International Journal of Computer Science and Applications, 2008, Vol. 5, No. 3a, pp 33 - 40

MODELING OF THE SUSPENDED PARTICULATE MATTER IN THE ALGERIAN COAST USING NEURAL NETWORKS AND MATHEMATICAL MORPHOLOGY

HAMOUD MERRAD, SALIHA LOUMI, AND BOUALEM SANSAL

University of Science and Technology of Houari Boumediene,

Faculty of Electronics and Computer, Laboratory of Image Processing and Atmospheric Radiation

P. O. Box 32 El-Alia Bab-Ezzouar, 16111, Algiers, Algeria

merradhamoud@yahoo.fr salihaloumi@yahoo.fr bsansal@hotmail.com

ABSTRACT

In this paper, we propose a methodology for the characterization of the suspended particulate matter along the Algiers's bay. An approach by multi-layer perceptron (MLP) with training by back propagation of the gradient optimized by the algorithm of Levenberg-Marquardt (LM) is used. The accent was put on the choice of the components of the base of training where a comparative study made for four methods: Random and three alternatives of classification by K-Means. The samples are taken from suspended matter image, obtained by an analytical model based on polynomial regression by taking account of in situ measurements. The mask which selects the region of interest (water in our case) was done by using a multi spectral classification with ISODATA algorithm. To improve the result of classification, a cleaning of this mask was carried out using the mathematical morphology tools.

KEYWORDS

Classification, Mathematical Morphology, Neural Network, Suspended Particulate Matter (SPM).

1. Introduction

The oceans cover more than 70% of the surface of the Earth. They play an important role in the balance of the planetary ecosystem. The observation of the ocean from space allowed oceanographers to have on the one hand a global view of the phenomena observed and on the other hand to establish the relation between their temporal and spatial variability [1]. The colour of the oceans is mainly due to the phenomena of the diffusion and absorption.

The sea water does not only consist of pure water, it also contains alive substances and various particles which modify its colour. The variation of concentration of these various components determines the optical properties of the ocean.

In coastal and estuary zones, the mineral particles are comparable with the suspended materials. Often, only the properties of diffusion of this suspended particulate matter (SPM) are considered and the properties of absorption are neglected. This approximation does not have a consequence in water of broad called water of the case 1 where the suspended materials are rare. But, in coastal water called water of case 2, the optical properties of water is much more complex to define because of diversity of the dissolved terrestrial contributions and materials in suspension.

The validation of the visible imagery and near infra-red in coastal water in general requires the description of empirical algorithms relating to a specific zone of study. Until the day of today, rare works are developed based on empirical algorithms, valid for particular sites [3]. It is within this framework that our work is registered insofar as our objective is the characterization and the cartography of the suspended particulate matter along the Algerian bay.

A significant fraction of the measurement observed by the satellite comes from the contribution of the atmosphere which overhangs the ocean. Marine reflectance accounts for only 10% of the total reflectance observed in blue and less still in the biggest wavelengths. It is thus essential to do a correction of the atmospheric effects. Georeferencing is also an essential step, it allows coinciding the positions of in situ measurements with the data images acquired by the sensors. These various treatments were carried out upstream of methodology [4].

2. Site of study and Sensor used

The region of our study is the bay of Algiers with geographical co-ordinates: from 2°45 ' E to 3°45 ' E and from 36°30 ' N to 36°60 ' N. The data used are multi spectral images acquired by the Algerian sensor Alsat-1 combined with in situ measurements.

The micro satellite Alsat-1[4] launched on November 28, 2002, belonged to a new generation of micro sensors with great control capacity on altitude and orbit, and a high flow of remote loading. It is the first of a series of 05 micro sensors launched within the framework of the DMC (Disaster Monitoring Constellation) for the period going from 2002 to 2005.

The imagery system used is the ESIS (Extended Swath Imaging System), It rests on three spectral bands: (Green: 0.523 µm - 0.605µm), (Red: 0.629 µm - 0.690µm) and (Near Infra-red: 0.774 µm - 0.900µm).

3. Image of Reference of the Suspended Matter

In coastal water considering the space-time variability of its components, is difficult to model the suspended matter parameter [6][7]. We propose by our paper to characterize this parameter by the neural networks model. The number of in situ measurements being reduced, we propose to create an image of reference of the suspended matter using in situ measurements and establishing an analytical model which will be used as reference image for our network. The model used is by polynomial regression integrating in situ measurements and the radiometric values of spectral band at the points of measurements. The spectral bands used are those offering the greatest value of the correlation coefficient which analytical expression is given by (1)

$$R = \frac{\sum (X - \overline{X})(Y - \overline{Y})}{\sqrt{\sum (X - \overline{X})^2} \cdot \sqrt{\sum (Y - \overline{Y})^2}} \qquad (1)$$

Where the vector Y represents in situ measurements of the suspended matter and Y its average, X is the standardized radiances of the various spectral bands at the point of measurements and Xits average.

- Creation of the Mask

The mask is obtained by using an unsupervised (no knowledge a priori) multi spectral classification with ISODATA algorithm. It consists in creating a partition of the multi spectral image (3 bands of the Alsat-1 image) in two classes (water/land). Zero value (0) is allotted to the land whereas the value one (1) is allotted to water.
To improve the result of this classification, a cleaning of the mask is necessary (Figure 1). We thus called upon the tools of mathematical morphology.

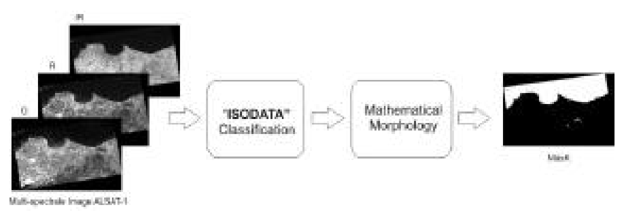

Figure 1: Diagram block of the determination of the mask.

4. Mathematical Morphology

Mathematical morphology [8][9] considers that the distribution of the values in the image forms objects. The morphological operations consist in transforming sets by the action of a structuring element B modifying the set by union, intersection or complementation. The choice of the structuring element depends on the type of required information.

5. Multilayer Perceptron Model (MLP)

We propose a multi-layer perceptron (MLP) as model for the determination of the concentration of the suspended matter [10]. The network model has a layer of input, a hidden layer and an output layer having 3, 5 and 1 neurons respectively. The training of the network is done by back propagation of the gradient optimized by the algorithm of Levenberg-Marquardt (LM) [11]. The result image obtained is multiplied by the mask previously given (Figure 2).

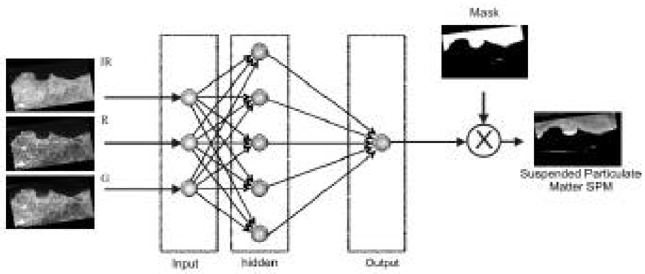

Figure 2: Modelling of SPM by MLP

In the multi-layer perceptron (MLP) with training by back propagation of gradient [12], the gradient gives the direction towards which it is necessary to move to find the minimum of the error, but does not give the step with which one must modify the weights of the network to make decrease as fast as possible this error [13], indeed this step is a fixed or adaptive coefficient (variable for each iteration). The algorithm of Levenberg-Marquardt [13][14][15] makes it possible to determine this step by using the second derivative expression of the average quadratic error.

a) Creation of the Training Base

The base of training of such neural network strongly influences the quality of the result. An optimal choice has a primary importance. This choice strongly depending on the number and the distribution of the samples, two steps was considered and compared

1- Selection in a random way of the samples.

2- Unsupervised multi spectral classification in K classes of the whole samples corresponding to the water part. In phase of training, we have fixed the average quadratic error to reach at 10^{-7}. In phase of generalization, the comparison criterion is the root mean square error (RMS) which has expressed by equation 2.

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Yref - Yest)^2} \qquad (2)$$

Where Yref and Yest are the reference and estimated images of SPM. N is the point's numbers of the image reference.

For the first step, we carry out a selection of a K number of samples among the whole of the samples corresponding to the water part and their correspondents in the SPM image of reference. It is obvious, by this manner of proceeding, that no control on the spectral distribution of the samples of this base is possible, and that only the control of the number can be it.

b) K-Means Classification

The base of training is made up after an unsupervised multi spectral classification of all the samples in K classes by using the K-Means algorithm of the pixels of the water part. We will have to consider the K nearest samples to the barycentre of K classes. The co-ordinates in row and column of each representative class will indicate SPM in the image of reference (Figure 3).

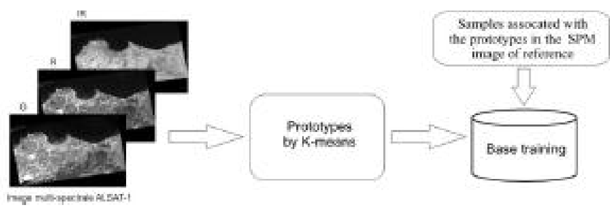

Figure 3: Determining of the training base.

Three alternatives of the technique were considered:

1. **With redundancy:** all the water pixels are considered. It is clear that the same pixel will be in its class as many once as it to appear in the image. The barycentre of this class will be strongly influenced.
2. **Without redundancy**: the same pixel is considered only once in classification.
3. **Without redundancy but with substitution of the identical samples**: this substitution will be done after classification while not taking (as far as possible) as representatives of classes only those which have the different SPM ones.

6. Results and Application

With our data the highest correlation is given in Table 1 by the spectral band (0.523 µm - 0.605µm). This band was used for the polynomial regression and the obtained model is given by equation 3.

$$\hat{y} = -635.3 + 2.57x - 0.1197x^2 \quad (3)$$

Band (µm)	0.523- 0.605	0.629-0.69	0.774 - 0.90
R^2	0.9049	0.8016	0.5486

Table 1: Correlation between *in situ* measurements and Alsat-1 image.

Classification by ISODATA gives a rather good separation between the sea and the land (almost smooth line of coast) to the detriment of some imperfections.

For the choice of the training base, a study according to the size of the base emphasizes (Table 2) that 100 samples taken in a random way give a result better than that of 500 samples and the best precision is obtained for 800 samples. The evolution of the error seems in contradiction with work affirming that more the number of samples is high, better is the training thus the precision. We sought for each case, the number of effective samples (those which do not have the same value of SPM) (Table 2). It is clear that the weakest error of generalization was obtained for 800 samples, and that this corresponds to a number of effective highest samples (37).

Samples	100	300	500	700	800	900	1000
RMS	0.934	1.333	1.315	0.480	0.468	0.507	0.556
Effective samples	22	30	29	31	37	37	34

Table 2: Influence of the size of the base in the random case.

As for the choice of the prototypes by the second alternative, the study was done for seven (07) bases of sizes: 10, 30, 50, 70, 80, 90 and 100 samples. Each element of a base corresponds to a representative of a class. The number of effective samples for each alternative is given by table 3.

Samples / Effective samples	10	30	50	70	80	90	100
With redundancy	9	16	21	17	24	23	26
Without redundancy	8	23	29	33	32	32	36
Without redundancy and substitution	10	30	49	46	50	50	50

Table 3: Effective samples for the three alternatives.

According to curves of the root mean square error in generalization (Figure 4), we note that classification without redundancy and substitution gives a better precision for a number of samples between 15 and 35. Beyond 65 samples, the best results were given by classification without redundancy. This error RMS varies very little beyond 80 samples. This result not being foreseeable, we then tried to find an explanation, by considering the two particular bases: 30 and 90 samples. For these two cases, the performances of the two steps quoted previously are reversed compared to the third alternative (classification with redundancy) (Figure 4).

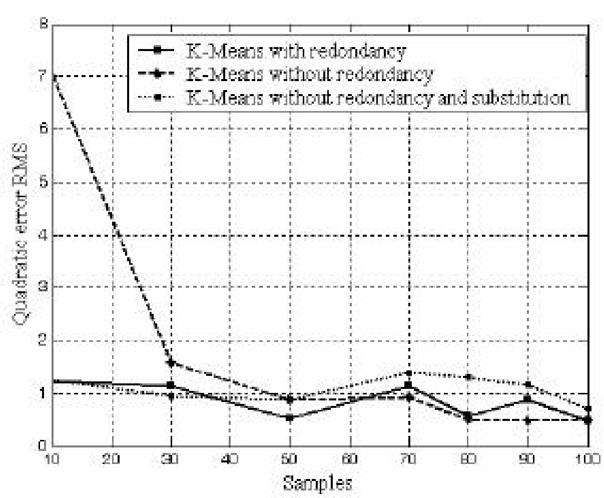

Figure 4: Influence of samples to the quadratic error for three alternatives.

We tested our methodology in a field of particular interest, and which lies within the scope of the natural disasters such as the floods and the seism, to explain certain direct or indirect consequences on coastal water. We used an image of the 24/5/2003 (03 days after the earth quake which struck the area of Boumerdes which is locate at 40km in the east of Algiers), acquired by the Alsat-1 sensor and which covers the zone touched by the seism. We cut out this rough image a scene of size 600x2400 pixels which we corrected and georeferenced.

We then applied the neural model determined previously to obtain the image of the suspended matter. The result image of generalization (Figure 5) reveals us a zone of strong suspended matter concentration localised close to the epicentre (white colour). This can be justified by the fact that under the effect of the jolt and the repeated counterparts, an agitation of sea-beds caused an increase of the suspended matter towards the surface of the sea.

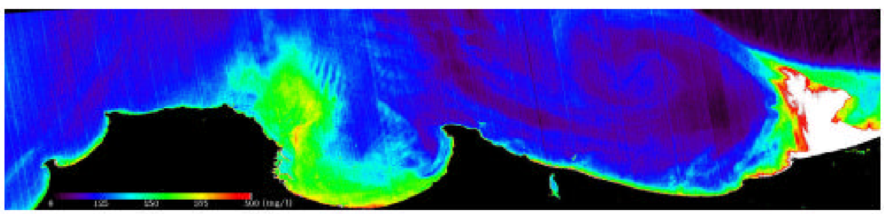

Figure 5: Image Result of the Suspended Particulate Matter corresponding to 24/05/2003.

7. Conclusion

This study is the first results of a methodology in which we combined multi spectral images of the ALSAT-1 sensor and in situ measurements for a characterization of the suspended particulate matter (SPM) along the Algerian coast. A model by the multi-layer perceptron (MLP) was implemented. An architecture corresponding to three (03) input neurons, a hidden layer of five (05) neurons and an output with a one neuron was adopted. In this step, the accent was laid on the choice of the prototypes of the training base where a comparative study was made for four different methods: a random selection of the samples and three alternatives of the selection by K-Means classification. Through this study, it appeared that for the constitution of the training base, two significant elements are to be taken into account:

• The Richness of the base which is expressed by the number of samples.
• The representativeness of the base which depends on the distribution of these samples.

A better representativeness is obtained when the samples occupy all the dynamics of the possible values. A particular care was to bring to the creation of the mask, where an unsupervised classification by ISODATA and cleaning by the quite selected tools of mathematical morphology gave full satisfaction. By the method suggested we could on the one hand improved the performances of the network by carrying out a good compromise between a number of samples and representativeness of the base and on the other hand highlighted the phenomenon of the increase of the suspended matter towards surface of the sea, had with the agitation of the sea-beds, caused by the repeated jolt and counterparts which touched the area of Boumerdes on May 21th, 2003.

Acknowledgment

The authors wish to thank the ASAL (Algerian Agency Space) Institute for the Alsat-1 images, and the ISMAL institute (Institute of Marine Science of the Algerian Littoral) for in situ measurements.

References

[1] D. K Hall., J. R. Key, Casey K. A., G. A. Riggs , D. J. Cavalieri., Sea Ice Surface Temperature Product from MODIS, IEEE Trans. Geoscience Remote Sensing, , 2004, vol. 42, no. 5, May pp. 1076-1087.
[2] L. Parkinson, Aqua: An Earth-Observing Satellite Mission to Examine Water and Other Climate Variables, IEEE Trans Geoscience Remote Sensing, 2003, vol. 41, no. 2, Feb pp. 173–183.
[3] J. M Froidefond , D. Doxaran, Télédétection Optique Appliquée à l'étude de eaux côtières, Télédétection, 2004, vol. 4, no. 2, pp. 157-174.
[4] H. Merrad, Caractérisation des Eaux Côtières à partir d'Images Multi spectrales de MODIS et ALSAT-1 : Application au littoral Algérien, Thèse de Magister, 2005, Université des Sciences et de la Technologie Houari Boumediene à Alger,.
[5] M. Bekhti, A. Oussedik, J.R. Cooksley. Alsat-1: Conception details and in orbit performance, Actes des Journées Techniques ALSAT 1/ Utilisateurs, 2003, Juillet 14 et 15 , Alger.
[6] D.B. Patissier, G.H. Tilstone, V.M. Vincente & G.F. Moore, Comparaison of bio-physical marine products from SeaWiFs, MODIS and a bio-optical model with in situ measurements from Northern European waters, .2004, Journal of Opitcs, pp. 875-889.
[7] D. Doxaran, J. M. Froidefond, S Lavender & P. Castaing, Spectral signature of highly turbid waters Application with SPOT data to quantify suspended particulate matter, Remote Sensing of Environment, 2002,81, 149-161,
[8] M.Coster ET J. L.Cherman, Précis d'analyse d'images, 1985, Éditions du CNRS.
[9] J. Serra. Image Analysis and Mathematical Morphology, 1982 Academic Press, London.S.
[10] Haykin, Neural Networks: A Comprehensive Foundation, 1999 Second edition, Prentice Hall.
[11] F. Alilat, S. Loumi, H. Merrad & B. Sansal, Nouvelle approche du réseau ARTMAP Flou Application à la classification multispectrale des images SPOT XS de la baie d'Alger, Revue Française de Photogrammétrie et de Télédétection SFPT, (2005-1),.no.177, pp 17-24.
[12] Z. Zhou, Chen & Z. Chen, FANNC: A Fast Adaptive Neural Network classifier, 2000, Knowledge and Information Systems, Vol. 2, N°1, pp. 115-129.
[13] P. Baldi, Gradient Descent Learning Algorithm Overview: A General Dynamical Systems Perspective, 1995, IEEE Transactions on Neural Networks, Vol. 6.
[14] M. Hagan T. & M .B. Menhaj, Training Feed forward Networks with the Marquardt Algorithm, , 1994, IEEE Transactions on Neural Networks, Vol.5 No 6.
[15] B.M. Wilamowskin, S. Iplikci, O. Kaynak & Efe, An Algorithm for fast Convergence in Training Neural Networks, 2001, IJCNN'01, Washington C. c July 15-19 , pp. 1778-1782.