APPLICATION OF ARTIFICIAL NEURAL NETWORK TECHNOLOGY IN WATER COLOR REMOTE SENSING INVERSION OF INLAND WATER BODY USING TM DATA

J. P. Wang[*], S. T. Cheng, H. F. Jia

The Department of Environmental Science and Engineering, Tsinghua University, 100084, Beijing, China
Wjp00@mails.tsinghua.edu.cn

KEY WORDS: Artificial Neural Network Model, Water Color Remote Sensing, Extraction, Lake, Landsat Image

ABSTRACT:

For water color remote sensing study of inland water body, terrestrial satellite is often a good data source, because it has high spatial resolution. However, the precision of water color remote sensing inversion limits its application to water environmental monitoring and pollution analysis. This paper firstly studied traditional regression arithmetic and found that it was difficult to extract a good combination to construct the regression model. In order to acquire good inversion results, the paper introduced an advanced nonlinear science, artificial neural network technology. On the basis of satellite synchronous monitoring experiment, a BP neural network model was constructed to inverse SS, COD_{Mn}, DO, T-N, T-P and chl-a from Landsat TM data. The accuracy was acceptable and the relative error could be controlled 25%. Moreover, the reasons of simulating error, ways of improving model and applications of the model were also analyzed in details. The result of this research showed that based on a small-scale of satellite synchronous experiment, the model could be applied successfully in investigation, analysis and estimation of water quality.

It comes through long time that water quality data collection often depends on traditional monitoring, which needs much time and labor. So it is impossible to realize real-time and quick data acquiring. Along with continuous development of environmental information technology, water color remote sensing is applied more and more widely in the water quality monitoring of oceanic, coastal and inland water body because it has many advantages, such as wide range, synchronization and low cost of data collecting (Campbell, 1988; Claudia, 2001; Zhao, 2000). However, in order to utilize water color remote sensing technology more deeply and widely, there are many aspects to be improved: water color remote sensor technology, atmospheric correction and inversing model, which are in accordance with many scientific fields, and this paper paid more attention to inversing arithmetic study.

1. OVERVIEW

Researches of water color remote sensing began at 1920s, which had given correct explanation of sea color and begun to study optical field of water body (Shuleikin, 1933). But only after spatial technology occurred, water color remote sensing developed truly and quickly. Morel & Prieur classified water body as two types: Case I water body, which is Open Ocean; Case II water body, which is coastal, estuary and inland water body (Morel, 1977). Now the inversion research of Case I water body is correspondingly mature and the accuracy of inversing model is relatively good, since which component is mainly chlorophyll and has little suspended solids (SS). And that of Case II water body is very difficult due to the interaction of many water components, such as SS, chlorophyll and yellow substance. Inversion research of Case II water body is a hot issue currently. Aiming at Case II water body inversion model,

both theoretical model (Bricaud, 1986; Cheng, 2002) and empirical model (Chen, 1996; Ekstrand, 1992; Kuang, 2002) is getting along in recent years. And TM or ETM data was used by most researches (Chen, 1996; Cheng, 2002; Zhan, 2000). At present, the theoretical research still cannot be applied to practical inversion, but it can provide useful information to direct and modify empirical inversion arithmetic. Empirical model is often the main choice of quantitative calculation. How to improve its precision has become a key issue.

Water color remote sensing satellites, such as SeaWiFS and OCTS, have high spectral resolution and good optical attributes of water body, which are ideal data sources for water color remote sensing research, but in practice, the small spatial resolution often limits their application in inland water body. Terrestrial remote sensing satellites, such as landsat7, which has good spatial resolution (about 30meters), is often adopted as data sources, but spectral resolution of terrestrial satellite is a little low and can't reflect optical attributes of water body well. All these reasons make it more difficult to identify suspended solids, chlorophyll and yellow substances, and also limit the application of water color remote sensing in inland water bodies.

Artificial neural network (ANN) technology is a kind of nonlinear science developed from 1980's, which tries to simulate some basic attributes of people, such as self-adapting, self-organizing and fault tolerance. ANN has been used in many fields, such as mode identification and system simulation. Integrating water color remote sensing and characteristics of ANN, the paper hoped that artificial neural network model could perform the research of water color remote sensing inversion well. For all these, the paper drew the following research plan (Figure 1).

---

[*] Corresponding author.
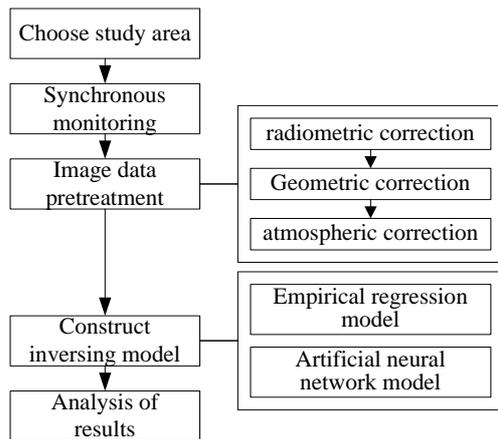
Figure 1. The Frame of Research Plan

2. DATA COLLECTION AND PROCESSING

2.1 Synchronous Monitoring

Water quality monitoring experiment synchronized with Landsat satellite was performed in PoYang Lake in July 8, 2001. The locations of sampling points were set as Figure 2.
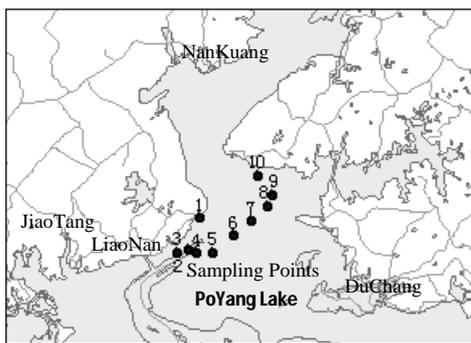
Figure 2. Distribution of sampling points

There were ten sampling points, at which SS, chl-a, TN, TP, COD_{Mn}, DO, temperature and pH were monitored. DO, temperature and pH were measured *in situ*, and the others in laboratory. Data of longitude and latitude were obtained by GPS at each point *in situ*.

2.2 Remote Sensing Data

Remote sensing data adopted synchronous Landsat 7 ETM+ data for its good spatial resolution, which path/row numbers were 121/40. The satellite image was clear and cloud-free.

2.2.1 Pre-processing of Remote Sensing Data

The remote sensing data needs several steps of pre-processing before an inversing model is applied, which include radiometric correction, geometric correction and atmospheric correction. Commonly the purchased image has been processed by radiometric correction and original geometric correction, so jobs that users need to do are accurate geometric correction and atmospheric correction. The accurate geometric correction in this study was accomplished by ground control points (GCPs),

whose precision was better than one pixel. The following part discussed the procedure of atmospheric correction.

2.2.2 Atmospheric Correction

PCI, a commercial image processing software package, provides a set of atmospheric correction tools for sensors of TM, MSS and SPOT, such as ATCOR0, ATCOR1 and ATCOR2. The flow chart of atmospheric correction is showed as Figure 3.
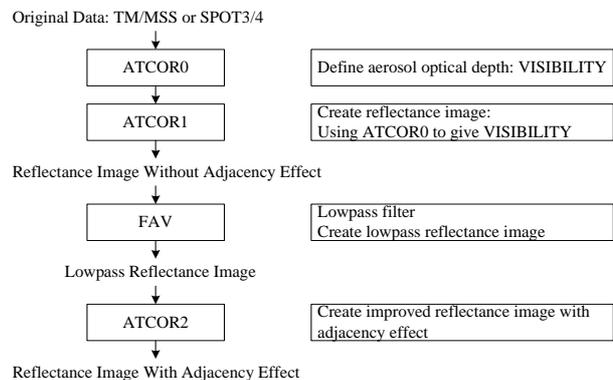
Figure 3. Flow chart of atmospheric correction provided by PCI

The theory of this method is given by Richter (Richter, 1990; Richter, 1996), whose main idea is that according to standard atmospheric categories, atmospheric dispersion has been calculated in different aerosol types, different sun zenith angles, different altitudes and different atmospheric visibilities, the results of which are stored in a directory, like as look-up table. In actual application atmospheric correction is performed according to this table. Categorizing basis of Richter method comes from middle resolution atmospheric transmission model - MODTRAN. Its arithmetic also considers and corrects the adjacent effect of ground reflection.

3. EMPIRICAL REGRESSION MODEL

Empirical Regression Model often sets remote sensing data (atmospheric correction or not) as independent variables and concentration of water quality components as dependent variable(s) to construct their relative equations. In order to review effects of empirical regression models, this study designed following combinations, and detailed description was showed in the previous study (Kuang, 2002) .

- Atmospheric correction: yes or no. two cases.
- Independent variables: 89 kinds of remote sensing band combination, such as R1, R1+R3, (R3*R5)/ln(R1*R2) and so on (Kuang, 2002).
- Dependent variables: SS and chl-a.

3.1 Results of Regression Model

The datasheet of water quality monitoring and remote sensing data were showed in Table 2. Calculating the relative coefficient of every above combination (the total mounts of combination were 356), results that had good correlation were listed in Table 1.

Atmospheric correction	dependent variables	good band combinations	relative coefficien t
No	SS	R1/ln(R3)	0.292
	Chl-a	ln(R1+R3)/ln(R2)	0.51
		ln(R1*R3)/ln(R2)	0.572
Yes	SS	R1/ln(R3)	0.411
		(R1+ln(R2))/ln(R3)	0.428
	Chl-a	ln(R1+R3)/ln(R2)	0.64

Table1. Results of Regression analysis

3.2 Analysis

Table 1 showed that:

- Most combinations had bad relative coefficients between concentration of water quality component and remote sensing data.
- As Table 1, only six combinations were extracted from the total 356, but their relative coefficient were still not good.
- Relative coefficient could be improved a little after atmospheric correction.
- It was hard to build an applicable regression model for quantitative analysis of Case II water body.

The main reason was:

- Water components interacted each other strongly. Relations between image bands and water components were crossed and non-corresponding relative.
- The wider band width also decreased their relativities.

4. THE ARTIFICIAL NEURAL NETWORK INVERSING MODEL

Theoretical research has proved that if an ANN includes biases, at least a S-style cryptic layer and a linear output layer, it can approach any rational function (Cong, 1998). And theoretical research of water color remote sensing indicates that remote sensing data are correlated with water body components and their concentrations directly after atmospheric effect is removed[*]. However, because water body ingredients affect each other, traditional inversing method, such as building a relative function between band data and components cannot solve it. The artificial neural network is suitable to simulate such complicated relationship. Therefore, theoretically, it is feasible to use ANN in water color remote sensing inversing research.

4.1 Model Structure

Figure 4 showed a concept chart of artificial neural network structure used in this study. Every input node represented a TM band, and these input data were distributed to each node of cryptic layer to operate. Output values of cryptic layer were inputted into output layer and operated again. The output values of output layer were parameters interested by users.

---

[*] Kuang, C., 1999. Identification of Water Color Remote Sensing Theoretical Model and Several Aspects in Its Solution. Tsinghua University doctoral dissertations, pp. 15~86.
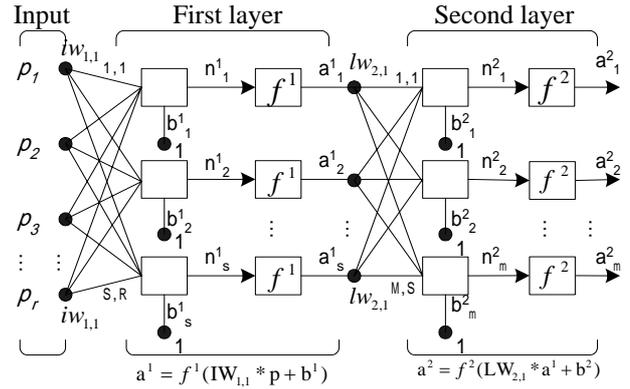
Figure 4. The structure of artificial neural network

In detail, the paper adopted a two layer BP artificial neural network.

p_i (i=1…r), r=5, represented five inputs which were band1 to band5 of TM image data;

$iw_{1,1}$ represented the weight between inputs and neural cells of first layer. $lw_{2,1}$ represented the weight between outputs of first layer and neural cells of second layer.

b^1 and b^2 were biases of neural cells of first and second layer respectively.

f^1 f^2 was transfer functions. f^1 often adopted S-style transfer function, in this paper the model used hyperbolic tangent S-style transfer function. f^2 adopted linear transfer function. Theoretical research has proved that an ANN, which has biases and at least a S-style cryptic layer and a linear output layer, can approach any rational function (Cong, 1998);

a^1 and a^2 were output values of first and second layer (see also Figure 4):

$$a^1 = f^1(IW_{1,1} * p + b^1)$$

$$a^2 = f^2(LW_{2,1} * a^1 + b^2)$$

m, the number of neural cells in the second layer, was determined by amounts of object of ANN model object. The paper hoped that this ANN model could simulate SS, COD_{Mn}, DO, TP, TN and chl-a, so m=6;

s, the number of neural cells in the first layer, was determined by neural network training process. According to results of the paper, s was equal to 10.

4.2 Training Results and Its Analysis

4.2.1 Data of Model Training and Verifying

Generally the input data should be normalized in order to improve training efficiency and accuracy of neural network (Zhan ,2000). Table 2 gave the input data and object values.

Item		Training samples							Verifying samples		
		1	2	3	4	5	6	7	8	9	10
Input	band-1	1.212	-0.736	-0.885	-0.587	-1.290	-0.141	1.672	1.131	-0.317	-0.060
	band-2	1.153	-0.628	-0.824	-0.677	-1.333	-0.176	1.637	1.235	-0.357	-0.029
	band-3	1.082	-0.313	-0.736	-0.684	-1.082	-0.456	1.825	1.258	-0.574	-0.319
	band-4	0.476	1.606	-0.714	-0.654	-1.071	-0.595	1.546	0.714	-0.654	-0.654
	band-5	-0.887	2.070	0.222	0.222	-0.795	-0.979	0.961	-0.979	0.499	-0.333
Object	SS	2.05	3.65	15.65	19.65	20.65	25.15	35	2.4	15.15	21.55
	CODmn	1.7	2.05	2.3	1.85	2	2	1.95	1.7	1.8	1.9
	DO	5.3	4.76	5.68	5.15	5.8	6.5	5.95	4.78	6.2	6.38
	T-P	0.055	0.03	0.04	0.035	0.03	0.03	0.055	0.05	0.04	0.025
	T-N	0.855	0.605	0.915	0.715	0.855	0.765	0.655	0.79	0.79	0.875
	chl-a/ug.L^{-1}	5	5	5	4	3	5	4	5	6	5

Table2. Original data of the artificial neural network model/mg.L^{-1}

In the training process, size of *s*, which was amounts of neural cell in cryptic layer, was justified gradually. Ultimately *s* was confirmed as 10 by analyzing accuracy of verifying samples, training time and iterative times.

The verifying results were showed in Table 3 after testifying samples were inputted to the having been trained ANN model.

Water quality parameters	SS	COD$_{Mn}$	DO	T-P	T-N	chl-a ug.L^{-1}
Inversing results	2.09	1.66	5.23	5.68	8.43	4.91
	20.29	1.64	4.98	3.59	6.21	4.11
	26.78	2.01	6.52	3.14	6.69	4.88
Synchronous monitoring results	2.40	1.70	4.78	5.00	7.90	5.00
	15.15	1.80	6.20	4.00	7.90	6.00
	21.55	1.90	6.38	2.50	8.75	5.00
Inversing errors	-13.01	-2.15	9.33	13.52	6.69	-1.84
	33.95	-8.81	-19.67	-10.26	-21.44	-31.46
	24.29	5.79	2.26	25.63	-23.50	-2.35

Table3. The results of inversion and validation of the model/mg.L^{-1}

It could be concluded that inversing errors could be controlled below 25%, except for SS and chl-a in second verifying data, which exceeded 30%. The results were satisfactory.

4.2.2 Result Analysis

Results indicated that the artificial neural network was well able to inverse water quality parameters from remote sensing image:

(1) Inversing effect of the model was good (Table 3). Inversing errors could be controlled less than 25%, except for SS and chl-a in second verifying data. If some measures as follows were taken in data collecting, processing and analyzing, inversing precision could be improved further.

(2) Ability of the model in simulating complicated relations was strong. ANN could realize nonlinear mapping between input and output parameter dimensions by the way of adjusting weights and biases of neural nodes (Cong, 1998). Results indicated that the ANN model had good simulating effects.

(3) The model inversed multi-water quality parameters simultaneously. The paper gave a model that SS, CODMn, DO, TP, TN and chl-a could be inversed from TM image using a trained ANN in the same time.

Reasons of one or two parameters having higher errors might be: synchronous experiment's internal errors, for example, boat stirring in shallow water or laboratory error; little sampling points were placed, which might result in water quality parameters bad-proportioned distribution in their concentration ranges; processing errors of remote sensing data, especially atmospheric correction errors; structure errors of the model.

The research showed that the first item error could be avoided or decreased by some steps, such as sampling after a moment of boat stopping, or increasing number of samples for each point; the second item error could be removed through steps of increasing sampling points or making these points reasonable arrangement; the third item error was difficult to remove. If it was possible, some optical experiments should be performed at the same time of satellite synchronous monitoring in order to decrease atmospheric correction errors. The last item error was inevitable, but adjusting neural network structure or comparing various amounts of neural cells could reduce the error.

5. CONCLUSIONS

Literature review indicated that water color remote sensing inversion often adopted empirical model for limits of water color remote sensor technology and atmospheric correction arithmetic.
This paper analyzed traditional regression model through combining different bands and operations. The relative coefficients between remote sensing data and water quality data was not good and could not satisfy the requirements of application. Analysis indicated that it was resulted from the interaction of many water quality components, such as SS, chlorophyll and yellow substances.

Both theoretical study and experimental analysis indicated that the ANN technology was feasible to water color remote sensing research, and the model had strong ability to simulate complicated inversing relation of second type water body.

On the basis of satellite synchronous monitoring experiment, a BP neural network model was constructed, by which concentrations of SS, CODMn, DO, T-N, T-P and chl-a were inversed from Landsat TM data and the accuracy was acceptable, the relative error could be controlled below 25%. And reasons of simulating error, ways of improving model and applications of the model were analyzed in details. The result of this research showed that based on a small-scale of satellite synchronous experiment, the model could be applied successfully in investigation, analysis and estimation of lake water quality.

REFERENCES:

Bricaud, A., and Morel, A., 1986. Light attenuation and scattering by phytoplankton cells: a theoretical modeling. *Appl. Opt.*, 25, pp. 571-580.

Campbell, J. W., O'Reilly, J. E., 1988. Role of satellites in estimating primary productivity on the northwest Atlantic continental shelf. *Cont. Shelf Res.*, 8, pp. 179-204.

Claudia, G., Monica, P., 2001. Detecting chlorophyll, Secchi disk depth and surface temperature in a sub-alpine lake using Landsat imagery. *The Science of the Total Environment*, 268, pp. 19-29.

Chen, C. Q., Shi, P., Mao, Q. W., 1996. Study on modeling chlorophyll concentration of surface coastal water using TM data. *Environmental Science and Technology*, 11(3), pp. 168-176(in Chinese).

Cheng, S. T., Kuang, C., Wang, J. P, 2002. Probe into the Water Color Remote Sensing Theoretical Model. *Journal of Tsinghua University*, 42(8), pp. 1027-1031(in Chinese).

Cong, S., 1998. *Neural network theory and application on MATLAB toolbox.* The university of science and technology of china press, Anhui, 11, pp. 45-60(in Chinese).

Ekstrand, S., 1992. Landsat TM based quantification of chlorophyll-a algae bloom in coastal waters. *Int. J. Remote Sens.*, 10, pp. 1913-1926.

Kuang, C., Wang, J. P., Cheng, S. T., etc., 2002. The Inversing Study of Water Color Remote Sensing in Sea Area of Macau. In: *The fourth symposium on environmental and city development*. Beijing, China, pp. 81-86 (in Chinese).

Morel, A., Prieur, L., 1977. Analysis of variations in ocean color. *Limnol. Oceanogr.*, 22, pp. 709-722.

Richter, R. A., 1990. Fast atmospheric correction algorithm applied to Landsat TM Images. *Int. J. Remote Sens.*, 11, pp. 159-166.

Richter, R. A., 1996. Spatially adaptive fast atmospheric correction algorithm. *Int. J. Remote Sens.*, 17, pp. 1201-1214.

Shuleikin, V. V., 1933. Data on the optics of a strongly scattering medium, applied to sea water, fog and cloud. *Geofizika,* 3, pp. 3-5.

Zhao, B. Y., He, B., Zhu, Y. Y., etc., 2000. Remote sensing water quality model of suspended sediments in the dianchi lake water bodies. *Environmental Science and Technology*, 3(94), pp. 16-18(in Chinese).

Zhan, H. G., Shi, P., Chen, C. Q., 2000. Inversing chlorophyll concentration of sea water using artificial neural network. *Chinese science bulletin*, 45(17), pp. 1779-1884(in Chinese).