval say to one fourth of the time period. Thus, for identifying a daily cycle, it may be desirable to have an interval no bigger than 6 hours.

4. Correlation

This statistic determines the degree of linear interrelation (that is, *scaled similarity*) between two time series.

A direct linear relation (that is, as one series rises, the other also rises and vice versa) is termed as *positive correlation.* An inverse linear relation (that is, as one series rises, the other declines and vice versa) is termed as *negative correlation.* If the rise of one series has apparently no effect on the other, the two series are known to be *uncorrelated*.

The two series must have the same data frequency and an adequately long overlap. A correlation between two series falling in different time spans can be computed by analysing the overlapping period only, that is, by curtailing one or both the series. If there is no overlapping period, the correlation can not be estimated. If the two series comprise data at different frequencies, it is necessary to manipulate one of the two series to ensure frequency-compatibility. Thus, either the series of higher frequency (say series of *DWLR* data) may be pruned, or the *missing* data in the series of low frequency (say series of manual data) may be interpolated.

This correlation is essentially a normalised covariance between the pivotal and the derived series. Thus, the correlation ® between two series (x_i, i = 1, 2, ……., n) and (y_i, i = 1, 2, ………, n) shall be given by the following equation:

$$\text{COR} = \frac{\sum_{i=1}^{n} (x_i - \bar{x}).(y_i - \bar{y})}{S_x.S_y}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}; \qquad \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

$$\text{S}_\text{x} = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2}\ ; \qquad \text{S}_\text{x} = \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}\ ;$$

Two positively correlated time series shall have a correlation greater than zero and it may go up to +1. A correlation of +1 indicates a perfect positive correlation, that is, a direct proportionality of the fluctuations in the two series and hence a perfect linearity with positive gradient between the two variables. Similarly, two negatively correlated time series shall have a correlation less than zero and it may go up to –1. A correlation of –1 indicates a perfect negative correlation, that is, an inverse proportionality and hence a perfect linearity with negative slope. Two uncorrelated series shall have a zero correlation.

5 Software

The above stated approach for optimising the monitoring frequency was conceptualised by technical consultants to the *Hydrology Project.* The approach has been assimilated in a software, christened as *DWLR-ANALYST.*

5. Overhead/flipchart master

6. Handout

7. Additional handout

These handouts are distributed during delivery and contain test questions, answers to questions, special worksheets, optional information, and other matters you would not like to be seen in the regular handouts.

It is a good practice to pre-punch these additional handouts, so the participants can easily insert them in the main handout folder.