Environ Monit Assess (2008) 139:205–215
DOI 10.1007/s10661-007-9828-6

Siting analyses for water quality sampling in a catchment

Jehng-Jung Kao · Pei-Hao Li · Chin-Lien Lin · Wen-Hsin Hu

Received: 13 November 2006 / Accepted: 18 May 2007 / Published online: 17 June 2007

Abstract Pollution loads discharged from upstream development or human activities significantly degrade the water quality of a reservoir. The design of an appropriate water quality sampling network is therefore important for detecting potential pollution events and monitoring pollution trends. However, under a limited budgetary constraint, how to site an appropriate number of sampling stations is a challenging task. A previous study proposed a method applying the simulated annealing algorithm to design the sampling network based on three cost factors including the number of reaches, bank length, and subcatchment area. However, these factors are not directly related to the distribution of possible pollution. Thus, this study modified the method by considering three additional factors, i.e. total phosphorus, nitrogen, and sediment loads. The larger the possible load, the higher the probability of a pollution event may occur. The study area was the Derchi reservoir catchment in Taiwan. Pollution loads were derived from the AGNPS model with rainfall intensity estimated using the Thiessen method. Analyses for a network with various numbers of sampling sites were implemented. The results obtained based on varied cost factors were compared and discussed. With the three additional factors, the chosen sampling network is expected to properly detect pollution events and monitor pollution distribution and temporal trends.

Keywords Nonpoint source pollution · Water quality sampling · Siting analysis · Multi-objective model · Optimization · Environmental systems analysis

Introduction

Reservoirs are vital water sources in Taiwan and significantly influence the livelihood of the society and national economy. However, the reservoir water quality is suffering adverse impacts from the non-point source (NPS) pollutants generated from upstream development and activities accelerating eutrophication and silting of the reservoir to affect the normal water use and increase the cost of water treatment. Establishing a proper water quality sampling network in a reservoir catchment is therefore essential for monitoring pollutant loads, distribution and events.

This study focuses on searching for an optimal spatial distribution of sampling sites in the upstream catchment of a reservoir. Various issues must be evaluated for determining the locations of water quality sampling sites. As described by Dixon and Chiswell (1996), the selection of sampling sites should consider information goals, indicators, data analyses, etc. In general, the numbers of sampling

J.-J. Kao (✉) · P.-H. Li · C.-L. Lin · W.-H. Hu
Institute of Environmental Engineering,
National Chiao Tung University,
75 Po-Ai Street, Hsinchu, 300 Taiwan,
Republic of China
e-mail: jjkao@mail.nctu.edu.tw

sites is limited by the available budget and cannot be widely and densely distributed. Dixon et al. (1999) also indicated that the cost involved in the subsequent investigation to find the source of a detected pollution event should be considered in determining proper sampling sites. Sanders et al. (1983) pointed that location is the most critical factor in designing a sampling network. Therefore, how to determine appropriate locations for installing an effective and representative sampling network is a research challenge. Sharp (1971) used the Shreve stream order number (Shreve 1967) and a successive division algorithm to find topologically optimum sampling sites. There were also some other related researches reported, but most of them had not applied an optimization-based approach (Warry and Hanau 1993; Loftis et al. 1991; Whitlach 1989; Lettenmaier et al. 1984; Sanders et al. 1983; Lettenmaier 1978; Lettenmaier and Burges 1977). Although some other researches had applied optimization models to determine the locations of sampling sites, the identification capability to locate the source of a detected pollution event is still not considered (Strobl et al 2006; Icaga 2005; Ning and Chang 2002, 2004, 2005). As described by Dixon et al. (1999), the method proposed by Sharp (1971) may not be capable of locating the optimal placement of sampling sites. Dixon et al. (1999) thus proposed a siting method based on the simulated annealing (SA) algorithm, graph theory, and a geographical information system. The cost function of their optimization model is formulated based on the expected cost for obtaining the information subsequent to detecting the water quality problem at a sampling site. Three cost functions they proposed are derived from the ratios of the number of reaches, upstream bank length, and subcatchment areas, respectively. However, these three ratios are not directly related to the pollution distribution characteristics. Sampling sites should be available for hot spots where the critical water quality is likely to occur, although sampling sites for other areas in a catchment are also required to avoid information bias.

Pollution loads originate from point or non-point source pollutants in a catchment are carried by runoff from storm events. In this study, the rainfall spatial variation is analyzed by using the Thiessen method (Lebel et al. 1987; Thiessen 1911). The Agricultural Non-Point Source Pollution (AGNPS) (Young et al. 1987) model is applied to simulate and estimate the distributions of total phosphorus, total nitrogen, and sediment. Based on the three pollutant distributions, three new cost functions are formulated. The model with the new cost functions is solved using the same SA algorithm for determining the locations of water quality sampling sites in Derchi reservoir catchment. Comparisons with results obtained from different cost functions including the three ones used by Dixon et al. (1999) are made and discussed.

Spatial variation of rainfall intensity

Nonpoint source pollutants are transported by runoff from storm events and thus are highly correlated with the rainfall intensity. There are 10 rainfall gauge stations in the study area; the rainfall spatial variation can be observed from the data monitored at these stations. For those subcatchments without a gauge station, their rainfall intensities are estimated using the weighting method (Lebel et al. 1987) as described below.

Let h_j denote the rainfall intensity at monitoring station j, and the weight of station j for subcatchment p is w_{p_j}, then the rainfall intensity in subcatchment p is estimated by:

$$h_p = \sum_{j}^{n} w_{p_j} h_j \tag{1}$$

Where the weight, w_{p_j}, is determined by:

$$w_{p_j} = \frac{a_{p_j}}{a_p} \tag{2}$$

Where a_{p_j} is the area of subcatchment p overlapped with the representative area covered by station j, and a_p is the area of subcatment p. The representative area covered by a gauge station is determined by using the Thiessen method (Thiessen 1911). All points in the representative area of a gauge station are closer to the station than any other stations.

Modeling simulation for pollution distribution

The AGNPS (Young et al. 1987) model is employed to simulate the NPSP loading distribution in the studied catchment. AGNPS is a physical- and grid-based model developed for evaluating upstream land erosion and water quality. In applying AGNPS, the catchment is divided into numerous rectangular land grids and parameters for each grid were collected on cover type, soil hydrologic group, field slope length, manning coefficient, coefficients used in the USLE equation, soil condition, soil texture, fertilizer amount, fertilizer incorporation, point source information, impoundment factor, channel indicator and slope, etc. The data are collected from various geographical maps, field investigations, and values suggested from previous research or the model manual.

Cost functions

Six cost factors are considered and formulated as the objective functions of the optimization model for selecting sampling sites. The first three cost functions are adopted from Dixon et al. (1999) and the other three are proposed in this study based on the pollutant distributions estimated using the AGNPS model. These cost functions are respectively described below.

REACH: The Ratio of The Number of Reaches

As described by Dixon et al. (1999) for this cost factor, the occurrence probability for potential pollution, which is defined to be proportional to the ratio of the number of reaches in a subcatchment to the total number of reaches in the entire catchment, is formulated as below:

$$P_i = \frac{m_i}{m_0} \tag{3}$$

where m_i is the number of reaches in subcatchment i and m_0 is the total number of reaches in the entire catchment. The cost function for this factor is expressed as:

$$E_{\text{cost}} = \sum_i P_i E_{w_i} \tag{4}$$

where E_{cost} is the expected cost; P_i the probability that the pollution source is in subcatchment i; E_{w_i} is the expected additional effort required to locate the source once it has been detected in subcatchment i. Assuming a binary search of the upstream reaches for locating the source, a good approximation of the mean number of samples required for the detection is $\log_2 m_i$. Therefore, the cost function can be formulated as follows.

$$E_{\text{cost}} = \sum_i \left(\frac{m_i}{m_0}\right) \log_2 m_i \tag{5}$$

Length: The ratio of the bank length

For this factor, the occurrence probability is assumed to be proportional to the ratio of the bank length in a subcatchment to the total bank length in the catchment (Dixon et al. 1999) and is defined as:

$$P_i = \frac{L_i}{L_0} \tag{6}$$

where L_i is the upstream bank length in subcatchment i and L_0 is the total bank length in the entire catchment. With the similar binary search assumption, the cost function can be formulated as follows.

$$E_{\text{cost}} = \left(\frac{L_i}{L_0}\right) \log_2 m_i \tag{7}$$

AREA: The sum of subcatchment areas squared

For this factor, the occurrence probability is assumed to be proportional to the subcatchment area, i.e. a larger the area leads to a greater probability of the pollution event; the expected effort required to locate a source is also proportional to the area (Dixon et al. 1999). The cost function can be formulated as follows.

$$E_{\text{cost}} = \sum_i a_i^2 \tag{8}$$

where a_i is the area of subcatchment i.

TP: Phosphorus loads

This cost factor is used to evaluate the pollution potential based on the magnitude of the phosphorus load generated from a subcatchment. The phosphorus load in the upstream catchment of a reservoir is primarily generated from distributed sources. The effort required to identify the source is expected to be proportional to the magnitude of the load as well as the downstream drainage boundary length, pollution type, and discharge frequency of a distributed source, e.g. an agricultural cropping area. Since the determination of downstream drainage boundary lengths and other factors of distributed sources is tedious, the expected effort required to identify the source is assumed to be proportional to the magnitude of the load estimated based on AGNPS modeling results. The expected cost function of each factor is calculated based on the probability of pollution event occurrence in a subcatchment multiplied by the expected effort required to locate the pollution source(s). For the three non-point source pollution load based cost functions, pollution loads are primarily generated from distributed sources, and a high pollution load indicates that a large area or a large number of pollution sources exists. A pollution event is likely to happen in a place with a high pollution load and its probability is assumed to be proportional to its estimated load. Furthermore, the effort required to identify the source when an event occurs is also expected to be proportional to the estimated load. Therefore, the cost functions for the proposed cost factors, such as phosphorous, can be expressed as follows.

$$E_{\text{cost}} = \sum_i p_i^2 \tag{9}$$

where p_i is the estimated phosphorus load generated from subcatchment i that is determined from the simulation result by using AGNPS. The cost function is intended to find a placement of sampling sites with each site monitoring an area with equal phosphorus load.

TN: Nitrogen loads

The magnitude of the estimated total nitrogen load generated from a subcatchment can be used as a cost factor too. The larger the estimated load, the higher the probability of a pollution event may occur. Therefore, similar to that for the phosphorus load factor, the cost function based on the total nitrogen load is formulated as follows.

$$E_{\text{cost}} = \sum_i n_i^2 \tag{10}$$

where n_i is the estimated total nitrogen load generated from subcatchment i, which is determined from the simulation result by using AGNPS.

SED: Sediment loads

Similar to those for total phosphorus and nitrogen loads, the magnitude of the estimated sediment load generated from a subcatchment can be used as a cost factor also and the associated cost function is formulated as follows.

$$E_{\text{cost}} = \sum_i s_i^2 \tag{11}$$

where s_i is the estimated sediment load generated from subcatchment i, which is determined from the simulation result by using AGNPS.

Simulated annealing algorithm

The simulated annealing algorithm used by Dixon et al. (1999) is applied to solve the optimization model with a computer program developed in this study. A cooling schedule is pre-defined before implementing the algorithm by setting an initial temperature (C_0), the number of iterations for each temperature cooling step, the factor (<1) for lowing temperature at each cooling step, and the number of desired cooling steps or final temperature. The decision variables used in the SA algorithm are the variables which represent the selected locations for the placement of sampling stations. An initial placement of sampling sites is randomly generated and the associated cost is computed according to a pre-specified cost function. Then, a sampling site is randomly selected to alter its location, if the network is improved after the alteration. Only one of adjacent upstream and downstream movements is allowed to alter the location of the selected site, and the movement is randomly selected. If the upstream movement is selected, a further random choice is made for which upstream branch to be moved to, if multiple upstream branches exist. Once a new placement is determined, the associated new cost is computed. If the new cost is better than the previous one, a subsequent iteration is then initiated to continue the task. If a worse placement was generated, another random number (r), whose value is between 0 and 1, is generated. The iteration is continued while r is less than a pre-specified function, $\exp\left(-^{\Delta}E/C\right)$ or the current placement is discarded and another placement is tried. This approach increases the possibility in locating the global optimum instead of being trapped at local optima. The procedure is repeated until the desired number of iterations is implemented and the temperature is cooled by one step. The whole procedure is repeated until the specified number of temperature cooling steps has been achieved.

Study area and NPSP distribution

The Derchi reservoir, storing water of approximately 2.5×10^8 m^3 in central Taiwan, is a major source of local drinking water as illustrated in Fig. 1. The catchment area of the reservoir is about 602 km^2 divided into 63 subcatchments. The major land use within the area is orchard, with approximately 26.2 km^2 under cultivation. The total phosphorus concentration of the reservoir water body ranges between 40 and 140 ppb. According to Carlson's (1977) trophic state index, this reservoir is eutrophic and requires attention for control of water quality.

The pollution distribution in the catchment had been estimated by the same modeling approach used by Kao and Tsai (1997) and Lin and Kao (2003) based on the simulation result provided by the AGNPS model (Young et al. 1987). Effective individual storm events were identified from the rainfall records monitored in 10 gauging stations. A storm event with rainfall intensity exceeding 12.7 mm (USDA, 1978) is regarded as an effective event capable of washing out a significant amount of NPSP. The rainfall intensity varies spatially and temporarily in the catchment, thus, estimation of the NPSP distribution in the entire area based on a single value may not be appropriate. Therefore, the Thiessen method (e.g., Lebel et al. 1987) was adopted in this study to estimate the rainfall intensity in each subcatchment.

Figure 2 illustrates the NPSP distributions of total phosphorus, total nitrogen, and sediment loads of subcatchments simulated by using the AGNPS model. The darker area in the figure indicates the higher NPSP load and vice versa. The distribution of total phosphorus loads is similar to that for total nitrogen loads, but the distribution of sediment loads is slightly different from the distributions of either phosphorus or nitrogen.

Result and discussion

The SA method was applied to solve the optimization model based on each of the six cost functions to select water quality sampling sites for the study area under

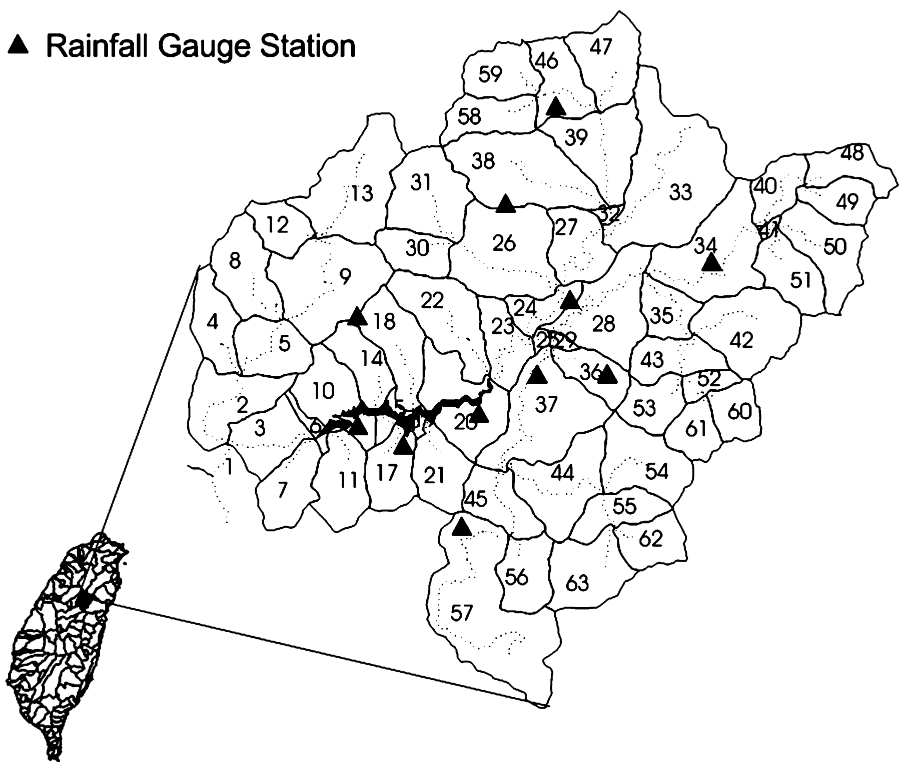

Fig. 1 Derchi reservoir in central Taiwan

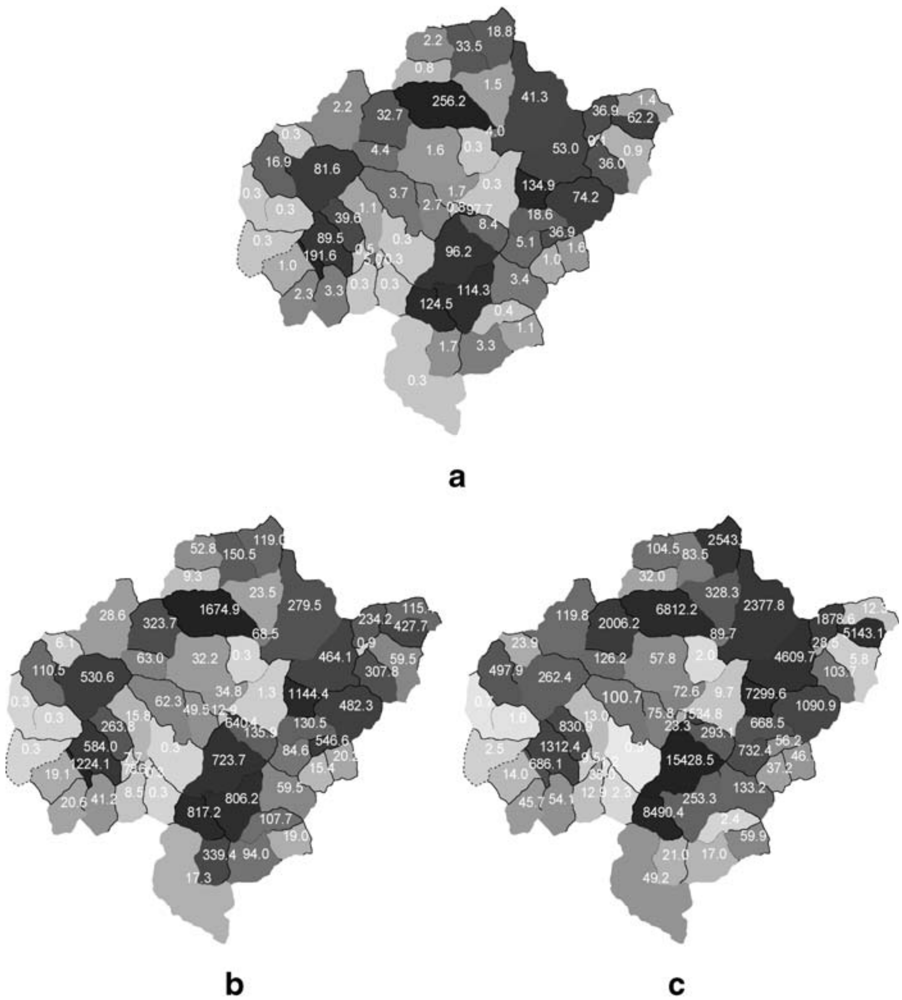

Fig. 2 (**a**) Total phosphorus; (**b**) total nitrogen; and (**c**) sediment loads (unit: kg) in each subcatchment

varied limits on the number of sites, i.e. five, six, seven, and eight. The initial temperature, the maximum number of iterations, the number of steps in each iteration, and the cooling rate used in this study are 109, 1000, 15, and 0.9, respectively. Table 1 lists the payoff table for the values of all cost functions for all solutions. The scales for each cost function in the table are different because the types and associated units of the different cost functions are also different. If a solution is obtained by using a specific cost function as the objective, the value for that cost function is expected to be smaller than those of other solutions. For example, the LENGTH cost function value for the five-site solution using the LENGTH cost function is 3.67, and is the smallest among all five-site solutions. Although this situation is not always true for some other sets of solutions, the difference between the one expected to have the optimal value and the smallest one obtained is generally trivial. Some possible reasons for such differences are that the SA method is a heuristic method and it may encounter a premature termination during the SA search procedure without locating the global optimum. Moreover, the number of iterative steps may be not numerous enough to cover the global optimum because the number of possible combinations of feasible locations which expands dramatically when the number of desired sites increases.

The sampling sites selected for five, six, seven, and eight-site solutions in the catchment are illustrated in Fig. 3. For all selected solutions for various numbers of sites, the sampling sites are placed at major branches on the stream network in several subcatchments, such as subcatchments 25, 29, and 37, because monitoring these major branches is critical to identify pollution events. However, significant differences can be observed among the solutions. For example, as

Table 1 The payoff table for solutions obtained based on different cost functions for varied limits on the number of sites

Cost function used	Cost function values of obtained solutions					
	REACH	LENGTH	AREA$_{(10^{16})}$	TP$_{(10^{7})}$	TN$_{(10^{5})}$	SED$_{(10^{8})}$
5 sites						
REACH	3.75	3.7	10.2	7.16	4.74	14
LENGTH	3.78	3.67	9.8	7.4	4.9	13.8
AREA	3.74	3.69	10.2	7.29	4.73	13.6
TP	3.99	4	13.7	7.7	4.96	15.1
TN	3.74	3.69	10.2	7.29	4.73	13.6
SED	3.73	3.71	10.4	7.16	4.74	14
6 sites						
REACH	3.51	3.44	8.85	7.32	4.8	13.8
LENGTH	3.69	3.48	8.99	6.44	4.32	14.1
AREA	3.44	3.45	8.85	7.06	4.63	13.9
TP	3.59	3.48	8.99	6.44	4.32	14.1
TN	3.63	3.55	9.31	6.38	4.28	12.9
SED	3.72	3.59	9.32	7.24	4.8	13.4
7 sites						
REACH	3.36	3.33	8.19	6.72	4.4	13.2
LENGTH	3.27	3.3	8.1	6.01	4.01	12.3
AREA	3.41	3.34	8.13	6.72	4.41	13.2
TP	3.34	3.51	8.33	6	4	12.3
TN	3.61	3.59	9.6	6.28	4	11.6
SED	3.39	3.39	8.86	5.72	3.69	10.6
8 sites						
REACH	3.15	3.14	7.89	5.33	3.45	10.2
LENGTH	3.07	3.15	7.47	5.77	3.7	12.7
AREA	3.34	3.31	9.03	6.09	4.19	11.4
TP	3.54	3.49	9.42	5.57	3.51	10
TN	3.54	3.49	9.42	5.57	3.51	10
SED	3.27	3.28	8.25	5.9	3.93	10.1

illustrated in Fig. 3(d), the SED and TN solutions place a sampling site in subcatchments 34 and 35, respectively. The pollution loads generated from the subcatchments upstream to subcatchment 28 are significantly high because there are numerous cropping activities in those subcatchments. Therefore, a sampling site is necessary to be placed in subcatchment 34 or 35 to balance the monitoring coverage of each site. Similar observations can be found in five-site, six-site, and eight-site solutions. Although three new cost factors are all pollution loads, their solutions are not the same. The sediment load is mainly caused by land erosion, while TP and TN loads are mainly determined by the type of cropping practices and fertilizer applications. Solutions obtained based on the three pollution cost factors are thus different. For instance, for six-site, seven-site and eight-site solutions, a sampling site is placed in subcatchment 45 because significant erosion frequently occurs in that area. On the other hand, the REACH and LENGTH cost factors represent geographic properties, sites selected based on these two factors are distributed more evenly in the catchment. Subcatchments located at the upstream ends of a river are not selected for these two factors to avoid unbalancing coverage of the selected monitoring site, except for the eight-site LENGTH solution that maybe a local optimum as illustrated in Fig. 3(d) which indicates that end subcatchments 22 and 57 were selected.

The planning results based on the three geographic factors proposed by Dixon et al. (1996) tend to give a set of sampling stations to cover areas of similar size without due consideration of the distribution of potential pollution sources and loads. Therefore, the

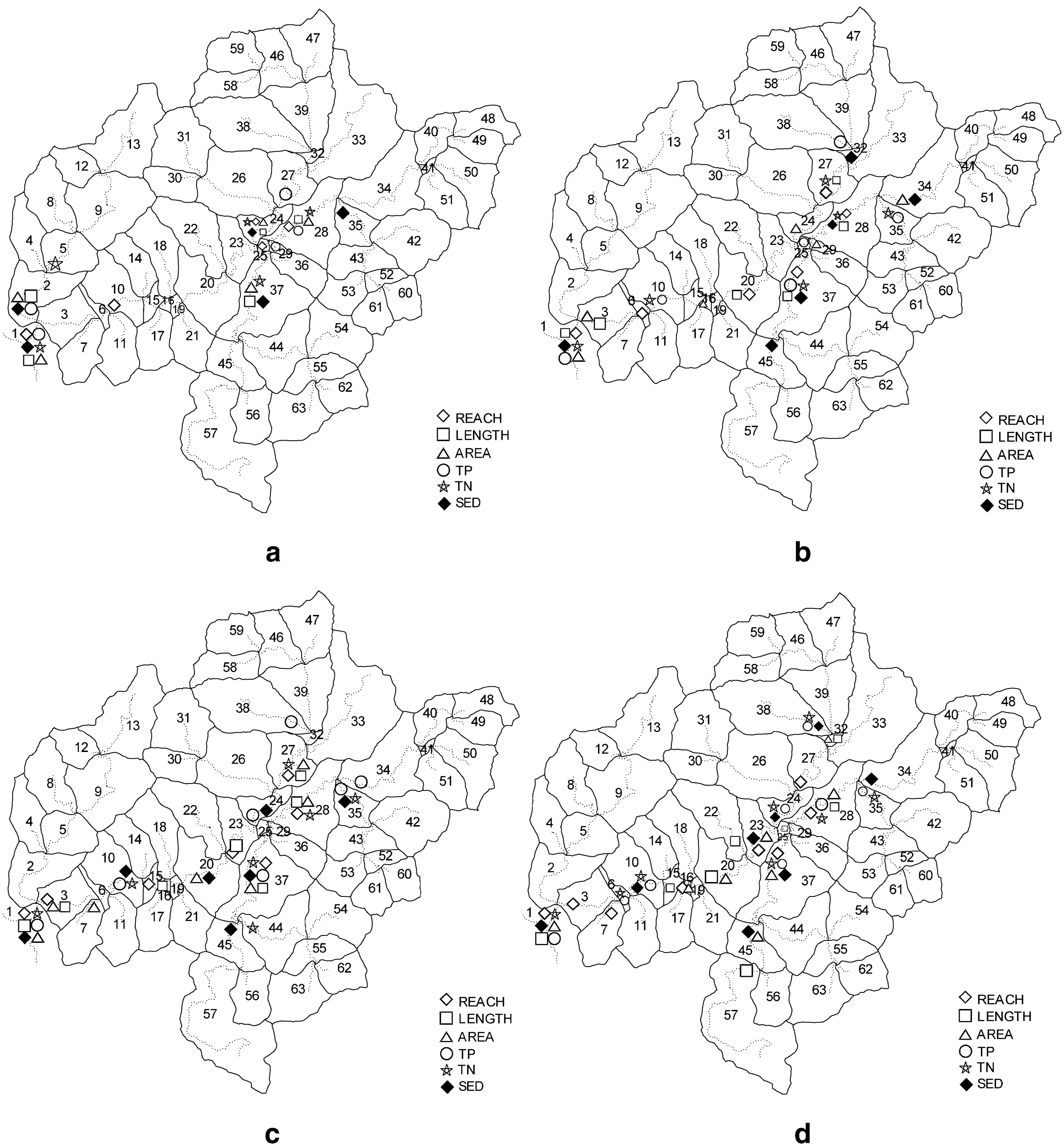

Fig. 3 (**a**) 5-site, (**b**) 6-site, (**c**) 7-site, and (**d**) 8-site solutions for the six cost functions

effort required to detect the pollution source when a water quality event occurs may be considerable. For example, subcatchment 28 has been selected as a location for a station by using one of the three geographic cost factors: REACH, LENGTH, or AREA. However, subcatchments 34 and 35, which both have significant cropping activities and high pollution potential, would be covered by this station and considerable effort would be required to locate the pollution source when an event occurs. Therefore, by applying the cost factor of TP or TN, in addition to subcatchment 28, subcatchments 34 or 35 is also selected to reduce the possible effort required to locate the pollution source when a water quality event is detected.

Figure 4 shows the cost function values of varied solutions obtained based on six cost functions. The cost to identify a pollution event is decreased while the number of sites increases because both the coverage of each site and the effort required to detect a pollution event are reduced. As illustrated in Fig. 4(a) and (b), the cost function values for all solutions are reduced when the number of sites increases, however, increasing the number of sites also increases the total cost for establishing the sites. Therefore, the available budget, the desired level of effort for detecting a pollution event, and the trade-off between the total cost and cost function values should be carefully examined for determining a proper number of sites. Increasing the number of sites introduces additional decision variables that may lead to a premature termination when using the same search parameters and thus may not locate the global

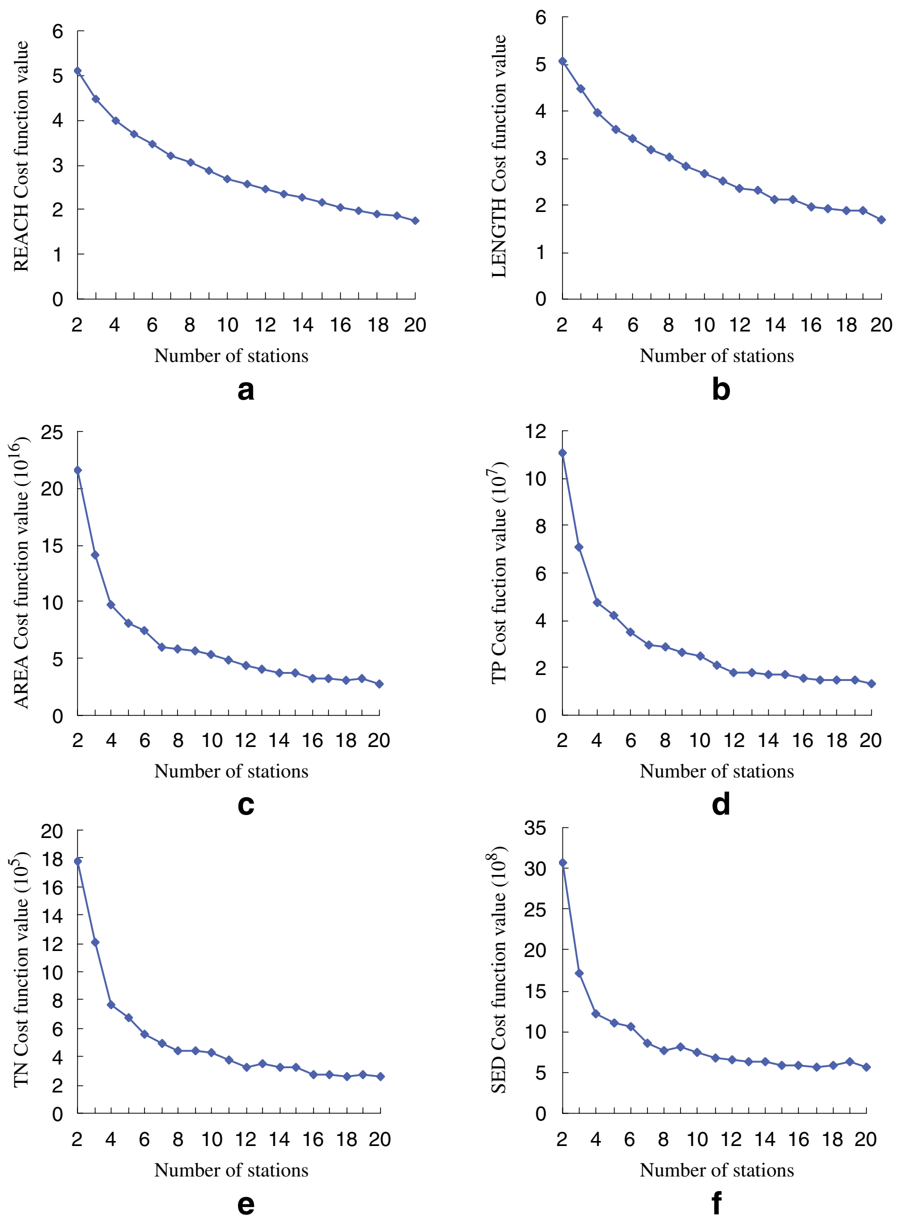

Fig. 4 Cost function values of varied solutions obtained based on (**a**) number of reaches; (**b**) bank length; (**c**) subcatchment areas squared; (**d**) phosphorus load; (**e**) nitrogen load; and (**f**) sediment load

optimum. For example, as illustrated in Fig. 4(f), the cost function value for the nine-site solution is even larger than that for the eight-site solution.

Conclusion

This study enhances the water quality sampling siting model proposed by Dixon et al. (1999) with three pollution load based cost factors. The REACH, LENGTH, and AREA cost functions proposed by Dixon et al. (1999) are not directly related to possible pollution distribution and thus the sites determined based on them may not be adequate for detecting pollution events. An area with a high potential pollution load indicates extensive development or human activities and a water quality event is likely to happen. Therefore, proper monitoring is required for such an area. Furthermore, the monitoring goal of a sampling network is not only focused on detecting unusual pollution events, but also on monitoring the pollution distribution and temporal trend. The sampling sites placed based on the proposed pollution load cost factors are appropriate to fulfill these monitoring goals.

If a significant number of human or agricultural cropping activities exist in the area studied, the three cost factors proposed in this study should be evaluated prior to the other three factors previously described by Dixon and Chiswell. (1996). The reason for this is that the estimated pollution loads are more suitable than the other three factors for locating a pollution source when a water quality event occurs. A sampling station monitors not only the subcatchment where it is located, but also the upstream subcatchments that are not covered by the other stations. The planning results based on the three factors proposed by Dixon and Chiswell. (1996) may include stations that cover subcatchments with high pollution loads which are generated by a large number or area of human activities. The expected effort needed to detect a pollution source when a water quality event occurs might thus be considerable.

To make the most informed decisions for the locations of sampling stations based on the obtained results, a further multi-factor analysis to evaluate the six cost factors simultaneously is desirable. Although such an analysis is beyond the scope of this research, several simple rules to reach a final decision can be suggested. For the study area, the phosphorus load is the main cause of eutrophication; however, the sediment load is also significant. Therefore, the TP and SED solutions have been recommended for use in the studied area. However, some differences exist between both TP and SED solutions. For example, the difference between both six-site TP and SED solutions is that the subcatchments 10 and 35 are selected when the TP cost factor is applied, while subcatchments 34 and 45 are selected when the SED cost factor is used. Although subcatchments 10 and 35 do not balance the sediment loads covered by the sample stations as effectively as do subcatchments 34 and 45, the deviation involved is still quite close to that of the SED optimal solution, only about 10% difference. Therefore, the TP alternative is recommended.

In case the TP and SED solutions are significantly different or other factors should also be evaluated, the following weighted cost function can be used.

$$\sum_i W_i \cdot c_i$$

where W_i is the weight of cost factor i, c_i, for which c_i's are the six cost factors. For example, an objective function such as TP+0.8 SED+0.2 TN+0.1 AREA may be applied. However, different sets of weights would generate different solutions and a further multi-factor analysis would be required to make the final decision.

Since the simulated annealing approach does not always provide the true optimal solution and may subsequently cause some problems in comparing the performance of varied cost functions, an on-going research is currently implemented to develop a mathematical model for remedying the situation. A preliminary model has been developed and will be tested in the near future and the results will be reported once they are available.

Acknowledgement The authors would like to thank National Science Council, R.O.C. for providing partial financial support of this research under Grant NO. NSC91-2211-E-009-015 and NSC93-2211-E-009-020. Valuable comments and suggestions provided by anonymous reviewers are also greatly appreciated. Without them, our manuscript can not be significantly improved.

References

Carlson, R. E. (1977). A trophic state index for lakes. *Limnology and Oceanography, 22*, 361–369.

Dixon, W., & Chiswell, B. (1996). Review of aquatic monitoring program design. *Water Research, 30*(9), 1935–1948.

Dixon, W., Smyth, G. K., & Chiswell, B. (1999). Optimized selection of river sampling sites. *Water Research, 33*(4), 971–978.

Icaga, Y. (2005). Genetic algorithm usage in water quality monitoring networks optimization in Gediz (Turkey) River Basin. *Environmental Monitoring and Assessment, 108*, 261–277.

Kao, J.-J., & Tsai, C.-H. (1997). Multiobjective zone TP reduction analyses for an off-stream reservoir. *Journal of Water Resources Planning and Management (ASCE), 123* (4), 208–215.

Lebel, T., Bastin, G., Obled, C., & Creutin, J. D. (1987). On the accuracy of a real rainfall estimation: A case study. *Water Resources Research, 23*(11), 2123–2134.

Lettenmaier, D. P. (1978). Design considerations for ambient stream quality monitoring. *Water Resources Bulletin, 14* (4), 884–992.

Lettenmaier, D. P., Anderson, D. E., & Brenner, R. N. (1984). Consolidation of a stream quality monitoring network. *Water Resources Bulletin, 20*(4), 473–481.

Lettenmaier, D. P., & Burges, S. J. (1977). Design of trend monitoring network. *Journal of Environmental Engineering Division (ASCE), 103*, 785–802.

Lin, Y.-C., & Kao, J.-J. (2003). Effects of seasonal variation in precipitation on estimation of non-point source pollution. *Water Science and Technology, 47*(7–8), 299–304.

Loftis, J. C., Taylor, C. H., & Chapman, P. L. (1991). Multivariate tests for trend in water quality. *Water Resource Research, 27* (7), 1419–1429.

Ning, S. K., & Chang, N. B. (2002). Multi-objective, decision-based assessment of a water quality monitoring network in a river system. *Journal of Environmental Monitoring, 4*, 121–126.

Ning, S. K., & Chang, N. B. (2004). Optimal expansion of water quality monitoring network by fuzzy optimization approach. *Environmental Monitoring and Assessment, 91*, 145–170.

Ning, S. K., & Chang, N. B. (2005). Screen the relocation strategies of water quality monitoring stations by compromise programming. *Journal of the American Water Resources Association, 41*(5), 1039–1052.

Sanders, T. G., Ward, R. C., Loftis, J. C., Steele, T. D., Adrian, D. D., & Yevjevich, V. (1983). *Design of networks for monitoring water quality*. Littleton, CO: Water Resources Publications.

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

Shreve, R. L. (1967). Infinite topologically random channel networks. *Journal of Geology, 75,* 178–186.

Strobl, R. O., Robillard, P. D., Shannon, R. D., Day, R. L., & McDonnell, A. J. (2006). A water quality monitoring network design methodology for the selection of critical sampling points: Part I. *Environmental Monitoring and Assessment, 112*, 137–158.

Thiessen, A. H. (1911). Precipitation averages for large areas. *Monthly Weather Review, 39*(7), 1082–1084.

Warry, N. D., & Hanau, M. (1993). The use of terrestrial ecosystems as a regional-scale screen for selecting representative sites for water quality monitoring. *Environment Management, 17*(2), 267–276.

Whitlatch, E. E. (1989). Spatial adequacy of Nasqan water quality data in Ohio River Basin. *Journal of Environmental Engineering (ASCE), 115*(1), 173–191.

Young, R. A., Onstad, C. A., Bosch, D. D., & Anderson, W. P. (1987). *AGNPS, Agricultural non-point source pollution model*. Minnesota, USA: USDA-ARS.