Frontiers of Environmental Science & Engineering
2025, 19(5): 69 https://doi.org/10.1007/s11783-025-1989-4

RESEARCH ARTICLE

A reagent-free on-site COD monitoring by a variable optical path UV-Vis spectrometer

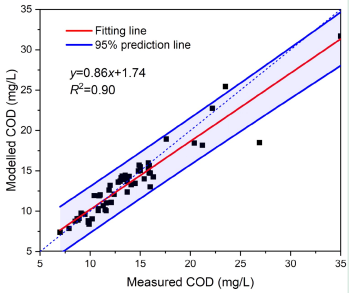
Xinnan Qian[1], Qiyun Zhu[1], Wenjun Sun[2], Xiaohong Zhou[1], Xiangyun Xiong[2], Siyu Zeng ✉ [1], Jianwu Sheng ✉ [3], Miao He[1]

1. School of Environment, Tsinghua University, Beijing 100084, China
2. Shenzhen Ecological and Environmental Monitoring Center of Guangdong Province, Shenzhen 518049, China
3. School of Municipal and Surveying Engineering, Hunan City University, Yiyang 413000, China

HIGHLIGHTS

- This spectrometer is characterized by variable optical path.
- Versatility and detection range significantly broadened for diverse applications.
- Enhanced sensitivity and accuracy for detecting subtle compositional changes.
- Ideal for monitoring water samples with diverse turbidity and concentration levels.

A variable optical path UV-Vis spectrometer



Field validation in the nearshore waters of Shenzhen Bay

ABSTRACT: The variation in pollutant concentrations among different water bodies poses a significant challenge for environmental surveillance. Traditional UV-Vis spectrometers, with fixed optical paths, face limitations in accurately determining Chemical Oxygen Demand (COD) and other water quality parameters. High concentrations surpass the detection limit, while low concentrations yield weak response signals, thereby compromising measurement accuracy. This study tackles these challenges by enhancing a UV-Vis spectrometer with a variable optical path. By utilizing a right-angle reflector for reflection and a stepping motor for control, measurements are conducted within the wavelength range of 190–700 nm. The instrument incorporates a spectral fusion algorithm to optimize spectral measurements within its operational range. Furthermore, a Partial Least Squares (PLS) model has been established for COD inversion by using laboratory standard solutions and field samples. The spectrometer has been tested in the nearshore waters of Shenzhen Bay, China, validating its applicability and the model's accuracy. The utilization of a variable optical path UV-Vis spectrometer facilitates the acquisition of precise monitoring data with wide measuring range, thereby enabling the prompt detection of anomalies and subsequent reduction in reaction time.

KEYWORDS: Variable optical path, UV-Vis spectrometer, COD, Reagent-free

---

✉ Corresponding authors. E-mails: shengjw@tsinghua.edu.cn (J. Sheng); szeng@tsinghua.edu.cn (S. Zeng)

Article history: Received 20 October 2024, Revised 9 January 2025, Accepted 25 February 2025, Available online 24 March 2025

1 Introduction

Water quality monitoring currently relies on collecting water samples locally and then transporting them to a laboratory for analysis, employing various chemical methods and conventional instrumental analytical techniques. As an example, the dichromate method can be used to determine the chemical oxygen demand (COD) content in water, and instruments such as high-performance liquid chromatography (HPLC), gas chromatography–mass spectrometry (GC–MS), and liquid chromatography tandem–mass spectrometry (LC–MS/MS), can be utilized to detect emerging pollutants like endocrine-disrupting chemicals (EDCs), toxicity, antibiotics etc. (Ji et al., 2020; Tan et al., 2022; Ji et al., 2024; Yang et al., 2024a; 2024b). Despite their accuracy and low detection limits, these traditional methods require expensive, sophisticated instrumentation and specialized personnel, which limits their applicability for online and real-time monitoring. Additionally, such methods are time-consuming and resource-intensive. Regular field sampling and chemical analysis not only present logistical challenges but also pose significant operational risks, and personnel safety, especially in remote or hazardous environments, remains a concern (Xing et al., 2022). Potential exposure to hazardous chemicals, biological agents, or contaminated environments may compromise the health and safety of field personnel (Kang et al., 2023), thus highlighting the need for reducing sampling frequency to minimize risk where feasible and exploring alternative, more efficient monitoring approaches.

Among the array of optical sensor technologies that are currently garnering significant attention, UV-Vis spectrometry stands out as a pivotal tool for quantifying pollutants. This analytical technique presents distinct advantages, including reagent-free operation, rapid detection capabilities, and the potential for continuous monitoring. Its prominence stems from its ability to utilize well-established models, correlating absorbance values with pollutant concentrations. UV-Vis spectrometers function on the principle of absorbance, quantifying the amount of light absorbed by a sample across the UV and visible spectrum. Within this range, compounds in water samples showcase distinctive absorption patterns, enabling the quantification of parameters such as COD, nitrate, total suspended solids (TSS), chlorophyll, and various others (Chen et al., 2014; Carré et al., 2017; Li and Hur, 2017; Dahlén et al., 2000; Sanchini and Grosjean, 2020; Li et al., 2023). This characteristic absorption behavior forms the basis for precise and reliable COD determination using UV-Vis spectrometry. By capitalizing on these advantages, UV-Vis spectrometry significantly improves the efficiency and efficacy of environmental monitoring practices across various applications. Furthermore, the versatility of UV-Vis spectrometry extends beyond simple pollutant detection. Its adaptability to various environmental matrices, coupled with advancements in instrument sensitivity and miniaturization, positions this technology for broader implementation in field-based monitoring systems. Continuous real-time monitoring of water bodies can now be achieved with portable or *in situ* UV-Vis spectrometers, offering immediate feedback on water quality changes without the need for extensive laboratory analysis. The cost-effectiveness, reduced sample preparation time, and non-destructive nature of UV-Vis spectrometry make it particularly attractive for long-term environmental surveillance, especially in regions where conventional monitoring methods may be challenging. As a result, UV-Vis spectrometry has gained recognition as a reliable method for measuring pollutants within the domain of water quality monitoring (Lepot et al., 2016).

Depending on the anticipated concentration range of the investigated water matrix, spectrometers employ optical paths typically pre-determined within the range of 0.5 to 100 mm. Although spectrometers with fixed optical paths are well-regarded, they face challenges when dealing with high-concentration samples, often encountering issues associated with exceeding-range measurements. Conversely, when analyzing low-concentration samples, these instruments may struggle with low signal-to-noise ratios. In the same geographical area, the concentrations of pollutants in various water matrices, such as rivers, lakes and pollution sources, can exhibit significant variability. Moreover, within certain studies, notable disparities in water matrix composition have been observed even among different segments of the same river and lake. For instance, TSS in four fish farming ponds situated in Glória de Dourados city range from 3 to 72 mg/L (Figueiró et al., 2018). Similarly, measurements of nitrate concentrations ranged from 2.46 to 164.91 mg/L in well samples collected from a small town (Savci, 2012). Additionally, within a drainage system located in the Lisbon district, COD concentrations ranges from 220 to 800 mg/L (Brito et al., 2014). Given the wide-ranging concentrations of pollutants in such diverse water matrices, precise measurement necessitates meticulous consideration of spectrometer design, particularly regarding the length of the optical path. Fixed optical path equipment may lack the necessary accuracy for reliable measurements across diverse pollutant concentrations. In contrast, spectrometers equipped with variable optical paths offer a more suitable solution under such varied working conditions.

Variable optical path UV-Vis spectrometry represents a cutting-edge advancement in analytical instrumentation and finds widespread application in fields such as aerospace, medicine (Tang et al., 2019) etc, where precision in measurement is critical. The use of spectral fusion algorithms in conjunction with variable optical paths offers significant advantages in data processing, as it allows for the combination of spectral data from different optical paths, resulting in improved measurement accuracy and reduced noise (Thomas and Burgess, 2017). Implementing variable optical paths in spectrometers significantly expands the range of pollutant concentrations that can be accurately measured, thus providing more versatile applications in environmental monitoring, industrial processes, and biomedical diagnostics. However, achieving variable optical paths often necessitates mechanically adjusting the positions of optical elements (Wang et al., 2019a), presenting four key challenges for device development: the susceptibility to optical system misalignment due to mechanical vibrations; the potential for increased instrument volume due to additional mechanical drives; the complexity of achieving waterproofing for moving parts and the necessity for the development of suitable spectrum fusion algorithms. Each of these challenges not only impacts the performance and reliability of the spectrometer but also poses significant hurdles in the context of field deployment, where conditions can be less than ideal. Hence, there is a pressing need to engineer a compact, reliable, waterproof spectrometer equipped with a variable optical path.

To our knowledge, no existing studies have employed a variable optical path UV-Vis spectrometer for on-site measurement of COD concentrations in surface water, sewage water, and industrial water. Considering the advantageous features of UV-Vis spectrometry, including high speed, low cost, and independence from specialized laboratory facilities, it is of practical significance to evaluate its applicability in this context. Therefore, the primary objectives of this study are twofold: first, to develop a variable optical path UV-Vis spectrometer specifically designed to acquire spectra from water samples with widely varying pollutant concentrations; and second, to establish local calibration models that will enable accurate measurement of pollutant concentrations across different aqueous environments. By addressing these objectives, this research aims to enhance the reliability and efficiency of water quality monitoring, thereby contributing to improved environmental management practices.

2 Materials and methods

2.1 Structure design of a variable optical path UV-Vis spectrometer

The variable optical path UV-Vis spectrometer developed in this study is constructed with a robust stainless-steel shell, complemented by a Teflon top cover and Teflon base, ensuring both waterproofing and exceptional corrosion resistance. Notably, this spectrometer is fully submersible in water, with dimensions of 400 mm in length and 72 mm in diameter, making it highly suitable for various aquatic applications. The design encompasses a UV-Vis spectral range that spans from 190 to 700 nm, enabling the analysis of a wide variety of compounds.

To meet the wavelength requirements traditionally covered by both deuterium and tungsten lamps, this spectrometer utilizes Xenon lamps. This choice not only streamlines the overall equipment design but also simplifies the control system, enhancing user accessibility. Furthermore, the use of Xenon lamps offers additional advantages, including broader spectral coverage, higher intensity output, and improved stability across a wide range of operating conditions. These features collectively enhance the overall performance and versatility of the spectrometer, positioning it as an efficient and reliable solution for diverse analytical applications in environmental monitoring and beyond. Internal details of the variable optical path spectrometer are shown in Fig. 1(a), which demonstrates the innovative operating principle of the variable optical path mechanism, underscoring its significance in advancing measurement accuracy and flexibility.

Moreover, in contrast to certain spectrometers where incident and reflected light are at specific angles, this spectrometer aligns incident light parallel to reflected light (Fig. 1(b)). This configuration reduces optical collimation challenges, enhances optical energy reception efficiency, and extends the optical path range. The adjustment of the optical path is accomplished by varying the distance between the right-angle reflector and the light source. This self-adjusting structure is mainly achieved by a small stepping motor and a rod, as shown in Fig. 1(a). The precise positioning of the angle reflector is monitored through voltage readings

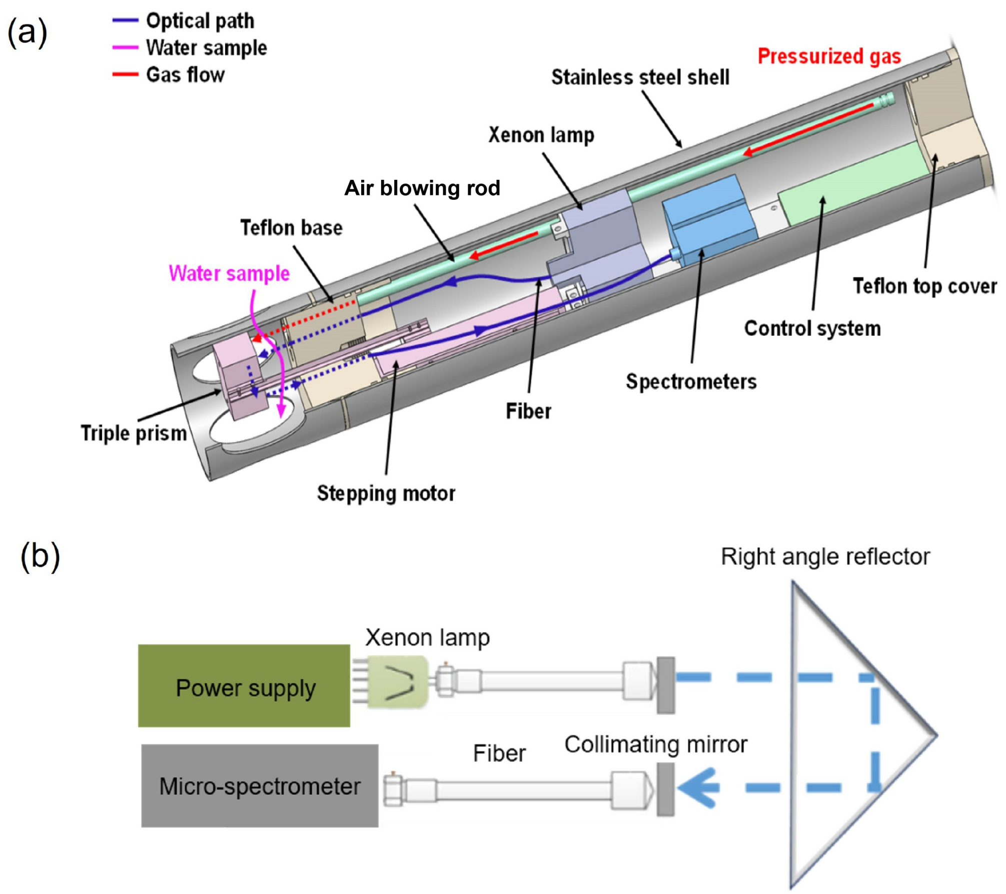


Fig. 1 Optical path optimization and reflection mechanisms in the self-developed UV/Vis spectrometry. (a) Internal details of the variable optical path UV/Vis spectrometer, (b) secondary reflection path of right-angle mirror.

from a voltage sensor, enabling accurate alignment of the optical components. During each motion adjustment, the optical path undergoes a reset process to ensure that the parallel alignment is maintained for continuous automatic monitoring. The control procedure comprises two scenarios: Firstly, if the occluder does not intercept the photoelectric switch upon receiving the detection instruction, the motor ascends until it contacts the switch, proceeding until it reaches the “reset distance.” Secondly, if the occluder obstructs the switch, prompting a signal interruption, the motor descends until the obstruction is removed, triggering a signal transition. It then halts and initiates the reset process from the first scenario. In both scenarios, the final step in resetting the stepping motor is executed when the photoelectric switch transitions from an unblocked to a blocked state, effectively minimizing system errors and enhancing the overall reliability of the spectrometer. This meticulous design approach ensures that the spectrometer maintains optimal performance, providing accurate measurements in varied operational conditions.

The change in optical path length is achieved by adjusting the position of the reflective prism. The position of the emitting light source and the position of the receiver are fixed, further we control the stepper motor through program control to drive the reflective prism to move up and down, when the moving distance becomes longer, the optical path increases, the light travels in the water sample increases, and the actual optical path increases twice the moving distance of the triangular prism.

The purpose of the tracheae is to direct high-pressure gas to the vicinity of the light window, and to blow the surface of the light window through a small hole to achieve the purpose of cleaning the light window.

As a result, the optical path of this spectrometer can be adjusted between 1 and 100 mm with the self-adjusting structure. Within each testing process, the spectrums for optical path 1, 10, and 100 mm can be generated, and combined together to obtain the fusion spectrum via spectrum synthesis algorithm.

2.2 Spectrum fusion algorithm

During the absorption spectrum detection across continuous wavelengths, a fixed optical path may not guarantee optimal absorbance performance for all wavelengths. This limitation can result in inaccurate readings and diminished sensitivity in certain spectral regions. To address this challenge, the variable optical path UV-Vis spectrometer developed in this study incorporates a self-adjusting function, enabling the acquisition of spectra at different optical path lengths (1, 10, and 100 mm).

Generally, when absorbance detection is performed in the laboratory, the detection result under a single optical path within the optimal absorbance detection range of the instrument will be selected by comparing the absorbance values under multiple optical paths, which represents the nature of the sample. For absorbance detection of a single wavelength, this detection method can obtain the best absorbance detection result for that wavelength, but if it involves multiple wavelengths of absorbance (or continuous absorption spectra of wavelengths), this method does not guarantee that there is an optical path length that can achieve the best detection of absorbance at different wavelengths. Generally, in the ultraviolet short wavelength region, it is easier to cause saturation absorption, and this happens to be a sensitive area for organic matter absorption, so it is advisable to use the short optical path mode; In the long-wavelength region of the near-infrared, absorption is not easily saturated, and parameters such as higher concentrations of suspended solids and turbidity may be measured. In summary, accurate acquisition of absorption spectral shape is particularly important in the functional requirements for pollution identification based on spectral shape features. Only by ensuring that the absorbance of each wavelength is within the optimal detection range of the instrument can critically shape information, such as the ratio of absorption peaks at different wavelengths, be identified. In other words, by reasonably controlling the optical path and selecting the optimal detection range, the short optical path low signal-to-noise ratio signal will be replaced by a good signal of the same wavelength with long optical path.

Basic principles and algorithm description of fusion algorithm: spectral fusion is to intercept the optimal part of the spectrum of different optical paths and splice them to form a complete spectral curve. It is important to exclude the portion of the spectrum measured for each path length that exceeds the detection limit and ensure that the remaining optimal detection interval covers all wavelength ranges. Due to the existence of baseline drift and random error, it is necessary to leave overlapping parts in the optimal detection interval of adjacent optical paths for fitting correction to prevent abrupt changes at the splicing site. The synthesis method of multi-path absorption spectroscopy needs to be based on Bill–Lambert’s law, and through reasonable process design and algorithm processing, to obtain high-fidelity synthesis spectra.

According to the principle and goal of variable path spectral synthesis analyzed in the previous section, the spectral synthesis was carried out in three steps:

1) Configure the optical path gradient according to the optimal absorbance detection interval.

2) Absorption spectra are collected and normalized to obtain spectra that can be spliced together.

3) Multi-path spectroscopic synthesis. Identify the overlap between the optimal range of different optical path spectra, and use the final spectrum of the overlap as a benchmark to calculate the average value of the corrected spectral overlap of each optical path after averaging. The correction parameters of the non-overlapping part of the optical path in the middle are the differential values of the correction parameters at both ends, so that the overlapping part and the non-overlapping part have better continuity in the complete spectrum. In this process, the mean correction of the constant baseline drift that occurs during different pathlength detection is also realized.

2.3 COD inversion model

The data processing and modeling involved in this study are coded in the Matlab R2020b environment. The COD chemical measurement method in this study is according to HJ 828 by the Ministry of Ecology and Environment of the People’s Republic of China.

Each test conducted by the spectrometer generated 208 absorbance values across a wavelength range of 190 to 700 nm, with intervals of 2.5 nm. These absorbance values, combined with the corresponding COD measurements obtained from laboratory analyses, constituted a comprehensive data set. In total, 53 such data sets were collected from a variety of water samples and subsequently divided into two groups. Two-thirds of the data sets were designated as a training set for the establishment of the COD inversion model, while the remaining one-third was reserved for validation to assess the model’s performance. Prior to modeling, all spectral data underwent preliminary data cleaning and principal component analysis (PCA) to eliminate redundant information, thereby aiding in the reduction of the number of model input variables (Thomas and Burgess, 2017; Guan et al., 2018).

Originally developed in the field of chemometrics, the Partial Least Squares (PLS) model has emerged as a versatile and widely applicable method for modeling complex relationships between predictor variables and response variables (Singh et al., 2007; Song et al., 2012; Nakar et al., 2020). Unlike traditional regression methods, Partial Least Squares (PLS) considers both predictor variables and response variables simultaneously, allowing for the extraction of latent factors or components that encapsulate the underlying structure of the data. As a result, PLS is particularly effective in situations with many correlated predictor variables or when dealing with high-dimensional data sets. Therefore, recognized as the most prevalent model for simulating COD concentrations, the PLS model is adopted in this study. Lepot et al. (2016) demonstrated that the number of variables in PLS model establishment can range from a few to thousands, indicating its flexibility in handling diverse data sets. Consequently, the performance of the PLS model were compared using three different sets of input variables: the 4 wavelengths exhibiting the strongest correlation with COD concentrations; 40 wavelengths between 302.5 and 400 nm, and all 208 wavelengths individually.

The accuracy of inversion model was evaluated using the parameters including the coefficient of determination (R^2), the mean relative error (MRE), the root means square error ($RMSE$), and the relative RMSE ($rRMSE$).

$$MRE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{C_{\mathrm{meai}}-C_{\mathrm{modi}}}{C_{\mathrm{meai}}}\right|, \tag{1}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(C_{\mathrm{meai}}-C_{\mathrm{modi}}\right)^2}, \tag{2}$$

$$rRMSE = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(C_{\mathrm{meai}}-C_{\mathrm{modi}}\right)^2}}{\frac{\sum_{i=1}^{N}C_{\mathrm{meai}}}{N}}\times 100\%, \tag{3}$$

where C_{meai} represents the measured COD concentration, C_{modi} denotes the modeled COD concentration, N is the number of data points.

3 Results

3.1 COD concentrations of the field samples

The summary statistics for COD concentrations from three different type of water samples showed that, significant differences are observed between the data sets collected from different type of samples. COD concentration is from 10 to 6060 mg/L, covering three orders of magnitude, which leads to challenges to the equipment performance. The COD concentration is obviously lower in surface water (27.80 mg/L) than sewer water (644.47 mg/L) and industrial wastewater (680.54 mg/L). Large standard deviations of industrial water (1162.84 mg/L) and sewer water (1150.70 mg/L) reveal that water matrix varies greatly between samples.

3.2 Fusion spectrum

To test the performance of spectrum fusion algorithm in Section 2.2, we conducted experiments using a 50 mg/L potassium hydrogen phthalate solution as samples. The spectra of optical paths 1, 10, and 100 mm were measured separately, as illustrated in Fig. 2(a). The results reveal that, at the 10 and 100 mm optical paths, absorption at certain wavelengths surpasses the optimal detection interval, potentially compromising measurement accuracy. In contrast, the 1 mm optical path yields relatively low signal values, which may also limit the quality of the spectral data. Remarkably, the fused spectrum produced from these measurements demonstrates a smooth curve with well-defined peak characteristics, highlighting the efficacy of the fusion approach. Specifically, between 200 and 270 nm, portions of the spectrum that exceed the optimal interval are eliminated, while retaining and enhancing the original signal values under the 1 and 10 mm optical paths. This careful refinement process leads to more pronounced peak characteristics and significantly increases measurement accuracy, thereby providing a clearer and more reliable representation of the sample's absorptive properties.

The same spectrum fusion algorithm detection was performed on water samples from Jiangsu, China (Fig. 2), including surface water and industrial water. Figures 2(b) and 2(c) illustrates the spectra of industrial wastewater with a COD concentration of 6600 mg/L and surface water with a COD concentration of 10 mg/L. Noticeable differences in spectral characteristics between the two sample types are apparent. The absorbance of industrial wastewater is consistently higher than that of surface water, indicating a more complex compositional matrix associated with industrial discharges. Additionally, the spectrum of industrial wastewater demonstrates a pronounced decline in absorbance within the range of 200 to 250 nm, contrasting with the relatively flat spectrum observed for surface water. Furthermore, in all instances, distinct areas of exceedance are observed in

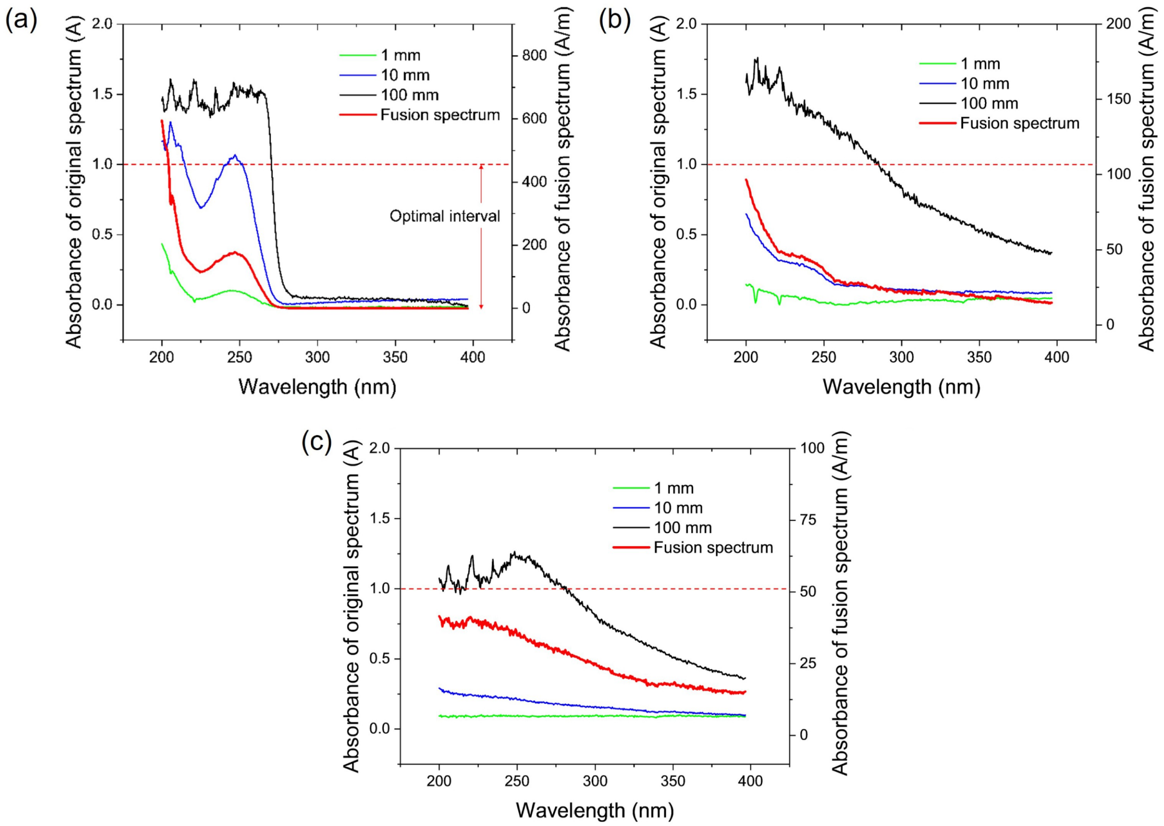


Fig. 2 Spectral analysis and sampling comparison of various water samples and standard solutions. (a) Spectrums of 50 mg/L potassium hydrogen phthalate solution, (b) spectrums of industrial water and (c) surface water.

the spectrum under the 100 mm optical path. However, following the fusion process, a fusion spectrum with the original characteristics is obtained, indicating the efficacy of the spectrum fusion algorithm for sample measurements.

In conclusion, the spectrum synthesis algorithm demonstrates significant efficacy in addressing challenges related to exceeding measurement ranges and low signal-to-noise ratios, thereby enhancing the reliability and robustness of the final measurements. By integrating spectra obtained from multiple optical paths, the algorithm effectively minimizes errors and uncertainties that may arise from individual measurements, leading to improved accuracy in quantifying analytes. Moreover, the algorithm offers flexibility in adjusting spectral range and parameters to cater to specific measurement needs, enhancing adaptability. This adaptability is particularly beneficial in dynamic environments where conditions may change frequently. Additionally, it compensates for environmental and instrument variations, ensuring consistent and reliable measurements across diverse scenarios.

3.3 Data processing

To facilitate data analysis, PCA was employed to condense 208 absorbance values into four principal components. The PCA results, shown in Fig. 3(a), indicate that the first principal component captures 90.4% of the variance, signifying strong correlation among input variables. This dimensionality reduction effectively addresses multicollinearity issues present among the 208 parameters, thereby allowing for the exclusion of less informative variables and facilitating the selection of optimal wavelengths through preliminary data analysis. Furthermore, Fig. 3(b) illustrates the correlation analysis results between COD concentrations and the UV-Vis spectrum, revealing that the sensitive band in this study spans from 302.5 to 402.5 nm (Correlation coefficient ⩾ 0.7). Specifically, the four variables with the strongest correlation are absorbances at 317.5, 330, 332.5, and 325 nm respectively. These findings not only substantiate the effectiveness of PCA in simplifying complex data sets but also emphasize its critical role in enhancing the accuracy of data interpretation.

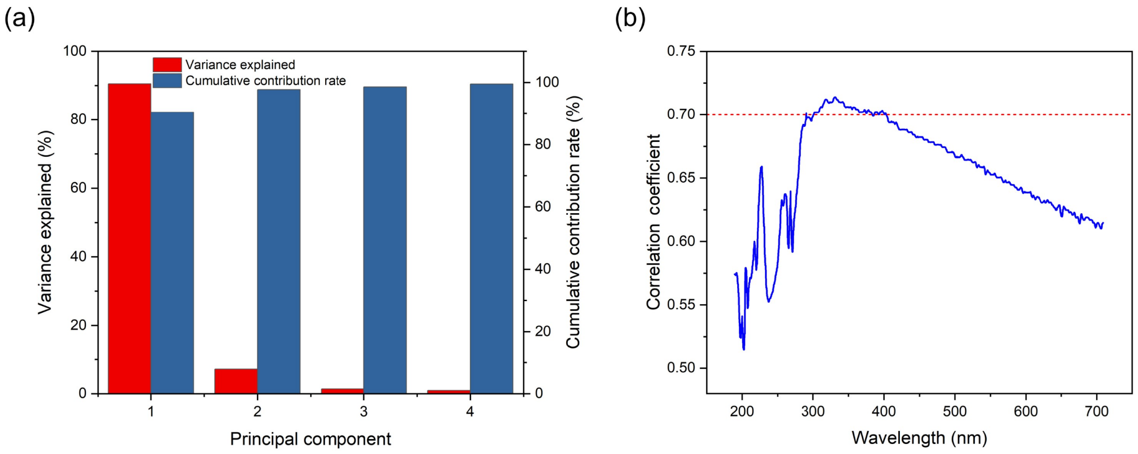


Fig. 3 Data processing. (a) Principal component analysis, (b) correlation coefficients of COD concentrations and absorption.

The insights gained from PCA serve as foundational support for the subsequent modeling processes, ensuring that the most relevant features are utilized in predictive modeling. By distilling the data into principal components that retain the majority of the variance, PCA facilitates a more focused approach to analysis, ultimately improving the reliability of predictions concerning COD concentrations.

3.4 PLS model building and validation

The PLS model, a multivariate regression technique, amalgamates aspects of PCA and multiple linear regression. Its theoretical foundation revolves around extracting latent variables, termed latent factors or components, from predictor variables (X) and response variables (Y) to characterize their relationship. The PLS method prioritizes the covariance between predictor variables (X) and response variables (Y), making it well-suited for situations where strong correlations exist between them. This focus on covariance enables PLS to effectively model both predictor and response variables simultaneously, enhancing its versatility and applicability across diverse domains.

In this study, the spectral data collected from the spectrometer serves as the predictor variable, while the COD value simulated by the PLS model serves as the response variable. Following a training period with known COD values, the PLS model provides accurate feedback ($R^2 \geqslant 0.90$). A linear regression equation between the COD values obtained from chemical analysis and the simulated values is established to assess the instrument's simulation efficacy, as depicted in Fig. 4.

As outlined in Section 2.3, separate PLS models were established using different input variables. Notably, the PLS model based on the four variables demonstrating the strongest correlation between COD concentrations and the UV-Vis spectrum was examined. The performance evaluation results indicate a high coefficient of determination ($R^2 = 0.90$), suggesting excellent agreement between measured results and modeled simulations across a wide range. The gap between the minimum and maximum values of the validation data set is approximately 6000, further highlighting the model's ability to explain variance in the response variable to a significant extent. These findings underscore the versatility and accuracy of the variable optical path spectrometer. To improve graph clarity, especially amidst dense low-concentration data and significant differences in high-concentration points, both measured and predicted results are plotted logarithmically, as illustrated in Fig. 4(a). Statistical metrics obtained from both the training and validation data sets are presented in Table 1, demonstrating the effectiveness of the PLS model.

In real-world scenarios, the model's applicability may be influenced by weather conditions, particularly rainfall, which can impact the volume and flow rate of water in rivers and sewage systems. Conversely, industrial water sources are typically less susceptible to environmental fluctuations due to consistent production procedures. Additional validation is warranted to expand the spectrometer's application range and enhance its flexibility. Moreover, ongoing optimization of these models can be achieved through frequent adjustments and retraining based on chemical measurements.

In December 2023, a field validation was taken in the nearshore waters of Shenzhen Bay to evaluate the robustness and generalizability of the PLS model. The study aimed to assess the model's performance across various validation data sets by simulating COD levels

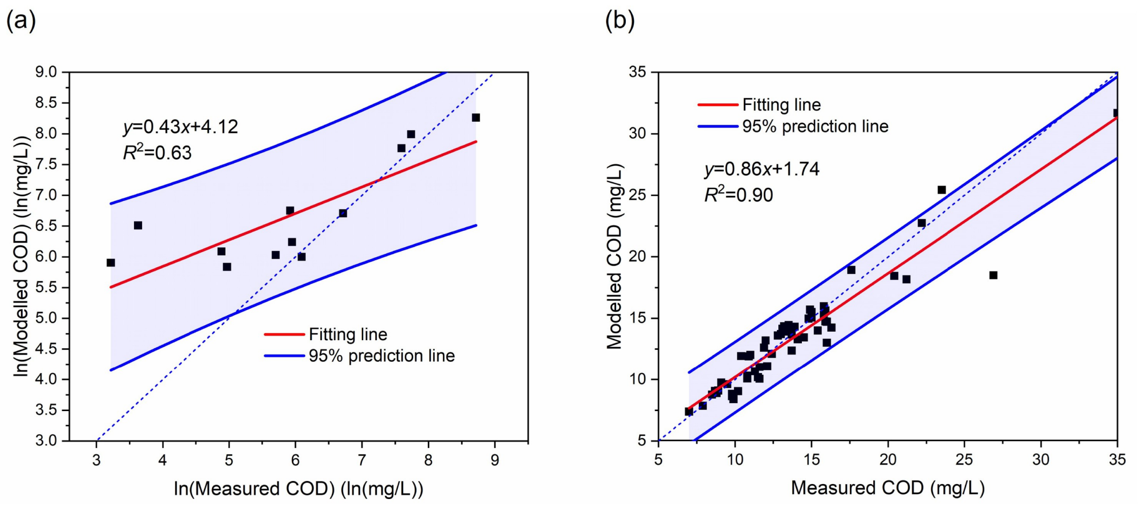


Fig. 4 Performance evaluation of the COD simulation using the PLS model. (a) Validation data from Jiangsu, China, (b) validation data from nearshore waters of Shenzhen Bay.

across a spectrum of water samples, including both oceanic and riverine environments. Automated sampling was conducted every two hours, with the results transmitted wirelessly to a PC for analysis.

The intricate dynamics of the ocean, including tides and other factors, contribute to significant daily fluctuations in concentration levels across a wide range. These unique characteristics of the ocean environment offer valuable opportunities to enhance the instrument's universality through comprehensive validation. Concentration levels at a unified sampling point demonstrate temporal variability throughout the day, fluctuating within a specific range. During the whole monitoring period, the concentrations span from a minimum of 7 mg/L to a maximum of 35 mg/L. The correlation between the COD concentrations simulated by the spectrometer and those determined through chemical analysis was depicted in Fig. 4(b), yielding an R^2 value of 0.90, indicating a strong linear association. This performance is consistent with previous validation data sets, affirming the reliability of the spectrometer's simulations. Statistical metrics derived from the validation data sets are summarized in Table 1, reaffirming the effectiveness of the PLS model.

The COD values of the Jiangsu water sample showed a higher root mean square error (RMSE) than the Shenzhen water sample. Through further experiments and analysis, the reasons were as follows:

1) the sample quantity of water samples in Jiangsu was obviously small.

2) The COD value of Jiangsu water samples varied widely, with the maximum value exceeding 6000 mg/L, while that of Shenzhen water samples was 7–35 mg/L.

3) The composition of water samples in Jiangsu is dominated by industrial wastewater, and the composition of different types of industrial wastewater is very different, and the absorption spectral characteristics are also very different. The principle of absorption spectroscopy to measure COD determines that it needs to be calibrated locally for the specific wastewater composition type of the point before it can be used in practice, and it needs to be calibrated again when it is replaced with another application. The Shenzhen water sample is a seawater and river water sample, and its concentration is not high, and the main components are relatively stable, so the fitting accuracy is very good.

Table 1 Statistical metrics of the data set

Metrics	Value		
	Training dataset	Validation dataset from Jiangsu	Validation dataset from Shenzhen
R^2	0.91	0.63	0.90
MRE (%)	2.87	2.73	7.17
RMSE (mg/L)	336.61	438.86	1.63
rRMSE (%)	7.63	19.65	11.78

In short, the specific factor that affects the reliability of the measurement results is the local calibration. When the water sample is a comprehensive water body such as a river, lake, ocean, etc., the fitting results of the model are very accurate. When the water sample is a different type of industrial wastewater with a large change in composition, local calibration is required, that is, the COD measured by the standard potassium dichromate oxidation method is used as the basis for the correction of the fitting COD value of the spectral probe.

The practical implementation within the nearshore waters of Shenzhen Bay validates the effectiveness of the variable optical path spectrometer. Its capability to swiftly detect and transmit high-frequency data establishes a solid foundation for potential future applications.

4 Discussion

The discharge and seepage of untreated industrial wastewater pose a significant challenge in water environments, acknowledged widely for their potential detrimental effects on ecosystems. The unregulated discharge of such wastewater poses a significant threat to the ecological equilibrium of surface water ecosystems. It leads to the contamination of groundwater, surface water bodies, and soil, while also undermining the effectiveness of sewage treatment facilities by impairing crucial bacterial processes. Furthermore, exposure to untreated industrial water pollutants presents grave health hazards to human populations. The public health implications of industrial wastewater pollution necessitate urgent and effective monitoring strategies to protect both environmental integrity and human well-being. To preemptively address this issue, onsite detection of COD using spectrometers in sensitive areas emerges as a pragmatic solution. Elevated COD concentrations detected in sewers or surface water often signal industrial discharge or leakage, enabling prompt intervention to mitigate pollution risks. By enabling real-time monitoring of water quality, this approach not only aids in the immediate identification of pollution sources but also informs regulatory compliance and environmental management efforts.

Traditional monitoring methods often entail time delays, yet many pollution incidents occur suddenly (Yi et al., 2020). Therefore, timely detection and swift response are imperative. On-site monitoring devices positioned at detection sites can significantly reduce the time needed to identify anomalies, enabling prompt action to mitigate adverse impacts. Henceforth, the future of local water quality monitoring holds vast potential for the utilization and proliferation of diverse sensor technologies employing photoelectric signals. In particular, optical sensors present an appealing alternative to conventional analytical techniques owing to their rapid response times, high sensitivity, and selectivity. These devices are capable of real-time detection, offering near-instantaneous results that enable faster intervention compared to laboratory-based methods (Zhou et al., 2013; Wang et al., 2019b).

Optical sensors also stand out for their cost-effectiveness, making them a more accessible option for continuous monitoring, particularly in resource-limited settings. Their ease of operation and relatively straightforward manufacturing process further enhance their practicality for widespread deployment in diverse environments. These advantages position sensor technologies as ideal tools for improving the efficiency and accuracy of water quality assessments. Furthermore, as advancements in materials science and nanotechnology continue to enhance the functionality and durability of these sensors, their ability to detect a broader range of contaminants—including emerging pollutants—will improve, ensuring more robust and adaptive environmental monitoring systems (Xia et al., 2019; Xue et al., 2022). Thus, optical sensors represent a key innovation in enabling swift, precise responses to environmental anomalies, facilitating more effective management of water resources in both urban and remote locations.

A common drawback of conventional spectrometers with fixed optical paths is their limited ability to detect wide-ranging COD concentrations, rendering their scope of application often narrow and undefined. As per prior research, these pre-fixed optical path spectrometers are predominantly utilized for detecting nitrogen only in slightly-polluted surface water and sea water without establishing COD simulation model (Kroeckel et al., 2011), or applied in studies of high-COD industrial effluents and low-COD treatment plant effluents (Langergraber et al., 2003). While Langergraber et al. (2003) did not specify the types of spectrometers used in their study, our experience suggests that a single spectrometer with a fixed optical path may encounter significant difficulties in accurately identifying spectral characteristics across the full range of concentration levels. It’s fortunate that the variable optical path UV/Vis spectrometer developed in this study demonstrates excellent performance across all samples collected from diverse locations, encompassing a wide range of COD concentrations. This versatility allows for more precise and reliable measurements, significantly streamlining the measurement process and enhancing the spectrometer’s applicability in real-world scenarios. The ability to adapt the optical path length according to the specific concentration of pollutants ensures that the spectrometer can maintain optimal performance, overcoming the limitations of traditional instruments.

The incorporation of variable optical path lengths represents a significant enhancement to the spectrometer's capabilities, notably increasing its versatility and enabling it to monitor concentrations across a wider spectrum of applications. This innovative feature is particularly beneficial for analyzing water samples characterized by varying turbidity and concentration levels, allowing for more nuanced evaluations of these samples. It facilitates a more thorough evaluation of such samples, thereby enhancing sensitivity and accuracy in detecting subtle compositional changes. Furthermore, the spectrometer is equipped with an embedded simulation model that provides real-time COD concentration values, thereby facilitating timely decision-making in various practical scenarios. The accuracy of the embedded model is continuously refined through periodic recalibration, utilizing actual COD data obtained from standard chemical analysis methods.

The importance of variable optical paths lies in their ability to dynamically adjust to the concentration levels of pollutants, ensuring optimal measurement accuracy across a broad spectrum of water quality scenarios. In high-concentration samples, a shorter optical path helps avoid issues like absorbance saturation, which can skew data and compromise the accuracy of pollutant quantification. On the other hand, for low-concentration samples, extending the optical path improves the sensitivity of the measurement by increasing the interaction between the light and the sample, thereby enhancing the signal-to-noise ratio. This flexibility makes variable-path spectrometers particularly well-suited for environments where pollutant levels fluctuate widely, such as industrial discharge sites, agricultural runoff, or complex aquatic ecosystems. Furthermore, variable-path systems enable more accurate and reliable long-term monitoring, as they can adapt in real time to the changing conditions of the water matrix. As environmental monitoring demands continue to evolve, the development and refinement of spectrometers with adjustable optical paths are likely to play a crucial role in advancing the precision and applicability of spectroscopic techniques in water quality assessment.

Additionally, the widespread adoption of spectrometers for water environment monitoring has been impeded by the necessity for stable AC 220 V power supply systems and cumbersome equipment. These limitations have often restricted the usability of such devices, particularly in remote or resource-limited settings where reliable electrical infrastructure may be lacking. However, the spectrometer developed in this study overcomes this limitation by utilizing a small battery for power, allowing for hand-held operation and significantly enhancing its portability and convenience. This innovative design not only alleviates the constraints associated with power supply but also promotes ease of use in various field conditions. By allowing for hand-held functionality, users can conduct real-time assessments of water quality without the encumbrance of heavy, fixed installations. Furthermore, the spectrum data, along with the analytical model, can be transmitted wirelessly to a mobile phone or PC, enhancing accessibility and usability. This feature enhances accessibility, making it easier for researchers, environmental professionals, and regulatory bodies to analyze and interpret data on-the-go.

The dual advantages of portability and wireless data transmission render this spectrometer a highly viable candidate for widespread implementation in water quality monitoring initiatives. Its user-centric interface and optimized operational functionalities enable a diverse array of stakeholders—including community organizations, regulatory agencies, and local governments—to actively participate in effective water management practices. By promoting the extensive application of this technology across various environmental contexts, it not only enhances the accuracy and efficiency of water quality assessments but also supports proactive strategies for pollution mitigation and ecological conservation. Consequently, the integration of this spectrometer into existing monitoring frameworks has the potential to significantly advance efforts in sustainable water resource management and environmental stewardship.

5 Conclusions

Our variable optical path UV/Vis spectrometer offers high temporal resolution, facilitating continuous monitoring at online stations. Moreover, its versatility in data acquisition enhances sensitivity and accuracy. Significantly, the continuous spectrum not only provides ongoing COD concentration values but also assists in analyzing pollution diffusion paths and identifying sources. This capability enables the identification and prediction of changes in the water environment, empowering decision-makers to proactively address unforeseen circumstances.

To sum up, the advancement of variable optical path spectrometers marks a substantial progress in analytical instrumentation, particularly within the context of environmental monitoring. The ability to conduct precise and efficient on-site detection, coupled with high-frequency monitoring of water quality, is crucial for ensuring water safety and preempting potential emergencies. By addressing the inherent limitations of traditional spectrometers, these innovative devices pave the way for more effective water quality evaluations and enhanced environmental monitoring practices. The integration of absorption spectrum data with hyperspectral information provides a comprehensive framework for monitoring rivers, lakes, oceans, and other extensive aquatic systems, fostering a holistic approach to water quality management. Look ahead, the potential synergy between these spectrometers and satellite remote sensing technologies holds promise for advancing our capabilities in environmental surveillance and management. This fusion is anticipated to lead to improved strategies for addressing emerging environmental challenges and promoting sustainable water resource practices.

CRediT Authorship Contribution Statement

Xinnan Qian: Formal analysis, Writing-original draft. **Jianwu Sheng:** Methodology, Writing-review and Funding acquisition. **Qiyun Zhu:** Methodology, Validation, Formal analysis, Data curation. **Xiaohong Zhou:** Supervision, Writing-review and editing. **Wenjun Sun:** Data collection, Funding acquisition. **Xiangyun Xiong:** Data collection, Funding acquisition. **Siyu Zeng:** Methodology, Supervision, Writing-review and editing, Funding acquisition. **Miao He:** Methodology, Supervision, Funding acquisition.

Conflict of Interests 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.

Acknowledgements The work was supported by the Bureau of Ecology and Environment of Shenzhen, Natural Science Foundation of Hunan Province (No. 2023JJ50346), and the Open Fund of National Engineering Research Center for Advanced Technology and Equipment of Water Environmental Pollution Monitoring (No. 2024KFJJ0106).

References

Brito R S, Pinheiro H M, Ferreira F, Matos J S, Lourenco N D (2014). *In situ* UV-Vis spectroscopy to estimate COD and TSS in wastewater drainage systems. Urban Water Journal, 11(4): 261–273

Carré E, Pérot J, Jauzein V, Lin L, Lopez-Ferber M (2017). Estimation of water quality by UV/Vis spectrometry in the framework of treated wastewater reuse. Water Science and Technology, 76(3): 633–641

Chen B, Wu H, Li S F Y (2014). Development of variable pathlength UV–Vis spectroscopy combined with partial-least-squares regression for wastewater chemical oxygen demand (COD) monitoring. Talanta, 120: 325–330

Dahlén J, Karlsson S, Bäckström M, Hagberg J, Pettersson H (2000). Determination of nitrate and other water quality parameters in groundwater from UV/Vis spectra employing partial least squares regression. Chemosphere, 40(1): 71–77

Figueiró C S M, Bastos de Oliveira D, Russo M R, Caires A R L, Rojas S S (2018). Fish farming water quality monitored by optical analysis: the potential application of UV–Vis absorption and fluorescence spectroscopy. Aquaculture, 490: 91–97

Guan L, Tong Y, Li J, Li D, Wu S (2018). Research on ultraviolet-visible absorption spectrum preprocessing for water quality contamination detection. Optik, 164: 277–288

Ji Q K, Zhang C H, Li D (2020). Influences and mechanisms of nanofullerene on the horizontal transfer of plasmid-encoded antibiotic resistance genes between *E. coli* strains. Frontiers of Environmental Science & Engineering, 14(6): 108–117

Ji Y T, Bai X, Tang J, Bai M, Zhu Y, Tang J W (2024). Photocathodic activation of peroxymonosulfate in a photofuel cell: a synergetic signal amplification strategy for a self-powered photoelectrochemical sensor. Analytical Chemistry, 96(8): 3470–3479

Kang M E, Weng Y Z, Liu Y, Wang H K, Ye L, Gu Y L, Bai X (2023). A review on the toxicity mechanisms and potential risks of engineered nanoparticles to plants. Reviews of Environmental Contamination and Toxicology, 261(5): 92–124

Kröckel L, Schwotzer G, Lehmann H, Wieduwilt T (2011). Spectral optical monitoring of nitrate in inland and seawater with miniaturized optical components. Water Research, 45(3): 1423–1431

Langergraber G, Fleischmann N, Hofstädter F (2003). A multivariate calibration procedure for UV/VIS spectrometric quantification of organic matter and nitrate in wastewater. Water Science and Technology, 47(2): 63–71

Lepot M, Torres A, Hofer T, Caradot N, Gruber G, Aubin J B, Bertrand-Krajewski J L (2016). Calibration of UV/Vis spectrophotometers: a review and comparison of different methods to estimate TSS and total and dissolved COD concentrations in sewers, WWTPs and rivers. Water Research, 101: 519–534

Li C, Shi Y, Luo D, Kang M E, Li Y J, Huang Y, Bai X (2023). Interventions of river network structures on urban aquatic microplastic footprint from a connectivity perspective. Water Research, 243, 120418.1–120418.11

Li P, Hur J (2017). Utilization of UV–Vis spectroscopy and related data analyses for dissolved organic matter (DOM) studies: a review. Critical Reviews in Environmental Science and Technology, 47(3): 131–154

Nakar A, Schmilovitch Z, Vaizel-Ohayon D, Kroupitski Y, Borisover M, Sela Saldinger S (2020). Quantification of bacteria in water using PLS analysis of emission spectra of fluorescence and excitation-emission matrices. Water Research, 169: 115197

Sanchini A, Grosjean M (2020). Quantification of chlorophyll a, chlorophyll b and pheopigments a in lake sediments through deconvolution of bulk UV–Vis absorption spectra. Journal of Paleolimnology, 64(3): 243–256

Savci S (2012). An agricultural pollutant: chemical fertilizer. International Journal of Environmental Sciences and Development, 3(1): 73–80

Singh K P, Malik A, Basant N, Saxena P (2007). Multi-way partial least squares modeling of water quality data. Analytica Chimica Acta, 584(2): 385–396

Song K, Li L, Tedesco L P, Li S, Clercin N A, Hall B E, Li Z, Shi K

(2012). Hyperspectral determination of eutrophication for a water supply source via genetic algorithm–partial least squares (GA–PLS) modeling. Science of the Total Environment, 426: 220–232

Tan J, Liu L, Li F, Chen Z, Chen G Y, Fang F, Guo J, He M, Zhou X (2022). Screening of endocrine disrupting potential of surface waters via an affinity-based biosensor in a rural community in the yellow river basin, China. Environmental Science & Technology, 56(20): 14350–14360

Tang W, Li G, Yan W, He G, Lin L (2019). Exploring the influence of concentration and optical path on nonlinearity in VIS&NIR dynamic spectrum. Infrared Physics & Technology, 103: 103103

Thomas O, Burgess C (2017). UV–Visible Spectrophotometry of Water and Wastewater. Amsterdam: Elsevier

Wang Q H, Xiao L, Liu C, Li L (2019a). Optofluidic variable optical path modulator. Scientific Reports, 9(1): 7082

Wang R, Zhang Q, Zhang Y, Shi H, Nguyen K T, Zhou X (2019b). Unconventional split aptamers cleaved at functionally essential sites preserve biorecognition capability. Analytical Chemistry, 91(24): 15811–15817

Xia K, Li Z, Zhou X (2019). Ultrasensitive detection of a variety of analytical targets based on a functionalized low-resistance AuNPs/β-Ni(OH)$_2$ nanosheets/Ni foam sensing platform. Advanced Functional Materials, 29(39): 1904922

Xing Y, Xue B, Lin Y, Wu X, Fang F, Qi P, Guo J, Zhou X (2022). A cellphone-based colorimetric multi-channel sensor for water environmental monitoring. Frontiers of Environmental Science & Engineering, 16(12): 155–166

Xue B, Yang Q, Xia K, Li Z, Chen G Y, Zhang D, Zhou X (2023). An AuNPs/mesoporous NiO/nickel foam nanocomposite as a miniaturized electrode for heavy metal detection in groundwater. Engineering, 27: 199–208

Yang S, Jin L, Shao Z, Zhang X, Han Y, Du B, Yang D, Gu A Z, Chen Y, Li D, et al. (2024a). *In situ* and rapid toxicity assessment of air pollution by self-assembly passive colonization hydrogel. Environmental Science & Technology, 58(41): 18109–18121

Yang S, Sarkar S, Xie X, Li D, Chen J (2024b). Application of optical hydrogels in environmental sensing. Energy & Environmental Materials, 7(3): 311–336

Yi X, Gao Z, Liu L, Zhu Q, Hu G, Zhou X (2020). Acute toxicity assessment of drinking water source with luminescent bacteria: impact of environmental conditions and a case study in Luoma Lake, East China. Frontiers of Environmental Science & Engineering, 14(6): 109

Zhou X H, Liu L H, Bai X, Shi H C (2013). A reduced graphene oxide based biosensor for high-sensitive detection of phenols in water samples. Sensors and Actuators. B, Chemical, 181: 661–667