

A fuzzy function model for remote evaluation of chlorophyll-a concentration in the Northern of South China Sea-Hong Kong

Zhao Xu¹, Xu Qian², Baoguo Shan¹, Jinhui Duan¹, Xiangdong Sun^{3*}

¹State Grid Energy Research Institute Co., Ltd., Changping District, Beijing 102209, China

²Ministry of Foreign Affairs of PRC, Beijing 100701, China

³The College of Economics and Management, Beijing University of Chemical Technology, Beijing, 100029, China

Abstract. Chlorophyll-a (Chla) concentration is an important indicator to monitor eutrophication, which is a persistent problem that affects the ecological health of coastal water or shallow lakes. We have applied a Gaussian-like fuzzy function model for the estimation of Chla concentration in Hong Kong Coastal Waters, based on the spectral signature using the HJ-1A /1B CCD images and measured data. The method of this paper is as follows. Firstly, reflectance peak and fluorescence band were used to figure out the bands or bands combination which impact the Chla concentration significantly; and then calculated the value of all the pixels without measured data according to their similarity degree with the measured points; finally, the full Chla concentration maps in the study area were produced by GIS spatial interpolation. It is concluded in this paper that the method can retrieve the Chla concentration accurately and the result of changes detected coincides with the observed result extremely, what's more, the maps generated are continuous and smooth which are quite different from traditional RS maps that can only accurate to pixel size.

1 Introduction

With the development of remote sensing technology, through the ocean color data of remote sensing technology has been widely applied to the field of climate, environment, fisheries and other, achieved good economic and social benefits, ocean color remote sensing is increasingly become an important part of the earth observation. Chla concentration is an important indicator of algae biomass size. In general, the unpolluted sea, Chla concentration is lower than 2 ug/L, while the average level in more than 10 ug/L indicates unacceptable eutrophication. In the waters of Hong Kong, Chla concentration and the site of most of the red tide frequency was significantly positive correlation, therefore, this article chooses Chla concentration as a sign of algal growth intensity and the degree of eutrophication, and established the model of calculating the concentration of Chla to detect the water quality in the Hong Kong area. Accurate assessment of Chla concentration in ocean waters has a great significance in the marine water quality monitoring. Principle of ocean color remote sensing is the change of the signal received by the satellite sensor, causing the content of each component of ocean color change to retrieve water, such as chlorophyll concentration, suspended sediment content, soluble organic matter content. All kinds of parameters of water quality remote sensing monitoring have an important significance on a large area. A variety of algorithms have been developed for estimating Chla in waters. A three-band reflectance model for remote estimation of Chla,

originally developed for estimating pigment contents in terrestrial vegetation. The SeaWiFS standard globally applicable Chla algorithm (OC4v4) made it possible to accurately quantify Chla in the open ocean to within $\pm 35\%$ of in situ measurements. Estimating Chla concentration using inherent optical properties of Chla models. A lot of semi-analytical algorithms and standard empirical algorithms were used. However, like many other models, they cannot solve the problem brought by variation of Chla specific absorption properly. At present, a main reason for the low accuracy of Chla estimation is the assumption of a constant for Chla specific absorption. Based on the research in the Northern of South China Sea-Hong Kong, a fuzzy membership function was established. Aiming at the uncertainty of data, it can dig deep relationship between the measured data and predicted data [1-4].

2 Materials and methods

2.1. Technology flow

As same as other Chla concentration calculation models, a few main steps, such as previous data preparation, algorithm model and accuracy verification, are needed for the foundation of the fuzzy membership function of chlorophyll concentration.

Fig.1 in the form of a flow chart is introduced which is about the modeling and the preparation process of the fuzzy function model: First is the pretreatment of data, in

* Corresponding author: qianxu1108@163.com

this paper, two independent datasets were used, CCD remote sensing image data and measured data. Remote sensing image preprocessing is an important preparation work; it is about to late the accuracy of the modeling. In order to achieve good accuracy, we used the methods, such as radiation correction, geometric correction and atmospheric correction, on CCD image to pretreat. We used ETM level5 data in geometric correction which have the same resolution and bands with CCD data and used the NIR-SWIR atmospheric correction algorithm in atmospheric correction. The next step after the pretreatment is the most important part of our study called fuzzy membership function model. This part is the central idea of this paper, the algorithm for the modeling process will be introduced in detail in subsequent chapters, so this is no need to repeat. Fig.1 shows clearly that the process to establish the model and the method based on Gaussian function. The last part is the output of the data results and the establishment of Chla concentration prediction data sets. One of the most important in this part is the accuracy of the results and this is also an indispensable part of the model. This paper is not only from the aspect of mathematical analysis of the precision of data processing to analyze, but also combined with the experience of predecessors' research on calculate of Chla concentration distribution in space and time scales [5,6].

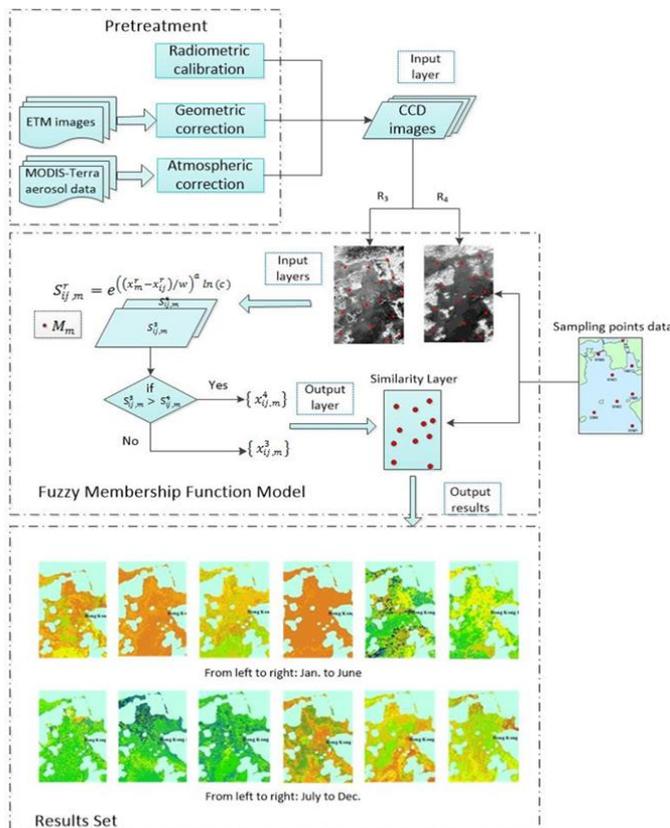


Fig. 1. Technology flow chart

2.2 Study area

The region under study ($113^{\circ} 54' - 114^{\circ} 14' E$, $22^{\circ} 11' - 22^{\circ} 25' N$, Fig.1) is located in the western part of the Hong Kong Victoria Harbor ocean and the northeastern part of the Lantau Island ocean, near the southeastern coast of China, including 26 monitoring stations, belonging to the four control region of Vitoria, southern, Northwestern, Western buffer (Fig.2). Waters in the study area mainly belongs to class 2 water, including fresh water by the Pearl River flows into the South China Sea. Hong Kong port is a world famous and very busy harbor, coastal zone water used to water recreation activities, aquaculture, industrial cooling water, waste water discharge and shipping, etc., is also the habitat of all kinds of Marine life. Therefore, Hong Kong coastal sea water quality assessment and water quality protection has important significance (Fig.3).

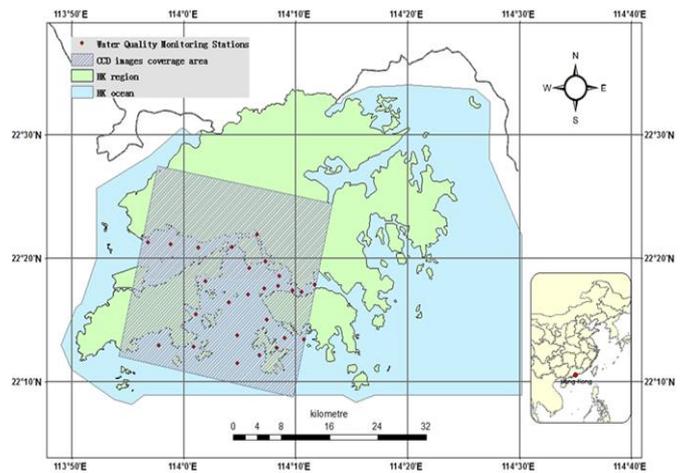


Fig. 2. Location and sampling locations of study area

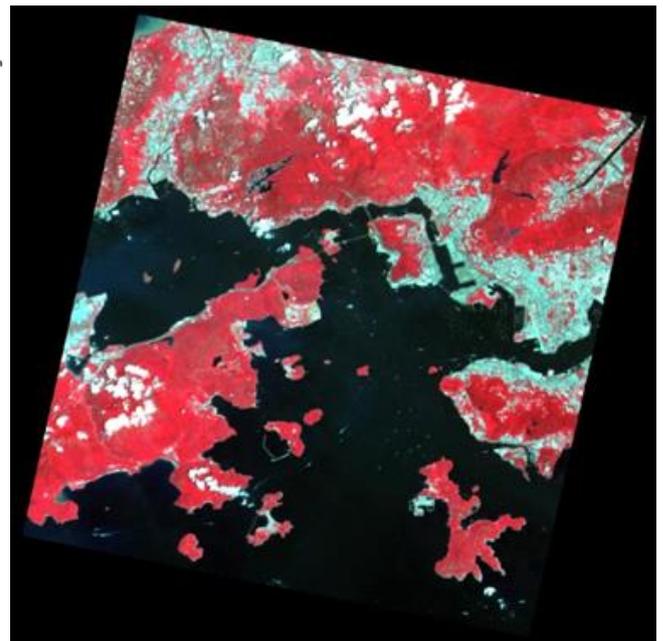


Fig. 3. The atmospheric correction CCD image in August 2010

3 Development of fuzzy membership function model

3.1. Data normalization

In this study, in order to avoid data saturation, the input variables in this model were normalized. We adopted the linear transfer function (Eq. (1))

Equations should be centred and should be numbered with the number on the right-hand side.

$$f(x) = (x_1, x_2, \dots, x_n), \quad x_{norm} = \frac{x_n - x_{min}}{x_{max} - x_{min}} \quad (1)$$

where x_n , x_{min} , x_{max} , and x_{norm} are the real-value input variable, the minimum input variable, the maximum input variable, and its normalized value, respectively.

3.2 Model building

The most crucial tasks for estimating Chla concentrations in sea water were the establishment of a data retrieval model and the selection of input variables for the model (Fig. 1). Fuzzy membership function describes how similarity between an unknown point and the sample points. The similarity value varies from 0 (which means that unknown points are very different from a given point) to 1 (which means that the unknown points are exactly the same with the given points). In this paper we used a Gaussian-like function as the basic form of fuzzy membership functions: First, calculated the similarity of the unknown points (i, j) to the sample points M_M ($m=1, \dots, n$; $n=26$, n is number of sample points) in a particular band R_r ($r=3, 4$) (Eq. (2)). Then, integrating the similarity of each unknown point at specific band (according to the principle of buckets takes the minimum value), so as to obtain the similarities of unknown points based on the sample points [7].

$$\begin{cases} S_{ij,m}^r = e^{\left(\frac{|x_{ij}^r - x_m^r|}{w_1}\right)^{a_1} \ln c}, & x_{ij}^r < x_m^r \\ S_{ij,m}^r = 1, & x_{ij}^r = x_m^r \\ S_{ij,m}^r = e^{\left(\frac{|x_{ij}^r - x_m^r|}{w_2}\right)^{a_2} \ln c}, & x_{ij}^r > x_m^r \end{cases} \quad (2)$$

where $S_{ij,m}^r$ is the similarity of the unknown points (i, j) to the sample points in a particular band; x_{ij}^r is the digital number (DN) of an unknown point in a particular band; x_m^r is the digital number (DN) of the sample points; c is a parameter for the user to control the function of shape, the average value of 0.5 in the paper; a is the parameter control the slope of curve, w control the width of the curve:

w_1 , a_1 are the parameters set for the user to control the scope of the function in $x_{ij}^r < x_m^r$ range; w_2 , a_2 are the parameters set for the user to control the scope of the function in $x_{ij}^r > x_m^r$ range.

$$S_{ij,m}^r = f(S_{ij,1}^r, \dots, S_{ij,n}^r), r = 3, 4, \dots \quad (3)$$

where $S_{ij,m}^r$ are the comprehensive similarity values for (i, j), $f()$ is the comprehensive function. We adopt minimum function in this paper.

So we got a data set: $x_{ij,m}^r$ then $x_{ij,m}^r$ as an input layer (we can think of it as a “band”) to equation 1 for calculating the similarities. The maximum values we were selected in order to improve the accuracy of this model.

$$S_{ij,m}^{x^r} = f(S_{ij,1}^{x^r}, \dots, S_{ij,n}^{x^r}) \quad (4)$$

where $S_{ij,m}^{x^r}$ are the comprehensive similarity values for $x_{ij,m}^r$, $f()$ is the comprehensive function. We adopt maximum function in this step.

At last, we estimated the Chla concentration as the function

$$Chla_{ij} = M_{m,max} S_{ij,m}^{x^r} \quad (5)$$

where $M_{m,max}$ is the sample point Chla concentration value corresponding $x_{ij,m}^{x^r}$.

4 Predictive Chl-a mapping using the constructed fuzzy membership functions

4.1 Predictive Chl-a mapping using the constructed fuzzy membership functions

Fuzzy membership functions were used under the matlab framework to generate fuzzy Chl-a maps for the area. Twelve fuzzy membership maps, about twelve months in 2010, were generated for the study area. Fig.3 shows Chla concentration maps from January to December 2010. In these maps we can definitely see the seasonal variation of Chla in study area. This article studied water belongs to the typical case 2 water in the ocean, so the Chl-a concentration was higher than case 1 water. In the case of the same year, the average phytoplankton biomass and Chl-a concentrations were higher in summer (June-September) than in winter (November-February), the maximum Chl-a concentrations typically occurred in August/July every summer and the more close to land, the higher chlorophyll content; the study area is situated at the Pearl River Estuary, so the seasonal distribution of chlorophyll more affected by Pearl River runoff, many scholars have been in-depth study on the Pearl River

estuary ecological environment, the distribution of Chla concentration in winter in coastal areas is above that in far from shore areas and in spring water temperature rises, the light time enhances, phytoplankton and chlorophyll into an upward trend, July and August reached the year's highest peak; after enter the autumn and winter, the Chla concentration gradually decreases. The changes in Chla concentration also corresponds to the Pearl River in rain season and dry season: Pearl River runoff taking lots of nutritive salt into the Pearl River Estuary in rain season; at the same time, the water temperature and solar radiation increased gradually, so the growth of phytoplankton accelerated, Chla concentrations show a rising trend, reached the peak in July and August, and then due to the water temperature and solar radiation decreased, Chla concentrations began to decrease in September. Enter the normal season, although water temperature and solar radiation conditions continue to decline, under these conditions, the growth of phytoplankton was still vigorous, and Chla concentrations were still maintaining a high level. Then, into the winter, water temperature and solar radiation decreased obviously, while the reduction of nutrient input to alleviate the phytoplankton growth, which makes the Chla concentration continued to decline, in the dry season is the lowest. In this research area, the change of Chla concentrations was also affected by fishing and tourism, but the overall trend was high in summer and low in winter [8,9].

Combining with the experience of scholars' research, we analyzed the predictive Chl-a mapping using the constructed fuzzy membership functions (Fig.4). We found that the distribution of chlorophyll concentrations consistent with previous research of the scholars: speaking from the season, Chla concentrations reached the highest values in summer and lowest in winter, spring and autumn are the transitional season for Chla. From the whole of the south China sea waters, the study area belongs to the Pearl River Estuary, this sea area is influenced by the Pearl River fresh water and terrigenous nutrients carried by influenced by Pearl River deeply, so the Chla concentrations are generally higher than the other sea areas in South China Sea. Therefore, in these ways, the remote estimation of Chla in a fuzzy function model is feasible [10-12].

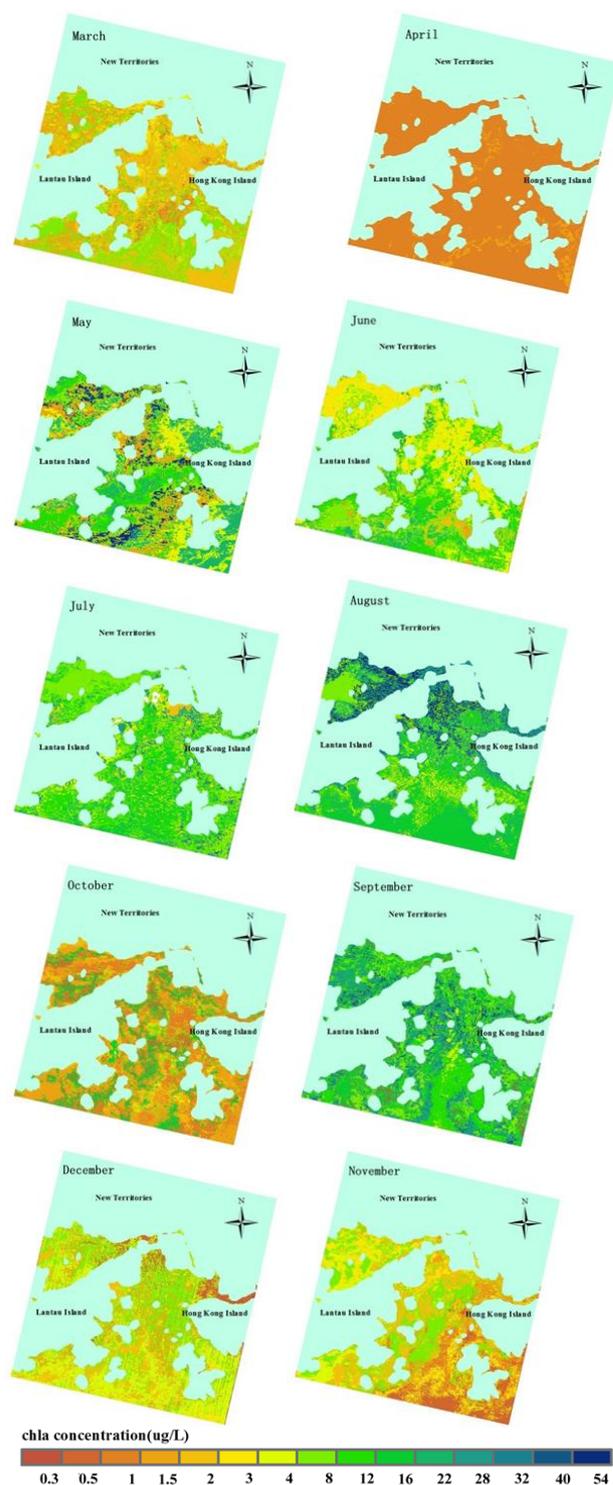


Fig. 4. Predictive Chl-a mappings from Jan. to Dec. in 2010

4.2 Validation and evaluation of the predictive maps

The objective of this study was to assess the performance of Chla algorithms available for CCD in coastal ocean waters in the Northern of South China Sea-Hong Kong. We used in situ and CCD $R_{rs}(\lambda)$ to estimate Chla under fuzzy logic and we also need to validate the precision of the model to prove the rationality and versatility of the model. In the accuracy of the model validation, firstly,

the sampling points were grouped, and then we used a combination of modeling and verification points to test the accuracy of the results of each month separately. Table 2 shows the validation results of each month: the fuzzy function model on chlorophyll concentration in the study area has good forecasting results. $R^2 \geq 0.88$, $ME \geq 0.85$, as mentioned above, this indicate a good prediction.

Table 1. Results of the fuzzy function model validation with the training samples.

Variables	R ²	RMSE; (RMSE %)	MAE; (MAE %)	ME	APE%
January	0.97			0.97	11.56
February	0.91	0.23(14.70)	0.09(5.76)	0.92	9.35
March	0.91	0.13(12.34)	0.08(7.39)	0.90	15.92
April	0.95	0.44(17.95)	0.32(12.87)	0.95	1.29
May	0.93	0.03(4.62)	0.007(1.24)	0.93	29.91
June	0.93	2.88(25.44)	1.54(13.63)	0.92	10.93
July	0.88	1.47(25.90)	0.53(9.34)	0.85	23.72
August	0.90	3.52(29.93)	2.22(18.87)	0.91	14.68
September	0.95	4.44(21.41)	2.63(12.69)	0.95	15.33
October	0.95	2.34(16.99)	1.33(9.54)	0.95	31.81
November	0.95	0.99(34.77)	0.39(13.90)	0.95	14.07
December	0.94	0.40(17.91)	0.24(10.58)	0.95	12.62

5 Conclusion and discussion

Through the above study, we found that the construction of membership functions for predictive Chla concentration under fuzzy logic have a good accuracy and the seasonal application. Different from using single exponential function modeling, this paper introduced the ideas of fuzzy logic. On the basis of the former research, we added the discriminant membership of pixels and introduced barrel principle to deeply mining data relationship. It greatly improved the accuracy of the model. Compared with the currently popular three-band model (The popular three-band model for Chla (Chla) estimation based on the absorbing and scattering properties of different water color matters, it combines reflectance values at three bands subtly and presents a function between the combination and Chla concentrations. It has a satisfactory precision and is easy to realize by avoiding the use or measurement of some environmental parameters and water optical parameters), it can dig deep relationship between the measured data and predicted data directly and we used this model twice, principle of buckets to significantly improve the accuracy of the model. And as no doubt it is very easy to establish and apply to the broad sea area. It has profound significance for the red tide on marine pollution monitoring.

We used CCD images provided by the China Centre for Resources Satellite Data and Application HJ1A/B satellites which are independently developed by China to calculate the Chla concentration in the Northern of South China Sea-Hong Kong, and have achieved good results. It also proved that HJ1A/B satellites application prospect in ocean color remote sensing.

In this paper, the results from the data selection and validation, evaluation of Chla concentration in the Northern of South China Sea-Hong Kong has achieved

very good results, but as the same with the other models it is also difficult to parameterize the model's coefficients uniformly for different waters, the same water with different regions, even the same region in different seasons. So next more research is necessary and we will test this model with different waters, different regions and different seasons. Make sure it will play a bigger role in the Marine pollution monitoring.

This work was supported by the Young Talents Programme of State Grid Energy Research Institute Co., Ltd. (Research on the new urbanization development indicators of energy and electricity supply and demand analysis based on night lighting remote sensing data for smart grid, No. XM2018020037579).

References

1. Gitelson, A. A.. The peak near 700 nm on reflectance spectra of algae and water: Relationships of its magnitude and position with chlorophyll concentration. *International Journal of Remote Sensing*, 13, **3367–3373** (1992).
2. Huang W, Mukherjee D, Chen S Assessment of Hurricane Ivan impact on chlorophyll-a in Pensacola Bay by MODIS 25m remote sensing. *Marine Pollution Bulletin* 62(3):**490-498** (2011).
3. Afanasyev YD, Nezlin NP, Kostianoy AG Patterns of seasonal dynamics of remotely sensed chlorophyll and physical environment in the Newfoundland region. *Remote Sensing of Environment* 76(2):**268-282** (2001).
4. Bresciani M, Stroppiana D, Odermatt D, Morabito G, Giardino C Assessing remotely sensed chlorophyll-a for the implementation of the Water Framework Directive in European perialpine lakes. *Sci. Total Environ.* 409(17):**3083-3091** (2011).
5. Cannizzaro JP, Carder KL Estimating chlorophyll a concentrations from remote-sensing reflectance in optically shallow waters. *Remote Sensing of Environment* 101(1):**13-24** (2006).
6. Le C, Hu C, Cannizzaro J, English D, Muller-Karger F, Lee Z. Evaluation of chlorophyll-a remote sensing algorithms for an optically complex estuary. *Remote Sensing of Environment* 129(0):**75-89** (2013).
7. Li F, Xu J, Ma R, Duan H, Zhang B. Chlorophyll-a retrieval in inland waters based on a regional three-band model. *Journal of Remote Sensing* 15(6):**1156-1170** (2011).
8. Liu M, Liu X, Wu M, Li L, Xiu L. Integrating spectral indices with environmental parameters for estimating heavy metal concentrations in rice using a dynamic fuzzy neural-network model. *Computers & Geosciences* 37(10):**1642-1652** (2011).
9. Mishra S, Mishra DR. Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. *Remote Sensing of Environment* 117(0):**394-406** (2012).

10. Spyrakos E, González Vilas L, Torres Palenzuela JM, Barton ED. Remote sensing chlorophyll a of optically complex waters (rias Baixas, NW Spain): Application of a regionally specific chlorophyll a algorithm for MERIS full resolution data during an upwelling cycle. *Remote Sensing of Environment* 115(10):**2471-2485** (2011).
11. Wang MH, Shi W, Tang JW. Water property monitoring and assessment for China's inland Lake Taihu from MODIS-Aqua measurements. *Remote Sensing of Environment* 115(3):**841-854** (2011).
12. Zhu AX, Yang L, Li B, Qin C, Pei T, Liu B. Construction of membership functions for predictive soil mapping under fuzzy logic. *Geoderma* 155(3-4):**164-174** (2010).