IOP Conference Series: Earth and Environmental Science

PAPER • **OPEN ACCESS**

Spatial assessment of pollutant loads for surface water quality management: a case study in Lai Chau city, Vietnam

To cite this article: T A Ngo *et al* 2024 *IOP Conf. Ser.: Earth Environ. Sci.* **1345** 012012

Spatial assessment of pollutant loads for surface water quality management: a case study in Lai Chau city, Vietnam

T A Ngo[1*], L See[2], S J Leisz[3,4], N B Tran[1], T H V Dinh[1] and T H Nguyen[5]

[1] Faculty of Natural Resources and Environment, Vietnam National University of Agriculture, Hanoi, VIETNAM; ntan@vnua.edu.vn, tnbang@vnua.edu.vn, dthvan@vnua.edu.vn
[2] International Institute for Applied Systems Analysis (IIASA), Laxenburg, AUSTRIA; see@iiasa.ac.at
[3] College of Arts and Sciences, VinUniversity, Ocean Park, Hanoi, Vietnam
[4] Department of Anthropology and Geography, Colorado State University, Fort Collins, CO 80523, US; steve.leisz@colostate.edu
[5] Lai Chau Department of Natural Resource and Environment, Lai Chau city, VIETNAM; ynhihanganh@gmail.com
[*]Correspondence: ntan@vnua.edu.vn

Abstract. The aim of this study is to present a method for estimating the pollutant load from different sources in an effort to provide improved information regarding water pollution and help control the surface water pollution, using Lai Chau city as a case study. The pollutant load was calculated in accordance with the Vietnam Environment Administration Decree No.154/2019 on the guidance for calculating the total pollutant load of river water. The pollutant sources include point sources (domestic wastewater, animal husbandry, industrial complexes and economic services) and surface sources (run-off from agricultural land uses) that generate wastes that potentially contaminate water bodies. The source locations were mapped and spatially joined with the drainage-basin map delineated from a Digital Elevation Model (DEM) to calculate the loads for the sub-basin units. Multivariate analysis then showed that the farming and domestic sources had the strongest positive loading factors for the sub-basins located in the city center and its fringe areas. Of these waste from animal husbandry account for up to 75.1% of total pollutant load. The main conclusion from the study's results is that the management approach should be changed from the total controlling mode, which is currently applied in the city, to a source specific approach based on the pollutant discharge loads and the allocated capacities.

Keywords: water quality, pollution sources, pollutant load mapping

1. Introduction

Assessing the water quality in the river basins by determining the main sources of pollution is part of a larger integrated water management process identified by the United Nation's Sustainable Development Goals ("SDG", i.e. Target 6.3) [1]. Although the sources of water pollution vary from country to country, they can generally be characterized into (i) point sources, e.g., those that arise from easily identifiable points such as factories or a sewage treatment plants, and (ii) non-point pollution sources (NPS), which are the combined set of pollutants that originate from a larger area and then end up in streams and rivers due to runoff [2]. Due to the fact that the distribution and composition of pollutant sources in watersheds have a great influence on the receiving water body's quality, "the source control" is an environmental management approach that has been widely applied [3, 4, 5]. Particularly in Vietnam, the source control of pollution from point and non-point sources has been stated clearly in the Environmental Protection law (2020) as a compulsory approach to surface water quality management planning (i.e., Article 9). However, the realities of how wastewater is treated in Vietnam complicate how this classification of wastewater source can be applied.

Worldwide, there are different methods for assessing water quality based on pollutant sources that range from detailed simulations using hydrodynamic modelling with tools such as SWAT (Soil and Water Assessment Tool), BATHTUB or CE-QUAL-W2, among several available tools [6, 7], to GIS-based approaches [5, 8, 9], which can be used depending on the purpose of the exercise. Providing a framework for the selection of water quality models, Chinyama et al. [7] clearly state that the model or approach to be adopted for management purposes should be determined by both the system that is to be modelled (and hence the data availability) as well as the legislation that is in place in the country where it will be used.

The study described in this paper took place in Vietnam, where legislation related to water pollution was passed in 2019 (Decree No.154/2019), along with recommendations for how to calculate the total pollutant load at the city scale [10]. The enclosed technical contents of Decree 154/2019 consist of the calculations and some coefficients that were adopted from the Guidance for Introducing the Total Pollutant Load Control System [11]. The pollutant sources that potentially contaminate water bodies required for the calculation include the following: domestic, agricultural and animal husbandry, industrial and those pollution sources arising from economic services such as from hotels and hospitals. These main sources of pollutants have been identified and tested in previous research studies [5, 11, 12]. The method of pollutant calculation has been formulated in legal documents. One key point is that unless there are centralized water treatment facilities for the area where calculations of water pollution are being done, the default is to assume that there is no water pollution treatment applied. Thus, the application in local practice within Vietnam of calculating water pollution levels in anything other than a very simplified way (e.g. general calculations for a large geographic area) is very limited.

One way that these calculations could be refined and made more geographically precise for specific areas within urban locations is to use geographic information system (GIS) software spatial analytical capacity to apply the specified method of pollutant calculation detailed in Decree No. 154/2019 to the specific areas within the city that are pollutant point sources. However, a review of literature shows that the application of GIS software (e.g. ArcGIS, BASINS) for such a purpose has been used very limitedly in Vietnam and other SE Asian countries. The main reason for this limitation is that the computational practice requires systematic data collection and knowledge of specialized spatial analysis techniques, which are currently lacking. Although the SWAT model has been used previously in Vietnam [13, 14, 15, 16], the aim of these studies was broader than the water quality assessment undertaken here, i.e., these studies determined the impacts of climate change and/or changes in land use on a broader set of variables including river discharge, sediment load and/or nitrogen/phosphorus loading, historically and in the future. For this broader purpose, a hydrodynamic model was required. Others have made use of variations of a GIS-based approach to monitor water quality. An example of a GIS-based approach is the study by Yan et al. [9], where water samples were collected at 67 sampling sites along the Honghe River in China. The authors then calculated a number of different single factor pollution indices that related the concentrations of dissolved oxygen, ammonia nitrogen, nitrate nitrogen, total nitrogen and total phosphorus based on a water quality classification as well as a combined factor pollution index. They then used spatial interpolation to create maps of pollution loading of these single factors and combined pollutants to calculate the percentage of the watershed that was polluted. However, when water samples are not available, GIS-based approaches focus on identifying waste sources by sub-basin and then calculating the cumulative loads where the discharge load conversion factor is area specific [5, 8]. For example, Lee et al. [8] showed that a GIS-based approach to calculating non-point sources (NPS) of pollution was more accurate than using administrative areas in combination with simple formulas for determining areas of sewage contribution. More recently, Su et al. [5] undertook an analysis of the density and intensity of chemical oxygen demand (COD) and N-total concentrations in Qingdao on the eastern coast of China. Pollution coefficients for the two NPS (agriculture and rural domesticity) were taken from manuals developed specifically for pollution calculations in the region and from a general survey of farming areas in Qingdao. After dividing the study area into 92 pollution source regions (PSR), they calculated the pollution density by dividing the pollutant concentration in each PSR by the area of each zone, and the pollution intensity by dividing the pollutant concentration in each PSR by the GDP of each zone. They then examined the distribution of the pollution intensity by calculating the Gini coefficient of each PSR. This approach allowed them to understand the structure of the pollution sources spatially, which led to recommendations for mitigation.

This study uses a GIS-based approach applied to the study area of Lai Chau city in northern Vietnam to combine information gathered using participatory methods regarding local, not centralized, treatment of wastewater within a spatial database to refine the precision of how pollution calculations are done. The objective of this study is to estimate the pollutant discharge loads from different point and non-point sources over sub-basins using accepted pollution coefficients and taking into account the local treatments of the wastewater that users report applying to the wastewater. Essentially, the study aims to use a GIS-based approach informed by local knowledge to identify pollution sources that significantly contributed to the pollutant loads at the study site; and to examine whether or not the current environmental management approach of the city, which applied the total controlling mode, is suitable for the pollution distribution pattern within the city.

2. Materials and methods

2.1. The study area

This study was conducted in Lai Chau city in northern Vietnam, which consists of five urban wards and two rural communes that bookend the urban center on the western and eastern sides of the city (Figure 1).

Lai Chau city is surrounded by mountain ranges, where the western part of the city has an uneven terrain with a maximum elevation of nearly 1300 m, which runs along the boundary of the city in the west. The remaining areas have less prominent slopes, with the lowest elevation at 895 m. The hydrological network flows from the northwest to the southeast. The surface water includes a lake in the city center and some small stream systems [28].

Characteristics of the main pollution sources are as follows:
Domestic (population) and land uses: In 2020, the total population of Lai Chau city was 43,294 people in 11,366 households, living mainly in the city center. According to provincial statistics [17], the agricultural land covers 293,332 ha, accounting for 47.5% of the total natural area and containing 41% cultivated land, water surfaces for aquaculture that occupy 0.2%, and forest land covering 24.2%.
Industrial: The industrial sector is still underdeveloped, especially in the field of industrial production. According to the provincial statistics [17], there is currently only one stone crushing plant located in the city for converting raw materials into cement production.
Commercial services: Commercial services are still in the early stages of development. The most significant commercial services are hotels. There are 42 hotels, located mainly in the city center.
Health care services (hospitals): Health care and medical services are still in the early stages of development in the city. There are 27 hospitals/clinics located mostly in the city center.
Livestock: Most livestock waste is only partially treated by the means of a biogas digester. The untreated and post-biogas waste is then discharged into various water bodies such as fishponds, lakes and irrigation canals. The number of concentrated farms (i.e., farms located in isolation from residential areas) is 62.

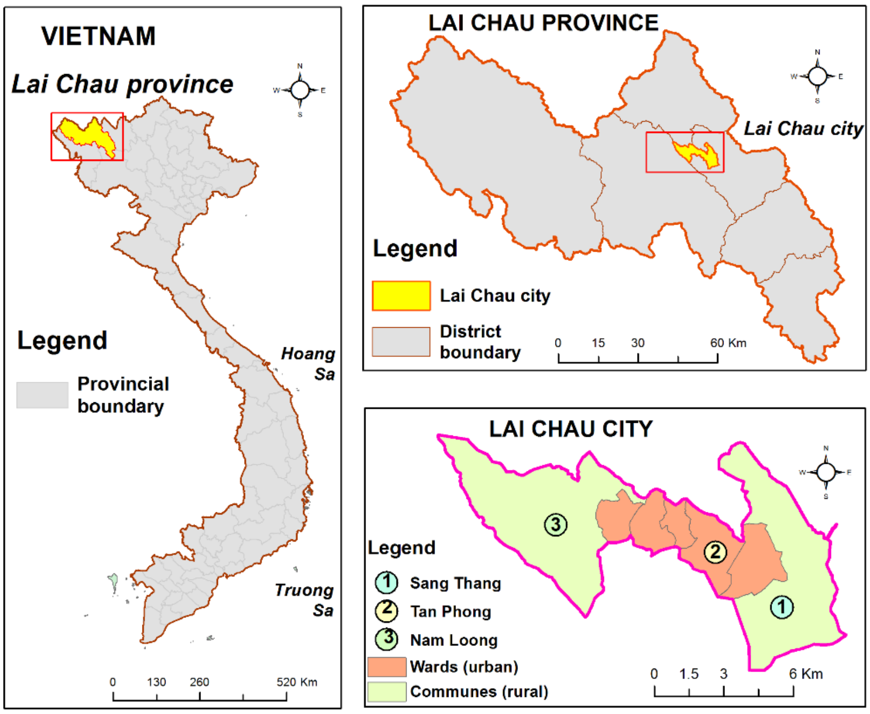

Figure 1. Lai Chau city and three selected administrative units for household interviews.

2.2. Data sources

The statistical data (in 2020) on industrial establishments, services, businesses, hospitals, livestock production (buffalos, cows, pigs and poultry) and population (domestic) were obtained from the Lai Chau Department of Agriculture and Rural Development (DARD) and the Lai Chau Statistical Office [17]. A digital elevation model (DEM) developed from Shuttle Radar Topography Mission (SRTM) and distributed as open access data was used to delineate the sub-basins using ArcGIS v.10.3, the methodology is described below in section 2.3. The administrative and land use map for 2020 was obtained from the Lai Chau Department of Natural Resources and Environment (DONRE).

The pollutant coefficients of the four major pollutant sources were estimated based on the requirements set out in the legislation in Vietnam (Decision No. 154/QD-TCMT) [10] as shown in Tables 1 and 2.
Participatory data: To obtain realistic model parameters, household surveys were conducted in three administrative units located in (1) the southeast (San Thang commune); (2) the center (Tan Phong ward); and (3) the northeast (Nam Loong commune) of the city (Figure 1, bottom right) to collect information about wastewater treatment at both the household and farm level. A total of 120 households were interviewed in 2020. In each commune, 30 households were randomly selected and all 62 animal farms in the city were interviewed. The questionnaire for the interview was designed in a structured form that focused on pollution removal rates related to the wastewater treatment facilities of the households and farms. The pollution removal rates at the industrial, business and hospital establishments, etc., were obtained from the Lai Chau's Department of Natural Resources and Environment (DONRE). These data were used as inputs for calculating the flow of pollutant loads for the whole city as presented in Figure 2.

Table 1. Pollutant coefficients from major point sources.

Discharge sources	COD	BOD	N-total	P-total
1. Domestic (human life) *(kg/person/year)*	27.302	14.454	0.964	0.272
2. Industries *(kg/m^3/year)*	0.150	0.050	0.040	0.006

3. Services, hotels (*kg/m³/year*)	0.220	0.135	0.113	0.030
4. Hospitals *(kg/m³/year)*	0.100	0.050	0.050	0.010
5. Livestock:				
Buffalo *(kg/head/year)*	295	164	43.8	11.3
Cow *(kg/head/year)*	295	164	43.8	11.3
Pig *(kg/head/year)*	59.2	32.9	7.3	2.3
Horse *(kg/head/year)*	263	146	95.3	16.4
Goat *(kg/head/year)*	60.7	33.7	13.5	3.7
Chicken *(kg/head/year)*	2.90	1.61	3.6	-

Source: VEA (2019) [10].

Table 2. Pollutant coefficients from surface run-off of major land use types.

Land use types	COD *(kg/ha/year)*	BOD *(kg/ha/year)*	N-total *(kg/ha/year)*	P-total *(kg/ha/year)*
Agriculture (crops)	16.8	30.2	17.9	1.1
Forest and fruit trees	72.8	131.0	4.4	0.3
Special uses	60.4	108.7	10.6	2.3
Residential (overflow)	56.0	100.8	9.0	2.2
Aquaculture	90.0	162.0	12.6	-

Source: VEA (2019) [10].

2.3. Mapping and GIS (spatial) analysis

Mapping and spatial analysis were performed using ArcGIS 10.3 and Basins 4.5, the US-EPA software developed specifically for terrain analysis [18].

Sub-basin delineation: In Vietnam Decision No.154/QD-TCMT specifies that the inventory of pollutant load should be carried out at the sub-basin level. The data used to delineate the sub-basin map was the DEM (SRTM 1 Arc-Second Global). The DEM was filtered using the Fill-Sinks method [19]. The flow direction and flow accumulation were then calculated for each pixel using the Top-down Deterministic-8 method. The flow network from high to low levels is essential for determining the hierarchy of the basins. In this research, we selected the limit of flow detection within 100 ha (equivalent to the area of a village) to identify the sub-basins using the automatic watershed delineation tool as recommended by previous research [20, 21].

For the point pollution sources, the exact location of each household is not known, only the total number of households by administrative ward and commune from the census. Therefore, the total number of households was randomly allocated across the residential areas based on the 2020 land use map. Attribute information, including the number of people and livestock (buffaloes, cows, pigs, and poultry) was assigned to each household based on the survey data (mean and standard error). The GPS coordinates of farms, industrial production facilities, services, and hospitals were used to map these locations.

Non-point sources were identified from a land use map. Land use types in the map were then assigned pollutant run-off coefficients (Table 2) for calculating the total pollutant load based on the legal requirements in Vietnam according to Decision No. 154/QD-TCMT [10].

Calculating the distribution of pollutant sources: The flow of pollution load from different sources is estimated according to VEA and guidelines from the Office for Environmental Management of Enclosed Coastal Seas [11]. In particular, the load generated from the waste sources was calculated according to the load coefficients (Table 1 and Table 2). The Pollutant discharge load is calculated as:

$$Pollutant\ discharge\ load = Pollutant\ load\ generation * (1 - Removal\ rate) \qquad (1)$$

where the Removal rate is defined as:

$$Removal\ rate = Rate\ of\ treated\ wastewater * Treatment\ efficiency \quad (2)$$

where the *Rate of treated water* and the *Treatment efficiency* were collected by the surveys and secondary data from the Lai Chau's DONRE. In particular, the interviewees were asked to estimate the wastewater discharged volume, the amount of treated water (i.e. wastewater passed through by septic tank, biogas digester). Treatment efficiencies were taken from DONRE assessments, which have been summarized by the administrative units (communes, wards). Finally, the pollutant loads were calculated for each source for each sub-basin, and then aggregated to the whole study area.

The methodology used to map the spatial distribution of pollutant loads is illustrated in Figure 2.

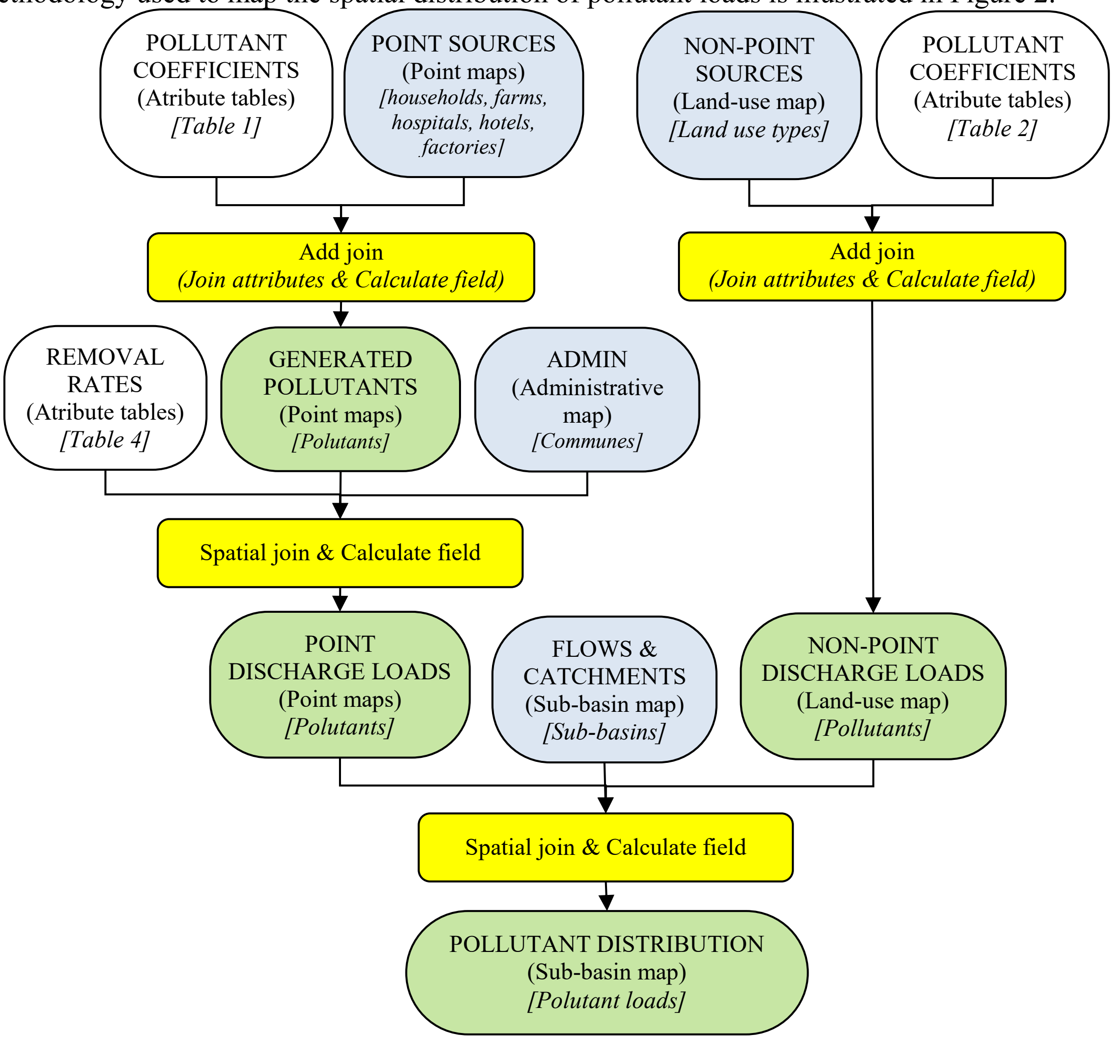

Figure 2. Methodology for calculating pollutant loads by sub-basin.

Multivariate analysis: Multivariate analysis was undertaken using a combination of Hierarchical cluster analysis (HCA), Discriminant analysis (DA) and Principal component/factor analysis (PCA/FA) using SPSS 20. The analyses were applied on normalized data ($\log_{10}$) as required by the statistical analysis (Box & Cox 1964). The HCA and DA were used for determining the spatial discriminants of the pollutant sources based on the discharge loads (COD, BOD, N-total, P-total) distributed by sub-basins. The HCA classifies objects (sub-basins), so that each sub-basin is similar to others in the cluster and different from those in the other clusters. As suggested from previous research [16, 22, 23, 24], Ward's method and the Euclidian distance measurement were used.

The DA determines the variables (i.e., pollutant sources) that best discriminate between groups of sub-basins produced by the HCA. Standard, backward stepwise and forward stepwise modes were employed in the DA to test whether the groups differ with regards to the mean of the variable. To achieve this, the clusters (of sub-basins) were selected as the dependent variable, while the estimated pollutant load parameters were the independent variables.

The PCA was applied to identify the most significant pollutant sources as suggested in previous research [19, 25, 26]. PCA determines the most significant parameters that best capture the variation in the data set by eliminating the least significant parameters with a minimal loss in the original variables [16]. This was achieved by rotating the axis defined by the PCA and constructing new variables known as varimax factors [16]. The varimax factor coefficients with a correlation greater than 0.75 are classified as having a strong significant factor loading, those that range between 0.50 - 0.75 have moderate factor loading, while a value of less than 0.50 is considered as a weak factor loading [27].

3. Results and discussion

3.1. Distribution of pollutant sources

By setting the network delineation threshold at 100 hectares, which is the approximate area of a village, Lai Chau city was divided into 65 sub-basins (Figure 3a). Each sub-basin is considered as a sink that locally accumulates pollutants from discharged sources before discharging them further into the stream and river systems.

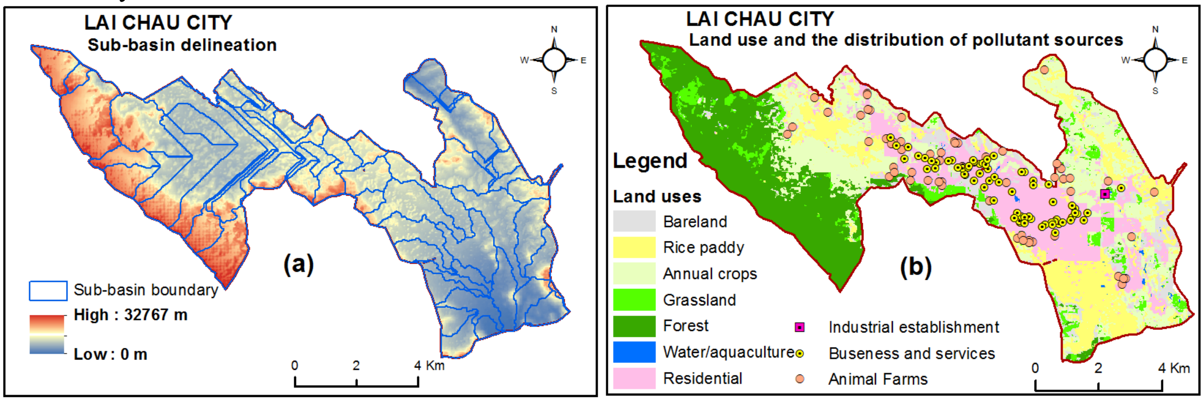

Figure 3. (a) Map of sub-basins delineated from the DEM; and (b**)** Location of pollutant sources determined from the land use map and the statistical data.

Based on the statistical data, the number of households and livestock were assigned to the residential clusters and other land use types from the land use map (2020), accordingly. The industrial establishments, services, businesses, and animal farms were added to the map according to their coordinates (Figure 3b).

Figure 3b shows that most of the pollutant discharging points are located in the city center. In addition, there are a few farms distributed sparsely on agricultural land outside the central area. The total pollutant sources that are presented on the map are summarized in Table 3.

Table 3. Main pollutant sources in Lai Chau city.

No.	Pollutant sources	Total
1	Households (# households):	11366
	Population (person)	*43294*
	Cattle (head)	*41059*
	Poultry (head)	*18217*
2	Animal farms (farm):	62

	Cattle (head)	*29960*
	Poultry (head)	*14524*
3	Industrial production facilities	1
4	Businesses and services (hotels & hospitals)	69
5	Land uses surface run-off (ha):	7020.0
	Bare land	*542.9*
	Rice paddy	*1176.1*
	Crops	*1734.5*
	Grassland	*404.3*
	Forest and woody trees	*1696.5*
	Surface water for aquaculture	*17.2*
	Urban and built-up	*1448.6*

3.2. Wastewater treatment

According to the survey data (2020) and the statistical report from the Lai Chau's DONRE [28], the current rate of treated wastewater is summarized by different sectors as shown in Table 4.

Table 4. Wastewater treatment rate by different sectors.

No.	Administration units (Ward/commune)	Treated wastewater by different sectors (%)			
		Animal husbandry	Domestic	Businesses and services	Industrial production
1	Doan Ket	88	92	95	-
2	Tan Phong	87	93	96	-
3	Dong Phong	59	89	93	-
4	Quyet Tien	23	86	90	-
5	Quyet Thang	52	85	92	-
6	San Thang	78	61	81	90
7	Nam Loong	10	59	78	-
	Average (±SD)	57(±24)	81(±12)	89(±6)	90(±0)

Table 4 shows that the rate of treated wastewater before discharging into water receiving bodies varies greatly between administrative units. Among the four sectors, animal husbandry has the lowest rate of treated wastewater (average of 57%) while the variation is very large, ranging from 10-88%. These data were used to provide parameters for estimating the discharge load to the environment based on equation (1).

3.3. Total pollutant load in Lai Chau City

In order to calculate the total pollution load for the entire Lai Chau city, the loads by sub-basins were first calculated according to the method described in Figure 2. The total pollution discharge load which is represented by four basic environmental parameters were then summarized by different sources (Table 5).

Table 5. Pollutant discharge load from major sources in Lai Chau city adjusted for local treatments.

Pollution discharge load	Parameters (*tons/year) and % in brackets*			
	COD	BOD	N-total	P-total
Point sources:				
Domestic (human)	382.5 (20.8)	202.5 (19.7)	13.5 (5.1)	3.8 (4.1)
Animals	1088.6 (59.1)	619.3 (60.3)	161.5 (61.4)	70.4 (75.1)
Industry	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Businesses and services	21.9 (1.2)	11 (1.1)	2.2 (0.8)	11 (11.7)
Non-point sources:				
Land use surface run-off	349.8 (19.0)	194.5 (18.9)	85.8 (32.6)	8.6 (9.2)
Total pollution load	1842.9 (100)	1027.3 (100)	262.9 (100)	93.8 (100)

IOP Conf. Series: Earth and Environmental Science 1345 (2024) 012012 doi:10.1088/1755-1315/1345/1/012012

Table 5 shows that the pollutant discharge load from animal husbandry (at households and farms) accounts for 59.1-75.1% compared to all the sources combined. This indicates that these sources create the greatest pressure for environmental management in the study area. As the total land area of Lai Chau city is 7020 ha (Table 3), pollutant load densities characterized by COD, BOD, N-total and P-total are 262.6, 146.3, 37.5 and 13.4 kg/ha/year, respectively. These loading levels from Lai Chau city are relatively low compared to that of other provinces in Vietnam, e.g. equivalent to about 80% of Bac Giang province [12]. This comparison is understandable as the population density and production intensity in Bac Giang are much higher than in Lai Chau city.

3.4. Spatial variation of pollutants

The distribution of pollutants over the sub-basins is presented in Figure 4 where darker colors represent higher pollutant loads accumulated in the sub-basins. The highest pollutant load is concentrated in some residential areas within the five wards of the city center. This is because the central area is both densely populated and where many farms still exist with a large number of livestock.

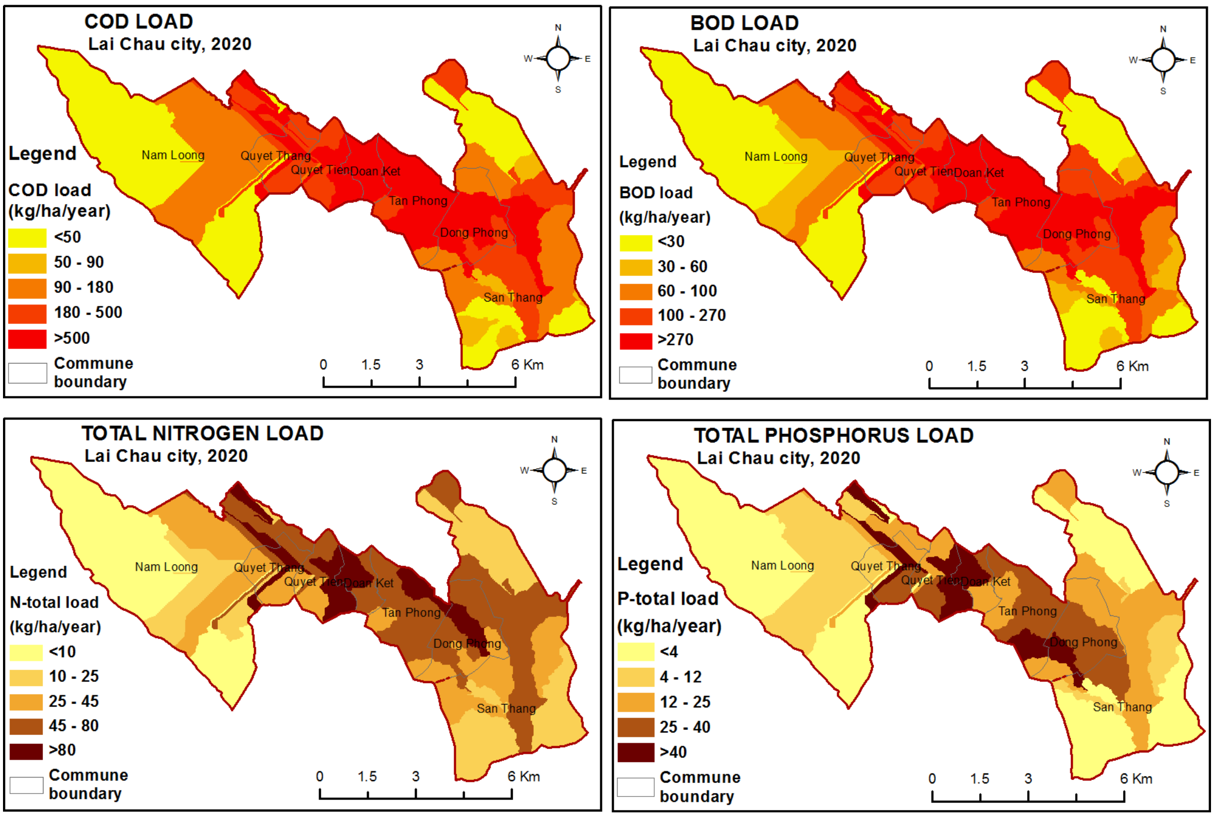

Figure 4. Maps of total pollutant loads in Lai Chau city.

In order to verify spatial differences with respect to characteristics of the pollutant sources, the HCA was applied. The 65 sub-basins were classified into three statistically significant clusters with a Euclidian distance (percentage of Dlink/Dmax) of greater than 5 (Figure 5a).

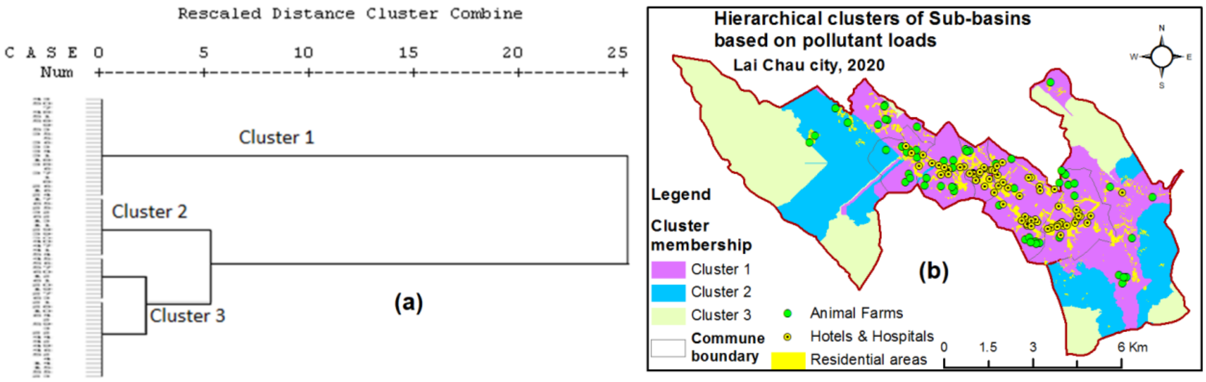

Figure 5. (a) Dendrogram of different sub-basin clusters using HCA Ward's Method; and (b) locations of HCA clusters in Lai Chau city

The HCA result is shown in Figure 5 (a). Cluster 1 is comprised of 28 sub-basins that cover residential areas, most animal farms, businesses, and services (i.e., hotels and hospitals) in the city center. This has the greatest pollutant load density among the three clusters. Cluster 2 includes 14 sub-basins where some small residential areas and animal farms are located in the fringe of the city center with medium pollutant loads. Cluster 3 corresponds to 23 sub-basins located in bare lands and forest of the remote areas where pollutant load is relatively low. The DA result indicates that there were 12 selected parameters that were significantly correlated with the three clusters classified by the HCA ($p<0.01$) as shown in Table 6.

Table 6. Significant parameters correlated with the three clusters selected by the DA

No.	Parameters (normalized data)	Pollutant sources
1	COD animal farming	Animal farming
2	BOD animal farming	
3	N-total animal farming	
4	COD domestic	Domestic
5	BOD domestic	
6	N-total domestic	
7	COD hotel and hospitals	Business and services
8	BOD hotel and hospitals	
9	COD surface run-off	Land uses (Surface run-off)
10	BOD surface run-off	
11	N-total surface run-off	
12	P-total surface run-off	

The overall accuracy for the spatial classification provided by the DA is 96.9%, indicating that the spatial variation of the pollutant sources is significant among the sub-basins.

3.5. Identification of significant pollutant sources

The data set of the 12 selected variables in Table 6 was further investigated by applying PCA with varimax rotation to identify the major possible source of variations in the study area. The KMO and Bartlett's sphericity tests were first performed on the correlation matrix of the parameters to examine the validity of using PCA. The tests provided a KMO value of 0.801 and Bartlett's significance ($p < 0.000$), both of which indicated that PCA may be useful in data reduction.

Three PCs were extracted from 12 variables with eigenvalues >1 explaining about 94.4% (or a nearly perfect fit) of the total variance in the original data structure. Within the list of 12 variables, only 8 of these parameters corresponded to varimax factor coefficients with correlations greater than 0.75, and hence, they were considered for further interpretation. The eight significant strong loading factors are COD, BOD and N-total from animal farming; COD and BOD from domestic; COD and BOD from business and services, and N-total run-off from agricultural land use (Table 7).

Table 7. Principal Component Analysis (PCA) of pollutant parameters.

No.	Parameters from different sources (Normalized data)	PC1	PC2	PC3
1	COD animal farming	0.945[a]	0.239	0.153
2	BOD animal farming	0.945[a]	0.250	0.146
3	N-total animal farming	0.942[a]	0.283	0.136
4	COD domestic	0.880[a]	0.448	0.092
5	BOD domestic	0.862[a]	0.483	0.081
6	N-total domestic	0.714	0.643	0.001
7	COD hotels and hospitals	0.285	0.906[a]	-0.135
8	BOD hotels and hospitals	0.259	0.903[a]	-0.138
9	COD surface run-off	0.513	0.723	0.347
10	BOD surface run-off	0.513	0.722	0.348
11	N-total surface run-off	0.061	-0.161	0.924[a]
12	P-total surface run-off	0.461	0.587	0.657

[a] Significant factors (varimax factor coefficients >0.75)

PC1 explains 72.3% of the total variance with strong positive loading for COD, BOD and N-total from animal farming, and COD, BOD and N-total from domestic sources. Therefore, PC1 represents farming & domestic sources. PC2 accounts for 12.9% of the total variance with a strong positive loading on COD and BOD from hotels and hospitals; thus, PC2 indicates business & service sources. PC3 accounts for 9.2% of the total variance. It is strongly and positively related to N-total from surface run-off, indicating pollutant run-off from agricultural land uses.

To identify significant pollutant sources for each cluster of sub-basins, the plot of scores for the three PCs associated with the three HCA pollutant clusters are presented in Figure 6. The plot of factor scores shows high associations of the pollutant sources and clusters, which can be further discriminated as follows:
Cluster 1 (red circles, which represent sub-basins in the city center) has very high PC1 and PC2 scores (in the top-right of Figure 6a), corresponding to a high pollutant load from animal farming & domestic, and business & service sources. This indicates that animal farming & domestic and businesses & services are two main sources of pollutant variation in the city center.
Cluster 2 (yellow triangles, which represent sub-basins in the fringe of the city center) tends to be concentrated in the middle of all PC1 (Figure 6b) but is low in PC2 (Figure 6a and Figure 6c). Therefore, cluster 3 has been moderately affected by pollutant sources from animal farming, domestic and surface run-off. The main sources of pollution for cluster 2 (the fringe of the city center) in this case are shared by both animal farming & domestic and nitrogen run-off.
Cluster 3 (cyan squares, which represent sub-basins in bare land and forest areas) is scattered in the bottom-left of Figure 6a, indicating that it has low scores or less pollutant loads in PC1 (farming & domestic) and PC2 (businesses & services). In Figure 6b, cluster 3 has some sub-basins located in the upper left, indicating high scores or high pollutant loads in PC3 (nitrogen run-off). In other words, nitrogen run-off is the main source of pollutant variation in the sub-basins located in the remote areas of the study region.

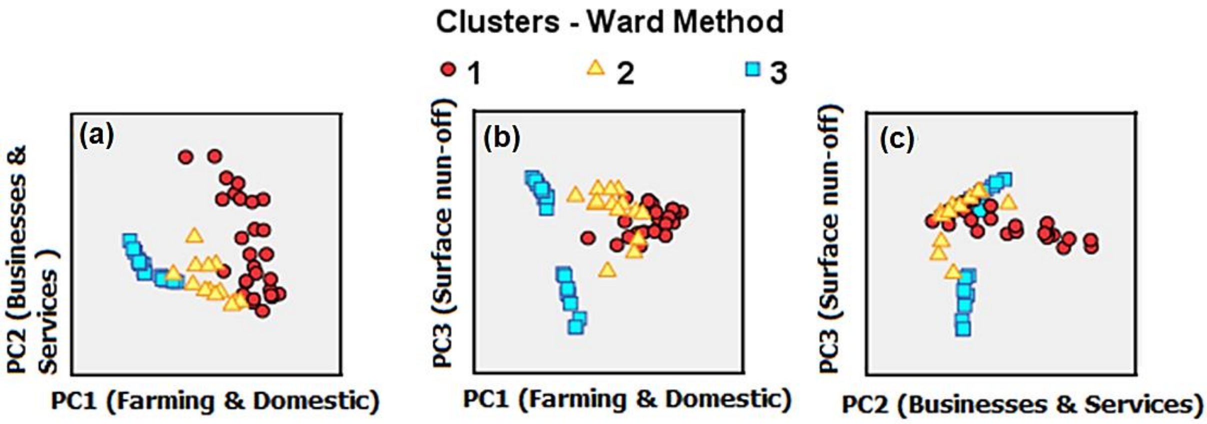

Figure 6. Component plot for sub-basin clusters in rotated space for the 3 principal components (PCs).

To visualize the pollution patterns, the strong positive loading parameters from animal farming (BOD), domestic (BOD), business & service (BOD) and surface run-off (N-total) are represented as maps in Figure 7. A visual comparison among the four selective parameter maps reveals a high consensus with the statistical analyses since almost all dark-colored areas from the map, which represents BOD from animal farming (Figure 7a) and BOD from domestic sources (Figure 7b), are mostly located within high pollution sub-basins (city center). Similarly, the map of BOD from the Business & Services sources is low in loading values but concentrated mostly in the city center. The map of N-total (Figure 7c), on the other hand, is quite dark around the map border, especially in the southeastern and northeastern sub-basins (cluster 1). These findings are similar to Su et al. [5], i.e., the majority of pollution arises from agricultural and domestic sources and are concentrated in the high-density populated areas.

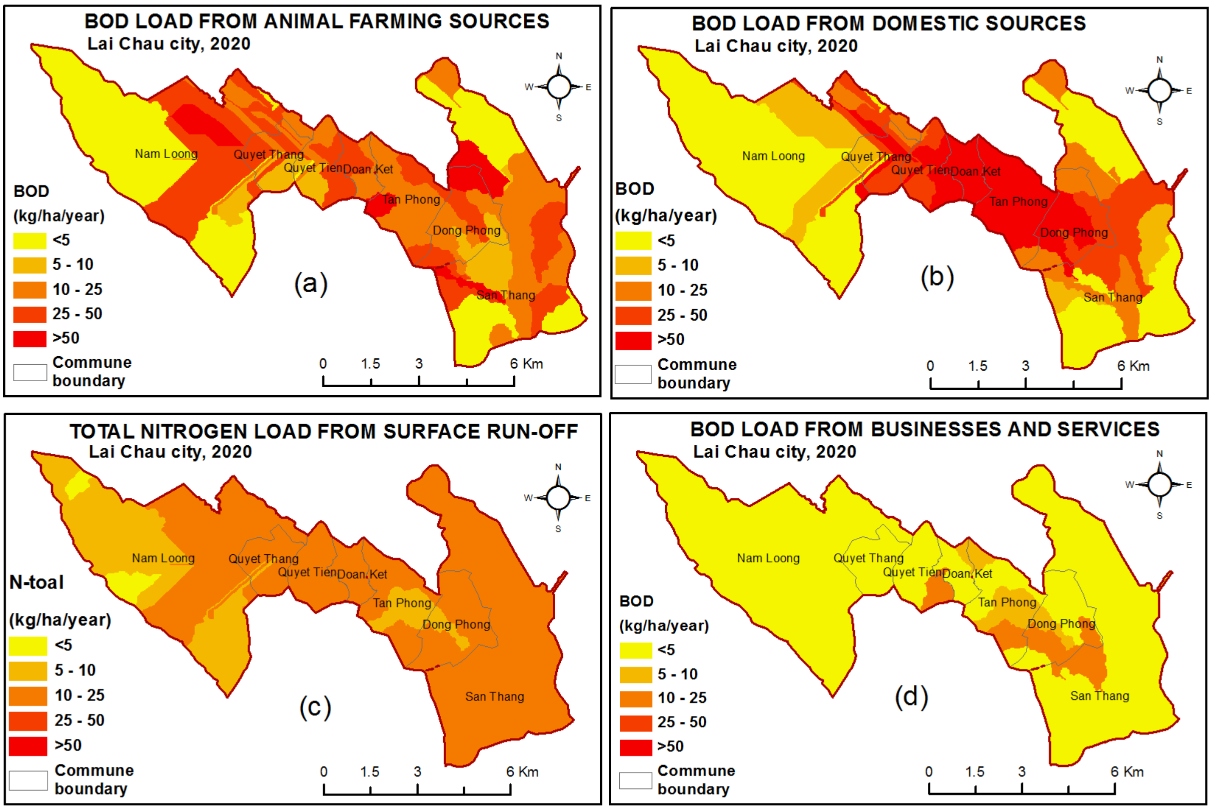

Figure 7. The pollutant loads from (a) animal farming, (b) domestic, (c) surface run-off and (d) business-service sources.

3.6. Policy implications

The above information demonstrates that environmental pressures with high pollutant loads are mainly concentrated in the downtown area of Lai Chau city, which is a commonly found phenomenon in most cities [12]. However, the difference between this study area and other cities is that the environmental pressures do not only come from urban domestic sources but also animal farming. The topographical conditions are one of the main reasons for this difference. Since most sub-basins located on the northwest and southwest borders tend to flow inward, the pollutant sources from farming and agricultural run-off are accumulated in the city center. In addition, another difference is that the source of wastewater from industrial production in the city is insignificant because industry in the city is newly developed. Statistics and spatial comparisons have shown that the role of this source creates insignificant pressure on the sub-basins.

In a conventional controlling pollutant approach, it is necessary to focus on sources with high loads. However, due to the topographical characteristics described above, sources with low loads outside the city center still need to be controlled in relation to the loading capacity of the inner-city sub-basins. In the context of Lai Chau, the total controlling mode, which is currently applied in the city [28] and many other provinces [29], is not a proper management approach. The existing production layout and differences in the pollutant allocated capacity in Lai Chau city requires differential proportional reduction targets for pollutant discharge between sub-basins based on their pollutant loads and the allocated capacities.

In order to reduce the pollutant yield from the high-density areas in the city center, it is necessary to adjust the spatial distribution of the pollutant sources, which means to reduce the amount and density of pollution from the sub-basins entering the city center while appropriately moving the amount and density of pollution to the lower loading sub-basins far from the center. For instance, when the domestic wastewater load is high and approaching the limit, the "at source" control approach must target the investment to solutions to increase the treatment rate and efficiency. Similarly, control solutions for animal farming sources can be combined to increase the treatment efficiency or relocate the farms to remote areas, i.e., on sub-basins that are experiencing low pollutant loads. Although the livestock zoning policy which relocates the concentrated farms out of residential areas has been implemented in many provinces [29], the synchronous implementation of this policy for Lai Chau still has many difficulties. One of the main reasons is that farm relocation requires large amounts of capital with significant financial support from the local government. However, Lai Chau is a remote province, mobilizing financial capital is not as easy as other provinces. Therefore, the fundamental solution of animal waste still relies on the "at source" control or the reuse of waste to produce organic fertilizer for agriculture. Of course, the reuse of waste needs to be balanced with the pollutant loads of neighboring clusters following the circular economy fashion applied throughout the city.

As Lai Chau is a young city, commonly applied techniques including building sewage treatment plants, increasing recycling and sewage disposal from point sources should be considered in the local land use planning. For non-point sources, improvement of planting patterns, reduction of fertilizer amount and other measures are also required to reduce and control the pollution loads.

4. Conclusions

Lai Chau city was spatially divided into 65 sub-basins, which locally accumulated pollutants from discharged sources before discharging into three river systems. Most pollutant discharging sources are located in sub-basins in the city center. Among the 5 main pollutant sources specified by the VEA, the pollutant load from animal husbandry accounts for 59.1-75.1% of total pollutant load. This indicates that these sources create the greatest pressure for environmental management in the city. The highest pollutant load was concentrated in some of the residential areas of the five wards, located in the city center.

The 65 sub-basins in the city were classified by HCA into three statistically significant clusters based on the similarity of pollutant factors. Cluster 1 includes sub-basins in the remote areas; cluster 2 consists of most sub-basins in the city center and cluster 3 has some sub-basins in suburban areas. PCA revealed the farming & domestic sources had the strongest positive loading factors for the sub-basins located in the city center and its fringe. N-total run-off is the main source of pollutant in the sub-basins located in the remote areas. Thus, the study demonstrates the efficacy of using a GIS-based approach informed by local knowledge to identify pollution sources that significantly contributed to the pollutant loads in different parts of the city. The results of the study indicate that the current environmental management approach of the city, applying the total controlling mode to manage pollutants, is not suitable for the pollution distribution pattern identified within the city.

The findings suggest that special attention should be paid to managing the waste treatment at animal farms and in residential areas to improve the water quality of the local community. The management approach should be changed from the total controlling mode, which is currently applied in the city, toward differential proportional reduction targets of pollutant discharge between sub-basins based on their pollutant loads and the allocated capacities. The policy solutions to reduce the pollutant yield from high-density areas include the adjustment of the spatial pollutant distribution, which requires legal enforcement to improve the treatment rate and efficiency of farming and domestic sources, or relocation of farms to remote areas on sub-basins that are experiencing low pollutant loads.

Acknowledgments

This research was funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 105.99-2018.318.

References

[1] UN Statistics Division 2020 SDG Indicators *Metadata repository* (From https://unstats.un.org/sdgs/metadata/)
[2] Novotny V 1995 *Non Point Pollution and Urban Stormwater Management* (CRC Press)
[3] Al Bakri D, Rahman S and Bowling L 2008 Sources and management of urban stormwater pollution in rural catchments, Australia *Journal of Hydrology* **356**(3), 299-311 (Doi: https://doi.org/10.1016/j.jhydrol.2008.04.017)
[4] Huang J, Li Q, Tu Z, Pan C, Zhang L, Ndokoye P and Hong H 2013 Quantifying land-based pollutant loads in coastal area with sparse data: Methodology and application in China *Ocean & Coastal Management* **81**, 14-28 (Doi: https://doi.org/10.1016/j.ocecoaman.2012.07.011)
[5] Su Y, Liang S, Zhang P and Wang X 2019 Spatial Distribution Assessment and Source Apportionment of Land-Based Pollutant Yield from Qingdao, China. *Journal of Ocean University of China* **18**(1), 151-158 (Doi: 10.1007/s11802-019-3548-8)
[6] Burigato Costa, da Silva Marques, Almeida A K, Leite I R and de Almeida I K 2019 Applicability of water quality models around the world—a review *Environmental Science and Pollution Research* **26**(36), 36141-36162 (Doi: 10.1007/s11356-019-06637-2)
[7] Chinyama A, Ochieng G M., Nhapi I and Otieno F A O 2014 A simple framework for selection of water quality models *Reviews in Environmental Science and Bio/Technology* **13**(1), 109-119 (Doi: 10.1007/s11157-013-9321-3)
[8] Lee J H., Ha S R and Bae M S 2010 Calculation of diffuse pollution loads using geographic information *Desalination and Water Treatment* **19**(1-3), 184-190 (Doi: 10.5004/dwt.2010.1911)

[9] Yan C A, Zhang W, Zhang Z, Liu Y, Deng C and Nie N 2015 Assessment of water quality and identification of polluted risky regions based on field observations & GIS in the Honghe River watershed, China *PloS one* **10**(3) (Doi: 10.1371/journal.pone.0119130)
[10] VEA 2019 *Decision number 154/QĐ-TCMT dated 15/2/2019 re. Issueing technical guideline on estimating Total maximal dialy load of rever water* Vietnam Environment Administration (VEA)
[11] Office of Environmental Management of Enclosed Coastal Sea 2011 Guidance for Introducing the Total Pollutant Load Control System (TPLCS) (from https://www.env.go.jp/en/water/ecs/pdf/english.pdf)
[12] An N T, Lan N P, Cong V H, Duong N H and Huong Giang N T 2020 Environmental Pressure from Pig Farming to Surface Water Quality Management in Yen Dung District Bac Giang Province *VNU Journal of Science: Earth and Environmental Sciences* **36**(1) (Doi: 10.25073/2588-1094/vnuees.4552)
[13] Khoi D N and Thang L V 2017 Climate change impacts on streamflow and non-point source pollutant loads in the 3S Rivers of the Mekong Basin *Water and Environment Journal* **31**(Issue 3), 8 (Doi.org/10.1111/wej.12256)
[14] Ngo T S, Nguyen D B, Rajendra P S 2015 Effect of land use change on runoff and sediment yield in Da River Basin of Hoa Binh province, Northwest Vietnam *Journal of Mountain Science* **12**(4), 1051-1064 (Doi: 10.1007/s11629-013-2925-9)
[15] Phan D B, Wu C C and Hsieh S C 2011 Impact of climate change on stream discharge and sediment yield in Northern Viet Nam *Water Resources* **38**(6), 827-836 (Doi: 10.1134/S0097807811060133)
[16] Shrestha S and Kazama F 2007 Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan *Environmental Modelling & Software,* **22**(4), 464-475 (Doi.org/10.1016/j.envsoft.2006.02.001)
[17] Lai Chau Statistics Office 2020 *Lai Chau Statistical Yearbook 2020* Lai Chau city: Statistical Publishing House of Lai Chau
[18] Wilson J P 2018 *Environmental Applications of Digital Terrain Modeling*: John Wiley & Sons Ltd.
[19] Wang Y B, Liu C W, Liao P Y and 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 (Doi: 10.1007/s10661-013-3492-9)
[20] Conrad O, Bechtel B, Bock M, Dietrich H, Fischer E, Gerlitz L and Böhner J 2015 System for Automated Geoscientific Analyses (SAGA) 2.1.4 *Geoscientific Model Development Discussions,* **8**, 2271-2312
[21] Deng F, Lin T, Zhao Y and Yuan Y 2017 Zoning and Analysis of Control Units for Water Pollution Control in the Yangtze River Basin, China *Sustainability* **9**(8), 1374
[22] Al-Odaini N A, Zakaria M P, Zali M A, Juahir H, Yaziz M I and Surif S 2012 Application of chemometrics in understanding the spatial distribution of human pharmaceuticals in surface water *Environmental Monitoring and Assessment* **184**(11), 6735-6748 (Doi: 10.1007/s10661-011-2454-3)
[23] Singh K P, Malik A and Sinha S 2005 Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques: a case study *Analytica Chimica Acta* **538**(1), 355-374 (Doi.org/10.1016/j.aca.2005.02.006)
[24] Spanos T, Ene A and Simeonova P 2015 Chemometric expertise of the quality of groundwater sources for domestic use *Journal of Environmental Science and Health, Part A* **50**(11), 1099-1107 (Doi: 10.1080/10934529.2015.1047646)
[25] Osman R, Saim N, Juahir H and Abdullah M P 2012 Chemometric application in identifying sources of organic contaminants in Langat river basin *Environmental Monitoring and Assessment* **184**(2), 1001-1014 (Doi: 10.1007/s10661-011-2016-8)

[26] Zhang X, Jiang H and Zhang Y 2013 Spatial distribution and source identification of persistent pollutants in marine sediments of Hong Kong *Environmental Monitoring and Assessment* **185**(6) 4693-4704 (Doi: 10.1007/s10661-012-2897-1)
[27] Liu C W, Lin K H and Kuo Y M 2003 Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. *Science of The Total Environment* **313**(1), 77-89 (Doi: https://doi.org/10.1016/S0048-9697(02)00683-6)
[28] Lai Chau DONRE 2020 Annual Report on Environmental protection missions of Lai Chau Province pp. 82 (Lai Chau city: Lai Chau department of Natural resources and Environment)
[29] Nguyen T H G, Ngo T A, Le T T H, Yabe M., Nguyen T T, Vu N H and & T S Cao 2021 Recycling Wastewater in Intensive Swine Farms: Selected Case Studies in Vietnam *Fac. Agr., Kyushu Univ.* **7** (Doi: https://doi.org/10.5109/4363559)