Transfer Of Information Among Water Quality Monitoring Sites: Assessment By An Optimization Method

Nilgun B. Harmancioglu¹, Cem Polat Cetinkaya¹ and Paul Geerders²

Abstract
Allocation of sampling sites is the initial and the most crucial step of the water quality monitoring network design and redesign process. Several different approaches have been used within the last 20-30 years in the selection of sampling sites. As yet, there are no standard design procedures to accomplish such a network. However, there are some scientific methods which may help to minimize the subjective aspects of design. This paper examines the application of an optimization method that can be used to assess an existing water quality monitoring for information transfer between its sampling sites. The method uses dynamic programming to evaluate the reduction of the number of sampling sites in a basin with respect to different monitoring objectives. The methodology is demonstrated in the case of the Gediz River basin in Turkey.

1. Introduction
1.1 General
Early in the 70’s, water pollution related to anthropogenic factors arose as an important problem in water resources. Water quality monitoring gained importance and water pollution has also been identified to lead to water scarcity. In order to determine the quality of available waters, monitoring networks have been established.

In recent years, problems observed in available water quality data and shortcomings of current monitoring networks have led designers and researchers to focus more critically on the design procedures used. Developed countries have felt the need to assess and redesign their monitoring programs after having run their net-

¹ Dokuz Eylul University, SUMER, Tinaztepe Campus, B게 35160 Izmir, Turkey, e-mail:nilgun.harmancioglu@deu.edu.tr, cem.cetinkaya@deu.edu.tr
² P. Geerders Consultancy, Kobaltpad 16, 3402JL, IJsselstein, the Netherlands, www.pgcons.nl
works for more than 20 years. Developing countries are still in the process of expanding their rather newly initiated networks; yet they also find it necessary to evaluate what they have accomplished so far and how they should proceed from this point on. In both cases of the developed and the developing countries, the major problem is that there are no universally confirmed guidelines to follow in the assessment and design of water quality monitoring networks. Upon this need, significant amount of research has been initiated to evaluate current design procedures and investigate effective means of improving the efficiency of existing networks (Ward et al., 1990; Chapman, 1992; Harmancioglu et al., 1992; Adriaanse et al, 1995; Ward, 1996; Timmerman et al. 1996; Niederlander et al., 1996; Dixon & Chiswell, 1996; Icaga, 1998).

In the view of the above-mentioned problems related to monitoring of water quality on rivers, most countries have started to assess and redesign their existing networks. Turkey, as a typical developing country, has established its water quality monitoring networks since the late 70’s. The government makes the investments for these networks, and the monitoring agencies have taken monitoring activity as one of their official tasks. Recently, these agencies have commenced to question the performance of their networks for their efficiency and cost-effectiveness and to inquire whether available data produce the required information. With respect to cost-effectiveness, there has been no major concern as the government used to pay for every investment; but recently, the government has foreseen a reorganization of all nation-wide activities, including monitoring, as dictated by increasing economic pressures (Harmancioglu et al., 1994; Harmancioglu, 1997).

1.2 Objectives and Scope of the Study

Lettenmaier et al. (1984) proposed a methodology based on dynamic programming as an optimization technique. The method used accomplishes the systematic consolidation of a fixed station water quality monitoring network using dynamic programming. The approach they developed uses a hierarchical structure; that is, monitoring stations are allocated to a weighted attribute score, and specific station locations within each subbasin are determined using a criterion based on stream order numbers. Lettenmaier et al. (1984) applied the method to reduce the number of stations in the fixed trend detection baseline network of the Municipality of Metropolitan Seattle. The results of their study helped to consolidate this network from 81 to 47 stations and led to annual savings of about $33,000.

This optimization method that can be used to assess an existing water quality monitoring for information transfer between its sampling sites has been employed in this paper. The methodology is demonstrated in the case of the Gediz River basin in Turkey with respect to different monitoring objectives.
With respect to the application of the methodology to Gediz River basin, the following general considerations are made: (a) alternative stations are obtained for the case of 14 stations to be retained in the network. Selection of the most appropriate one requires two issues to be accomplished: (1) relevant costs of each alternative solution must be analyzed; (2) the monitoring agency has to delineate its own specific objectives for monitoring and then select the solution that best suits these objectives; (b) if the monitoring agency prefers to expand the existing network, the results of the study can be used to select the locations of new stations. The methodology further shows which station(s) must be selected when one wants to retain different numbers of stations in the network.

2. Methodology

The criteria to be used for ordering stations are essentially based on data and information requirements of water management objectives in a basin. In the case investigated by Lettenmaier et al. (1984), the objectives of the Municipality of Metropolitan Seattle had been to protect and enhance swimmability and fishability of the waters within the selected drainage basins. The station retention algorithm described by Lettenmaier et al. (1984) uses a weighted sum of transformed values of the above criteria. The specific values of the criteria associated with the stations considered for retention are called station attributes and are denoted by “l” as the attribute index.

Icaga (1998) modified the above methodology in the sense of alternative basin water quality management objectives such as the delineation of point and nonpoint pollution discharges. Through the application of the methodology, the optimum combination of stations in a reduced network for each management objective has been analyzed.

The station allocation algorithm is employed in two steps. First, the basin is divided into subbasins or “primary basins”, and for each primary basin, the algorithm determines the preferred sets of station combinations for each possible number of stations ranging from zero to the pre-existing number of stations. There may be very large numbers of station combinations; therefore, a method based on stream order numbers (Sharp, 1970) is used to limit the number of alternative station configurations within each primary basin. Thus, the preferred sets of station combinations is determined by maximizing the sum of the stream order numbers for each station retained and, for a fixed number of stations, by breaking ties through maximizing the score sums for weighted (transformed) attributes. In the second step, a dynamic program, using primary basins as stages, and stations within each primary basin as states, determines the combination of station allocations to the various primary basins, resulting in a total network of given size having maximum total score, where the total score is the sum of station scores within each primary basin for the selected station configurations determined in first step (Figure 1).
2.1 Determination of Subbasins

The network reduction problem is approached first by dividing the river basin into \( N \) subbasins with \( k \) denoting the subbasin index as \( k = 1, \ldots, N \). This division does not have to be hydrologic. Basin properties such as topography, geology, meteorology, land use, industry, population density, junctions of tributaries, etc., may be used as criteria for segregating the basin into subbasins. One criterion that must be satisfied is that each subbasin must have at least one monitoring station.

Regarding each subbasin \( k \), \( P_k \) denotes the pre-existing number of stations in the \( k^{th} \) primary basin and \( R_k \), the number of stations to be retained in that subbasin.

2.2 Alternative Combinations of Stations

Lettenmaier et al. (1984) suggest the use of a method based on stream order numbers (Sharp, 1970 and 1971) to limit the number of alternative station combinations considered within each subbasin. Figure 2 shows the application of stream order numbers in a hypothetical subbasin. Here, each exterior link of a river reach is numbered as “1”. The stream order number of an interior link is found as the sum of the orders of the upstream exterior links. The method reduces the number of alternative number of combinations from 63 to 8 in the hypothetical basin of Figure 2.
2.3 Determination of Attributes and Management Objectives for Monitoring Stations

Each station in a monitoring network must be identified with its attributes \( l \) that comply with management objectives in the basin. In this study, 6 basic attributes are chosen as drainage area, population and irrigation area considering the physical situation effecting monitoring station and number of observations, length of the observation period and observed variables regarding the characteristics of station itself.

The above attributes are defined by their numerical values (or scores). These scores are represented by \( SR_{j(i)k} \) with \( k \) representing the primary basin index; \( l \), the attribute index; \( i \), the station index within primary basin \( k \); and \( j(i) \), index number of the \( i \)th station in primary basin \( k \). The determination of the numerical values for some attributes (i.e., drainage area, population, irrigation area, number of samples, and period of observation) is straightforward. For some attributes like observed variables, appropriate ones must be selected to comply with management objectives. The selection process is also complicated due to the presence of a large number of variables observed at each station. It is desirable to specify the scores \( SR_{j(i)k} \) to have an approximately uniform distribution for each attribute \( l \). If this can be achieved, the dominance of the weighted score by any one attribute on the basis of its relative magnitude alone can be avoided.

---

R\(_i\): Number of stations to be retained in the primary basin

<table>
<thead>
<tr>
<th>Station Name</th>
<th>Stream Order No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(SON):</td>
</tr>
<tr>
<td>a</td>
<td>7</td>
</tr>
<tr>
<td>b</td>
<td>4</td>
</tr>
<tr>
<td>c</td>
<td>3</td>
</tr>
<tr>
<td>d</td>
<td>1</td>
</tr>
<tr>
<td>e</td>
<td>1</td>
</tr>
<tr>
<td>f</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( R_k )</th>
<th>Combination(s)</th>
<th>Sum of SON</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>a</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>ab</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>abc, abf</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>abef</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>abcef, abcdcf</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td>abcddef</td>
<td>19</td>
</tr>
</tbody>
</table>
The system adopted by Lettenmaier et al. (1984) and by Icaga (1998) had the property that the SRj(i)kl were approximately uniformly distributed on the interval (0, 100) and was achieved as follows: for a normally distributed random variable (e.g., attribute), a uniform distribution can easily be achieved by an inverse normal transformation, since the cumulative distribution function of a random variable is by definition uniform in the range (0, 1). Alternately, if a variable can be made approximately normal by transformation (e.g., logarithm or power function), a two-step process can be followed to derive a score, which is approximately uniform. Thus, tests for normality should be carried out and normalization for non-normal scores must be made.

Once variables are selected and all attributes SRj(i)kl are identified and their uniformization (SUj(i)kl) is carried out, the significance of each attribute (or the preference to be given to each) must be specified. This is done by assigning weights w_l to each attribute. The selection of attribute weights is subjective and reflects different management concerns. Essentially, the stations to be allocated within each primary basin are a function of the weights; therefore, the sensitivity of station allocation to different weights may be tested (Lettenmaier, et al., 1984).

2.4 Optimization

For the number of stations to be retained in the entire river basin (TR_N), it is necessary to determine the number of stations to be retained in each primary basin (R_k). Accordingly, the numbers of stations that will be retained in each primary basin should be determined such that the sum of the uniformized attribute data

\[( SU_{j(i)kl} )\] of these stations must be maximum.

For each station combination in primary basin k, the sum of the uniformized attribute data (SU_{j(i)kl}) is shown as TS_{j(i)k}:

\[ TS_{j(i)k} = \sum_{l=1}^{L_k} (w_1 \times SU_{j(i)kl}) \]  \hspace{1cm} (1)

Equation 1 gives the total attribute value in primary basin k and for j(i)th station combination. In each primary basin k, there will be different station combinations depending on R_k so that the TS_{j(i)k} values of these station combinations will be different. It is a general logical rule to select those combinations with the highest TS_{j(i)k} values (MTS_{j(i)k}).

When determining the TR_N stations to be retained in the entire basin, the R_k combinations will be those that give the maximum MTS_{j(i)k} attribute values:
\[ SMTS = \max \sum_{k=1}^{N} \sum_{i=1}^{R_k} MTS_{j(i)k} \]  

where \( SMTS \) is the sum of the \( MTS_{j(i)k} \), which give the maximum value for \( TR_N \); \( N \), the number of primary basins; \( R_k \), the number of stations to be retained in primary basin \( k \); \( MTS_{j(i)k} \), maximum uniformized total attribute value for \( j(i)^{th} \) station combination in primary basin \( k \). When equation 2 is examined, it is obvious that the problem has two dimensions. Hence, it is not possible to determine the station combinations to be retained in the basin by ranking the \( MTS_{j(i)k} \) values from the largest to the smallest and taking the first \( TR_N \) value reached. Dynamic programming method is a tool to be used to solve this problem as there are several alternative combinations of maximum total scores \( MTS_{j(i)k} \).

The objective of the station allocation problem is to find the combination of stations that maximizes the \( MTS_{j(i)k} \) corresponding to the determined \( TR_N \) (total number of stations which will be retained), so that the objective function of the problem can be determined as:

\[ V = \max \sum_{k=1}^{N} \sum_{i=1}^{R_k} MTS_{j(i)k} \]  

The constraints of the problem are:

\[ \sum_{k=1}^{N} R_k = TR_N \quad ; \quad 0 \leq R_k \leq TR_N \]  

\[ 1 \leq j(i) \leq P_k \quad ; \quad j(i) \neq j(h), i \neq h \]  

where \( V \) is the objective function; \( N \), the total number of primary basins; \( R_k \), the number of stations which will be retained in primary basin \( k \); \( k \), primary basin index; \( i \), station index within primary basin \( k \); \( j(i) \), index number of \( i^{th} \) station in primary basin \( k \); \( TR_N \), the total number of stations to be retained in entire basin; and \( P_k \), pre-existing number of stations in primary basin \( k \).

### 3. Application to Gediz Basin

Gediz River basin is located in West Anatolia between the Aegean Sea and Kucuk Menderes and Bakircay river basins. The basin is covered with mountains of 2000 m high from the north, south, and the east, and it reaches the Aegean Sea from the west. The length of the river is 276 km and the total drainage area is 16 775 km². The main tributaries are Deliniş, Selendi, Demirci, Nif, Alaşehir and Kumçay streams. Adala, Ahmetli, Menemen, Akhisar, Selendi, Kapaklı, and Alaşehir plains
make up the lower Gediz plains, which are subject to extensive agricultural practices with large irrigation schemes covering around 110 000 hectares. There are many municipalities and small towns with increasing population in the basin. Around 1.6 million population lives within the borders of basin; furthermore, the basin provides groundwater supply for domestic purposes for the city of Izmir. There are two important industrial areas in the basin. The largest is in the Nif Valley immediately east of Izmir in Kemalpasa municipality. There is also a growing industrial estate in the western edge of Manisa.

Water quality in the Gediz River has been observed by State Hydraulic Works (DSI) at a number of sites since 1980. These observations started at 6 locations and were increased to 33 by 1993. After 1993, DSI reduced the number of sampling sites to 14 (Figure 3). The application presented here covers an assessment of alternative combinations of monitoring sites when a total of 14 stations are to be retained in the network.
a) 33 Water quality monitoring stations in Gediz Basin operated until 1993
b) Existing 14 water quality monitoring stations in Gediz Basin

Regarding the previous network, which contains 33 stations, the above explained methodology is applied to the Gediz Basin. The basin has been divided into 7 subbasins (Figure 4) and according to Sharp’s (1970) procedure, alternative station combinations and their hierarchical levels have been determined. The above mentioned attribute scores such as drainage area, population, irrigated area, number of observations and operation period of the station have been calculated for every station. On the other hand, variables observed at the stations have been taken into account, regarding the monitoring policies of DSI. Although DSI’s monitoring goal along the basin is not clearly identified, three possible goals are considered: a) to detect point pollution (P), b) to detect non-point pollution (NP) and c) to detect both (P,NP). Related variables observed at the stations are divided into three groups and are given in Table 1.
### Table 1

<table>
<thead>
<tr>
<th>Variables considered, regarding pollution sources</th>
</tr>
</thead>
</table>

All attributes including observed variables have been normalized and uniformized and their scores have been calculated. Regarding management goals, different weights are attached to the scores in order to have final scores of attributes. For point pollution, drainage area and population are weighted 2, and the others have been considered as 1. In the case of non-point pollution, drainage area and irrigated area had 2 as the weight, and the other attributes were rated 1. Finally, considering both pollution types, drainage area, population and irrigated area had a weight of 2 and the others, 1.

### 4. Results

The optimization procedure explained in section 2 has been carried out for 7 subbasins and for three different management goals; the network with 33 stations has been reduced to 14 stations and the results are shown in Table 2. In Table 2, the numbers indicate the station identification numbers which are available in Figure 3a.

### Table 2

Results of optimization procedure: 33 stations has been reduced to 14. The numbers are indicating the station identification numbers in Figure 3a.
Only for the case of point pollution, station “47” replaces station “30”. The operational and laboratory analysis costs per measurement are calculated for non-point and point & non-point pollution choices and have been found as 4440 US dollars. For the point pollution choice, the cost rises to 4460 US dollars due to the location of station “47”, which is on the far upstream side of the basin. Costs of alternative station combination choices do not differ from each other; the main reason for this is that the laboratory analysis costs are almost the same for each station and the traveling costs for sampling are small as compared to laboratory analysis costs.

The existing network with 14 stations has been analyzed for point & non-point pollution detection goals. 10 stations (from downstream to upstream; “5”, “6”, “22”, “26”, “30”, “2”, “24”, “41”, “43” and “49”) should be retained in the network, according to the optimization method. Stations “38”, “23”, “44” and “47” should be closed and instead of these stations, the previously operated four other stations, “25”, “3”, “28” and “39” should be put into operation.

Bibliography


versity, Graduate School of Natural and Applied Sciences, Izmir, 102 p.,
(Advisor: Prof. Dr. N. B. Harmancioğlu).

Stream Quality Monitoring Network, Water Resources Bulletin, AWRA, 20,
p. 473-481.

Niederlander, H.A.G., Dogterom, J., Buijs, P.H.L., Hupkes, R., & Adriaanse, M.
(1996) UN/ECE Task Force on Monitoring & Assessment, Working Pro-
gramme 1994/1995, Volume:5: State of the Art on Monitoring and Assess-
ment of Rivers, RIZA report: 95.068.

sources Research 6(3), 919-926.

Sharp, W.E. (1971) A topologically optimum water - sampling plan for rivers and
streams, Water Resources Research 7(6), 1641-1646.

Quality Assurance,. RIZA report: 95.067.
