Environmental Pollution 308 (2022) 119611

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Indices and models of surface water quality assessment: Review and perspectives☆

Tao Yan [a,b], Shui-Long Shen [a,*], Annan Zhou [b]

[a] MOE Key Laboratory of Intelligent Manufacturing Technology, Department of Civil and Environmental Engineering, College of Engineering, Shantou University, Shantou, Guangdong, 515063, China

[b] Discipline of Civil and Infrastructure, School of Engineering, Royal Melbourne Institute of Technology (RMIT), Victoria, 3001, Australia

ARTICLE INFO

Keywords:
Influential factors
Assessment indices
Assessment models
Management
Surface water

ABSTRACT

Many technologies have been designed to monitor, evaluate, and improve surface water quality, as high-quality water is essential for human activities including agriculture, livestock, and industry. As such, in this study, we investigated water quality indices (WQIs), trophic status indices (TSIs), and heavy metal indices (HMIs) for assessing surface water quality. Based on these indices, we summarised and compared water assessment models using expert system (ES) and machine learning (ML) methods. We also discussed the current status and future perspectives of water quality management. The results of our analyses showed that assessment indices can be used in three aspects of surface water quality assessment: WQIs are aggregated from multiple parameters and commonly used in surface water quality classification; TSIs are calculated from the concentrations of different nutrients required for algae and bacteria, and employed to evaluate the eutrophication levels of lakes and reservoirs; HMIs are mainly applied for human health risk assessment and the analysis of correlation of heavy metal sources. ES- and ML-based assessment models have been developed to efficiently generate assessment indices and predict water quality status based on big data obtained from new techniques. By implementing dynamic monitoring and analysis of water quality, we designed a next-generation water quality management system based on the above indices and assessment models, which shows promise for improving the accuracy of water quality assessment.

1. Introduction

Water resources and quality are critical to human health, economic development, and the environment (Alver, 2019; Lin et al., 2020a). Global freshwater use, including by reservoirs, municipalities, industries and agriculture, has grown rapidly over the past 100 years (UNESCO, 2021). However, water quality deterioration has become a problem worldwide (Gad et al., 2021), with water pollution occurring in various regions and countries (Kim et al., 2021). Human activities and natural processes, including rock weathering, erosion, and climate change, affect water quality (Lan et al., 2020; Lyu et al., 2021). Fig. 1 shows algal blooms in lakes determined from imagery captured by remote sensing satellites in China. Surface water pollution is posing a serious challenge for water quality management (Zhang et al., 2021a). Assessing water quality is essential for water resource management (Renouf et al., 2017; Salerno et al., 2018). Various surface water properties should be evaluated when developing water resource management plans. The pollution of water bodies is threatening the ecological environment and human health (Le Moal et al., 2019); therefore, many indices for assessing surface water quality (e.g., water quality indices (WQIs), trophic status indices (TSIs), and heavy metal indices (HMIs)) based on water quality parameters (WQPs) have been designed to assess water quality.

The initial WQI was constructed by aggregating the physical and chemical factors of water bodies (Horton, 1965; Hurley et al., 2012). The WQI provides a more accurate overview of water quality variability in specific areas and can be used to effectively depict water quality (Rangeti et al., 2015; Tyagi et al., 2013). However, no universal WQI exists for evaluating surface water quality, though many modifications have been considered for generating different WQIs based on the situation in specific areas (Sutadian et al., 2016; Tyagi et al., 2013). The initial WQI was proposed by the National Sanitation Foundation (NFS); another form of the WQI was defined by the Canadian Council of Ministers of the

☆ This paper has been recommended for acceptance by Charles Wong.

* Corresponding author.
E-mail addresses: yantao@stu.edu.cn (T. Yan), shensl@stu.edu.cn (S.-L. Shen), annan.zhou@rmit.edu.au (A. Zhou).

https://doi.org/10.1016/j.envpol.2022.119611
Received 22 March 2022; Received in revised form 9 June 2022; Accepted 9 June 2022
Available online 15 June 2022
0269-7491/

Abbreviations

ABS_g	gastrointestinal absorption factor
ADD	average dose per day
ADD_{ing}	ADD by direct ingestion
ADD_{derm}	ADD by skin absorption
AHP	analytic hierarchy process
ANN	artificial neural network
AI	artificial intelligence
AT	average time
BEI	bacterial eutrophic index
BOD_5	5-day biochemical oxygen demand
BW	average body weight
C_w	heavy mental concentration
CA	cluster analysis
CART	classification decision tree
CCME	Canadian Council of Ministers of the Environment
CSF	cancer slope factor
CTSI	Carlson trophic status index
chl-a	chlorophyll-a
COD	chemical oxygen demand
COD_{Mn}	permanganate index
CR	carcinogenic risk
DL	deep learning
DO	dissolved oxygen
DOE	Department of Environment of Malaysia
DT	decision tree
EC	electrical conduction
ES	expert system
ESA	exposed skin area
ET	exposure time
f_e	exposure frequency
FA	factor analysis
F. coli	fecal coliform
GIS	geographic information system
HI	hazard index
HQ	hazard quotient
IoT	Internet of Things
ITSI	improved Carlson trophic status index
K_p	dermal permeability constant
LDA	linear discriminant analysis
ML	machine learning
MTSI	modified trophic status index
NFS	National Sanitation Foundation
NH_3–N	ammonia nitrogen
NO_3–N	nitrate nitrogen
NO_2–N	nitrite-nitrogen
OOB	out-of-bag
PCA	principal components analysis
PI	pollution index
pH	pondus hydrogenii
R_i	ingestion rate
RF	random forest
RfD	corresponding reference dose
RS	remote sensors
SA	sensitivity analysis
SD	secchi depth
SS	suspended solids
SVM	support vector machine
t_{ed}	exposure duration
T	temperature
TDI	trophic diatom index
TN	total nitrogen
TP	total phosphorus
TRIX	Vollenweiner trophic index
TS	total solids
TSI	trophic status index
TLI	trophic level index
TOPSIS	technique for order preference by similarity to an ideal solution
UAV	unmanned aerial vehicles
USEPA	United States Environmental Protection Agency
WQI	water quality index
WQP	water quality parameter

Environment (CCME) (CCME, 2001; Noori et al., 2019). Other WQIs have been modified or improved based on the NFS and CCME WQIs (Bhateria and Jain, 2016; Gao et al., 2020; Khan and Jhariya, 2017; Sutadian et al., 2016). Multivariate models based on various technologies (e.g., remote sensors and spectral signatures) have also been established to evaluate the water quality in different countries according to the WQP characteristics (Elsayed et al., 2021; El Osta et al., 2022; Gad et al., 2022). Carlson (1977) developed an effective method using TSI to evaluate the eutrophication of surface water. The China National Environmental Monitoring Center adopts the trophic level index (TLI), based on the WQPs and local properties of Chinese lakes to evaluate the eutrophication in lakes and reservoirs (CEMS, 2001; Ding et al., 2021).

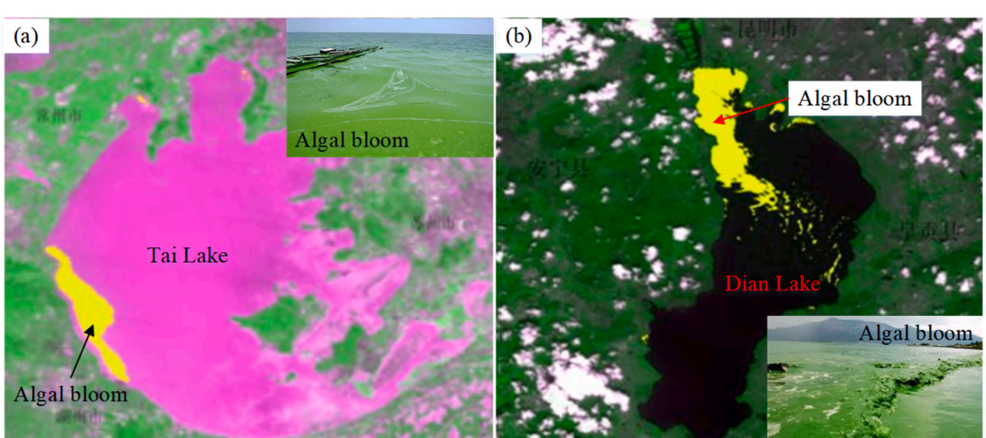

Fig. 1. Algal bloom monitoring in lakes via remote sensing satellites, (a) Tai Lake, and (b) Dian Lake (recreated from MEE, 2012).

The TLI has been recommended for estimating the eutrophication of lakes (Xiong et al., 2016; Zhou et al., 2020). Heavy metal pollution is another threat to water quality. Excess amounts of heavy metals not only affects human health, but also disrupts the aquatic ecosystems (Saha et al., 2016, 2017). Heavy metal distribution in surface water should be identified, and the risk of exposure, affecting human health by ingestion and dermal contact, should be controlled (Saha et al., 2017). The most widely used method for analysing the exposure risk of heavy metals was proposed by the United States Environmental Protection Agency (USEPA) (Alves et al., 2014; USEPA, 1989). Assessing the risk posed to human health by heavy metals involves evaluating ingestion and dermal adsorption with average dose per day (ADD), noncarcinogenic risk assessment with hazard quotient (HQ), hazard index (HI), and carcinogenic risk (CR) estimation (Alver, 2019; Ustaoğlu et al., 2021). Additionally, the HMIs were modified from the WQIs and have been employed for water quality classification based on geographical location and pollution indicators (Gad et al., 2021). Different water quality evaluation indices can be selected to assess the water quality levels based on the actual situation in an area.

Artificial intelligence (AI) technology has recently been introduced for water quality assessment (Akhtar et al., 2021). Expert systems (ESs) and machine learning (ML) are two important branches of AI technology that are widely used for water quality assessment (Tan, 2017). With the help of AI, indices for surface water quality assessment can be effectively generated by integrating ES methods and ML models with big data collected via sensors (Babbar and Babbar, 2017; Norouzi and Moghaddam, 2020; Zhao et al., 2020). Water quality assessment models have been enhanced using AI (Chou et al., 2018; Tiyasha et al., 2021). Recently, water quality prediction and classification models have been developed based on ES and ML technologies (Gad and El-Hattab, 2019; Li et al., 2022). The assessment models exhibit excellent performance in practical water quality management applications (Yu et al., 2022; Zhuang et al., 2022). However, the water quality dynamically changes, these methods can only evaluate water quality during a given period based on back-calculation or after-thought of water pollution (Shah et al., 2021). As such, recognising, monitoring, and expressing the quality of water, and integrating it into a decision-making system is critical for achieving sustainable water resource management. Therefore, an automated water quality management system capable of dynamically measuring and analysing water quality is urgently required so water can be quickly treated to improve its quality. Here, we provide an overview of the approaches for assessing water quality and some key approaches applied in water quality management. In this study, we aimed to (i) review current studies on water quality assessment approaches; (ii) compare existing assessment indices (WQIs, TSIs, and HMIs) and assessment models (ES methods and ML models) used for water quality evaluation; and (iii) discuss the potential for an early warning system based on AI and Internet of Things (IoT) technology to ensure the sustainability of water resource management.

2. Brief bibliometric analysis

We briefly analysed the bibliometric data collected from Web of Science. We retrieved relevant papers within the water quality field to prepare a database for review (Pan et al., 2021). Fig. 2 shows a flowchart of our bibliometric analysis of water quality studies. We used review papers to select keywords and increase the efficacy of our search (Sutadian et al., 2016; Wang and Yang, 2019). As a result, we selected keywords ((water quality index OR water quality risk OR water quality assessment OR water quality evaluation) AND (river OR lake)) to search for relevant publications. We input keywords into the database. We then selected the articles and analysed them according to the refining conditions. Finally, we selected 21,962 publications for the bibliometric analysis. The VOSviewer was used to analyse the keyword co-occurrence and cocitation relationships (Darko et al., 2020).

Fig. 3 visualises the keyword co-occurrence network based on the number of times an item occurred. We set the minimum number of occurrences threshold to 10. We also extracted items related to water quality. Based on the clustering results from VOSviewer, we divided the keywords into three research clusters (Fig. 3). Cluster 1 focused on WQPs. Chemical factors such as chemical oxygen demand (COD), total phosphorus (TP), and bioelements such as chlorophyll-a (chl-a) and total nitrogen (TN) are usually used to aggregate WQIs and assess water quality in rivers and lakes (Nong et al., 2020). Water quality levels can also be classified based on WQIs (Pennino et al., 2020). Cluster 2 focused on the contamination of surface water. The water status was assessed based on a single factor (the highest contaminant concentration level). Heavy metals are the primary pollutants in rivers and seriously threaten ecological balance. Therefore, an ecological risk assessment based on the risk indices of heavy metals was proposed to manage pollutant release (Rampley et al., 2020). Cluster 3 focused on water quality assessment methods. Principal components analysis (PCA) has been combined with cluster analysis (CA) to classify water quality levels (Jehan et al., 2020; Jabbar and Grote, 2019).

Fig. 4 presents the number of publications on the water quality assessment of surface water from 2016 to 2020. The number of publications rapidly increased in this period. Table 1 shows the top ten journals in terms of number of publications of papers related water quality assessment from 2016 to 2020 in the prepared database. The rapid growth in the number of publications indicated that water quality assessment has been a research hotspot for the past few years. Thus, a review of the latest developments in water quality assessment is required to promote further understanding and the application of evaluating methods in water quality management.

3. Factors influencing water quality

The factors that influence the quality of water bodies are complicated and include several WQPs (Bhateria and Jain, 2016). The WQPs can be

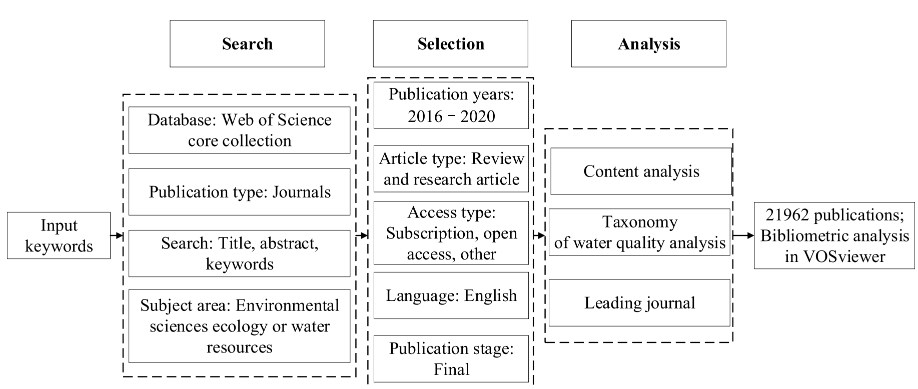

Fig. 2. Flowchart of bibliometric analysis.

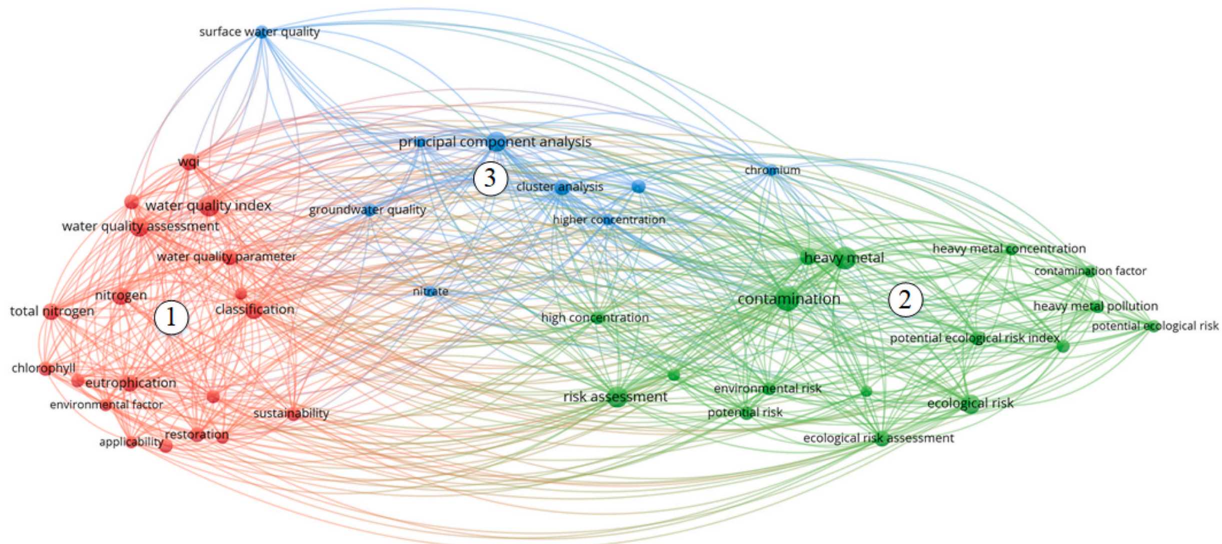

Fig. 3. Keywords co-occurrence network visualization.

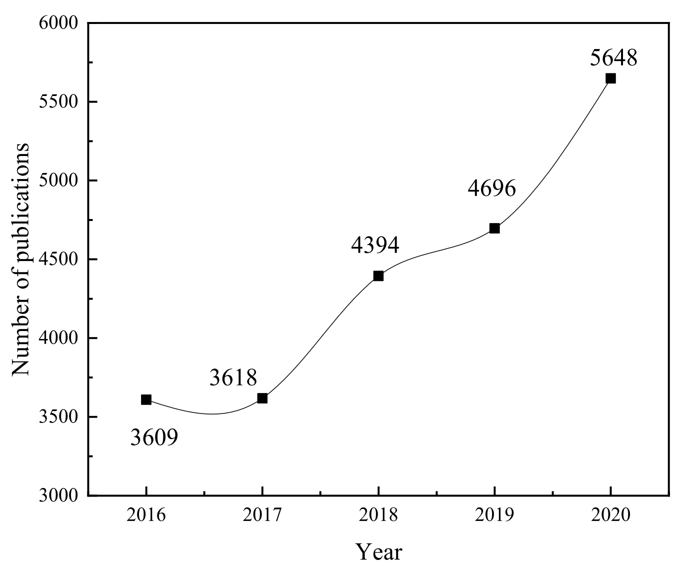

Fig. 4. Number of publications on surface water quality assessment.

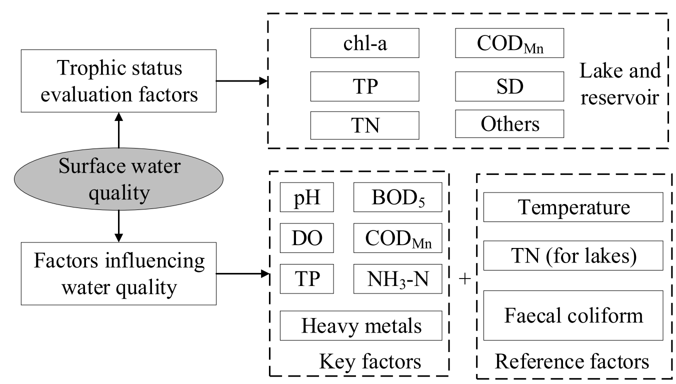

Fig. 5. Factors influencing trophic status and water quality.

Table 1
Top 10 journals in terms of publication numbers from 2016 to 2020.

Journals	Publications
Science of the Total Environment	1378
Water	1209
Environmental Science and Pollution Research	1192
Environmental Monitoring and Assessment	580
Environmental Earth Sciences	562
Journal of Hydrology	532
Ecological Indicators	444
Environmental Pollution	384
Environmental Science & Technology	347
IOP Conference Series: Earth and Environmental Science	321

directly measured by collecting water samples. According to their applications, WQPs are divided into influential and trophic status evaluation factors (Lin et al., 2020a). Fig. 5 shows the factors influencing water quality and trophic status. In China, the factors influencing water quality are key and reference factors (MEE, 2002): pH, permanganate index (COD_{Mn}), TP, 5-day biochemical oxygen demand (BOD_5), heavy metals, and other pollutants. Temperature (T), TN (for lakes), and faecal coliform (*F. coli*) were selected as reference factors to separately evaluate surface water quality (CEMS, 2001). In addition, chl-a, TN, TP, COD_{Mn}, and Secchi depth (SD) were selected as the primary factors for evaluating the eutrophication of lakes and reservoirs. In general, water temperature strongly affects the eutrophication of established water systems. Temperature can adjust the rates of chemical reactions in water bodies, consequently affecting the photosynthesis of aquatic plants and fish growth (Zhang et al., 2021b). Water oxygen content also decreases as the temperature increases. Water quality assessment is based predominantly based on multiple factors, and only a few assessment indices include *F. coli* as a factor when evaluating water quality (Sutadian et al., 2016). Water temperature and *F. coli* are the reference indicators for assessing water quality in Chinese surface water standards (MEE, 2002).

4. Overview of water quality assessment indices and methods

4.1. Water quality assessment indices

The initial WQI was proposed in 1965 on the basis of physical and chemical factors to evaluate water quality status (Horton, 1965; Hurley et al., 2012). Various scientists and experts have modified the WQI concept in consideration of different factors (Tyagi et al., 2013). Different forms of WQI equations have been developed as technology has advanced to increase the accuracy of water quality assessments (Sutadian et al., 2016). Table 2 shows the development of the main

Table 2
Development of water quality indices (WQIs) and water quality classification.

Items	Reference	Factors selection	Subindices and interpretation	*WQI* calculation	Water quality assessment
1	Brown and McClelland (1970) Kumar et al. (2019)	11 factors: DO, *F. coli*, pH, BOD_5, T, TP, TN, total solids, turbidity, pesticides, and toxic elements.	Q_i: subindex for ith water quality factor; W_i: weight of ith factor; n: number of factors.	NFS $WQI = \sum_{i=1}^{n} Q_i W_i$	91-100: Excellent; 71-90: Good; 51-70: Medium; 26-50: Bad; 0-25: Very bad.
2	CCME (2001); Khan et al. (2003) Lumb et al. (2006)	At least 4 factors; Not specified maximum number of factors.	Scope (F_1): number of variables whose objectives are not met; Frequency (F_2): number of times by which the objectives are not met; Amplitude (F_3): amount by which the objectives are not met; *nse*: normalized sum of excursions; n: number of factors.	CCME $WQI = 100 - \left(\frac{\sqrt{F_1^2 + F_2^2 + F_3^2}}{1.732}\right)$ $F_1 = \frac{Number\ \ failed\ \ variables}{Total\ \ number\ \ of\ \ variables} \times 100$ $F_2 = \frac{Number\ \ failed\ \ tests}{Total\ \ number\ \ of\ \ tests} \times 100$ $F_3 = \frac{nse}{0.01nse + 0.01}$ $nse = \frac{\sum_{i=1}^{n} excursions_i}{Number\ \ of\ \ tests}$ $excursions_i = \frac{Failed\ \ test\ \ value_i}{Objective_j} - 1$ $or\ \ excursions_i = \frac{Objective_j}{Failed\ \ test\ \ value_i} - 1$	95-100: Excellent; 80-94: Good; 60-79: Fair; 45-59: Marginal; 0-44: Poor.
3	Balan et al. (2012); Sutadian et al. (2016)	8 factors: T, DO, BOD, pH, NH_3–N, TP, TS, and *F. coli*.	n: number of subindices; *SI*: sub-index of ith factor.	$WQI = \sqrt{\frac{n}{\sum_{i=1}^{n} \frac{n}{SI_i^2}}}$	90-100: Excellent; 85-89: Good; 80-84: Fair; 60-79: Poor; 0-59: Very poor.
4	Pesce and Wunderlin (2000)	20 factors: ammonia, BOD_5, calcium, chloride, COD, DO, hardness, magnesium, nitrates, oil and greases, pH, phosphorus, dissolved solids, total solids, sulfates, T, *F. coli*, turbidity.	C_i: normalized value of ith factor; P_i: weight of ith factor; n: number of factors; k: subjective constant; C_{DO}: value due to *DO* after normalization; C_{cond}: value due to either conductivity or dissolved solids after normalization; C_{turb}: value due to turbidity after normalization.	$WQI = \frac{\sum_{i=1}^{n} C_i P_i}{\sum_{i=1}^{n} P_i}$ $WQI_{sub} = k\frac{\sum_i C_i x P_i}{\sum_i P_i}$ $WQI_{\min} = \frac{C_{DO} + C_{cond} + C_{turb}}{3}$	*WQI*: 91-100: Excellent; 71-90: Good; 51-70: Moderate; 26-50: Low; 0-25: Bad. WQI_{sub}: 1 = without apparent contamination; 0.75 = light contamination; 0.5 = contamination; 0.25 = highly contaminated.
5	Ma et al. (2013)	11 factors: T, DO, NH_3–N, nitrate nitrogen (NO_3–N), nitrite-nitrogen (NO_2–N), turbidity, pH, BOD_5, COD, dissolved inorganic phosphorus, and chl-a.	W_k: weight of kth factor; VF_k: score of principal component analysis (PCA). a_{ki}: value of ith factor component on VF_k; i: number of factors; j: maximum permissible concentration status.	$WQI = \sum_{k=1}^{n} (W_k VF_k)$ $VF_k = \sum_{i=1}^{n} (a_{ki} p_{ij})$	>0.4: Excellent; 0.3–0.4: Good; 0.2–0.3: Medium; 0.1–0.2: Poor; 0–0.1: Bad.
6	Zhang et al. (2015)	10 factors: Pb, Ni, Cd, Co, Hg, As, Cu, Mn, Zn, and Cr.	I: over-limit ratio of heavy metals; C_i: tested single heavy metal concentration; S_i: evaluation standard of heavy metals; n: number of factors; *Max*: maximum concentration value of a heavy metal.	$WQI = \sqrt{\frac{\left(Max\frac{C_i}{S_i}\right)^2 + \frac{1}{n}\left(\sum_{i=1}^{n}\frac{C_i}{S_i}\right)^2}{2}}$ $I = \frac{C_i}{S_i} \times 100\%$	1-2: Light pollution; 2-3: Pollution; 3-5: Heavy pollution; >5: Malignant pollution
7	Bhateria and Jain (2016)	10 factors: pH, DO, turbidity, BOD, conductivity, hardness, alkalinity, nitrate, and nitrite.	*RW*: relative weight; *AW*: assigned weight of each factor; n: number of factors; Q_i: a quality rating scale; C_i: factors value obtained from laboratory analysis; S_i: factors value from WHO; V_i: ideal value when pH = 7 and DO = 14.6; SI_i: subindices for each factor.	$WQI = \sum_{i=1}^{n} SI_i$ $SI_i = RW \times Q_i$ $Q_i = \frac{C_i}{S_i} \times 100$ $Q_{\mathrm{pH,\ DO}} = [(C_i - V_i)/(S_i - V_i)] \times 100$ $RW = AW_i / \sum_{i=1}^{n} AW_i$	>300: Unsuitable; 200-300: Very poor; 100-200: Poor; 50-100: Good; <50: Excellent.
8	Khan and Jhariya (2017)	8 factors: pH, hardness, alkalinity, chloride, nitrate, fluoride, calcium, magnesium.	W_i: relative weight; w_i: weight of each factor; n: number of factors;		<35: Excellent; 35-45: Good; 45-55: Moderate;

(*continued on next page*)

Table 2 (*continued*)

Items	Reference	Factors selection	Subindices and interpretation	*WQI* calculation	Water quality assessment
			C_i: concentration of each factor in a sample; C_{io}: ideal value of the factor in pure water; S_i: standard value; SI_i: sub-index of ith factor.	$WQI = \sum SI_i$ $SI_i = W_i Q_i$ $W_i = \frac{w_i}{\sum_{i=1}^{n} w_i}$ $Q_i = (C_i - C_{io})/(S_i - C_{io}) \times 100$	55-65: Poor; 65-75: Very poor; >75: Undrinkable water.
9	Nong et al. (2020)	16 factors: DO, BOD_5, COD_{Mn}, NH_3–N, TP, TN, pH, *F. coli*, T, SO_4^{2-}, F^-, Hg, As, Cu, Zn, Se.	T_i: measured concentration of ith factor; $S_{i,k}$ and $S_{i,k+n}$: standard thresholds of ith factor at level K and K + n; $I_{i,k}$ standard normalization value of factor classification; *n*: number of the equal values of the threshold. n = 1 when no equal threshold exists; C_i: normalized value of the ith factor.	$WQI_{min-nw} = \frac{\sum_{i=1}^{n} C_i}{n}$ $C_i = \begin{cases} 100 - \left[\frac{(T_i - S_{i,k})}{(T_{i,k+n} - S_{i,k})} \times 20n + I_{i,k}\right], & T_i \in (S_{i,k}, S_{i,k+n}] \\ 100 - \frac{T_i}{S_{i,k+n}} \times 20n, \quad T_i \in [0, S_{i,k}) \end{cases}$	81-100: Excellent; 60-80: Good; 41-60: Fair; 21-40: Poor; 0-20: Very poor.
10	Gao et al. (2020); Ustaoğlu et al. (2021)	12 factors: pH, DO, electrical conduction (EC), and 9 heavy metals.	W_i: weight of each factor and calculated based on eigenvalues by PCA. C_i: concentration of each factor; S_i: the limit value of drinking water for each heavy metal.	$WQI = \sum \left[W_i \times \left(\frac{C_i}{S_i}\right)\right] \times 100$	≥300: Undrinkable water; 200-300: Very poor; 100-200: Poor; 50-100: Good; <50: Excellent;

WQIs. In general, the WQIs are developed in three steps: (1) selecting fundamental water quality factors; (2) determining the quality function of each factor considered a subindex, and (3) aggregating subindices using a mathematical expression (Tyagi et al., 2013).

Fig. 6 shows a flowchart for generating WQIs. Selecting key factors is an essential step when generating a WQI because the subindex and WQI are calculated using the values of the selected parameters. WQIs consider various numbers of factors, ranging from four to twenty, as shown in Table 2 (Khan et al., 2003; Pesce and Wunderlin, 2000). However, selecting the influential factors remains a challenge because the initial water quality factors involve subjective assessment during index generation (Lumb et al., 2006; Rangeti et al., 2015). To address this problem, some approaches have been applied to reduce the uncertainty and inaccuracy in selecting influential factors (Akhtar et al., 2021). Expert judgement and statistical analysis are methods commonly used for factor selection (Akhtar et al., 2021). For expert judgement, three approaches (individual interviews, interactive groups, and the Delphi method) are employed to evaluate the influence of various factors. The Delphi method is widely used for factor selection owing to its effectiveness (Kumar et al., 2019). Statistical analysis, including Pearson

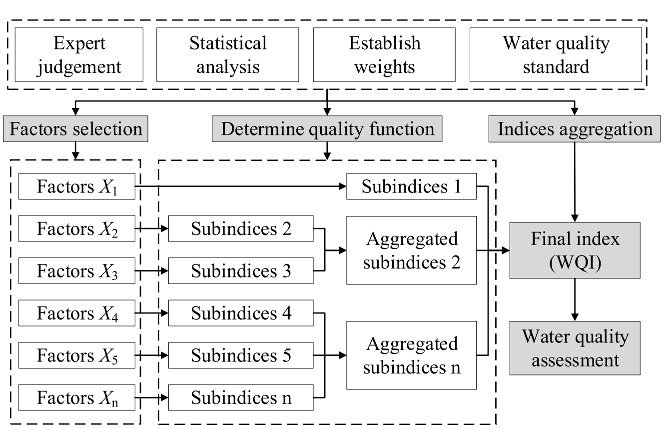

Fig. 6. Flowchart of WQI generation (recreated from Sutadian et al., 2016).

coefficient of correlation, PCA, and factor analysis, is another widely used approach for selecting influential parameters (Gao et al., 2020; Ustaoğlu et al., 2021). After parameter selection, the subindices are transformed to nondimensional scale values based on the selected parameters because the values of the selected factors have different units (Tyagi et al., 2013). In some WQIs, primary factors are directly considered subindices to produce the final index. However, in general, aggregated subindices can be generated based on primary factors or subindices. The subindices functions and rating curves are developed according to the weights assigned via expert judgement, statistical analysis, and water quality standards in different countries and regions (Sutadian et al., 2016). Then, the final WQI can be established based on the subindices and aggregated subindices (Sutadian et al., 2016; Tyagi et al., 2013). The water quality classification level is determined according to the value of the WQI. NFS proposed an initial WQI using the Delphi method to determine the influential factors (NSF WQI) (Noori et al., 2019). Additionally, CCME proposed another form of the WQI equation (CCME WQI) (Abbasi et al., 2012). The NSF and CCME WQIs are widely used for water quality assessment, and the WQIs shown in Table 2 have usually been developed and improved based on the NFS and CCME WQIs (Balan et al., 2012; Sutadian et al., 2016). Various methods for assigning weights to selected factors to improve the accuracy of WQI have been proposed (Pesce and Wunderlin, 2000; Zhang et al., 2015). PCA is commonly used to determine the weights of various factors (Ma et al., 2013). Additionally, WQIs have been further modified using the standard value of the factor in pure water based on situation in specific areas (Bhateria and Jain, 2016; Khan and Jhariya, 2017; Nong et al., 2020).

4.2. *Trophic status assessment indices*

Water eutrophication severely threatens aquatic ecosystems, as it results in the death of aquatic organisms (Hamilton et al., 2018). Therefore, evaluating eutrophication status based on physical and chemical water parameters can effectively provide a theoretical basis and technical guidance for engineers for predicting water quality levels (Ji et al., 2020). Many trophic status assessment methods involving simple, single water factors and comprehensive indices have been proposed to evaluate water eutrophication levels (Wang et al., 2019; Opiyo et al., 2019). We summarise the TSIs of the water bodies in this section.

Table 3 shows the representative trophic status assessment indices. Owing to various influential factors and situations, eutrophication assessment indices for water bodies have been developed in different areas. The simplest method of evaluating water trophic status involves using a single factor index (I_i) (Li et al., 2020). The single-factor index is used to compare the measured with the standard concentration value for each factor in trophic level classification, which can quickly describe the trophic level of water and the main factors influencing eutrophic water status. However, the single-factor method only reflects the water eutrophication of individual factor, and ignores the comprehensive influence of different indicators (Li et al., 2020). Therefore, the Nemerow synthetic pollution index (*PI*) was developed based on a single-factor index (Huang et al., 2018). Although some factor concentrations exceeded the standard values and induced water eutrophication, the average values of the indicators did not exceed the threshold value. The

Table 3
Trophic status assessment indices and classification.

Methods	Trophic status index	Parameters	Tropic status
Single factor index (I_i) (Li et al., 2020)	$I_i = \frac{C_i}{C_{i0}}$	C_i: actual measured value of ith factor; C_{i0}: standard value of ith factor.	$I_i \leq 1$: not eutrophic; $1 < I_i \leq 2$: light eutrophic; $2 < I_i \leq 3$: medium eutrophic; $3 < I_i$: heavy eutrophic.
Nemerow synthetic pollution index (*PI*) (Huang et al., 2018)	$PI = \sqrt{\frac{(I_{imax}^2 + I_{iave}^2)}{2}}$	I_{imax}: maximum value of ith single factor index; I_{iave}: average value of ith single factor index.	$PI \leq 0.7$: safe level of eutrophic; $0.7 < PI \leq 1$:minimum eutrophic; $1 < PI \leq 2$:light eutrophic; $2 < PI \leq 3$: medium eutrophic; $3 < PI$: heavy eutrophic.
Trophic diatom index (TDI) (Kelly, 1998)	$TDI = (25 \times WMS) - 25$ $WMS = \sum_{j=1}^{m} a_j s_j v_j / \sum_{j=1}^{m} a_j v_j$	*WMS*: weight mean sensitive; a_j: abundance of species j in sample; s_j: pollution sensitivity of species j; v_j: indicator value.	0-100: low nutrient concentrations to very high nutrient concentrations.
Vollenweiner trophic index (TRIX) (Vollenweider et al., 1998; Morales-Ojeda et al., 2010)	$TRIX = \frac{(\log 10[chl - a \cdot \mathrm{aD\%O} \cdot \mathrm{DIN} \cdot \mathrm{P}] + x)}{m}$	P: soluble reactive phosphorus; DIN: dissolved inorganic nitrogen; aD%O: absolute % of oxygen deviation from saturation; *x* and *m*: scale coefficients.	$TRIX \leq 4$: oligotrophic; $4 < TRIX \leq 5$: mesotrophic; $TRIX > 5$: eutrophic.
Comprehensive trophic level index (*TLI*) (Wang et al., 2019)	$TLI(\Sigma) = \sum W_j \cdot TLI(j), \quad W_j = \frac{r_{ij}^2}{\sum_{j=1}^{m} r_{ij}^2}$ $TLI(chl - a) = 10 \times (2.5 + 1.086 \ln chl - a)$ $TLI(TP) = 10 \times (9.436 + 1.624 \ln TP)$ $TLI(TN) = 10 \times (5.453 + 1.694 \ln TN)$ $TLI(SD) = 10 \times (5.118 - 1.941 \ln SD)$ $TLI(COD_{Mn}) = 10 \times (0.109 + 2.661 \ln COD_{Mn})$	W_j: relative weight of *TSI* for jth factor; r_{ij}: correlation coefficient between chl-a and jth factor; *m*: number of factors; *TLI*(*j*): *TSI* of jth factor;	$TLI(\sum) < 30$: oligotrophic; $30 \leq TLI(\sum) \leq 50$: mesotrophic; $TLI(\sum) > 50$: eutrophic; $50 < TLI(\sum) \leq 60$: light eutrophic; $60 < TLI(\sum) \leq 70$: moderate eutrophication; $TLI(\sum) > 70$: hypereutrophic.
Carlson trophic status index (*CTSI*) (Carlson, 1977; Opiyo et al., 2019)	$CTSI = \frac{TSI(SD) + TSI(chl - a) - TSI(TP)}{3}$ $TSI(chl - a) = 10 \times \left(6 - \frac{2.04 - 0.68 \ln chl - a}{\ln 2}\right)$ $TSI(TP) = 10 \times \left(6 - \frac{\ln(48/TP)}{\ln 2}\right)$ $TSI(SD) = 10 \times \left(6 - \frac{\ln SD}{\ln 2}\right)$	*TSI*(*j*): trophic status index of jth factor.	$CTSI < 40$: oligotrophic; $40 \leq CTSI < 50$: mesotrophic; $50 \leq CTSI < 70$: eutrophic. $70 \leq CTSI < 100$: hypereutrophic.
Modified Carlson trophic status index (*MTSI*) (Aizaki et al., 1981; Wen et al., 2019)	$MTSI = 0.297 TSI(SD) + 0.54 TSI(chl - a) + 0.163 TSI(TP)$ $TSI_M(chl - a) = 10 \times \left(2.46 + \frac{\ln chl - a}{\ln 2.5}\right)$ $TSI_M(TP) = 10 \times \left(2.46 + \frac{6.71 + 1.15 \ln(TP)}{\ln 2.5}\right)$ $TSI_M(SD) = 10 \times \left(2.46 + \frac{3.69 - 1.52 \ln SD}{\ln 2.5}\right)$	$TSI_M(j)$: trophic status index of jth factor.	$MTSI < 30$: oligotrophic; $30 \leq MTSI \leq 50$: mesotrophic; $MTSI > 50$: eutrophic.
Improved Carlson trophic status index (*ITSI*) (Yu et al., 2010)	$ITSI = TSI(\Sigma) = \sum_{j=1}^{m} [W_j \times TSI(j)]$ $TSI_I(TP) = 10 \times (9.436 + 1.488 \ln TP / \ln 2.5)$ $TSI_I(TN) = 10 \times (5.453 + 1.694 \ln TN / \ln 2.5)$ $TSI_I(COD) = 10 \times (0.109 + 2.438 \ln COD / \ln 2.5)$ $TSI_I(BOD) = 10 \times (2.118 + 2.363 \ln BOD)$ $TSI_I(NH_3 - N) = 10 \times (7.77 + 1.511 l \ln NH_3 - N / \ln 2.5)$	W_j: entropy weight for jth factor; *m*: number of factors; $TSI_I(j)$: trophic status index of jth factor.	$ITSI \leq 30$: oligotrophic; $30 < ITSI \leq 50$: mesotrophic; $50 < TLI(\sum) \leq 60$: eutrophic; $60 < TLI(\sum) \leq 60$: super eutrophic; $70 < TLI(\sum) \leq 100$: hypereutrophic.
Bacterial eutrophic index (BEI) (Ji et al., 2020)	$BEI = \frac{\text{Cyano} - \text{A}}{\text{Actino} - \text{A}} f\left(\frac{Y1}{Y2}, T\right)$	Cyano-A and Actino-A: abundances of Cyanobacteria and Actinobacteria; Y1 and Y2: functions fitted by using Cyano-A, Actino-A, and T; *f*(a): function fitted by using Y1/Y2 and T.	$BEI \leq 0.25$: oligotrophic; $0.25 < BEI \leq 3$: oligo-mesotrophic; $0.3 < BEI \leq 0.35$: mesotrophic; $0.35 < BEI \leq 0.65$: light eutrophic; $0.65 < BEI \leq 1.3$: middle eutrophic; $BEI > 1.3$: hypereutrophic.

method of assessing the Nemerow PI combining the average trophic factor value with the maximum value, was applied to evaluate water quality (Xu et al., 2010). The different trophic indicators have different influences on water quality and ecological systems. An improved Nemerow synthetic pollution index was established by assigning weights to trophic factors (Huang et al., 2018). The Nemerow synthetic pollution index highlights the impact of the factor that mostly affects trophic status, water quality, and the ecological system. The results of the Nemerow synthetic pollution index cannot be used to determine the eutrophication level of water in some cases because of the low sensitivity of the water factors. Therefore, more comprehensive TSIs have been proposed for assessing water eutrophication.

The Carlson trophic status index (CTSI) and comprehensive TLI are widely used to classify the eutrophication status of lakes and reservoirs (Carlson, 1977; Lin et al., 2020b; Liu et al., 2020). Water eutrophication can be qualitatively evaluated based on colour characterisation (CEMS, 2001). Table 4 presents a qualitative evaluation system of water quality and its levels. The colour of the water body is primarily influenced by three WQPs: TP, SD, and chl-a. Therefore, these three WQPs were analysed using their TSI equations, and the value of the final CTSI was obtained together with the TSI results for these three indicators (Opiyo et al., 2019). Subsequently, the final CTSI value was used to identify the eutrophication of lakes based on TSI values ranging from 0 to 100, as shown in Table 3. In the CTSI, transparency is considered a relative index of algal biomass. However, Aizaki et al. (1981) claimed that transparency is partially influenced by factors other than algal biomass. Therefore, a modified trophic status index (MTSI) based on chl-a concentration rather than transparency was proposed. Subsequently, more synthesised tropic status index models were established and improved based on CTSI, MTSI, and more water quality factors (Yu et al., 2010). In some lake ecosystems, the nitrogen concentration and COD_{Mn} may limit phytoplankton growth (Qin et al., 2020). Hence, TN and COD_{Mn} were also considered in a water eutrophication assessment (MEE, 2011). The TLI is a comprehensive index applied to determine the eutrophication of water in China. The TLI involves five parameters, that is, TN, SD, TP, chl-a, and COD_{Mn}, which are assigned relative weights calculated by the correlation coefficient between chl-a concentration and WQPs shown in Table 5. Additionally, the TSI of water quality based on diatoms, bacteria, and species, that is, Trophic diatom index (TDI), Vollenweiner trophic index (TRIX), and Bacterial eutrophic index (BEI), were also designed to evaluate water eutrophication in different water ecological systems.

4.3. Heavy metal evaluation and human health risk assessment

Heavy metal pollution globally poses a threat to the environment (García-Carmona et al., 2017). Heavy metals dissolved in water are crucial factors affecting water quality and potentially pose human health risks because they enter the food chain during plant and animal growth (Huang et al., 2018). Therefore, the concentrations and sources of heavy metals must be understood. Identifying the human health risks caused by heavy metals is also crucial (Saha et al., 2017). Heavy metal pollution can be evaluated based on WQIs and TSIs (e.g., the single-factor index, and Nemerow synthetic pollution index) (Li et al., 2020). Additionally, heavy metals that enter the environment can affect aquatic ecosystems and human health (Saha et al., 2017). Therefore, potential exposure risks in terms of human health by contaminated water should be estimated and applied in practice to protect human health (Alves et al., 2014).

Fig. 7 shows the risk assessment procedure for water polluted with heavy metals in China. Heavy metal risk assessment consists of four steps: hazard identification, exposure evaluation, risk characterisation, and control value calculation (MEE, 2014). In hazard identification step, environmental survey, water use, pollutants information should be collected and analysed to determine the exposed population. According to hazard identification results, the exposure routes, exposure model, and model parameters should be established, and the heavy metals exposure value should be calculated. Additionally, human health risks should be analysed based on carcinogenic and noncarcinogenic risk equations (USEPA, 2004). Then, the risk assessment model is used to calculate the CR and HQ for uncertainty analysis. Finally, the calculated value of the heavy metal risk is estimated as either an acceptable or unacceptable value. Next, risk management measures can be applied to ensure the safety of water use. The most widely used approach was proposed in USEPA guidance regarding human health exposure risk analysis during heavy metal assessment (USEPA, 1989).

The risk posed by the exposure to heavy metals to human health may occur through ingestion by the mouth or dermal absorption through skin contact (Ustaoğlu et al., 2021). Four health risk assessment indices, ADD, HI, HQ, and CR, were proposed to characterise the hazard of human body exposure to heavy metals. The ADD by direct ingestion (ADD_{ing}) and skin absorption (ADD_{derm}), which consider two population groups, the general population (adults) and sensitive population (children), are calculated as follows (Amiri et al., 2020; USEPA, 2004; Ustaoğlu et al., 2021):

$$\begin{aligned} ADD_{ing} &= \frac{f_e \times R_i \times t_{ed} \times C_w \times ABS_g}{AT \times BW} \\ ADD_{derm} & \frac{SA \times K_p \times ED \times ET \times CF \times EF \times C_w}{AT \times BW} \end{aligned} \quad (1)$$

where f_e is the exposure frequency; R_i is the ingestion rate; t_{ed} is the exposure duration; C_w is the heavy metal concentration; ABS_g is the gastrointestinal absorption factor; K_p is the dermal permeability constant; SA is the exposed skin area; ET is the exposure time; AT is the average time; BW is the average body weight; CF is 10.

The noncarcinogenic risks of heavy metals are generally estimated using HQ and HI. HQ is the ratio of ADD of each pollutant for a single exposure pathway to the corresponding reference dose (RfD). HI is the sum of the HQs for different heavy metals from all possible pathways. HI can be used to estimate the total noncarcinogenic risks of multiple exposure pathways (Li and Zhang, 2010; Zeng et al., 2019). The following equations are used to calculate HQ and HI (Alver, 2019).

Table 4
Systems of qualitative evaluating water quality (MEE, 2011).

Classification	Water quality	Characterisation of colour	Water quality function types
Level I–II	Excellent	Blue	Drinking water source level I protection zone, habitat for rare aquatic life, spawning grounds for fish and shrimp, etc.
Level III	Good	Green	Drinking water source level II protection zone, winter grounds, and migration channels for fish, shrimp, etc.
Level IV	Light pollution	Yellow	General industrial water and recreational water not in direct contact with the human body.
Level V	Middle pollution	Orange	Agricultural water and general landscape water.
Inferior level V	Heavy pollution	Red	Local climate regulation and poor use function.

Table 5
Correlation coefficient between water quality factors and chl-a in China (MEE, 2011).

Parameters	chl-a	TP	TN	SD	COD_{Mn}
r_{ij}	1	0.84	0.82	−0.83	0.83
r_{ij}^2	1	0.7056	0.6724	0.6889	0.6889
W_j	0.2663	0.1879	0.1790	0.1834	0.1834

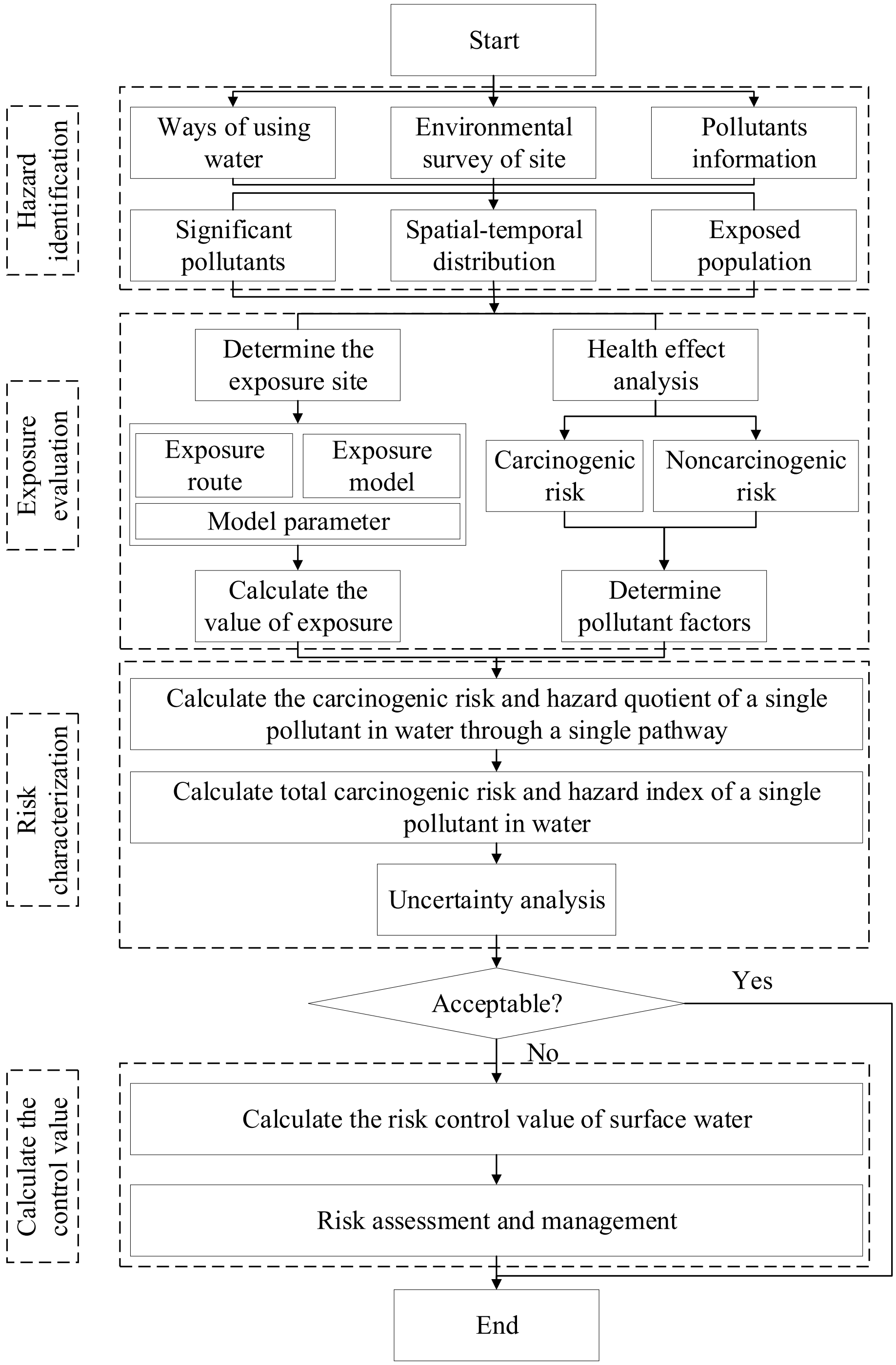

Fig. 7. Procedure of assessing the risk posed by heavy metals in polluted water.

$$HQ_{ing} = \frac{ADD_{ing}}{RfD_{ing}}, \quad HQ_{derm} = \frac{ADD_{derm}}{RfD_{derm}}$$
$$RfD_{derm} = RfD_{ing} \times ABS_g \tag{2}$$
$$HI = \sum (HQ_{ing} + HQ_{derm})$$

Based on the HQ and HI values, the noncarcinogenic risk can be categorised into four levels, which is presented in Table 6. When the value of HQ or HI exceeds one, potential noncarcinogenic health risk might pose a risk to human health. Additionally, CR is applied to assess the risk of developing cancer during a human lifetime because of exposure to carcinogenic heavy metals. The CR can be obtained using the following formula:

$$CR = ADD \times CSF \tag{3}$$

where CR is unitless; CSF is the cancer slope factor, (0.0015 and 0.00366 μg/kg/d for ingested and dermal As, respectively). The acceptable CR values range from 1×10^{-6} to 1×10^{-4}. When CR $< 1 \times 10^{-6}$, heavy metals might not pose a serious health hazard, where the carcinogenic risk is unacceptable when CR $> 1 \times 10^{-4}$ (Ustaoğlu et al., 2021). Risk management measures should be implemented to reduce risks and protect human health. Human health risk assessments often result in uncertainty (Alver, 2019). Therefore, Monte Carlo simulation is a widely used method for determining the probability of carcinogenic and noncarcinogenic risks (Amiri et al., 2020; Ustaoğlu et al., 2021).

Table 6
Noncarcinogenic risk level classification.

Risk level	Very low risk	Low risk	Medium risk	Very high risk
Values	HQ < 0.1 or HI < 0.1	HQ ≥ 0.1 or HI < 1	HQ ≥ 1 or HI < 4	HQ ≥ 4 or HI ≥ 4

5. Water quality assessment models

With the rapid development of computer science, AI technology has been introduced in assessment models (including ES and ML-based models) to assess water quality. ES methods and ML models use computers to improve the decision-making ability of human experts to solve complex problems with specialised knowledge and experience (Akhtar et al., 2021). Water quality assessment models have been developed based on ES and ML methods (Chou et al., 2018). Fig. 8 presents a flowchart of water quality assessment based on ES and ML methods. First, water parameters are collected from water samples using laboratory analysis, advanced remote sensing mapping technology, and a real-time monitoring system (Raju and Varma, 2017). The data are then input into the ES and ML methods to estimate the water quality and human health risk. Fig. 8 shows several ES methods frequently used to evaluate water quality, including fuzzy set, rough set, analytic hierarchy process (AHP), technique for order preference by similarity to an ideal solution (TOPSIS) and entropy methods (Baghapour et al., 2020; Lyu et al., 2021). Data–driving techniques and ML methods, such as sensitivity analysis (SA), CA, PCA, and random forest (RF), are applied in practice to identify the correlation between water quality factors and determine the water quality level and potential risk level (Akhtar et al., 2021; Marín et al., 2018; Norouzi and Moghaddam, 2020). Finally, the results of the water quality assessment are presented using maps produced via a geographic information system (GIS).

5.1. Water quality assessment with ESs

ES methods can emulate scientists in estimating the probability of alternatives for objectives, providing water management suggestions to engineers. The AHP is a widely used method in practice. Fig. 9 shows the water quality assessment process based on the AHP method. The AHP method based on multicriteria decision-making was developed by Saaty (1977). The AHP is a valuable tool for dealing with group or individual decision-making problems, and consists of three layers: objective, criteria, and alternative layers. In a study, the selected objective was water quality assessment. To evaluate water quality, multiple criteria for drinking, domestic, irrigation, and industry were selected based on the guidance of experts and scientists (Akhtar et al., 2021). The alternatives consisted of multiple WQPs. Additionally, fuzzy and rough set theories can be integrated into the AHP to improve its efficiency and effectiveness (An et al., 2014; Akhtar et al., 2021). Then, the relative weights of available alternatives can be determined using the AHP method. The subindices can be aggregated as a WQI based on the relative weights of the WQPs. Finally, the water quality can be classified by the aggregated WQI based on the AHP method.

The AHP method can also be directly applied to water quality assessment (Shi et al., 2020). To illustrate the capability of the AHP method, the process of assessing water quality for Liao River in the Liaoning Province of China is demonstrated as a case study. The water environment risk in Liao River was divided into cumulative and sudden environmental risks. The comprehensive evaluation system was established by the cumulative environmental risk assessment method from the USEPA, which was used to identify the cumulative risk indicators (W_i). The sudden risk indicators (X_j) were identified using the water environmental emergency risk index system from the stress-state-response model. Finally, twenty-four cumulative risk indicators and ten sudden risk indicators were selected as alternatives for

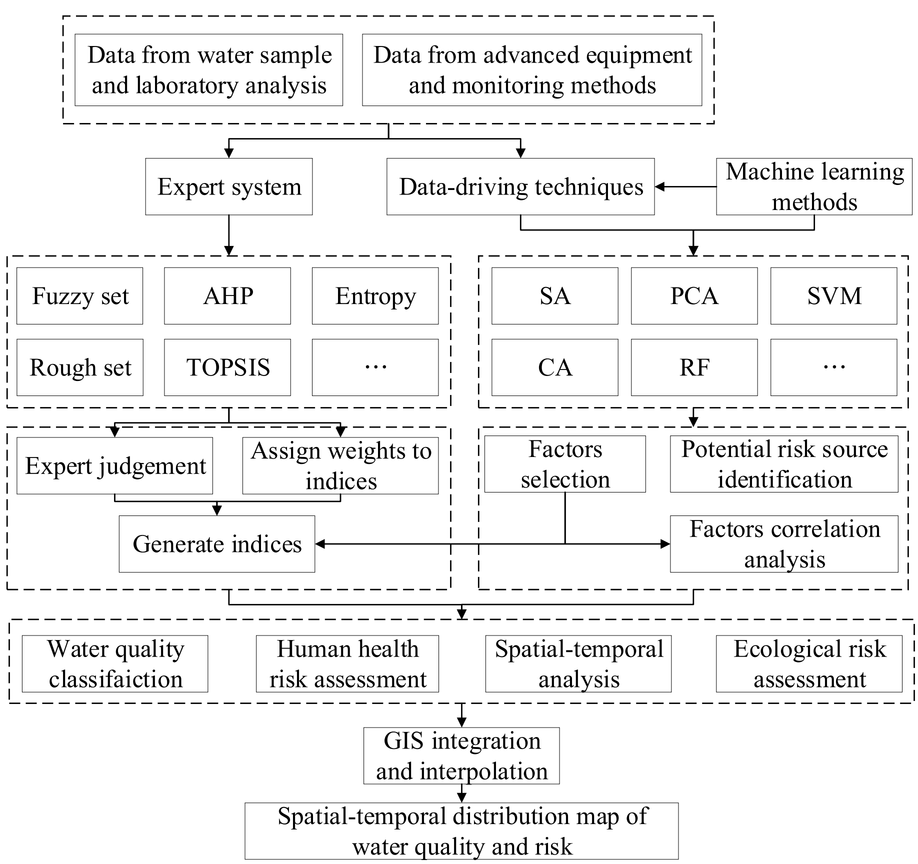

Fig. 8. Flowchart of water quality assessment based on models.

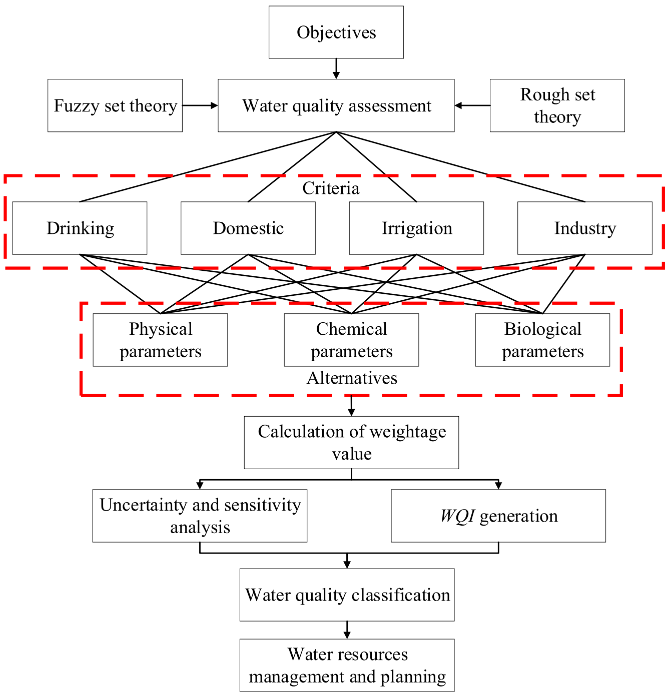

Fig. 9. Process of water quality assessment based on AHP method (recreated from Akhtar et al., 2021).

the AHP method. The relative weights of each indicator were calculated as shown in Fig. 10a and b. According to the measured WQP values from the water samples, the index scores were determined based on the water environment risk rating and classification standard. Finally, the value of the cumulative and sudden risks can be calculated using the risk values for each indicator, as shown in Fig. 10c. Engineers for water quality management may implement countermeasures when the water environment risk level exceeds the standard value.

5.2. Water quality assessment with ML methods

ML methods have been rapidly developed owing to advances in computing techniques, and can efficiently deal with complex nonlinear problems with knowledge obtained from the real world (Khan and See, 2016). ML has been applied for water quality assessments (Shah et al., 2021), and numerous ML methods, such as cluster algorithm, ANN, PCA, RF, and deep learning (DL), have been proposed to solve different problems (Abba et al., 2020; Banadkooki et al., 2020). Cluster algorithms and PCA can be used to select features and identify primary pollutants and their distributions (Das et al., 2020). ANN, RF, and DL can be applied for water status level classification and regression based on large datasets (Wong et al., 2021).

Fig. 11 presents a flowchart and perspectives on water quality assessment using ML methods. The first step in water quality assessment is WQP monitoring using various techniques. The WQPs and water status are obtained using sensors and GIS processing. Next, the monitoring data are analysed and transformed into a WQI before setting up the input and output sets. In this case, relevant data features are extracted using PCA or linear discriminant analysis (LDA) (Ratolojanahary et al., 2019). Then, the WQPs, such as COD, TP, TN, chl-a, and heavy metals, are set as the input data for the ML methods. Next, the water quality status is divided into five categories, excellent, good, moderate, poor, and bad, which are set as the classification targets. RF is the most widely used ML method for classification and regression (Norouzi and Moghaddam, 2020; Tiyasha et al., 2021). Breiman (2001) developed RF based on the bagging method. Fig. 12 shows the basic structure of a RF. The database, including the WQPs and status, is split into two sections: training and test data. The variable features are randomly selected, and original samples in the training data are drawn in a bag using the resampling bootstrap method to develop classification decision trees (CARTs) (Davison and Hinkley, 1997). The data placed into another category are called out-of-bag (OOB), which are used to assess the accuracy of the RF model (Norouzi and Moghaddam, 2020). The Gini index is then used to generate the nodes by feature determination (Lin et al., 2021). Finally, *N* CART trees are established, and the classification results are determined by voting based on the results of the CART trees. After water quality status classification, the importance of different variables for WQPs is analysed using SA. Spearman rank correlation coefficient method is commonly used in SA to reveal the relationships between heavy metals (Zhang et al., 2015). Additionally, Monte Carlo simulation is commonly used in uncertainty analysis to determine the uncertainty of the process in water quality assessment (Saha et al., 2017). Finally, the spatial-temporal distribution of water quality and contamination can be visualised on a map using GIS technology.

To illustrate the application of ML methods to water quality assessment, the prediction of WQI classification in Klang River in Malaysia and surface water quality assessment in the Tianshan Mountains in China are demonstrated as two case studies, respectively.

(1) For Klang River, three ML methods, RF, DL, and decision tree (DT), are used to predict the WQI classification. In the first step, we gathered raw data from monitoring stations. Second, we determined the input and output datasets via data processing prior to establishing the

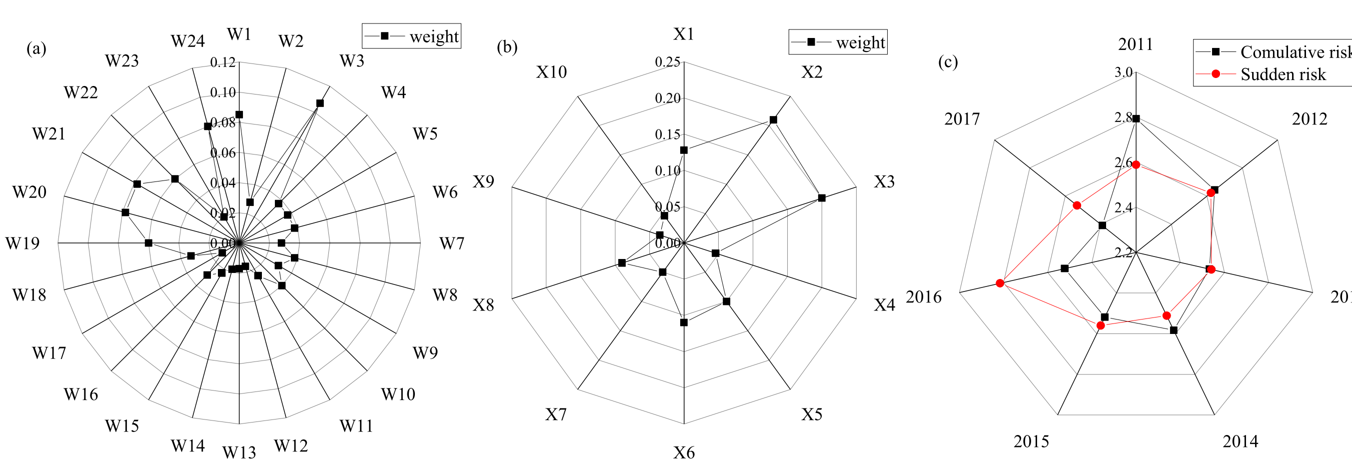

Fig. 10. Water environment risk assessment based on AHP, (a) cumulative risk weight of comprehensive indicators (W_i), (b) sudden risk weight of comprehensive indicators (X_j), and (c) risk level from 2011 to 2017 (recreated from Shi et al., 2020).

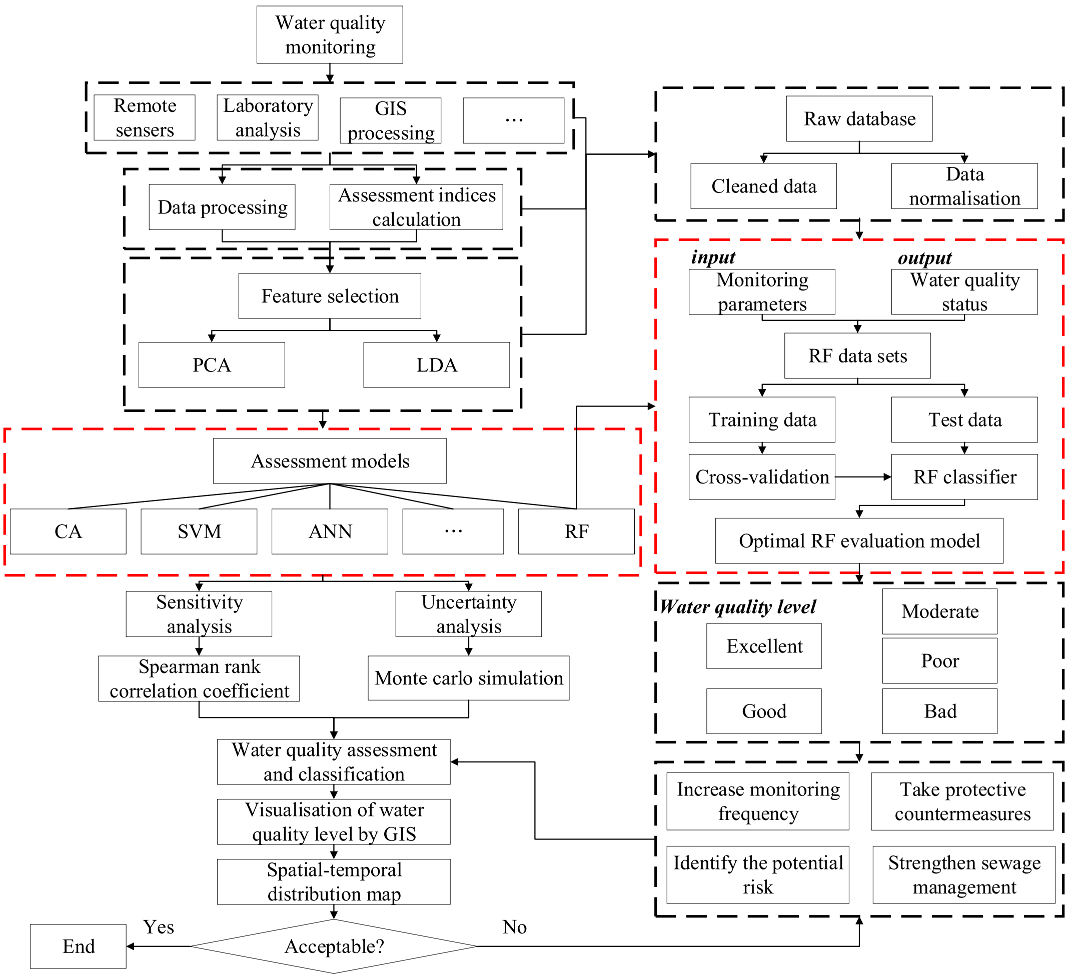

Fig. 11. Flowchart of water quality assessment based on machine learning methods.

database. In this case study, we set the collected WQPs, such as DO, pH, COD, BOD, suspended solids (SS), and NH_3–N, as the input data set. Our classification target was the water quality status, which was classified into five categories based on the DOE-WQI values in the water quality standards set by the Department of Environment of Malaysia (DOE) (DOE, 2014). Finally, we established a database based on the monitoring parameters and water quality status. The database was split into training and test sets at 80% and 20%, respectively. Then, the ML models for water quality status classification were established by training and verifying the classification models with the training and test sets. Fig. 13 presents the performance of the three classification models. The performance metrics showed that the DL model was more accurate with less classification error than the RF and DT models (Fig. 13a). The precision for each class is shown in Fig. 13b. The precision of the DL model for each class was higher than that of the RF and DT models, except for class I. In general, the DL model was the most accurate of the three models. After the WQI is predicted using ML methods, countermeasures, such as increasing monitoring frequency, identifying potential risks, and enhancing sewage treatment can be applied to improve water quality.

(2) For the Tianshan Mountains in China, SA, PCA, and CA are applied to reveal the correlations and sources of heavy metals. Table 7 shows Spearman correlation coefficients of heavy metals. The SA result showed that the groups of Cu, Zn, and Pb; Cr, Mn, Hg, Zn, and Pb; and Ni, As, Co, and Cr were positively related. Therefore, the three groups were influenced by three factors (Zhang et al., 2015). Contrarily, the negative correlation coefficients of the heavy metals indicated that these heavy metals may have different sources. Therefore, the PCA and CA methods were used to determine the origins of these heavy metals. Fig. 14 shows the results of PCA and CA for heavy metal groups. The first principal component (PC1) consisted of Cu, Zn, and Pb; the second principal component (PC2) of Co, Cr, As, and Ni; and the third principal component (PC3) consisted of Mn, Cd, and Hg (Fig. 14a) (Zhang et al., 2015). After analysing the water samples from sampling sites, heavy metals, such as Zn, Hg, and Mn, in PC1 and PC3 may have been emitted from industrial areas and human activities near townships (Kurun et al., 2010; Li et al., 2012).

The heavy metals in PC2 (Co, Cr, and As) mainly derived from rock weathering and dry river floods, and could have entered rivers and lakes (Hu et al., 2012; Li et al., 2012). The results of CA were consistent with those of PCA, where the heavy metals were categorised into three groups (Fig. 14b): Cu, Zn, and Pb; Co, Ni, Cr, and As; and Mn, Cd, and Hg. Combined with the analysis of heavy metal sources, the second group of heavy metals were mainly derived from the natural environment, and the first and third groups mainly derived from human activities.

6. Summary of and perspective on water quality management

6.1. Summary of water quality assessment methods

This study introduces several aspects of water quality assessment using different assessment methods, including assessment indices (WQIs, TSIs, and HMIs) and assessment models (ES methods and ML models). Fig. 15 shows a Venn diagram of the objectives of water quality assessment and the relationships among assessment indices (WQIs, TSIs, and HMIs) and assessment models (ES methods and ML models). Different indices have their own assessment expressions for evaluating

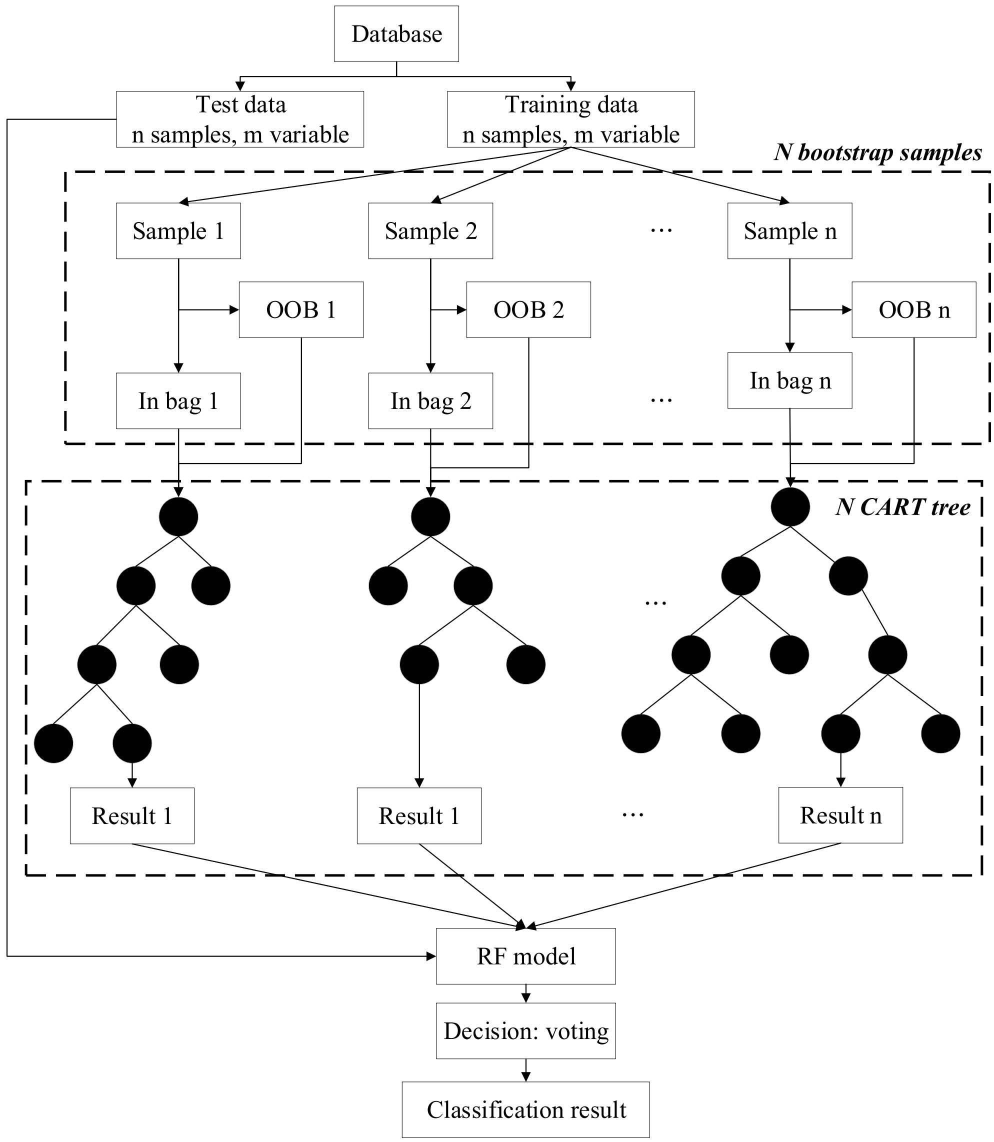

Fig. 12. Basic structure of random forest (recreated from Norouzi et al., 2020).

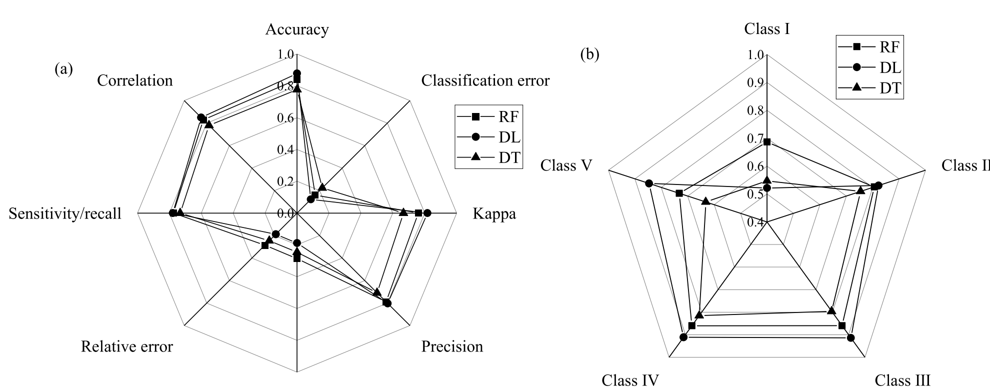

Fig. 13. Performance of each model in classifying individual *WQI* classes: (a) performance metrics; (b) precision of each class (data from Tiyasha et al., 2021).

various aspects of water quality (i.e., water quality level, eutrophic status, and human health risk). The overlapping region in Fig. 15a indicates that some indices (e.g., the single-factor index and Nemerow synthetic pollution index) can be used for all aspects of surface water quality evaluation. Additionally, assessment models have been developed by integrating assessment indices (WQIs, TSIs, and HMIs) based on the weights of WQPs to efficiently predict water status, which can be employed in various situations according to geographical location and pollution indicators.

WQIs have become a valuable tool for analysing water quality trends

Table 7
Correlation coefficients of the concentrations of heavy metals in surface water (recreated from Zhang et al., 2015).

	As	Cd	Co	Cr	Cu	Hg	Mn	Ni	Pb	Zn
As	1									
Cd	−0.322[a]	1								
Co	0.351	−0.213[a]	1							
Cr	0.475	−0.421[a]	0.561[a]	1						
Cu	0.327	0.391[a]	0.417	0.361	1					
Hg	−0.336[a]	0.688[b]	−0.283[a]	−0.342[a]	0.312	1				
Mn	−0.251[a]	0.479[b]	−0.274[a]	−0.512[a]	0.293	0.431	1			
Ni	0.528	−0.231	0.494[a]	0.995[a]	−0.252	−0.322[a]	−0.481[a]	1		
Pb	−0.261[a]	0.405[b]	−0.317[a]	−0.311	0.341[a]	0.546	0.322	−0.433	1	
Zn	−0.427[a]	0.557[b]	−0.221[a]	−0.212	0.324[a]	0.455	0.184	−0.251[a]	0.422	1

[a] Correlation significance at the 0.01 level.
[b] Correlation significance at the 0.05 level (one tail).

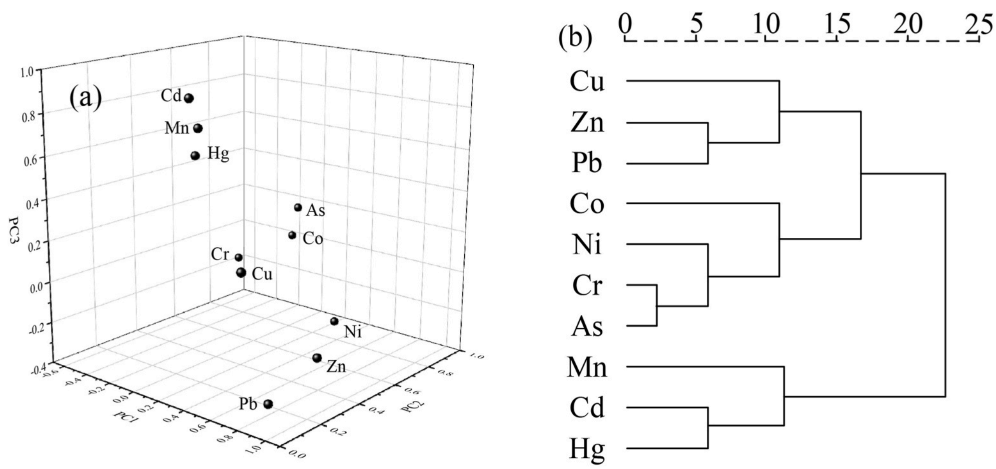

Fig. 14. Analyses results of heavy metal groups: (a) PCA; (b) CA (recreated based on Zhang et al., 2015).

and describing overall water quality (Rangeti et al., 2015). However, WQIs depict the composite influence of various WQPs. Some countries use aggregated WQPs to develop WQIs for water quality classification, because of the absence of globally acceptable WQIs (Tyagi et al., 2013). Selecting the crucial parameters for developing WQIs is challenging because of the subjective nature of the assessment of some of the parameters. Additionally, TSIs and HMIs have been developed to assess water eutrophication and contamination levels, respectively, which also aggregated WQPs for each aspect of water quality. Compared with WQIs, TSIs can be developed using more parameters (e.g., concentrations of bacteria and algae) for assessing eutrophication. Furthermore, heavy metal pollution assessments consist of not only water quality classification but also human health risk assessment and the analysis of correlations and sources of heavy metals (Fig. 15b). With advances in computer technology, AI methods have been applied to assess water quality (Wong et al., 2021). Water quality assessment models based on AI technology have been designed to generate assessment indices, identify the sources of contaminants, and classify water quality levels. WQP data have been collected by remote sensors, laboratory analysis, and GIS processing and input to assessment models to help researchers more quickly and comprehensively analyse water quality. Finally, the analysis results for the spatial-temporal distribution of water quality levels can be visualised on maps, which provide valuable information for decision-makers.

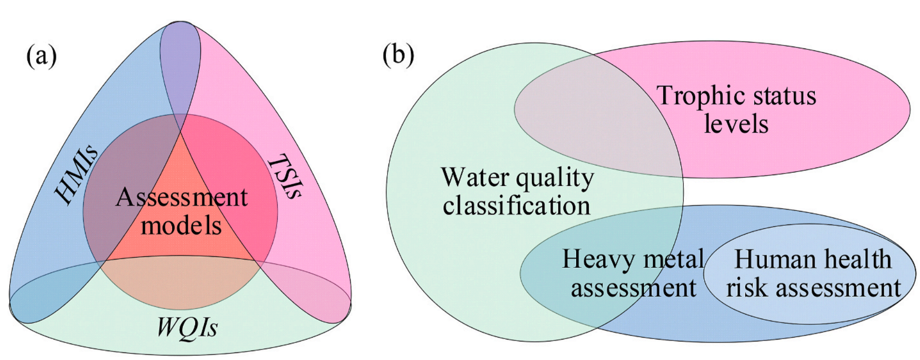

Fig. 15. Venn diagram of (a) relationship among assessment indices and assessment models; (b) objectives of water quality assessment.

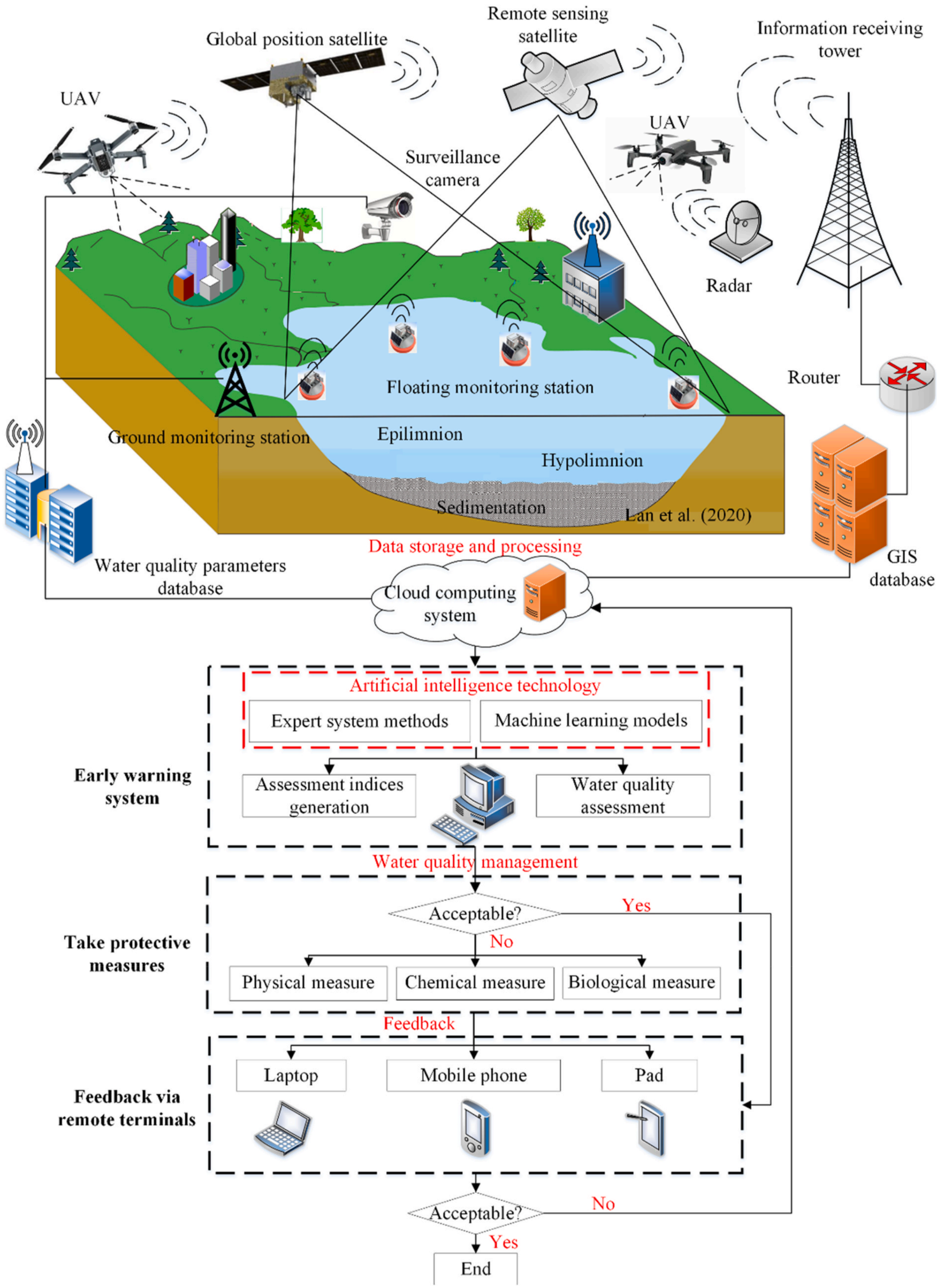

Fig. 16. Flowchart of water quality management system.

Table 8
Comparison of various techniques for water quality management.

References	Assessment method	Parameters	Objectives	Advantages	Limitations
Zhuang et al. (2022)	Intelligent algorithm	TN for agricultural runoff based on multi-parameters sensors	Real-time measurement of TN	High precision; Suitable for rainy, cloudy, or night-time conditions	Single WQP measurement
de Paul Obade and Moore (2018)	Single-value WQI	Harmful algal blooms, turbidity, and water content indices from remote sensing, GIS, and GPS	Proposed operational tools and models for monitoring water quality	Automatically monitoring WQPs using remote sensing	Not suitable for rainy, and cloudy condition
Elsayed et al. (2021)	ANNs models and spectral reflectance indices	TN, ammonium, orthophosphate, and COD obtained from ground-based remote sensing measurements	Estimating WQPs in lakes	High precision; Effectively characterize spatial-temporal variability for lake system	High total cost; Not suitable for cloudy condition
Kim et al. (2021)	Machine learning models based on adaptive sampling method	Water quality, hydrodynamic, and meteorological variables collected from monitoring stations	Early warning of harmful algal blooms	Reliability; Avoid imbalance of observed data	Back-calculation of algal blooms
This study	Comprehensive assessment indices based on actual condition of the site	Full WQPs obtained by sensors, laboratory test, GIS, field test, and UAV	Establish a water quality management platform	Suitable for all conditions; Automatically evaluate water quality and provide suggestions by IoT platform	Require large amount of investment

6.2. Current status and future perspectives for water quality management

Currently, more than half of the water resources in the world are polluted, which poses a threat to human health, the environment and climate change (Saravanan et al., 2021). Surface water quality must urgent be restored to combat its negative influence on human health, economic development, and the ecological environment (Zhang et al., 2021a). However, current water quality management methods are commonly based on integrating different WQPs into various indices to evaluate water quality, which is time-consuming, inefficient and cannot meet the demands for improving water quality in a rapidly developing society. The worth of water should be recognised, measured and integrated into the decision-making systems of water quality managers (UNESCO, 2021). In addition, pollution should be prevented instead of treating contaminated water. Effectively and efficiently identifying pollution sources and making decisions regarding water management are bottlenecks in water quality protection. Moreover, the current water management system lacks a platform that integrates real-time monitoring, identification, feedback, and decision-making for surface water based on the recognition of its worth.

China knows its lucid waters and lush mountains are invaluable assets for economic development. Many of its financial resources have been devoted to water resource protection, which requires not only the development of composite indices, but also a complete water quality management system. This system consists of identifying pollution sources, monitoring WQPs, evaluating water quality, treating water pollution, and restoring the ecological environment. The most crucial aspect of water quality management is the establishment of an early warning system that includes monitoring, evaluation, and feedback. Therefore, technical countermeasures can be implemented to prevent the deterioration of water quality. Finally, society will benefit from the increased availability of water resources, achieved through an innovative water quality management system.

With the development of computing technology, many technologies and equipment are being used to provide early water quality warning. Fig. 16 presents a flowchart of a water quality management system. Remote sensors (RS) are valuable measure tools used for obtaining WQPs. Several sensors can be fixed on a floating water quality monitoring station to obtain the WQPs, such as conductivity, T, pH, and SD. These data can then be collected in a cloud system using wireless transmission technology (Wang and Yang, 2019). Surveillance cameras and unmanned aerial vehicles (UAVs) can be used to obtain pictures of the water area. A GIS can be utilised to generate real-time water quality maps and obtain geographic positions, monitoring information, and variation in WQPs. A real-time water quality monitoring system has been built up by integrating above techniques. However, it cannot satisfy the requirements of water quality management. As such, assessment methods and decision-making systems for water quality treatment should also be integrated into water quality management systems. Then, the data collected by the above techniques can be stored in a cloud system, which can be used to develop water quality assessment models. In the previous section, we presented several ES and ML models used for water quality evaluation. Water quality information can be processed and predicted using water quality assessment models. If the water quality status is not acceptable, technical measures can be applied to improve water quality (Fig. 11). Early water quality warning systems can be integrated into IoT, which provides a powerful platform for water quality management. IoT technology can enable the dynamic monitoring and analysis of water quality and implement water treatment during water quality management. Feedback can be visualised in remote terminals, such as laptops, smartphones, and computers.

Previous studies for water quality assessment were compared with this framework to demonstrate its technical feasibility and applicability (see Table 8) (Zhuang et al., 2022; Kim et al., 2021). The implementation of the framework for water quality management can enable all-directional water quality monitoring. Based on water use, suggestions and countermeasures for different water statuses can be provided for water quality management, thereby considerably contributing to the recognition of the quality of water. With this framework, water resources can be more effectively used for human activities and to regulate climate, which will be directly valuable to society and the global environment (Elsayed et al., 2021). The water quality status in the country can be presented in an IoT platform and in remote terminals for the public providing an opportunity for society to supervise the protection of surface water quality.

7. Concluding remarks

In this study, we reviewed major surface water quality assessment indices (WQIs, TSIs, and HMIs) and assessment models (ES and ML-based assessment models). Water resources are among the most critical factors affecting human survival. To ensure the availability of water resources, we established a water quality classification model and management system to prevent and manage water pollution. The summarisation of indices and assessment models can assist researchers in developing water quality management systems.

The review of these indices showed that existing water quality assessment indices, including WQIs, TSIs, and HMIs, can depict different aspects of water quality (surface water quality classification, eutrophication levels, and human health risk). The procedure for calculating these indices by aggregating various WQPs is a time-consuming process for evaluating water quality. Thus, assessment models (ES and ML models) can be applied to efficiently generate these indices and predict water quality status based on big data obtained from advanced equipment and monitoring methods, which can guide engineers in constructing and implementing proper measures to improve water quality.

The value of water should be recognised, measured, and integrated into water quality management decision-making systems. Thus, a complete water quality management system that consists of identifying pollution sources, monitoring WQPs, evaluating water quality, treating water pollution, and restoring the ecological environment, should be developed. With the development of technology, numerous types of equipment (e.g., RS, GIS, and UAV) have been used for the real-time monitoring of WQPs in China. However, financial investment is still required to develop a water quality management platform based on AI and IoT technology to meet the future demands for water as the economy rapidly develops. Future studies will focus on integrating water treatment methods into a water quality management system to enable automated decision-making based on ESs and provide suggestions for governments. Additionally, water quality status can be presented in an IoT platform and in the remote terminals for the public, providing an opportunity for society to supervise the protection of surface water quality.

Credit author statement

Tao Yan: Investigation, Data curation, Writing-Original draft; Shui-Long Shen: Conceptualization, Methodology, Supervision, Funding acquisition; Annan Zhou: Investigation, Supervision, Reviewing and Editing;

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgment

The research work was funded by "The Pearl River Talent Recruitment Program" in 2019 (Grant No. 2019CX01G338), Guangdong Province, China and the Research Funding of Shantou University for New Faculty Member (Grant No. NTF19024-2019), China.

References

Abbasi, T., Abbasi, S.A., 2012. Chapter 1 - why water-quality indices. In: Tasneem Abbasi, S.A. Abbasi, Water Quality Indices, 2012. Elsevier, pp. 3–7. https://doi.org/10.1016/B978-0-444-54304-2.00001-4.

Abba, S.I., Hadi, S.J., Sammen, S.S., Salih, S.Q., Abdulkadir, R.A., Pham, Q.B., Yaseen, Z. M., 2020. Evolutionary computational intelligence algorithm coupled with self-tuning predictive model for water quality index determination. J. Hydrol. 587 (2020), 124974 https://doi.org/10.1016/j.jhydrol.2020.124974.

Aizaki, M., Otsuki, A., Fukushima, T., Hosomi, M., Muraoka, K., 1981. Application of Carlson's trophic state index to Japanese lakes and relationships between the index and other parameters. SIL Proceedings 21 (1), 675–681. https://doi.org/10.1080/03680770.1980.11897067, 1922-2010.

Alves, R.I., Sampaio, C.F., Nadal, M., Schuhmacher, M., Domingo, J.L., Segura-Muñoz, S. I., 2014. Metal concentrations in surface water and sediments from Pardo River, Brazil: human health risks. Environ. Res. 133, 149–155. https://doi.org/10.1016/j.envres.2014.05.012, 2014.

Alver, A., 2019. Evaluation of conventional drinking water treatment plant efficiency according to water quality index and health risk assessment. Environ. Sci. Pollut. Control Ser. 26, 27225–27238. https://doi.org/10.1007/s11356-019-05801-y.

Amiri, V., Berndtsson, R., 2020. Fluoride occurrence and human health risk from groundwater use at the west coast of Urmia Lake, Iran. Arabian J. Geosci. 13, 921. https://doi.org/10.1007/s12517-020-05905-7.

Akhtar, N., Ishak, M.I.S., Ahmad, M.I., Umar, K., Md Yusuff, M.S., Anees, M.T., Qadir, A., Ali Almanasir, Y.K., 2021. Modification of the water quality index (*WQI*) process for simple calculation using the multi-criteria decision-making (MCDM) method: a review. Water 13 (7), 905. https://doi.org/10.3390/w13070905.

An, Y., Zou, Z., Li, R., 2014. Water quality assessment in the harbin reach of the songhuajiang river (China) based on a fuzzy rough set and an attribute recognition theoretical model. Int. J. Environ. Res. Publ. Health 11 (4), 3507–3520. https://doi.org/10.3390/ijerph110403507.

Babbar, R., Babbar, S., 2017. Predicting river water quality index using data mining techniques. Environ. Earth Sci. 76, 504. https://doi.org/10.1007/s12665-017-6845-9, 2017.

Baghapour, M.A., Shooshtarian, M.R., Zarghami, M., 2020. Process mining approach of a new water quality index for long-term assessment under uncertainty using consensus-based fuzzy decision support system. Water Resour. Manag. 34, 1155–1172. https://doi.org/10.1007/s11269-020-02489-5, 2020.

Balan, I.N., Shivakumar, M., Kumar, P.D.M., 2012. An assessment of groundwater quality using water quality index in Chennai, Tamil Nadu, India. Chronicles Young Sci. 3 (2), 146–150. https://doi.org/10.4103/2229-5186.98688.

Banadkooki, F.B., Ehteram, M., Panahi, F., Sammen, S.S., Othman, F.B., EL-Shafie, A., 2020. Estimation of total dissolved solids (TDS) using new hybrid machine learning models. J. Hydrol. 587 (2020), 124989 https://doi.org/10.1016/j.jhydrol.2020.124989.

Bhateria, R., Jain, D., 2016. Water quality assessment of lake water: a review. Sustain. Water Resour. Manag. 2, 161–173. https://doi.org/10.1007/s40899-015-0014-7.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Brown, R.M., McClelland, N.I., Deininger, R.A., Tozer, R.G., 1970. Water quality index-do we dare? Water Sewage Works 117 (10), 339–343.

Carlson, R.E., 1977. A trophic state index for lakes1. Limnol. Oceanogr. 22, 361–369. https://doi.org/10.4319/lo.1977.22.2.0361.

CCME, 2001. Canadian environmental quality guidelines for the protection of aquatic life, CCME water quality index: technical report 1.0.

CEMS, 2001. Lakes (Reservoirs) Eutrophication Assessment Methods and Classification Technology Requirements. China Environmental Monitoring Station, Beijing, 2001.

Chou, J.S., Ho, C.C., Hoang, H.S., 2018. Determining quality of water in reservoir using machine learning. Ecol. Inf. 44, 57–75. https://doi.org/10.1016/j.ecoinf.2018.01.005, 2018.

de Paul Obade, V., Moore, R., 2018. Synthesizing water quality indicators from standardized geospatial information to remedy water security challenges: a review. Environ. Int. 119, 220–231. https://doi.org/10.1016/j.envint.2018.06.026, 2018.

Das, P., Naganna, S.R., Deka, P.C., Pushparaj, J., 2020. Hybrid wavelet packet machine learning approaches for drought modeling. Environ. Earth Sci. 79 (2020), 221. https://doi.org/10.1007/s12665-020-08971-y.

Darko, A., Chan, A.P.C., Adabre, M.A., Edwards, D.J., Hosseini, M.R., Ameyaw, E.E., 2020. Artificial intelligence in the AEC industry: scientometric analysis and visualization of research activities. Autom. ConStruct. 112, 103081 https://doi.org/10.1016/j.autcon.2020.103081, 2020.

Davison, A.C., Hinkley, D.V., 1997. Bootstrap Methods and Their Application (Cambridge Series in Statistical and Probabilistic Mathematics). Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511802843.

Ding, Y., Zhao, J.Y., Peng, W.Q., Zhang, J., Chen, Q.C., Fu, Y.C., Duan, M.Q., 2021. Stochastic trophic level index model: a new method for evaluating eutrophication state. J. Environ. Manag. 280 (2021), 111826 https://doi.org/10.1016/j.jenvman.2020.111826.

Doe, M., 2014. National Water Quality Standards for malaysia Parameter Source : EQR2006. National Water Quality Standards For Malaysia, pp. 1–5.

Elsayed, S., Ibrahim, H., Hussein, H., Elsherbiny, O., Elmetwalli, A.H., Moghanm, F.S., Ghoneim, A.M., Danish, S., Datta, R., Gad, M., 2021. Assessment of water quality in lake Qaroun using ground-based remote sensing data and artificial neural networks. Water 13 (2021), 3094. https://doi.org/10.3390/w13213094.

El Osta, M., Masoud, M., Alqarawy, A., Elsayed, S., Gad, M., 2022. Groundwater suitability for drinking and irrigation using water quality indices and multivariate modeling in makkah Al-Mukarramah province, Saudi Arabia. Water 14 (2022), 483. https://doi.org/10.3390/w14030483.

Gad, M., El-Hattab, M., 2019. Integration of water pollution indices and DRASTIC model for assessment of groundwater quality in El Fayoum Depression, Western Desert, Egypt. J. Afr. Earth Sci. 158, 103554 https://doi.org/10.1016/j.jafrearsci.2019.103554, 2019.

Gad, M., Abou El-Safa, M.M., Farouk, M., Hussein, H., Alnemari, A.M., Elsayed, S., Khalifa, M.M., Moghanm, F.S., Eid, E.M., Saleh, A.H., 2021. Integration of water quality indices and multivariate modeling for assessing surface water quality in Qaroun lake, Egypt. Water 13 (2021), 2258. https://doi.org/10.3390/w13162258.

Gad, M., Saleh, A.H., Hussein, H., Farouk, M., Elsayed, S., 2022. Appraisal of surface water quality of Nile river using water quality indices, spectral signature and multivariate modeling. Water 14 (2022), 1131. https://doi.org/10.3390/w14071131.

Gao, S., Wang, Z., Wu, Q., Zeng, J., 2020. Multivariate statistical evaluation of dissolved heavy metals and a water quality assessment in the Lake Aha watershed, Southwest China. PeerJ 8, e9660. https://doi.org/10.7717/peerj.9660.

García-Carmona, M., Romero-Freire, A., Sierra Aragón, M., Martínez Garzón, F.J., Martín Peinado, F.J., 2017. Evaluation of remediation techniques in soils affected by residual contamination with heavy metals and arsenic. J. Environ. Manag. 191, 228–236. https://doi.org/10.1016/j.jenvman.2016.12.041, 2017.

Hamilton, H.A., Ivanova, D., Stadler, K., Merciai, S., Schmidt, J., Zelm, R.V., Moran, D., , Wood, Trade, R., 2018. Trade and the role of non-food commodities for global eutrophication. Nat. Sustain. 1 (6), 314–321. https://doi.org/10.1038/s41893-018-0079-z, 2018.

Hu, Y., Qi, S.H., Wu, C.X., Ke, Y.P., Chen, J., Chen, W., Gong, X., 2012. Preliminary assessment of heavy metal contamination in surface water and sediments from Honghu Lake, East Central China. Front. Earth Sci. 6 (1), 39–47. https://doi.org/10.1007/s11707-012-0309-z.

Huang, Y., Chen, Q.Q., Deng, M.H., Japenga, J., Li, T.Q., Yang, X., He, Z.L., 2018. Heavy metal pollution and health risk assessment of agricultural soils in a typical peri-urban area in southeast China. J. Environ. Manag. 207, 159–168. https://doi.org/10.1016/j.jenvman.2017.10.072, 2018.

Horton, R.K., 1965. An index number system for rating water quality. J. Water Pollut. Control Fed. 37 (3), 300–306.

Hurley, T., Sadiq, R., Mazumder, A., 2012. Adaptation and evaluation of the Canadian Council of Ministers of the Environment Water Quality Index (*CCME WQI*) for use as an effective tool to characterize drinking source water quality. Water Res. 46 (11), 3544–3552. https://doi.org/10.1016/j.watres.2012.03.061.

Jabbar, F.K., Grote, K., 2019. Statistical assessment of nonpoint source pollution in agricultural watersheds in the Lower Grand River watershed, MO, USA. Environ. Sci. Pollut. Control Ser. 26, 1487–1506. https://doi.org/10.1007/s11356-018-3682-7.

Jehan, S., Ullah, I., Khan, S., Muhammad, S., Khattak, S.A., Khan, T., 2020. Evaluation of the Swat River, Northern Pakistan, water quality using multivariate statistical techniques and water quality index (*WQI*) model. Environ. Sci. Pollut. Control Ser. 27, 38545–38558. https://doi.org/10.1007/s11356-020-09688-y.

Ji, B., Liang, J., Chen, R., 2020. Bacterial eutrophic index for potential water quality evaluation of a freshwater ecosystem. Environ. Sci. Pollut. Control Ser. 27, 32449–32455. https://doi.org/10.1007/s11356-020-09585-4.

Khan, A.A., Paterson, R., Khan, H., 2003. Modification and Application of the *CCME WQI* for the Communication of Drinking Water Quality Data in Newfoundland and Labrador, Proceedings of the 38th Central Symposium on Water Quality Research. Canadian Association on Water Quality, Burlington, Canada, 2003.

Khan, Y., See, C.S., 2016. Predicting and analyzing water quality using Machine Learning: a comprehensive model. In: 2016 IEEE Long Island Systems, Applications and Technology Conference (LISAT), pp. 1–6. https://doi.org/10.1109/LISAT.2016.7494106.

Khan, R., Jhariya, D.C., 2017. Groundwater quality assessment for drinking purpose in Raipur city, Chhattisgarh using water quality index and geographic information system. J. Geol. Soc. India 90, 69–76. https://doi.org/10.1007/s12594-017-0665-0, 2017.

Kurun, A., Balkis, N., Erkan, M., Balkis, H., Aksu, A., Erşan, M.S., 2010. Total metal levels in crayfish Astacus leptodactylus (Eschscholtz, 1823), and surface sediments in Lake Terkos, Turkey. Environ. Monit. Assess. 169 (1–4), 385–395. https://doi.org/10.1007/s10661-009-1181-5.

Kelly, M.G., 1998. Use of the trophic diatom index to monitor eutrophication in rivers. Water Res. 32 (1), 236–242. https://doi.org/10.1016/S0043-1354(97)00157-7.

Kumar, D., Alappat, B.J., 2019. NSF-water quality index: does it represent the experts' opinion? Pract. Period. Hazard. Toxic, Radioact. Waste Manag. 13 (1) https://doi.org/10.1061/(ASCE)1090-025X(2009)13:1(75). January 2009.

Kim, J.H., Shin, J.K., Lee, H.K., Lee, D.H., Kang, J.H., Cho, K.H., Lee, Y.G., Chon, K.M., Baek, S.S., Park, Y.G., 2021. Improving the performance of machine learning models for early warning of harmful algal blooms using an adaptive synthetic sampling method. Water Res. 207 (2021), 117821 https://doi.org/10.1016/j.watres.2021.117821.

Lan, J.H., Wang, T.L., Chawchai, S., Cheng, P., Zhou, K., Yu, K.K., Yan, D.N., Wang, Y.Q., Zang, J.J., Liu, Y.J., Tan, L.C., Ai, L., Xu, H., 2020. Time marker of 137Cs fallout maximum in lake sediments of Northwest China. Quat. Sci. Rev. 241 (2020), 106413 https://doi.org/10.1016/j.quascirev.2020.106413.

Le Moal, M., Gascuel-Odoux, C., Menesguen, A., Souchon, Y., Etrillard, C., Levain, A., Moatar, F., Pannard, A., Souchu, P., Lefebvre, A., Pinay, G., 2019. Eutrophication: a new wine in an old bottle? Sci. Total Environ. 651 (1), 1–11. https://doi.org/10.1016/j.scitotenv.2018.09.139.

Li, C., Cao, J., Yao, L., Wu, Q.Y., Lv, J.S., 2020. Pollution status and ecological risk of heavy metals in the soils of five land-use types in a typical sewage irrigation area, eastern China. Environ. Monit. Assess. 192, 471. https://doi.org/10.1007/s10661-020-08384-1.

Li, G., Liu, G., Zhou, C., Chou, C.L., Zheng, L., Wang, J., 2012. Spatial distribution and multiple sources of heavy metals in the water of Chaohu Lake, Anhui, China. Environ. Monit. Assess. 184 (5), 2763–2773. https://doi.org/10.1007/s10661-011-2149-9.

Li, S., Zhang, Q., 2010. Risk assessment and seasonal variations of dissolved trace elements and heavy metals in the Upper Han River, China. J. Hazard Mater. 181, 1051–1058. https://doi.org/10.1016/j.jhazmat.2010.05.120, 2010.

Li, Z.L., Zhang, C., Liu, H.X., Zhang, C., Zhao, M.K., Gong, Q., Fu, G.T., 2022. Developing stacking ensemble models for multivariate contamination detection in water distribution systems. Sci. Total Environ. 828 (2022), 154284 https://doi.org/10.1016/j.scitotenv.2022.154284.

Lin, S.S., Shen, S.L., Zhou, A., Lyu, H.M., 2020a. Sustainable development and environmental restoration in Lake Erhai, China. J. Clean. Prod. 258 (2020), 120758 https://doi.org/10.1016/j.jclepro.2020.120758.

Lin, S.S., Shen, S.L., Zhou, A., Xu, Y.S., 2020b. Approach based on TOPSIS and Monte Carlo simulation methods to evaluate lake eutrophication levels. Water Res. 187 (2020), 116437 https://doi.org/10.1016/j.watres.2020.116437.

Lin, S.S., Shen, S.L., Lyu, H.M., Zhou, A., 2021. Assessment and management of lake eutrophication: a case study in Lake Erhai, China. Sci. Total Environ. 751 (2021), 141618 https://doi.org/10.1016/j.scitotenv.2020.141618.

Liu, D., Du, Y.X., Yu, S.J., Luo, J.H., Duan, H.T., 2020. Human activities determine quantity and composition of dissolved organic matter in lakes along the Yangtze River. Water Res. 168, 115132 https://doi.org/10.1016/j.watres.2019.115132, 2020.

Lumb, A., Halliwell, D., Sharma, T., 2006. Application of CCME water quality index to monitor water quality: a case of the Mackenzie river basin, Canada. Environmental. Monitoring and Assessment 113 (1–3), 411–429. https://doi.org/10.1007/s10661-005-9092-6.

Lyu, H.M., Shen, S.L., Zhou, A., 2021. The development of IFN-SPA: a new risk assessment method of urban water quality and its application in Shanghai. J. Clean. Prod. 282 (2021), 124542 https://doi.org/10.1016/j.jclepro.2020.124542.

Ma, Z., Song, X.F., Wan, R., Gao, L., 2013. A modified water quality index for intensive shrimp ponds of Litopenaeus vannamei. Ecol. Indicat. 24, 287–293. https://doi.org/10.1016/j.ecolind.2012.06.024, 2013.

Marín Celestino, A.E., Martínez Cruz, D.A., Otazo Sánchez, E.M., Gavi Reyes, F., Vásquez Soto, D., 2018. Groundwater quality assessment: an improved approach to K-means clustering, principal component analysis and spatial analysis: a case study. Water 10 (4), 437. https://doi.org/10.3390/w10040437.

MEE, 2002. Environmental Quality Standards for Surface Water. Ministry of Ecology and Environment of the People's Republic of China. June 1, 2002, URL: https://www.mee.gov.cn/ywgz/fgbz/bz/bzwb/shjbh/shjzlbz/200206/t20020601_66497.shtml.

MEE, 2011. Measures for Surface Water Environmental Quality Assessment. Ministry of Ecology and Environment of the People's Republic of China. March 9, 2011, URL: https://www.mee.gov.cn/gkml/hbb/bgt/201104/t20110401_208364.htm.

MEE, 2012. Algal Bloom Mornitoring by Remote Sensing Satellite. Ministry of Ecology and Environment of the People's Republic of China. October 10, 2012, URL: https://www.mee.gov.cn/home/ztbd/rdzl/wxyg/ygyy/201210/t20121010_237320.shtml.

MEE, 2014. Technical Guidelines for Risk Assessment of Contaminated Sites. Ministry of Ecology and Environment of the People's Republic of China. July 1, 2014, URL: https://www.mee.gov.cn/ywgz/fgbz/bz/bzwb/trhj/201912/t20191224_749893.shtml.

Morales-Ojeda, S.M., Herrera-Silveira, J.A., Montero, J., 2010. Terrestrial and oceanic influence on spatial hydrochemistry and trophic status in subtropical marine near-shore waters. Water Res. 44 (20), 5949–5964. https://doi.org/10.1016/j.watres.2010.07.046, 2010.

Nong, X.Z., Shao, D.G., Zhong, H., Liang, J.K., 2020. Evaluation of water quality in the South-to-North Water Diversion Project of China using the water quality index (*WQI*) method. Water Res. 178, 115781 https://doi.org/10.1016/j.watres.2020.115781, 2020.

Norouzi, H., Moghaddam, A.A., 2020. Groundwater quality assessment using random forest method based on groundwater quality indices (case study: Miandoab plain aquifer, NW of Iran). Arabian J. Geosci. 13, 912. https://doi.org/10.1007/s12517-020-05904-8, 2020.

Noori, R., Berndtsson, R., Hosseinzadeh, M., Adamowski, J.F., Abyaneh, M.R., 2019. A critical review on the application of the National sanitation Foundation water quality index. Environ. Pollut. 244, 575–587. https://doi.org/10.1016/j.envpol.2018.10.076, 2019.

Opiyo, S., Getabu, A.M., Sitoki, L.M., Shitandi, A., Ogendi, G.M., 2019. Application of the Carlson's trophic state index for the assessment of trophic status of lake Simbi ecosystem, a deep Alkaline-Saline lake in Kenya (August 11, 2019). Int. J. Fish. Aquatic Stud. 7 (4), 327–333. https://doi.org/10.2139/ssrn.3451145. https://ssrn.com/abstract=3451145. Available at: SSRN:

Pan, Y., Zhang, L.M., 2021. Roles of artificial intelligence in construction engineering and management: a critical review and future trends. Autom. ConStruct. 122 (2021), 103517 https://doi.org/10.1016/j.autcon.2020.103517.

Pennino, M.J., Leibowitz, S.G., Compton, J.E., Hill, R.A., Sabo, R.D., 2020. Patterns and predictions of drinking water nitrate violations across the conterminous United States. Sci. Total Environ. 722, 137661 https://doi.org/10.1016/j.scitotenv.2020.137661, 2020.

Pesce, S.F., Wunderlin, D.A., 2000. Wunderlin, Use of water quality indices to verify the impact of Córdoba City (Argentina) on Suquía River. Water Res. 34 (11), 2915–2926. https://doi.org/10.1016/S0043-1354(00)00036-1, 2000.

Qin, B.Q., Zhou, J., Elser, J.J., Gardner, W.S., Deng, J.M., Brookes, J.D., 2020. Water depth underpins the relative roles and fates of Nitrogen and Phosphorus in lakes. Environ. Sci. Technol. 54 (6), 3191–3198. https://doi.org/10.1021/acs.est.9b05858.

Rangeti, I., Dzwairo, B., Barratt, G.J., Otieno, F.A.O., 2015. Ecosystem-specific water quality indices. Afr. J. Aquat. Sci. 40 (3), 227–234. https://doi.org/10.2989/16085914.2015.1054341.

Ratolojanahary, R., Houé Ngouna, R., Medjaher, K., Dauriac, F., Sebilo, M., 2019. Groundwater quality assessment combining supervised and unsupervised methods. IFAC-PapersOnLine 52 (10), 340–345. https://doi.org/10.1016/j.ifacol.2019.10.054.

Raju, K.R.S.R., Varma, G.H.K., 2017. Knowledge Based Real Time Monitoring System for Aquaculture Using IoT, 2017 IEEE 7th International Advance Computing Conference. IACC), pp. 318–321. https://doi.org/10.1109/IACC.2017.0075, 2017.

Rampley, C.P.N., Whitehead, P.G., Softley, L., Hossain, M.A., Jin, L., David, J., Shawal, S., Das, P., Thompson, I.P., Huang, W.E., Peters, R., Holdship, P., Hope, R., Alabaster, G., 2020. River toxicity assessment using molecular biosensors: heavy metal contamination in the Turag-Balu-Buriganga river systems, Dhaka, Bangladesh. Sci. Total Environ. 703, 134760 https://doi.org/10.1016/j.scitotenv.2019.134760, 2020.

Renouf, M.A., Neumann, S.S., Kenway, S.J., Morgan, E.A., Choy, D.L., 2017. Urban water metabolism indicators derived from a water mass balance – bridging the gap between visions and performance assessment of urban water resource management. Water Res. 122, 669–677. https://doi.org/10.1016/j.watres.2017.05.060, 2017.

Saaty, T.L., 1977. A scaling method for priorities in hierarchical structures. J. Math. Psychol. 15 (3), 234–281. https://doi.org/10.1016/0022-2496(77)90033-5.s.

Saha, N., Rahman, M.S., Jolly, Y.N., Rahman, A., Abdus Sattar, M., Abdul Hai, M., 2016. Spatial distribution and contamination assessment of six heavy metals in soils and their transfer into mature tobacco plants in Kushtia District, Bangladesh. Environ.

Sci. Pollut. Control Ser. 23, 3414–3426. https://doi.org/10.1007/s11356-015-5575-3, 2016.
Saha, N., Rahman, M.S., Ahmed, M.B., Zhou, J.L., Ngo, H.H., Guo, W., 2017. Industrial metal pollution in water and probabilistic assessment of human health risk. J. Environ. Manag. 185, 70–78. https://doi.org/10.1016/j.jenvman.2016.10.023, 2017.
Salerno, F., Gaetano, V., Gianni, T., 2018. Urbanization and climate change impacts on surface water quality: enhancing the resilience by reducing impervious surfaces. Water Res. 144, 491–502. https://doi.org/10.1016/j.watres.2018.07.058, 2018.
Saravanan, A., Senthil Kumar, P., Jeevanantham, S., Karishma, S., Tajsabreen, B., Yaashikaa, P.R., Reshma, B., 2021. Effective water/wastewater treatment methodologies for toxic pollutants removal: processes and applications towards sustainable development. Chemosphere 280 (2021), 130595. https://doi.org/10.1016/j.chemosphere.2021.130595.
Shah, M.I., Javed, M.F., Alqahtani, A., Aldrees, A., 2021. Environmental assessment based surface water quality prediction using hyper-parameter optimized machine learning models based on consistent big data. Process Saf. Environ. Protect. 151 (2021), 324–340. https://doi.org/10.1016/j.psep.2021.05.026.
Shi, E., Shang, Y.C., Li, Y.F., Zhang, M., 2020. Evaluation of Liao river estuary with cumulative and sudden risk assessment methods. IOP Conf. Ser. Earth Environ. Sci. 612 (2020), 012064 https://doi.org/10.1088/1755-1315/612/1/012064.
Sutadian, A.D., Muttil, N., Yilmaz, A.G., Perera, B.J.C., 2016. Development of river water quality indices—a review. Environ. Monit. Assess. 188, 58. https://doi.org/10.1007/s10661-015-5050-0, 2016.
Tan, H.C., 2017. IOP Conf. Ser. Mater. Sci. Eng. 242, 012111 https://doi.org/10.1088/1757-899X/242/1/012111, 2017.
Tiyasha, Tung, T.M., Yaseen, Z.M., 2021. Deep learning for prediction of water quality index classification: tropical catchment environmental assessment. Nat. Resour. Res. https://doi.org/10.1007/s11053-021-09922-5.
Tyagi, S., Sharma, B., Singh, P., Dobhal, R., 2013. Water quality assessment in terms of water quality index. Am. J. Water Resour. 1 (3), 34–38. https://doi.org/10.12691/ajwr-1-3-3.
UNESCO, 2021. The United Nations World Water Development Report 2021: Valuing Water, Published in 2021 by the United Nations Educational, Scientific and Cultural Organization (UNESCO), 7, place de Fontenoy, 75352 Paris 07 SP. France.
Ustaoğlu, F., Taş, B., Tepe, Y., Topaldemir, H., 2021. Comprehensive Assessment of Water Quality and Associated Health Risk by Using Physicochemical Quality Indices and Multivariate Analysis in Terme River, Turkey. Environmental Science and Pollution Research. https://doi.org/10.1007/s11356-021-15135-3. Published online.
USEPA, 1989. Risk-assessment guidance for superfund. In: Human Health Evaluation Manual. Part A. Interim Report (Final), vol. 1. US Environmental Protection Agency, Washington, DC, 1989.
USEPA, 2004. Risk Assessment Guidance for Superfund Volume I: Human Health Evaluation Manual (Part E, Supplemental Guidance for Dermal Risk Assessment). US Environment Protection Agency, Washington DC, 2004.
Vollenweider, R.A., Giovanardi, F., Montanari, G., Rinaldi, A., 1998. Characterization of the trophic conditions of marine coastal waters with special reference to the NW Adriatic Sea: proposal for a trophic scale, turbidity and generalized water quality index. Environmetrics 9, 329–357. https://doi.org/10.1002/(SICI)1099-095X(199805/06)9:3<329::AID-ENV308>3.0.CO;2-9.
Wang, X.Y., Yang, W., 2019. Water quality monitoring and evaluation using remote sensing techniques in China: a systematic review. Ecosys. Health Sustain. 5 (1), 47–56. https://doi.org/10.1080/20964129.2019.1571443.
Wang, J.L., Fu, Z.S., Qiao, H.X., Liu, F.X., 2019. Assessment of eutrophication and water quality in the estuarine area of Lake Wuli, Lake Taihu, China. Sci. Total Environ. 650 (1), 1392–1402. https://doi.org/10.1016/j.scitotenv.2018.09.137.
Wen, Z.D., Song, K.S., Liu, G., Shang, Y.X., Fang, C., Du, J., Lyu, L.L., 2019. Quantifying the trophic status of lakes using total light absorption of optically active components. Environ. Pollut. 245, 684–693. https://doi.org/10.1016/j.envpol.2018.11.058, 2019.
Wong, Y.J., Shimizu, Y., Kamiya, A., Maneechot, L., Bharambe, K.P., Fong, C.S., Sulaiman, N.M.N., 2021. Application of artificial intelligence methods for monsoonal river classification in Selangor river basin, Malaysia. Environ. Monit. Assess. 193, 438. https://doi.org/10.1007/s10661-021-09202-y, 2021.
Xu, G., Xie, J., Zhang, Y., Zhao, C., Wu, Q., 2010. Application of Nemerow Pollution Index in Landscape River Water Quality Assessment of Tianjin, 2010 4th International Conference on Bioinformatics and Biomedical Engineering, pp. 1–4. https://doi.org/10.1109/ICBBE.2010.5514830, 2010.
Xiong, J.Q., Wang, X.C., Zhang, Q.Q., Duan, R., Wang, N., 2016. Characteristics of a landscape water with high salinity in a coastal city of China and measures for eutrophication control. Ecol. Indicat. 61 (2), 268–273. https://doi.org/10.1016/j.ecolind.2015.09.026.
Yu, F.C., Fang, G.H., Ru, X.W., 2010. Eutrophication, health risk assessment and spatial analysis of water quality in Gucheng Lake, China. Environ. Earth Sci. 59, 1741–1748. https://doi.org/10.1007/s12665-009-0156-8.
Yu, J.W., Kim, J.S., Li, X., Jong, Y.C., Kim, K.H., Ryang, G., 2022. Water quality forecasting based on data decomposition, fuzzy clustering and deep learning neural network. Environ. Pollut. 303 (2022), 119136 https://doi.org/10.1016/j.envpol.2022.119136.
Zeng, J., Han, G., Wu, Q., Tang, Y., 2019. Geochemical characteristics of dissolved heavy metals in Zhujiang River, Southwest China: spatial-temporal distribution, source, export flux estimation, and a water quality assessment. PeerJ 7, e6578. https://doi.org/10.7717/peerj.6578.
Zhao, Y.P., Wu, R., Cui, J.L., Gan, S.C., Pan, J.C., Guo, P.R., 2020. Improvement of water quality in the Pearl River Estuary, China: a long-term (2008–2017) case study of temporal-spatial variation, source identification and ecological risk of heavy metals in surface water of Guangzhou. Environ. Sci. Pollut. Control Ser. 27, 21084–21097. https://doi.org/10.1007/s11356-020-08378-z, 2020.
Zhang, X., Zhang, Y., Shi, P., Bi, Z.L., Shan, Z.X., Ren, L.J., 2021a. The deep challenge of nitrate pollution in river water of China. Sci. Total Environ. 770 (2021), 144674 https://doi.org/10.1016/j.scitotenv.2020.144674.
Zhang, Z.Y., Abuduwaili, J., Jiang, F.Q., 2015. Heavy metal contamination, sources, and pollution assessment of surface water in the Tianshan Mountains of China. Environ. Monit. Assess. 187, 33. https://doi.org/10.1007/s10661-014-4191-x, 2015.
Zhang, Y., Li, M., Dong, J., Yang, H., Van Zwieten, L., Lu, H., Alshameri, A., Zhan, Z., Chen, X., Jiang, X., Xu, W., Bao, Y., Wang, H., 2021b. A critical review of methods for analyzing freshwater eutrophication. Water 13, 225. https://doi.org/10.3390/w13020225.
Zhou, Y.W., Song, K., Han, R.M., Riya, S.H., Xu, X.G., Yeerken, S., Geng, S.X., Ma, Y., Terada, A., 2020. Nonlinear response of methane release to increased trophic state levels coupled with microbial processes in shallow lakes. Part B Environ. Pollut. 265, 114919. https://doi.org/10.1016/j.envpol.2020.114919, 2020.
Zhuang, Y., Wen, W., Ruan, S., Zhuang, F., Xia, B., Li, S., Liu, H., Du, Y., Zhang, L., 2022. Real-time measurement of total nitrogen for agricultural runoff based on multiparameter sensors and intelligent algorithms. Water Res. 210 (2022), 117992 https://doi.org/10.1016/j.watres.2021.117992.s.