A Decision Fusion Method Based on Classification Models for Water Quality Monitoring

Mohamed LADJAL (✉ mohamed.ladjal@univ-msila.dz)

University of M'sila <https://orcid.org/0000-0003-2730-0030>

Mohamed BOUAMAR

University of M'sila

Youcef BRIK

University of M'sila

Mohamed DJERIOUI

University of M'sila

Research Article

Keywords: Surface water quality monitoring, Principal Component Analysis, Features selection, Support Vector Machine, Artificial Neural Network, Decision Templates.

Posted Date: December 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1013779/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A decision fusion method based on classification models for water quality monitoring

Mohamed LADJAL✉*, Mohamed BOUAMAR, Youcef BRIK, Mohamed DJERIOUI

*LASS, Laboratory of Analysis of Signals and Systems,
Department of Electronics, Faculty of Technology, University of M'sila, Algeria
PB N° 166 Avenue Ichbilia, 28000 M'sila, Algeria*

✉ mohamed.ladjal@univ-msila.dz, mohamed.bouamar@univ-msiladz, youcef.brik@univ-msila.dz mohamed.djerioui@univ-msila.dz

Abstract— Monitoring of water quality is one of the world's main intentions of countries. In this paper we present the use of Principal Component Analysis (PCA) combined with Support Vector Machines (SVM) and Artificial Neural Network (ANN) based on Decision Templates combination data fusion method. SVM and ANN are employed in classification stage. Decision Templates is applied to increase accuracy of the water quality classification compared to others combination data fusion methods. This work concerned the water quality assessment from Tiledit dam (Algeria) that it permitted us to acquire additional knowledge and information about study area and to obtain an intelligent monitoring system. The Multi-Layer Perceptron network (MLP) and the One-Against-All strategy for SVM method are have been widely used.

The training step is performed in this paper using these techniques to classify water quality from various physicochemical parameters such as temperature, pH, electrical conductivity and turbidity, etc. Eight of them were collected in the period 2009-2018 from the study area. The selection of the excellent parameters of the used models can be improving the performance of classification process. In order to assess their results, an experiment step using collected data set corresponding to the accuracy and running time of training and test phases, and robustness, is carried out. Various scenarios are examined in comparative study to obtain the most results of decision step with and without features selection of the input data. The combination by Decision Templates of two classifiers enhanced expressively the results of the proposed monitoring framework that had prove a considerable ability in surface water quality assessment.

Keywords— Surface water quality monitoring, Principal Component Analysis, Features selection, Support Vector Machine, Artificial Neural Network, Decision Templates.

I. INTRODUCTION

The quality of surface water plays a crucial and strategic role in people's health, sustainable development and ecological systems (Y. Wang et al., 2013). Efficient information about the location and quality of surface water helps for approving the performances and assessment in diverse scientific fields, such as the surface water survey and management, water resources assessment, and environmental monitoring (ZHOU & WU, 2008). Water quality control and monitoring play an important role in ecological running management and it presents a considerable concern for conservation and rational utilization of water supplies in the world (BOUAMAR & LADJAL, 2012). However, some methods for water quality assessment are usually employed with incapacity due to complicated relationships between water quality monitoring parameters and qualitative status. The methods used for the water quality monitoring, are: Matter Element Model, Fuzzy Synthetic Evaluation, Gray Analysis Method, Logistic Curve Model, Attribute Recognition Model and Fuzzy Logic (Liu & Zou, 2012 ; Wang et al., 2019; ZOU et al., 2006; W. Zhang et al., 2018; Jin et al., 2003; Wang, Wang & Zou, 2008; WANG & YANG, 2010). The Artificial Neural Networks (ANN) method is frequently considered as a solution to this kind of modeling process among others the water quality monitoring (Areerachakul & Sanguansintukul, 2010). Methods like Support Vector Machines (SVM) and Artificial Neural Networks (ANN) were efficiently used in financial analysis (Wu et al., 2007), chemical process (Liao et al., 2011), image processing and face recognition (Déniz et al., 2003), biomedical (Mookiah et al., 2012) and Arabic text and speech analysis (Hend & Amani, 2010). In this study, the monitoring process is a multiclass classification problem. To be efficient, the preparation of data sets requires a certain treatment, which ensures that the classifier models are well decided. Recently, the use of features selection for data sets treatment in classification applications has received significant attention which can increase the performance of classifier (Widodo & Yang, 2007; Cao et al., 2003). However, we need to select features to prevent the curse of dimensionality phenomenon and redundancy, since irrelevant features would decrease the classifier's accuracy (Widodo et al., 2007).

The curse of dimensionality with high-dimensional data refers to the phenomena that occur when classifying data in order to train a precise model. Therefore, the selection of subset of relevant and useful features without any transformation is desirable (YANG et al., 2006). Many approaches to feature selection are based on linear methods such as Principal Component Analysis (PCA) (Widodo et al., 2007). The use of features selection can improve the performance of classification results by reduction in the number of data input needed to attain training with significant data base and reducing running time (Kumar et al., 2005; Widodo et al., 2007). The purpose of this paper is to integrate PCA in combination with SVM and ANN in the evaluation of the quality of water. The

* Corresponding author: Mohamed LADJAL
E-mail address: mohamed.ladjal@univ-msila.dz

PCA is employed precisely here as a features selection method that to describe easily and with the best way the correlations between a list of variables (Ladjal et al., 2020), by generating a set of orthogonal Principal Components, i.e. not correlated; thereby reducing the dimensionality of the original data set. SVM and ANN are employed as multiclass classifiers based on three classes of water quality. A comparative study is examined with and without features selection process. The combination data fusion by Decision Templates of two classifiers is performed. Nevertheless, there were no references which applied the fusion of these methods by Template decision in water quality monitoring. Decision Templates combines seamlessly the outputs of the best models of both SVM and ANN classifiers to enhance accuracy of the water quality classification.

This study is particularly limited to an experiment work carried out using dataset collected from study area. The choice of the suitable approach is the object of this work by the application of SVM and ANN multi-class methods combined with PCA and using Decision Template's rule combination data fusion with respect to recognition rates, training times and sensitivity to the noise.

II. MATERIALS AND METHODS

2.1. Study site and data descriptions

The study area (Tiledit dam-Fig. 1 and 2), is situated in the region of Bechloul, 20 km southeast of Algeria's Bouira Department (Ladjal et al., 2016). It is located approximately 122 km east of Algiers (35° 13' 22' ''N 4° 14' 23' '' E) (Fig. 3). The research area is characterized by a semi arid climate. Mean annual temperatures range from 20.4 °C to 37.9 °C. Yearly precipitation averages is about 440-660 mm/year.

Fig.1. Tiledit dam [Google Earth].

Fig. 2. Photo of the study area.

Fig. 3. Map showing the Tiledit dam – Bouira – Algeria [Google Maps].

The volume of reservoir was evaluated in March 2007 at 167 million m³. Water from the dam needs to be designed to curb the tension in water distribution in 12 cities. The transfer of water including the launch of the construction works was scheduled for early 2011. Work is underway to connect many towns of Bouira Department. A processing plant with a capacity of 74000 m³/day is equipped. Water collected in the dam is pumped to the treatment plant. This being at the same place, is commissioned since 2009. It performs the purification process through the five processing levels: pre-treatment, pre-oxidation, clarification, disinfection and refining. The clarification step is performed by the method of coagulation-flocculation, decantation and filtration through a phase separator and a sand filtration stage.

In this paper, we search to develop our framework approach of control and monitoring of Water quality using several descriptors provided in a water production plant by certain physical sensors. These parameters are collected during three years from the Tiledit production plant (2009-2018). The parameters like pH, Temperature (T°), Electrical Conductivity (EC) and Turbidity (TU) are collected by sensors installed in all treatment process of

the station (Ladjal et al., 2016; Ladjal et al., 2020). Every week in the lab some chemical parameters are examined such as: Magnesium (Mg), Bicarbonate (B), Total Hardness (TH) and Full Title Alkaline (FTA). The above-mentioned collected data will be applied to check the water quality assessment model. Summary table of Statistical characteristics of the collected parameters of water under study are given by:

Table 1. Statistical characteristics of the collected parameters.

Variables	Min	Max	Average	Standard deviation
pH	7,15	8,30	7,567	0,25
EC	414,00	624,00	585,393	36,278
T°	9,70	24,20	16,13	3,483
TU	1,320	23,81	3,835	2,392
Mg	7,290	47,628	22,268	4,931
B	158,620	289,14	222,497	23,213
TH	0,00	168,00	32,287	23,029
FTA	130,00	237,00	181,845	18,703

2.2. Proposed framework

Water quality assessment can be examined in this study as a multiclass classification modeling problem. In general, it includes data acquisition and processing, selection of features, and classification of water quality. Figure 4 illustrates our proposed framework, based first on preparing data sets using PCA algorithms for feature selection before entering into ANN and SVM classifiers. Decisions concerning water quality status are obtained by using only a subset of appropriate and usable characteristics, without any transformation. In order to enhance the recognition rates, Decision Templates combination data fusion is performed.

Fig. 4. Flowchart of combining data fusion based approaches with Decision Templates.

The aim is to classify the quality of water into three separate classes from selected variables: pH, Temperature (T°), Turbidity (TU) and Electrical Conductivity (EC). The next sections contain a short description of the principal methods employed in this paper.

2.2.1. Principal Component Analysis based on features selection

PCA is a technique commonly used to reduce multivariate problem dimensions. It was used in many areas like feature extraction and selections, High-dimensional data visualization, cluster analysis, pattern recognition,

classification and regression. Principal Components (PCs) have a minimum loss of input data information in this method.

Without any transformation, PCA transforms a subset of features that may be correlated into smaller uncorrelated features or variables called PCs (Areerachakul & Sanguansintukul, 2010). All of these factors are orthogonal to each other, so no redundant features exist. By the following equation, PCs can be defined: (Jolliffe, 2010):

$$z_{ij} = a_{i1}x_{1j} + a_{i2}x_{2j} + \dots + a_{im}x_{mj} \quad (1)$$

Where z_{ij} represent PCs, a_{im} the related eigenvectors and x_{mj} input features, i is the component number, j is the sample number and m is the total number of features (Ladjal et al., 2020). This information is obtained through the resolution of the equation (Semmlow 2004):

$$|R - I\lambda| = 0 \quad (2)$$

Where I is the unit matrix, λ is the eigenvector and R is the variance-covariance matrix.

PCA results are generally evaluated by means of component values (values of the transformed features that match a certain data point) or loading values, also known as factor scores (weight to multiply each standardized original function in order to achieve the component score).

We have N samples of M -dimensional data:

- *Step 1:* To implement PCA, we should first calculate the variance-covariance matrix.
- *Step 2:* Search for the matrix of eigenvectors and diagonal matrix components as variance-covariance matrix values.
- *Step 3:* Sort the PC eigenvectors in the decreasing order of importance of eigenvalues.
- *Step 4:* By taking the dot product between the determined data and eigenvectors, project the data input into the directions of sorted eigenvectors.
- *Step 5:* Based on the containment of a specified percentage of variability, select the first few PCs.

The PCs have the following characteristics (Jolliffe, 2010):

- They are uncorrelated.
- They have sequentially maximum variance.
- The mean-squared approximation error in the representation by the first, several PCs of the initial inputs are minimal.

2.2.2. Artificial Neural Network

As shown in figure 5, Artificial Neural Networks (ANN) are multi-layer, completely connected neural networks. ANN architectures have been classified into different types based on their training mechanisms and other features. By using the Multi-Layer Perceptron (MLP) architecture (Jin et al., 2003), non linear data classification is carried out. Many recent studies are used effective ANN models for water quality monitoring (Liu & Zou, 2012; ZOU et al., 2006). The supervised learning process consists of calculating the weights that reduce the differences in the training set between the target output values y_r and the computed output values y_d .

Fig. 5. Exemple of a Multi-Layer Perceptron

A minimum of the following problem of quadratic optimization can be found:

$$C_w = \frac{1}{N} \sum_{i=1}^N (y_{r_i} - y_{d_i})^2 \quad (3)$$

N presents the samples number of the learning dataset.

We used in this application the algorithm of Levenberg-Marquardt that known as 2^{nd} order method and it is rather better because they supply however much good results.

2.2.3. Support Vector Machines

The SVM method developed by Vapnik has been extensively used to classification, regression and density estimation (Vapnik, 2000; Schölkopf & Smola, 2018). In this method, through the construction of the optimal hyperplane, which is evaluated to optimize the generalization potential of the classifier, an initial input data space is mapped in a higher dimension space by selecting some non-linear functions, called kernel functions (Übeyli, 2009).

Non linear SVM classification

Using kernel methods, the initial input data should be implicitly mapped to a typically higher-dimensional feature space. In this mapping space, the classification method is then performed via the construction of the optimal linear separating hyperplane (Figure 6). In this mapping space, the classification process is then carried out by building the optimal linear separating hyperplane (Figure 6) with a maximization margin between it and the nearest point to obtain high generalization capacity via the quadratic optimization problem. The problem of quadratic SVM optimization for binary classification was established by the following dataset:

$$(x_i, y_i), y_i \in \{-1, +1\}, i = 1, \dots, n \quad (4)$$

n is the number of observations, $x \in \mathbb{R}^d$ and y_i is a distribution and the corresponding class label respectively.

Fig. 6. The optimal hyperplane and Margin of a binary SVM.

The optimal separating hyperplane is determined by the vector of weight w and a constant b , defined by (Bae et al., 2010; Ladjal et al., 2020):

$$w.x + b = 0 \quad (5)$$

Under constraints

$$\begin{aligned} (w.x_i) + b &\geq +1, \quad \text{if } y_i = +1 \\ (w.x_i) + b &\leq -1, \quad \text{if } y_i = -1 \end{aligned} \quad (6)$$

The primal quadratic minimization problem according to w, b is given by (Singh et al., 2011):

$$\begin{cases} \min_w & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\ \text{with} & y_i (wx + b) \geq 1 - \xi_i \quad \xi_i \geq 0, \quad i = 1, \dots, n \end{cases} \quad (7)$$

The dual quadratic maximization problem using multipliers of Lagrange α_i is given by (Hend & Amani, 2010; Singh et al., 2011):

$$\begin{cases} \max_{\alpha_i} & L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j x_i x_j \\ \text{with} & \sum_{i=1}^n \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C \quad i = 1, \dots, n \end{cases} \quad (8)$$

C is a penalization parameter that control the level of errors in classification.

The nonlinear mapping $\prec \Phi \succ$ is carried out via a kernel function $K(x_i, x_j)$ from an input space to some higher dimensional feature space (Wu et al., 2007; Ladjal et al., 2020). The new dual quadratic maximization problem is defined by (Hend & Amani, 2010; Horng, 2009):

$$\begin{cases} \max_{\alpha_i} & L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ \text{with} & \sum_{i=1}^n \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C \end{cases} \quad (9)$$

Thus, the vector solution: $\alpha^0 = (\alpha_1^0, \dots, \alpha_n^0)$. Using the theorem of *Karush-Kuhn-Tucker (KKT)* for an optimal weight vector α , is (Horng, 2009):

$$\alpha_i^0 \{y_i [(w_0 x_i) + b_0] - 1\} = 0, i = 1, \dots, n \quad (10)$$

This means $\alpha_i^0 = 0$ or $y_i [(w_0 x_i) + b_0] = 1$, the letter corresponds to the Support Vectors (SVs) present the nearest point to optimal hyperplane, which is equivalent to (Horng, 2009):

$$SVs = \{x_i \text{ that } \alpha_i > 0\} \quad (11)$$

The function of decision is defined by (Singh et al., 2011; Horng, 2009):

$$f(x) = \text{sign}(\sum_{SVs} \alpha_i y_i K(x_i, x) + b) \quad (12)$$

If $f(x) < 0$, then x belongs to class -1; if not, it belongs to class 1, as b is the solution of the equation (Ladjal et al., 2020).

We can use all functions that satisfy Mercer's theorem like as kernel function with appropriate parameters selection for more performances. The widely employed kernel functions are (Vapnik, 2000; Schölkopf & Smola, 2018; Abedi et al., 2012):

The polynomial function:

$$k(x, x') = (\gamma \cdot x^T \cdot x' + c)^d \quad (13)$$

with $c \geq 0$ and $d \in \mathbb{N}$

The radial basis function (RBF):

$$k(x, x') = \exp(-\frac{\|x - x'\|^2}{2\sigma^2}) \quad (14)$$

One-Against-All approach

The "One-Against-All" approach is the one most common used in multiclass classification problems (Burges, 1998 ; Deng et al., 2011). Consider a multi-class k -class problem, where we can have N examples of training set: $\{(x_1, y_1), \dots, (x_N, y_N)\}$. Here $x_i \in \mathbb{R}^k$ is a k -dimensional input features vector and $y_i \in \{1, 2, \dots, k\}$ is the corresponding class output. The One-Against-All approach (OAA) constructs k binary SVM models where the number of classes needed is k . With all the training samples in the i^{th} class with positive class and all the other

samples with negative class, the i th SVM is trained. The i^{th} SVM defines the following optimization issue that results in the final decision function: $f_i(x) = w_i^T \phi(x) + b_i$ (Wu et al., 2007; Deng et al., 2011):

$$\begin{aligned} \underset{w^i, b^i, \xi^i}{\text{minimise}} \quad & \frac{1}{2} \|w^i\|^2 + C \sum_{i=1}^N \xi_j^i (w^i)^T \\ \text{subject to} \quad & (w^i)^T \phi(x_j) + b^i \geq 1 - \xi_j^i, \quad \text{if } y_j = i, \\ & (w^i)^T \phi(x_j) + b^i \leq -1 + \xi_j^i, \quad \text{if } y_j \neq i, \\ & \xi_j^i \geq 0, \quad j = 1, \dots, N, \end{aligned} \quad (15)$$

where $\hat{y}_j = 1$ if $y_j = i$ and $\hat{y}_j = -1$ otherwise.

Sample x is classified as in class i^* at the classification step, whose f_{i^*} producing the largest value (WANG & YANG, 2010):

$$i^* = \arg \max_{i=1, \dots, k} f_i(x) = \arg \max_{i=1, \dots, k} (w_i^T \phi(x) + b_i). \quad (16)$$

2.2.4. Methods of Classifier Combination fusion

Integrating information from multiple sources and making combined decisions from these sources is becoming a common task across several disciplines and applications. Simple set of well-known combination data fusion methods such as: Minimum, Maximum, Majority voting and Average compared with Decision Templates have been broadly applied to build a multiple classifier model for our proposed approach.

Conventional Fusion Methods

- **Average, Maximum and Minimum:** There is a similar idea to these approaches. The maximum method selects the largest value for each class among the outputs of the classifiers. Then the limit is compared and a class with a greater value is chosen. It is calculated as follows for a multi-class problem (M) with L classifier models (Kuncheva et al., 2001):

$$\max_{z=1, \dots, M} = \left\{ \max_{y=1, \dots, L} \{d_{y,z}(x)\} \right\} \quad (17)$$

Here, $d_{y,z}(x)$ is the degree of support determined, by the y^{th} classifier for the example x of the class z . The average and the minimum methods are the same as the maximum method except that the smallest values are compared as (Kuncheva et al., 2001; Jun & Sung., 2007):

$$\max_{z=1, \dots, M} = \left\{ \min_{y=1, \dots, L} \{d_{y,z}(x)\} \right\} \quad (18)$$

for the average, and the maximum methods compares the mean values as (Kuncheva et al., 2001):

$$\max_{z=1, \dots, M} = \left\{ \text{avg}_y \{d_{y,z}(x)\} \right\}, \text{avg}_y \{d_{y,z}(x)\} = \frac{1}{L} \sum_{y=1}^L d_{y,z}(x) \quad (19)$$

- **Majority Voting:** The theory of this system is that the votes obtained from each classifier are counted and that the class with the highest number of votes is affected (Ruta & Gabrys, 2005).

Decision Templates

The application of decision templates (DT) as a method of classifier combination fusion was proposed by Kuncheva (Kuncheva et al., 2001). DT is a method which makes employ of all the base classifiers used on each of the m templates (or m datasets - one per class -) with the same training set that is used for the set of classifiers (Haghighi et al., 2011). For the m multi-class problem, the classifier decisions can be organized in an output profile (DP(x)) as a matrix. The DP(x) for an example x is a matrix composed by the $d_{t,j} \in [0,1]$ elements representing the support defined by the t^{th} classifier to class ϕ_j . Decision templates DT_j are the averaged output decision profiles obtained from X_j , the set of training examples belonging to the class ϕ_j (W. Zhang et al., 2014; Chen et al., 2010a):

$$DP(x_i) = \begin{bmatrix} d_{1,1}(x_i) & \dots & d_{1,M}(x_i) \\ & d_{y,z}(x_i) & \\ d_{L,1}(x_i) & & d_{L,M}(x_i) \end{bmatrix} \quad (20)$$

$$DT_j = \frac{1}{|X_j|} \sum_{x \in X_j} DP(x) \quad (21)$$

where $d_{y,z}(x_i)$ is the degree of support defined by the y^{th} classifier for the example x_i of the class z and L is the number of classifiers in an ensemble. When decision output profiles are generated, the template of the class m is predicted as follows (W. Zhang et al., 2014; Jun & Sung.,2007):

$$DT_m = \begin{bmatrix} dt_m(1,1) & \dots & dt_m(1,M) \\ & dt_m(y,z) & \\ dt_m(L,1) & & dt_m(L,M) \end{bmatrix} \quad (22)$$

$$dt_m(y,z) = \sum_{l=1}^n u_{m,l} d_{y,z}(x_l) / \sum_{l=1}^n u_{m,l} \quad (23)$$

The similarity S between the decision template DT_j for a class ϕ_j and the decision output profile for a defined test example x is:

$$S_j(x) = 1 - \frac{1}{T \times C} \sum_{t=1}^T \sum_{k=1}^C [D_j(t,k) - d_{t,k}(x)]^2 \quad (24)$$

The last final decision of the ensemble is determined by assigning the test example to the class with the biggest similarity:

$$D(x) = \arg \max_j S_j(x) \quad (25)$$

The similarity between the decision output profile of a test example and each prototype is identified in the test process. In the class of the most comparable prototype, the example is then affected. Kuncheva (Kuncheva et al., 2001) studied DT with various distance measurements and achieved great success in classification compared to traditional combination data fusion techniques (Jun & Sung.,2007).

III. RESULTS AND DISCUSSION

3.1. Data features selection

For data features selection, the PCA method is used with 80-90% variation of eigenvalues, without any transformation of the resulting components which are uncorrelated (Héctor, 2006). A total of 1800 samples from eight physicochemical water quality parameters are used in this phase (Figure 7).

Fig. 7. Evolution section of the water quality variables.

A variance-covariance matrix is formed by using PCA on input variables. Eight eigenvalues are obtained after solving the equation (2). Table 2 presents the PCA results and statistical parameters such as: eigenvalues, cumulative variance proportion and variance proportion. The four PCs indicates 84.68% of the total input samples variance proportion and eliminates the remaining components, as set out in Table 3. These PCs calculate mainly the initial data variance.

Table 2. Statistical characteristics of the resulted PCs.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Eigenvalues								
	2.57	2.23	1.12	0.85	0.57	0.33	0.27	0.06
Total variance proportion (%)								
	32.07	27.93	14.05	10.63	7.16	4.10	3.32	0.74
Cumulative variance proportion (%)								
	32.07	60.00	74.05	84.68	91.84	95.93	99.26	100
Variables of eigenvectors obtained by applying PCA								
pH	-0.57	-0.46	-0.25	0.52	-0.08	0.29	0.20	0.00
EC	0.74	0.30	-0.37	0.08	0.26	0.33	-0.21	0.00
T°	-0.04	0.78	-0.10	-0.45	-0.28	0.22	0.22	0.00
TU	-0.26	-0.47	0.73	-0.27	0.07	0.29	-0.10	0.00
Mg	0.38	0.48	0.43	0.48	-0.44	0.03	-0.13	-0.01
B	0.87	-0.41	0.11	-0.03	0.01	0.01	0.20	-0.17
TH	-0.03	0.70	0.42	0.27	0.46	-0.02	0.20	0.02
FTA	0.85	-0.45	0.09	-0.00	-0.07	-0.01	0.16	0.17

In addition, PCA applications are used to obtain eigenvectors to evaluate the coefficients for training of PCs. The correlations between each variable and the main acquired components are shown in Table 2. In this table, most effective parameters in PCs training are exposed by bold font. The total variance in the dataset account for 84.68% of the first four principal components together. The first component (PC1) is 32.07%, with 27.93% being the second component (PC2), 14.05% being the third component (PC3) and 10.63% of the total variance being the fourth component (PC4). In general, the *EC*, *B*, and *FTA* concentrations are obvious to be the most effective for the PC1 and represent more than 32 % of input variable variance proportions. Furthermore, the *T°* and *TH* concentrations also have the most effect on the PC2, which contains more than 27% of input variables variance proportions. Moreover, *TU* and *pH* concentrations are affected by PC3 and PC4 respectively (Ladjal et al., 2020).

In table 2, the rapid decay of eigenvalues is apparent. For the evaluation of prevailing physicochemical processes, the eigenvalues of the first fourth principal components (PC1-PC4) can be used (Bhardwaj et al., 2010; Ayeni, 2013). The *EC*, *B*, and *FTA* concentrations are highly positive (0.74 - 0.87), while the *Mg* concentration is low positive for the first component (0.38). *T°* and *TH* have high positive loads in the PC2 (0.70 - 0.78), and the other concentrations show low positive loads (0.3 - 0.48). The concentrations of *TU* in the PC3 have high positive loadings (0.73), while concentrations of *FTA* have low positive loads (0.09). The *pH* concentrations for the PC4 show high positive charges (0.52), while the *Mg* display moderate positive charges (0.48), *TU* and *TH* show low positive charges (0.08 - 0.27) (Ladjal et al., 2020).

From Table 2, the first four PCs are the input features of the evaluated multi-class classifiers. Variables retained are: pH, Temperature (*T°*), Electrical Conductivity (*EC*) and Turbidity (*TU*). These results are equivalent to the results obtained in literature (Ladjal et al., 2016; Ladjal et al., 2020) with different database and period, which adopts the same selected variables and these parameters were measured in the field, using the station's sensors. In any case, this solution is not final, a relearning system should probably be conducted periodically so that situations that might arise can be taken into account and continuously adjusted to change water quality.

3.2. Samples classification with SVM and ANN

To carry on the training and classification process, data sets for the training and test phases are developed and arranged in three separate classes of water quality (I: Excellent, II: middle, III: mediocre) according to the local

environmental water quality guidelines (Décret, 2011). A collected data collection of 1800 samples (Table 1) was used.

In this work, diverse architectures of ANN and activation functions have been applied to establish the suitable number of hidden layers and neurons. The SVM using OAA approach is used to carry out the multiclass classification process. C , d and σ are the three parameters associated with the SVM kernels functions. The parameter d related to the polynomial degree, σ for RBF function and C is the penalization factor. Therefore, good choice of all parameters in the two models ANN and SVM, can show excellent results and in the opposite case can cause under fitting or over fitting problem (Widodo et al., 2007).

To assess the used methods, 10-fold cross-validation has been performed in training and testing phases. The cross-validation process can stop the overfitting problem that is very important in subsamples random selection used for testing and training datasets. We can develop classification models with high performance and accuracy through the use of cross-validation.

Table 3 indicate the results of ANN multi-class models using data input with and without features selection. The different parameters of training, such as: the global number of neurons and hidden layers and the recognition rates (Training and Testing) which are evaluated all samples by using the correct classification rate.

Table 3. Classification results using ANN models.

Number of hidden layers	Number of neurons in hidden layers	Recognition rates (%)			
		Without PCA features selection (8 Variables)		With PCA features selection (4 Variables)	
		Training	Testing	Training	Testing
1	(4)	97.56 %	86.63 %	99.83 %	97.94 %
1	(8)	100 %	84.18 %	97.83 %	98.11 %
2	(4-8)	100 %	84.15%	98.92 %	98.07 %
2	(10-10)	98.78 %	83.86 %	99.83 %	98.42 %
3	(4-8-12)	97.56 %	86.61 %	99.75 %	99.13 %

Table 4 indicate the results of samples classification with SVM multi-class models. The performance criteria, such as: Number of Support Vectors (NSV), the recognition rates for training and testing phases are determined for different kernel functions and its parameters and values of the factor C .

Table 4. Classification results of SVM models and selected kernel parameters.

Kernel parameters	Recognition rates (%)					
	Without PCA features selection (8 Variables)			With PCA features selection (4 Variables)		
	Training	Testing	NSV	Training	Testing	NSV
Linear ($d = 1, \gamma = 1, C = 1000$)	96.34 %	89.09 %	19	98.50 %	98.02 %	49
Polynomial 2 ($d = 2, \gamma = 1, C = 1000$)	100 %	84.42 %	28	98.67 %	98.00 %	53
Polynomial 3 ($d = 3, \gamma = 1, C = 100$)	100 %	85.78 %	26	98.42 %	98.20 %	73
Polynomial 4 ($d = 4, \gamma = 1, C = 100$)	100 %	85.24 %	37	98.92 %	<u>97.80 %</u>	50
Gaussian RBF ($\sigma = 0.1, C = 1000$)	100 %	86.88 %	42	99.00 %	98.48 %	63
Gaussian RBF ($\sigma = 1, C = 1000$)	100 %	85.82 %	45	99.50 %	98.04 %	40
Gaussian RBF ($\sigma = 2, C = 1000$)	100 %	84.14 %	53	98.83%	98.07%	67

In Tables 3 and 4, the recognition rates on the original dataset without features selection process is up than 96% in training step and among 83.86 % until 89.09% in testing step for the two models (ANN and SVM). The existence of irrelevant and useless features decreases the performance of classification process (Widodo & Yang, 2007). Then, as shown in cited tables, the recognition rate with PCA features selection ranged from 97.83% to 99.83% in training step and among 97.80 % until 99.13% in testing step for the two models (ANN and SVM). It is better than the precedent classification without features selection.

As shown in table 4, the effect of choice of architecture and parameters of networks is important. Indeed, the good choice of the ANN architecture characteristics can improve the performances of classification. The best model for this application is the network with three hidden layers using the original feature set with and without features selection process. This architecture is characterized by a recognition rate in testing step with features selection process of 99.13 %. For SVM model, the feature selection step increased of the performances of

classification process. It can be compared with Tables 3 and 4 in the case of with and without features selection by PCA. In Table 4, the recognition rate in training phase for linear kernel is usually lower than Polynomial and Gaussian RBF kernel with and without features selection. Even though the degrees of polynomial are 2, 3 and 4 in the process with features selection. However, the recognition rate reaches 98.48% using Gaussian RBF kernel ($\sigma = 0.1$, $C = 1000$) due to good quality of data input after features selection process. This model is characterized by the recognition rates in training and testing steps which are 99% and 98.48 % respectively. Gaussian RBF kernel has shown to be a best choice for this application (Widodo et al., 2007; BOUAMAR & LADJAL, 2012; Djerioui et al., 2018). Therefore, a good choice of kernel function and its parameters C and σ or d and γ can achieve a best performance in classification steps (Widodo et al., 2007).

The recognition rates of each model with features selection process are high. The features selection process searches the uncorrelated components from the input data using PCA that is useful to increase the performance of classification. Generally, as listed in the Tables 4 and 5, the SVM models using strategy of OAA are better than ANN models mainly when using the original data without features selection process in high-dimensional data classification (Widodo & Yang, 2007). Moreover, using kernel parameters selection will increase the performance of classification. However, the ANN models are better than SVM models mainly when using the original data set with features selection process as listed in the table against recognition rate in testing step with a small improvement.

3.3. Robustness

It is important to evaluate the robustness of techniques used to see its impact in the system decision (BOUAMAR & LADJAL, 2008). In this step, all input data value were standardized and normalized between 0 and 1 to avoid having more weight being assigned to features with larger values. The normalization step is defined by the following equation (25) (Msiza et al., 2008):

$$x_{\text{new}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (25)$$

Where x is the initial datasets, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values, respectively. The results showed that normalized data is easier to process (Liao et al., 2011). This is simulated by artificially adding a white noise to the initial data input. A Signal to Noise Ratio (SNR) is calculated using five different levels of white noise of: 10, 20, 40 and 60 dB. Table 5 present the results associated with ANN and SVM multi-class models using reduced real data input with additive white noise. The various parameters, such as: the recognition rates according to the SNR which are calculated by using the correct classification rate of real data and the Minimum Square Error (MSE) for ANN model, are presented.

The results showed that normalized data is easier to process but the recognition rate has dropped. Compared to ANN, we found a remarkable resistance to the white noise of SVM multi-class models. When the SNR is equal to or greater than 20 dB, the test vectors using SVM are absolutely insensitive to the different noise levels applied. This explains why this approach enjoys exceptional immunity. For the ANN model, when the SNR decreases, we observe a strong deterioration of the recognition rate. In addition, for a ratio lower than or equal to 10 dB, there is a really clear immunity limitation. The MSE explains this situation. However, it has been noted that when this SNR is equal to or greater than 40 dB, the test vectors do not exhibit concrete degradation. We may assume that the ANN model offers an appropriate resistance an SNR above 40 dB. Ultimately, this model tends to be more noise sensitive and therefore less stable than the SVM model. Table 6 summarizes the characteristic results corresponding to the two models carried out on the suggested real data.

Table 5. Recognition rates according to the SNR of PCA-SVM and PCA-ANN multi-class models.

Input Variables	Models with features selection	SNR (dB)	Methods	Recognition rate according to the SNR	MSE
(4 variables) (EC, T°, TU, pH)	Without Noise		ANN	96.33 %	0.04
			SVM	97.24%	
	ANN (4-8-12)	10	ANN	-	1.4
			SVM	87.80 %	
		20	ANN	83.17 %	0.15
			SVM	97.17%	
	SVM Gaussian RBF ($\sigma = 2^{-1}$, $C = 1000$)	40	ANN	95.83 %	0.05
			SVM	97.17%	
		60	ANN	96.33 %	0.04
			SVM	97.24	

Input Variables	Models with features selection	Characteristics			
		Training time (sec)	Recognition rate (%)		Robustness
			Training	Testing	
(4 variables) (EC, T°, TU, pH)	ANN (4-8-12)	65.29	99.75 %	99.13 %	Good
	SVM Gaussian RBF ($\sigma = 0.1$, C = 1000)	1.52	99.00 %	98.48 %	Excellent

It appears that the two models perform good results on the decisional level with recognition rates of more than 98 % in training and testing phases with the features selection process. In the training step, the SVM model is rather better positioned on the computing time, which gives it the benefit of integration into a dynamic monitoring system. With SNR above or equal to 20 dB, the two multi-class models have slightly strong immunity. Furthermore, the ANN model suffers from a significant handicap related to its apparent noise sensitivity. The SVM model removes this limitation because of its excellent noise robustness. Finally, we can conclude that the classification method performed using the SVM technique on a real Tilesdit dam data provides the best performance and the acceptable solution combined with the PCA features selection strategy. These results are equivalent to the results obtained by Achmad et al. (Widodo et al., 2007). This result is important because it reflected an economic impact on the overall cost of the control and monitoring system (less training time and reduced of physical sensors).

3.4. Classification using Decision Template rule combination

The results were combined after obtaining the classification accuracy of the classifiers. In this study, the fusion of primary classification results is carried out in comparison with a set of well-known combination methods such as Majority voting, Minimum, Maximum, Average, and Bayes, using the rules of the Decision Template for classifier combination. These techniques are noted as the best among all combination data fusion methods in pattern recognition (Polikar, 2006; L. Kuncheva, 2014). Because of this, it is important to find the viability of the combined structure preferred in this work. Table 7 illustrates the hybrid data fusion of the two ANN and SVM classifiers and the overall accuracy for the fusion approaches, and also of the suggested framework.

	Models with PCA features selection (4 variables: EC, T°, TU, pH)						
	ANN (4-8-12)	SVM Gaussian RBF ($\sigma = 2^{-1}$, C = 1000)	Majority voting	Maximum	Minimum	Average	Decision Template
Recognition rate (%) in testing steps	99.13%	98.48%	98.80%	99.13%	98.98%	99.20%	99.24%

Referring to Table 7, it can be seen that the fusion of the classifiers proposed significantly increased the accuracy of the classification and enhanced the efficiency of the multi-class system proposed. These results are equivalent to the results obtained by L. I. Kuncheva et al. (Kuncheva et al., 2001), Wei Chen et al. (Chen et al., 2010b), and B. Bigdelia et al. (Bigdeli et al., 2015) in different database and applications. By comparison between results shown in this table, it can be found that the two models ANN and SVM have a good ability in water quality monitoring. The classifier combination data fusion is used to increase the classification accuracy and efficacy of the proposed process, meaning that the ANN and SVM results obtained have been combined. It can be shown that with the features selection process, the classification precision increased by up to 98%. It can be shown that when we use the decision template technique, our approach offers greater classification precision, offering an increase in the recognition rate. The approach of fusing multi-classifiers with a decision template at the decision phase is more powerful compared to other approaches. This result denotes the high capability of using the classifier combination. Furthermore, it is useful for practical purposes, so that the proposed technique can be used efficiently for monitoring water quality. This means that the precision of water quality can be greatly enhanced by applying several classifiers. This growth is a strong explanation for the efficacy of the process of

data fusion and the principle of the decision template. The results obtained underline the use of multiple sources of knowledge for accurate monitoring of water quality.

IV. CONCLUSION

In this work, we have provided a performance assessment for intelligent water quality monitoring of ANN and SVM multi-class models. The research area is the Tilesdit dam in Algeria. An adequate intelligent procedure was proposed based on surface water physicochemical variables. It included PCA features selection, ANN, SVM and data fusion method. These techniques have shown good results with respect to accuracies. PCA was successfully applied to the features selection process; however, we carried out this approach to exclude irrelevant and redundant features. Therefore, the use and implementation in the field of water quality control of these approaches are well justified. We have used a cross-validation procedure particularly in this study that can prevent overfitting problems by selecting random subsamples used for training and testing datasets. We trained the ANN and SVM models onto the real data input without and with features selection to show the importance of this process. With the use of PCA as a reducing technique of the input variables, the obtained results showed clearly excellent performances with a slight improvement in terms of recognition rates. Indeed, the use of PCA features to select the number of input parameters is decreased, indicating a small number of sensors. That means, the PCA technique reflected an economic impact on the overall cost of the monitoring system. With this reduction operation, we can say that the storage of data in memory can be considered advantageous for the enrichment of the database collected by the expert system. In general, since it depends on several climatic and geographical parameters, continuous enrichment of the database is necessary. The computational time assigned to the training dataset for PCA-SVM model is extremely fast, which confers the advantage of integration in a dynamic multi-sensor monitoring system. The ANN model, however, suffers from a handicap related to its apparent noise sensitivity. However, due to its best robustness, this limitation is eliminated by the SVM model in particular. The principle of the optimization algorithm is another significant feature of SVM relative to ANN; the solution has a global optimum, eliminating the use of gradient-based search techniques that can cover a local optimum. However, there are no clear rules for fixing the number of neurons and hidden layers in the ANN technique, which is a major concern for obtaining an optimal architecture. The results obtained showed that using a one-against-all multi-class approach, SVM can achieve high performance in classification in terms of recognition rate, training time, and robustness. Overall, the present application demonstrates SVM's promising results. The use of multiple sources of knowledge is highly successful in enhancing classification accuracy. For this analysis, Decision Template (DT) has high potential. The precision of the system decision can be enhanced by using the Decision Template fusion method. With this method, the power of each classifier to achieve a more powerful classifier was combined. The Decision Template has shown more success than each classifier. Furthermore, the recognition rate of 99.24% was obtained. Hence, the use of various sensors and multiple classifiers and subsequently combining them is strongly recommended for classification in water quality monitoring. The use of the DT algorithm for final decision-making improves the system's efficiency and the accuracy of the approved classification process. When chemical parameters are unable to be continuously measured, the precision of the system decision can be increased by using new input parameters or soft sensors. It is also important to assess the robustness of the proposed solution concerning noise. It should be noted that the domain's sensitivity and unexpected threats need greater efforts to optimize the system's immunity and to make more changes to reduce the risks to public health.

V. ACKNOWLEDGEMENT

This work is supported by the General Directorate of Scientific Research and Technological Development, Ministry of Higher Education and Scientific Research of Algeria. The authors thank the editor and reviewer for many helpful and constructive suggestions and remarks about an earlier draft of this article which improved the paper quality considerably. The authors would like to thank the engineers from Tilesdit dam direction for their supports and providing the facilities for this investigation and free access to databases and valuable guidance for the field sampling.

Ethics declarations

Authors Contributions

All authors contributed to the study conception and design through Material preparation, data collection, and analysis. All authors read and approved the final manuscript.

Mohamed LADJAL : Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data Curation and collection, Writing - Original Draft, Writing - Review & Editing, Visualization

Mohamed BOUAMAR: Supervision, Project administration, Conceptualization, Formal analysis, Writing - Review & Editing, Visualization.
Youcef BRIK: Conceptualization, Software, Formal analysis, Investigation, Visualization
Mohamed DJERIOUI : Software, Formal analysis, Investigation.

Availability of data and materials

All data generated or analysed during this study are included in this published article, they are available from the corresponding author on reasonable request. For the purposes of privacy, all used data are confidential and cannot be made available.

Compliance with Ethical statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

This article does not contain any studies with any participants performed by any of the authors.

Informed consent and Funding

None.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent to Participate and Publish: none.

REFERENCES

- Abedi, M., Norouzi, G. H., & Bahroudi, A. (2012). Support vector machine for multi-classification of mineral prospectivity areas. *Computers & Geosciences*, 46, 272-283. <https://doi.org/10.1016/j.cageo.2011.12.014>
- Areerachakul, S., & Sanguansintukul, S. (2010). Classification and Regression Trees and MLP Neural Network to Classify Water Quality of Canals in Bangkok, Thailand. *International Journal of Intelligent Computing Research*, 1(2), 30-37. <https://doi.org/10.20533/ijicr.2042.4655.2010.0004>
- Ayeni, O. (2013). Interpretation of surface water quality using principal components analysis and cluster analysis. *Journal of Geography and Regional Planning*, 6(4), 132-141. <https://doi.org/10.5897/jgrp12.087>
- Bae, M. H., Wu, T., & Pan, R. (2010). Mix-ratio sampling : Classifying multiclass imbalanced mouse brain images using support vector machine. *Expert Systems with Applications*, 37(7), 4955-4965. <https://doi.org/10.1016/j.eswa.2009.12.018>
- Bhardwaj, V., Singh, D. S., & Singh, A. K. (2010). Water quality of the Chhoti Gandak River using principal component analysis, Ganga Plain, India. *Journal of Earth System Science*, 119(1), 117-127. <https://doi.org/10.1007/s12040-010-0007-8>
- Bigdeli, B., Samadzadegan, F., & Reinartz, P. (2015). Fusion of hyperspectral and LIDAR data using decision template-based fuzzy multiple classifier system. *International Journal of Applied Earth Observation and Geoinformation*, 38, 309-320. <https://doi.org/10.1016/j.jag.2015.01.017>
- Bouamar, M., Ladjal, M. (2012). Performance Evaluation Of Three Pattern Classification Techniques Used For Water Quality Monitoring. *International Journal Of Computational Intelligence And Applications*, 11(02), 1250013. <https://doi.org/10.1142/s1469026812500137>
- Bouamar M., Ladjal M., (2008). A comparative study of RBF neural network and SVM classification techniques performed on real data for drinking water quality. *Proceedings of Systems, Signals and Devices, 5th Int. Multi-Conf. Systems, Signals & Devices IEEE SSD'08*. Amman, Jordan, 20-23 July 2008, pp. 201.
- Burges C. J.C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*. Vol. 2. 121-167.
- Cao, L., Chua, K., Chong, W., Lee, H., & Gu, Q. (2003). A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. *Neurocomputing*, 55(1-2), 321-336. [https://doi.org/10.1016/s0925-2312\(03\)00433-8](https://doi.org/10.1016/s0925-2312(03)00433-8)
- Chen, W., Zhang, S. W., Cheng, Y. M., & Pan, Q. (2010a). Prediction of protein-protein interaction types using the decision templates based on multiple classifier fusion. *Mathematical and Computer Modelling*, 52(11-12), 2075-2084. <https://doi.org/10.1016/j.mcm.2010.01.025>
- Chen, W., Zhang, S. W., Cheng, Y. M., & Pan, Q. (2010b). Prediction of protein-protein interaction types using the decision templates based on multiple classifier fusion. *Mathematical and Computer Modelling*, 52(11-12), 2075-2084. <https://doi.org/10.1016/j.mcm.2010.01.025>

- Décret exécutif N° 11-125 du 22 Mars 2011 relatif à la qualité de l'eau de consommation humaine, *Journal officiel de la République Algérienne* N° 18.
- Deng, S., Lin, S. Y., & Chang, W. L. (2011). Application of multiclass support vector machines for fault diagnosis of field air defense gun. *Expert Systems with Applications*, 38(5), 6007-6013. <https://doi.org/10.1016/j.eswa.2010.11.020>
- Déniz, O., Castrillón, M., & Hernández, M. (2003). Face recognition using independent component analysis and support vector machines. *Pattern Recognition Letters*, 24(13), 2153-2157. [https://doi.org/10.1016/s0167-8655\(03\)00081-3](https://doi.org/10.1016/s0167-8655(03)00081-3)
- Djerioui, M., Bouamar, M., Ladjal, M., & Zerguine, A. (2018). Chlorine Soft Sensor Based on Extreme Learning Machine for Water Quality Monitoring. *Arabian Journal for Science and Engineering*, 44(3), 2033-2044. <https://doi.org/10.1007/s13369-018-3253-8>
- Haghighi, M. S., Vahedian, A., & Yazdi, H. S. (2011). Extended decision template presentation for combining classifiers. *Expert Systems with Applications*, 38(7), 8414-8418. <https://doi.org/10.1016/j.eswa.2011.01.036>
- Héctor Ricardo Hernández De León. (2006). Supervision et diagnostic des procédés de production d'eau Potable, Thèse de Doctorat, Institut National des Sciences Appliquées de Toulouse.
- Hend S. Al-Khalifa and Amani A. Al-Ajlan. (2010). Automatic readability measurements of the arabic text: an exploratory study. *The Arabian Journal for Science and Engineering*. Vol. 35 N° 2C. 103-124.
- Hornig, M. H. (2009). Multi-class support vector machine for classification of the ultrasonic images of supraspinatus. *Expert Systems with Applications*, 36(4), 8124-8133. <https://doi.org/10.1016/j.eswa.2008.10.030>
- Jin J L, Liu L, Ding J, Fu Q. (2003). Logistic curve model of groundwater quality evaluation. *Environmental Pollution and Control*, Vol. 25 N° 1. 46-48.
- Jun-Ki Min, Sung-Bae Cho. (2007). Multiple Classifier Fusion Using k-Nearest Localized Templates. *IDEAL 2007 International Conference on Intelligent Data Engineering and Automated Learning*. pp 447-456. https://doi.org/10.1007/978-3-540-77226-2_46
- Jolliffe, I. T. (2010). *Principal Component Analysis (Springer Series in Statistics)* (Softcover reprint of the original 2nd ed. 2002 éd.). Springer.
- Kumar, R., Jayaraman, V., & Kulkarni, B. (2005). An SVM classifier incorporating simultaneous noise reduction and feature selection : illustrative case examples. *Pattern Recognition*, 38(1), 41-49. <https://doi.org/10.1016/j.patcog.2004.06.002>
- Kuncheva, L. (2014). *Combining Pattern Classifiers : Methods and Algorithms* (2^e éd.). Wiley.
- Kuncheva, L. I., Bezdek, J. C., & Duin, R. P. (2001). Decision templates for multiple classifier fusion : an experimental comparison. *Pattern Recognition*, 34(2), 299-314. [https://doi.org/10.1016/s0031-3203\(99\)00223-x](https://doi.org/10.1016/s0031-3203(99)00223-x)
- Ladjal Mohamed, Bouamar Mohamed, Djerioui Mohamed, Brik Youcef (2016). Performance Evaluation Of ANN And SVM Multiclass Models For Intelligent Water Quality Classification Using Dempster-Shafer Theory. *International Conference on Electrical and Information Technologies (ICEIT)*. Tangiers, Morocco. DOI: 10.1109/EITech.2016.7519588
- Ladjal Mohamed, OUALI Mohammed Assam, Djerioui Mohamed (2020). Optimization Of SVM Parameters With Hybrid PCA-PSO Methods For Water Quality Monitoring. *2020 International Conference on Electrical Engineering (ICEE)*. Istanbul, Turkey. DOI: 10.1109/ICEE49691.2020.9249881.
- Zhang, W., Gao, H., & Sun, H. (2018). Application and Analysis of Bayesian Method and Grey Relational Analysis in Marine Water Quality Evaluation. *IOP Conference Series : Earth and Environmental Science*, 182, 012007. <https://doi.org/10.1088/1755-1315/182/1/012007>
- Liao, Y., Xu, J., & Wang, W. (2011). A Method of Water Quality Assessment Based on Biomonitoring and Multiclass Support Vector Machine. *Procedia Environmental Sciences*, 10, 451-457. <https://doi.org/10.1016/j.proenv.2011.09.074>
- Liao, Y., Xu, J. Y., & Wang, Z. W. (2012). Application of biomonitoring and support vector machine in water quality assessment. *Journal of Zhejiang University SCIENCE B*, 13(4), 327-334. <https://doi.org/10.1631/jzus.b1100031>
- Liu, D., & Zou, Z. (2012). Water quality evaluation based on improved fuzzy matter-element method. *Journal of Environmental Sciences*, 24(7), 1210-1216. [https://doi.org/10.1016/s1001-0742\(11\)60938-8](https://doi.org/10.1016/s1001-0742(11)60938-8)
- Mookiah, M. R. K., Rajendra Acharya, U., Lim, C. M., Petznick, A., & Suri, J. S. (2012). Data mining technique for automated diagnosis of glaucoma using higher order spectra and wavelet energy features. *Knowledge-Based Systems*, 33, 73-82. <https://doi.org/10.1016/j.knsys.2012.02.010>
- Msiza, I. S., Nelwamondo, F. V., & Marwala, T. (2008). Water Demand Prediction using Artificial Neural Networks and Support Vector Regression. *Journal of Computers*, 3(11), 1-8. <https://doi.org/10.4304/jcp.3.11.1-8>
- Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine*, 6(3), 21-45. <https://doi.org/10.1109/mcas.2006.1688199>
- Ruta, D., & Gabrys, B. (2005). Classifier selection for majority voting. *Information Fusion*, 6(1), 63-81. <https://doi.org/10.1016/j.inffus.2004.04.008>
- Schölkopf, B., & Smola, A. J. (2018). *Learning with Kernels : Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning series)* (Reprint éd.). The MIT Press.
- Semmlow, J. L. (2004). *Biosignal and Medical Image Processing (Signal Processing and Communications, 22)* (1^{re} éd.). CRC Press.
- Singh, K. P., Basant, N., & Gupta, S. (2011). Support vector machines in water quality management. *Analytica Chimica Acta*, 703(2), 152-162. <https://doi.org/10.1016/j.aca.2011.07.027>
- Übeyli, E. D. (2009). Analysis of electrocardiographic changes in partial epileptic patients by combining eigenvector methods and support vector machines. *Expert Systems*, 26(3), 249-259. <https://doi.org/10.1111/j.1468-0394.2009.00478.x>
- Vapnik, V. (2000). *[(The Nature of Statistical Learning Theory)] [Author : Vladimir Vapnik] [Dec-2000]*. Springer-Verlag New York Inc.
- Wang, L J, Zou Z H. (2008). Application of improved attributes recognition method in water quality assessment. *Chinese Journal of Environmental Engineering*, Vol. 2 N° 4. 553-556.
- Wang, Q., Li, S., & Li, R. (2019). Evaluating water resource sustainability in Beijing, China : Combining PSR model and matter-element extension method. *Journal of Cleaner Production*, 206, 171-179. <https://doi.org/10.1016/j.jclepro.2018.09.057>
- Wang, Y., Wang, P., Bai, Y., Tian, Z., Li, J., Shao, X., Mustavich, L. F., & Li, B. L. (2013). Assessment of surface water quality via multivariate statistical techniques : A case study of the Songhua River Harbin region, China. *Journal of Hydro-environment Research*, 7(1), 30-40. <https://doi.org/10.1016/j.jher.2012.10.003>
- WANG, Z. Y., & YANG, Y. F. (2010). Multi-class cluster support vector machines. *Journal of Computer Applications*, 30(1), 143-145. <https://doi.org/10.3724/sp.j.1087.2010.00143>
- Widodo, A., & Yang, B. S. (2007). Application of nonlinear feature extraction and support vector machines for fault diagnosis of induction motors. *Expert Systems with Applications*, 33(1), 241-250. <https://doi.org/10.1016/j.eswa.2006.04.020>
- Widodo, A., Yang, B. S., & Han, T. (2007). Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors. *Expert Systems with Applications*, 32(2), 299-312. <https://doi.org/10.1016/j.eswa.2005.11.031>
- Wu, C. H., Tzeng, G. H., Goo, Y. J., & Fang, W. C. (2007). A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy. *Expert Systems with Applications*, 32(2), 397-408. <https://doi.org/10.1016/j.eswa.2005.12.008>
- Yan, H., Zou, Z., & Wang, H. (2010). Adaptive neuro fuzzy inference system for classification of water quality status. *Journal of Environmental Sciences*, 22(12), 1891-1896. [https://doi.org/10.1016/s1001-0742\(09\)60335-1](https://doi.org/10.1016/s1001-0742(09)60335-1)

- YANG, B. S., HAN, T., & YIN, Z. J. (2006). Fault Diagnosis System of Induction Motors Using Feature Extraction, Feature Selection and Classification Algorithm. *JSME International Journal Series C*, 49(3), 734-741. <https://doi.org/10.1299/jsmec.49.734>
- Zhang, S. W., Liu, Y. F., Yu, Y., Zhang, T. H., & Fan, X. N. (2014). MSLoc-DT : A new method for predicting the protein subcellular location of multispecies based on decision templates. *Analytical Biochemistry*, 449, 164-171. <https://doi.org/10.1016/j.ab.2013.12.013>
- Zhang, W., Gao, H., & Sun, H. (2018). Application and Analysis of Bayesian Method and Grey Relational Analysis in Marine Water Quality Evaluation. *IOP Conference Series : Earth and Environmental Science*, 182, 012007. <https://doi.org/10.1088/1755-1315/182/1/012007>
- ZHOU, W., & WU, B. (2008). Assessment of soil erosion and sediment delivery ratio using remote sensing and GIS : a case study of upstream Chaobaihe River catchment, north China. *International Journal of Sediment Research*, 23(2), 167-173. [https://doi.org/10.1016/s1001-6279\(08\)60016-5](https://doi.org/10.1016/s1001-6279(08)60016-5)
- ZOU, Z. H., YUN, Y., & SUN, J. N. (2006). Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment. *Journal of Environmental Sciences*, 18(5), 1020-1023. [https://doi.org/10.1016/s1001-0742\(06\)60032-6](https://doi.org/10.1016/s1001-0742(06)60032-6)

699 **Highlights**

- 700 • New intelligent water quality classification is performed using data combination fusion and features selection.
- 701 • ANN and SVM methods have been proposed for water quality classification status.
- 702 • The final decision is performed using Decision Templates rule combination based on probabilistic output from
703 both the two classifiers.
- 704 • Real database from Tilesdit dam (Algeria) are used for evaluation.
- 705 • Results demonstrate the proposed approach has a superior performance in water quality monitoring