Science of the Total Environment 777 (2021) 146113

Contents lists available at ScienceDirect

Science of the Total Environment

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

Optimal sampling strategy of water quality monitoring at high dynamic lakes: A remote sensing and spatial simulated annealing integrated approach

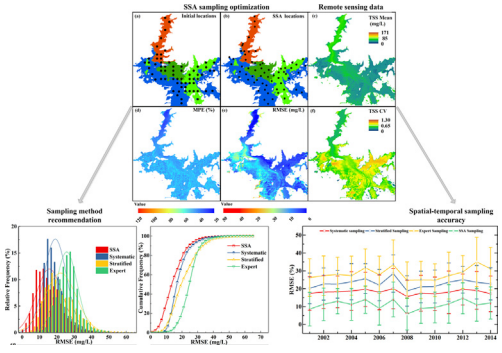
Jian Li [a], Liqiao Tian [b], Yihong Wang [c], Shuanggen Jin [a,e,*], Tingting Li [b], Xuejiao Hou [d]

[a] School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China
[b] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
[c] Jiangsu Hydraulic Research Institute, Nanjing 210029, China
[d] School of Environmental Science and Engineering, Southern University of Science and Technology of China, Shenzhen 5180055, China
[e] Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China

HIGHLIGHTS

- Remote sensing and spatial sampling annealing integrated approach is developed for sampling optimization
- The annual total suspended sediment (TSS) estimation errors dropped about 50% compared to traditional methods
- Sampling accuracy was affected by the sampling strategy and water quality variations
- Remote sensing provides valuable prior-knowledge for efficient sampling design.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 7 November 2020
Received in revised form 6 February 2021
Accepted 22 February 2021
Available online 1 March 2021

Edited by: Martin Drews

Keywords:
Water quality
Sampling method
Spatial sampling annealing
Remote sensing
Poyang Lake

ABSTRACT

An efficient and precise spatial sampling design is critical to capture spatial and temporal water quality variations under cost and labor constraints. Therefore, it is practically essential to optimize the sampling locations using limited sampling numbers to obtain the most comprehensive water quality monitoring results considering both the spatial and temporal dynamics. Existing sampling methods were restricted due to lacking pre-information and specific sampling objective function. This paper proposed an optimal sampling strategy using remote sensing (RS) big data and spatial sampling annealing (SSA) integrated approach for sampling design. The proposed method involved spatial-temporal clustering of the total suspended sediment (TSS) using long-term remote sensing data (Terra/Aqua MODIS, 2000–2014), determining the required sampling numbers using geostatistical analysis, and SSA simulation following the objective function of minimization of the spatial-temporal mean estimation error using remote sensing data as references. Taking total suspended sediment (TSS) observations at Poyang Lake, China, as the case study and application region. Results showed that the RS + SSA sampling approach is superior to conventional sampling methods such as systematic, stratified, and expert sampling, concerning spatial and temporal sampling accuracy. TSS estimation errors of the whole lake were reduced by 18.11% and 29.34% on average when compared to systematic and stratified sampling under the same sample size. The annual TSS estimation errors were dropped by approximately 50%. The sampling accuracy was affected by the synthetic effects of sampling strategy (station numbers and spatial distributions) and water quality variations (coefficient of variation, CV). Sampling optimization is more efficient to improve the sampling accuracy than increasing sampling size, which requires more cost and human resources. Remote sensing showed great potential as ideal

* Corresponding author.
E-mail addresses: sgjin@shao.ac.cn, sg.jin@yahoo.com, sgjin@nuist.edu.cn (S. Jin).

https://doi.org/10.1016/j.scitotenv.2021.146113

means to provide spatially contiguous and comprehensive data as prior-knowledge for efficient sampling design. This paper provides solutions and recommendations for evaluating existing monitoring stations in their representation of water quality or optimizing a new sampling network for future implications of more efficient and precise water quality sampling and routine monitoring.

1. Introduction

Lakes, estuaries, and reservoirs worldwide have experienced dramatic changes under pressure from climate change and anthropic activities over the past several decades (UNEP, 2016). A global exacerbation of bloom conditions was observed with increasing summertime bloom intensity in 68% of the lakes worldwide (Ho et al., 2019). The U.S. Environmental Protection Agency (EPA) reported that only 28% of U.S. rivers and streams are in good biological condition, while 46% are low (USEPA, 2017). To maintain the sustainability of the aquatic functions to food, security, and society (Read et al., 2017), a set of regulations and policies on water quality protections have been implemented at regional, national, and even global scales. For instance, a Water Framework Directive (WFD) was introduced for all European Union (EU) member states countries since 2000, aiming to protect the water quality and ecosystem status in the inland and coastal waters (Destouni et al., 2017). To address the requirement for standardized water quality monitoring, a Water Quality Portal (WQP) has been developed and established, under the joint effort of the U.S. Geological Survey (USGS), the U.S. EPA, and the National Water Quality Monitoring Council. China has placed the China 13th (2016–2020) and 14th (2021–2025) National Five-Year Science and Technology Development Plan for Environmental Protection has focused on prevention, treatment, and monitoring of water quality nationwide, which promote the construction and implementation of the "space-and-ground" integrated monitoring network for pollution monitoring and emergency response to environmental assessment (Fujiang, 2017). These actions are related to achieving the Sustainable Development Goals (SDGs) and the 2030 Agenda for Sustainable Development to improve water quality by 2030 (Alcamo, 2019; Hering, 2017).

Implementation of these policies requires continuous and consistent water quality monitoring at lakes, regional, national, and even global scales (UNEP, 2016). A proper monitoring approach should provide a systematic assessment of the water quality's evolution to support water resources assessment and management (Chen et al., 2012). Varied water quality sampling approaches and techniques have been adopted, including traditional field station sampling methods (Abd-Elrahman et al., 2011), automatic network (Pule et al., 2017), and remote sensing (IOCCG, 2000), or some new optical tools proposed for the real-time diagnosis of water quality without traditional sampling and laboratory physico-chemical analysis (Krapivin et al., 2017; Varotsos and Krapivin, 2018; Varotsos et al., 2019; Varotsos et al., 2020), each with its pros and cons. In-situ sampling is still one of the most common methods adopted by policy-makers, water conservation organizations, research institutes, and the general public for water quality monitoring. The three most widely used sampling strategies including random, systematic, and stratified sampling (Brus and Knotters, 2008). However, the field sampling method requires field water sample collection and laboratory analysis to determine water quality parameters, which are labor and cost-intensive, time-consuming, and limited spatial coverage (Koparan et al., 2018). Random sampling is the most straightforward strategy, using a random number generator to determine the sample locations (Tammi et al., 1999). However, the sampling uncertainty is generally more considerable than other methods for the same sampling numbers unless the study region was sampled exhaustively (Croft and Chow-Fraser, 2009). The systematic sampling strategy uses a regular grid to select the sample locations at a fixed, periodic interval (Agency, 2002). In contrast, stratified sampling was designed to use prior information, such as water quality classification, to divide the whole region into relatively homogeneous clusters and ensure that the sample is distributed across significant gradients (Catherine et al., 2008). However, water quality is typically spatial heterogeneous and not uniform across space, the optimal sampling design to be more efficient and effective is required to satisfy monitoring objectives.

Remote sensing technology is superior to traditional field sampling regarding broader spatial and temporal coverage at a low cost (Liu et al., 2003). With more synoptic and frequent observations, remote sensing sensors like Terra/Aqua MODIS, Landsat TM/ETM+/OLI, etc. could significantly improve comprehensive water quality assessments more effectively and efficiently (IOCCG, 2013; Palmer et al., 2015). These data have been commonly adopted to complement field sampling to expand water quality monitoring's spatial and temporal coverage. Such abundant information from remote sensing big data is essential to meet many of the challenges of information asymmetry, and data gaps from field sampling, especially in developing countries but is recommended to integrate with traditional sampling methods and field surveying to improve its precision (Arabi et al., 2020). Under ideal conditions, more sampling stations are better for higher quality monitoring data. However, it is essential to determine sampling sizes and locations accurately and efficiently to monitor the water quality variations (Alilou et al., 2018), which desires the optimal sampling design of water quality monitoring. Taking advantages of the remote sensing data, some pioneer studies have made some efforts on sampling network optimizing based on remote sensing imagery (Karabork, 2009; Kiefer et al., 2015). However, research on methods and applications of remote sensing big data-based sampling optimization is still less. Existing studies mainly focused on the statistical indicators (such as mean, the variance of the water quality index) obtained from single or limited images. Simultaneously, the spatial-temporal variations were not considered comprehensive information and prior-knowledge to design or optimize a sampling network.

Efforts have been made to identify representative sampling stations, reduce sampling numbers, or optimize sampling distributions using statistical methods or models. Most research on water quality sampling focused on reducing sampling numbers of the existing sampling network using statistical spatial assessment methods, such as cluster analysis (Wang et al., 2014). Besides, some optimal sampling approaches have been proposed to optimize spatial distributions of water quality sampling stations, including principal component analysis (Ouyang, 2005), matter element analysis (Wu et al., 2010), fuzzy clustering (Karamouz et al., 2009a), entropy analyses (Karamouz et al., 2009b), and genetic algorithms (Park et al., 2006). However, these sampling methods are based on sufficient data from an existing sampling network to obtain statistically significant results. Such prior information could be missing if sampling networks are not properly located and water quality data not sufficiently collected (Karabork, 2009). The ideal approach for sampling optimization requires high spatial-temporal resolution water quality data under different hydrological conditions (Bendoricchio and De Boni, 2005), which are often limited due to economic and labor constraints (Letcher et al., 2002).

This study proposed an optimal sampling strategy based on long-term spatial-temporal remote sensing data to resolve the issues above for water quality monitoring by taking total suspended sediment (TSS) observations at the Poyang Lake, China, as a region of interest.

The optimal sampling strategy involved the following questions to be addressed: (Abd-Elrahman et al., 2011) How to describe the spatial-temporal variations of water quality index quantitively, such as TSS, using long-term remote sensing data (Terra/Aqua MODIS, 2000–2014)? (Agency, 2002) How to determine the required number of samplings objectively; and (Alcamo, 2019) How to optimize the sampling distributions efficiently for water quality monitoring. Besides, traditional sampling methods' precision and efficiency, including systematic, stratified, and expert sampling, will be assessed as well as the effect factors of water quality on the sampling accuracy. This study aims to provide a novel approach and recommendations for future implications of water quality sampling, monitoring, and assessment.

2. Materials and datasets

2.1. Poyang Lake

Located in the middle of the Yangtze River basin, the maximum area of the Poyang Lake could reach up to 4000 km^2 during the wet season (28°22′–29°45′N, 115°47′–116°45′E, Fig. 1). Listed as the National Nature Reserve and a Ramsar site (www.ramsar.org/), Poyang Lake and adjacent wetland are of great importance to local and global ecological conservations (Dronova et al., 2012). The lake was selected as the case study region for its dramatic variations in suspended sediments induced by anthropogenic activities, including impacts of the Three Gorges Dam and sand mining, and thus needs effective monitoring and water quality regulation (Feng et al., 2013) (see Fig. 2).

Field survey-based routine monitoring of water quality at Poyang lake was conducted by the Water Resources Department of Jiangxi Province monthly. In general, only 19 sampling stations are set around the lake (Guo and Wang, 2014), where data are collected and then interpolated to represent the water quality of the whole lake. Extensive research primarily focused on remote sensing of water quality monitoring using Terra/Aqua MODIS (Feng et al., 2012), Landsat TM/ETM+/OLI (Wu et al., 2013), or HJ-1 CCD, GF-1 (Li et al., 2015) at different scales. The lake's long-term variations have been well documented (Feng et al., 2012; Wu and Cui, 2008), while how the sampling should be designed and optimized are still blank.

2.2. Remote sensing dataset and processing

The long-term remote sensing Terra/Aqua MODIS (2000–2014) derived TSS products were used in this study to analyze the spatial-temporal characteristic of TSS. The detailed information about the TSS data are provided in our previous study (Hou et al., 2017) and briefly summarized here: First, the MOD09Q1 and MYD09Q1 with spatial resolutions of 250 m and MOD09A1 and MYD09A1 with quality control flags of 500 m were obtained from the NASA Land Processes Distribution Active Archive Center (https://ladsweb.nascom.nasa.gov/), which have been atmospherically corrected, and a total of 343 cloud-free images were available for the Poyang Lake. The sharpen method was used to re-sample 500 m data to 250 m for comparison and elimination of bad quality pixels (Pohl and Van Genderen, 1998). Then, the land adjacency effects on the water pixels were removed by calculating relative differences between adjacent pixels using MODIS surface reflectance bands at 555 and 645 nm along with the land-water profiles (Feng et al., 2012). The NDWI (Normalized Difference Water Index, NDWI =

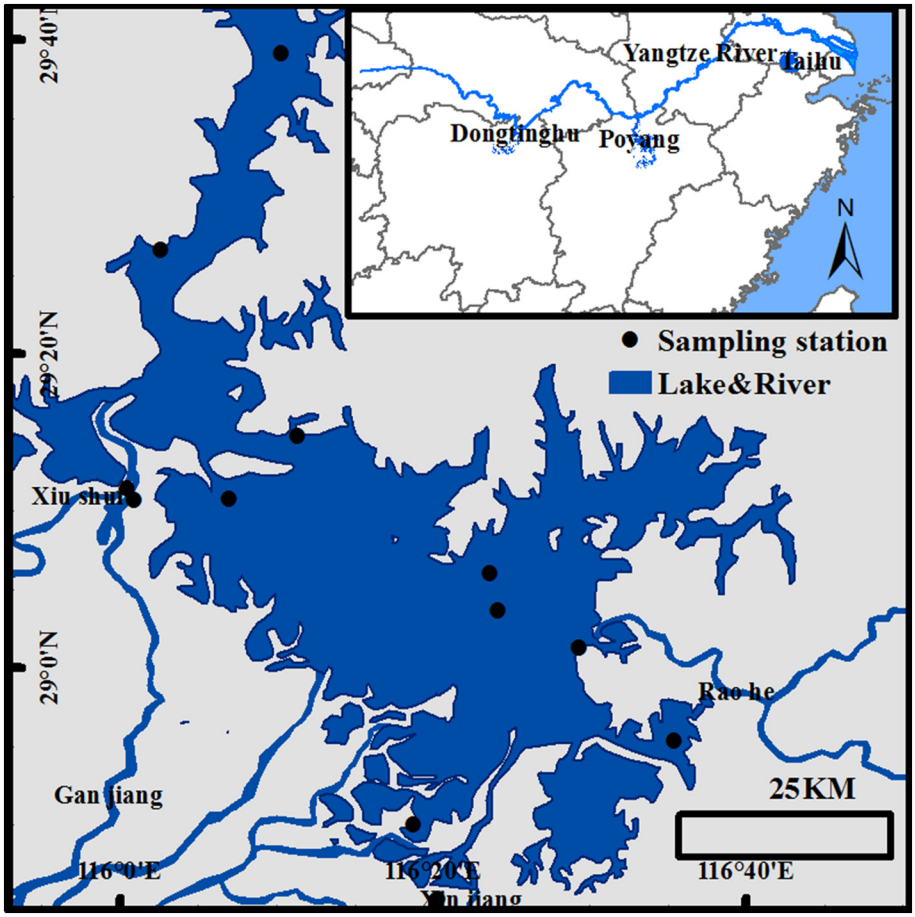


Fig. 1. Geographic setting of Poyang Lake and existing water quality monitoring sampling sites (black circles) of Water Resources Department of Jiangxi Province.

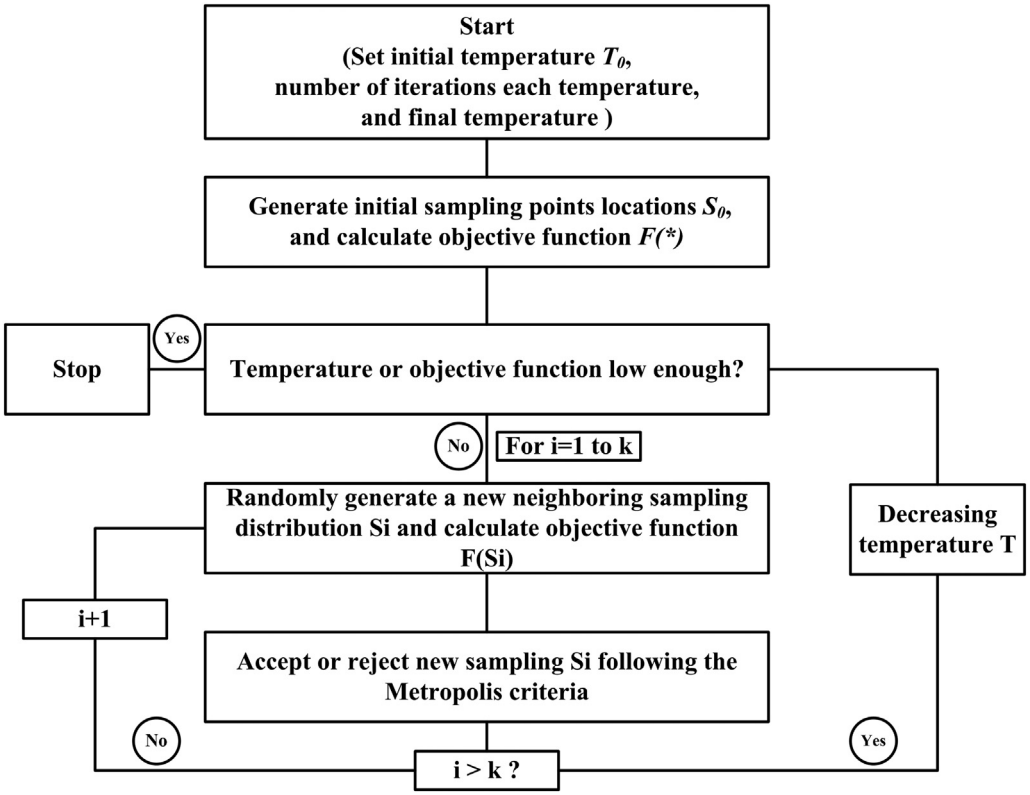


Fig. 2. Flowchart of SSA spatial sampling optimization algorithm.

$(R_{green} - R_{nir}) / (R_{green} + R_{nir})$) was used to build the water mask since it enlarges the signal contrast between water and land.

Third, field measurements were performed from 13 July of 2009 to 10 October of 2017 by cruise surveys with a total of 138 sampled stations, ranged from turbid to clear waters (TSS of 1–300 mg $L-1$). At each station, water samples were collected and filtered immediately on a pre-weighed Whatman filter (GF/F or Cellulose Acetate Membranes) with a diameter of 47 mm. The filter was stored in a desiccator, burned at 550 °C for three h, and weighed again in the laboratory. The TSS concentration was determined according to the weight difference normalized by the filtered water volume. 79 matchups of MODIS daily surface reflectance data and in-situ sampling data were selected to develop a remote-sensing TSS retrieval model. The model was based on MODIS green (555 nm) and red (645 nm) bands extracted from the above MODIS products, with the formula as: TSS (mg/L) = 132.83 × (R645/R555)2–52.618 × (R645/R555). This model produced the best precision with the highest determination coefficient ($R2 = 0.88$, $P < 0.05$) and the lowest RMSE (34.2%) compared with other methods, validated using field-collected data.

Additionally, the pixels affected by cloud shadow and atmospheric correction failure were removed using the Quality Control (QC) flag. The processing of the MODIS data and statistic analysis, as well as the sampling approach were programmed and implemented in ENVI/IDL software.

3. Methods

3.1. How to describe the spatial-temporal variations of water quality?

The K-mean clustering method was selected to assign pixels with similar variation trends into one cluster, and TSS variations were obtained from remote sensing TSS data from 2000 to 2014. K-means clustering is one of the simplest and widely used unsupervised machine learning algorithms (Hartigan and Wong, 1979). The purpose of clustering is to locate pixels with high similarity of spatial-temporal patterns, and thus one cluster could be represented by interpolation of limited sampling stations. Thus, pixels of high similarity to each other were determined to the same cluster and could be effectively represented using typical sampling observations.

One critical issue to be resolved when k-means clustering is used is to determine the optimal number of clusters. A standard measure named Sum of Squared Error (SSE) was used to determine k by trial and error. SSE is the sum of the squared differences between each observation and its cluster center, thus being used to measure variation within a cluster. The definition of SSE was given in the eq. 1, in which x is a data point in a cluster, C_i, m_i is the clustered index and the center of cluster i.

$$\mathrm{SSE} = \sum_{i=1}^{k} \sum_{x \in C_i} \mathrm{dist}^2(m_i, x) \tag{1}$$

From the definition of SSE, it is reasonable to choose clusters with the smallest SSE. To determine the proper k, we run k-means clustering on the dataset for a range of values of k. Since SSE tends to decrease toward 0 as k increases, a small k with a low SSE will be chosen. When clustering the time series TSS data from remote sensing imagery, each pixel was considered as one data point with 15 dimensions (TSS values from 2000 to 2014). Thus, pixels with highly similar spatial-temporal variations would be classified into the same cluster.

3.2. How to determination of optimal sampling size?

An optimal sampling size requires prior spatial-temporal information on the TSS variations. In this study, a geostatistical approach based on semi-variogram analysis is used to analyze the spatial-temporal variations of TSS. The semi-variogram analysis has been

widely used in environmental science, including soil mapping (Chaney et al., 2015), vegetation monitoring (Zeng et al., 2014), meteorological data (Gebremichael and Krajewski, 2005), and water quality (Caeiro et al., 2003). The semi-variogram plots the semivariance against the distance between the measurements, which indicates distance separations beyond which the measurements become irrelevant (Aurin et al., 2013; Ibrahim et al., 2014). The experimental semivariance (Eq. 2) was calculated as half the squared TSS difference at different distances (denoted as spatial lag h):

$$r(h) = \frac{1}{2N(h)} \sum_{j=1}^{N(h)} [z(x_j) - z(x_j + h)]^2 \quad (2)$$

In which $z(x_j)$ is the TSS value at pixel x_j. h is the distance of two different pixels, and $N(h)$ is the data pairs of pixels at a distance of h. Then, the semivariance with a set of h could be obtained, as shown in Supplementary Fig. S2.

Three parameters could be provided from the semi-variogram: range (a), sill (c), and nugget ($c0$). The sill (c) indicates the total variance of the TSS variations, and the nugget ($c0$) indicates random measurement uncertainties. The range (a) is the distance at which observations between two locations become independent over this distance. Thus, semi-variograms could provide 1) an estimate of the TSS variance and 2) an estimate of the minimum distance required for samples to be considered spatially independent, which can be used to inform sample size for designing a robust sampling strategy. The range represents how the sampling points were dependent on each other (Journel and Huijbregts, 1991). In this study, the commonly used spherical model was adopted to analyze the TSS variation's spatial scale using an ordinary least squares fit method (OLS) (Li et al., 2018).

3.3. How to obtain the optimal sampling distribution?

Spatial simulated annealing (SSA) is a classical optimization algorithm based on Monte Carlo simulation, which has been widely used to solve sampling design problems in the environmental sciences. A standard simulated annealing algorithm was adopted in this study as developed in Kirkpatrick (Kirkpatrick and Toulouse, 1985), but expanded to the spatial scale following Brus and Heuvelink (Brus and Heuvelink, 2007). Optimal spatial sampling's main idea is to obtain the best sampling allocations with high precision and low cost by allocating sampling locations randomly to achieve minimizing prediction errors or called an objective function. In practice, SSA randomly moves the spatial location of a candidate sampling point at a time and calculating whether the updated spatial sample distribution is better than the previous one by comparing the objective function. Compared to other local optimal methods such as the hill-climbing method, the SSA is more robust since it provides chances of accepting a worse spatial sampling distribution so that the algorithm can escape from local optima solutions. Moreover, the probability of accepting a worse spatial sample configuration will decrease with the optimization procedure's iteration. Therefore the SSA can get very close to the global optimum configuration.

The objective function used in this study to determine whether a candidate sampling distribution is accepted or rejected during an iteration of the SSA, is defined as the Minimization of the Spatial-temporal mean Error (MSTE). The MSTE is the temporal average of spatial interpolated TSS error compared to observed data over the time-series observations.

$$\mathrm{MSTE} = \sqrt{\sum_{i=1}^{t} \sum_{j=1}^{s} \frac{1}{N_s} \frac{1}{N_t} \left(\mathrm{TSS}^{i,j}_{\mathrm{observed}} - \mathrm{TSS}^{i,j}_{\mathrm{inferred}}\right)^2} \quad (3)$$

where $\mathrm{TSS}^{i,j}_{\mathrm{observed}}$, $\mathrm{TSS}^{i,j}_{\mathrm{inferred}}$ were observed and inferred TSS at location j and time i using current sampling allocation, N_s and N_t is the total count of spatial pixels and days.

In practice, SSA can be described step by step as; i): First, initialize the simulation by setting initial temperature T_0, the number of iterations each temperature, and final temperature. ii): Randomly allocating n sampling locations S_0 for each spatially clustered region and calculating the objective function MSTE from these random locations. iii): Update the sampling locations by generating a random neighboring sampling distribution S_i and calculating corresponding objective function $F(S_i)$. iv): Start iteration until the temperate or objective function is low enough. For each iteration, the probability of changing the current sampling locations S_i to new S_{i+1} following the Metropolis algorithm's acceptance rule (Metropolis et al., 1953). If the objective function $F(S_{i+1})$ is smaller than $F(S_i)$ then the probability P is 1, which means the sampling distribution S_{i+1} is accepted with an improvement to the final sampling design and is stored for memorizing. If $F(S_{i+1})$ is smaller than $F(S_i)$ then the current design is accepted only with a certain probability

$$P\{S_i \rightarrow S_{i+1}\} = \begin{cases} 1, & if, F(S_i) \leq F(S_{i+1}) \\ e^{\left(\frac{F(S_i) - F(S_{i+1})}{T}\right)}, & if, F(S_i) > F(S_{i+1}) \end{cases} \quad (4)$$

where T is temperature, which is set large at the beginning of the algorithm, the larger T indicates a higher probability of accepting a worsening sampling design. The algorithm then iterates by decreasing T and k times iterations are performed for each T. Therefore, the probability of accepting a worse sampling design could prevent the algorithm from being trapped in local minima of the objective function, and high overall accuracy can be achieved. In this study, the initial and cooling temperatures were set as 2 and 0.01, and the temperature decreasing rate was set as 0.95. For each temperature, 500 iterations were performed to obtain a robust result.

3.4. Comparison and accuracy assessment of varied sampling methods

Three commonly used conventional sampling methods were selected and evaluated to compare the proposed sampling approach, including systematic sampling, stratified sampling, and expert sampling. Detailed information on each sampling method was provided as follows.

Systematic sampling is frequently adopted in the research community for its simplicity and quality and not requiring any prior knowledge (Haining, 2015). The samples are selected evenly and randomly, which means each location has an equal and independent chance to be selected. Therefore, the distribution of systematic sampling could be determined while the sampling size or sampling interval was determined for a specific study region by knowing the spatial area and geometry shape (Van Niel and Laffan, 2003). The systematic sampling with the sampling size of 45, 55, 65, 80, 90, 120, 160 for the Poyang Lake was obtained, which corresponding to a distance of 1.5, 2.0, 2.5, 3.0, 3.3, 3.6, and 4.0 km between two sampling stations. An illustration of the sampling distributions was presented in Supplementary Fig. S3 (a, b, c) at the sample size of 160, 90 and 45, respectively. Besides, the random number generator was run for 100 times for each sampling size with different initial points, and the averaged results were then used for evaluations of each sampling strategy.

Stratified sampling is one sampling approach designed by dividing the whole study region into independent, non-overlapping subregions, called clusters or strata, then applying a simple random sample within each stratum (Crepin and Johnson, 1993). Strata are obtained by clustering or classification based on spatio-temporal characteristics of the study area and thus are more heterogeneous within strata than among strata (Wang et al., 2012). Stratified random sampling is appropriate, especially when the study region is heterogeneous and can be classified with ancillary information such as remote sensing data. The method's precision depends on each stratum's distinct, which is generally higher than the simple random or systematic sampling method (Michalcová et al., 2011). An example of the stratified sampling was

presented in Supplementary Fig. S3 (d, e, f) with the total sample size of 160, 90 and 45, respectively. The strata were obtained using the method in Section 3.1. Like systematic sampling design, the random number generator was run 100 times for each sampling size with different initial points. The averaged results were then used for evaluations of each sampling strategy.

Expert sampling is also known as judgmental or purposive sampling, in which the sampling design, including size and positions, is based on expert knowledge or professional judgment. The expert sampling is efficient and cost-effective than systematic or random designs with expert knowledge of regions of interest. However, its precision of water quality estimates depends on personal judgment, which may not quantify the investigation's level of uncertainty accurately. Fig. 1 exhibited the spatial distribution of expert sampling locations in Poyang Lake, with a sampling size of 17 in total.

For each sampling strategy, the whole lake's TSS values were obtained by interpolating limited sampling values. The spatial interpolation was conducted using the Inverse Distance Weighting (IDW) method in this study. The IDW is a widely used interpolation method based on geography's first law that unsampled locations are more similar to the sampled point closer to it than those further away (Tobler, 2004). Thus, TSS values of unsampled locations were estimated as a linear combination of nearby observations where weight was an inverse function of the distance from the unsampled points to the nearby sampling locations (Burrough et al., 2015).

To evaluate the sampling strategy precision, remote sensing TSS products were used as a reference or measured data. The root mean square errors (RMSE, Equation .5) and mean percentage error (MPE, Eq. 6) were calculated between measured data and interpolated data from sampling points.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{i=N}(P_i - Q_i)^2} \tag{5}$$

$$\mathrm{MPE} = 100 \times \frac{1}{N}\sum_{i=1}^{i=N}\left|\frac{P_i - Q_i}{Q_i}\right| \tag{6}$$

P_i stands for the interpolated TSS value at location i using the sampling strategy, Q_i is the measured TSS values at the same location. The time series estimated TSS errors from 2000 to 2014 were obtained and averaged at different time scales (monthly, annually) to predict statistical errors for water quality monitoring using different sampling methods.

4. Results and analysis

4.1. Spatial-temporal variations of TSS

Fig. 3 revealed the trends of SSE, and the total explained variance with the number of cluster K increased. The clustering algorithm converges to relatively stable values when the number of cluster K larger enough, for instance, K > 8. A significant drop in SSE was found at 4 clusters, and a knee point was also found correspondingly with 4 clusters, with 93.4% variance explained. By increasing the number to 5 or even higher, the SSE and total explained variance were not improved much; therefore, the final number of clusters was determined to be 4.

The TSS variations of the Poyang Lake could be divided into 4 clusters/classes which could be reasonably recognized (Fig. 4 and Table 1). Class 1 was mainly located in the south and east-south area of Poyang Lake. Class 2 and 3 were mainly located in the part of the east and middle lake. Class 4 was distributed in the narrow water channel in the north of Poyang lake. Descriptive statistics of long-term TSS variations in 4 clusters were listed in Table 1. Cluster 1 represents 54.36% of the total area with the lowest values and variations of TSS. The TSS values of class 1 ranged from 0.83 to 24.6 mg/L, with a mean value of

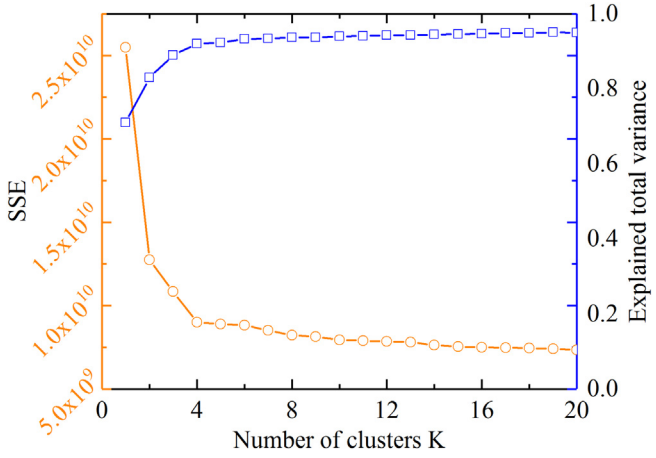


Fig. 3. Variations of Sum of Squared Error (SSE) and total explained variance with increasing K.

6.73 mg/L. Class 2 occupied 22.89% of the Poyang Lake, representing the medium concentration of TSS, with the mean values of 28.12 mg/L. TSS became higher for cluster 3, and the highest TSS was found in cluster 4 and cluster 3. The mean and maximum TSS in cluster 4 were 59.02 and 136.46 mg/L, which were about eight times higher than the lowest TSS in cluster 1. Therefore, it could be inferred that the spatial-temporal variations of TSS in Poyang Lake were significant, which cloud has also been indicted by the standard deviation (STD) for each class. The STD ranged from 2.72 to 25.91 from class 1 to class 4, which means class 1 was relatively stable, and high variations were found in class 4.

15-year long-term TSS trends in Poyang Lake observed by Terra/Aqua MODIS from 2000 to 2014, were shown in Fig. 4. Generally, significant spatial and temporal variability were revealed for the whole study area and each class. Class 1 was dominated by low TSS values and low inter-and inner-annual variations (STD). Class 2 to 4 showed an increasing annual TSS trend, and the largest TSS values and variations were also found in Class 4. The higher values in Class 3 and 4 were induced by intensive human activities such as sand dredging in Poyang Lake. Moreover, there were evident drops in TSS around 2008 or 2009, which were caused by the sudden decrease of sand dredging due to the sand dredging ban by the local authorities (Feng et al., 2012; Li et al., 2014).

4.2. Required number of sampling sites

Fig. 5 presents the experimental semivariance plots and the fitted theoretical variogram models to quantify the spatial variability of TSS ($R^2 > 0.9, P < 0.01$) for each class. Table 2 listed the range, sill, and nugget values of different classes corresponding to different levels of TSS variations, as described in the previous section. A larger variogram sill (C1) was observed in Class 4 (approximately 3.8) and Class 2, indicating more considerable TSS variations, and a smaller sill in Class 2 as expected. Similar patterns were also revealed in the nugget effect (C0) for these different classes, with increased nugget values from 0.47 for class 1 to 1.18 for class 4.

The spatial range in the fitted variograms, whereby a suitable sampling distance should be retained to capture the TSS variations. For the TSS variation in class 1, the mean spatial range was approximately 4.2 km. A smaller spatial range of TSS variation was 1.5 km in class 4, indicating a smaller sampling distance was required to monitor the TSS variation. Therefore, each class's sampling number was determined as the spatial range proportion to the class area while considering the class's shape. From class 1 to class 4, the minimum number of samples required was 25, 18, 11 and 11.

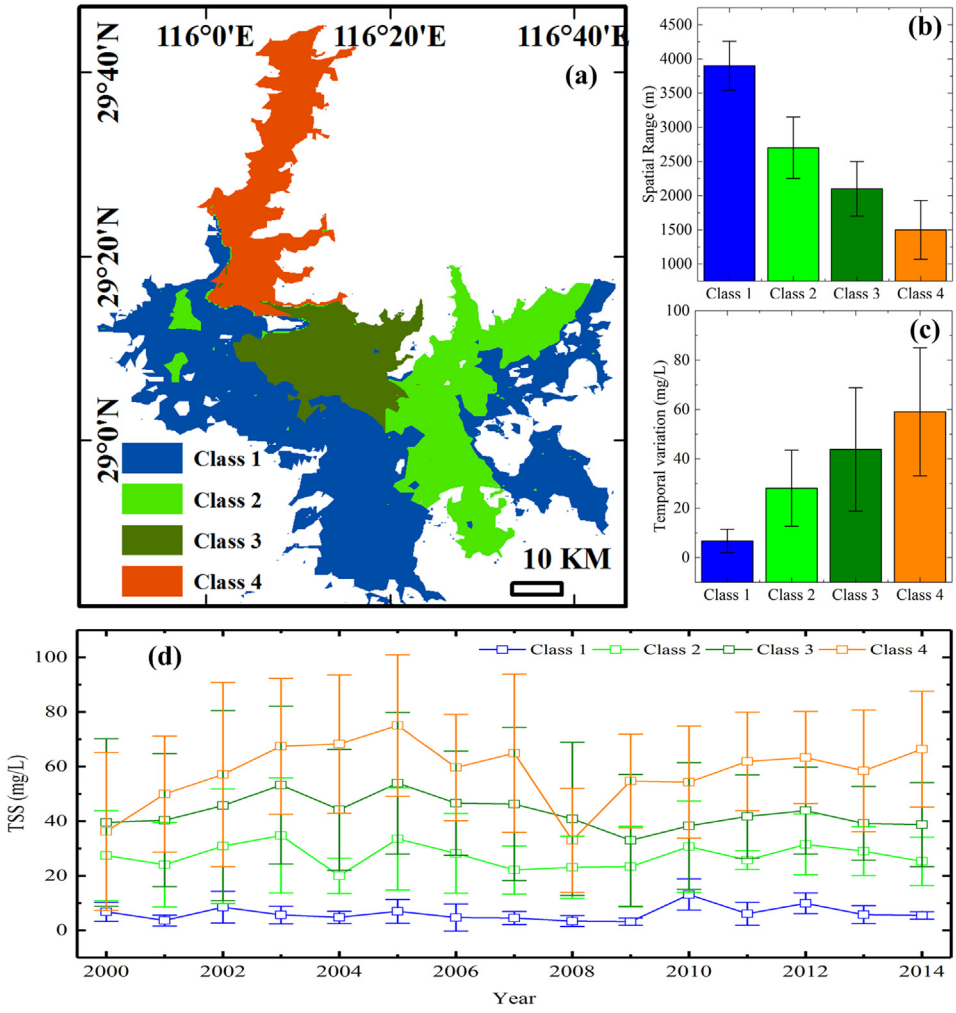


Fig. 4. Spatial-temporal clusters of TSS variations in Poyang Lake using 15 years MODIS data.

4.3. Optimal spatial distributions of sampling sites

Fig. 6 shows the decrease of the RMSE and MPE of the predicted TSS during the SSA. Some worse sampling designs could also be accepted at the earlier stage of the process. After this initial phase, the RMSE and MPE decreased steadily and almost stable after 10,000 iterations. No further reduction of prediction errors was achieved after 24,000 iterations, which means that the algorithm reached an acceptable optimum design and quitted the simulation after 34,900 iterations. The results were confirmed by running the algorithm 200 times with different initial sampling locations and obtaining similar patterns (sensitivity analysis in Section 5.2). Overall, the RMSE and MPE dropped from 19.1 mg/L, 61.9% to 11.1 mg/L and 35.3%, representing an improvement of TSS prediction of about 42.1% and 42.9%, respectively.

Table 1
Descriptive statistics of long-term TSS variations in 4 clusters of Poyang Lake.

	Class 1	Class 2	Class 3	Class 4
Mean (mg/L)	6.73	28.12	43.80	59.02
STD	2.72	15.43	25.01	25.92
Max (mg/L)	24.60	93.22	141.13	136.46
Min (mg/L)	0.83	7.53	5.52	0.09
Median (mg/L)	5.30	23.94	39.26	60.52
Area percentage	54.36%	22.89%	17.08%	3.71%

Fig. 7 presented an example of the initial and optimized TSS sampling locations. The optimized design showed a higher density in regions with higher TSS and coefficient variation (CV) of TSS. For instance, more samples were located along the water channel in the middle and north lake and higher CV regions in the south-east lake. Fig. 7 presented the histograms of TSS and TSS's CV estimated using the entire image pixels, initial sampling locations, and SSA optimized sampling locations. It is worth noticing that the range of both TSS and CV of TSS from SSA sampling was more consistent with the results from the entire image pixels and the shape of the distributions. In contrast, initial sampling results were prone to miss the higher and lower value ranges of TSS estimations. The improvement of SSA sampling was more specific, as listed in Table 3. The maximum, mean, minimum, upper 95% confidence interval (CI) of mean and lower 95% CI of the mean of both TSS and TSS CV was more consistent between SSA sampling an entire image data. Thus, it is reasonable to conclude that the optimized sampling could improve the overall accuracy of TSS estimations.

The spatial distributions of RMSE (mg/L) and MPE (%) of predicted TSS errors using SSA sampling were displayed in Fig. 7 (d) and (e). The RMSE ranged from 0.01 to 47.94 mg/L, with the mean (±standard deviation) value of 16.65 (±7.54) mg/L. Higher RMSE was found in the middle and south lakes, as well as some parts of the north outlet, and these areas were also high CV regions, but not vice versa. Besides, regions in some lake bays and lakeshore showed higher RMSE, which may be caused by irregular lake shape, limited sampling locations, as well as high CV due to intensive land-lake interaction. Similar spatial

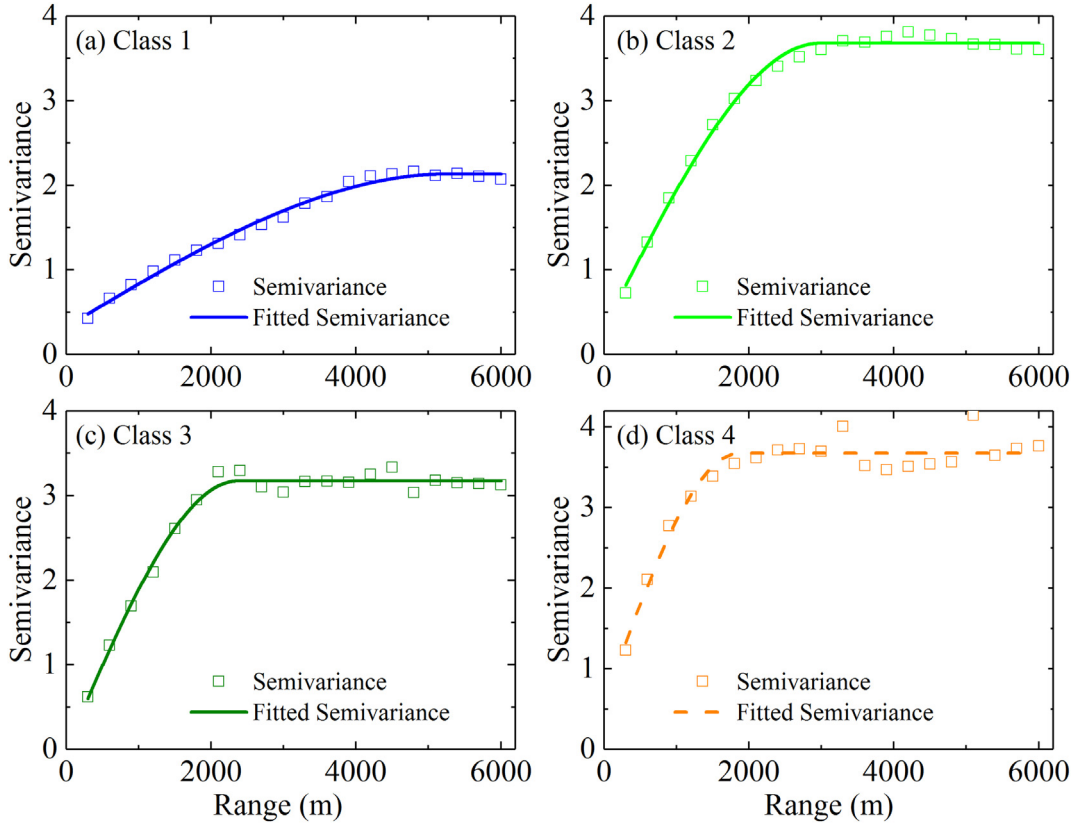


Fig. 5. Semi-variograms of the TSS calculated from remote sensing TSS products. The solid lines represent the fitted spherical models.

Table 2
Statistics of the semi-variogram analysis of TSS variations at Poyang Lake.

Class	Area (km^2)	Range (m)	Nugget (C_0)	Sill (C1)	Sampling numbers
Class 1	2120.04	4200	0.47	2.34	25
Class 2	892.71	2400	0.78	3.61	18
Class 3	666.12	2100	0.52	3.27	11
Class 4	144.69	1500	1.18	3.79	11

patterns were found with MPE, with the maximum, mean, and minimum values of 93.4%, 20.8% and 4.8%, respectively. To prove SSA-optimized sampling's improvement, we compared the spatial and temporal trends of TSS prediction errors to conventional methods in the next section.

4.4. Improvement compared to conventional methods

The spatial distribution of the average TSS estimation errors (RMSE and MRE) using conventional sampling methods, including systematic, stratified, and expert sampling, were presented in Fig. 8. RMSE and MRE's spatial patterns from three conventional methods were quite like SSA sampling, but with higher values indicating conventional methods' lower performances. For instance, higher RMSE was found for all three methods in the middle, south lake, and parts of north outlet and some lake bays and lakeshore, where TSS CVs were also high.

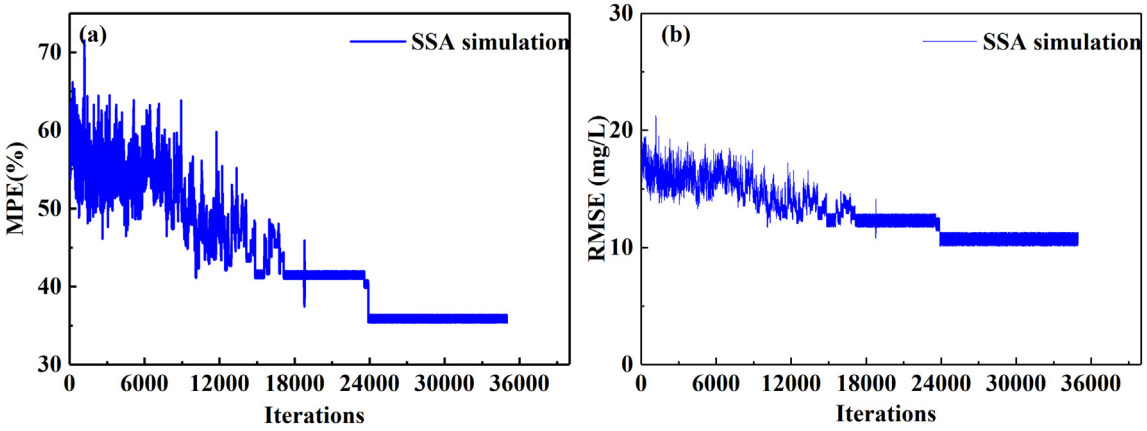


Fig. 6. Trace of the minimization objective function, MPE and RMSE for TSS estimations, during SSA (for a case of 34,900 iterations).

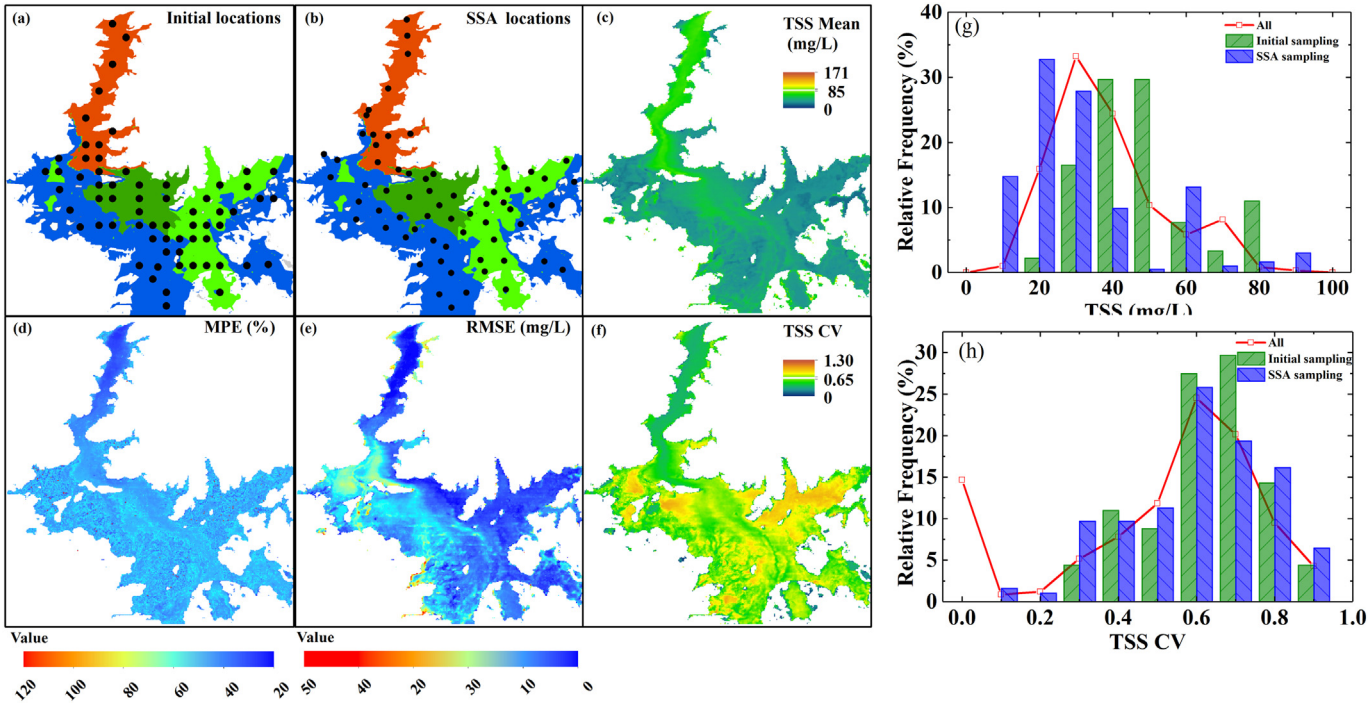


Fig. 7. Optimization results of sampling locations using SSA method: (a) Initial locations; (b) Optimized locations; (c) TSS mean of the Poyang Lake from 2000 to 2014; (d) MPE (%) of TSS estimation from optimized samples; (e) RMSE (mg/L) of TSS estimation from optimized samples; (f) Coefficient of variation (CV) TSS of the Poyang Lake from 2000 to 2014; (g) Histograms of TSS and (h) CV of TSS estimated using the entire image pixels (All), initial sampling locations, and SSA optimized sampling locations.

Table 4 showed the histograms and statistical comparisons of the proposed SSA sampling method and three conventional methods. These results were obtained and compared under the same sampling size of 65 for SSA, systematic, and stratified sampling method, but only 17 samples were used for expert sampling. The improvement of SSA sampling was apparent, with the mean (±STD) RMSE of 15.65 (±6.54) mg/L, while the mean RMSE for systematic, stratified and expert sampling was 19.11 (±6.02), 22.15 (±9.05), 25.69 (±6.07) mg/L, respectively. These results implied that by optimizing sampling locations using SSA, the whole lake's TSS estimations errors were reduced by 18.11% and 29.34% on average, compared to systematic and stratified sampling under the same sample size. Besides, the maximum and minimum RMSE of SSA sampling was also reduced. The SSA's superiority was more pronounced when the MPE of these methods was compared as listed in Table 4.

Moreover, the improvement of SSA was further demonstrated by an index called percentage less than a defined RMSE of MPE threshold. For instance, the percentage of RMSE less than 5 mg/L was 5.83%, 0.07%, 0.49%, 0.03% for SSA, systematic, stratified, and expert sampling method, respectively, and when the RMSE threshold increased to 20 mg/L, the percentage met the criteria for each method was 82.19%, 75.43%, 55.90%, 26.14%. A similar but more significant improvement of SSA was also revealed from MPE comparisons, with 91.52% of all data has MPE of less than 30%, while the MPE percentage of systematic, stratified, and expert sampling method was 15.95%, 17.57%, 8.64%.

Furthermore, we compared the performance of these sampling methods under the scenario of long-term TSS observations, taking annual TSS observations as a case. The assessments were achieved using the remote sensing TSS product images as "ground truth" references, and the annual predicted TSS values were obtained by interpolation using sampling locations of each sampling method. Finally, the annual mean RMSE and MPE were calculated and presented in Fig. 9. SSA's advantage over another three sampling methods in long term TSS monitoring was with lower time series RMSE and MPE. RMSE for SSA was maintained at approximately 10% from 2001 to 2014, while the systematic method was around 18%, the stratified method of about 21%, and the expert method of about 28%. All provided results proved SSA sampling's superiority in both lower spatial and long-term TSS prediction errors. Besides, it is interesting to notice that the spatial distributions and temporal trends of both RMSE and MPE showed similar patterns for different methods or different sampling sizes, which implied that these methods' sampling accuracy might be affected by similar or even the same factors.

Table 3
Statistics and comparison of TSS and CV of TSS estimated using the entire image pixels (All), initial sampling locations, and SSA optimized sampling locations.

		All	Initial	SSA
TSS (mg/l)	Max	124.8	76.3	92.0
	Mean	43.5	45.1	44.3
	Min	4.6	17.9	15.4
	Upper 95% CI of Mean	43.46	46.14	45.37
	Lower 95% CI of Mean	43.19	39.96	41.29
TSS CV	Max	1.07	0.97	0.94
	Mean	0.68	0.68	0.66
	Min	0.005	0.37	0.17
	Upper 95% CI of Mean	0.66	0.71	0.70
	Lower 95% CI of Mean	0.62	0.65	0.61

5. Discussions

5.1. Factors of influencing sampling accuracy

Two factors, including sampling method (size and locations) and TSS variations, were assumed to influence TSS observations' sampling accuracy. Fig. 10 provided RMSE and MPE trends under different sampling sizes from 35 to 165 using systematic and stratified sampling methods, together with expert sampling with a sampling size of 17. Clear decreasing trends of RMSE and MPE were both observed with the increasing sampling size, as expected. For instance, the RMSE of the systematic

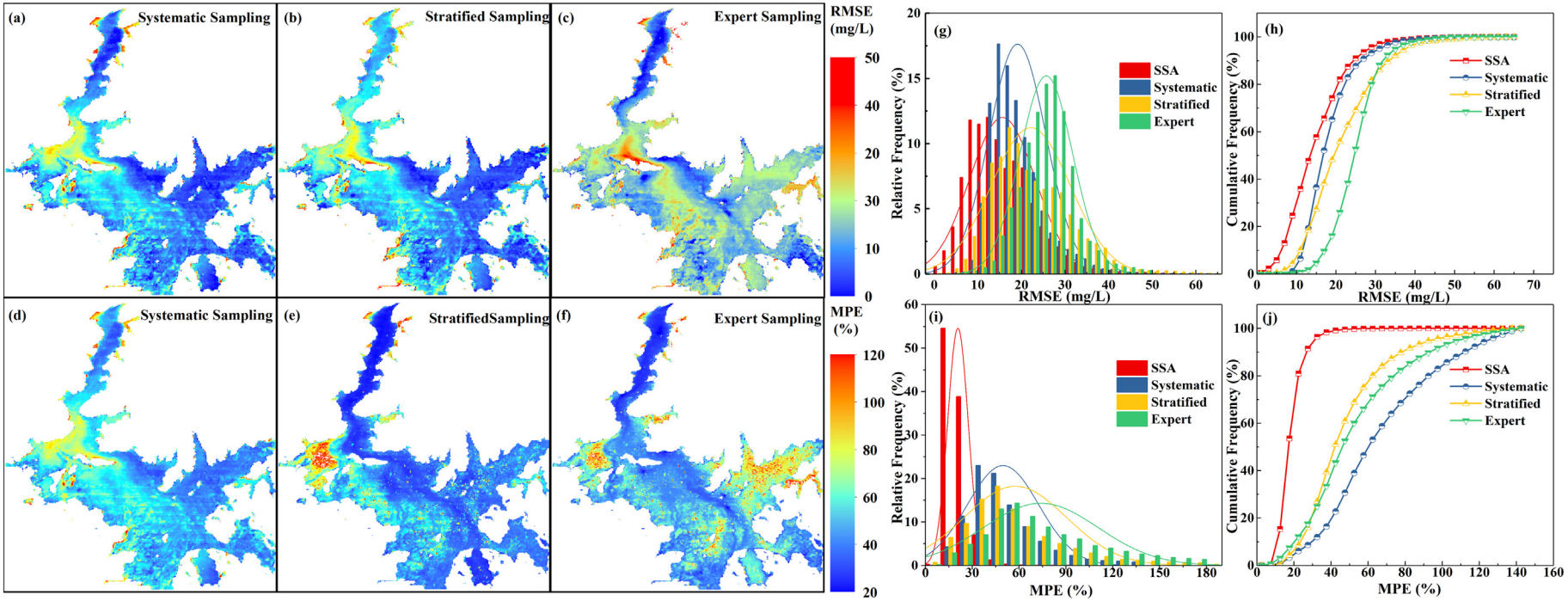


Fig. 8. RMSE (mg/L) and MPE (%) of TSS estimation using conventional sampling methods: (a), (b), (c) were RMSE of systematic, stratified, and expert sampling; (d), (e), (f) were MPE of systematic, stratified and expert sampling methods under the sampling size of 65. (g), (h), (i), (j) were the histograms and comparisons of RMSE (mg/L) and MPE (%) of TSS estimated using different sampling methods under the sampling size of 65.

Table 4
Statistics and comparisons of RMSE (mg/L) and MPE (%) of TSS estimated using different sampling methods.

		SSA	Systematic	Stratified	Expert
RMSE (mg/L)	Mean	15.65	19.11	22.15	25.69
	STD	6.54	6.52	9.05	6.07
	Min	0.01	2.80	0.62	1.79
	Max	54.95	64.97	64.94	69.99
	Percentage (%) < 5 mg/L	5.83	0.07	0.49	0.03
	Percentage (%) < 10 mg/L	24.66	1.24	4.39	0.36
	Percentage (%) < 15 mg/L	57.78	36.58	27.32	4.67
	Percentage (%) < 20 mg/L	82.19	75.43	55.90	26.14
MPE (%)	Mean	20.77	73.88	49.70	57.56
	STD	6.61	36.93	22.93	32.87
	Min	4.76	3.22	3.56	4.53
	Max	93.40	179.98	139.98	184.88
	Percentage (%) < 10%	0.27	0.04	0.73	0.17
	Percentage (%) < 20%	53.59	4.38	7.47	3.31
	Percentage (%) < 30%	91.52	15.95	17.57	8.64

method dropped from 21.65 to 19.31 mg/L, while the sampling size enlarged from 35 to 165. This meant the TSS estimation errors only decreased by approximately 10.81% and 10.04%, respectively. Similar results could be seen for the stratified sampling method. Therefore, the influence of sampling locations or distributions matters more than sampling sizes. Fig. 10 also implied the number of samples required to meet a certain sampling accuracy for conventional methods. For instance, at least 120 samples were required in Poyang Lake to obtain a long-term TSS accuracy of RMSE less than 20 mg/L using a systematic method or more than 140 samples required for the stratified sampling method.

The spatial distributions and temporal trends of both RMSE and MPE showed similar patterns for different methods, indicating the sampling accuracy may be affected by some common factors. We compared and listed the RMSE, together with TSS CV, sampling density, and CV's ratio to sampling density for different methods and TSS cluster classes in Table 5. The sampling density was defined as the ratio of sampling

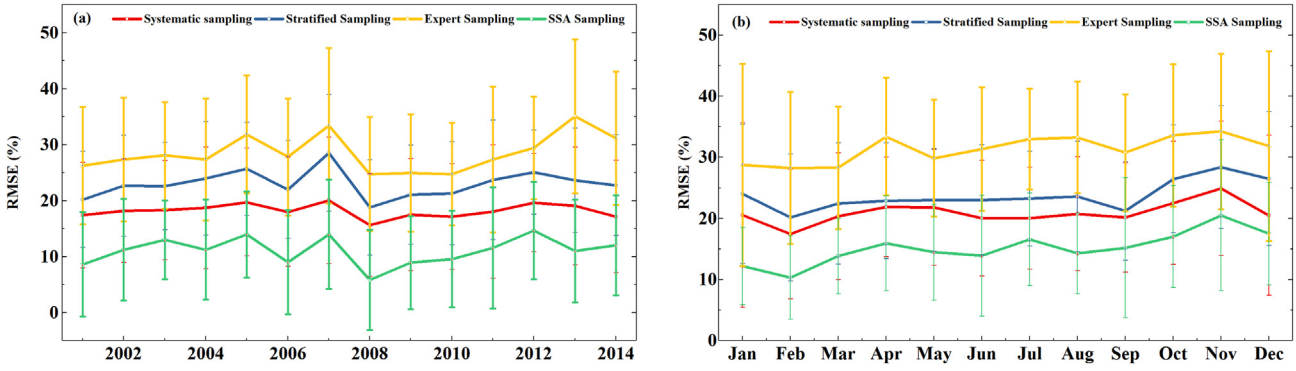


Fig. 9. Temporal trends of RMSE and MPE of TSS prediction errors using different sampling methods: (a) Annual trends of RMSE; (b) Monthly trends of RMSE.

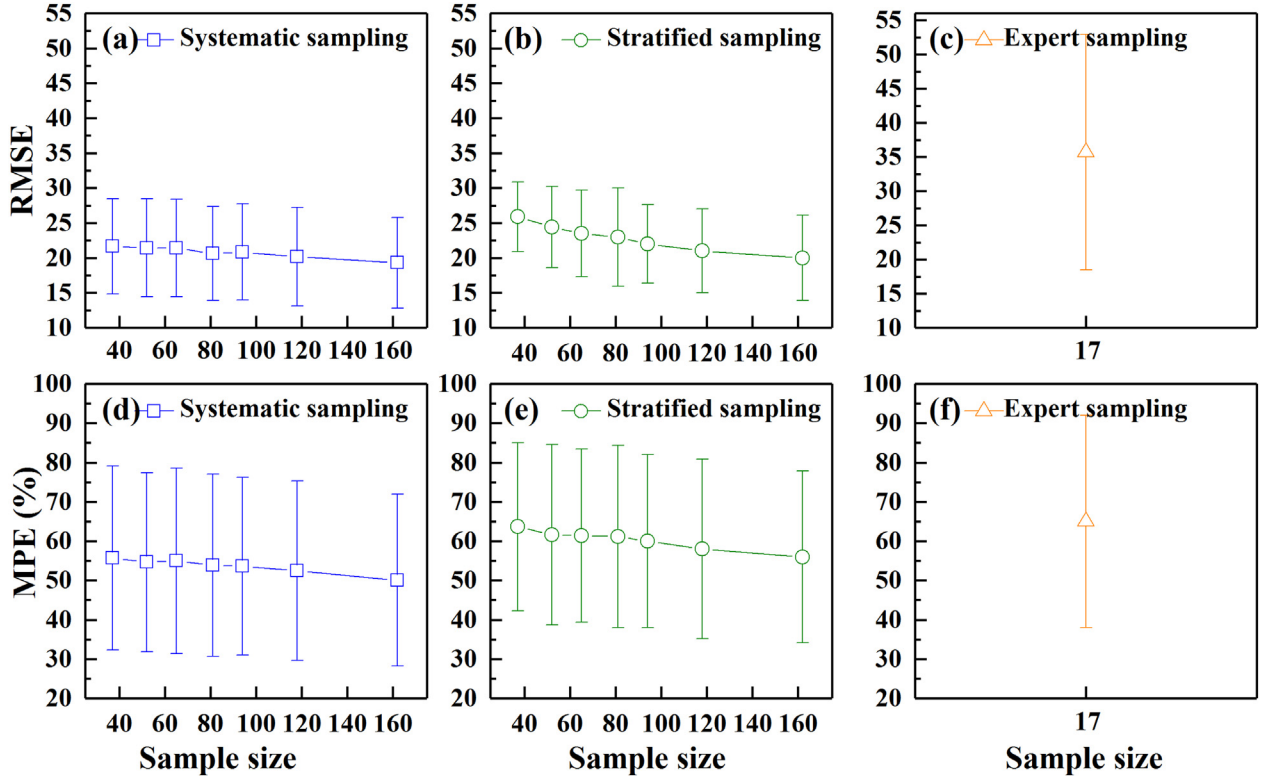


Fig. 10. Variations of RMSE and MPE of TSS predictions errors under different sampling sizes of different sampling methods.

Table 5
Influence factors analysis on sampling accuracy of different sampling methods[a].

		Class 1	Class 2	Class 3	Class 4
TSS	CV	0.58 (0.17)	0.49 (0.16)	0.56 (0.15)	0.42 (0.14)
Systematic	RMSE (STD)	18.74 (2.35)	16.05 (1.41)	18.86 (1.97)	19.90 (2.91)
	Sampling density	0.007	0.019	0.019	0.069
	Ratio	78.24	26.05	30.03	6.08
Stratified	RMSE (STD)	24.86 (3.51)	23.69 (2.57)	23.92 (2.91)	22.35 (3.71)
	Sampling density	0.014	0.069	0.054	0.318
	Ratio	40.86	7.00	10.29	1.32
Expert	RMSE (STD)	23.53 (3.67)	23.98 (6.68)	23.20 (2.48)	22.94 (3.95)
	Sampling density	0.0028	0.0034	0.0015	0.0207
	Ratio	204.29	144.72	370.32	20.26
SSA	RMSE (STD)	22.87 (4.46)	20.36 (4.03)	24.84 (3.44)	22.71 (3.08)
	Sampling density	0.012	0.020	0.017	0.076
	Ratio	49.03	24.12	33.67	5.53

[a] Ratio = CV/sampling density; the units of sampling density: samples per km^2.

size to the total area of each class. In general, the trends between RMSE and single factor were not harmonious, such as RMSE and CV, or RMSE and sampling density. For instance, the RMSE of the systematic method was 18.87, 16.05, 18.86 and 19.90 mg/L from class 1 to 4, showing first a decreasing and then an increasing trend. In contrast, the TSS CV showed a decreasing, increasing, and then increasing trend, and the sampling density with an increasing trend from class 1 to 4. However, the comparison between RMSE and TSS CV's ratio to sampling density showed better consistency from class 1 to 4. For example, the trends between RMSE and ratio using a systematic or SSA sampling method were identical, thus proven that the sampling accuracy was under the joint influence of multi factors.

The relationships between the TSS prediction errors and TSS CV's ratio to sampling density were further confirmed using regression analysis. Strong correlations were found among all four sampling methods, as shown in Supplementary Fig. S4. The squared coefficient of correlation, R^2, was 0.61, 0.73, 0.42, and 0.78 for systematic, stratified, expert, and SSA sampling methods, respectively. A correlation greater than 0.7 is generally described as a high or strong correlation and a correlation between 0.3 and 0.7 as a moderate correlation. We are confident to conclude that the sampling accuracy was significantly determined by the ratio of TSS variations to sampling density, at the 0.05 level (P-value <0.05), and most of the sampling errors fall within the 95% confidence range. Therefore, for regions or lakes with high TSS variations, it is crucial to increase sampling size and optimize the sampling distributions to improve the sampling accuracy. Considering the cost and human resources of increasing sampling size, sampling optimization would be more efficient to enhance the sampling accuracy, especially in the long-term TSS or routine TSS observations.

5.2. Sensitivity analysis of SSA simulation

The initial sampling locations, initial temperature, cooling rate, and the number of iterations per new temperature may affect the simulation stability (Van Groenigen, 1997). To ensure that the annealing optimization's best fitness was reached, the cooling rate and the number of iterations per new temperature were set to large values as 0.95 and 500, and the cooling temperature was set low enough as 0.01. The procedure was repeated 500 times, using a combination of 50 different initial sampling locations and ten initial temperatures from 5 to 0.5 with an interval of 0.5.

Each simulation's mean center location was obtained as the mean longitude and latitude of all sampling points for each TSS class. The distributions of 500 mean center locations were displayed in Supplementary Fig. S4, together with each class's directional distributions. Sensitivity analysis indicated that the SSA method and results were relatively stable because the spatial distribution of all mean centers of 500 simulations was concentrated for each class. The percentage of the mean centers' distribution ellipse to the total area was 0.92%, 1.49% and 2.70% for class 2 to 4, while a larger percentage of 4.17% was found for class 1. Besides, The STD of the mean center locations could serve as an indicator of variations. As listed in **Supplementary Table S1**, the STD values were smaller than 0.03, except for class 1. The causes of more fluctuations during class 1, maybe its irregular shape and more considerable TSS variations (Table 5). However, these analyses proved that the SSA simulation method is stable, and the results were reliable.

The robustness of SSA lies in the possibility of accepting a worse spatial sampling distribution so that the algorithm can escape from local optima solutions. The probability of accepting a worse spatial sample configuration will decrease with the iteration of the optimization procedure. Therefore, the SSA can get very close to the global optimum configuration (van Groenigen, 2000). A previous study proved that the global solution could always be found if infinite calculation time was given (Aarts and Korst, 1989). The SSA sampling schemes outperformed the traditional sampling methods. The most prominent disadvantages of the SSA are the computational resources and time required to find the best solutions (Öhman and Eriksson, 2002). Therefore, there is a trade-off between simulation accuracy and efficiency, and improvement could be made by cloud computing or high-performance parallel computation systems.

5.3. Implications and recommendations

Spatial sampling is a crucial issue in environmental research and environmental protection departments because the sample configuration influences both costs and effectiveness of a survey, especially for long-term routine monitoring. However, existing sampling methods, including systematic, stratified sampling, often lack or limited prior knowledge. Although the previous study proposed to reduce the sampling numbers or sampling frequency by using remote sensing data to evaluate the representatives of existing sampling stations, this paper showed how remote sensing big data could serve as available pre-information to optimize the sampling scheme for the first time, to our knowledge.

Synoptic information on water quality parameters is hard to maintain from the routine in situ monitoring network due to the spatial inhomogeneity and temporal variations of the water quality parameter. However, integration of remote sensing, in situ data, and even water quality models could significantly improve water quality monitoring (Dekker et al., 2001). In the past few decades, remote sensing water quality monitoring has been prosperous regarding remote sensing technology and model development, to derive water quality data with sufficient accuracy on a regional basis, as well as to detect the spatial-temporal evolutions of the water quality (Aurin et al., 2013; Barnes et al., 2014). One main concern that should be addressed is to reduce

uncertainties of water quality to integrate remote sensing and in-situ measurements, which also required a reliable sampling method to determine the most representative sampling stations and monitoring network. Such work is more crucial for model development/validation, product validation of remote sensing water quality monitoring, or other land surface parameters, especially for validation over heterogeneous land or water surfaces (Wang et al., 2015; Wu et al., 2016), to obtain reliable remote sensing water quality products. Such water quality products are the great interest of many sectors such as the scientific community, government, and public, such as the 2030 Agenda for Sustainable Development (the SDGs) (Ga, 2015), the European Union (EU) Water Framework Directive (WFD) (De Stefano, 2010) and the United Nations Framework Convention on Climate Change (UNFCCC) (Bodansky, 1993). The optimal sampling design will also benefit the design and implementation of water quality monitoring networks, such as the National Ground-Water Monitoring Network of the US (https://acwi.gov/sogw/ngwmn_framework_report_july2013.pdf), the National Automatic Surface Water Quality Monitoring Station Network project (http://www.gov.cn/xinwen/2018-07/15/content_5306539.htm). The remote sensing big-data based sampling optimization would help design and re-assess the monitoring network's effectiveness.

6. Conclusions

This paper proposed a remote sensing and spatial simulated annealing integrated method for the optimal sampling design for water quality monitoring at lakes, which has two critical advantages over the conventional sampling approach. First, utilizing remote sensing big-data overcame the drawbacks of field sampling-based analysis, with limited spatial coverage and temporal range. Second, the spatial simulated annealing optimization algorithm based on the Monte Carlo simulation was adopted following the objective function of the minimization of the spatial-temporal mean error. By evaluating and comparing the proposed approach to the conventional sampling method, we have shown that the remote sensing-based SSA sampling design could significantly improve the accuracy of water quality monitoring at both spatial and temporal terms. TSS estimation errors of the whole lake were reduced by 18.11% and 29.34% on average, compared to systematic and stratified sampling under the same sample size. The annual TSS estimation errors were dropped by approximately 50%. Remote sensing showed great potential as an ideal means to provide spatially contiguous and comprehensive data as prior knowledge for efficient sampling design. Its capability has been expanded with more satellite data. Using this data, the water quality monitoring's sampling design could be of more precision and efficiency by evaluating existing monitoring stations concerning their representation of water quality or optimizing a new sampling network in future study and routine monitoring.

CRediT authorship contribution statement

Jian Li: Conceptualization, Methodology, Writing- Original draft preparation. **Liqiao Tian**: Software. **Tingting Li**: Visualization, Investigation. ***Shuanggen Jin****:* Supervision. **Yihong Wang**: Software, Validation. **Xuejiao Hou:** Data processing.

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.

Acknowledgments

This work was supported by the Strategic Priority Research Program Project of the Chinese Academy of Sciences (Grant No. XDA23040100), the National Key R&D Program of China (2018YFB0504904, 2016YFC0200900), the National Natural Science Foundation of China (Nos. 41571344, 41701379), the Open Research Fund of State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Grant NO.IWHR-SKL-KF201809.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2021.146113.

References

Aarts, E., Korst, J., 1989. Simulated annealing and Boltzmann machines: a stochastic approach to combinatorial optimization and neural computing. John Wiley & Sons, Inc.
Abd-Elrahman, A., Croxton, M., Pande-Chettri, R., Toor, G.S., Smith, S., Hill, J., 2011. In situ estimation of water quality parameters in freshwater aquaculture ponds using hyperspectral imaging system. ISPRS J. Photogramm. Remote Sens. 66, 463–472.
Agency, U.E.P., 2002. Guidance on Choosing a Sampling Design for Environmental Data Collection. United States Environmental Protection Agency Washington, DC.
Alcamo, J., 2019. Water quality and its interlinkages with the sustainable development goals. Curr. Opin. Environ. Sustain. 36, 126–140.
Alilou, H., Moghaddam Nia, A., Keshtkar, H., Han, D., Bray, M., 2018. A cost-effective and efficient framework to determine water quality monitoring network locations. Sci. Total Environ. 624, 283–293.
Arabi, B., Salama, M.S., Pitarch, J., Verhoef, W., 2020. Integration of in-situ and multi-sensor satellite observations for long-term water quality monitoring in coastal areas. Remote Sens. Environ. 239, 111632.
Aurin, D., Mannino, A., Franz, B., 2013. Spatially resolving ocean color and sediment dispersion in river plumes, coastal systems, and continental shelf waters. Remote Sens. Environ. 137, 212–225.
Barnes, B.B., Hu, C., Holekamp, K.L., Blonski, S., Spiering, B.A., Palandro, D., Lapointe, B., 2014. Use of Landsat data to track historical water quality changes in Florida keys marine environments. Remote Sens. Environ. 140, 485–496.
Bendoricchio, G., De Boni, G., 2005. A water-quality model for the lagoon of Venice, Italy. Ecol. Model. 184, 69–81.
Bodansky, D., 1993. The United Nations framework convention on climate change: a commentary. Yale J. Int'l l. 18, 451.
Brus, D.J., Heuvelink, G.B., 2007. Optimization of sample patterns for universal kriging of environmental variables. Geoderma 138, 86–95.
Brus, D.J., Knotters, M., 2008. Sampling design for compliance monitoring of surface water quality: a case study in a polder area. Water Resour. Res. 44.
Burrough, P.A., McDonnell, R., McDonnell, R.A., Lloyd, C.D., 2015. Principles of Geographical Information Systems. Oxford university press.
Caeiro, S., Painho, M., Goovaerts, P., Costa, H., Sousa, S., 2003. Spatial sampling design for sediment quality assessment in estuaries. Environ. Model Softw. 18, 853–859.
Catherine, A., Troussellier, M., Bernard, C., 2008. Design and application of a stratified sampling strategy to study the regional distribution of cyanobacteria (Ile-de-France, France). Water Res. 42, 4989–5001.
Chaney, N.W., Roundy, J.K., Herrera-Estrada, J.E., Wood, E.F., 2015. High-resolution modeling of the spatial heterogeneity of soil moisture: applications in network design. Water Resour. Res. 51.1, 619–638.
Chen, Q., Wu, W., Blanckaert, K., Ma, J., Huang, G., 2012. Optimization of water quality monitoring network in a large river by combining measurements, a numerical model and matter-element analyses. J. Environ. Manag. 110, 116–124.
Crepin, J., Johnson, R.L., 1993. Soil sampling for environmental assessment. Soil sampling and methods of analysis 5–18.
Croft, M.V., Chow-Fraser, P., 2009. Non-random sampling and its role in habitat conservation: a comparison of three wetland macrophyte sampling protocols. Biodivers. Conserv. 18, 2283–2306.
De Stefano, L., 2010. Facing the water framework directive challenges: a baseline of stakeholder participation in the European Union. J. Environ. Manag. 91, 1332–1340.
Dekker, A.G., Vos, R.J., Peters, S.W.M., 2001. Comparison of remote sensing data, model results and in situ data for total suspended matter (TSM) in the southern Frisian lakes. Sci. Total Environ. 268, 197–214.
Destouni, G., Fischer, I., Prieto, C., 2017. Water quality and ecosystem management: data-driven reality check of effects in streams and lakes. Water Resour. Res. 53, 6395–6406.
Dronova, I., Gong, P., Clinton, N.E., Wang, L., Fu, W., Qi, S., Liu, Y., 2012. Landscape analysis of wetland plant functional types: the effects of image segmentation scale, vegetation classes and classification methods. Remote Sens. Environ. 127, 357–369.
Feng, L., Hu, C., Chen, X., Tian, L., Chen, L., 2012. Human induced turbidity changes in Poyang Lake between 2000 and 2010: Observations from MODIS. Journal of Geophysical Research 117.
Feng, L., Hu, C.M., Chen, X.L., Zhao, X., 2013. Dramatic inundation changes of China's two largest freshwater lakes linked to the three gorges dam. Environmental Science & Technology 47, 9628–9634.
Fujiang, Z., 2017. A review of development of environmental protection information during the 12th five-year plan period and its outlook. Meteorological and Environmental Research 8, 67.
Ga, U., 2015. Transforming our World: The 2030 Agenda for Sustainable Development. Division for Sustainable Development Goals, New York, NY, USA.

Gebremichael, M., Krajewski, W.F., 2005. Effect of temporal sampling on inferred rainfall spatial statistics. J. Appl. Meteorol. 44, 1626–1633.
van Groenigen, J.W., 2000. The influence of variogram parameters on optimal sampling schemes for mapping by kriging. Geoderma 97, 223–236.
Guo, Y., Wang, S., 2014. Construction and Exploration of Ecolo-Hydrological Monitoring System in the Poyang Lake.
Haining, R., 2015. Spatial sampling. In: Wright, J.D. (Ed.), International Encyclopedia of the Social & Behavioral Sciences, Second edition Elsevier, Oxford, pp. 185–190.
Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: a k-means clustering algorithm. J. R. Stat. Soc.: Ser. C: Appl. Stat. 28, 100–108.
Hering, J.G., 2017. Managing the 'monitoring imperative'in the context of sdg target 6.3 on water quality and wastewater. Sustainability 9, 1572.
Ho, J.C., Michalak, A.M., Pahlevan, N., 2019. Widespread global increase in intense lake phytoplankton blooms since the 1980s. Nature 574, 667–670.
Hou, X., Feng, L., Duan, H., Chen, X., Sun, D., Shi, K., 2017. Fifteen-year monitoring of the turbidity dynamics in large lakes and reservoirs in the middle and lower basin of the Yangtze River, China. Remote Sens. Environ. 190, 107–121.
Ibrahim, E., Adam, S., De Wever, A., Govaerts, A., Vervoort, A., Monbaliu, J., 2014. Investigating spatial resolutions of imagery for intertidal sediment characterization using geostatistics. Cont. Shelf Res. 85, 117–125.
IOCCG, 2000. Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters. IOCCG Dartmouth, Canada.
IOCCG, 2013. Mission Requirements for Future Ocean-Colour Sensors. IOCCG Dartmouth, Canada.
Journel, A., Huijbregts, C., 1991. Mining Geostatistics. Academic Press Limited, London 600 p.
Karabork, H., 2009. Selection of appropriate sampling stations in a lake through mapping. Environ. Monit. Assess. 163, 27–40.
Karamouz, M., Kerachian, R., Akhbari, M., Hafez, B., 2009a. Design of river water quality monitoring networks: a case study. Environmental Modeling & Assessment 14, 705.
Karamouz, M., Nokhandan, A.K., Kerachian, R., Maksimovic, Č., 2009b. Design of on-line river water quality monitoring systems using the entropy theory: a case study. Environ. Monit. Assess. 155, 63.
Kiefer, I., Odermatt, D., Anneville, O., Wüest, A., Bouffard, D., 2015. Application of remote sensing for the optimization of in-situ sampling for monitoring of phytoplankton abundance in a large lake. Sci. Total Environ. 527, 493–506.
Kirkpatrick, S., Toulouse, G., 1985. Configuration space analysis of travelling salesman problems. J. Phys. 46, 1277–1292.
Koparan, C., Koc, A.B., Privette, C.V., Sawyer, C.B., 2018. In situ water quality measurements using an unmanned aerial vehicle (UAV) system. Water 10, 264.
Krapivin, V.F., Varotsos, C.A., Nghia, B.Q., 2017. A modeling system for monitoring water quality in lagoons. Water Air Soil Pollut. 228, 397.
Letcher, R.A., Jakeman, A.J., Calfas, M., Linforth, S., Baginska, B., Lawrence, I., 2002. A comparison of catchment water quality models and direct estimation techniques. Environ. Model Softw. 17, 77–85.
Li, J., Tian, L., Chen, X., li, x., Huang, J., Lu, J., & Feng, L., 2014. Remote sensing monitoring for spatio-temporal dynamics of sand dredging activities at Poyang Lake in China. Int. J. Remote Sens. 35, 6004–6022.
Li, J., Chen, X., Tian, L., Huang, J., Feng, L., 2015. Improved capabilities of the Chinese high-resolution remote sensing satellite GF-1 for monitoring suspended particulate matter (SPM) in inland waters: radiometric and spatial considerations. ISPRS J. Photogramm. Remote Sens. 106, 145–156.
Li, J., Tian, L., Song, Q., Sun, Z., Yu, H., Xing, Q., 2018. Temporal variation of chlorophyll-a concentrations in highly dynamic waters from unattended sensors and remote sensing observations. Sensors 18, 2699.
Liu, Y., Islam, M.A., Gao, J., 2003. Quantification of shallow water quality parameters by means of remote sensing. Prog. Phys. Geogr. 27, 24–43.
Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21, 1087–1092.
Michalcová, D., Lvončík, S., Chytrý, M., Hájek, O., 2011. Bias in vegetation databases? A comparison of stratified-random and preferential sampling. J. Veg. Sci. 22, 281–291.
Öhman, K., Eriksson, L.O., 2002. Allowing for spatial consideration in long-term forest planning by linking linear programming with simulated annealing. For. Ecol. Manag. 161, 221–230.
Ouyang, Y., 2005. Evaluation of river water quality monitoring stations by principal component analysis. Water Res. 39, 2621–2635.
Palmer, S.C.J., Kutser, T., Hunter, P.D., 2015. Remote sensing of inland waters: challenges, progress and future directions. Remote Sens. Environ. 1–8.
Park, S.-Y., Choi, J.H., Wang, S., Park, S.S., 2006. Design of a water quality monitoring network in a large river system using the genetic algorithm. Ecol. Model. 199, 289–297.
Pohl, C., Van Genderen, J.L., 1998. Review article Multisensor image fusion in remote sensing: Concepts, methods and applications. International Journal of Remote Sensing 19 (5), 823–854. https://doi.org/10.1080/014311698215748.
Pule, M., Yahya, A., Chuma, J., 2017. Wireless sensor networks: a survey on monitoring water quality. Journal of applied research and technology 15, 562–570.
Read, E.K., Carr, L., De Cicco, L., Dugan, H.A., Hanson, P.C., Hart, J.A., Kreft, J., Read, J.S., Winslow, L.A., 2017. Water quality data for national-scale aquatic research: the water quality portal. Water Resour. Res. 53, 1735–1745.
Tammi, J., Lappalainen, A., Mannio, J., Rask, M., Vuorenmaa, J., 1999. Effects of eutrophication on fish and fisheries in Finnish lakes: a survey based on random sampling. Fish. Manag. Ecol. 6, 173–186.
Tobler, W., 2004. On the first law of geography: a reply. Ann. Assoc. Am. Geogr. 94, 304–310.
UNEP, 2016. Snapshot of the World's Water Quality: Towards a Global Assessment. United Nations Environment Programme, Nairobi.
USEPA, 2017. National Water Quality Inventory—Report to Congress.
Van Groenigen, J., 1997. Spatial Simulated Annealing for Optimizing Sampling. geoENV I—Geostatistics for Environmental Applications. Springer, pp. 351–361.
Van Niel, K., Laffan, S.W., 2003. Gambling with randomness: the use of pseudo-random number generators in GIS. Int. J. Geogr. Inf. Sci. 17, 49–68.
Varotsos, C.A., Krapivin, V.F., 2018. Pollution of Arctic waters has reached a critical point: an innovative approach to this problem. Water Air Soil Pollut. 229, 343.
Varotsos, C.A., Krapivin, V.F., Mkrtchyan, F.A., 2019. New optical tools for water quality diagnostics. Water Air Soil Pollut. 230, 177.
Varotsos, C.A., Krapivin, V.F., Mkrtchyan, F.A., Gevorkyan, S.A., Cui, T., 2020. A novel approach to monitoring the quality of lakes water by optical and modeling tools: Lake Sevan as a case study. Water Air Soil Pollut. 231, 435.
Wang, J.-F., Stein, A., Gao, B.-B., Ge, Y., 2012. A review of spatial sampling. Spatial Statistics 2, 1–14.
Wang, Y.-B., Liu, C.-W., Liao, P.-Y., Lee, J.-J., 2014. Spatial pattern assessment of river water quality: implications of reducing the number of monitoring stations and chemical parameters. Environ. Monit. Assess. 186, 1781–1792.
Wang, C.M., Meng, Q.Y., Zhan, Y.L., Peng, J., Wei, X.Q., Yang, J., Li, J., 2015. Ground sampling methods for surface soil moisture in heterogeneous pixels. Environ. Earth Sci. 73, 6427–6436.
Wu, G., Cui, L., 2008. Remote sense-based analysis of sand dredging impact on water clarity in Poyang Lake. Acta Ecol. Sin. 28, 6113–6120.
Wu, W., Chen, Q., Li, J., Chen, G., 2010. Optimization of river water quality monitoring sections. Acta Sci. Circumst. 30, 1537–1542.
Wu, G., Cui, L., Duan, H., Fei, T., Liu, Y., 2013. An approach for developing Landsat-5 TM-based retrieval models of suspended particulate matter concentration with the assistance of MODIS. ISPRS J. Photogramm. Remote Sens. 85, 84–92.
Wu, X., Wen, J., Xiao, Q., Liu, Q., Peng, J., Dou, B., Li, X., You, D., Tang, Y., Liu, Q., 2016. Coarse scale in situ albedo observations over heterogeneous snow-free land surfaces and validation strategy: a case of MODIS albedo products preliminary validation over northern China. Remote Sens. Environ. 184, 25–39.
Zeng, Y., Li, J., Liu, Q., Li, L., Xu, B., Yin, G., Peng, J., 2014. A Sampling Strategy for Remotely Sensed LAI Product Validation over Heterogeneous Land Surfaces.