ANALYSIS OF SEAWATER POLLUTION USING NEURAL NETWORKS AND CHANNELS RELATIONSHIP ALGORITHMS*

JENI VASILESCU, LUMINITA MARMUREANU, EMIL CARSTEA

National Institute of Research and Development for Optoelectronics, INOE 2000, 409 Atomistilor Street, Magurele, Ilfov, Romania, jeni@inoe.inoe.ro

Received November 4, 2009

The aim of this paper is to establish the most accurate and fastest method of Lidar signal analysis used for fluorescent water compounds identification. This study analyzes data acquired in 2007 late summer campaign on Romanian part of the Black Sea. Artificial Neural Network and channels relationship methods were evaluated in order to identify the best choice for fluorescence compounds detection and characterization.

The Artificial Neural Network processing method proved to be useful for online recognition of oil spills, the sensitivity is dependent on training data set and training rule. It was found that is suitable to use channels relationship in order to reveal the distribution of fluorescence water components on a path and is suitable to use Artificial Neural Network to identify the type of fluorescence water components.

Key words: fluorescence Lidar signal, Artificial Neural Network, Channels Relationship.

1. INTRODUCTION

Fluorescence Lidar is a useful tool in water pollution monitoring. Based on fluorescence signature of each contaminant, the Lidar can be used to detect organic matter on sounding volume with a very high accuracy and in a very short time. During the last decades studies concerning phytoplankton distribution, pollution levels, hydrographic parameters temporal evolution have been carried out using Lidar systems [1] [2] [3] [4].

Main advantage of Lidar is the real time observation of fluorescent water components as oil spills, natural and anthropogenic organic matter and pigments, which can be further used to identify the contaminants concentration along the detection trajectories. The technique is non-invasive with detection limits from ppb to ppm, depending on fluorophore.

* Paper presented at the "Optoelectronic Techniques for Environmental Monitoring" (OTEM-2009), September 30–October 2, 2009, Bucharest, Romania.

The fluorescence signals are complex data, with fluorophore fingerprints overlapped. Their processing take long time and require powerful hardware and software.

This study is focused on the analyses of the accuracy and percentage of classification of laser induced fluorescence LIDAR data in real time using the Artificial Neural Network or Channels Relationship algorithms, which are considered proper for water fluorescence pattern recognition [5] [6].

We used data measured with a fluorescence monostatic Lidar during a field campaign along the Romanian part of the Black Sea coastal zone, from Danube Delta Sf. Gheorghe discharge branch to Agigea harbour and around oil platforms. The experimental cruises were performed during August-September 2007, on seven days, onboard of Mare Nigrum research vessel. The areas of investigation contain zones with intensive anthropogenic influences and oils sources.

Three transects along costal zone were done, first being on 20m bathymetric line in order to sense the anthropogenic and Danube Delta influences. Also round circles were done around six oil platforms for oil spills and emulsion identification from platforms or pipelines liking (fig. 1). The Lidar was used to specifically detect oil derivates, dissolved organic matter and chlorophyll a.

Fig. 1 – Lidar measurements map for August-September 2007 campaign.

2. METHODOLOGY

2.1. FLUORESCENCE LIDAR SIGNAL FOR SEAWATER

The intensity of the returned Lidar signal is proportional to the fluorescence molecule concentration when a homogeneous water column with constant attenuation coefficient and fluorescence efficiency is considered [7], [8], [9]. For a pulsed, monostatic system [10], the fluorescence Lidar equation is given by:

$$\Delta P(\lambda, R) = T_r(\lambda)\frac{A_r}{R^2}T(\lambda, R)\xi(R)A(R)W(\lambda, R)\Delta R\Delta\lambda \tag{1}$$

where λ represents the wavelength sensed by the detector, R distance to the target, $\Delta P(\lambda, R)$ signal power received by the detector, $W(\lambda, R)$ spectral radiance per unit, the other factors are functions which describe the system, atmosphere transmission and geometry of measurements configuration.

In the case of a fluorescent substance into water medium the laser radiation excite a number of molecules, which will jump on the excited energy state. The variation of excited number of molecules can be described as follows:

$$\frac{\mathrm{d}N(R,t)}{\mathrm{d}t} = \frac{\lambda_L \sigma^A(\lambda_L)}{hc}N(R,t)I(R,t) - \frac{N(R,t)}{\tau} \tag{2}$$

where $\sigma^A(\lambda_L)$ is the aborption cross section, $N(R,t)$ the number of molecules on the ground state, $I(R,t)$ laser beam power density, τ the fluorescent species lifetime. The Lidar equation specific for fluorescence targets takes into account the number of molecules excited, meaning the concentration of fluorescence substances.

The LIDAR signal can be used to investigate water quality parameters by evaluating the main water components through their spectral fluorescence fingerprints: phytoplankton dissolved organic matter (DOM), proteins and petroleum products. DOM contains amino acids, nucleic acids, hidrocarbours, fenolic compounds and humic substances [11]. These are natural macromolecules from detritus, phytoplankton, or leafs decomposition, with colors from yellow to brawn. Also the DOM quantity can be increased due to human activities and pollution. The proteins contain three fluorescent amino acids: tryptophan, tyrosine and phenylalanine. First two are usually found in aquatic medium, mostly as components of bacteria membranes [12]. The refined oils products contain hydrocarbons, which can be excited using UV wavelengths and fluorescence on visible interval (400-650 nm) [13].

2.2. METHODS FOR WATER COMPONENTS IDENTIFICATION

During the last years the Artificial Neural Networks and Channels Relationships methods have been frequently used to process spectral signals [6]. An Artificial Neural Network (ANN) represents a mathematical projection of the human neural network, based on neurons, axons and synapses. The information is propagated as a neural influx and classified.

The architecture and learning rule of a neural network depends on application and on complexity of data set for analysis. The network can contain tens or hundreds of neurons, the output of the first becoming the input to the next layer [14], [15]. The most used artificial neuron is the Perceptron. The activation function for this is step function and the weights are adjustable.

$$S = \sum_{i=1}^{n} x_i w_i \tag{3}$$

where n represents the number of input signals, x the input signal, w the synaptic weight, with following activation function:

$$Y = \{+1, S \geq \theta \text{`sau`} -1, S \leq \theta \tag{4}$$

The Perceptron permits classifications of data into classes of patterns, splitted by a hyperplane defined by the following equation:

$$\sum_{i=1}^{N} x_i w_i - \theta = 0 \tag{5}$$

In this paper, a *Multilayer Perceptron* with supervised learning is used, comparing the output result with the desired one. The learning process implies changing connection weights after each data processing based on the amount of error in the output compared to the expected result.

The *channels relationship method* (CRM) is a simple method, computing the relation of relative integral intensities of fluorescence signal [6]. The fluorescence signal is divided into an arbitrary number of channels depending on needed sensitivity or distinct spectral fluorescence fingerprints. In this paper the fluorescence spectra were divided into a Raman peak and other 50 channels in order to interpret well the oil spills and emulsions.

3. RESULTS AND DISCUSSIONS

A typical Lidar signal after using a 308 nm excitation wavelength of seawater column has a maximum at 344 nm due to water Raman scattering, a maximum around 450 nm corresponding to the dissolved organic matter (DOM), and sometimes peaks due to oil derivates presence (fig.2). A Lidar signal typical for 460 nm excitation wavelength shows the Raman band mixed with blue-green fluorescence and chlorophyll *a* at 685 nm.

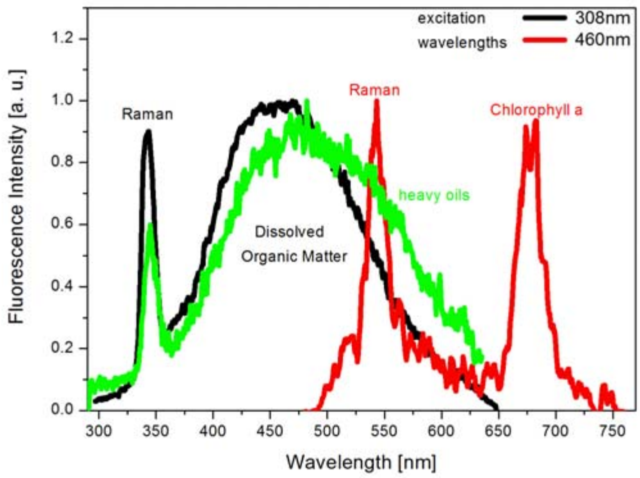

Fig. 2 – Typical Fluorescence Lidar signal for 308 and 460 nm excitation wavelengths.

3.1. OIL SPILL IDENTIFICATION

The data analyzed were obtained with a monostatic Lidar during the field campaign on the Romanian Black Sea coast, on harbors and around oil platforms. The CRM algorithm was applied for oil emulsion and spills detected around platforms (fig. 3).

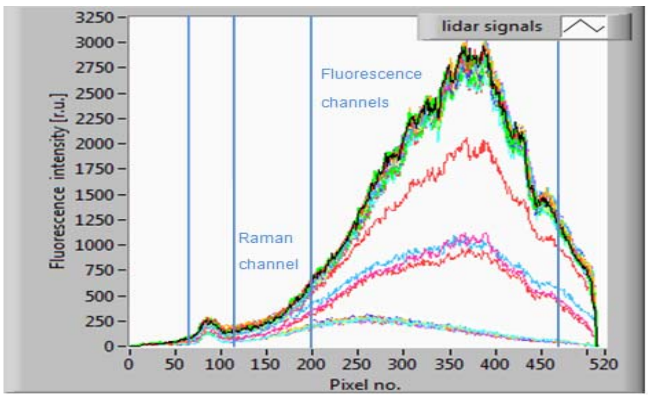

Fig. 3 – Lidar signal measured around oil platforms; oil emulsion can be emphasized.

The presence of oil derivates on seawater modified the shapes of dissolved organic matter, being difficult to recognize the fluorescence pattern. The Channels ratio on fig. 4 represents the relation between oil emulsion fluorescence integral area (500–580 nm) and water Raman signal, respectively between oil emulsion and DOM (425–450 nm) fluorescence integral area.

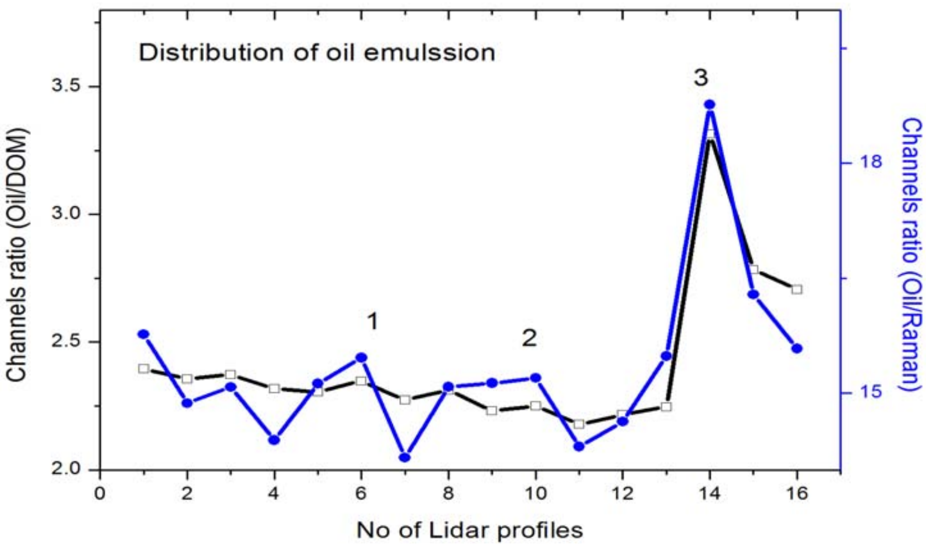

Fig. 4 – Lidar signal analyzed using Channels relationship.

The oil presence was detected in all areas around platforms, but is significant for three Lidar profiles where the ratio is approximately 1.2–1.5 times higher than in previous measurements. The oil Raman ratio can better identify small amounts of oil emulsions. The differences between values are more important comparing with those obtained with the second ratio because oils attenuated the incident laser beam and a small amount of water molecules are excited.

The CRM doesn't allow identification of signals with very similar spectral fingerprints like oil spills and emulsion, but is proper to be used for quick mapping of relative oil spills distribution on a path [16].

The data acquired during campaign were further analyzed using ANN to identify the presence of typical seawater DOM or oils. For input, 5 channel files with 100 mixed data sets, containing fluorescence fingerprint of DOM, oil emulsion and oil spills were used.

The channel number 3 is the most important for signal classification as can be seen in fig. 5, because here the specific differences between those three patterns for processing can be pinpointed. In some cases the channel number 3 represents the sum between the other channels. An increase number of channels can take into account all specific features of fluorescence fingerprints, the ANN being capable to identify more complex patterns.

The oil fingerprint is not well identified by ANN, being the most complex spectra with high difference of information on all five channels. The ANN is capable to well classify the simplest spectra which has very similar information on all channels. For best pattern recognition is necessary to enlarge the number of channels and database for training.

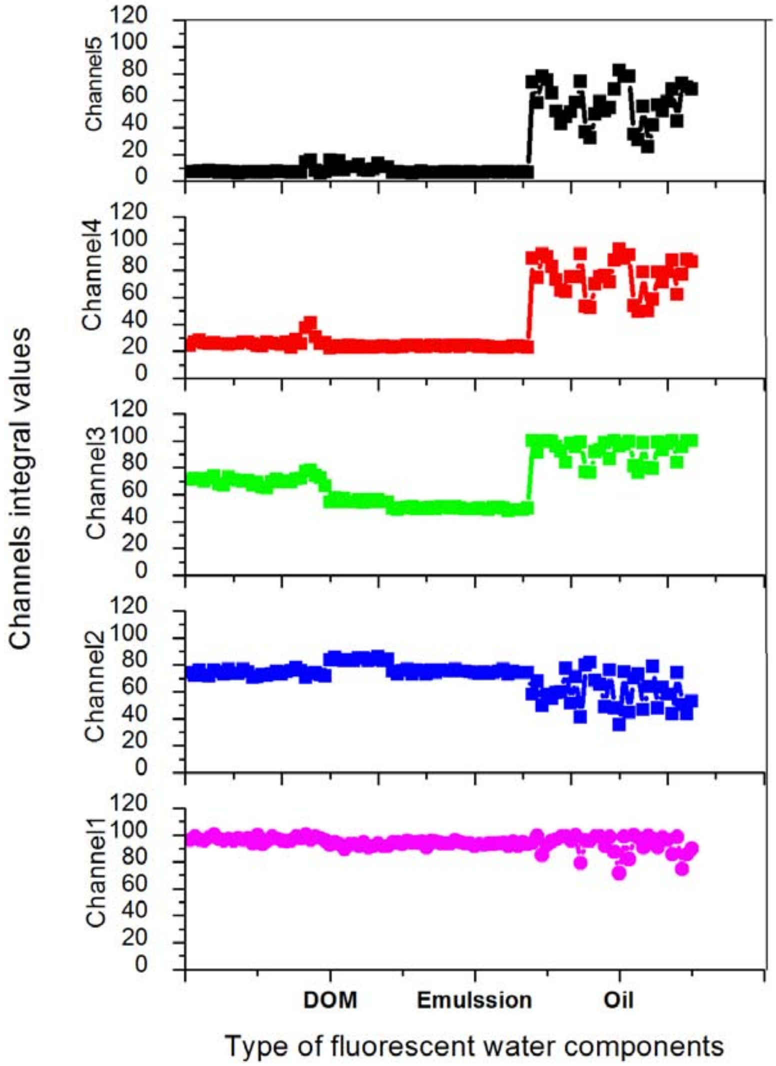

Fig. 5 – Data inputs used for Artificial Neural Network training.

The smallest learning errors were observed for MLP with three hidden layers. Even if the learning and testing errors are small not all the signals are well classified, only 90%. The ANN recognizes all the signals and the errors are diminishing after several retraining.

The ANN processing method is useful for online recognition of oil spills; the sensitivity is dependent on training data set and training rule. The percent of well classified fluorescence fingerprints are dependent on the number of files used for ANN training process, but also on complexity of fluorescence spectra (fig. 6).

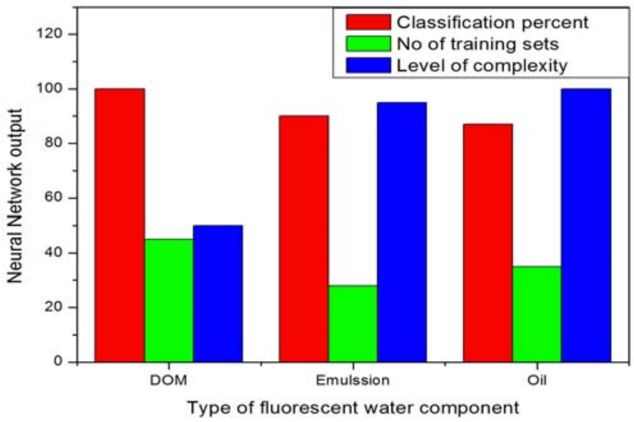

Fig. 6 – Percent of Lidar data well classified using Multilayer Perceptron.

3.2. RELATIONSHIP BETWEEN FLUORESCENCE INTENSITY AND PHYSICAL-CHEMICAL PARAMETERS

The Lidar system can be used to relatively sense the physical water parameters taking into account that fluorescence signal is direct or inverse proportional with oxygen concentration and salinity.

The microalgae and phytoplankton represent a trophyc bases for primer consumers, but after degradation become suspended organic matter food for zooplankton and benthos organisms. The amount of algae can be evaluated using the fluorescence intensity for chlorophyll a.

Good correlations between fluorescence intensity and water quality parameters have been demonstrated. The chlorophyll a fluorescence signal and standard measurements are well linked with oxygen concentration for the water surface layer (fig. 7, fig. 8).

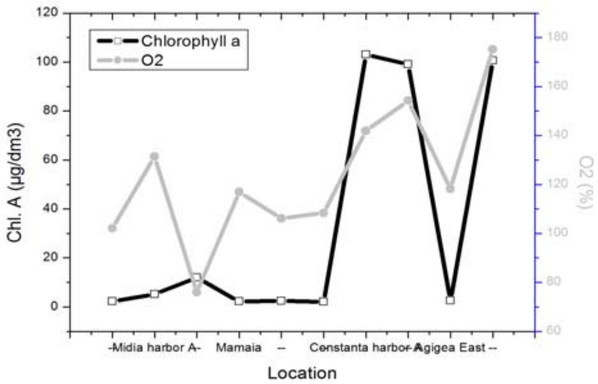

Fig. 7 – Relationship between chlorophyll fluorescence and oxygen.

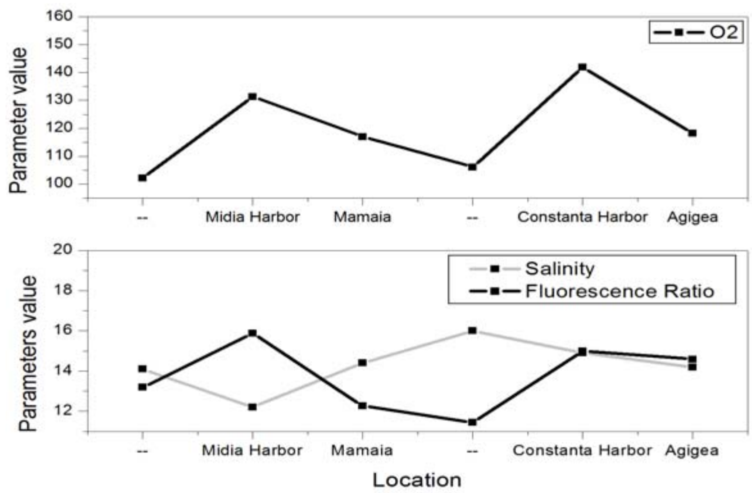

Fig. 8 – Spatial distribution of oceanographic variables.

4. CONCLUSIONS

During the 2007 Black Sea campaign the physical water parameters presented normal values characteristic to late summer season. The Black Sea represents a water basin with a high biological productivity and high level of rivers inputs.

The microalgae concentration and higher phytoplankton values were observed on harbors berths. The quantitative analysis showed the existence of positive correlations between physical – chemical water parameters and also between these and phytoplankton concentration.

The most important characteristics of the fluorescence Lidar method is the real time observation of fluorescence water components in a non-invasive mode, which can be further used to identify the oil spills distribution along the ship trajectories. The sensitivity of oil derivatives detection depends on device parameters and on data processing and analysis tools.

Two processing algorithms were used to analyze the oil spills and emulsions observed during 2007 Black Sea campaign: channels relationship and artificial neural networks. The complexities of ANN training data and of the fluorescence spectra determine a good classification process. The ANN proved to be a proper algorithm for classification of close fluorescence fingerprints, while CRM is good to reveal the fluorescent marine water components spatial distribution.

Acknowledgements. The LIDAR data classification was carried out using the NeuroSolutions software (www.nd.com). The authors wish to acknowledge National RTD Agency for LIDIA and IDEA contracts 64/2005, respective 32-148/2008.

REFERENCES

1. R. Reuter, D. Diebel and T. Hengstermann, *Measurements of hydrographic fronts in the German Bight and in the nothern Adriatic Sea*, Int. J. Remote Sens., 14, p. 823–148, 1993
2. V. Drozdowska, W Walczowski, R. Hapter, J. Ston, M. Irczuk, T. Zielinski, J. Piskozub, *Fluorescence characteristics of the upper water layer of the Arctic seas based on Lidar, spectrophotometric, and optical methods*, EARSeL eProceedings, 3, 1, p. 136–142, 2004
3. F. E. Hoge and R. N. Swiftt, *Airborne simultaneous spectroscopic detection of laser induced water Raman backscatter and fluorescence from chlorophyll a and other naturally occurring pigments*, 20 (18) Applied Optics, 1981
4. R. Karpicz, A. Dementjev, Z. Kuprionis, S. Pakalnis, R. Westphal, R. Reuter and V. Gulbinas, *Oil spill fluorosensing lidar for inclined on shore or shipboard operation*, Applied Optics, 45 (25), p. 6620–6625, 2006
5. A.K. Jain, R.P.W. Duin and J. Mao, *Statistical Pattern Recognition: A Review*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22, p. 4–37, 2000
6. K.M. Almhdi, P. Valigi, V. Gulbinas, R. Westphal and R. Reuter, *Classification with Artificial Neural Networks and Support Vector Machines: Application to oil fluorescence spectra*, EARSeL eProceedings 6, 2, p. 115–129, 2007
7. R. Measures, *Laser remote sensing: fundamentals and applications*, Krieger Publishing Company, Malabar, Florida, 1992
8. T. Hengstermann and R. Reuter, *Lidar fluorosensing of mineral oil spills on the sea surface*, Applied Optics, 29 (22), p. 3218–3227, 1990
9. E. V. Browell, *Analysis of Laser Fluorosensor Systems for Remote Algae Detection and Quantification*, NASA Technical Note, TN D-8447, NASA, Washington, D.C., 1977
10. R. Measures, *Lidar equation analysis allowing for target lifetime, laser pulse duration, and detector integration period*, Applied Optics 16(4), p. 1092–1103, 1977
11. Matthews B.J.H., Jones, A.C., Theodorou N.K.,Tudhope, A.W., *Excitation emission-matrix fluorescence spectroscopy applied to humic acid bands in coral reefs*, Marine Chemistry, 55, 317–332,1996
12. Lakowicz, J.R., *Principles of Fluorescence Spectroscopy*, Plenum Press, New York, 1983
13. Ellingsen G., Fery-Forgues S., *Application de la spectroscopie de fluorescence à l'étude du pétrole : le défi de la complexité*, Revue de l'institute Fancais du Petrole, vol. 53, nr.2, 201–216, 1998
14. P.J. Braspenning, F. Thuijsman, A. J. M.M. Weijters, *Artificial Neural Network - An Introduction to ANN Theory and Practice*, Springer, ISBN 3-540-59488-4, p.11–18,1995
15. J. Vasilescu, T. Onciu, L. Jugaru, L. Belegante, *Remote estimation of fluorescence marine components distribution*, Rom. Rep. Phys., vol 61(3), 2009