# THE ARTIFICIAL BEE COLONY ALGORITHM IN TRAINING ARTIFICIAL NEURAL NETWORK FOR OIL SPILL DETECTION

*Coskun Ozkan*,[*] *Celal Ozturk*,[†] *Filiz Sunar*,[‡] *Dervis Karaboga*[†]

**Abstract:** Nowadays, remote sensing technology is being used as an essential tool for monitoring and detecting oil spills to take precautions and to prevent the damages to the marine environment. As an important branch of remote sensing, satellite based synthetic aperture radar imagery (SAR) is the most effective way to accomplish these tasks. Since a marine surface with oil spill seems as a dark object because of much lower backscattered energy, the main problem is to recognize and differentiate the dark objects of oil spills from others to be formed by oceanographic and atmospheric conditions. In this study, Radarsat-1 images covering Lebanese coasts were employed for oil spill detection. For this purpose, a powerful classifier, Artificial Neural Network Multilayer Perceptron (ANN MLP) was used. As the original contribution of the paper, the network was trained by a novel heuristic optimization algorithm known as Artificial Bee Colony (ABC) method besides the conventional Backpropagation (BP) and Levenberg-Marquardt (LM) learning algorithms. A comparison and evaluation of different network training algorithms regarding reliability of detection and robustness show that for this problem best result is achieved with the Artificial Bee Colony algorithm (ABC).

Key words: *Oil spill, Artificial Neural Network (ANN), Artificial Bee Colony (ABC)*

*Received: November 22, 2010*
*Revised and accepted: October 4, 2011*

## 1. Introduction

Petroleum is a major energy source for modern life. The industrialized countries usually obtain most of the required petroleum by tanker ships from overseas sources (Sabins 1997). This crowded marine transportation has increased marine pollution

[*]Coskun Ozkan
Geomatics Engineering, Erciyes University, Kayseri, Turkiye, cozkan@erciyes.edu.tr
[†]Celal Ozturk, Dervis Karaboga
Computer Engineering, Erciyes University, Kayseri, Turkiye, celal@erciyes.edu.tr, karaboga@erciyes.edu.tr
[‡]Filiz Sunar
Geomatics Engineering, Istanbul Technical University, Turkiye, fsunar@itu.edu.tr

because of oil spill discharges. It has been estimated that the annual amount of spilled oil worldwide is more than 4.5 million tons (Bava *et al.* 2002). An oil spill, accidental petroleum release into the environment, is usually localized if it originates from land-based sources and thus its impact can be eliminated relatively easily. In contrast, marine oil spills, mainly associated with oil transportation by tankers and pipelines, may result in oil pollution over large areas and present serious biological and economic impacts. Oil spill discharges generally occur during tanker and cargo ship operations, where ballast waters, tank washing residues, fuel oil sludge and machinery space bilge deliberately are discharged into the sea (Harahsheh *et al.* 2004). Even small discharges can affect marine wildlife. Mostly birds and mammals are affected by sticky crude oil and bunker fuels that float on the sea surface. The transformation of these oil types by microorganisms and sea hydrodynamics takes a long time. Although refined petroleum products are not sticky and floating so long as crude oil and bunker fuels, they are more poisonous for marine environment (AMSA – Australian Maritime Safety Authority). Since it is becoming increasingly important to detect and monitor oil spills at sea for marine ecosystem, today remote sensing technology becomes the most efficient technology (Sabins 1997). Synthetic Aperture Radar (SAR) remote sensing is superior to optical sensors due to the all weather and all day operation capabilities (Solberg and Theophilopoulos 1997). The backscatter energy level for oil-spilled areas is too low since oil dampens the capillary waves of the sea surface, which causes dark areas. However, Synthetic Aperture Radar (SAR) images must be processed carefully since the dark areas might occur because of some natural phenomena without oil like smooth water (low wind areas), organic films, wind front areas, areas sheltered by land, rain cells, grease ice, internal waves and shallow bathymetric features (Sabins 1997, Alpers *et al.* 1991, Hovland *et al.* 1994). The procedure steps of oil spill detection in Synthetic Aperture Radar (SAR) data can be generalized as segmentation (dark object extraction), feature extraction and classification (determination oil) stages (Pavlakis *et al.* 2001, Brekke and Solberg 2005a, Brekke and Solberg 2005b, Solberg *et al.* 2007, Shi *et al.* 2008, Topouzelis *et al.* 2009). A detailed survey for these steps and imaging systems can be found in Topouzelis (2008).

In this study, the dark image areas, which were either oil spills or look-alikes, were segmented (Xiaoying 2009). The features like shape, contrast and textural characteristics were extracted from the segmented parts (Topouzelis *et al.* 2009). By using the features, statistical, artificial intelligence, rule-based or boosting algorithms were used for identification of the dark areas in a manner of binary classification, i.e. oil or look-alike (Fiscella *et al.* 2000, Del Frate *et al.* 2000, Solberg *et al.* 1999, Keramitsoglou *et al.* 2005, Topouzelis *et al.* 2009). In this study, a well-known Multilayer Perceptron (MLP) structure of Artificial Neural Networks (ANN) was used. Since Multilayer Perceptron (MLP) is a very powerful and flexible classifier, many different remote sensing studies were reported in literature (Paola and Schowengerdt 1995, Sunar and Özkan 2001, Kavzoglu and Mather 2003, Foody 2001) and a detailed review of remote sensing applications by Artificial Neural Network (ANN) is given in Mas and Flores (2008). The Backpropagation (BP) algorithm (Rumelhart *et al.* 1986) is the most well known training method to optimize weight and bias parameters of Multilayer Perceptron Artificial Neural

Network (MLP ANN). While the Backpropagation (BP) algorithm converges in the first order derivatives, the Levenberg-Marquardt (LM) algorithm (Hagan and Menhaj 1994) converges using the second order of derivatives. Therefore, in most of the training applications Levenberg-Marquardt (LM) is faster than the Backpropagation (BP) (Topouzelis *et al.* 2009) and outperforms the Backpropagation (BP) algorithm (Kermani *et al.* 2005).

This study proposes a Multilayer Perceptron Artificial Neural Network (MLP ANN) training model using Artificial Bee Colony (ABC) algorithm which is described by Karaboga (2005). The algorithm, based on the foraging behavior of honeybees, was first tested on the numerical optimization problems (Karaboga and Basturk 2007a). The Artificial Bee Colony (ABC) algorithm was used to train feed-forward neural networks on classification of machine learning community benchmark test problems (Karaboga and Ozturk 2009) and inertial sensor based terrain classification (Kurban and Bedok 2009) where algorithm showed superior performance against other well-known gradient-based and population-based optimization techniques. As the original contribution of the paper, it is aimed to train Multilayer Perceptron Artificial Neural Network (MLP ANN) by Artificial Bee Colony (ABC) algorithm for oil spill classification purpose while comparing the performance of this new generation algorithm by the Backpropagation (BP) and the Levenberg-Marquardt (LM) conventional techniques.

## 2. Artificial Bee Colony Algorithm (ABC)

Artificial Bee Colony (ABC) algorithm is a new meta-heuristic population based swarm intelligence algorithm developed by Karaboga (2005). The ABC algorithm mimics the intelligent foraging behavior of honeybee swarms. The first researches about ABC algorithm focused on examining the effectiveness of ABC for constrained and unconstrained problems against other well-known modern heuristic algorithms such as Genetic Algorithm (GA), Differential Evolution (DE), and Particle Swarm Optimization (PSO) (Karaboga and Basturk 2007b, Karaboga and Akay 2009). Later on, ABC has been used for ANN classifier training and clustering problem (Karaboga and Ozturk 2009, Karaboga and Ozturk 2011) where some benchmark classification problems were tested, and the results were compared with those of other widely-used techniques.

The ABC algorithm consists in a set of possible solutions $x_i$ (the population) represented by the positions of food sources where the nectar amount of a food source corresponds to the quality (fitness) of the associated solution. The basic idea of the ABC algorithm is to assign artificial bees to investigate the search space searching the feasible solutions. The artificial bees collaborate and exchange information so that bees concentrate on more promising solutions in terms of certain evaluation criteria. A set of artificial bees is used to collaboratively search for the optimal solution. The ABC algorithm basically uses three types of bees in the colony: (i) employed bees, (ii) onlooker bees and (iii) scout bees. While employed bees are chosen from half of the colony at the beginning, the other half is assigned as onlookers. Employed bees go to the food sources, and then inform onlooker bees about the nectar and the positional information of the food sources, meanwhile onlooker bees wait on the dance area to determine to choose a food source. An employed bee is transformed into a scout bee if its food source is abandoned. As a simulation element, the scout bees carry out random search in the simulation model space.

In the ABC algorithm, the number of employed bees equals the number of food sources plus the number of onlooker bees. There is only one employed bee for each food source where the initial positions of employed bees are randomly generated. At each of following iterations, each of employed bees searches a new food source neighboring to its current food source by using Equation (1), and then computes the amount of nectar of the new food source. On the basis of greedy selection, if the amount of nectar of the new food source is higher than that of its current food source, then the employed bee continues with that food source, otherwise it keeps the current one.

$$v_{ij} = x_{ij} + \theta_{ij}(x_{ij} - x_{kj}), \tag{1}$$

where $v_i$ is a candidate solution, $x_i$ is the current solution, $x_k$ is a neighbor solution and $\theta$ is a random number between [-1,1] that controls the production of neighbor food sources around $x_{ij}$.

After this search process is completed for all employed bees, they share the information about their food sources with the onlooker bees. An onlooker bee analyzes the nectar information and selects a food source in terms of a probability related to the nectar amount of the sources, computed by Equation (2). These empirical probabilities enable a roulette wheel selection which produces better solution candidates to have a greater chance of being selected.

$$p_i = \frac{fit_i}{\sum\limits_{n=1}^{SN} fit_n}, \tag{2}$$

where *fit*$_i$ is the fitness value of solution $i$, which is proportional to the nectar amount of the food source at position $i$, and *SN* is the number of food sources, which is equal to the number of employed bees.

The food source to be assigned to an onlooker bee is controlled by a random number which is between 0 and 1 on the basis of a comparison of the random number and the probability value. Once all onlookers have been assigned to food sources, each of them searches within a new neighboring food source by Equation (1) of its assigned food source and computes its nectar amount. If the amount of nectar of the new source is higher than that of the assigned one, then the onlooker bee memorizes the new position and forgets the old one.

If a food source cannot be improved enough through a predetermined number of cycles by the related employed and onlooker bees, it is assigned as an abandoned source and then the employed bee becomes a scout bee when the cycle number exceeds the critical number of cycles, called limit parameter. The scout bee generates a new random solution by Equation (3). Assume that $x_i$ is the abandoned source and $j \in \{1,2,...,D\}$, where $D$ is the dimensionality of the solution vector, the scout discovers a new food source which will be replaced with $x_i$:

$$x_i^j = x_{min}^j + rand(0,1)(x_{max}^j - x_{min}^j), \tag{3}$$

where $j$ is determined randomly, to be different from $i$.

The steps of the ABC algorithm are as follows (Karaboga and Ozturk 2009):

- Generate initial population $x_i$, *i=1...SN*
- Evaluate the population
- Set *cycle to 1*
- REPEAT

    FOR each employed bee

        Produce new solutions $v_i$ by using Equation (1)
        Calculate fitness
        Apply the greedy selection process

    FOR each onlooker bee

        Choose a solution $x_i$ depending on $p_i$ calculated by Equation (2)
        Produce new solutions $v_i$ by using Equation (1)
        Calculate fitness
        Apply the greedy selection process

- If there is an abandoned solution for the scout, then

        Replace it with a new solution produced by Equation (3)

- Memorize the best solution achieved so far
- Assign cycle to *cycle + 1*
- UNTIL maximum cycle number is reached

There are three control parameters in the ABC algorithm: (i) the number of food sources which is equal to the number of employed and also onlooker bees (*SN*), (ii) the limit parameter, and (iii) the maximum cycle number.

## 3. Study Area and Feature Extraction

Radarsat-1 images covering the oil pollution occurred on the Lebanese coast in July 2007 were used as a dataset. Radarsat-1 images were acquired by ITU-CSCRS (Istanbul Technical University – Centre for Satellite Communication and Remote Sensing) during the event. Depending on the wind and current condition, the oil pollution spread 100 miles along the coast and affected also Syria's shoreline. In the East Mediterranean region, Lebanon, neighboring Syria and Israel, has 225 km coastline (Fig. 1). This oil pollution occurred due to the bombing of a power plant at Jiyeh, 12 miles south of Beirut, and affected approximately 1/3 of the whole Lebanese coastline, nearly 70-80 km north of the power plant which is illustrated with blue points on Fig. 1. Depending on the weather conditions, it could have been also a serious threat to the neighboring Mediterranean countries such as Turkey, Cyprus and Syria.

Features can be branched into three categories (Solberg and Theophilopoulos 1997, Del Frate *et al.* 2000, Karathanassi *et al.* 2006, Brekke and Solberg 2005a) as given below:

- The geometric characteristics of oil spills such as area, perimeter, complexity.
- The physical behavior of oil spills such as mean, standard deviation or max backscatter value.

- The oil spill context in the image such as number of other dark formations, presence of ships and proximity to and route of ships.

In addition to these features, textural information may be seen as another category, etc. Haralick features (Haralick 1979). Texture can be defined as a variation of the pixel intensities in a specific image area and some features based on texture measures may increase the classification performance (Törma *et al.* 2004). Haralick features describing the texture features can be computed using Haralick's co-occurrence matrix based on image intensities (Haralick *et al.* 1973). A co-occurrence matrix is a two dimensional histogram of intensity values for a pair of image pixels which are separated by a fixed spatial relationship. The texture measures based on co-occurrence matrix can be Angular Second Moment, Contrast, Correlation, Dissimilarity, Entropy, Homogeneity, Mean and Standard Deviation (Assilzadeh and Mansor 2001).

The features based on the above characteristics represent objects instead of pixels. Especially geometric, physical and contextual characteristics are computed, based on the objects (dark formations). Therefore, before the features are computed, the dark objects must be extracted. It is obvious that an object-oriented approach is more convenient for SAR imagery classification because of both the coherent nature imagery of SAR and the closeness of the dark area intensity values. The detailed explanations of the sub-features of these general categories can be found in Topouzelis *et al.* (2009), which examines different combinations of a total of 25 sub-features on a basis of maximizing the oil spill detection performance by

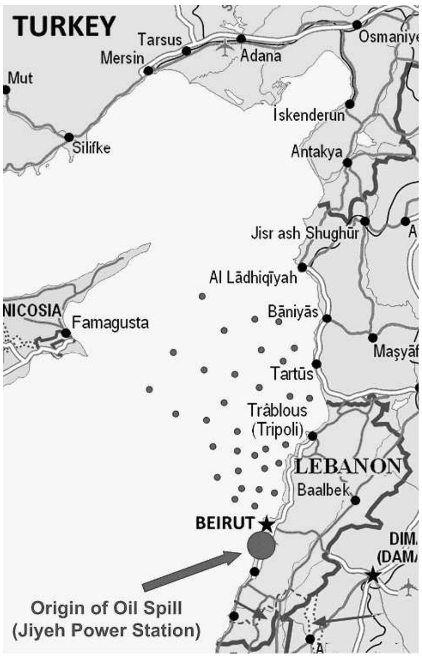


**Fig. 1** *Study area: Big point represents the origin of oil spill and small points represent the dispersion of oil spill.*

using a Genetic Search Algorithm (GA). They reported that 10 of the features as the optimal group of features represented the full classification accuracy according to their proposed method. Although many combinations of features can be selected, the following features are preferred to be used as the feature set according to the characteristics and our analysis. In fact, finding an optimal set of features is not the main motivation of the paper. Moreover, the feature set might be preferred not to obtain full accuracy in order to be able to test various optimization algorithms. Consequently, the selected features are SP2, BSd, ConLa, Opm, Opm/Bpm, OSd, P/A, C and THm (Topouzelis *et al.* 2009):

- SP2 means Shape factor II (SP2), which describes the general shape of the object. It is also called as 'first invariant planar moment', 'form factor', and 'asymmetry'.
- BSd means Background standard deviation, which is the standard deviation of the intensity values of the pixels belonging to the region of interest, selected by the user surrounding the object.
- ConLa means Local area contrast ratio, which is the ratio between the mean backscatter value of the object and the mean backscatter value of a window centered at the region.
- Opm means object power to mean ratio, which is the ratio between standard deviation of the object and the mean of the object.
- Opm/Bpm means of the power to mean ratio, which is the ratio between the object power to mean ratio and the background power to mean ratio.
- OSd means object standard deviation, which is the standard deviation of the object.
- P/A means perimeter to area ratio, which is the ratio between the perimeter (P) and the area (A) of the object.
- C means object complexity, which describes how simple (or complex) the geometrical objects are.
- THm means mean Haralick texture, which is the mean Haralick texture based on the average of the grey level co-occurrence matrices of the sub-objects.

The features can be categorized as: SP2, P/A, and C are geometrical characteristics; BSd, ConLa, Opm, Opm/Bpm, and OSd refer to physical characteristics, and THm is textural characteristic of dark formations (Topouzelis *et al.* 2009).

In order to extract the above features on the basis of low computational burden, the dark areas are firstly windowed in proper extents. Fig. 2 shows windowed dark areas of determined oil spill and look-alike.

The objects are the dark formations in the windows where a total of 68 oil and 53 look-alike objects detected by CSCRS is windowed. The objects were segmented on the basis of binary image. Binary image enables to extract the pixel coordinates of object elements. Using binary images as a mask, the preferred 9 features are

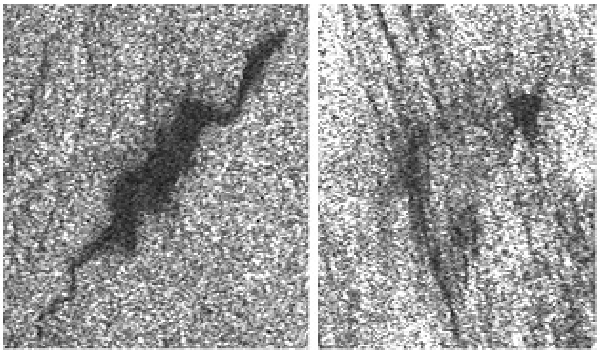

**Fig. 2** *Oil (left) and look-alike (right) windows.*

computed from the objects and the backgrounds for each window. The statistics of oil and look-alike data are given in Tab. I.

The features are randomly divided into train and test data for both oil and look-alike objects: 35 train and 30 test objects for oil data and 30 train and 23 test objects for look-alike data, etc. all in all 65 objects as a train set and 56 objects as a test set.

| Features | Oil | | | | Look-Alike | | | |
|---|---|---|---|---|---|---|---|---|
| | Min. | Max. | Mean | Std. | Min. | Max. | Mean | Std. |
| **SP2** | 0.053 | 1 | 0.547 | 0.202 | 0.269 | 1 | 0.666 | 0.173 |
| **BSd** | 13.635 | 68.04 | 40.453 | 14.677 | 23.475 | 68.066 | 41.780 | 9.551 |
| **ConLa** | 0.373 | 0.85 | 0.547 | 0.088 | 0.440 | 0.663 | 0.570 | 0.062 |
| **Opm** | 0.177 | 0.457 | 0.257 | 0.066 | 0.169 | 0.297 | 0.207 | 0.025 |
| **Opm/Bpm** | 0.454 | 5.424 | 1.071 | 0.685 | 0.433 | 1.410 | 0.808 | 0.181 |
| **OSd** | 7.009 | 53.165 | 20.275 | 9.309 | 8.972 | 29.838 | 16.791 | 4.349 |
| **P/A** | 0.095 | 0.881 | 0.486 | 0.221 | 0.173 | 0.673 | 0.448 | 0.110 |
| **C** | 2.020 | 51.252 | 13.250 | 9.873 | 9.213 | 74.762 | 23.898 | 11.588 |
| **THm** | 11.400 | 36.202 | 22.346 | 5.890 | 12.400 | 35.577 | 21.437 | 5.312 |

**Tab. I** *Statistics of oil and look-alike data.*

# 4. Network Model

The design of a topologic structure (number of hidden layers and number of neurons within the layers) of a Multilayer Perceptron Artificial Neural Network (MLP ANN) is important (Dayhof 1990, Haykin 1999). Generally, one hidden layer is capable of learning (generalizing) the relations between input and output well enough (Lawrence 1993, Bishop 1995). The number of neurons in a hidden layer is so flexible that any number of neurons (not less than the output dimension) may be easily employed with enough training iterations depending on the input-output data. If the topology is not enough, poor results are obtained because of lack of generalization. On the contrary, if the topology is overdetermined, then generalization transforms into memorizing; this is also known as overfitting. So, these topologic parameters are usually determined by trial and error (Kavzoglu and Mather 2003). After many tests, our topology was selected as 9-6-2, one hidden layer with 6 neurons. Input and output layers were fixed by the dimension of input and output patterns. Two classes (oil and look-alike) were represented by 2 neurons in the output layer and nine different features were represented by 9 neurons in the input layer. The logarithmic sigmoid transfer function was employed at hidden and output layer neurons. The topology covered 74 total unknown parameters of weights and biases to be optimized (9*6+6+6*2+2).

All the parameters of BP, LM and ABC were determined by trial and error from many tests. The learning rate parameter for BP was 0.4. For LM applications the blending factor was 0.01 and its increase and decrease values were set to 10 and 0.1, respectively. Colony size, maximum cycle number and limit value of the ABC algorithm were set 20, 1000 and 500, respectively. The working interval of parameters (ANN weights and biases) was chosen as [-10 10]. The input data were normalized (i.e. [-1 , 1] closed interval) to increase the classification performance and to decrease the convergence time before the data given to the network.

# 5. Methodology

Since the optimal values of the weight and bias parameters which store information are unknown, the training phase (updating of weights and biases) is naturally an iterative process. The parameters are randomly initialized at the first iteration. Then, each run of training phases with a sufficient number of iterations provides a different solution space. The general strategy in training phase is using validation data to prevent overfitting and then testing the optimized parameters (Kavzoglu and Mather 2003, Weigend 1994). Furthermore, assuming to have a reasonable number of iterations on the basis of generalization, it is believed that not only generalization but also robustness is important. From the viewpoint of the study, robustness can be defined as the performance precision of multiple runs, having a small dispersion of performances. When the performance or error values are clustered in a narrow interval, it means that the algorithm is resistant to different initial conditions and the algorithm is stable. Since the artificial intelligence techniques are mostly heuristic algorithms, i.e. a solution can always be found but there is not a guarantee to be optimal, robustness is one of the main criteria to be checked especially in the comparison of such type algorithms. Therefore, BP, LM, and ABC algorithms are compared to each other in terms of descriptive statistics obtained from 30 independent runs. Afterwards, these groups of 30 elements are statistically tested on the basis of the inferential statistical approaches. In statistical inference, 30 is assumed as a threshold for reasonable statistical analysis (Johnson and Bhattacharyya 2000).

The value of 1000 was determined as an iteration number for all runs of BP, LM, and ABC algorithms. In spite of the fact that the number of iterations of ABC runs might have been assigned a higher value because of its heuristic nature, ABC had a moderate convergence speed so that higher iterations were not needed in the problem.

In the experiments, the test data were simulated by the network models after the training phase to obtain the classification results which were used to compute the overall, procedure and user accuracies (Congalton and Green 1999) and the mean, minimum, maximum, standard deviation, skewness and kurtosis statistics. Skewness is a form measure of shape of distribution. A nonzero skewness means the stack of the values lies to the left or right of the mean, etc. asymmetric tail formation. A zero value means nearly even distributed on both sides of the mean. Kurtosis illustrates the flatness of the distribution with respect to Normal Curve, so it is informative about the tail behavior of distribution. Nonzero values mean more of the variance is the result of infrequent or frequent extreme deviations.

The frequency distributions of overall accuracies were illustrated graphically in terms of bar graphics (kind of histogram) with fitting Normal distribution. Let us analyze visually the shapes of the graphics and the fitted Normal curves to find whether the variation is within a reasonable range or not, etc. the bigger standard deviation (a flattened Normal curve), the less robustness (precision). Finally, the mean overall accuracy performances obtained from 1000 iterations were plotted for each of BP, LM and ABC algorithms to examine the convergence speed.

# 6. Results

The results obtained from the simulations were used to compute overall, procedure and user accuracies of oil and look-alike test data. The classification accuracies are basically given in two types: the accuracies obtained from the 1000 iterations for each run after 1000 training iterations and accuracies obtained from the optimal iterations for each run. The optimal iteration of a run is determined according to the highest overall accuracy in that run. The optimal number of iterations is computed since a better accuracy can be obtained by the parameter values which are determined in the earlier iterations of the training process than the accuracy of the last iteration. In order to be able to evaluate the performances, the weight and bias parameters of each of iteration were saved. For example, in LM training, a total of 1000 iterations give a 74x1000 parameter matrix. After running LM 30 times independently, the matrix becomes a 74x1000x30 matrix. By loading the stored parameters into the network model, accuracies for oil and look-alike were computed for each iteration at each run. Consequently, five (one overall, two procedure and two user accuracies) 1000x30 sized performance matrices were produced. Then, the optimal iteration numbers which reflect the highest overall accuracies in each of the 30 runs were determined. According to the optimal iterations, the classification performances were extracted from the performance matrices. The results are given in Tabs. II–IV. In these tables, OA, PA and UA abbreviations stand for overall accuracy, procedure accuracy and user accuracy, respectively. The descriptive statistics of 1000 and optimal iterations are given in Tabs. V–VI. At a glance, the results clearly point out the vantage of ABC over BP and LM.

The two-sided tests for skewness and kurtosis illustrate the normality of overall accuracies of algorithms from 30 runs, i.e. groups. The null hypothesis ($H_0$) claims that the skewness and kurtosis values are zero. The alternative hypothesis ($H_1$) generally claims that skewness and kurtosis are not equal to zero. For significance

| # of Run | Iteration | OA | PA | | UA | | Iteration | OA | PA | | UA | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Oil | Alike | Oil | Alike | | | Oil | Alike | Oil | Alike |
| **1** | 1000 | 69.64 | 69.70 | 69.57 | 76.67 | 61.54 | 943 | 69.64 | 69.70 | 69.57 | 76.67 | 61.54 |
| **2** | 1000 | 80.36 | 87.88 | 69.57 | 80.56 | 80.00 | 927 | 80.36 | 87.88 | 69.57 | 80.56 | 80.00 |
| **3** | 1000 | 78.57 | 81.82 | 73.91 | 81.82 | 73.91 | 790 | 80.36 | 84.85 | 73.91 | 82.35 | 77.27 |
| **4** | 1000 | 80.36 | 84.85 | 73.91 | 82.35 | 77.27 | 488 | 82.14 | 87.88 | 73.91 | 82.86 | 80.95 |
| **5** | 1000 | 78.57 | 84.85 | 69.57 | 80.00 | 76.19 | 920 | 80.36 | 87.88 | 69.57 | 80.56 | 80.00 |
| **6** | 1000 | 80.36 | 78.79 | 82.61 | 86.67 | 73.08 | 498 | 82.14 | 78.79 | 86.96 | 89.66 | 74.07 |
| **7** | 1000 | 78.57 | 84.85 | 69.57 | 80.00 | 76.19 | 499 | 80.36 | 81.82 | 78.26 | 84.38 | 75.00 |
| **8** | 1000 | 75.00 | 72.73 | 78.26 | 82.76 | 66.67 | 483 | 80.36 | 84.85 | 73.91 | 82.35 | 77.27 |
| **9** | 1000 | 82.14 | 84.85 | 78.26 | 84.85 | 78.26 | 546 | 82.14 | 81.82 | 82.61 | 87.10 | 76.00 |
| **10** | 1000 | 75.00 | 72.73 | 78.26 | 82.76 | 66.67 | 925 | 75.00 | 72.73 | 78.26 | 82.76 | 66.67 |
| **11** | 1000 | 83.93 | 93.94 | 69.57 | 81.58 | 88.89 | 763 | 87.50 | 96.97 | 73.91 | 84.21 | 94.44 |
| **12** | 1000 | 75.00 | 69.70 | 82.61 | 85.19 | 65.52 | 979 | 75.00 | 69.7 | 82.61 | 85.19 | 65.52 |
| **13** | 1000 | 58.93 | 100 | 0 | 58.93 | **NaN** | 86 | 58.93 | 100 | 0 | 58.93 | **NaN** |
| **14** | 1000 | 82.14 | 90.91 | 69.57 | 81.08 | 84.21 | 926 | 83.93 | 93.94 | 69.57 | 81.58 | 88.89 |
| **15** | 1000 | 55.36 | 63.64 | 43.48 | 61.76 | 45.45 | 9 | 60.71 | 100 | 4.35 | 60.00 | 100 |
| **16** | 1000 | 66.07 | 87.88 | 34.78 | 65.91 | 66.67 | 267 | 66.07 | 90.91 | 30.43 | 65.22 | 70.00 |
| **17** | 1000 | 78.57 | 81.82 | 73.91 | 81.82 | 73.91 | 325 | 82.14 | 84.85 | 78.26 | 84.85 | 78.26 |
| **18** | 1000 | 82.14 | 87.88 | 73.91 | 82.86 | 80.95 | 952 | 82.14 | 87.88 | 73.91 | 82.86 | 80.95 |
| **19** | 1000 | 78.57 | 78.79 | 78.26 | 83.87 | 72.00 | 685 | 80.36 | 78.79 | 82.61 | 86.67 | 73.08 |
| **20** | 1000 | 76.79 | 78.79 | 73.91 | 81.25 | 70.83 | 962 | 76.79 | 78.79 | 73.91 | 81.25 | 70.83 |
| **21** | 1000 | 78.57 | 75.76 | 82.61 | 86.21 | 70.37 | 672 | 78.57 | 78.79 | 78.26 | 83.87 | 72.00 |
| **22** | 1000 | 75.00 | 78.79 | 69.57 | 78.79 | 69.57 | 773 | 75.00 | 78.79 | 69.57 | 78.79 | 69.57 |
| **23** | 1000 | 83.93 | 84.85 | 82.61 | 87.50 | 79.17 | 820 | 85.71 | 87.88 | 82.61 | 87.88 | 82.61 |
| **24** | 1000 | 75.00 | 78.79 | 69.57 | 78.79 | 69.57 | 450 | 82.14 | 87.88 | 73.91 | 82.86 | 80.95 |
| **25** | 1000 | 69.64 | 63.64 | 78.26 | 80.77 | 60.00 | 534 | 69.64 | 54.55 | 91.30 | 90.00 | 58.33 |
| **26** | 1000 | 60.71 | 54.55 | 69.57 | 72.00 | 51.61 | 953 | 60.71 | 54.55 | 69.57 | 72.00 | 51.61 |
| **27** | 1000 | 78.57 | 78.79 | 78.26 | 83.87 | 72.00 | 676 | 80.36 | 81.82 | 78.26 | 84.38 | 75.00 |
| **28** | 1000 | 62.5 | 84.85 | 30.43 | 63.64 | 58.33 | 980 | 64.29 | 87.88 | 30.43 | 64.44 | 63.64 |
| **29** | 1000 | 78.57 | 78.79 | 78.26 | 83.87 | 72.00 | 936 | 78.57 | 78.79 | 78.26 | 83.87 | 72.00 |
| **30** | 1000 | 78.57 | 84.85 | 69.57 | 80.00 | 76.19 | 907 | 80.36 | 87.88 | 69.57 | 80.56 | 80.00 |

**Tab. II** *Classification accuracies of BP from 1000 (left) and optimal iterations (right) (%). (OA: Overall Accuracy, PA: Procedure Accuracy, UA: User Accuracy).*

| # of Run | Iteration | OA | PA | | UA | | Iteration | OA | PA | | UA | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Oil | Alike | Oil | Alike | | | Oil | Alike | Oil | Alike |
| **1** | 1000 | 78.57 | 78.79 | 78.26 | 83.87 | 72.00 | 15 | 82.14 | 84.85 | 78.26 | 84.85 | 78.26 |
| **2** | 1000 | 58.93 | 57.58 | 60.87 | 67.86 | 50.00 | 12 | 83.93 | 87.88 | 78.26 | 85.29 | 81.82 |
| **3** | 1000 | 78.57 | 84.85 | 69.57 | 80.00 | 76.19 | 5 | 83.93 | 87.88 | 78.26 | 85.29 | 81.82 |
| **4** | 1000 | 75.00 | 69.70 | 82.61 | 85.19 | 65.52 | 10 | 83.93 | 87.88 | 78.26 | 85.29 | 81.82 |
| **5** | 1000 | 82.14 | 84.85 | 78.26 | 84.85 | 78.26 | 19 | 85.71 | 87.88 | 82.61 | 87.88 | 82.61 |
| **6** | 1000 | 80.36 | 78.79 | 82.61 | 86.67 | 73.08 | 9 | 82.14 | 84.85 | 78.26 | 84.85 | 78.26 |
| **7** | 1000 | 80.36 | 78.79 | 82.61 | 86.67 | 73.08 | 31 | 82.14 | 81.82 | 82.61 | 87.10 | 76.00 |
| **8** | 1000 | 75.00 | 69.70 | 82.61 | 85.19 | 65.52 | 8 | 80.36 | 84.85 | 73.91 | 82.35 | 77.27 |
| **9** | 1000 | 75.00 | 72.73 | 78.26 | 82.76 | 66.67 | 6 | 82.14 | 84.85 | 78.26 | 84.85 | 78.26 |
| **10** | 1000 | 83.93 | 81.82 | 86.96 | 90.00 | 76.92 | 9 | 85.71 | 84.85 | 86.96 | 90.32 | 80.00 |
| **11** | 1000 | 76.79 | 75.76 | 78.26 | 83.33 | 69.23 | 4 | 83.93 | 81.82 | 86.96 | 90.00 | 76.92 |
| **12** | 1000 | 71.43 | 72.73 | 69.57 | 77.42 | 64.00 | 5 | 83.93 | 87.88 | 78.26 | 85.29 | 81.82 |
| **13** | 1000 | 67.86 | 63.64 | 73.91 | 77.78 | 58.62 | 6 | 85.71 | 90.91 | 78.26 | 85.71 | 85.71 |
| **14** | 1000 | 75.00 | 72.73 | 78.26 | 82.76 | 66.67 | 4 | 80.36 | 96.97 | 56.52 | 76.19 | 92.86 |
| **15** | 1000 | 73.21 | 72.73 | 73.91 | 80.00 | 65.38 | 6 | 83.93 | 81.82 | 86.96 | 90.00 | 76.92 |
| **16** | 1000 | 67.86 | 63.64 | 73.91 | 77.78 | 58.62 | 6 | 83.93 | 84.85 | 82.61 | 87.50 | 79.17 |
| **17** | 1000 | 76.79 | 75.76 | 78.26 | 83.33 | 69.23 | 3 | 83.93 | 96.97 | 65.22 | 80.00 | 93.75 |
| **18** | 1000 | 76.79 | 75.76 | 78.26 | 83.33 | 69.23 | 8 | 83.93 | 87.88 | 78.26 | 85.29 | 81.82 |
| **19** | 1000 | 80.36 | 75.76 | 86.96 | 89.29 | 71.43 | 33 | 85.71 | 87.88 | 82.61 | 87.88 | 82.61 |
| **20** | 1000 | 78.57 | 78.79 | 78.26 | 83.87 | 72.00 | 5 | 85.71 | 87.88 | 82.61 | 87.88 | 82.61 |
| **21** | 1000 | 75.00 | 72.73 | 78.26 | 82.76 | 66.67 | 35 | 78.57 | 78.79 | 78.26 | 83.87 | 72.00 |
| **22** | 1000 | 80.36 | 81.82 | 78.26 | 84.38 | 75.00 | 12 | 85.71 | 87.88 | 82.61 | 87.88 | 82.61 |
| **23** | 1000 | 75.00 | 69.70 | 82.61 | 85.19 | 65.52 | 8 | 85.71 | 96.97 | 69.57 | 82.05 | 94.12 |
| **24** | 1000 | 83.93 | 81.82 | 86.96 | 90.00 | 76.92 | 14 | 83.93 | 81.82 | 86.96 | 90.00 | 76.92 |
| **25** | 1000 | 80.36 | 84.85 | 73.91 | 82.35 | 77.27 | 5 | 87.50 | 96.97 | 73.91 | 84.21 | 94.44 |
| **26** | 1000 | 73.21 | 69.70 | 78.26 | 82.14 | 64.29 | 6 | 89.29 | 90.91 | 86.96 | 90.91 | 86.96 |
| **27** | 1000 | 82.14 | 84.85 | 78.26 | 84.85 | 78.26 | 12 | 83.93 | 100 | 60.87 | 78.57 | 100 |
| **28** | 1000 | 80.36 | 81.82 | 78.26 | 84.38 | 75.00 | 5 | 82.14 | 90.91 | 69.57 | 81.08 | 84.21 |
| **29** | 1000 | 80.36 | 81.82 | 78.26 | 84.38 | 75.00 | 15 | 82.14 | 84.85 | 78.26 | 84.85 | 78.26 |
| **30** | 1000 | 75.00 | 72.73 | 78.26 | 82.76 | 66.67 | 7 | 78.57 | 81.82 | 73.91 | 81.82 | 73.91 |

**Tab. III** *Classification accuracies of LM from 1000 (left) and optimal iterations (right) (%). (OA: Overall Accuracy, PA: Procedure Accuracy, UA: User Accuracy).*

| # of Run | Iteration | OA | PA | | UA | | Iteration | OA | PA | | UA | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Oil | Alike | Oil | Alike | | | Oil | Alike | Oil | Alike |
| **1** | 1000 | 82.14 | 87.88 | 73.91 | 82.86 | 80.95 | 457 | 85.71 | 84.85 | 86.96 | 90.32 | 80.00 |
| **2** | 1000 | 80.36 | 87.88 | 69.57 | 80.56 | 80.00 | 495 | 83.93 | 93.94 | 69.57 | 81.58 | 88.89 |
| **3** | 1000 | 80.36 | 87.88 | 69.57 | 80.56 | 80.00 | 291 | 87.50 | 90.91 | 82.61 | 88.24 | 86.36 |
| **4** | 1000 | 83.93 | 87.88 | 78.26 | 85.29 | 81.82 | 367 | 89.29 | 96.97 | 78.26 | 86.49 | 94.74 |
| **5** | 1000 | 80.36 | 84.85 | 73.91 | 82.35 | 77.27 | 468 | 87.50 | 93.94 | 78.26 | 86.11 | 90.00 |
| **6** | 1000 | 83.93 | 93.94 | 69.57 | 81.58 | 88.89 | 688 | 85.71 | 93.94 | 73.91 | 83.78 | 89.47 |
| **7** | 1000 | 89.29 | 96.97 | 78.26 | 86.49 | 94.74 | 891 | 91.07 | 96.97 | 82.61 | 88.89 | 95.00 |
| **8** | 1000 | 85.71 | 87.88 | 82.61 | 87.88 | 82.61 | 558 | 89.29 | 90.91 | 86.96 | 90.91 | 86.96 |
| **9** | 1000 | 80.36 | 87.88 | 69.57 | 80.56 | 80.00 | 425 | 87.50 | 90.91 | 82.61 | 88.24 | 86.36 |
| **10** | 1000 | 83.93 | 90.91 | 73.91 | 83.33 | 85.00 | 264 | 85.71 | 90.91 | 78.26 | 85.71 | 85.71 |
| **11** | 1000 | 80.36 | 84.85 | 73.91 | 82.35 | 77.27 | 293 | 87.50 | 90.91 | 82.61 | 88.24 | 86.36 |
| **12** | 1000 | 85.71 | 93.94 | 73.91 | 83.78 | 89.47 | 164 | 89.29 | 96.97 | 78.26 | 86.49 | 94.74 |
| **13** | 1000 | 80.36 | 84.85 | 73.91 | 82.35 | 77.27 | 695 | 87.50 | 96.97 | 73.91 | 84.21 | 94.44 |
| **14** | 1000 | 82.14 | 78.79 | 86.96 | 89.66 | 74.07 | 437 | 82.14 | 81.82 | 82.61 | 87.10 | 76.00 |
| **15** | 1000 | 80.36 | 81.82 | 78.26 | 84.38 | 75.00 | 109 | 82.14 | 84.85 | 78.26 | 84.85 | 78.26 |
| **16** | 1000 | 82.14 | 84.85 | 78.26 | 84.85 | 78.26 | 470 | 85.71 | 87.88 | 82.61 | 87.88 | 82.61 |
| **17** | 1000 | 85.71 | 96.97 | 69.57 | 82.05 | 94.12 | 553 | 91.07 | 93.94 | 86.96 | 91.18 | 90.91 |
| **18** | 1000 | 83.93 | 84.85 | 82.61 | 87.5 | 79.17 | 323 | 85.71 | 87.88 | 82.61 | 87.88 | 82.61 |
| **19** | 1000 | 83.93 | 87.88 | 78.26 | 85.29 | 81.82 | 192 | 85.71 | 93.94 | 73.91 | 83.78 | 89.47 |
| **20** | 1000 | 87.50 | 90.91 | 82.61 | 88.24 | 86.36 | 211 | 89.29 | 93.94 | 82.61 | 88.57 | 90.48 |
| **21** | 1000 | 89.29 | 96.97 | 78.26 | 86.49 | 94.74 | 970 | 91.07 | 96.97 | 82.61 | 88.89 | 95.00 |
| **22** | 1000 | 83.93 | 84.85 | 82.61 | 87.50 | 79.17 | 162 | 91.07 | 96.97 | 82.61 | 88.89 | 95.00 |
| **23** | 1000 | 83.93 | 87.88 | 78.26 | 85.29 | 81.82 | 666 | 87.50 | 90.91 | 82.61 | 88.24 | 86.36 |
| **24** | 1000 | 87.50 | 90.91 | 82.61 | 88.24 | 86.36 | 961 | 89.29 | 93.94 | 82.61 | 88.57 | 90.48 |
| **25** | 1000 | 83.93 | 90.91 | 73.91 | 83.33 | 85.00 | 514 | 87.50 | 90.91 | 82.61 | 88.24 | 86.36 |
| **26** | 1000 | 83.93 | 87.88 | 78.26 | 85.29 | 81.82 | 995 | 83.93 | 87.88 | 78.26 | 85.29 | 81.82 |
| **27** | 1000 | 83.93 | 84.85 | 82.61 | 87.50 | 79.17 | 763 | 85.71 | 87.88 | 82.61 | 87.88 | 82.61 |
| **28** | 1000 | 85.71 | 90.91 | 78.26 | 85.71 | 85.71 | 305 | 85.71 | 93.94 | 73.91 | 83.78 | 89.47 |
| **29** | 1000 | 76.79 | 75.76 | 78.26 | 83.33 | 69.23 | 694 | 83.93 | 90.91 | 73.91 | 83.33 | 85.00 |
| **30** | 1000 | 87.50 | 87.88 | 86.96 | 90.63 | 83.33 | 439 | 89.29 | 90.91 | 86.96 | 90.91 | 86.96 |

**Tab. IV** *Classification accuracies of ABC from 1000 (left) and optimal iterations (right) (%). (OA: Overall Accuracy, PA: Procedure Accuracy, UA: User Accuracy).*

| | Minimum | Maximum | Mean | Std. | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| **BP** | 55.36 | 83.93 | 75.24 | 7.56 | -1.2516 | 0.5878 |
| **LM** | 58.93 | 83.93 | 76.61 | 5.29 | -1.3237 | 2.4329 |
| **ABC** | 76.79 | 89.29 | 83.63 | 2.97 | -0.0043 | -0.3326 |

**Tab. V** *Descriptive statistics of the overall accuracies from 1000 iterations (%).*

| | Minimum | Maximum | Mean | Std. | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| **BP** | 58.93 | 87.50 | 76.73 | 7.76 | -1.0537 | -0.0121 |
| **LM** | 78.57 | 89.29 | 83.69 | 2.38 | -0.1972 | 0.3041 |
| **ABC** | 82.14 | 91.07 | 87.14 | 2.54 | -0.1516 | -0.7134 |

**Tab. VI** *Descriptive statistics of the overall accuracies from optimal iterations (%).*

level of 0.05 ($\alpha$) and sample size of 30, the non-rejection regions of $H_0$ for skewness and kurtosis are [-0.847 0.847] and [-1.08 2.12], respectively (MVP program 2010). According to both skewness and kurtosis tests through Tabs. V–VI, ABC distribution is the nearest one to a Normal Distribution, which supports that the performance of ABC algorithm is the most stable and robust.

In order to show that three algorithms are statistically different from each other, their paired groups (not paired samples) were statistically tested by two-sided Student's t-test with significance level ($\alpha$) of 5%. Since true standard error (standard deviation of sampling distribution) is unknown, using a Student's t-test is more appropriate than a z-test. In the Student's t-test, the null hypothesis ($H_0$) claims that the means of the groups are equal to each other and the alternative hypothesis ($H_1$) claims that they are not equal, which are given in Equation (4):

$$H_0\text{: } \mu_1 = \mu_2 \text{ versus } H_1\text{: } \mu_1 \neq \mu_2 \text{ with rejection region R: } |T| \geq \ t_{\alpha/2} \text{ or } P \leq \alpha \tag{4}$$

The main assumptions for the Student's t-test are that the groups are independent and normally distributed. Since the size of the groups is 30, the normal distribution assumption is accepted as valid, based on the central limit theorem. Even if this assumption is violated, the tests about a population mean are relatively robust. The detailed explanations can be found in Johnson and Bhattacharyya (2001). According to the equal and unequal variances, the test statistics and P values are given in Tabs. VII–VIII.

The results pointed out that the ABC group was completely different from the other ones while the LM and BP groups were statistically even, which supports the superior performance of ABC against the others.

The distribution graphics (Figs. 3–5) clearly show the robustness of ABC over BP and LM. In other words, the scale of the curve of ABC is narrower than and the location of the curve of ABC is bigger than the ones for LM and BP. Presenting a higher density of the normal curve means the higher probability of producing

| Test Pairs | T statistics | Degrees of Freedom | Critical Value $\lvert t_{\alpha/2} \rvert$ | P Value | Decision |
|---|---|---|---|---|---|
| ABC – BP | 5.66 | 58 | 2.00 | $4.88 \times 10^{-7}$ | Rejection of $H_0$ |
| ABC – LM | 6.34 | 58 | 2.00 | $3.72 \times 10^{-8}$ | Rejection of $H_0$ |
| BP- LM | 0.81 | 58 | 2.00 | 0.42 | Non-Rejection of $H_0$ |

**Tab. VII** *Student's t-test parameters with equal variance assumption ($\alpha$=0.05).*

| Test Pairs | T statistics | Degrees of Freedom | Critic value $\lvert t_{\alpha/2} \rvert$ | P Value | Decision |
|---|---|---|---|---|---|
| ABC – BP | -5.66 | 37.74 | 2.00 | $1.69 \times 10^{-6}$ | Rejection of $H_0$ |
| ABC – LM | -6.34 | 45.60 | 2.00 | $9.25 \times 10^{-8}$ | Rejection of $H_0$ |
| BP- LM | 0.81 | 51.93 | 2.00 | 0.42 | Non-Rejection of $H_0$ |

**Tab. VIII** *Student's t-test parameters with unequal variance assumption ($\alpha$=0.05).*

results in that interval. However, the curves of optimal iterations of LM and ABC are close (2.38, 2.54), the location of the curve of ABC is bigger than of LM (83.69, 87.14). In fact, it is reasonable to take into account the results from 1000 iterations to decide which algorithm is better rather than optimal performances. Though, the distribution bars and fitted curves of performances from the last and optimal iterations point out that ABC is the best as illustrated by the numerical results. The results of LM and BP were statistically equivalent. However, it can be claimed that LM is better than BP from observing the bars and the curves.

The general convergence characteristics of algorithms can be seen from Fig. 6, which illustrates the mean performances of the algorithms from 30 runs. From Fig. 6, it is evident that LM accelerates the fastest within 100 iterations, it then gets stuck at a local minimum. The convergence speed of the ABC algorithm is fast within the first 300 and then it slows down but evolves with a low convergence rate to the global minimum (optimum solution). The BP algorithm is the slowest one in terms of the convergence rate and seems to evolve beyond 1000 iterations. However, tests for more than 1000 iterations (up to 20000 iterations) were studied and did not produce any further gain.

# 7. Conclusions

Oil spill is an important threat to marine environment, for which early warning systems have been developed. The basic problem is to decide which dark formations are oil and which ones are not. After appropriate features are obtained, classification becomes the vital phase. ANN classifiers are the most popular and superior

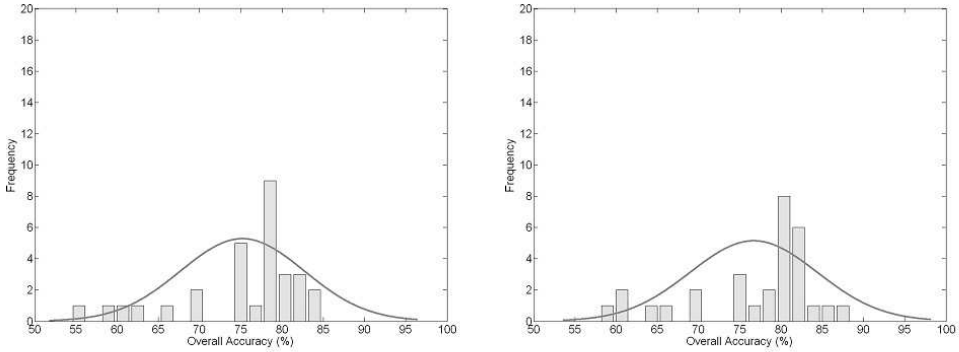


**Fig. 3** *The accuracy distributions of BP from (a) 1000 iteration and (b) the optimal number of iterations.*

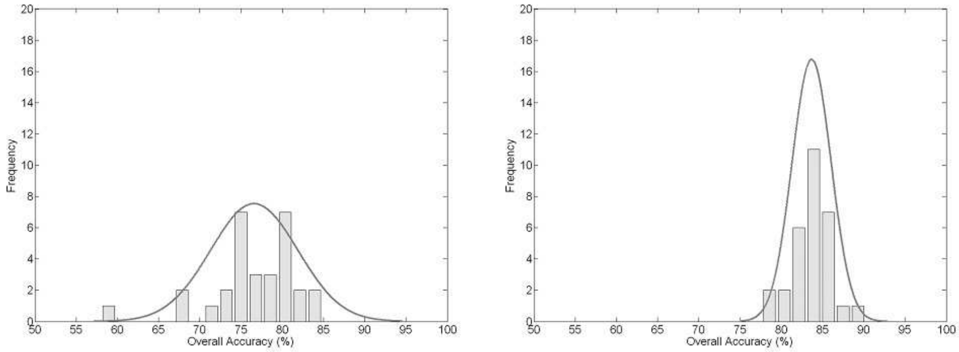


**Fig. 4** *The accuracy distributions of LM from (a) 1000 iteration and (b) the optimal number of iterations.*

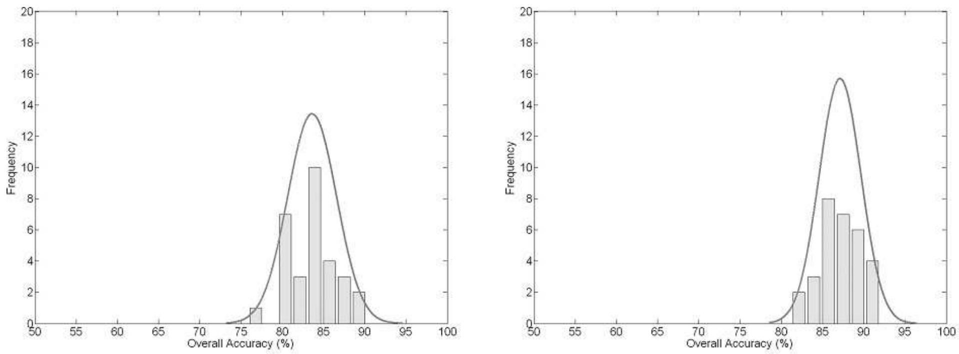


**Fig. 5** *The accuracy distributions of ABC from (a) 1000 iteration and (b) the optimal number of iterations.*

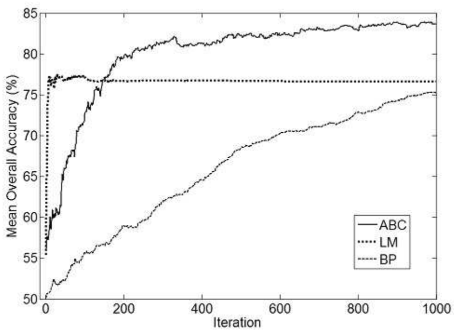


**Fig. 6** *Mean accuracies of ABC, LM, and BP algorithms from 30 runs.*

models to discriminate oil spills candidates. The performance of the ANN models is highly depended on the training algorithm. Among the conventional methods, LM is the most preferred algorithm because of convergence speed and performance, on the other hand, derivative based algorithms present the danger of tackling local minima. In this study, a new generic heuristic optimization algorithm, Artificial Bee Colony (ABC), was tested on training MLP ANN for oil spill classification against BP and LM algorithms.

Radarsat-1 SAR images of the Lebanese power plant case were used in the experiments, which include a total of 121 dark objects (68 oil and 53 look-alike objects). The 9 features of the geometric, physical and textural characteristics were extracted as input. The performance analysis was based on the 30 different runs of algorithms. According to the results, the ABC algorithm statistically gave different results than the BP and LM algorithms, and ABC has performed not only the best mean overall accuracy from 30 runs, but also the most robust algorithm against BP and LM. Even though BP seemed having comparable results with LM, procedure and user accuracies imprinted that LM was better than BP.

Based on the good performance of the ABC algorithm, the general conclusion is that ABC can certainly be used to train an MLP ANN on oil spill classification instead of BP or LM with a high confidence.

As a future work, it is planned to work with multiclass classification problems (more than two classes) of the multispectral optical remote-sensing data for land cover-land use. Besides, the ABC algorithm will be studied for unsupervised classification against conventional clustering techniques.

**Acknowledgments**

The authors would like to thank to ITU-CSCRS for providing the satellite images used in this study.

# References

[1] Alpers W., Wismann V., Theis R., Huhnerfuss H., Bartsch N., Moreira J., Lyden J.: The damping of ocean surface waves by monomolecular sea slicks measured by airborne multi-

frequency radars during the SAXON-FPN experiment. Proceedings of the International Geoscience and Remote Sensing Symposium, Helsinki, Finland, 1991, pp. 1987-1990.

[2] AMSA, http://www.amsa.gov.au/marine_environment_protection/educational_resources_and_information/teachers/the_effects_of_oil_on_wildlife.asp

[3] Assilzadeh H., Mansor S. B.: Early warning system for oil spill using SAR images. Proceedins of ACRS 2001—22nd Asian Conference on Remote Sensing, Singapore, 2001, pp. 460-465.

[4] Bava J., Tristan O., Yasnikouski J.: Earth observation applications through systems and instruments of high performance computer. ASI/CONEA training course, September-December, Matera, Italy, 2002.

[5] Bishop C.: Neural Networks for Pattern Recognition (Oxford University Press), 1995.

[6] Brekke C., Solberg A. H. S.: Feature extraction for oil spill detection based on SAR images. SCIA 2005, LNCS, **3540**, 2005a, pp. 75-84.

[7] Brekke C., Solberg A. H. S.: Oil spill detection by satellite remote sensing. Remote Sensing of Environment, **95**, 2005b, pp. 1-13.

[8] Congalton R., Green K.: Assessing the Accuracy of Remotely Sensed Data: Principles and Practices (CRC/Lewis Press, Boca Raton, FL.), 1999.

[9] Dayhof J.: Neural-Network Architectures: An Introduction (New York: Van Nostrand Reinhold), 1990.

[10] Del Frate F., Petrocchi A., Lichtenegger J., Calabresi G.: Neural networks for oil spill detection using ERS-SAR data. IEEE Transactions on Geoscience and Remote Sensing, **38**, 2000, pp. 2282-2287.

[11] Fiscella B., Giancaspro A., Nirchio F., Trivero P.: Oil spill detection using marine SAR images. International Journal of Remote Sensing, **21**, 2000, pp. 3561-3566.

[12] Foody G. M.: Thematic mapping from remotely sensed data with neural networks: MLP, RBF and PNN based approaches. Journal of Geographical Systems, **3**, 2001, pp. 217-232.

[13] Hagan M. T., Menhaj M. B.: Training feedforward networks with the Marquardt algorithm, IEEE Transactions on Neural Networks, **6**, 1994, pp. 989-993.

[14] Harahsheh H., Essa S., Shiobarac M., Nishidaid T., Onumad T.: Operational satellite monitoring and detection for oil spill in offshore of United Arab Emirates. XXth ISPRS Congress, Istanbul, Turkey, Proceedings of Commission VII, 2004, pp. 658-663.

[15] Haralick R. M., Shanmugan K., Dinstein I.: Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics, **3**, 1973, pp. 610-621.

[16] Haralick R.: Statistical and structural approaches to texture. Proceedings of IEEE, **5**, 1979, pp. 786-803.

[17] Haykin S. (2nd Ed.): Neural Networks: A Comprehensive Foundation (Upper Saddle River, NJ: Prentice Hall), 1999.

[18] Hovland H. A., Johannessen J. A., Digranes G.: Slick detection in SAR images. Proceedings of the International Geoscience and Remote Sensing Symposium, Pasadena CA, 1994, pp. 2038-2040.

[19] Johnson R. A., Bhattacharyya G. K. (4th Ed.): Statistics: Principles and Methods (John Wiley & Sons Inc.), 2000.

[20] Karaboga D.: An idea based on honey bee swarm for numerical optimization. Technical Report-TR06, Erciyes University, Computer Engineering Department, 2005.

[21] Karaboga D., Akay B.: Comparative study of Artificial Bee Colony algorithm. Applied Mathematics and Computation, **214**, 2009, pp. 108-132.

[22] Karaboga D., Basturk B.: A powerful and efficient algorithm for numerical function optimization: Artificial Bee Colony (ABC) algorithm. Journal of Global Optimization, **3**, 2007a, pp. 459-171.

[23] Karaboga D., Basturk B.: Artificial Bee Colony (ABC) optimization algorithm for solving constrained optimization problems. LNCS: Advances in Soft Computing: Foundations of Fuzzy Logic and Soft Computing, **4529**, 2007b, pp. 789-798.

[24] Karaboga D., Ozturk C.: Neural networks training by Artificial Bee Colony algorithm on pattern classification. Neural Networks World, **3**, 2009, pp. 279-292.

[25] Karaboga D., Ozturk C.: A novel clustering approach: Artificial Bee Colony (ABC) algorithm. Applied Soft Computing, **1**, 2011, pp. 652-657.

[26] Karathanassi V., Topouzelis K., Pavlakis P., Rokos D.: An object-oriented methodology to detect oil spills. International Journal of Remote Sensing, **27**, 2006, pp. 5235-5251.

[27] Kavzoglu T., Mather P. M.: The use of backpropagating artificial neural networks in land cover classification. International Journal of Remote Sensing, **24**, 2003, pp. 4907-4938.

[28] Keramitsoglou I., Cartalis C., Kiranoudis C.: Automatic identification of oil spills on satellite images. Environmental Modeling and Software, **21**, 2005, pp. 640-652.

[29] Kermani B. G., Schiffman S. S., Nagle H. T.: Performance of the Levenberg Marquardt neural network training method in electronic nose applications. Sensors and Actuators B: Chemical, **1**, 2005, pp. 13-22.

[30] Kurban T., Besdok E.: A comparison of RBF neural network training algorithms for inertial sensor based terrain classification. Sensors, **9**, 2009, pp. 6312-6329.

[31] Lawrence J.: Introduction to Neural Networks, (California Scientific, Nevada City, CA), 1993.

[32] Mas J. F., Flores J. J.: The application of artificial neural networks to the analysis of remotely sensed data. International Journal of Remote Sensing, **3**, 2008, pp. 617-663.

[33] MVP Program, 2010,
(http://mvpprograms.com/help/mvpstats/distributions/SkewnessCriticalValues)
(http://mvpprograms.com/help/mvpstats/distributions/KurtosisCriticalValues).

[34] Paola J. D., Schowengerdt R. A.: A detailed comparison of backpropagation neural network and maximum-likelihood classifiers for urban land use classification. IEEE Transactions on Geoscience and Remote Sensing, **33**, 1995, pp. 981-996.

[35] Pavlakis P., Tarchi D., Sieber A.: On the monitoring of illicit vessel discharges, A reconnaissance study in the Mediterranean Sea. European Commission, EUR 19906 EN, 2001.

[36] Rumelhart D. E., Hinton G. E., Williams R. J.: Learning representations by backpropagation errors. Nature, **323**, 1986, pp. 533-536.

[37] Sabins F. F.: (3rd Ed.): Remote Sensing, Principles and Interpretation (W. H. Freeman and Company, New York), 1997.

[38] Shi L., Ivanov A., Yu. H. M., Zhao C.: Oil spill mapping in the western part of the East China Sea using synthetic aperture radar imagery. International Journal of Remote Sensing, **21**, 2008, pp. 6315-6329.

[39] Solberg A., Brekke C., Husoy P. O.: Oil spill detection in Radarsat and Envisat SAR images. IEEE Transactions on Geoscience and Remote Sensing, 45, 2007, pp. 746-755.

[40] Solberg A., Storvik G., Solberg R., Volden E.: Automatic detection of oil spills in ERS SAR images. IEEE Transactions on Geoscience and Remote Sensing, 37, 1999, pp. 1916-1924.

[41] Solberg R., Theophilopoulos N. A.: Envisys – A solution for automatic oil spill detection in the Mediterranean. Proceedings of 4th Thematic Conference on Remote Sensing for Marine and Coastal Environments, Environmental Research Institute of Michigan, Ann Arbor, Michigan, 1997, pp. 3-12.

[42] Sunar F., Ozkan C.: Forest fire analysis with remote sensing data. International Journal of Remote Sensing, **22**, 2001, pp. 2265-2277.

[43] Topouzelis K. N.: Oil spill detection by SAR images: Dark formation detection, feature extraction and classification algorithms. Sensors, **10**, 2008, pp. 6642-6659.

[44] Topouzelis K., Stathakis D., Karathanassi V.: Investigation of genetic algorithms contribution to feature selection for oil spill detection. International Journal of Remote Sensing, **3**, 2009, pp. 611-625.

[45] Törmä M., Lumme J., Pyysalo U., Patrikainen N., Luojus K.: Tree species classification using ERS SAR and MODIS NDVI images, XXth ISPRS Congress, Istanbul, Turkey, Proceedings of Commission VII, 2004, pp. 927-932.

[46] Weigend A.: On Overfitting and the efective number of hidden units. Proceedings of the Connectionist Models Summer School, (Hillsdale, NJ), 2004, pp. 335-342.

[47] Xiaoying J.: Segmentation-based image processing system. ITT Manufacturing Enterprises, Inc., US Patent 20090123070, 2009.