INNSPUB

Journal of Biodiversity and Environmental Sciences (JBES)
ISSN: 2220-6663 (Print) 2222-3045 (Online)
Vol. 22, No. 5, p. 1-8, 2023
http://www.innspub.net

RESEARCH PAPER OPEN ACCESS

Assessment of spatial variations in surface water quality of Laguna Lake stations using cluster analysis approach

Joey S. Estorosos*, John Michael B. Genterolizo, Wilfred V. Rios, Kristilynn C. Saavedra, Mark Ian C. Andres, Ricky B. Villeta, Jonecis A. Dayap

School of Arts and Sciences, University of San Jose-Recoletos, Cebu, Philippines

Article published on May 05, 2023

Key words: Laguna lake, Water pollution, Water quality, Cluster analysis

Abstract

The Philippine archipelago consists of various natural water resources, an asset that helps bring speed to nation-building activities considering the country is a developing nation. Laguna lake is a prime example where economic activity and human settlement impact the natural resources, which leads to consequential problems relating to water pollution and fish kill. Applying the multivariate cluster analysis that take into account salient critical factors in the assessment of the spatial variation of water quality in Laguna Lake can give a better understanding of the problem. Cluster analysis identified two state clusters (high polluted station, and less polluted station) that are present in the lake, having unique water qualities associated with geographically located Laguna Lake water stations. Cluster analysis results revealed the presence of sewage systems in high polluted stations, specifically the Manggahan floodway. High polluted station is characterized by high concentration of inorganic phosphate and ammonia due to sewage effluents runoff. These monitoring stations in Cluster 1 are located further inland. In contrast, a less polluted station is known to be in areas near from commercial factories and households. Less polluted station is determined by lesser concentrations of inorganic phosphate and ammonia which leads to sufficient concentration of Dissolve Oxygen that is suitable for aquatic organisms.

***Corresponding Author:** Joey S. Estorosos ✉ joey.estorosos@usjr.edu.ph

Introduction

Human development is a fundamental phenomenon (Bronfenbrenner & Morris, 2007) that affects the marine biological ecosystem, the human economy, and degradation of the environment in general (Dayap *et al.*, 2023). The devastating impact of human activity on natural water resources, such as rivers and lakes (Twardochleb & Olden, 2016), is a cause for great concern, given that these bodies of water have played a critical role in the emergence and sustenance of civilizations throughout history. It is essential that we take urgent action to preserve and protect these vital resources for future generations. The Philippines is a famous archipelagic paradise, popularly surrounded by rivers and lakes, which are affected by various human economic activities whose settlement near and around the vicinity brings a huge impact on the life of the exquisite marine ecosystem in the country (Boquet, 2017). For example, the Laguna Lake in Luzon is a prominent tourist destination that offers wide business opportunities for tourists (Historical Sea Surface Temperature for Laguna Beach, 2013), locals, and the rest of the nearby population which directly impacts the quality of naturally abundant water resources (Barril & Tumlos, 2002) (M, Dp, Babilonia, M, & Rebancos, 2015) that are being exploited for the benefit of the locality (Concepcion, 1976). While various studies are conducted in Laguna lake (Jalbuena *et al.*, 2019) (Pariset, 1977)to bring awareness (Barril *et al.*, 1994) of how this impacts the lake's water quality (Schröder *et al.*, 2016), little is known to address the impact of spatial variations within the lake (García-Ferrer *et al.*, 2003). Thus, a study is needed for the assessment of water quality in the Laguna Lake conducted in various strategic locations.

This study aims to classify the water quality of Laguna Lake on well-located government-controlled (Masuda, 2019) water stations, using known seven critical parameters that were measured to assess water quality by means of cluster analysis. These critical parameters are listed as follows; BOD (Biological Oxygen Demand), DO (Dissolved Oxygen), pH levels, presence of fecal coliform, concentrations of ammonia, nitrates, and inorganic phosphates. The identification of these clusters will effectively provide a better understanding on how spatial variations can contribute to the quality of water and its impact on the natural marine environment. The results of this study can have an effect on scientific-based government actions to help develop human economic activities surrounding the Laguna Lake (Kada, 2014)and generate ideas to mitigate problems found within and near it (Ancog *et al.*, 2009). Finally, this study aims to contribute to the development of other Lakes in the country and bring about ideas for creating sound scientific government policies on lakes.

Materials and methods

Data Collection

This study employed exploratory data analysis, also known as data mining. It aimed to uncover new information without relying on a pre-existing hypothesis (Sim, 2003). To assess the spatial variations in Laguna Lake's surface water quality, we utilized the Laguna Lake Development Authority's 2018-2019 dataset (LLDA, n.d.) on water quality parameters for the nine monitoring stations: Central West Bay (I), East Bay (II), Central Bay (IV), Northern West Bay (V), South Bay (VIII), West Bay-San Pedro (XV), West Bay-Sta Rosa (XVI), Fish Sanctuary (XVII), and Pagsanjan (XVIII). The dataset is included in the LLDA's quarterly water monitoring report, which includes seven water quality parameters: Biological Oxygen Demand (BOD), Minimum Dissolved Oxygen (DO), Level of Acidity or pH, Nitrate, Inorganic Phosphate, Ammonia, and Fecal Coliform. Multivariate cluster analysis was performed for analyzing and interpreting the dataset using Minitab version 16.

Data Treatment

The identified parameters were subjected to multivariate cluster analysis in order to identify major clusters among the nine monitoring stations with high similarities in water quality. Cluster analysis was used to determine whether monitoring stations can be clustered into distinct water quality groups. In this study, Q-mode hierarchical cluster analysis (HCA) was applied to classify the sampling stations. For this procedure, the Wards linkage method with squared Euclidean distance was used to create a dendrogram that determines the similarities among the monitoring stations.

Data Analysis

The result of hierarchical clustering is a dendrogram, which is a tree-like diagram that illustrates the order in which the clusters were merged. We can utilize the dendrogram to determine the number of clusters we desire (Baker, R., 2010). After acquiring the clusters, a comparison of the water quality parameters of each cluster was conducted to determine whether there were any differences or similarities which can provide valuable insights about the condition of the lake and identify potential sources of pollution or areas that need further monitoring and protection. It can help local government efforts to improve water quality and ensure the sustainability of lake resources.

Results and discussion

Using multivariate cluster analysis, out of 9 sampling stations, two major clusters were identified as shown in Fig. 1 and Table 1. These sampling stations were clustered according to the similarity of the listed station's water quality (Yu *et al.*, 2013). This allows better classification of water stations having similar or closely related water qualities. For example, stations I, V, and VIII are found to closely resemble each other's water quality as shown in Fig. 1. It follows that such a unique cluster, which contained three stations sharing the same water quality will be labeled according to its level of classification. Moreover, their water quality is as quantitatively shown in table 1 as cluster 1 qualitatively labeled as the "Highly Polluted Stations" of Laguna Lake. This is expected since close introspection of the raw data and set locations of such stations show that it contained a significantly large amount of pollution.

According to (Poté, *et al.*, 2009) there are five major findings that revealed how spatial-temporal distribution contributes to water pollution. One of the main contributors is associated with regions near the river mouth, where entry of such pollutants is expectedly carried by the runoff water, especially during wet seasons (Dayap *et al*, 2023). Since stations, I, V, and VIII are near or close to the mouth of the Marikina River and Manggahan floodway; their locations would expectedly be polluted.

In contrast, those stations far from the mouth of the rivers are expectedly low in pollution. This is also clearly shown by the results found in Fig. 1, where the rest of the stations are located and are found to have a significantly low amount of pollution.

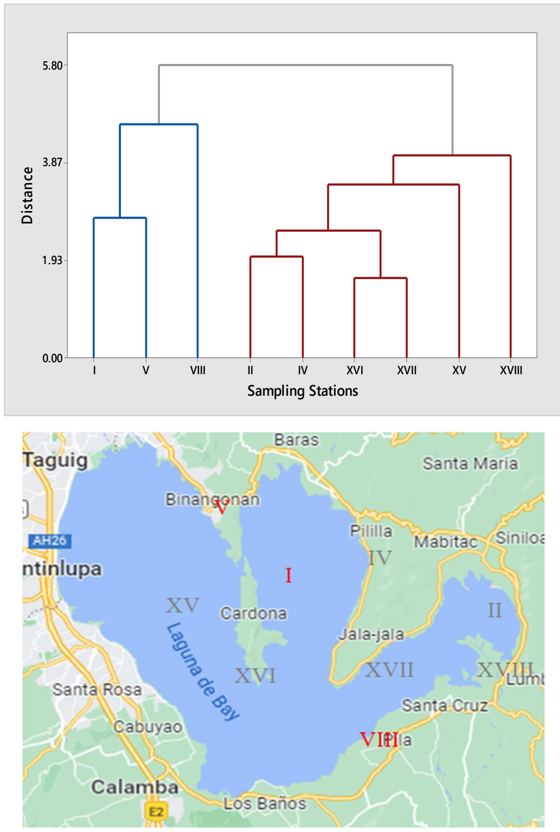

Fig. 1. a). Dendrogram showing clustering of water stations according to water quality parameters of Laguna Lake stations and b) Map of the Laguna Lake and the corresponding locations of the 9 sampling stations.

Table 1. Final Partition Number of clusters: 2.

	Number of Observations	Within cluster Sum of squares	Average distance From centroid	Maximum distance From centroid
Cluster 1 (HPR)	3	16.7120	2.3210	2.92669
Cluster 2 (LPR)	6	19.7493	1.7284	2.4864

Table 2. Cluster Centroids.

Variable	Cluster1	Cluster2	Grand centroid
BOD	0.68577	-0.342883	-0.0000000
DO	-0.42855	0.214277	0.0000000
Fecal Coliform	0.87020	-0.435101	0.0000000
pH	-0.80475	0.402377	-0.0000000
Ammonia	0.91297	-0.456483	-0.0000000
Nitrate	0.65606	-0.328032	-0.0000000
Inorganic Phosphate	1.00955	-0.504775	-0.0000000

Table 3. Distances Between Cluster Centroids.

Cluster1	Cluster2	Grand centroid
Cluster1	3.125600	.00000
Cluster2	3.12560	0.00000

The results of a Cluster Analysis undertaken to define the loadings of the 9 parameters in the water quality of the various locations of Laguna Lake and its tributaries from 2018 to 2019 data suggest the formation of two clusters. Table 2 shows that the first cluster for water quality in Laguna Lake during 2018 to 2019 had high positive loadings for Ammonia, Nitrate, Inorganic Phosphate, Fecal Coliforms, and BOD, with values of 0.91297, 0.65606, 1.00955, 0.87020, and 0.68577, respectively. Moreover, this cluster exhibits negative pH and DO loadings, with values of -0.80475 and -0.42855, respectively.

The second cluster of data for Laguna Lake's water quality in 2018–2019 shows positive loadings for DO and pH at 0.214277 and 0.402377, respectively. While there are negative loadings for the other factors. The current state of both clusters indicates that both locations are experiencing eutrophication, which implies that both areas are polluted as a result of the high nutrient levels found at each study site. However, as shown by the cluster analysis, the first cluster has a larger concentration of effluents such as inorganic phosphates, ammonia, and nitrates. Locations aggregated in the first cluster was classified and designated as "High Polluted Stations". Conversely, the sampling stations under the second cluster were categorized and designated as "Less Polluted Stations".

According to the cluster analysis results, Cluster 1, comprising of monitoring stations I, V, and VIII, has greater nutrient levels as evidenced by positive loadings of Ammonia, Nitrate, and Inorganic Phosphates in comparison to its negative loadings with cluster 2. These nutrients contributed to the high eutrophication level in the Cluster 1, creating an ideal habitat for algae and bacteria such as fecal coliforms to grow (Filik *et al.*, 2008, Serrano *et al.,* 2008, Kundu *et al.,* 2017, Vadde *et al.,* 2018). Algae feed on excess nutrients like phosphates and nitrogen, which leads to an algal bloom (Carpenter *et al.,* 1998, Paerl *et al.,* 2001, Chislock *et al.,* 2013). The sunlight is blocked by the large amount of developing algae, which also slows the photosynthesis of aquatic plants, lowering the levels of soluble oxygen in the affected areas (Conley, *et al.*, 2009). Furthermore, as these algae die, the organic matter is broken down by bacteria found in the river system enteric bacteria and other fecal coliforms flourish in this environment (Oswald *et al.,* 1953, Priyadharshini *et al.,* 2021).

As these fecal coliforms carry out their metabolic process by decomposing, they increase the need for biological oxygen by consuming even more of the little oxygen that is still dissolved in the environment, which creates the hypoxic situation known as a dead zone (Gerardi 2003, Shmeis 2018, Meshesha *et al.,* 2020). Fecal coliform, as a facultative anaerobe, can also benefit during eutrophic events by using nitrate as final electron acceptor once dissolve oxygen is depleted (Unden and Dunnwald 2008). As fecal coliform conducts cellular respiration, one of its by-products, CO_2, dissolves in water and forms carbonic acid and other inorganic carbon. Accumulation of dissolved inorganic carbon therefore decreases pH and cause acidification in natural waters (Weathers, Strayer, and Likens 2021).

As a result of the decomposition occurring within Cluster 1 and the Laguna Lake's location in a tropical nation with warm temperatures, the environment becomes more acidic than a typical and healthy river system. Although it is still regarded as contaminated in comparison to the monitoring stations in Cluster 2, the negative ammonia, nitrate, and inorganic phosphate loadings imply that it has lower nutrient levels than the other stations. This observation is further supported by its positive loading between pH and dissolve oxygen levels, which indicates that among the sampling stations that are included in this clustering, Cluster 2 has less pollution than Cluster 1 because it still has a manageable level of dissolve oxygen level and a negative loading of biological oxygen demand, implying that the eutrophication and decomposing processes in these sites are less severe.

The differing condition of monitoring stations from cluster 1 and cluster 2 has something to do with the areas as to which these stations were sampled. Most monitoring stations that are from cluster 2 are considered to be near the inland, though it might be the area that is close to contact with solid wastes during the dry season these areas are left to be sundried and baked under the heat of the sun (Sun, *et al.*, 2010). Despite the large levels of inorganic and organic nutrients that may be present in the sediments, the season's intense sunlight is too hot for algae and bacteria to survive (Sollie *et al.,* 2008, Tilahun and Kifle 2021). But, once the wet season begins, the water begins to rise, which is where the bacterial and algal communities grow and where the eutrophication process begins (Wen *et al.,* 2016, Chen *et al.,* 2022). This is however cut short when the wet season is on its full-swing. As rain deposits more precipitates in these stations, the algae and bacteria are wash away and the nutrient-rich condition of these sites are mixed with influx of water lowering its concentration level (Herrera and Nadaoka 2021). Furthermore, these stations are also prone to flooding which could also be a cause as to why the eutrophication and decomposition in these stations is not as intense as observed with the stations present in cluster 1. The most likely explanation for the positive nutrient loadings in Cluster 1 monitoring stations is that, during the dry season, they become the immediate recipient of polluted waters flowing from Marikina and Pasig River (Santos-Borja 1994, Herrera *et al.,* 2015, Paronda *et al.,* 2019), while, during the wet season, they receive industrial and agricultural runoffs from the Manggahan River floodways (Dela Pena and Pael 2009). Also, monitoring stations in Cluster 1 are located further inland, in which during the dry season, dirty waters from Marikina are mixed in these waters due to a backflow, and because the location is far from the land and the regular supply of the backflow, the waters from this area generally never completely dry up. This results in the mixing of nutrient-rich effluents at these locations, which, when accumulated over time, causes significant eutrophication and pollution in these areas.

As for the wet seasons, the concentration of these nutrient-rich effluents is supposed to decrease due to the addition of water in the form of rain, but because these areas act as catch basins for industrial and agricultural runoff from the Manggahan River, the concentration of these nutrients only serves to contaminate the waters of these sites further.

Conclusion

The cluster analysis identifies two distinct clusters that include unique factors taken from 7 variables associated with water quality, namely: high polluted, and less polluted stations. A high polluted station is known to be in areas near the mouth of rivers and floodways where large amounts of sewer sludges, bacteria, and algae are abundantly found, while geographically favoring the production of dissolved oxygen. High polluted station is characterized by high concentration of inorganic phosphate and ammonia due to sewage effluents runoff. In contrast, a less polluted station is known to be in areas far from outlets of rivers and creeks, while geographically promoting large-scale fish farms, teeming with various aquatic life forms and favoring the production of dissolved oxygen. Less polluted station is determined by lesser concentrations of inorganic phosphate and ammonia which leads to sufficient concentration of Dissolve Oxygen that is suitable for aquatic animals.

Acknowledgement

The authors would like to express their profound appreciation to the Laguna Lake Development Authority (LLDA) for providing the vital water quality data that served as the foundation for this work.

References

Ancog RC, Briones ND, Florece LM, Alcantara AJ, Alaira SA. 2009. Expansion of Environmental Usersâ€™ Fee System To Households For Enhanced Water Pollution Control in Laguna de Bay, Philippines. Journal of Environmental Science and Management **11(2),** Retrieved 4 10, 2022, from https://journals. uplb.edu.ph/index.php/jesam/article

Baker R. 2010. "Data mining for education," in International encyclopedia of education, Vol. 7, eds B.mcGaw, P. Peterson, and E. Baker (Oxford: Elsevier) 112-118. DOI: 10.1016/B978-0-08-044894-7.01318-X

Barril CR, Tumlos ET. 2002. Water quality trends and trophic state assessment of Laguna de Bay, Philippines. Aquatic Ecosystem Health & Management **5(2),** 115-126. Retrieved 4 10, 2022, from https://eurekamag.com/research/011/648/01164845

Barril CR, Madamba LS, Tumlos ET. 1994. Water quality of Laguna de Bay: status and trends. Retrieved **4(10),** 2022, from http://kimika. philippinechem.org/index.php/kimika/article/view

Boquet Y. 2017. Conclusion: Towards Sustainable Development in the Philippines? Retrieved **410,** 2022, from https://link.springer.com/chapter/10.1007/978-3-319-51926-5_23

Bronfenbrenner U, Morris P. 2007. The Bioecological Model of Human Development. Retrieved **410,** 2022, from https://onlinelibrary.wiley.com /doi/abs/10.1002/9780470147658.chpsy0114

Carpenter SR, Caraco NF, Correll DL, Howarth RW, Sharpley AN, Smith VH. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological applications **8(3),** 559-568.

Chen K, Duan L, Liu Q, Zhang Y, Zhang X, Liu F, Zhang H. 2022. Spatiotemporal Changes in Water Quality Parameters and the Eutrophication in Lake Erhai of Southwest China. Water **14(21),** 3398.

Chislock MF, Doster E, Zitomer RA, Wilson AE. 2013. Eutrophication: causes, consequences, and controls in aquatic ecosystems. Nature Education Knowledge **4(4),** 10.

Concepcion M. 1976. Children of the lakeshore : studies on the value of children in selected Laguna Lake towns. Retrieved **4(10),** 2022, from https://idl-bnc-idrc.dspacedirect.org/handle/10625/359

Conley DJ, Paerl HW, Howarth RW, Boesch DF, Seitzinger SP, Havens KE, Likens GE. 2009. Controlling eutrophication: nitrogen and phosphorus. Science **323(5917),** 1014-1015.

Filik Iscen C, Emiroglu Ö, Ilhan S, Arslan N, Yilmaz V, Ahiska S. 2008. Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat Lake, Turkey. Environmental monitoring and assessment **144(1-3),** 269-276.

Dayap J, Gentallan GR, Villeta R, Dionsay J. 2023. Unveiling the driving forces affecting carbon dioxide emissions in the Philippines. Journal of Biodiversity and Environmental Sciences **22(2),** 1-8.

Dayap J, Rios W, Estorosos J, Genterolizo JM, Villeta R, Andres MI. 2023. Assessment of seasonal variations in surface water quality of Laguna Lakestations using factor analysis. Journal of Biodiversity and Environmental Science **22(2),** 63-73.

García-Ferrer I, Camacho A, Armengol X, Miracle MR, Vicente E. 2003. Seasonal and spatial heterogeneity in the water chemistry of two sewage-affected saline shallow lakes from central Spain. Hydrobiologia 101-110. Retrieved **410,** 2022, from https://link.springer.com/article/10.1023/b:hydr.00 00008593.32525.1d

Herrera EC, Nadaoka K. 2021. Temporal dynamics and drivers of lake ecosystem metabolism using high resolution observations in a shallow, tropical, eutrophic lake (Laguna Lake, Philippines). Journal of Great Lakes Research **47(4),** 997-1020.

Herrera E, Nadaoka K, Blanco A, Hernandez EC. 2015. Hydrodynamic investigation of a shallow tropical lake environment (Laguna Lake, Philippines) and associated implications for eutrophic vulnerability. ASEAN Journal of Chemical Engineering **4(1),** 48-62.

Kada R. 2014. Environmental Risks to Food and Health Security in the Laguna Watershed. Retrieved **410,** 2022, from https://econpapers.repec.org/repec:sa

Kundu R, Aura CM, Nyamweya C, Agembe S, Sitoki L, Lung'ayia HB, Werimo K. 2017. Changes in pollution indicators in Lake Victoria, Kenya and their implications for lake and catchment management. Lakes & Reservoirs: Research & Management **22(3),** 199-214.

Masuda T. 2019. Interactive Governance for Sustainable Resource Use and Environmental Management: A Case Study of Yaman ng Lawa Initiative in the Laguna Lake Watershed, Philippines. Retrieved **410,** 2022, from https://link.springer.com /chapter/10.1007/978-981-13-2399-7_7

Meshesha TW, Wang J, Melaku ND. 2020. Modelling spatiotemporal patterns of water quality and its impacts on aquatic ecosystem in the cold climate region of Alberta, Canada. Journal of Hydrology **587,** 124952.

Oswald WJ, Gotaas HB, Ludwig HF, Lynch V. 1953. Algae symbiosis in oxidation ponds: III. Photosynthetic oxygenation. Sewage and Industrial Wastes 692-705.

Paerl HW, Fulton RS, Moisander PH, Dyble J. 2001. Harmful freshwater algal blooms, with an emphasis on cyanobacteria. The Scientific World Journal **1,** 76-113.

Pariset E. 1977. Laguna de Bay Water Resources Development. Journal of the Environmental Engineering Division **103(6),** 1021-1038. Retrieved 410, 2022, from https://cedb.asce.org/cedbsearch/ record.jsp?dockey=0007778

Paronda GRA, David CPC, Apodaca DC. 2019. River flow pattern and heavy metals concentrations in Pasig River, Philippines as affected by varying seasons and astronomical tides. In IOP conference series: Earth and environmental science (**Vol. 344, No. 1, p.** 012049). IOP Publishing.

Priyadharshini SD, Babu PS, Manikandan S, Subbaiya R, Govarthanan M, Karmegam N. 2021. Phycoremediation of wastewater for pollutant removal: a green approach to environmental protection and long-term remediation. Environmental Pollution **290,** 117989.

Poté J, Goldscheider N, Haller L, Zopfi J, Khajehnouri F, Wildi W. 2009. Origin and spatial-temporal distribution of faecal bacteria in a bay of Lake Geneva, Switzerland. Environmental Monitoring and Assessment **154(1),** 337-348. Retrieved 4 4, 2022, from https://link.springer. com/article/10.1007/s10661-008-0401-8

Santos-Borja AC. 1994. The control of saltwater intrusion into Laguna de Bay: Socioeconomic and ecological significance. Lake and Reservoir Management **10(2),** 213-219.

Schröder T, Melles M, Hoff J, Reicherter K. 2016. Sediments of the salt lake Laguna Salada (Southern Spain) & their application for climate studies. Retrieved 4 10, 2022, from https://publications. rwth-aachen. de/record/681298

Shmeis RMA. 2018. Water chemistry and microbiology. In Comprehensive analytical chemistry (Vol **81,** pp. 1-56). Elsevier.

Serrano L, DeLorenzo ME. 2008. Water quality and restoration in a coastal subdivision stormwater pond. Journal of environmental management **88(1),** 43-52.

Sollie S, Coops H, Verhoeven JT. 2008. Natural and constructed littoral zones as nutrient traps in eutrophicated shallow lakes. Hydrobiologia **605(1),** 219-233.

Sun Q, Wang S, Zhou J, Chen Z, Shen J, Xie X, Chen P. 2010. Sediment geochemistry of Lake Daihai, north-central China: implications for catchment weathering and climate change during the Holocene. Journal of Paleolimnology **43,** 75-87.

Tilahun S, Kifle D. 2021. Atmospheric dry fallout of macronutrients in a semi-arid region: an overlooked source of eutrophication for shallow lakes with large catchment to lake surface area ratio. Earth Systems and Environment **5,** 473-480.

Twardochleb LA, Olden JD. 2016. Human development modifies the functional composition of lake littoral invertebrate communities. Hydrobiologia **775(1),** 167-184.

Vadde KK, Wang J, Cao L, Yuan T, mc Carthy AJ, Sekar R. 2018. Assessment of water quality and identification of pollution risk locations in Tiaoxi River (Taihu Watershed), China. Water **10(2),** 183.

Weathers KC, Strayer DL, Likens GE. 2021. Fundamentals of ecosystem science. Academic Press.

Wen G, Li X, Qiu X, Cheng Y, Sun Y, Huang T. 2016. Characteristics of water pollution in typical reservoirs. Water pollution and water quality control of selected Chinese Reservoir Basins 25-94.

Yu S, Shang J, Zhao J, Guo H. 2003. Factor analysis and dynamics of water quality of the Songhua River, Northeast China. In Water, Air, and Soil Pollution (pp. **144(1),** 159-169.). Northeast China.