文章编号：1001-9014(2022)01-0323-14

DOI:10.11972/j.issn.1001-9014.2022.01.029

A neural networks based method for suspended sediment concentration retrieval from GF-5 hyperspectral images

LIU Yi-Ming[1], ZHANG Lei[2], ZHOU Mei[1,3,4], LIANG Jian[5], WANG Yan[1], SUN Li[1,3], LI Qing-Li[1,3,4*]

(1. Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China;
2. Beijing Tracking and communication Technology Institute, Beijing 100094, China;
3. Engineering Center of SHMEC for Space Information and GNSS, Shanghai 200241, China;
4. Engineering Research Center of Nanophotonics & Advanced Instrument, Ministry of Education, East China Normal University, Shanghai 200241, China;
5. Nantong Academy of Intelligent Sensing, Nantong 226000, China)

Abstract: The suspended sediment concentration (SSC) is an extremely important property for water monitoring. Since machine learning technology has been successfully applied in many domains, we combined the strengths of empirical algorithms and the artificial neural network (ANN) to further improve remote sensing retrieval results. In this study, the neural network calibrator (NNC) based on ANN was proposed to secondarily correct the empirical coarse results from empirical algorithms and generate fine results. A specialized regularization term has been employed in order to prevent overfitting problem in case of the small dataset. Based on the Gaofen-5 (GF-5) hyperspectral remote sensing data and the concurrently collected SSC field measurements in the Yangtze estuarine and coastal waters, we systematically investigated 4 empirical baseline models and evaluated the improvement of accuracy after the calibration of NNC. Two typical applications of NNC models consisting baseline model calibration and temporal calibration have been tested on each baseline models. In both applications, results showed that the calibrated D'Sa model is of highest accuracy. By employing the baseline model calibration, the root mean square error (RMSE) decreased from 0.1495 g/L to 0.1436 g/L, the mean absolute percentage error (MAPE) decreased from 0.7821 to 0.7580 and the coefficient of determination (R^2) increased from 0.6805 to 0.6926. After implementation of the temporal calibration, MAPE decreased from 0.8657 to 0.7817 and R^2 increased from 0.6688 to 0.7155. Finally, the entire GF-5 hyperspectral images on target date were processed using the NNC calibrated model with the highest accuracy. Our work provides a universal double calibration method to minimize the inherent errors of the baseline models and a moderate improvement of accuracy can be achieved.

Key words: GF-5, hyperspectral, water quality, suspended sediment concentration

基于人工神经网络的高分五号高光谱影像悬浮泥沙浓度反演方法

刘一鸣[1], 张 磊[2], 周 梅[1,3,4], 梁 建[5], 王 妍[1], 孙 力[1,3], 李庆利[1,3,4*]

(1. 华东师范大学 上海市多维度信息处理重点实验室,上海 200241;
2. 北京跟踪与通信技术研究所,北京 100094;
3. 华东师范大学 空间信息与定位导航上海高校工程研究中心,上海 200241;
4. 华东师范大学 纳光电集成与先进装备教育部工程研究中心,上海 200241;
5. 南通智能感知研究院,江苏 南通 226000)

Received date: 2021-01-18, **revised date**: 2021-08-18　　**收稿日期**:2021-01-18,**修回日期**:2021-08-18

Foundation items: Supported by the National Natural Science Foundation of China (61975056, 61901173), the Shanghai Natural Science Foundation (19ZR1416000), the Science and Technology Commission of Shanghai Municipality (20440713100, 19511120100, 18DZ2270800)

Biography: LIU YI-Ming(1997-), male, Shanghai, PhD. Research area involves Hyperspectral Imaging. E-mail:840122115@qq.com.

* **Corresponding author**: E-mail: qlli@cs.ecnu.edu.cn

摘要:悬浮泥沙浓度是水体监测中极为重要的指标。本论文基于神经网络具有弥补传统经验算法固有误差的潜力,设计并开发了基于人工神经网络的神经网络校正器来对经验反演结果进行二次校正。为了防止在小数据集的情况下出现过拟合问题,采用了特殊设计的正则化项。基于高分五号高光谱遥感数据以及在长江口和沿海水域同时收集的悬浮泥沙浓度实地测量结果,研究了4种基线经验模型,并评估了使用神经网络校正器后的精度。在每个基线模型上都测试了神经网络校正器模型的两个典型应用,包括基线模型校正和时间校正。在这两种应用中,结果均表明,经校正的D′Sa模型具有最高的准确性。通过使用基线模型校正,均方根误差从0.1495 g/L降低至0.1436 g/L,平均绝对百分比误差从0.7821降低至0.7580,决定系数从0.6805升高至0.6926。实施时间校正后,平均绝对百分比误差从0.8657降低至0.7817,决定系数从0.6688升高至0.7155。最后,基于神经网络校正器校正后精度最高的模型处理了整幅高分五号高光谱图像。本论文结果为各种经验反演算法提供了一种通用的二次校正方法,以最大程度地减少基线模型的固有误差,并且保证了反演精度。

关 键 词:神经网络校正器;机器学习;高光谱;高分五号卫星;水质;悬浮泥沙浓度

中图分类号:TN215 **文献标识码**:A

Introduction

Due to the optical reflectance, scattering and absorption of different substances, distinct optical properties of the surface reflectance have the great capabilities of extracting the information of the water quality parameters[1-6]. With the advances of computer calculation and the development of the remote sensing technology, especially spaceborne optical sensors, it is possible to achieve both the wider spatial range and the higher precision of water quality parameters (WQPs) retrieval[7-9].

The suspended sediment concentration (SSC) is an extremely important property for water monitoring, which is the consequence of aquatic degradation and soil erosion for deforestation and urbanization. The SSC is typically defined as the total concentration (g/L or mg/L) of both organic and inorganic matter suspended in the water because of the turbulence[10]. SSC can also be denoted as the total suspended matter (TSM), suspended particulate matter (SPM) and total suspended solids (TSS), etc[11]. The correlation between SSC and turbidity makes it easy to set certain landmark with known clarity in the map to intuitively evaluate the retrieval accuracy. According to the historical development of retrieval algorithms of the WQPs from remotely sensed data, there are broadly three approaches: analytical, semi-analytical and empirical approaches[12]. The empirical algorithms have been widely used because of the advantages of the easy implementation and the requirement of less fieldwork. A widespread empirical algorithm early developed in the International Ocean-Colour Coordinating Group (IOCCG) is the color-ratio algorithm based on the exponential function and band ratio of reflectance in order to compensate the intensity deviation of the light at different spots[13]. D'Sa et al. presented an empirical algorithm designed for moderately turbid waters based on the band ratio between the wavelength 670 and 555 nm for remote sensing reflectance (R_{rs}), which has been found highly correlated[14]. Nechad et al. argued that the use of a single band algorithm can also provide accurate results with the proper selection of the band[15, 16]. Chen et al. tested various SSC retrieval models such as single band and difference of two bands based on moderate resolution imaging spectroradiometer (MODIS) in a wide-range SSC concentration[16]. However, traditional empirical algorithms have the limitations for using the fixed mathematical formula forms, such as exponential function and band ratio, and thus can be further optimized by diminishing complex model errors.

Recently, with the improvement of computational power and the development of machine learning technology, rather complex WQP retrieval problems can be solved. The artificial intelligence technology holds the advantage of retrieving different water parameters based on a single machine learning algorithm. Plenty of implementations of machine learning algorithms such as multilinear regression[17], support vector machine[1] and artificial neural network (ANN)[18-20] have achieved high accuracy in WQP inversion problems. Wei et al. employed the least-squares support vector machine parameterized through the particle swarm optimization algorithm for SSC estimation based on unmanned aerial vehicle-borne hyperspectral images[1]. Hafeez et al. evaluated the retrieval potential by comparing several machine learning algorithms and extracted the relative variable importance which is an indicator for further research[21]. In spite of the black-box essence and difficulty of deconstructing, machine learning algorithms are still extensively and successfully applied in remote sensing retrieval of WQPs.

In recent years, many new satellites equipped with advanced imagers, which can obtain increasing spatial coverage, spectral resolution and spectral range, have been launched for general or specific purposes[22]. Gaofen-5 (GF-5) satellite equipped with 6 payloads aims at geographic and atmospheric monitoring, which was successfully launched on May 9, 2018 in China[23, 24]. The Advanced Hyper Spectral Imager (AHSI), one of the payloads, obtains 330 high-resolution bands in the spectral range from 400 to 2500 nm with the swath width of 60 km, which is highly progressive in the world[25].

In this study, we systematically investigated the SSC retrieval in the Yangtze estuarine and coastal waters by implementing several empirical baseline models based on the GF-5 hyperspectral images and SSC field measurements collected simultaneously. A neural network calibrator (NNC) for double calibration was proposed to com-

bine the advantages of ANN and the traditional empirical algorithms. This combination can compensate the inherent errors of the empirical models and reduce the data that ANN requires. In order to prevent the overfitting problem, an identity function was pretrained and a specialized regularization term was employed. Two typical applications of the NNC model including baseline model calibration and temporal calibration have been investigated based on 4 baseline algorithms. With the small size of dataset, a moderate improvement of accuracy has been achieved in both applications. Finally, the entire hyperspectral images on target date were processed using the algorithms with the highest accuracy to analyze the distribution of SSC and finish the reality check. This paper provides a universal secondary calibration method based on ANN to minimize the inherent errors of baseline models.

1 Materials and methods

1.1 Locations of SSC measurements

The Yangtze Estuary is selected as the area to investigate SSC retrieval algorithms. The Yangtze River, the longest river in Euro-Asian continent, rises in the Tibetan Plateau, flows generally 6300 km to the East China Sea and generates the Yangtze Estuary. The prosperous Yangtze Estuary, the geographically largest, most densely populated and industrialized area of China, plays an important role in geochemical cycles for a considerable amount of sediment suspended in the Yangtze River. The suspended sediment load per year from the Yangtze River reaches approximately 480 million tons and nearly 40% of the load is deposited in the Yangtze Estuary making it an extremely highly turbid region[26]. Besides, the Yangtze Estuary is characterized by the optical complexity because of the low salinity, high levels of nutrients, tidal currents and the biogeochemical environment[27-29]. Thus, the researches on the Yangtze coastal and estuarian waters are challengeable and valuable.

The Yangtze Estuary starts from Xuliujing and ends at the East China Sea, presenting a "three-order bifurcation and four outlets into the sea" pattern. The Yangtze Estuary is firstly divided by the Chongming Island and Hengsha Island into the North and the South Branch. Then the South Branch is secondly separated by Changxing Island and Hengsha Island into the North and South Channel. Finally, the South Channel is split into the North Passage and the South Passage by Jiuduansha wetland[23, 30]. With respect to the calibration and validation, the field measurements were concurrently collected according to the satellite overpassing time so as to provide a universal evaluation of accuracy. A total number of 14 water samples consisting 4 samples from 27 March, 3 samples from 24 May and 7 samples from 31 October 2019 were taken as shown in Figure 1. The field measurements collected by the buoy stations were obtained using the optical backscattering sensors (OBS) at 10-minute intervals and the linear interpolation was used to estimate SSC value at the satellite overpassing time based on the neighboring collected data. The field measurements collected by ships were obtained simultaneously at the satellite overpassing time using the weighing method with drying and filtration[31].

1.2 The GF-5 hyperspectral images

GF-5 satellite, launched on May 9 2018, denotes a polar-orbiting satellite of a series of China High-resolution Earth Observation System (CHEOS) satellites of the China National Space Administration, which has taken an AHSI designed and developed by Shanghai Institute of Technical Physics (SITP), Chinese Academy of Sciences[24]. The main characteristics of the GF-5 AHSI are shown in Table 1 according to the report from the China Centre for Resources Satellite Data and Application. Note that the GF-5 AHSI collects 330 spectral bands in total from 400 to 2500 nm with a very high spectral resolution (i. e., 5 and 10 nm for visible and near-infrared (VNIR) and short-wavelength infrared (SWIR) bands, respectively), meanwhile covering a viewing width of 60 km with high signal-to-noise ratio (SNR). Both its viewing width and number of spectral bands exceed other onboard spaceborne hyperspectral imager, such as EO-1 Hyperion and HICO of USA, HysIS of India and DESIS of Germany[25]. In this paper, the hyperspectral images were taken on 27 March, 24 May and 31 October 2019. Note that 3 cloud-free hyperspectral images taken on 31 October 2019 with spatial coverage of the Yangtze estuarian and coastal waters were selected for best visualization of final results.

1.3 Hyperspectral image preprocessing

The space-borne hyperspectral images were preprocessed in ENVI software as follows: orthorectification, radiometric calibration, atmospheric correction, masking and water extraction[23]. By preprocessing, digital number (DN) values of origin images were translated to surface reflectance.

1)Orthorectification

The GF-5 hyperspectral images contain the necessary information, i. e., the Rational Polynomial Coefficients (RPCs), to complete the photogrammetric processing. The ENVI RPC Orthorectification tools use RPC information and a high-resolution digital elevation model (DEM) to create a geometrically corrected image.

2)Radiometric calibration

The conversion from the quantized DN of raw imagery into at-aperture radiance ($W \cdot m^{-2} \cdot sr^{-1} \cdot \mu m^{-1}$) is a linear transformation described in Eq. (1) based on the gain and offset coefficients from the auxiliary information provided in GF-5 AHSI. In our experiment, the vicarious calibration coefficients mentioned in [32] were used.

$$Radiance = Gain * DN + Offset \quad . \quad (1)$$

3)Atmospheric correction

The next step is atmospheric correction which removes or decreases the influence of the atmospheric scattering, absorption and reflection and translates the at-aperture radiance to the surface reflectance signature[33]. The Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) model was used in this paper. In this research of the Yangtze coastal and estuarine waters, the atmospheric model and aerosol model were set

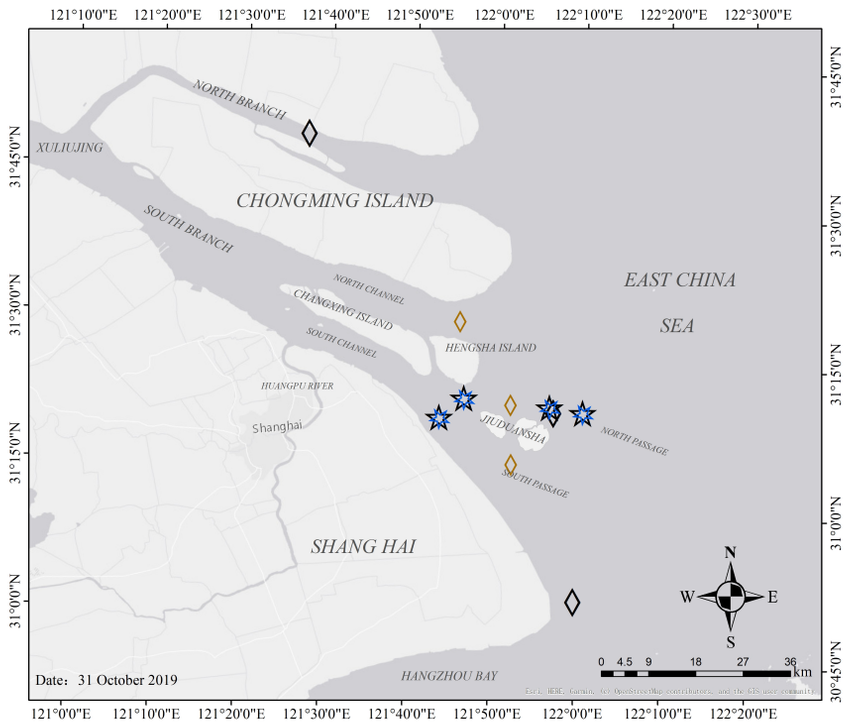

Fig. 1 Locations of 14 SSC field measurements on March 27 (blue), May 24 (brown) and 31 October (black) 2019 near the Yangtze estuarine and coastal waters. The stars and diamonds represent the field measurements collected by the buoy stations and ships, respectively

图1 长江河口卫星地面SSC同步实测位置示意图

Table 1 Main parameters for GF-5 AHSI

表1 高分五号高光谱卫星的主要参数

Parameters	Capability
Spatial Coverage	60 km
Spectral Range	400~ 2500 nm
Spectral Resolution	VNIR: 5 nm, SWIR: 10 nm
Spatial Resolution	30 m
Signal to Noise	100~200

as "Mid-Latitude Summer" and "Rural", respectively.

4) Masking and water extraction

Open water body can be identified via the Normalized Difference Water Index (NDWI) method, as in [34]. The NDWI can be calculated as follows:

$$NDWI = \frac{Green - NIR}{Green + NIR} \quad , \quad (2)$$

where *Green* and *NIR* represent the surface reflectance of green and near-infrared (NIR) bands, respectively. In our experiment, wavelengths of 895 and 565 nm were selected as the NIR and green bands respectively by observing and comparing the surface reflectance curves of water body with those of other terrain types.

1.4 Retrieval method

Based on the preprocessed AHSI data and field measurements, the entire procedures of the SSC retrieval are shown in Figure 2. In the input procedure, variables are initialized in order to prepare the required data for the next steps. The weights in ANN are initialized by simulating the identity function as pre-training. λ is a hyperparameter used to test the generalization ability of the ANN model. Dataset partition is employed to be ready for the implementation of the improved k-fold cross-validation method. The processing procedure is achieved by two steps. First, we use empirical algorithms as baseline models to generate the coarse results. Next, the NNC model is employed for secondary correction to compensate the non-linear components which cannot be described by the empirical algorithms. In this step, the NNC model takes the coarse results as inputs and outputs the corrected fine results. Besides, an evaluation of 3 assessment parameters is utilized to test the generalization performance of trained ANN under different λ. Finally, the network parameters and hyperparameter λ of the ANN with the best generalization ability are selected. More details are mentioned later in this section.

1.4.1 Baseline models

There are generally three approaches for quantitative remote sensing of WQPs: the empirical, analytical and semi-analytical approaches[12]. The empirical algorithms take the key advantages of easy implementation, computational simplicity and requirement of less fieldwork based merely on the simultaneous field measurements and remote sensing data. These algorithms generally provide robust accuracy for the calibrated area due to the assumptions that the inverse modeling of water and atmosphere remains the same throughout this region. Considering the same regional assumptions and computationally consuming features held by the NNC, empirical algo-

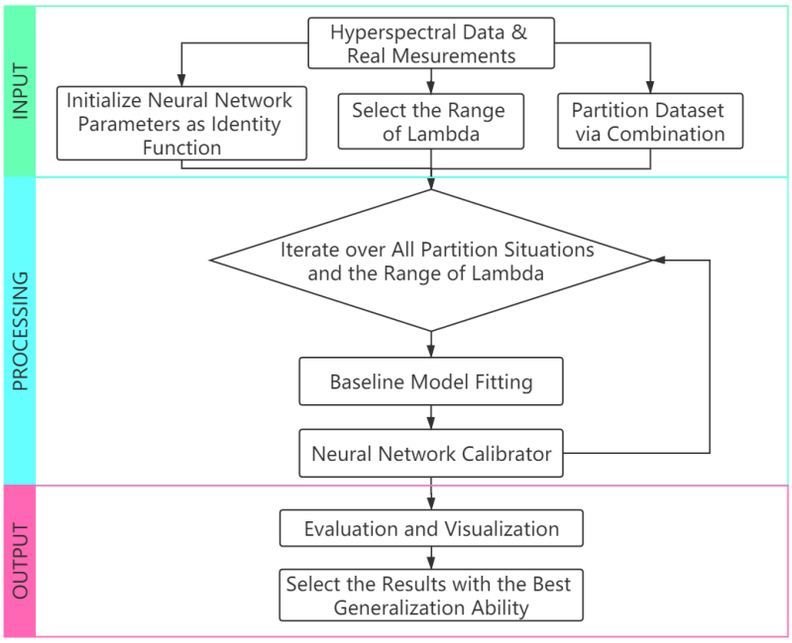

Fig. 2 Flow diagram for the entire SSC retrieval process.
图2 SSC反演流程图

rithms were chosen for the basic inverse modelling. Considering the Yangtze Delta being highly turbid, according to the research of Freeman et al. [11], 4 typical algorithms were employed in this paper, which took surface reflectance as input. Besides, these baseline algorithms utilized either single band, band ratio or band arithmetic as the independent variable in the form of linear and nonlinear models, including linear, exponential and power function models. The first algorithm was developed in the northern Gulf of Mexico by D'Sa utilizing R_{rs} for SeaWIFS sensor[14]. Developed as a band ratio power function model based on two SeaWIFS bands (555, 670 nm), a wide range of SSC and high tolerance for different preprocessing methods can be achieved via this model:

$$SSC = A*(\frac{R_{rs}(670)}{R_{rs}(555)})^{B} \quad , \quad (3)$$

where A and B represent the fitting coefficients.

Nechad et al. presented that the single band model can provide a robust SSC retrieval accuracy for case II turbid waters based on appropriate band selection around 700 nm. The recommended linear form of this algorithm is as follows[15]:

$$SSC = A*R_{rs}(Best\ Band) + B \quad , \quad (4)$$

where *Best Band* denotes the band selected using exhaustive search method. In order to translate this algorithm from MERIS, MODIS and SeaWIFS sensors to GF-5 AHSI, we tested entire 48 bands from 600 - 900 nm to locate the *Best Band*.

Similar to the Nechad model, Ruhl et al. derived and tested a single band exponential algorithm measured in the very turbid San Francisco Bay, California[35]:

$$SSC = Ae^{B*R_{rs}(Best\ Band)} \quad . \quad (5)$$

In this research, the algorithm was built based on field measurements collected from 1994 to 1998 with SSC values ranging from 0 to over 400 mg/L. This algorithm obtained R^2=0.59.

Considering the model developed by Loisel et al. in the highly turbid Mekong River Delta with SSC maximum values over 5000 mg/L, three bands (489, 557 and 668 nm) are utilized here to adapt to the GF-5 AHSI[36]:

$$SSC = 10^{A + B(R_{rs}(557) + R_{rs}(668)) - C(R_{rs}(489)/R_{rs}(557))} \quad , \quad (6)$$

where A, B and C are the fitting parameters.

1.4.2 Neural network calibrator

Our intuition of designing NNC is combining the complementary advantages between empirical models and ANN. Compared to empirical models which lack certain complex nonlinear features, NNC obtains the great capability of the ANN in extracting potential features and generating highly complex nonlinear functions. However, the ANN model requires a large dataset to prevent overfitting problem, which is hard to be satisfied in the field of remote sensing. In order to prevent the overfitting problem, the simple empirical models with just a few parameters can help ANN to reduce the required parameter number. By using transfer learning, our ANN is first trained to learn an identity function, aiming at learning the hypothesis of baseline models which require fewer parameters. Following this intuition, we proposed the NNC which takes the coarse results of baseline models as input and generate the calibrated fine results.

Usually, the ANN model consists of a collection of the connected neurons (or nodes) and corresponding weights assigned with links in the multilayer structure which typically includes an input layer, one or more hidden layers and an output layer. In this work, we aim to

secondarily calibrate the baseline SSC results and generate more precise results. In detail, the input of ANN is one baseline retrieval result and the output takes the corresponding field measurement as the label. Thus, a classical three-layer feed forward network with one node in the input layer and one node in the output layer was employed to update each input to a better output. Further, the number of nodes in the hidden layer should be small in order to reduce network parameters and prevent the overfitting problem. In our experiment, a hidden layer containing 10 nodes was selected for the small size of parameters and enough nonlinear expression ability. Finally, a sigmoid function was added after the output layer for activation. Below, we formulate the general form of ANN. In the feed forward process of prediction, the node vector of the former layer is multiplied with corresponding network parameters, added to a bias and then activated by the sigmoid function to obtain the node vector of the latter layer, as follows:

$$g(z) = 1/(1 - e^{-z}) \quad , \quad (7)$$

$$a^{(l+1)} = g(a^{(l)}\theta^{(l)} + \theta_0^{(l)}) \quad , \quad (8)$$

$$h_\theta(x) = a^{(3)} \quad , \quad (9)$$

where $g(z)$ is the sigmoid function, $a^{(l)}$ is the activated node vector of layer l, $\theta^{(l)}$ is the network parameter matrix from layer l to $(l+1)$, $\theta_0^{(l)}$ is a bias value from layer l to $(l+1)$ and $h_\theta(x)$ is the hypothesis value of the output layer with x being the input value. Specifically, given S_l nodes in the layer l, the shape of $a^{(l)}$ is $1 \times S_l$ and the shape of $\theta^{(l)}$ is $S_l \times S_{l+1}$.

The cost function (or loss) describes the error between the prediction values and the ground truth. The back propagation (BP) algorithm has been employed to iteratively minimize the cost function and complete the training process. Furthermore, we developed a distinct cost function with the purpose of optimizing the baseline model accuracy. The basic cost function is shown in Eq. (10).

$$Cost\ Function = \frac{1}{N}\sum_{i=1}^{N}\left[-y^{(i)}\log\left(h_\theta\left(x^{(i)}\right)\right) - \left(1 - y^{(i)}\right)\log\left(1 - h_\theta\left(x^{(i)}\right)\right)\right] \quad , \quad (10)$$

where N represents the number of training data, h_θ is the hypothesis values in the output layer, x is the baseline predictions in the input layer and y is the values of the field measurements. The regularization term is often used to penalize network parameters and improve the generalization ability of the ANN model. Here, a specialized regularization term is added to the cost function:

$$Cost\ Function\ += \frac{\lambda}{2N}\sum_{l=1}^{L}\sum_{j=1}^{S_{l+1}}\sum_{p=1}^{S_l+1}\left(\Theta_{j,p}^{(l)} - \Theta_{j,p,init}^{(l)}\right)^2 \quad , \quad (11)$$

where λ is the regularization hyperparameter controlling the degree of penalty, L is the layer number of the input and hidden layers, S_l is the number of nodes in the layer l and $\Theta_{j,p}^{(l)}$ represents the network parameter linking the layer l node p to the layer $(l+1)$ node j. An extra network based on identity function, i. e. inputs equal to outputs, was pre-trained to obtain the initial parameters Θ_{init} which provide the initial hypothesis based on baseline models to guarantee accuracy improvement after the secondary calibration.

A systematical investigation of two typical applications of NNC including baseline model calibration and temporal calibration has been presented. As for the baseline model calibration, aiming to compensate the inherent errors of the baseline models, the field measurements from 31 October 2019 were used both for fitting the baseline model and the secondary calibration of the NNC model. As for the temporal calibration, the purposes are the specialization of the parameterized historical model to adapt to the specific new field measurement data and the correction of inherent baseline errors. In this case, the baseline model was fitted as the historical model based on the in situ data from 27 March and 24 May 2019. Then, an extra linear calibration (LC) model was fitted to assign prediction results of the historical model to results on the specific date by using the data from 31 October 2019. Finally, the NNC model was trained based on the data from 31 October 2019 to secondarily calibrate the historical model to adapt to the specific date.

1.4.3 Statistical analysis

In order to gain better understanding of the various models, the accuracy for calibration and validation can be statistically evaluated by the three indices, root mean square error (RMSE), the mean absolute percentage error (MAPE) and the coefficient of determination (R^2). RMSE and MAPE are defined as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(X_{Est,i} - X_{Mea,i}\right)^2} \quad , \quad (12)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|X_{Est,i} - X_{Mea,i}\right|}{X_{Mea,i}} \quad , \quad (13)$$

where N is the total number of samples, $X_{Est,i}$ is the estimated value and $X_{Mea,i}$ is the field measurement value. The RMSE maintains the same unit as the in situ data and thus is intuitive and representative of the size of error. Besides, because of the disproportionate weight given by the squaring process, the RMSE is sensitive to occasional large errors and performs well in the situation with no outliers. As expressed in relative meaning, this statistical measurement can be compared widely across distinct data ranges. It is notable that the MAPE puts a heavy penalty on errors of small SSC values due to the ratio form, which leads to a significant complement of the RMSE. R^2 is defined as:

$$R^2 = 1 - \frac{RSS}{SST} = \frac{SSR}{SST} \quad , \quad (14)$$

where SSR is the sum of squares for regression, RSS is the residual sum of squares, SST is the total sum of squares and R^2 is defined as the ratio of SSR to SST. R^2 generally provides a replicated percent of the model for fitting the observation outcomes. With respect to the small size of in situ dataset, an improved k-fold cross validation method has been designed and implemented to precisely calculate 3 accuracy assessment parameters.

The method can be described in three steps.

(1) The first step is to select a reasonable number of training data. Mention that the number should be greater than the free degree of baseline models and less than the total number of the dataset minus 3 to obtain the valid R^2. Here the size of the training set is selected as 4.

(2) The next step is to find out all the possible situations via combination to pick training data from the total dataset and the number of situations here is $C_7^4 = 35$.

(3) After completing the division to different training and validation groups, each statistical parameter for validation of the groups can be calculated and the average is taken as the final evaluated accuracy.

It is indicated that every possible combination of the training and test data groups can account for the final average accuracy. However, due to fast growing rate of factorial function, this improved k-fold cross validation method can only be considered in the small size of dataset.

2 Results

2.1 Field measurements and spectral reflectance

In our research, field measurements of SSC have been collected concurrently to the GF-5 overpass based on the aforementioned method. The total 14 in situ SSC data measured on buoy stations and ships using drying and filtration process and optical backscattering method respectively was statistically analyzed in a line chart as shown in Figure 3. The SSC values of samples 1, 2, 5 and 7 are high over 0.35 g/L and the sample 7 achieves the highest concentration of 0.76 g/L. The higher SSC values on 31 October 2019 is probably caused by the nearly highest tide of the day according to the official tide table. The solid yellow line indicates the trend of sorted 7 in situ data on 31 October 2019. Note that the 7 field measurements are distributed relatively evenly in the spatial domain and the SSC measurements spread relatively equally ranging from 0.026 to 0.76 g/L. Hence, the field measured data of 31 October 2019 has the capability of representing the real SSC features in a wide range in spite of low number, which can provide more information for ANN to learn from, comparing the highly centralized dataset of the same size.

The preprocessed surface reflectance curves extracted in the highly likely estuarine spots of low, middle and high SSC values on each individual date are shown in Figure 4 (a), (b) and (c). Besides, some preprocessed surface reflectance curves of typical ground objects on 31 October 2019 are depicted for comparison in Figure 4 (d). Notice that in the radiometric calibration process vicarious and onboard parameters are applied for the images of 31 October 2019 and the other dates, respectively. In terms of low and middle turbid waters, a similar bimodal reflectance shape can be concluded through all 3 graphs. The two distinct spectral peaks appear in the vicinities of 580 and 820 nm. The first peak is jointly caused by the strong absorptive effect by CDOM and phytoplankton at shorter wavelengths and the exponential increasing absorption by water molecules[37, 38]. The second peak is generated for the combined effect of the high amount of suspended sediments and the strong absorption of water molecules between 700 ~ 750 nm, which is high-

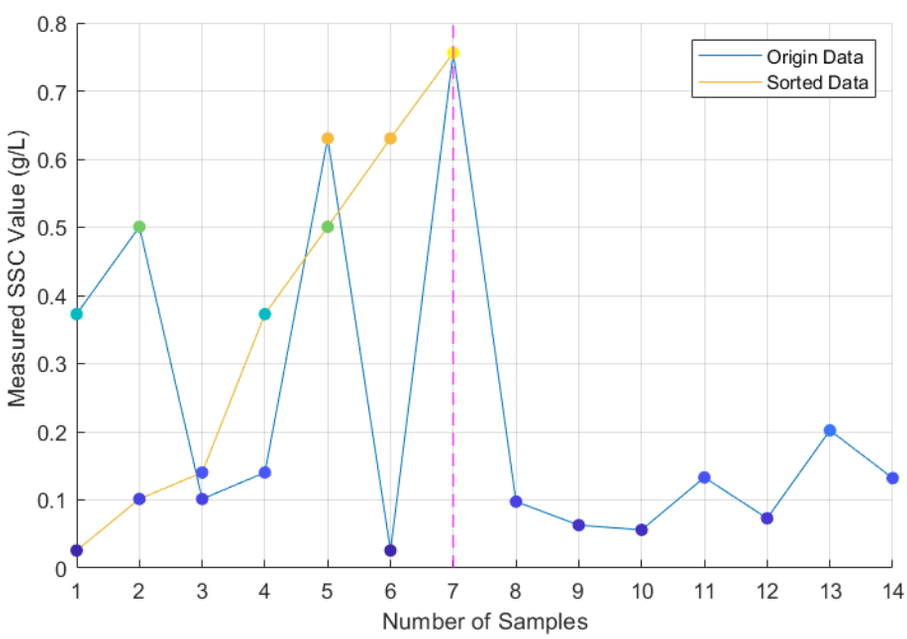

Fig. 3 Line chart of total in situ SSC data. The number 1~7, 8~10, 11~14 samples were measured on 31 October, 24 May and 27 March 2019, separately. A separation line (purple) is plotted to highlight the water samples 1~7 used for the final retrieval. The blue to yellow colors of dots intuitively show the low to high SSC levels. The lines drew in blue and orange represent the origin SSC values of all 3 days and sorted SSC values of 31 October 2019, respectively

图3 实测SSC数据. 点1~7, 8~10, 11~14分别代表了2019年10月31日,5月24日和3月27日的结果. 紫色分隔线用于突出点1~7,点的蓝色到黄色反映了从低到高的SSC浓度,蓝色和橙色线分别代表了原始SSC数值和排序后的2019年10月31日SSC数值.

ly correlated to SSC. As the SSC increases, a new peak near 681 nm emerges due to phytoplankton and the mutual cancellation of the high reflectance of suspended particles and the strong absorption of water molecules[37, 39].

According to documented locations of the 7 water samples on 31 October 2019, the preprocessed surface reflectance curves of GF-5 images are shown in Figure 5. The similar bimodal spectral characteristics and the general trend of increasing reflectance with the increment of suspended particles at approximate 820 nm can also be obtained.

2.2 Retrieval results of baseline model correction

In normal case that only in situ data on targeting date is available, NNC can be easily implemented to improve the accuracy of baseline models through compensating the inherent errors of the baseline models. By selecting the in situ data of 31 October 2019 as the whole dataset and using the improved k-fold cross validation method with 4 as the size of the training dataset, the SSC retrieval results of the baseline models and NNC are shown in the Table 2. Note that the band selection of the baseline Nechad and Ruhl models was accomplished by exhaustive search.

From the results, it is noticeable that the accuracy of the baseline model has been enhanced moderately for all RMSE, MAPE and R^2 after the double calibration of NNC. Because of the high sensitivity of RMSE and MAPE in terms of the high and low SSC values respectively, the calibrated results perform better in both SSC ranges, which indicates the effectiveness of our proposed NNC method. The calibrated D'Sa model achieved the highest accuracy. After calibration, RMSE decreased from 0.1495 to 0.1436 g/L, MAPE decreased from 0.7821 to 0.7580 and R^2 increased from 0.6805 to 0.6926. Besides, the highest improvement of accuracy was achieved by the Loisel model which had the worst performance in our limited dataset. After calibration, RMSE decreased by 19.2% from 0.4941 to 0.3993 g/L, MAPE decreased from 2.5812 to 2.1995 and R^2 increased from 0.2914 to 0.3992.

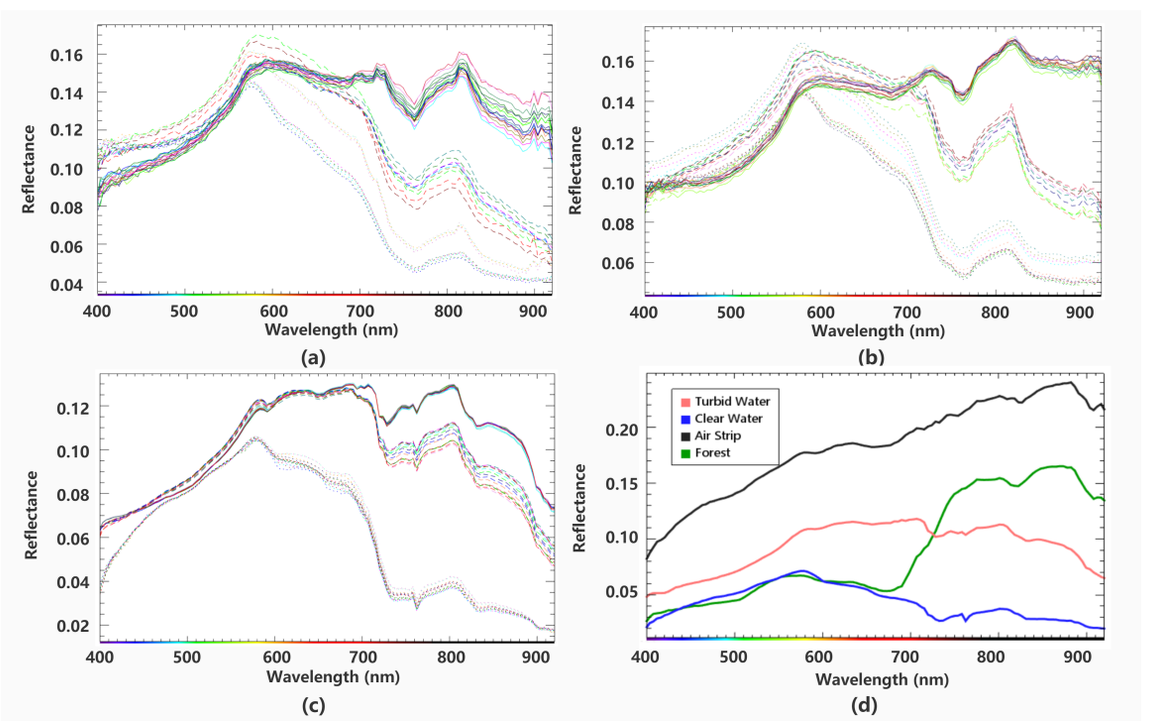

Fig. 4 Spectra of the surface reflectance in the research region on 27 March (a) 24 May (b) and 31 October (c) 2019. The dotted, dashed and solid lines represent the low, middle and high SSC values, respectively (d) some surface reflectance spectra extracted from different typical ground objects on 31 October 2019

图4 不同时间的光谱反射率曲线. 2019年(a) 3月27日 (b) 5月24日 (c) 10月31日. 点线、虚线、实线线型分别代表了低、中、高的SSC浓度值 (d)不同典型地物的反射率光谱曲线(2019年10月31日)

Table 2 Comparison between baseline and NNC results in the application for baseline model calibration.
表2 基线模型校正应用中的基线和NNC结果比较

Modeling Method	Independent Variables (nm)	Baseline			NNC		
		RMSE (g/L)	MAPE	R^2	RMSE (g/L)	MAPE	R^2
D'Sa	668,549	0.1495	0.7821	0.6805	0.1436	0.7580	0.6926
Nechad	758	0.1587	0.8049	0.6729	0.1567	0.7657	0.6772
Ruhl	745	0.2104	1.1142	0.6039	0.1939	0.9849	0.6336
Loisel	557, 489, 668	0.4941	2.5812	0.2914	0.3993	2.1995	0.3992

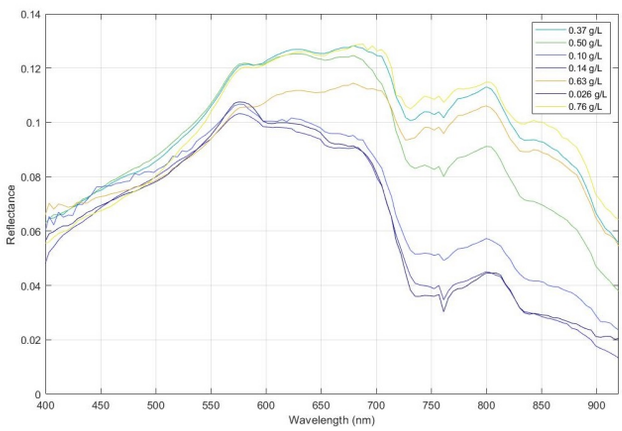

Fig. 5 The 7 examples of preprocessed surface reflectance spectra for different SSCs measured on 31 October 2019
图5 7种SSC浓度的反射率光谱曲线(2019年10月31日)

In order to overcome the problem of overfitting, a wide range of hyperparameter λ in the regularization term has been employed to test the generalization ability of the NNC model and thus the optimum λ of the best generalization performance of NNC is selected. The dependence relationships of λ and corresponding RMSE, MAPE, R^2 are displayed in the Figure 6. In general, the curves decrease at the beginning and then increase with the increasing of λ for RMSE and MAPE and an inverse trend is for R^2. The up-down shapes of the accuracy are associated with underfitting and overfitting problems. Hence, the λ corresponding to the general extremum of the curves is chosen to gain the optimum generalization performance. Note that the smaller hyperparameters always lead to a better improvement after the secondary calibration of NNC. However, rather big hyperparameters may be selected due to the small size of the training dataset so as to prevent overfitting. Besides, in order to visualize the NNC model, the relationships of the predicted values and the field measurements have been plotted as shown in Figure 7. The nonlinear errors of the baseline models can be visualized according to the NNC curves which may be clues for modifying the original baseline models. In spite of the slight difference between the NNC curve and the initial identity function due to the huge regularization hyperparameter used for preventing overfitting, a clearly moderate improvement can be available which is the evidence of the effectiveness of the application for the baseline model calibration.

2.3 Retrieval results of temporal calibration

When the baseline models calibrated and validated based on extra historical data are available, the NNC model can be used to adjust the existing model to adapt to specific date with an extra LC step. Our intuitive of the

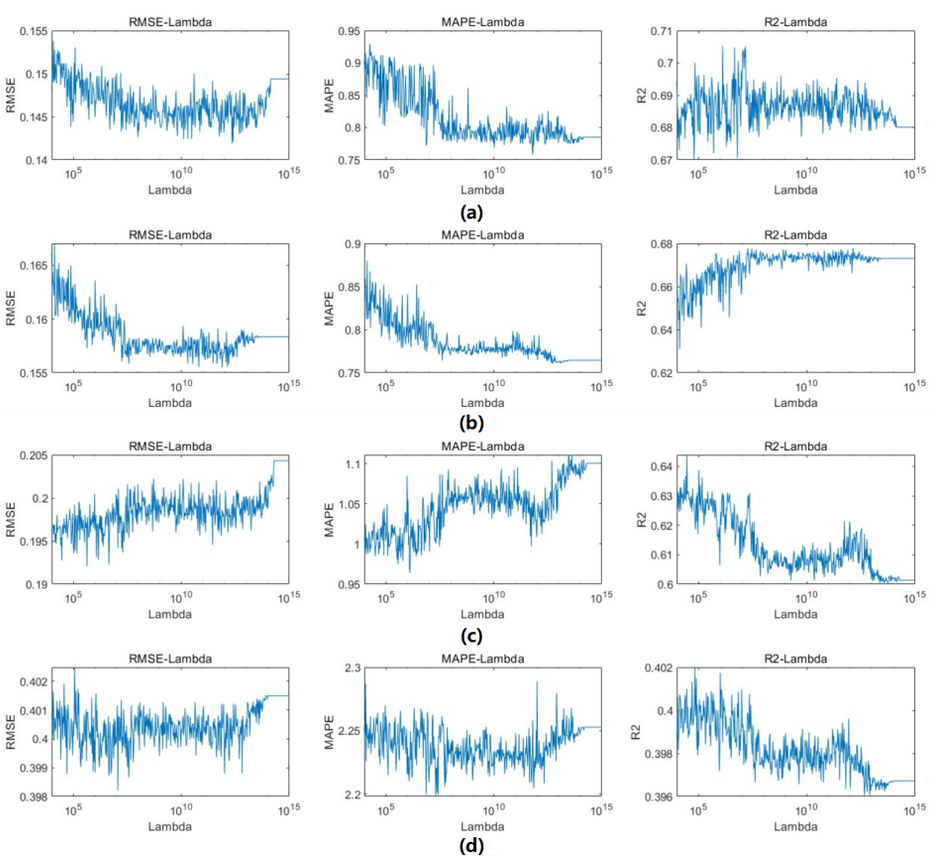

Fig. 6 The relationships between the regularization hyperparameter λ, RMSE, MAPE and R^2 for D'Sa (**a**) Nechad (**b**) Ruhl (**c**) and Loisel (**d**) models in the application for baseline model calibration
图6 在基线模型校正应用中的正则化参数λ, RMSE, MAPE and R2 (a)D'Sa模型 (b)Nechad模型 (c)Ruhl模型 (d)Loisel模型

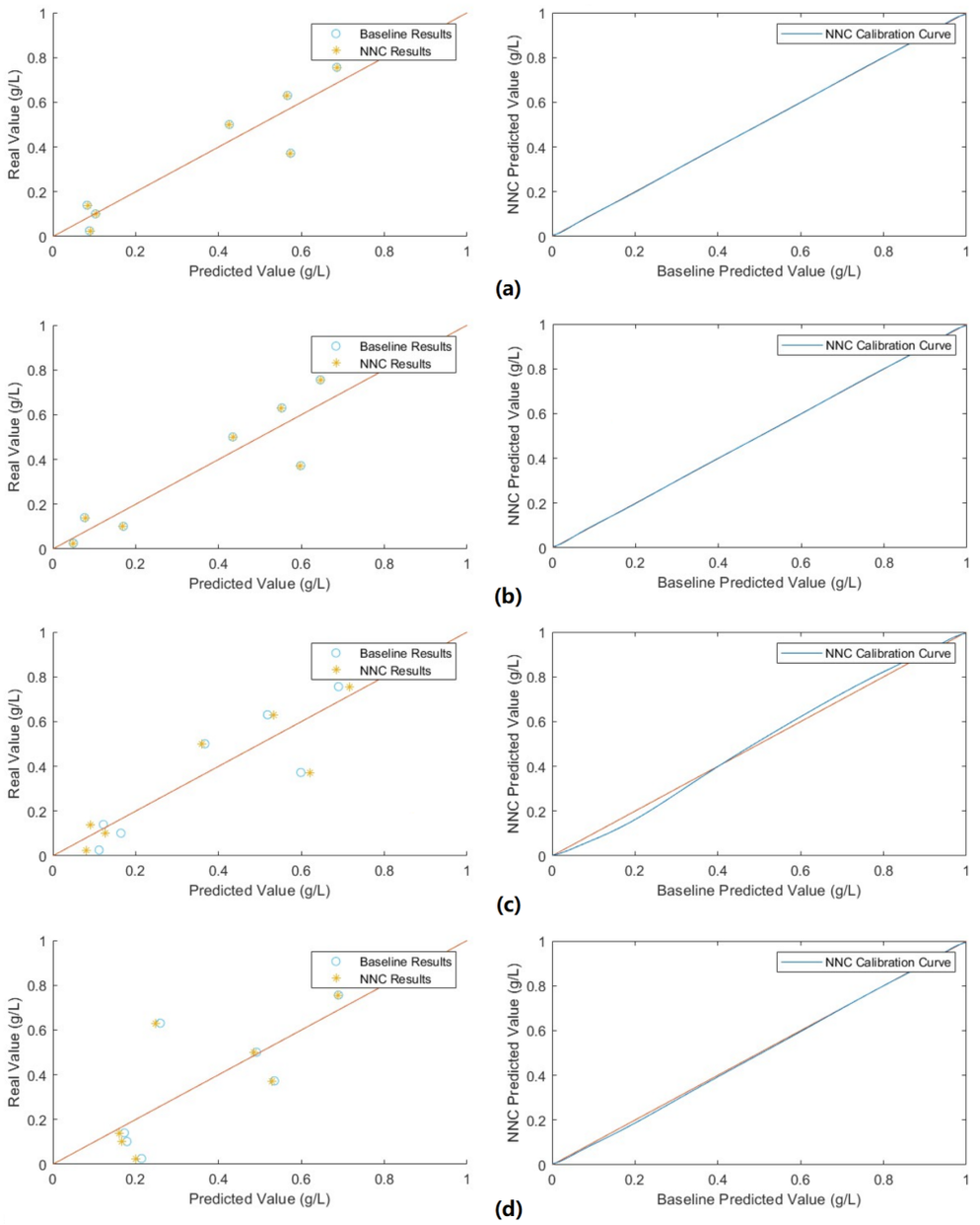

Fig. 7 The scatter diagrams (**left**) between the predicted values and field measurement values and the NNC calibration curves (**right**) for D'Sa (**a**), Nechad (**b**), Ruhl (**c**) and Loisel (**d**) models in the application for baseline model calibration

图7 在基线模型校正应用中的预测值与实测值散点图(左)和NNC校正曲线(右),(a)D'Sa模型,(b)Nechad模型,(c)Ruhl模型,(d)Loisel模型

temporal calibration is that adding extra historical information may generate better results. By selecting 4 as the size of the training set and using the improved k-fold cross-validation method, the application for temporal calibration was tested. The results of SSC retrieval based on historical baseline models are shown in Table 3.

Significant improvement of RMSE and R2 in most models can be obtained after the double calibration. Although the RMSE in D'Sa increases (from 0.1218 to 0.1352 g/L) after NNC in temporal calibration, the value decreases (from 0.1436 to 0.1352 g/L) compared to the result in baseline model calibration. The R2 in Loisel decreases (from 0.3685 to 0.3037), mainly because the great error after LC cannot be well calibrated by NNC. Specifically, the complex non-monotonic Loisel model leads to the overfitting problem in our small dataset and causes the great error after LC. Besides, the NNC only has limit calibration ability due to the small dataset and prevention of overfitting. Thus, a drop in R2 is observed in Loisel model. With a larger dataset, the NNC may achieve better results in temporal calibration. In terms of MAPE, the MAPE of most models decreases because the redistribution in the LC process may cause big relative errors when predicting small SSC values. Aiming at the visualization of the NNC model, the relationships of the predicted values and the field measurements for each

baseline model have been plotted in Figure 8. It can be indicated that linear errors including bias and scale can be simply calibrated by LC while the nonlinear errors can be further calibrated by NNC to obtain better results.

2.4 AHSI image inversion based on NNC

From the inverse results of the two applications, the D'Sa model of the temporal calibration with the highest accuracy (RMSE=0.1352 g/L, MAPE=0.7817 and R^2=0.7155) was selected for the SSC retrieval of the entire GF-5 images. Figure 9 (a) and (b) show the results of the SSC retrieval in the temporal calibration application based on the baseline model fitting and the NNC secondary calibration with the addition of LC, respectively. It is observed that the baseline model fitted with historical data is secondarily calibrated to match the SSC characteristics of the targeting date such as the SSC range. Some places with confidentially known SSC levels are used for reality check so as to further verify the retrieval accuracy in an intuitive way. The Changxing Reservoir, around 31°26′ N and 121°38′ E, has a rather low level of suspended concentration which is in great accordance with the inverse estimation of SSC. The turbid water quality of the Hangzhou Bay, around 30°44′ N and 121°50′ E, is well reflected by the red region in the graph. Beizhi, around 31°47′ N and 121°29′ E, is also a highly turbid region according to Gu et al. [23], and the fact is also consistent with the retrieval results.

3 Discussion

This study shows that the great learning capability of the ANN can be utilized to improve the accuracy in the SSC retrieval process. As mentioned above, moderate improvement can be observed, indicating the effectiveness of NNC. By employing the baseline model calibration, all three assessment parameters in four models obtain increment in precision. By employing the temporal calibration, RMSE and R^2 in most models obtain better results, despite the increment in MAPE due to the simple LC process.

Generally, the ANN model requires substantial data to drive and even very complicated models can be extracted by the great learning and reasoning abilities of ANN. However, considering the limitation of the dataset size, there may be the risks of overfitting. Hence, in order to prevent the overfitting problem, several aforementioned methods have been designed and employed. First, our proposed NNC takes the advantage of the small size of parameters of the simple baseline models. By using transfer learning, our NNC is first trained to learn an identity function, which reduces the data size that ANN requires. Second, a regularization term is added in the loss function of ANN to test the generalization ability. Third, the best hyperparameter λ is selected to obtain the model with the best generalization performance. Fourth, the improved k-fold cross-validation method is used to obtain low-variance accuracy estimation results and avoid the high-variance risks due to the limited dataset. In addition, 4 baseline models of different types and 3 accuracy assessment parameters were tested to ensure the reliability of our research.

4 Conclusion

This study shows that the great learning capability of the ANN can be utilized in the double calibration process to improve the accuracy of the SSC retrieval. In this paper, the proposed double calibration system is able to correct both linear and nonlinear errors of the baseline models based on ANN with a specialized regularization term. Our method obtained a moderate improvement of accuracy in both applications. For the two typical applications including baseline model calibration and temporal calibration, 4 distinct baseline models and corresponding NNC models have been systematically investigated using the GF-5 AHSI images and the concurrently collected field measurements. The results show D'Sa model is of highest accuracy in both applications. By employing the baseline model calibration, RMSE decreased from 0.1495 g/L to 0.1436 g/L, MAPE decreased from 0.7821 to 0.7580 and R^2 increased from 0.6805 to 0.6926, indicating NNC can compensate the inherent errors of the baseline models. After implementation of the temporal calibration, RMSE changed from 0.1218 g/L to 0.1352 g/L, MAPE decreased from 0.8657 to 0.7817 and R^2 increased from 0.6688 to 0.7155, which means the information from the historical field measurements can be extracted by NNC and provide a better initial hypothesis which probably leads to better accuracy compared with the baseline model calibration. The shortcoming of this experiment is the lack of concurrent SSC field measurements. Due to the small dataset, the huge hyperparameter λ was selected to prevent overfitting, which limited the improvement of accuracy. Thus, on the basis of this experiment, the concurrent collection process will be optimized in the future study to obtain more data. Also, only empirical algorithms were tested in this paper. Therefore, the effect of utilization of NNC on different model types can be tested for the future research.

Table 3 Comparison between baseline and NNC results in the application for temporal calibration
表3 时间校正应用中的基线和NNC结果比较

Modeling Method	Independent Variables (nm)	Baseline			NNC		
		RMSE (g/L)	MAPE	R^2	RMSE (g/L)	MAPE	R^2
D'Sa	668, 549	0.1218	0.8657	0.6688	0.1352	0.7817	0.7155
Nechad	762	0.3166	0.7016	0.4083	0.1588	0.7683	0.6670
Ruhl	762	0.2993	0.5867	0.3978	0.1804	0.9947	0.6456
Loisel	557, 489, 668	0.4160	0.6972	0.3685	0.3615	3.558	0.3037

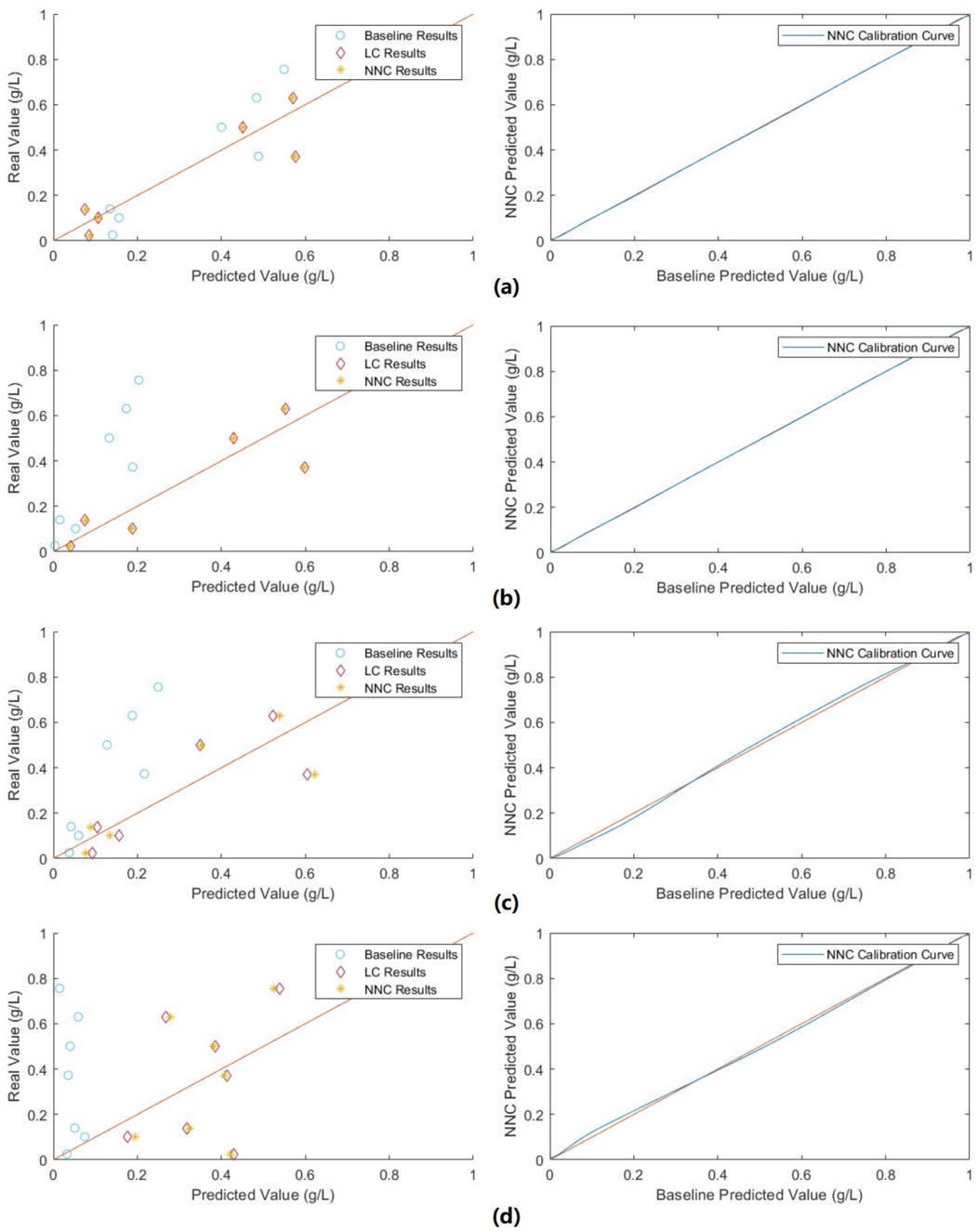

Fig 8 The scatter diagrams (**left**) between the predicted values and field measurement values and the NNC calibration curves (**right**) for D'Sa (**a**), Nechad (**b**), Ruhl (**c**) and Loisel (**d**) models in the application for temporal calibration

图 8 在时间校正应用中的预测值与实测值散点图(左)和NNC校正曲线(右),(a)D'Sa模型,(b)Nechad模型,(c)Ruhl模型,(d)Loisel模型

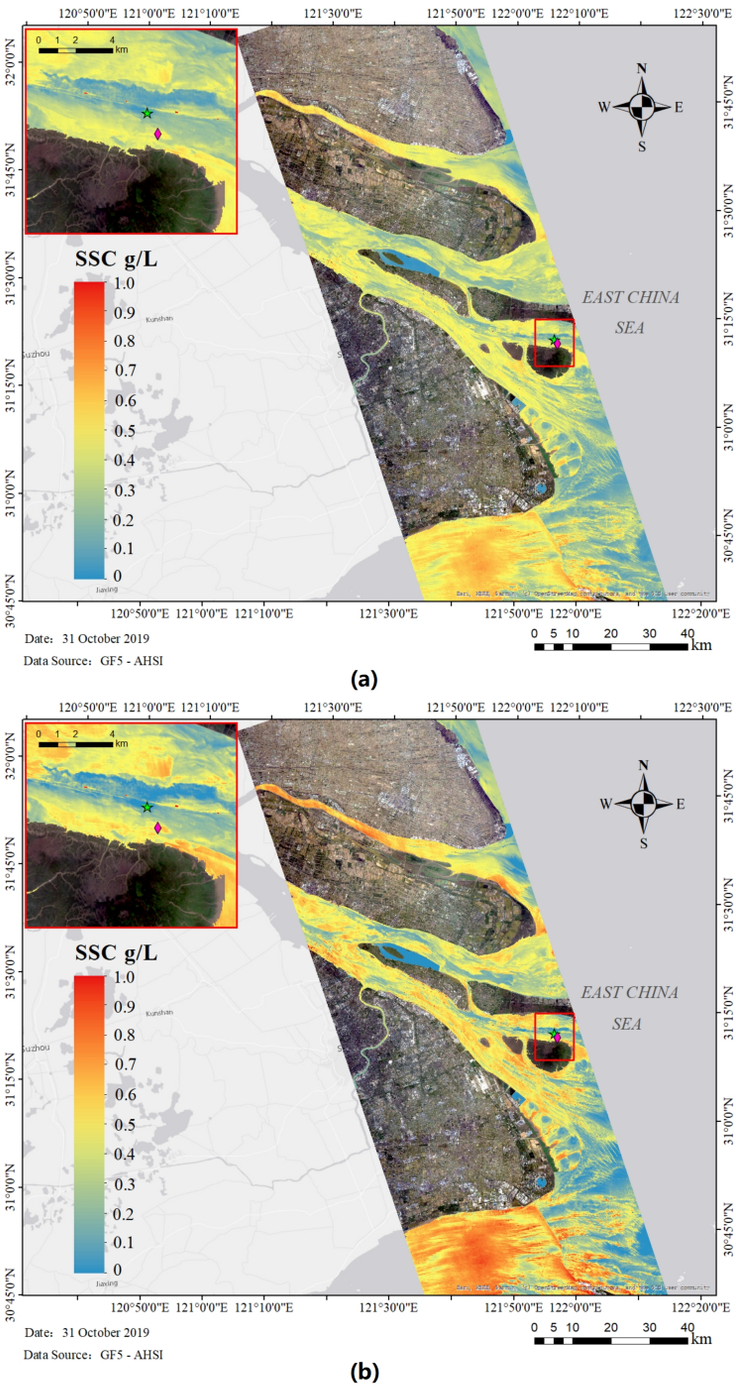

Fig. 9 SSC retrieval results of the baseline model (**a**) and NNC double calibration (**b**) using the D'Sa model in the application of temporal calibration based on the GF-5 images in the Yangtze estuarine and coastal waters on 31 October 2019. For result comparison, the magnified images of the region of interest (ROI) labelled in the red area are provided in the top left of each picture. The green star and pink diamond denote the samples with 0. 14 and 0. 63 g/L SSC values, respectively

图9 2019年10月31日长江河口的SSC反演结果,(a)基线D'Sa模型,(b) NNC校正模型,图中红色框内区域放大于左上角,其中绿色和粉色标注分别代表了0. 14 g/L和0. 63 g/L反演值

References

[1] Wei L, Huang C, Zhong Y, *et al*. Inland Waters Suspended Solids Concentration Retrieval Based on PSO-LSSVM for UAV-Borne Hyperspectral Remote Sensing Imagery [J]. *Remote Sensing*, 2019, **11**(12).

[2] Lei G, Zhang Y, Pan D, *et al*. Parameter selection and model research on remote sensing evaluation for nearshore water quality [J]. *Acta Oceanologica Sinica*, 2016, **35**(1): 114-117.

[3] Qamar A M, Mumtaz S, Anwar H, *et al*. Water quality monitoring: from conventional to emerging technologies [J]. *Water Supply*, 2020, **20**(1): 28-45.

[4] Arabi B, Salama M S, van der Wal D, *et al*. The impact of sea bottom effects on the retrieval of water constituent concentrations from MERIS and OLCI images in shallow tidal waters supported by radia-

tive transfer modeling [J]. *Remote Sensing of Environment*, 2020, 237.
[5] Cao Z G, Duan H T, Cui H S, *et al*. Remote estimation of suspended matters concentrations using VIIRS in Lake Hongze, China [J]. *Journal of Infrared and Millimeter Waves*, 2016, **35**(4): 462-469.
[6] Miao S, Wang R, Li J C, *et al*. Retrieval algorithm of phycocyanin concentration in inland lakes from Sentinel 3A-OLCI images [J]. *Journal of Infrared and Millimeter Waves*, 2018, **37**(5): 621-630.
[7] Al-Fahdawi A A, Rabee A M, Al-Hirmizy S M. Water quality monitoring of Al-Habbaniyah Lake using remote sensing and in situ measurements [J]. *Environ Monit Assess*, 2015, **187**(6): 367.
[8] Saberioon M, Brom J, Nedbal V, *et al*. Chlorophyll-a and total suspended solids retrieval and mapping using Sentinel-2A and machine learning for inland waters [J]. *Ecological Indicators*, 2020, 113.
[9] Menon H B, Adhikari A. Remote Sensing of Chlorophyll-A in Case II Waters: A Novel Approach With Improved Accuracy Over Widely Implemented Turbid Water Indices [J]. *Journal of Geophysical Research: Oceans*, 2018, **123**(11): 8138-8158.
[10] Bilotta G S, Brazier R E. Understanding the influence of suspended solids on water quality and aquatic biota [J]. *Water Research*, 2008, **42**(12): 2849-2861.
[11] Rhea W J, Ackleson S G, Freeman L A. Comparison of remote sensing algorithms for retrieval of suspended particulate matter concentration from reflectance in coastal waters [J]. *Journal of Applied Remote Sensing*, 2017, **11**(04).
[12] Matthews M W. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters [J]. *International Journal of Remote Sensing*, 2011, **32**(21): 6855-6899.
[13] Sarthyendranath S. Remote sensing of ocean colour in coastal, and other optically-complex, waters [J]. 2000.
[14] D'Sa E J, Miller R L, McKee B A. Suspended particulate matter dynamics in coastal waters from ocean color: Application to the northern Gulf of Mexico [J]. Geophysical Research Letters, 2007, **34**(23): (n/a-n/a.
[15] Nechad B, Ruddick K G, Park Y. Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters [J]. *Remote Sensing of Environment*, 2010, **114**(4): 854-866.
[16] Chen S, Han L, Chen X, *et al*. Estimating wide range Total Suspended Solids concentrations from MODIS 250-m imageries: An improved method [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2015, **99**: 58-69.
[17] Alparslan E, Aydoner C, Tufekci V, *et al*. Water quality assessment at Omerli Dam using remote sensing techniques [J]. *Environ Monit Assess*, 2007, **135**(1-3): 391-398.
[18] Sudheer K, Chaubey I, Garg V J J J o t A W R A. Lake water quality assessment from landsat thematic mapper data using neural network: an approach to optimal band combination selection1 [J]. 2006, **42**(6): 1683-1695.
[19] Sun D Y, Li Y M, Wang Q, *et al*. A neural-network model to retrieve CDOM absorption from in situ measured hyperspectral data in an optically complex lake: Lake Taihu case study [J]. *International Journal of Remote Sensing*, 2011, **32**(14): 4005-4022.
[20] Volpe V, Silvestri S, Marani M. Remote sensing retrieval of suspended sediment concentration in shallow waters [J]. *Remote Sensing of Environment*, 2011, **115**(1): 44-54.
[21] Hafeez S, Wong M, Ho H, *et al*. Comparison of Machine Learning Algorithms for Retrieval of Water Quality Indicators in Case-II Waters: A Case Study of Hong Kong [J]. *Remote Sensing*, 2019, **11**(6).
[22] Gholizadeh M H, Melesse A M, Reddi L. A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques [J]. *Sensors (Basel)*, 2016, **16**(8).
[23] Gu Q, Li Q, Zhou M. 2019 12th International Congress on Image and Signal Processing, *BioMedical Engineering and Informatics (CISP-BMEI)*, 2019: 1-5.
[24] Tang B-H. Nonlinear Split-Window Algorithms for Estimating Land and Sea Surface Temperatures From Simulated Chinese Gaofen-5 Satellite Data [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, **56**(11): 6280-6289.
[25] Sun W, Ren K, Yang G, *et al*. *2019 10th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp)*, 2019: 1-4.
[26] Song D, Wang X H. Suspended sediment transport in the Deepwater Navigation Channel, Yangtze River Estuary, China, in the dry season 2009: 2. Numerical simulations [J]. *Journal of Geophysical Research: Oceans*, 2013, **118**(10): 5568-5590.
[27] Shen F, Suhyb Salama M H D, Zhou Y-X, *et al*. Remote-sensing reflectance characteristics of highly turbid estuarine waters – a comparative experiment of the Yangtze River and the Yellow River [J]. *International Journal of Remote Sensing*, 2010, **31**(10): 2639-2654.
[28] Wu H, Shen J, Zhu J, *et al*. Characteristics of the Changjiang plume and its extension along the Jiangsu Coast [J]. *Continental Shelf Research*, 2014, **76**: 108-123.
[29] Chang P-H, Isobe A. A numerical study on the Changjiang diluted water in the Yellow and East China Seas [J]. *Journal of Geophysical Research: Oceans*, 2003, **108**(C9).
[30] Wu S, Cheng H, Xu Y, *et al*. Riverbed Micromorphology of the Yangtze River Estuary, China [J]. *Water*, 2016, **8**(5).
[31] Chen J, D'Sa E, Cui T, *et al*. A semi-analytical total suspended sediment retrieval model in turbid coastal waters: a case study in Changjiang River Estuary [J]. *Opt Express*, 2013, **21**(11): 13018-13031.
[32] Tan K, Wang X, Niu C, *et al*. Vicarious Calibration for the AHSI Instrument of Gaofen-5 With Reference to the CRCS Dunhuang Test Site [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, **59**(4): 3409-3419.
[33] Dall'Olmo G, Gitelson A A, Rundquist D C. Towards a unified approach for remote estimation of chlorophyll-a in both terrestrial vegetation and turbid productive waters [J]. *Geophysical Research Letters*, 2003, **30**(18).
[34] McFeeters S K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features [J]. *International Journal of Remote Sensing*, 2007, **17**(7): 1425-1432.
[35] Ruhl C A, Schoellhamer D H, Stumpf R P, *et al*. Combined Use of Remote Sensing and Continuous Monitoring to Analyse the Variability of Suspended-Sediment Concentrations in San Francisco Bay, California [J]. *Estuarine*, Coastal and Shelf Science, 2001, **53**(6): 801-812.
[36] Loisel H, Mangin A, Vantrepotte V, *et al*. Variability of suspended particulate matter concentration in coastal waters under the Mekong's influence from ocean color (MERIS) remote sensing over the last decade [J]. *Remote Sensing of Environment*, 2014, **150**: 218-230.
[37] Kutser T. Passive optical remote sensing of cyanobacteria and other intense phytoplankton blooms in coastal and inland waters [J]. *International Journal of Remote Sensing*, 2009, **30**(17): 4401-4425.
[38] Kutser T, Paavel B, Verpoorter C, *et al*. Remote Sensing of Black Lakes and Using 810 nm Reflectance Peak for Retrieving Water Quality Parameters of Optically Complex Waters [J]. *Remote Sensing*, 2016, **8**(6).
[39] Sun D, Li Y, Le C, *et al*. A semi-analytical approach for detecting suspended particulate composition in complex turbid inland waters (China) [J]. *Remote Sensing of Environment*, 2013, **134**: 92-99.