

An Optimized Fuzzy System for Coastal Water Quality Mapping Using Remote Sensing Data

Bahia Lounis^(✉) and Aichouche Belhadj-Aissa

Laboratory of Image Processing and Radiation, Faculty of Electronics and Computer Science,
USTHB University, BP 32, El-Alia–Beb-Ezzouar, 16111 Algiers, Algeria
lounisbahia@yahoo.fr, h.belhadj@mailcity.com

Abstract. In this paper, we propose coastal water quality mapping process using the combination of in-situ measurements and remote sensing data. Water maps is processed by an hybridization of fuzzy model and genetic algorithm which exploits remotely sensed multispectral reflectances to estimate coastal water quality. The relation between the water parameters and the subsurface reflectances is modeled by a set of fuzzy rules extracted automatically from the data through two steps procedure. First, fuzzy rules are generated by unsupervised fuzzy clustering of the input data. In the second step, genetic algorithm is applied to optimize the rules. Our contribution has focused on the use of several water parameter maps to construct a graphical tool named Pollution Signature Draw (PSD) in order to characterize the water quality. Water characterization is then evaluated by analyzing several types of PSDs related to typical sites selecting the most representatives' ones. After, the selected PSDs are introduced in a classifier system to generate a pollution map (PM) associated the studied area. The proposed approach was tested on Algiers bay and has highlighted four pollution levels corresponding to High Pollution (HP), Medium Pollution (MP), Low Pollution (LP) and Clear Water (CW).

Keywords: Remote sensing data · Fuzzy clustering · Genetic optimization · Pollution signature · Coastal water classification

1 Introduction

Under the pressure of the human and industrial activities, the Algerian coast had undergone profound environmental changes affecting the quality of its water. The current environment's ministry methods for establishing water pollution are analysis and measurements done in laboratories. These techniques are very accurate, however they remain insufficient and very expensive for assessing and monitoring water quality on Algerian coast that extends on 1200 km [1]. Actually, remote sensing can overcome this constraint by providing an alternative for water quality monitoring over a range of temporal and spatial scales. Imagery from recent satellites with improved spectral and spatial resolution and the integration of the Geographic Information System (GIS) technologies offer a valuable tool for developing management plans for water pollution and thus allowing quicker and more effective actions to be taken. In this work, we are interested to this

type of data to, first, map water quality indices. Then construct Pollution Signature Draw PSD in order to characterize the quality of this water. Finally, we propose a new water quality classification to highlight most polluted sites of Algiers bay.

Remote sensing of water quality monitoring is evaluated by several substances which affect its optical properties. Conventionally, three main components are used to estimate marine water quality in coastal areas. These components are: Suspended Particulate Matters “SPM”, chlorophylls “Chl” and dissolved organic matter “DOM” [2]. To these components, we are also interested to water Turbidity (Turb) and its transparency measured by Secchi Disk Depth (SDD).

Various approaches have been developed to estimate coastal water quality parameters from remote sensing data. Due to the simultaneous presence of the three main water components, the relation between water components and remotely subsurface reflectances is complicate and is considered as no-linear. So, the first developed algorithms were empirical and semi-analytical models [3, 4]. With the increase of spectral information and the complexity to solve the inverse model, new estimation algorithms inspired from natural phenomena have emerged. Neural networks were successfully used to implement the inverse model and to properly address the non-linearity problem [5]. Besides neural networks, fuzzy systems have proved to be particularly effective in identifying non-linear models too [6]. The most popular approach to fuzzy modeling is based on the identification of fuzzy rules, which describe in linguistic terms the water parameters concentration/marine subsurface reflectance relationship. Furthermore, an optimization process is usually added to tune these fuzzy rules so that the fuzzy method implements the desired inverse model. Using this approach, a mapping process of coastal water classification is proposed in order to estimate the boundaries of coastal zones according to different water types. To attend this goal, we have followed the next steps (see Fig. 1):

- **Water quality estimation maps:** water quality maps are estimated using fuzzy systems. It consists of identifying fuzzy rules [7] which describe the relationship between water quality parameters and their corresponding reflectance. These rules, extracted from unsupervised clustering of training data, determine the actions that water parameter must perform if some conditions on multispectral reflectance are satisfied. After, an accurate definition of the rules coefficients is assigned to genetic algorithm (GA). This latter searches the best chromosomic structure that codifies the different parts of the fuzzy rules to give best optimization results.
- **PSD analysing:** in this part, we focused our interest on the use of PSD tool to water quality characterization which is evaluated by analysing several types of PSDs related to typical sites in order to select the most representative’s ones [1].
- **Pollution map:** the main PSDs of typical sites are selected and introduced in maximum likelihood classifier in order to generate a pollution map related to Algiers bay. For this site, four pollution levels corresponding to “High pollution”, “Medium Pollution”, “Few Pollution” and “Clear Water” are considered.

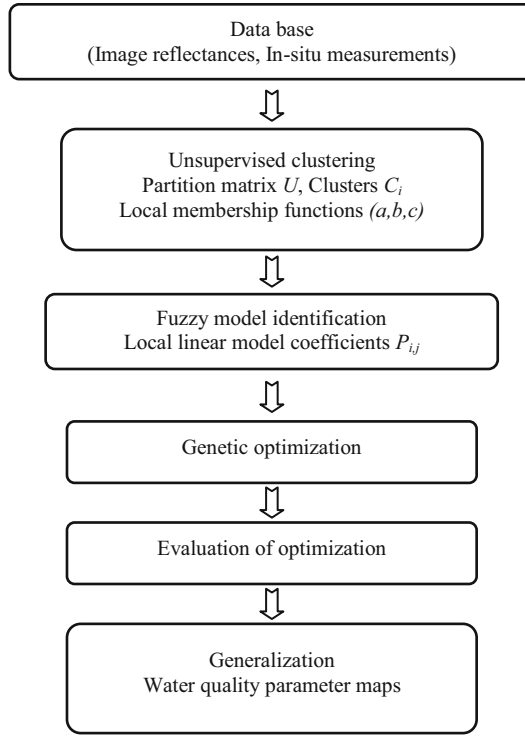


Fig. 1. Flowchart of estimation map process.

Our paper is structured as follows; in Sect. 2 the estimation principle of water parameters maps is described where an hybridization of fuzzy model and genetic algorithm is used. The PSD tool and water characterization are outlined in Sect. 3. Water pollution map is presented in Sect. 4. Finally, in Sect. 5, conclusions on the developed process are presented in order to show the contribution of remote sensing data to monitor the coastal water pollution.

2 Water Quality Estimation Maps

In the fuzzy estimation, Takagi Sugeno Kang (TSK) model was adapted to our application where the relation between water component concentration y_i and marine Subsurface reflectance R_i is considered as locally linear [8] (see Fig. 2). Thus, the input space is divided into crisp subspaces or clusters C_i . In each cluster C_i , y_i is assumed as a linear local function that models the variation of R_i in this cluster ($y_i = p_{i,0} + p_{i,1}R_i$, where $p_{i,0}, p_{i,1}$ are real numbers). The global model Y connects y_1 and y_2 according to their local activation degrees β_i and gives $Y = \beta_1 y_1 + \beta_2 y_2$. Whereas, in the optimization step, the fuzzy coefficients are applied to a genetic algorithm in order to tune them and improve the estimation map results. To implement this monetization, we took the following steps:

- Processed data: to carry out our approach, we have used one image corresponding to ETM+ of Landsat7 satellite covering Algiers bay. The image has been, first, geometrically corrected and transformed into radiances. After, it was corrected from atmospheric contribution by the dark point technique, converted into corresponding reflectance R and masked from the land areas. We have also used a set of in-situ measurements of SPM, Chl, Turb and SDD collected in Algiers bay. 300 samples of punctual measurements were provided by Institut National des Sciences de la Mer et de l'Aménagement du littoral (ISMAL) covering the studied area. They were collected two weeks before the satellite pass. The green points of Fig. 3 show their geographical locations on ETM+ image. According to the available data $y_i = (SPM, Chl, Turb, SDD)$, we have built a set of data (R_i, y_i) relevant to subsurface reflectances R_i and in situ concentrations y_i . Finally, we randomly split the data into two subsets, the training-set composed of the 2/3 of initial data is used to define the fuzzy model whereas the test-set to evaluate its performances.

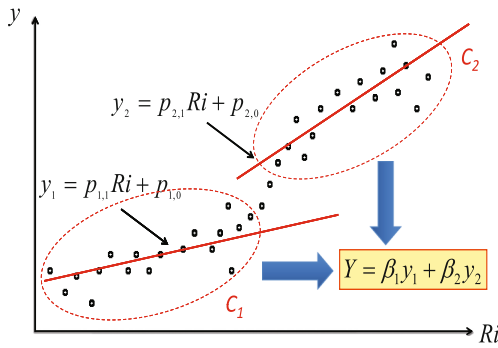


Fig. 2. Fuzzy estimation principle (Takagui-Sugeno-Kang model)



Fig. 3. Composed color of ETM + image (Red: TM1, Green: TM2, Blue: TM3) acquired on Algiers bay. (Color figure online)

- Determination of fuzzy input subspaces: using the Weighted Fuzzy C-means (WFCM) algorithm, we divide the training data-base into crisp subspaces or clusters C_i which share the same spectral characteristics [8]. To each cluster C_i , we define a fuzzy set A_i and associate a triangular membership function characterized by its parameters (a, b, c), where b corresponds to the abscissa of triangle vertex and a and c are deduced as the intersection of the abscissas axis with the lines on the left and the right sides of b ;
- Determination of fuzzy model coefficients: To each cluster C_i , the output y_i of in-situ concentration variation is assumed to be locally linear and expressed as following relation: $y_i = p_{0,i} + p_{i,1}R_1 + p_{i,2}R_2 + \dots + p_{i,M}R_M$, where R_M are the M subsurface reflectances and $(p_{i,0}, p_{i,1}, p_{i,2}, \dots, p_{i,M})$ are real numbers. The estimation model is, then, expressed as a set of conditions which assign each pixel to a water type under the form of Takagi-Sugeno-Kang (TSK) rules. The premises of these rules depend on the fuzzy sets $A_{i,j}$ defined on the reflectance domain and the consequences present the local linear y_i model form, like shown by Eq. 1 [8].

$$\begin{aligned} \text{Rule}_i: & \text{ If } (R_1 \text{ is } A_{i,1}) \ \& \ \dots \ \& \ (R_M \text{ is } A_{i,M}) \\ \text{Then } y_i & = p_{i,0} + p_{i,1}R_1 + \dots + p_{i,M}R_M \end{aligned} \tag{1}$$

- Each local model is associated with a degree of activation β_i expressed as $\beta_i = \min(u_{A_{i,1}}(R_1), u_{A_{i,2}}(R_2), \dots, u_{A_{i,M}}(R_M))$, where $u_{A_{i,j}}(R_j)$ is the membership function associated to the fuzzy set $A_{i,j}$. The global Y concentration is, then, computed by aggregating the conclusions inferred from r individual rules as followed:

$$Y = \frac{\sum_{i=1}^r \beta_i \cdot y_i}{\sum_{i=1}^r \beta_i} \tag{2}$$

- Optimization of fuzzy coefficients: the obtained fuzzy coefficients are applied to genetic algorithm. For our application, the chromosome codifies the membership function (a, b, c) of the fuzzy sets coefficients [1]. We defined the fitness value as the inverse of the Mean Square Error MSE. Also, we apply the arithmetic crossover and the uniform mutation operators to generate a new population [9]. Chromosomes to be mated are chosen by using the well-known roulette wheel selection method, which associates to each chromosome a probability proportional to its fitness value. We fixed the probability of crossover and mutation to 0.9 and 0.1, respectively. When the average of the fitness values of all the individuals in the population is greater than 99% of the fitness value of the best individual or a prefixed number of iterations has been executed, the GA is considered to have converged.
- Generalization: After optimization, we used the optimum coefficients found and applied them to the total images to generate water maps related to each water parameter [8].

The estimation process of Fig. 1 was implemented in Interactive Data Language (IDL) and applied to Algiers bay Landsat7 ETM+ image acquired on 03 june 2001 to generate water maps related to each water parameter.

Table 1 presents an example of statistics results of SPM estimation before and after optimization. We estimated the SPM concentrations and evaluated the performances of the fuzzy model by calculating the mean square error MSE and the correlation coefficient ρ . We noticed that fuzzy estimation gives satisfactory results. Furthermore, the optimization process improves these results, the correlation coefficient is increased from 91% to 98%.

Table 1. Statistic comparison of SPM estimation process.

	Fuzzy model		Genetic optimization	
	<i>MSE</i>	ρ	<i>MSE</i>	ρ
Training data	0.0081039	0.9142	0.0015308	0.9842
Test data	0.0084515	0.9093	0.0021842	0.9762

After optimization, the fuzzy coefficients were applied to full ETM image. The resulting maps represent the spatial variability of water constituents (SPM, Chl, Turb and SDD) in coastal and marine surfaces. To examine this variability, we present in Fig. 4 the maps in pseudo-colours.

For all maps, it is obvious notice that the high parameters values are located near the coast and this phenomenon is more important for the Algiers bay. For example, the concentration of the SPM parameter is closed to 1300 mg/l near the coast. In fact, the data partition revealed nine clusters (C_i) which correspond to 9 nine rules (Eq. 1). This means that this site has a high risk of pollution and presents a complex system requiring a high number of rules to interpret this wide variability. This information motivates us to exploit the results maps and propose a new tool to characterize the water quality, to localize the sources of pollution and identify the most polluted sites of Algiers bay.

3 PSD Analysing

The pollution signature draw exploits the obtained maps of Fig. 4 and provides a tool to evaluate the water quality at any point of the bay. This signature gathers in the same graph the marine indices in the following order: SPM, Chl, Turb and SDD and presents the normalized concentration of each index.

Figure 5(a) shows an example of PSD taken at different distances of Oued Elhar-rach. In the same graph, we reported a PSD taken at the large of the bay and considered as a reference PSD (purple color). Compared to this latter, we note high concentrations of SPM, Chl, Turb recorded just at the embouchure of the oued. The discharges poured at the oued are transported at several meters, and even kilometres from their source as it is illustrated by the PSDs taken at 300 m, 900 m, 1500 m and 3000 m from the oued.

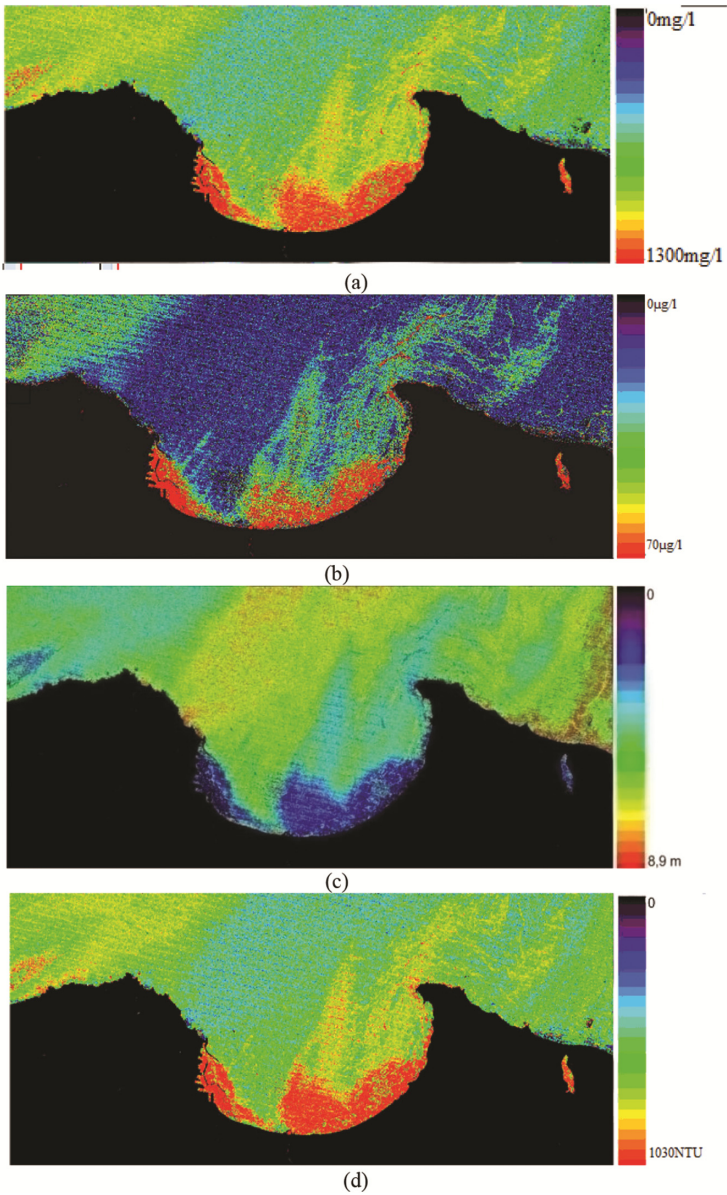


Fig. 4. Coastal water indices maps for Algiers bay. (a) SPM map, (b) Chl map, (c) SDD map and (d) Turb map.

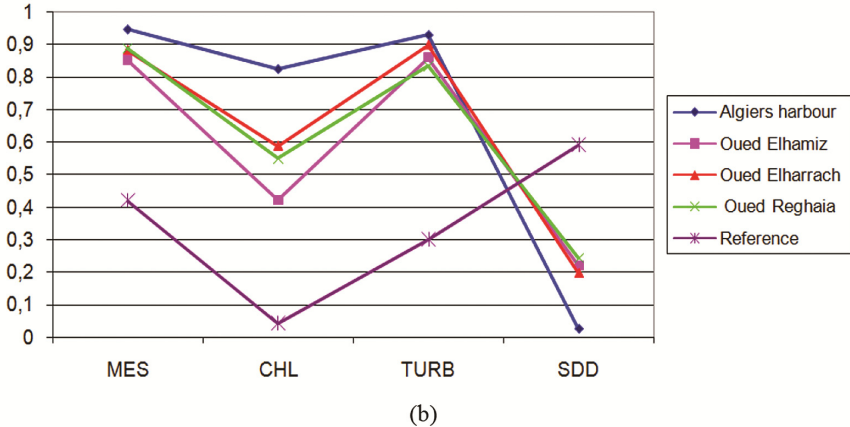
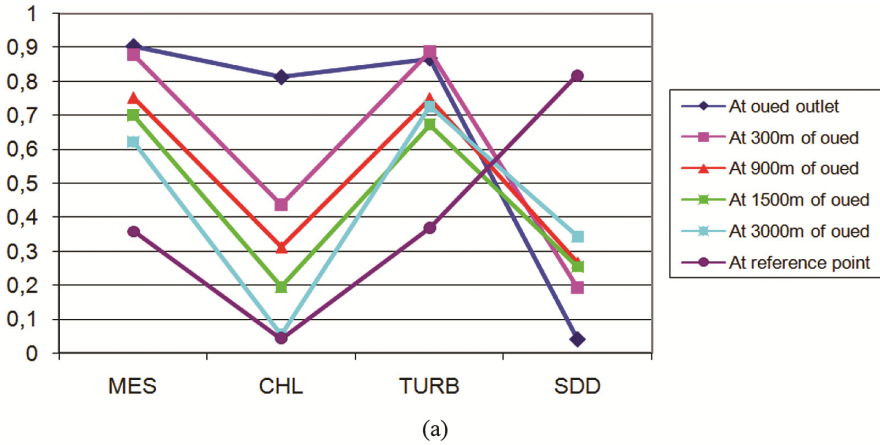


Fig. 5. Example of pollution signature. (a) PSDs taken at different distance from oued Elharrach, (b) PSDs taken at different sites.

Furthermore, the comparison between the PSDs of different sites provides an appreciation of water quality by its degree of pollution and hence, allows the detection of the risky sites. Figure 5(b) shows the average PSDs taken at risky sites (Algiers harbour, Oued Elharrach, oued Elhamiz and oued Reghaia). We also reported in the same graph a reference PSD taken at the large of the bay. It can be seen that Algiers harbour has the highest signature.

4 Pollution Map

Water classification based on the nature of waters components was proposed by Morel and Prieur [10]. Waters of the Case-I are those for which phytoplankton has a role in determining the water optical properties. Whereas the Case-II are determined or strongly influenced by the particulate and dissolved organic matter including the

main water constituents as SPM, Chl and DOM. This classification is widely used until nowadays. In this context and from forementioned findings, we typify Algiers Bay coastal waters on the basis of the PSD characterization.

4.1 Selection of Risky Sites

According to the PAC's reports, Algiers's bay gathered about 70% of industrial and urban dismissals of PAC's basins [11]. It is considered as point of meeting of all rejects of rivers (Oued el-harrach, Oued El-Hamiz and Oued Reghaia). Around 110.4 Hm³/year of domestic wastewater are discharged in this bay. Add to this, 5.4 millions m³/an of industrial wastewater coming from industrial cities which surround the bay (Hussein Dey, Bordj ElKiffan, Elhamiz and Rouiba). It should be noted that the water purification rate is around 47% for PAC region and it mainly concern Réghaia and Rouiba basin's without missing the oil pollution due to Algiers's harbour. These remarks reveal a high risk on some sites which need the integration of remote sensing tool in order to follow as well temporal as spatial of their pollution degree.

4.2 PSD Analysis

The analysis of some PSD related to risky sites (Fig. 5(b)) allows the evaluation of the pollution degree of these sites. Indeed, the PSDs of oued Elharrach and Oued Reghaia are similar. These sites are characterized by the same type of discharges (industrial discharges and wastewater). Also, oued Elhamiz has a low PSD value compared to those of oued Elharrach and oued Reghaia, but it remains significant when compared to the reference PSD. Thus, Algiers harbour is assigned to case-I (Strong Pollution), Oued Elharrach and Oued Reghaia are affected to the case-II (Average Pollution) while Oued Elhamiz is considered as case-III (Slight Pollution).

From risky sites relevant to different cases of water quality (Strong pollution, Average Pollution and Slight Pollution), we have selected some ROIs (region Of Interest) and extract their corresponding value in SPM, Chl, SDD and Turb images (see Fig. 6(a)). Also, we selected an ROI in clear water and urban area to construct a training data of five classes. Using Maximum Likelihood [12], the classification result is shown in Fig. 6. The pollution map highlights the most affected sites which present a high pollution risk. To evaluate our results and with a lake of ground realty, we compare our map to PAC (Plan d'Action Cotier) reports established on Algiers bay [11]. The obtained map highlights the risky sites which are cited in PAC reports.

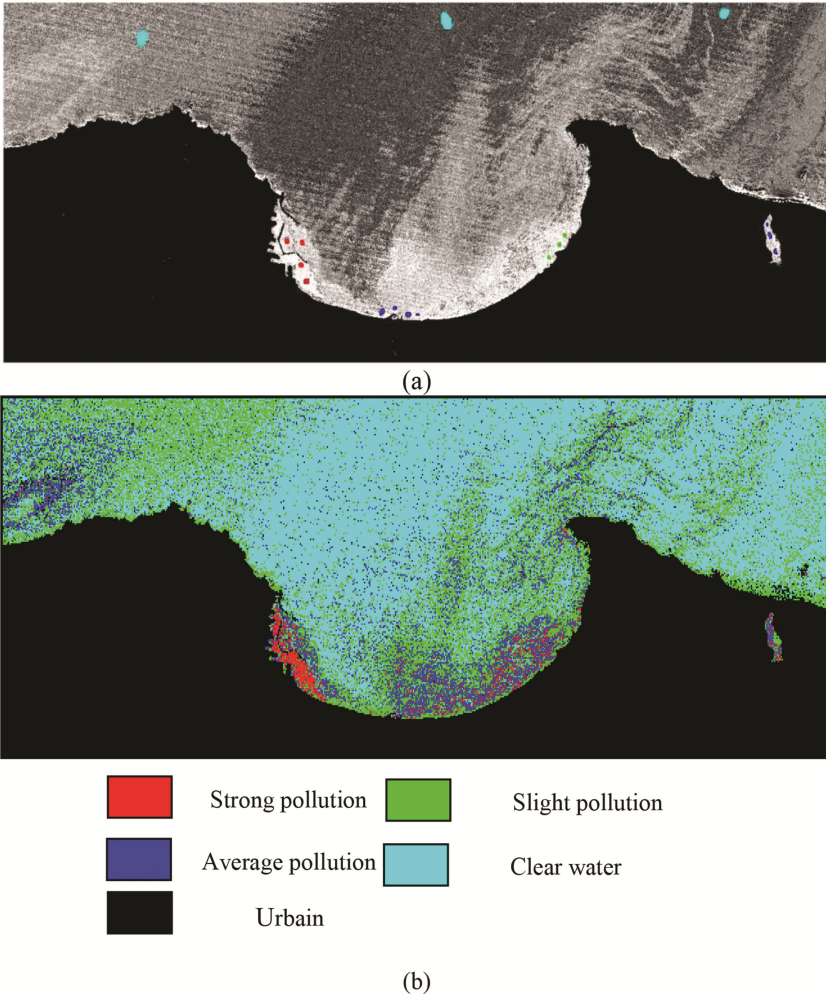


Fig. 6. Algiers bay water classification. (a) ROIs selection, (b) Pollution map

5 Conclusion

In this work, we proposed a coastal water quality using the combination of multispectral imagery and in-situ measurements. The first part is focused on the estimation and mapping of coastal water indices using an hybridization of fuzzy modeling and genetic optimization. The generalization of this hybridization to full images provides a map representing the spatial variability of water indices in the coastal and marine areas. Second and from these maps, Pollution Signature Draw “PSD” is constructed to characterize the water quality and to determine the boundaries of coastal zones with different water types. Finally, the comparison of some PSDs taken at risky sites of Algiers bay

highlighted the pollution degree of each site and allowed a water quality classification. The latter was coherent with the ground truth of the studied site. Moreover, and using multi-temporal remote sensing images, a PSDs analysis can give an important information for coastal water quality monitoring in order to determine seasonal and yearly changes. These objectives are under consideration for future works.

References

1. Lounis, B., Belhadj Aissa, A., Rabia, S., Ramoul, A.: Hybridisation of fuzzy systems and genetic algorithms for water quality characterisation using remote sensing data. *Int. J. Image Data Fus.* **4**(2), 171–196 (2013). doi:[10.1080/19479832.2011.617318](https://doi.org/10.1080/19479832.2011.617318). ISSN 1947-9832 (Print), 1947-9824 (Online)
2. Sathyendranath, S.: Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex Waters, IOCCG Report Number 3, Edited by Bedford Institute of Oceanography, Canada (2000)
3. Froidefond, J.M., Doxaran, D.: Télédétection optique appliquée à l'étude des eaux côtières. *Télédétection* **4**(2), 579–597 (2004)
4. Acker, J., Ouillon, S., Gould, R., Arnone, R.: Measuring marine suspended sediment concentrations from space: history and potential. In: 8th International Conference on Remote Sensing for Marine and Coastal Environments, Halifax, Canada, 17–19 May 2005
5. Zhang, Y., Pulliainen, J., Koponen, S., Hallikainen, M.: Application of an empirical neural network to surface water quality estimation in the Gulf of Finland using combined optical data and microwave data. *Remote Sens. Environ.* **81**, 327–336 (2002)
6. Cococcioni, M., Corsini, G., Lazzarini, B., Marcelloni, F.: Approaching the ocean color problem using fuzzy rules. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **34**(3), 1360–1373 (2004)
7. Takagi, T., Sugeno, M.: Fuzzy identification of systems and its application to modeling and control. *IEEE Trans. Syst. Man Cybern.* **15**, 116–132 (1985)
8. Lounis, B., Rabia, S., Ramoul, A., Belhadj Aissa, A.: An adaptive method using genetic fuzzy system to estimate suspended particulates matters SPM from Landsat and MODIS data. In: Fourth International Conference on Image and Signal Processing, ICISP 2010, Trois-Rivières, Québec, Canada (2010). ISBN 364213680-X, 9783642136801
9. Man, K.F., Tang, K., Kwong, S., Halang, S.: Genetic for Control and Signal Processing. Springer, London (1997)
10. Morel, A., Prieur, L.: Analysis of variations in ocean color. *Limnol. Oceanogr.* **22**(4), 709–722 (1977)
11. Mohamed, L.: Analyse de durabilité dans le cadre du PAC: Zone côtière algéroise (Algérie). Rapport du 4ème atelier, Alger, pp. 16–17, mai 2004. www.planbleu.org/publications/pac_alger_atelier4.pdf
12. Cocquerez, J.P., Philipp, S.: Analyse d'images: filtrage et segmentation, Edition Masson (1997)