A novel approach for selecting sampling points locations to river water quality monitoring in data-scarce regions

Hossein Alilou[1], Alireza Moghaddam Nia[2*], Mohsen Mohseni Saravi[3], Ali Salajegheh[4], Dawei Han[5], Bahram Bakhtiarienayat[6]

[1] Faculty of Natural Resources, University of Tehran, Iran.
Email: h.aliloo@ut.ac.ir

[2] Faculty of Natural Resources, University of Tehran, Iran.
(Corresponding author) Email: a.moghaddamnia@ut.ac.ir

[3] Faculty of Natural Resources, University of Tehran, Iran.
Email: msaravi@ut.ac.ir

[4] Faculty of Natural Resources, University of Tehran, Iran.
Email: salajegh@ut.ac.ir

[5] Department of Civil Engineering, University of Bristol, Bristol, BS8 1TR, UK.
Email: dawei.han@bristol.ac.uk

[6]Faculty of Natural Resources, University of Tehran, Iran.
Email: bakhtiare@ut.ac.ir

Abstract

In order to rationalize a surface water quality monitoring network (WQMN), it is critical to appropriately design surface water quality sampling locations. This is due to high installation, operational, and maintenance costs for each sampling representative of the whole water system conditions. The main objective of this study was to propose an integrated method to determine the most appropriate sampling points in the Khoy watershed northwest of Iran, where financial resources and water quality data are limited. Multi criteria evaluation method including analytic network process (ANP) and Fuzzy logic were incorporated in River Mixing Length (RML) procedure in order to identify exact locations of sampling points. Based on RML procedure, 15 candidate sampling points were identified to suitably select sampling points based on budget deficiency. Relative weights for 12 criteria and 10 sub-criteria related to non-point sources and surficial rocks as well as criteria of topography were then calculated by the ANP method. According to the obtained results, a new total potential pollution score (TPPS) was presented to prioritize 15 candidate sampling points. Then, the values of TPPS were classified and fuzzified to distinguish real differences between scores. Based on current monitoring stations and budget deficiency, the hierarchy value, and Fuzzy rank, six points are proposed as the most appropriate locations for surface water quality monitoring. Furthermore, four points are identified as the second most appropriate for enhancing a robust WQMN in the study area in order for an expansion plan in the future. The results of this study should be valuable for water quality monitoring agencies looking for a cost-effective approach for selecting exact sampling locations.

Keywords: Water quality monitoring network; cost-effective siting sampling locations; Mixing Length procedure; TPPS; ANP; Fuzzy logic.

1. Introduction

Since industrial revolution, human activities have had negative repercussions on both quality and quantity of the surface water resources. Most countries and researchers have attempted to develop a variety of approaches in order to assess, evaluate, and monitor water quantity and quality in the watersheds (Baltacı et al., 2008; Behmel et al., 2016). Water quality monitoring (WQM) is a technical term and good example of this scientific endeavor in the realm of water research (Sanders et al., 1983; Chapman, 1996). The main purposes of WQM include understanding of watershed health and dynamic, current conditions and trends in the surface water system, and providing reliable information to help decision-makers to interpret and manage the stakeholders' health risk (Park et al, 2006; Strobl et al., 2006 a; Baltacı et al., 2008; Telci et al., 2009; Xiaomin et al., 2016). According to the literature on the WQMN design, every surface water monitoring network has the main tasks including definition of monitoring goals, proper locations of sampling points, selection of parameters and methods under consideration, and determination of sampling recurrence and frequencies (Telci et al., 2009; Behmel et al., 2016). For this purpose, the location of the most appropriate sampling points is a vital factor in the WQMN design (Sanders et al., 1983; Varekar et al., 2015b). In addition, appropriate monitoring sites play a key role in integrated watershed management (IWM), management of the total maximum daily load (TMDL), and improving water quality models (Park et al, 2006; Telci et al., 2009; Chen et al., 2012). More specifically, from the cost and time-efficient perspective, it is essential to appropriately locate sampling points for assessment and evaluation of temporal and spatial changes of water quality (Kovacs et al., 2016; Behmel et al., 2016).

A comprehensive literature exists on selection and optimization of sampling sites in WQMN design. According to the literature, the multivariate statistical techniques (Ouyang, 2005; Chilundo et al., 2008; Noori et al., 2010; Wang et al., 2014) and the Genetic Algorithm(GA) have been employed to select representative sampling points (Icaga, 2005; Park et al., 2006; Chen et al., 2008; Karamouz et al., 2009; Telci et al., 2009; Liyanage et al., 2016). In recent years, a combination of numerical models, experiments, and matter-element analysis has been used to assess WQMNs (Chen et al., 2012, Keum and Kaluarachchi, 2015). In addition, multi objective analysis has been increasingly utilized to optimize and propose monitoring sites (Ning and Chang, 2002; Khalil et al., 2011; Aboutalebi et al., 2016). The other methods like Geostatistical techniques (Beveridge et al., 2012) and Entropy approach (Memarzadeh et al., 2013; Mahjouri and Kerachian, 2011) have been employed to appropriately locate sampling points. With presence of numerous frameworks and guidelines on selection and optimization of the sampling point numbers as well as WQM program, most of them have still not been universally utilized or accepted to date (Varekar et al., 2015; Behmel et al., 2016). It is worth mentioning that most of researches conducted on WQMN concentrate chiefly on mathematical facets (Do et al., 2012). In addition to above stated approaches, some researchers have introduced alternative techniques for properly designing WQMN and locating sampling points. Sharp's procedure, as a systematic approach (Sharp, 1971) was modified by Sanders et al. (1983) in order to identify exact locations of sampling points. In Sharp's procedure, river network is subdivided into equal segments which are selected as sampling points by identifying the centroids, while in Sanders' procedure, pollution loadings and the number of outfalls are employed (Varekar et al., 2015b). Although several studies have been conducted by using both these methods (Park et al., 2006; Do et al., 2011; Varekar et al., 2015a, b), there are some limitations in employing these methods for a river without tributaries as well as short or long rivers. Moreover, in order to use these procedures, reliable and long-term data collection of water quality must be in place, which is often not feasible in developing country (e.g. Iran).

Do et al. (2012), in turn, proposed a new procedure by modifying Sanders' procedure and taking river mixing length (Day, 1977) into account for removing the aforementioned problems. This modified approach was incorporated with pollution potential of each land-use by considering event mean concentration (EMC) and human activities. It is highly suitable for a river system suffering from inaccurate and unreliable data on hydraulic and flow characteristics (Do et al., 2012). However, limitations of study conducted by Do et al. (2012) was to use analytic hierarchy process (AHP) which do not consider inter-dependencies of criteria (correlation between water quality variables). Since water quality parameters are not independent of each other (Newell et al., 1992; Barid et al., 1996; Harper, 1998; Baldys et al., 1998; Line et al., 2002), a robust multi-criteria decision making approach such as the ANP method is needed to consider their relationship. Considering distance zone based on linear surface ground was the other limitations (Varekar et al., 2015a). More importantly, watershed geology (surficial rocks), which plays a vital role in the chemistry of water bodies (McCartan et al., 1998; Zektser et al, 2007; Olson, 2012), has not been considered in the recent studies (Sanders., 1983; Park et al., 2006; Do et al., 2011; Do et al., 2012; Varekar et al., 2015 a, b, Alilou et al., 2018). In addition, under case study of the Xiangxi River in China and the Portland Metropolitan area in the USA, Ye et al. 2009 and Pratt and Chang, 2012 showed that watershed topography is responsible for 25% of water quality variation mostly in dry season and more specifically some water quality variables, for example, nitrate nitrogen (NO3-N) and total phosphorus (TP) having negative and positive correlation with topography, respectively (D'Arcy and Carignan, 1997), was neglected. Therefore, it is of vital importance to consider watershed geology, topography, and interdependencies of water quality variables when it comes to design sampling points. Also, the lack of reliable and regular long-term water quality data collection as well as data on hydraulic and flow characteristics in Iranian watersheds motivate us to present this study.

The main objective of this study is to apply robust multi-criteria evaluation approach in the first stage by using analytic network process (ANP) in order to determine relative weights of water quality variables as well as proposing a new pollution potential for non-point sources, surficial rocks, and watershed topography. The second stage involves using the modified approach (Do et al., 2012) to select potential sampling points. The third stage is to determine total potential pollution scores (TPPS) for each candidate sampling point to rank priority for setting new monitoring stations. Last but not least, in order to prioritize ranking and distinguish real differences of each sampling point's scores, the fuzzy theory is applied in this study.

2. Material and methods

2.1. Description of the study area

The Khoy watershed located in West Azerbaijan province in Iran (Fig.1) has a drainage area of approximately 3166 km^2; its elevation varies significantly from about 938 m to 3670 m above sea level, with an average slope of 23.16 %. It consists of three rivers: Qutor Chai as the main stream (110.13 km long), Qudox Bogan (98 km long), and Gazan Chai (around 40 km long) flowing from Turkey Mountains to the Caspian Sea. The Khoy watershed has a semiarid climate with annual precipitation of 281.92 mm, which decreases from approximately 400 mm in the west with high elevation to about 190 mm in the north east. Nowadays, these rivers are facing several environmental issues and mismanagement: 1) rangeland overgrazing, which drives rapid erosion and transfer of sediment into rivers; 2) industrialization and land-use changes along the rivers, especially industrial park founded in upstream of the Gazan Chai; 3) currently irregular data collection and inappropriate location of current stations (please see Fig. 1). Fig. 1 indicates that the number of currently operated monitoring sites is six. Two of them are located in the south west of the study area where there is no highly populated area and anthropogenic activities. The water quality sampling frequency is one per month or one per season. Moreover, water quality variables measured are mainly concentrated on physical characteristics (e.g. temperature and the total solids content), chemical characteristics (e.g. major cations and anions), and inorganic Indicators (e.g. hardness and conductivity). However, organic materials (e.g. total phosphorus, total nitrogen, and nitrate nitrogen) and organic indicators (biochemical oxygen demand) are not sampled because of financial problems.

Thus, these issues made an urgent need to design and select robust sampling points for water quality monitoring, based on the traditional, recent policies and technologies objectives of monitoring networks, listed as follow (Liebetrau, 1979; Lettenmaier, 1979; Park et al., 2006): 1) understanding temporal variations of water quality parameters in short and/or long-term trends; 2) supporting application of water resources; 3) testing short-term changes in water quality; 4) early detection of pollution; 5) calculation of pollution loads of a given area to accomplish TMDL analyses; 6) creation of data-base system for water resources management. To achieve the aforementioned objectives of monitoring program, appropriate locations of monitoring networks paly main role.

2.2. Design of potential sampling point locations

Do et al., (2012) proposed the RML method to remove limitation of Sanders' procedure (Sanders et al., 1983) and identify sampling point locations in more detail. Fig.2 illustrates the differences between these two methods. In the mixing length method to compensate the lack of tributes and differences in the length of branches, rivers and branches are divided into small segments, which are equal to mixing length of rivers. "River mixing length is a distance over which an upstream water parcel will keep its original properties before dispersing those characteristics into the surrounding downstream water" (Do et al., 2012). Thus, each of the segments (river's mixing length) is considered as a potential sampling point. First, we calculated the mixing lengths for each branch or river only using a simple equation, L = 25W (Day, 1977; Do et al., 2012). To do so, Google earth software is used to measure the stream width because of its spatial resolution (15m-15cm) (http://earth.google.com). After that, 100 bridges over the rivers are measured by field works (Telci et al., 2009) to ensure the accuracy of the measured stream width. Then, Arc-GIS 9.3 is applied to divide a river into small segments, equal to the river mixing length with different lengths. Subsequently, the total number of segments of a branch or river is determined by Eq. (1) (Do et al., 2012). Finally, Eq. (2) is applied to determine the total number of segments for an entire river network or the number of total potential (Do et al., 2012).

$$N_j = \frac{l_j}{L_j} = \frac{l_j}{25W_j} \tag{1}$$

$$N = \frac{1}{25}\sum_{j=1}^{n}\frac{l_j}{W_j} \tag{2}$$

where lj is the total length of river j; n name of rivers; Lj indicates river's mixing length of each segment; Wj is the stream width; Nj is the total number of segments of river j; and N is the total number potential sampling points of entire river system.

In the second step, we used Eq. (3) introduced by Sanders et al., (1983) was used to determine the number of stations needed in the study area. In this study, based on existing stations and budget limitations of the regional water authority, i is assumed as four. Therefore, the number of stations need is 15.

$$S_i = 2^i - 1 \tag{3}$$

where *i* is hierarchy of sampling points and S_i is the number of stations; i is a natural number. A low-hierarchy value point has a higher priority than a high-hierarchy value point in selecting sampling points (Sanders et al., 1983) (Fig. 2). In the third step, the locations of 15 sampling points with different ith hierarchy values are identified by Eqs. (4) – (5) (Do et al., 2012).

$$M_i = \frac{N-k+1}{2} = \frac{(\frac{1}{25}\sum_{j=1}^{n}\frac{l_j}{W_j})-k+1}{2} \tag{4}$$

$$M_i + 1 = \frac{M_i+1}{2} \tag{5}$$

where K is the total number of junctions and M*i* is the river mixing length's magnitude at the i^{th} hierarchy. After determining segments that should be placed as sampling points with different i^{th} hierarchy, these points are named as “candidate sampling points”. Each candidate sampling points is given a code C1 to Cn, n stands for name of candidate points.

2.3. Contributing area

It is obvious that the land unit areas being far away from river cannot have pollution potential for surface water bodies (DO et al., 2012). Sivertun and Prange (2003) proposed that pollutants produced at the distance more than 1000 meter cannot reach to the river and influence water quality of the rivers. Therefore, to precisely estimate the contributing area affecting water quality a buffer zone (1000 m) is applied. In the present study, flow length of each unit area, which has less than 1000 meter length, is considered because it would remove linear surface ground problem(simple buffer zone) mentioned in Do et al., (2012) study. The buffer zone between the candidate points is divided into watersheds with different pollution sources, affecting water quality changes. The area of each pollution source in each watershed is then calculated using Arc-GIS.

2.4. Multi-criteria evaluation

Multi-criteria evaluation is an efficient approach for considering all factors (pollution sources) and prioritizing their effect on WQMN designs (Chang and Lin, 2014a). Therefore, in this section, we first determined pollution sources (criteria and sub-criteria), and then potential pollution weights were calculated for criteria by ANP approach.

2.4.1. Selection of criteria and sub-criteria

Based on literature reviewed and expert opinions (Vieux and Farajalla, 1994; Chapman, 1996; McCartan et al., 1998; Strobl et al., 2006 a; Chang and Lin, 2014b), non-point sources, lithology, and topography were selected as factors, indicating the chemical and physical characteristics of water quality for the rivers under study. One of the most important contributors to the degradation of water quality is non-point source pollution (Chang and Lin, 2014a). In present study, unlike the previous studies, six non-point sources were used as criteria such as residential, agriculture, rangeland, forest/wooded, water bodies, and highway/road (Fig. 3). Furthermore, event mean concentrations (EMC) of each non-point sources, which represents the concentration of a specific pollutant contained in runoff coming from a particular non-point source within a watershed, including total phosphorus (TP), total nitrogen (TN), total suspended solids (TSS), biochemical oxygen demand (BOD), and nitrate nitrogen (NO3-N), were used as sub-criteria (Table 1).

Among critical factors affecting river water quality in the absence of anthropogenic activities, watershed geology (surficial rocks) plays a vital role in the chemistry of water bodies (McCartan et al., 1998; Zektser et al, 2007; Olson, 2012). According to the nature of surficial rocks/criteria (sedimentary, metamorphic, and igneous rocks), under natural conditions (chemical weathering), dissolved elements from a given lithological unit would enter into and effect water quality of river systems (Meybeck, 1987).Therefore, dissolved elements of different surficial rocks under the study area are divided into three main water quality variables/sub-criteria including: trace elements (e.g. heavy metals), major ions (e.g. salinity and alkaline), and nutrients (Meybeck, 1987; McCartan et al., 1998; Chapman, 1996; Zektser et al., 2007). In addition, relative erosion rate, reflecting dissolved elements value derived from chemical weathering of each rock-type, is considered as a sub-criterion in order to more precisely compute the pollution weight (Meybeck, 1987) (Table 2 and Fig. 4).

As it mentioned in the introduction section, apart from the watershed geology and land-use, the position of each land unit plays a main role in transporting pollutants and their pollution potential (Strobl et al., 2006a). Hydrological process of pollutant transporting is similar to the sediment transport (Sivertun and Prange, 2003). Since the factors influencing sediment transport can affect pollutant transport, the most common topography indices including sediment transport index (STI), stream power index (SPI), and topographic wetness index (TWI), which are used to calculate soil loss, can be employed to identify pollution weight (Fig. 5). These indices can be easily computed by Arc-GIS (Lanni et al., 2012).The effect of topography on soil loss has been particularly determined by sediment transport index (STI) (Moore and Burch, 1986). It reflects the capacity of overland flow in transporting sediment (Pourghasemi et al., 2012) and this index shows the total phosphorus (TP) transporting mechanism (Strobl et al., 2006 a). STI can be calculated with the following relation (Moore and Burch, 1986):

$$STI = \left(\frac{A_s}{22.13}\right)^{0.06} * \left(\frac{\sin\beta}{0.0896}\right)^{1.3} \quad (6)$$

where A_s is the area of a given watershed (m^2), β is the slope (in degree), STI is sediment transport index (dimensionless) (Strobl et al., 2006 a).

TWI as a well-accepted indicator reflecting soil moisture distribution at different position for surface runoff generation (Beven and Kirkby, 1979; Pourghasemi et al., 2012; Conoscenti et al., 2014) is used in this study. TWI is defined as (Beven and Kirkby, 1979):

$$\mathrm{TWI} = \ln\left(\frac{A_s}{tan\,\beta}\right) \tag{7}$$

where TWI is topographic wetness index; A_s and β were introduced in Eq. (6). High TWI indicates that a given cell can generate more runoff than the other cells having low TWI (Beven and Kirkby, 1979). Therefore, generated runoff can carry more particles from soil and affect the water quality (Dube et al., 2014). The other index is stream power index (SPI); it indicates the erosive power of overland flow (Moore et al., 1993).

$$SPI = A_s * \tan\beta \tag{8}$$

where SPI is stream power index (unit less) (Strobl et al., 2006a). High value of SPI reflects the area being more prone to runoff erosive power (Moore et al., 1993). All in all, 12 criteria and 10 sub-criteria are selected to identify pollution potential in present study.

2.4.2. Identification of pollution potential

After selection of criteria, the ANP method was implemented with SuperDecisions software as a multi-criteria evaluation to determine potential pollution weights for non-point sources, different kind of surficial rocks, and topography. Potential pollution weights show relative effect of each criterion on water quality. Among multi-criteria decision-making (MCDM) approaches (e.g., AHP, DEA, and TOPSIS), the ANP method is the most appropriate method (Saaty and Vargas's, 2006; Kucukaltan et al., 2016), as it takes into account the criteria's dependencies and the calculation of their relative weights (Lin et al., 2009). For this purpose, the interdependency (correlation) of water quality variables (sub-criteria) was firstly determined based on the experts' opinions and literature review (Table 3). Then, questionnaires based on Fig. 3, Tables 1(criteria) and 3(sub-criteria) were designed and gave out to 20 experts (hydrologists and geologists) in order to do pair-wise comparison and calculate the relative weights of each criteria by ANP method.

2.5. Scoring sampling points

In present study, to prioritize and select sampling points, the weighted method, which has been used for solving the multiple criteria evaluation problem (Chang and Lin, 2014b), is selected. Therefore, new total potential pollution scores (TPPS) was introduced to prioritize candidate sampling points (Eq. 9). The high value of TPPI indicates high priority for candidate sampling points.

$$TPPS = (W_i * NPP) + (W_i * GPP) + (W_i * TPP) \quad (9)$$

$$NPP = \sum_{i=1}^{6} W_i * A_i, \qquad then \gg Normalized\ between\ 0 - 1 \quad (10)$$

$$GPP = \sum_{i=1}^{3} W_i * A_i, \qquad then \gg Normalized\ between\ 0 - 1 \quad (11)$$

$$TPP = (W_i * TWI_n) + (W_i * SPI_n) + (W_i * STI_n) \quad (12)$$

where W_i is the potential pollution weight of each criterion achieved by the ANP; A_i is the percentage of each non-point sources/surficial rocks in the buffer zone between candidate sampling points; TWI_n, SPI_n, and STI_n are normalized value of topographic indices; and NPP, GPP, and TPP are non-point sources pollution potential, geological pollution potential, topographic pollution potential, respectively.

2.6. Fuzzy logic theory and ranking sampling points

In order to rank the priorities of candidate sampling points, the fuzzy logic theory is applied. It can help to differentiate real differences between the estimated TPPS for candidate points. The natural break approach and fuzzy logic theory proposed by Chang and Lin, (2014b) were employed to classify the candidate sampling points into four grades and data classification. This section contains three following steps. First, the values of each candidate point estimated by Eq. (9) are normalized and fixed between 0 and 1. The normalized value of TPPS is symbolized as $TPPS_n$.

Second, this study applied three fuzzy membership functions proposed by Chang and Lin, (2014b), as indicated in Fig. 7. According to Fig. 7, it shows that each of candidate points has the values of low (l ($TPPS_n$)), medium (m ($TPPS_n$)), and high (h (TPPSn)). They are calculated by Eqs. (15), (16), and (17). The total fuzzy score for 15 candidate points, Fj (j =1~15), is calculated by using Eq. (16) (Chang and Lin, 2014b).

$$l(TPPS_n) = \begin{cases} -2\,TPPS_n + 1 & TPPS_n < 0.5 \\ 0 & TPPS_n > 0.5 \end{cases} \tag{13}$$

$$m(TPPS_n) = \begin{cases} 2\,TPPS_n & TPPS_n < 0.5 \\ -2\,TPPS_n + 2 & TPPS_n > 0.5 \end{cases} \tag{14}$$

$$h(TPPS_n) = \begin{cases} 0 & TPPS_n < 0.5 \\ 2\,TPPS_n - 1 & TPPS_n > 0.5 \end{cases} \tag{15}$$

$$F_j = 0 * l(TPPS_n) + 5 * m(TPPS_n) + 10 * h(TPPS_n) \qquad F_j(j = 1{\sim}15) \tag{16}$$

The candidate points can be classified into four grades based on the F value. First grade is classified between 7.5 and 10, as it shows the most appropriate sampling point. The sampling points are classified as second grade with the value of larger than 5 to less than 7.5. The third and the fourth grades are classified from larger than 2.5 to less than 5, and less than 2.5, respectively. Finally, in order to identify the most appropriate sampling points, low value of both hierarchy(Sanders et al., 1983) and the fuzzy rank for each candidate point (Chang and Lin, 2014b), and considering high anthropogenic activities through land-use maps (Do et al., 2012; Varekar et al., 2015a) are combined. Graphical diagram shows an outline of the full study (Fig. 8).

3. Results and discussion

3.1. Location of potential sampling points

Based on existing stations as well as considering budget deficiency in the study area (Fig. 1), the number of candidate sampling points is 15 (Eq. 3). The main rivers with differences in width were divided into different reaches. Guotor Chai was divided into three different sections in the upstream, middle, and downstream with the average river widths of 33.5m, 74.6m, and 28.1m, respectively. The average river widths for Gudox Bogan and Gazan Chai were 26.4m and 19.0m, respectively. Therefore, the total number of 360 potential sampling points and their locations determined based on Eqs. (1) − (2) (Fig. 10a). Eqs. (4)– (5) were applied to discern the location of 15 candidate sampling points at different ith hierarchy and Mi (Fig. 10b and Table 4). Also, the catchments between candidate points, which are identified by flow length, are shown in Fig.10b.

The results are similar to the findings of the research conducted by Sanders et al. (1983) and Do et al., (2012) that sampling points are located in the downstream and upstream sections of the watershed. Sampling point locations are generally subdivided into microlocations for critical points and macrolacations for routine monitoring (strobl and Robillard, 2008). Macroloctions are systematically designed; moreover, microlocations are functions of macroloctions. Since 15 sampling points are evenly distributed in both the downstream and upstream of the study watershed and are systematically designed, they will partially help critical point monitoring (emergency monitoring). In addition, if there be a sudden water pollution reached to the rivers, the 360 potential sampling points can be used to recognize pollution sources (including amount of pollutant and source location), since they are designed based on river mixing length theory. The result showed that a buffer zone between candidate points should be achieved by considering the flow length of each unit area because areas calculated by flow length were reduced 17%and 27% in C1 and C4, respectively, in contrast to linear surface ground buffer zone. In the light of results, findings corroborated that river mixing length procedure was the best method to the study area which suffer from the lack of reliable data collection on hydraulic and flow characteristics (Day, 1977, Do et al., 2012).

3.2. The results obtained from multi criteria evaluation

3.2.1. Relative weights computed by ANP for criteria

Relative potential pollution weights for non-point sources, calculated by ANP method, are shown in Fig. 6a. According to the Fig. 9a, the residential hits a peak of 0.27 pollution weights which is more than triple the pollution weights of water bodies (0.07). The weights of Highway/road, agriculture, rangeland, and forest/wooded are 0.24, 0.18, 0.13, 0.120, respectively. Furthermore, the results also show that TSS and BOD have the highest relative weights among sub-criteria for non-point sources, with the totals of 0.26 and 0.25, respectively (Fig. 9a). High weights of TSS and BOD are the result of inter-relationship between sub-criteria. Based on the literature review, suspended sediment (e.g. TSS) plays a main role in transporting pollutants (Chapman, 1996) and BOD has a good correlation with the other water quality variables (Ouyang, 2005) (Fig. 6 and

Table 3). The pattern of high relative pollution weights for residential areas and agriculture was demonstrated in Do et al., (2012); however, in the mentioned study its pollution weight was achieved by AHP method. In addition, highway/road, NO3-N, and TKN were not considered as well as the inter-relationship between sub-criteria calculated by the ANP approach.

For surficial rocks (Fig. 9b), it is sedimentary rocks which stands out with far more relative pollution weight than the other two rocks (0.43). In contrast, igneous rocks represent the second relative weight (0. 49), which is followed by metamorphic rocks with the weight of 0.0.36. In addition, the results for sub-criteria demonstrated that relative erosion rate has a major effect on the total potential weight (0.41) (Fig. 9b). The relative weights for major ions, nutrients, and trace elements are 0.23, 0.16, and 0.18, respectively. The high relative weight of sedimentary rocks is due to the high amount of relative erosion rate and major ions in sedimentary rocks (Table 2). There is no study conducted to determine the pollution weights for surficial rocks affecting surface water quality as well as to involve their weights in WQMN design. However, the ANP method helped to propose a new pollution weight for surficial rocks in the present study. According to the results for topography's criteria (Fig. 9c), the relative weight of TWI, which is 0.41, is the largest relative weight among the other criteria. In contrast, SPI and STI have the relative weights of 0.36 and 0.23, respectively. All in all, non-point sources, geology, and topography have its own relative weight in determining sampling points (Eq. 9). Relative weights by TPPS's criteria are 0.41, 0.34, and 0.26 for NPP, GPP, and TPP, respectively (Fig. 9d).

3.2.2. Potential pollution score

Although the study area is dominated by rangeland, a high percentage of anthropogenic activities are seen in the downstream area (Do et al., 2012). C4, C13, and C14 accounted for 7.28, 14.46, and 12.49 percent of residential area in these catchments. Moreover, a majority of agricultural activities are in lower catchments, especially, in C4 to C8 and C13 to C15. Furthermore, in the catchments of C2, C3, and C10 more forest/wooded are located with the percentages of 2.27, 5.71, and 2.18, respectively. Unlike the previous studies conducted (please refer to Strobl et al., 2008 a; Do et al., 2011; Do et al., 2012; Chang and Lin, 2014b), highway/road is taken into account as individual land-use in the present study, as it can be shown that highways/roads are mainly close to the studied rivers. Thus, it is essential to consider highway/road as individual land-use. After calculating NPP by Eq. (10), the normalized scores of NPP for individual candidate sampling points are given in Table 4. According to the table, catchments of C13 and C14 represent the highest number of NPPn (about 0.70), while C3 and C10 accounted the smallest number of NPPn (about 0.38) among candidate points. Variations between the scores for the sampling points could be explained by the more human activities in the lower catchments than the upper catchments.

According to the results, 67 percent of the study area is occupied by the outcrop of sedimentary rocks, which is more than twice of igneous rocks (26%), while metamorphic rocks is accounted about 7 percent of the surficial rocks. The same as the results for non-point sources, the percentage of the surficial rocks in buffer zone (catchments) was significantly different from the whole study area. The results show that 100 percent of C4, C6 to C8, and C13 to C15 is sedimentary rocks. There are more igneous rocks in C9 and C10 with the total of 51.40 and 87.47 percent, respectively. In contrast, only C2 recorded highest amount of metamorphic rocks, being about 33 percent of this catchment. According to the analysis given in Table 4, the highest normalized scores of GPPn are addressed for those candidate points which have high percentage of the sedimentary rocks. Furthermore, C10, C9, and C2 accounted smallest normalized number of GPPn (0.224, 0.544, and 0.586, respectively) due to the presence of different surficial rocks (Table 4).

Based on the analysis carried out in Table 4, the catchments of C9 and C10 have the lowest normalized values of TPPn, 0 and 0.11, respectively. In contrast, the highest normalized value of TPPn is related to C7. These differences between normalized values are due to the variation in the slope gradient in the upstream and downstream of the study area (Fig. 1). In recent years, topography indices have been applied to determine the main role of topography in natural events such as flood and erosion (Dube et al., 2014; Conoscenti et al., 2014). Nevertheless, there is little literature to identify the major role of topography in selecting the sampling points (Strobl et al., 2006a). By representing the TPP method in this study, a novel method is put forward to precisely determine right locations of sampling points.

The values of TPPS and their normalized values, based on Eq. (9), for individual candidate points are given in Table 4. According to the obtained results, candidate points of C6, C7, and C8 stand out, accounting for 0.83, 0.82, and 0.80 of TPPS value, respectively. On the contrary, the lowest values of TPPS are recorded for C10 (0.23), C9 (0.38), and C11 (0.43). Moreover, these numbers indicate that pollution sources between C10, C9, and C11 pose lower risk than the other candidate points for river water quality in the study area.

3.3. Selection of appropriate sampling points for water quality monitoring

To distinguish real differences between the values of TPPS for individual candidate points, they need to be classified by reliable methods. The classified data, which are calculated by the Fuzzy method and natural break approach, are given in Table 4 for candidate points. Table 4 indicates the spatial variability of the candidate points with different fuzzy ranks attached according to the real need for enhanced water quality sampling points. In order to propose appropriate sampling points for water quality monitoring, the low fuzzy rank, low value of hierarchy (Do et al., 2012), and considering high anthropogenic activities (Sanders et al., 1983; Varekar et al., 2015a) are combined. As a result, the most appropriate locations were determined for water quality sampling. Based on the results, six sample points including C4, C6, C8, C12, C14, and C15 are chosen as the most appropriate locations for WQM (Table 4 and Fig. 10c). The fuzzy rank for these points is 1, and their hierarchy values are 2, 3, 1, 2, 3, and 4, respectively. Furthermore, the priorities of C2, C5, C7, and C13, as shown in Table 4 with two stars, were also proposed as the second most appropriate locations in order for WQM sampling points expansion plan in the future. The other catchments indicate the least appropriate locations (Fig. 10c).

All in all, it is clear that the most appropriate locations have the highest values of TPPS and fuzzy rank; as a result of human activities (Do et al., 2012; Varekar et al., 2015a), and existing sedimentary rocks in the catchments between candidate points. This research also highlights that to properly monitor water quality in the study area, six appropriate points according to the current stations; and four points in the future are needed, providing that the budget limitation in the regional water authority could be solved or there will be an expansion plan (Fig. 10, black and red stars). In addition, with a combination of fuzzy ranks and hierarchy values, the selection and priority of appropriate sampling points for WQM becomes an easy task. Therefore, our findings complement previous results by Do et al. (2012) and Chang and Lin, (2014b), with the results that sampling points are evenly distributed in the upstream and downstream and, especially, the catchments which really need WQM. The proposed points have shown that none of current stations are located in appropriate locations in order for WQM in the study area (Figs. 1 and 10). The proposed sampling points will be able to better track pollution sources because the present study has used the natural processes and human activities to enhance sampling points of WQM (Baird et al., 1996; Park et al., 2006; strobl et al., 2006b). Previous studies such as Sanders et al., 1983;

Karamouz et al., 2009; Telci et al., 2009; Chen et al., 2012; Varekar et al., 2015 a, b are too complicated and too case specific for a watershed manager to implement easily (Behmel et al., 2016). This study can be classified in cost-effective method to determine sampling points, since it uses only available watershed data, technical and expert resources. Proposed framework will be useful for regional water authorities struggling with limited financial resources and looking for a method to determine sampling points location for the first time, in particular, for developing countries like Iran.

4. Conclusion

This study, conducted on the Khoy watershed, describes a novel methodology in order to appropriately locate the existing stations and proposing new sampling points for surface water quality monitoring. 12 criteria (residential area, agriculture, rangeland, forest /wooded, water bodies, highway/road, sedimentary rocks, metamorphic rocks, igneous rocks, TWI, TPI, and STI) and 10 sub-criteria (TSS, TP, TN, TKN, BOD, NO3-N, major ions, trace elements, nutrients, and relative erosion rate) have been selected to determine the suitable locations of sampling points for WQM.

It can be concluded that an integrated application of the multi criteria evaluation methods including ANP can assist the identification of exact relative pollution weights of factors involved in appropriately locating sampling sites. Relative pollution weights for residential, highway/road, agriculture, rangeland, forest/wooded, and water bodies are 0.27, 0.24, 0.18, 0.13, 0.12, and 0.07, respectively. In the light of the results, TSS and BOD are shown as the most important parameters in identifying relative pollution weights for non-point sources. This study also introduces new relative pollution potential weights for surficial rocks so that sedimentary, igneous, and metamorphic rocks' pollution weights are derived as 0.49, 0.28, and 0.23, respectively. In addition, relative erosion rate and major ions are addressed as sub-criteria having more effect on pollution weights for outcropped rocks. Furthermore, weights for topography indices of TWI, SPI, and STI are 0.41, 0.36, and 0.23, respectively.

Pollution potential scores for non-point sources, surficial rocks, and topography are combined by the weighted procedure to introduce new total potential pollution index, when extensive watershed information is available but there is the lack of water quality data. This index is classified and ranked by the fuzzy theory for each candidate point. A combination of mixing length method, fuzzy rank and hierarchy value assists us in prioritizing and proposing new locations of the sampling sites in the whole river system. In summary, six points as the most appropriate (current situation) and four points as the second most appropriate sampling sites (in the future) are proposed in order to relocate the current stations and enhancing WQMNs in the study area. The present study provides a novel prescription and practical recommendation for water quality monitoring agencies, which have been suffered from reliable water quality data and cost-effective method for selecting the exact location of sampling sites. The proposed methodology has a huge potential to be applied in other countries around the world, especially in developing countries with limited financial resources. The method also does not require water quality data as an input, so could be applied in settings where those data are scarce.

References

Aboutalebi, M., Bozorg-Haddad, O., Loáiciga, H.A., 2016. Multiobjective Design of Water-Quality Monitoring Networks in River-Reservoir Systems. Journal of Environmental Engineering: 04016070.

Ahn, B.S., Park, K.S., 2008. Comparing methods for multiattribute decision making with ordinal weights. Computers & Operations Research, 35(5): 1660-1670.

Alilou, H., Nia, A.M., Keshtkar, H., Han, D. and Bray, M., 2018. A cost-effective and efficient framework to determine water quality monitoring network locations. *Science of the Total Environment*, *624*, pp.283-293.

Alvarez-Vázquez, L.J., Martínez, A., Vázquez-Méndez, M., Vilar, M., 2006. Optimal location of sampling points for river pollution control. Mathematics and Computers in Simulation, 71(2): 149-160.

Aragonés-Beltrán, P., García-Melón, M., Montesinos-Valera, J., 2017. How to assess stakeholders' influence in project management? A proposal based on the Analytic Network Process. International Journal of Project Management, 35(3): 451-462.

Baird, C., Jennings, M., Ockerman, D., Dybala, T., 1996. Charaterization of Nonpoint Sources and Loadings to the Corpus Christi Bay National Estuary Program Study Area.

Baldys, S., Raines, T., Mansfield, B., Sandlin, J., 1998. Urban stormwater quality, event-mean concentrations, and estimates of stormwater pollutant loads, Dallas-Fort Worth area, Texas, 1992--1993, Geological Survey, Water Resources Div., Austin, TX (United States); North Central Texas Council of Governments, Arlington, TX (United States).

Baltacı, F., Onur, A.K., Tahmiscioğlu, S., 2008. Water quality monitoring studies of Turkey with present and probable future constraints and opportunities. Desalination, 226(1-3): 321-327.

Behmel, S., Damour, M., Ludwig, R., Rodriguez, M., 2016. Water quality monitoring strategies—A review and future perspectives. Science of the Total Environment, 571: 1312-1329.

Beven, K., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrological Sciences Journal, 24(1): 43-69.

Beveridge, D. et al., 2012. A geostatistical approach to optimize water quality monitoring networks in large lakes: application to Lake Winnipeg. Journal of Great Lakes Research, 38: 174-182.

Böhlke, J., Denver, J., 1995. Combined use of groundwater dating, chemical, and isotopic analyses to resolve the history and fate of nitrate contamination in two agricultural watersheds, Atlantic coastal plain, Maryland. Water Resources Research, 31(9): 2319-2339.

Burrough, P.A., McDonnell, R., McDonnell, R.A., Lloyd, C.D., 2015. Principles of geographical information systems. Oxford University Press.

Chang, C.-L., 2013. Evaluation of basin environmental vulnerability: the weighted method compared to the compromise method. International Journal of Environmental Science and Technology, 10(5): 1051-1056.

Chang, C.-L., Lin, Y.-T., 2014. A water quality monitoring network design using fuzzy theory and multiple criteria analysis. Environmental monitoring and assessment, 186(10): 6459-6469.

Chapman, D.V., 1996. Water quality assessments: a guide to the use of biota, sediments, and water in environmental monitoring.

Chen, L., Tan, C.-H., Kao, S.-J., Wang, T.-S., 2008. Improvement of remote monitoring on water quality in a subtropical reservoir by incorporating grammatical evolution with parallel genetic algorithms into satellite imagery. Water Research, 42(1): 296-306.

Chen, Q., Wu, W., Blanckaert, K., Ma, J., Huang, G., 2012. Optimization of water quality monitoring network in a large river by combining measurements, a numerical model and matter-element analyses. Journal of environmental management, 110: 116-124.

Chilundo, M., Kelderman, P., 2008. Design of a water quality monitoring network for the Limpopo River Basin in Mozambique. Physics and Chemistry of the Earth, Parts A/B/C, 33(8): 655-665.

Conoscenti, C. et al., 2014. Gully erosion susceptibility assessment by means of GIS-based logistic regression: a case of Sicily (Italy). Geomorphology, 204: 399-411.

D'Arcy, P., Carignan, R., 1997. Influence of catchment topography on water chemistry in southeastern Quebec Shield lakes. Canadian Journal of Fisheries and Aquatic Sciences, 54(10): 2215-2227.

Day, T.J., 1977. Observed mixing lengths in mountain streams. Journal of Hydrology, 35(1-2): 125-136.

Do, H.T., Lo, S.-L., Chiueh, P.-T., Thi, L.A.P., 2012. Design of sampling locations for mountainous river monitoring. Environmental Modelling & Software, 27: 62-70.

Do, H.T., Lo, S.-L., Chiueh, P.-T., Thi, L.A.P., Shang, W.-T., 2011. Optimal design of river nutrient monitoring points based on an export coefficient model. Journal of Hydrology, 406(1): 129-135.

Dube, F. et al., 2014. Potential of weight of evidence modelling for gully erosion hazard assessment in Mbire District– Zimbabwe. Physics and Chemistry of the Earth, Parts A/B/C, 67: 145-152.

Harper, H.H., 1998. Stormwater chemistry and water quality. Environmental Research & Design, Inc.: Orlando, FL, USA.

Icaga, Y., 2005. Genetic algorithm usage in water quality monitoring networks optimization in Gediz (Turkey) river basin. Environmental Monitoring and Assessment, 108(1): 261-277.

Karamouz, M., Kerachian, R., Akhbari, M., Hafez, B., 2009. Design of river water quality monitoring networks: a case study. Environmental Modeling & Assessment, 14(6): 705.

Keum, J., Kaluarachchi, J.J., 2015. Development of a decision-making methodology to design a water quality monitoring network. Environmental monitoring and assessment, 187(7): 466.

Khalil, B., Ouarda, T., St-Hilaire, A., 2011. A statistical approach for the assessment and redesign of the Nile Delta drainage system water-quality-monitoring locations. Journal of Environmental Monitoring, 13(8): 2190-2205.

Kovacs, Z. et al., 2016. Water spectral pattern as holistic marker for water quality monitoring. Talanta, 147: 598-608.

Kucukaltan, B., Irani, Z., Aktas, E., 2016. A decision support model for identification and prioritization of key performance indicators in the logistics industry. Computers in Human Behavior, 65: 346-358.

Lanni, C., Borga, M., Rigon, R., Tarolli, P., 2012. Modelling shallow landslide susceptibility by means of a subsurface flow path connectivity index and estimates of soil depth spatial distribution. Hydrology and Earth System Sciences, 16(11): 3959.

Lelikov, E.P., 1988. Metamorphism in the formation of the plicated foundation of marginal seas around the Pacific Ocean. Journal of Physics of the Earth, 36(Proceeding2): S169-S178.

Lettenmaier, D.P., 1979. Dimensionality problems in water quality network design. Water Resources Research, 15(6), pp.1692-1700.

Liebetrau, A.M., 1979. Water quality sampling: some statistical considerations. Water Resources Research, 15(6), pp.1717-1725.

Lin, Y.-H., Tsai, K.-M., Shiang, W.-J., Kuo, T.-C., Tsai, C.-H., 2009. Research on using ANP to establish a performance assessment model for business intelligence systems. Expert Systems with Applications, 36(2): 4135-4146.

Line, D.E., White, N.M., Osmond, D.L., Jennings, G.D., Mojonnier, C.B., 2002. Pollutant export from various land uses in the Upper Neuse River Basin. Water Environment Research, 74(1): 100-108.

Liyanage, C.P., MARASINGHE, A., YAMADA, K., 2016. Comparison of Optimized Selection Methods of Sampling Sites Network for Water Quality Monitoring in a River. International Journal of Affective Engineering, 15(2): 195-204.

Mahjouri, N., Kerachian, R., 2011. Revising river water quality monitoring networks using discrete entropy theory: the Jajrood River experience. Environmental monitoring and assessment, 175(1): 291-302.

McCartan, L., Peper, J.D., Bachman, L.J., Horton, J.W., 1998. Application of geologic map information to water quality issues in the southern part of the Chesapeake Bay watershed, Maryland and Virginia, eastern United States. Journal of Geochemical Exploration, 64(1): 355-376.

Memarzadeh, M., Mahjouri, N., Kerachian, R., 2013. Evaluating sampling locations in river water quality monitoring networks: application of dynamic factor analysis and discrete entropy theory. Environmental earth sciences, 70(6): 2577-2585.

Meybeck, M., 1987. Global chemical weathering of surficial rocks estimated from river dissolved loads. American journal of science, 287(5): 401-428.

Moore, I., Burch, G., 1986. Modeling erosion and deposition: topographic effects. Transactions of the ASAE, 29(6): 1624-1630.

Moore, I., Turner, A.K., Wilson, J.P., Jenson, S.K., Band, L.E., 1993. GIS and land-surface-subsurface process modeling. Environmental modeling with GIS, 20: 196-230.

Mustapha, A., Aris, A.Z., Juahir, H., Ramli, M.F., Kura, N.U., 2013. River water quality assessment using environmentric techniques: case study of Jakara River Basin. Environmental Science and Pollution Research, 20(8): 5630-5644.

Newell, C.J., Rifai, H.S., Bedient, P.B., 1992. Characterization of nonpoint sources and loadings to Galveston Bay.

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(1): 121-126.

Noori, R., Sabahi, M.S., Karbassi, A., Baghvand, A., Zadeh, H.T., 2010. Multivariate statistical analysis of surface water quality based on correlations and variations in the data set. Desalination, 260(1): 129-136.

Olson, J.R., 2012. The influence of geology and other environmental factors on stream water chemistry and benthic invertebrate assemblages.

Ouyang, Y., 2005. Evaluation of river water quality monitoring stations by principal component analysis. Water research, 39(12): 2621-2635.

Park, S.-Y., Choi, J.H., Wang, S., Park, S.S., 2006. Design of a water quality monitoring network in a large river system using the genetic algorithm. Ecological modelling, 199(3): 289-297.

Pourghasemi, H.R., Pradhan, B., Gokceoglu, C., 2012. Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. Natural hazards, 63(2): 965-996.

Pratt, B., Chang, H., 2012. Effects of land cover, topography, and built structure on seasonal water quality at multiple spatial scales. Journal of hazardous materials, 209: 48-58.

Saaty, T.L., Takizawa, M., 1986. Dependence and independence: From linear hierarchies to nonlinear networks. European Journal of Operational Research, 26(2): 229-237.

Saaty, T.L., Vargas, L.G., 2006. Decision making with the analytic network process. Springer.

Sanders, T.G., 1980. Principles of network design for water quality monitoring. Colorado State University.

Sanders, T.G., 1983. Design of networks for monitoring water quality. Water Resources Publication.

Sivertun, Å., Prange, L., 2003. Non-point source critical area analysis in the Gisselö watershed using GIS. Environmental Modelling & Software, 18(10): 887-898.

Strobl, R., Robillard, P., Day, R., Shannon, R., McDonnell, A., 2006. A water quality monitoring network design methodology for the selection of critical sampling points: Part II. Environmental Monitoring and Assessment, 122(1): 319-334.

Strobl, R., Robillard, P., Shannon, R., Day, R., McDonnell, A., 2006. A water quality monitoring network design methodology for the selection of critical sampling points: Part I. Environmental Monitoring and Assessment, 112(1): 137-158.

Telci, I.T., Nam, K., Guan, J., Aral, M.M., 2009. Optimal water quality monitoring network design for river systems. Journal of environmental management, 90(10): 2987-2998.

Tourtelot, H.A., 1971. Chemical compositions of rock types as factors in our environment. Geological Society of America Memoirs, 123: 13-18.

Varekar, V., Karmakar, S., Jha, R., 2016. Seasonal rationalization of river water quality sampling locations: a comparative study of the modified Sanders and multivariate statistical approaches. Environmental Science and Pollution Research, 23(3): 2308-2328.

Varekar, V., Karmakar, S., Jha, R., Ghosh, N., 2015. Design of sampling locations for river water quality monitoring considering seasonal variation of point and diffuse pollution loads. Environmental monitoring and assessment, 187(6): 376.

Vieux, B.E., Farajalla, N.S., 1994. Capturing the essential spatial variability in distributed hydrological modelling: Hydraulic roughness. Hydrological Processes, 8(3): 221-236.

Volkova, N., Budanov, V., 1999. Geochemical discrimination of metabasalt rocks of the Fan–Karategin transitional blueschist/greenschist belt, South Tianshan, Tajikistan: seamount volcanism and accretionary tectonics. Lithos, 47(3): 201-216.

Wang, Y.-B., Liu, C.-W., Liao, P.-Y., Lee, J.-J., 2014. Spatial pattern assessment of river water quality: implications of reducing the number of monitoring stations and chemical parameters. Environmental monitoring and assessment, 186(3): 1781-1792.

Xiaomin, Z., Jianjun, Y., Xiaoci, H., Shaoli, C., 2016. An Ontology-based Knowledge Modelling Approach for River Water Quality Monitoring and Assessment. Procedia Computer Science, 96: 335-344.

Ye, L., Cai, Q.-h., Liu, R.-q., Cao, M., 2009. The influence of topography and land use on water quality of Xiangxi River in Three Gorges Reservoir region. Environmental Geology, 58(5): 937-942.

Young, R.G., Quarterman, A.J., Eyles, R.F., Smith, R.A., Bowden, W.B., 2005. Water quality and thermal regime of the Motueka River: influences of land cover, geology and position in the catchment. New Zealand Journal of Marine and Freshwater Research, 39(4): 803-825.

Zektser, I.S., Marker, B., Ridgway, J., Rogachevskaya, L., Vartanyan, G., 2007. Geology and ecosystems. Springer Science & Business Media.