KSOM and MLP neural networks for on-line estimating the efficiency of an activated sludge process

Stéphane Grieu *, Frédérik Thiery, Adama Traoré, Tien Phong Nguyen, Mathieu Barreau, Monique Polit

Laboratoire de Physique Appliquée et d'Automatique, Université de Perpignan, 52 Avenue Paul Alduy, 66860 Perpignan, France

Abstract

This work is devoted to the study of the Saint Cyprien (south of France) activated sludge WasteWater Treatment Plant (WWTP) process and to the on-line estimation of chemical parameters (influent and effluent chemical oxygen demand, ammonia and suspended solids) not easily measurable on-line. Their knowledge makes it possible to estimate the process efficiency and to provide reliable information for the plant monitoring.

A tool including Kohonen's self-organizing maps and a multi-level perceptron is used. The Kohonen's self-organizing maps neural network is applied to analyze the multi-dimensional Saint Cyprien process data and to diagnose the inter-relationship of the process variables in an activated sludge WWTP. The multi-level perceptron is used as estimation tool.

The obtained results are satisfactory. The information provided by the estimation procedure is sufficiently reliable and precise to be exploitable by operators in charge of the plant monitoring and maintenance. It allows understanding how the system is evolving. The whole procedure (Kohonen's self-organizing maps and multi-level perceptron) uses tools which proved to be efficient and complementary.

1. Introduction

Since the implementation of European Directive 91/271/CEE, nutrient removal of WasteWater Treatment Plants (WWTPs) has been progressively adopted by existing and new facilities. Nevertheless, when looking at nutrient removal, and specially nitrogen removal, on-line knowledge of influent quality could improve plant efficiency by implementing control schemes [1]. It is possible to implement on-line ammonia or COD sensors, but all of these sensors are based on indirect measurements. The mechanism knowledge involved in wastewater treatment, related to carbon and nitrogen removal, has been widely studied since the presentation of ASM1 in 1986 [2]. From the actual process knowledge, it is possible to clearly identify variables affecting plant performances and the way the process is affected by variations in influent quality and quantity.

In France, the dumping of effluents is subjected to the existing legislation related to classified plants. Thus, the law no. 76-663 of July 19, 1976 fixes quality standards for effluent water. Limiting values of parameters for effluent water were fixed by the decree of February 2, 1998, but even more strict limiting values can be decided by means of specific decrees. The dumping of effluents must also respects the objectives established by the law of January 3, 1992 in order to preserve a balanced management of water resources. WWTPs operations must be monitored by means of a rigorous follow-up of the various treatments efficiency. That is why the law requires the implementation of a controlled self-monitoring [3].

First, the objective of this work is to highlight correlations between physicochemical parameters characterizing the activated sludge process carried out at the Saint Cyprien wastewater treatment plant (south of France), then to analyze the evolution of such a system using a Kohonen's self-organizing maps neural network. From a data set and an unsupervised learning algorithm generating a taxonomy, using Kohonen's self-organizing maps (KSOM) makes it possible to simply visualize in two dimensions all the interactions existing between variables. Thus, this tool is an invaluable help for operators in charge of a WWTP monitoring by its capability of data analysis, information simplification and visual representation.

* Corresponding author. Tel.: +33 4 68 66 22 40; fax: +33 4 68 66 22 87.
  E-mail address: grieu@univ-perp.fr (S. Grieu).

Kohonen's self-organizing maps are also a very effective clustering tool, by their ability to search for common properties in a data file [4]. Next, the results obtained thanks to self-organizing maps are used to carry out an estimation, using a multi-level perceptron neural network, of parameters describing the good behavior of the Saint Cyprien organic matter and nitrogen removal biological process. These parameters, not easily measurable on-line, are the chemical oxygen demand (COD), the ammonia (NH_4) and the suspended solids (SS) [5].

The correlations highlighted by the Kohonen's self-organizing maps neural network made it possible to effectively choose the variables used, for their predictive character, as multi-layer perceptron inputs. The clustering allows to carry out a split training and makes it possible to facilitate, for the MLP network, the comprehension of the various involved phenomena. It allows to minimize the frequent risks of overfitting when neural networks are used. Neural networks, a statistical tool for data analysis [6], could be applied to establish a relation between variables describing a process state and different measured quantities. This relation depends, in a not always obvious way, on the predictive variables used. The main characteristic of neural networks is their capability to automatically establish relations between variables by means of a mechanism called training or learning [7]. Neural networks are designed for a specific application and, after a training phase, are able to generate an estimation, applying the relationship developed during the learning period.

2. Saint Cyprien WWTP

2.1. Plant presentation

The experimental data were provided by the Saint Cyprien WWTP employees. Three towns are connected to this "activated sludge" plant able to treat the wastewater of 80,000 inhabitants: Saint Cyprien (8653 inhabitants), Alenya (2339 inhabitants) and Latour Bas Elne (1711 inhabitants). This WWTP is a weak charges biological plant: denitrification and biological dephosphatation with physical and chemical complement. It has the particularity of being able to adapt operation to the seasonal population variations. The town of Saint Cyprien is a seaside resort and its population increases considerably during the summer season. Therefore, the plant consists of parallel treatment modules characterizing its adaptability. At the end of the wastewater treatment sequence, a sludge composting unit allows the sludge re-use in agricultural applications. The plant is divided into a water line and a sludge line. The water line consists of the primary treatment, the biological treatment tanks and the clarifiers. The sludge line consists of the sludge dehydration and composting, the treated water filtration and finally the UV disinfection.

2.2. Water line

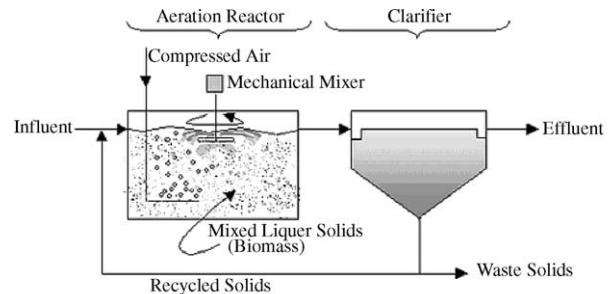- *Primary treatment*: Cleaning by tilted rake and grease and oil removal.

Fig. 1. Schematic diagram of the activated sludge process.

- *Biological treatment tanks*: Four parallel tanks that can be used individually or jointly according to the period of the year and the quantity of water to be treated (total volume of $15,600 \, m^3$).
- *Clarifiers*: Two individually usable parallel clarifiers (total volume of $3200 \, m^3$).

2.3. Sludge line and complementary treatments

- *Sludge dehydration*: Three presses placed in a closed and ventilated room (increase by 16% for dryness).
- *Sludge composting*: Compost used in agriculture (dehydrated sludge and wood plates mixture).
- *Treated water filtration*: Two sand filters. Total capacity: $150 \, m^3/h$.
- *UV disinfection*: For bacterial decontamination with two possible files of, respectively, 850 and $150 \, m^3/h$ throwing out water into the "canal d'Elne".

2.4. The activated sludge process

In general, the activated sludge process is a continuous or semi continuous (fill and draw) aerobic method for biological wastewater treatment, including carbonaceous oxidation and nitrification [8]. This process is based on the aeration of wastewater with flocculating biological growth, followed by separation of treated wastewater from this growth. Part of this growth is then wasted, and the remainder is returned to the system. Usually, the separation of the growth from the treated wastewater is performed by settling (gravity separation) but it may also be done by flotation and other methods (Fig. 1). The biological component of the activated sludge system is comprised of microorganisms. Bacteria, fungi, protozoa and rotifers constitute the biological component or biological mass, of activated sludge. In addition, some metazoa, such as nematode worms, may be present. However, the constant agitation in the aeration tanks and sludge recirculation are deterrents to the growth of higher organisms. The microorganisms that are of greatest numerical importance in activated sludge are bacteria, which occur as microscopic individuals from one micron in size to visible aggregations or colonies of individuals. Some bacteria are strict aerobes (they can only live in the presence of oxygen), whereas others are anaerobes (they are active only in the absence of oxygen).
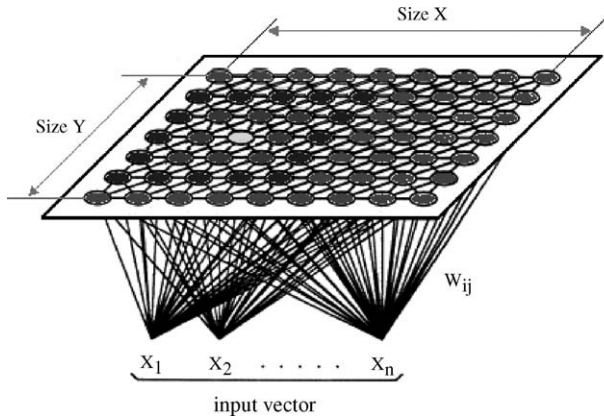
Fig. 2. Kohonen network structure.

3. Materials and methods

3.1. Kohonen's self-organizing maps

Kohonen's self-organizing maps are very popular in the field of data processing. They belong to the category of unsupervised neural networks allowing pattern classification without preliminary knowledge. The pattern classification application can be regarded as a specific case of the function approximation. The mapping is done from the input space to a finite number of output classes.

Kohonen's self-organizing maps can be used both to highlight groups of elements with similar characteristics and for projecting the data non-linearity onto a lower dimension display (Fig. 2). They create effectively spatially organized internal representations of input signals features.

The number of groups is unknown and the network has to identify elements common characteristics. The network is adjusted according to statistical regularities and builds intern representations in order to create groups [9]. It is an unsupervised learning. Thus, the Kohonen analysis makes it possible to represent data, preserving a topology. Close data (in the input space) will have close representations in the output space and will be placed in the same class or in close classes. Feature maps are trained on input vectors using rules that aim to iteratively arrange the neurons in the maps so as to preserve input data characteristics.

Kohonen networks consist of an input layer and an output layer, called competitive layer or Kohonen layer.

All the input neurons are connected to each neuron of the output layer (balanced connections).

Every individual to be classified is represented by a multidimensional vector (input vector). To each individual is affected an output neuron which represents the center of the class. The neurons of the output layer compete among themselves: for each neuron corresponds an output space, generally a two dimensions space defined by a grid.

Only the best gains, it is the "winner takes all" method. The training is described as competitive. Kohonen's self-organizing maps are connection networks with a fixed dimensionality allowing to simply carry out a projection of a data space on a 2D grid.

3.2. Kohonen training algorithm

The Kohonen training algorithm is a simple procedure which consists of randomly selecting a training pattern, determining the winning node, updating the weights of all nodes within the winning neighborhood and modifying the training parameters [9]. In determining the winning unit, a certain distance metric is used, and the node, which registers the least distance relative to the current training pattern is referred to as the winning unit (the best matching unit). Changes in the weights will subsequently involve only nodes in the region surrounding this winning unit. The distance metric to be used in the algorithm is left unspecified, although the most common metrics used are the Euclidean distance, the Manhattan distance and the cosine of the angle of the current input vector and each of the nodes weight vectors. The output of each neuron can be computed according to:

$$y_i = \sum_{j=1}^{n} m_{i,j} x_j \tag{1}$$

where n is the number of inputs, $m_{i,j}$ a synaptic weight and x_j is the jth component of the input vector X. The learning algorithm can be described as follows:

1. Determine both map topology and number of neurons. The number of neurons is selected as big as possible with the neighborhood size controlling the smoothness and generalization of the mapping.
2. Initialize the weight vectors with small random values.
3. Select an input pattern X at random from the data set.
4. Find the best matching neurons m_c with an appropriate distance metric, usually the Euclidian distance. Also find the set of neurons that composes the neighborhood N_c about m_c. The best matching neurons m_c is the closest to the input vector $X(t)$ according to:

$$||X - m_c|| = \min_i \{||X - m_i||\} \tag{2}$$

where $||\cdot||$ is the distance measure. The best matching unit, denoted as c, is the unit whose weight vector has the greatest similarity with the input pattern X.
5. Modify the weights of the neurons within the winning neighborhood using the following learning rule:

$$m_i(t + 1) = m_i(t) + \alpha(t)[X(t) - m_i(t)] \quad \forall i \in N_c(t) \tag{3}$$

$$m_i(t + 1) = m_i(t) \quad \forall i \notin N_c(t) \tag{4}$$

where $\alpha(t)$ is a suitable decreasing function of time, $0 \leq \alpha(t) \leq 1$. $\alpha(t)$ is called "gain" and it decreases as the number of training cycles increases. It controls the amount of weight change that will take place per training cycle.

The neighborhood size is also decreasing in value as the number of training cycles increases. This winning neighborhood is the set of nodes surrounding the winning unit which will undergo weight update. Outside nodes in the map will not change their weights. Note that since the winning unit

would depend on the training pattern presented, the winning neighborhood moves around the map throughout the training phase.

The neighborhood size is typically denoted by a "radius", which is the number of "hops" from one node to another. In a rectangular map, the radius between a given node and all its eight direct neighbor nodes is 1. Its distance to the 16 next nearest nodes has a radius of 1 and so on. There are no hard and fast rules as to the initial value of the neighborhood size. But setting it initially to be equal to the size of longest map dimension (height or width of map in terms of number of nodes) is adequate. This will decrease in value as explained below.

6. Stop, if some optimal number of iteration steps is done, or else continue from step 3.

3.3. Multi-level perceptron

Linear approximation networks are too restrictive to treat such a process and non-linear approximation networks offer much greater capability. Thus, the used neural network is a multilevel perceptron (MLP). This type of neural network is known as a supervised network because it requires a desired output in order to learn. It is based on the classical perceptron network introduced by Rosenblatt [10]. It includes a number of active neurons, the basic building block [11], connected together to form a network. The used multi-level perceptron used consists of one layer of linear output neurons and one hidden layer of non-linear neurons (Fig. 3) [12].

According to our tests, more than one hidden layer proved to be unproductive. It causes slower convergence in the backpropagation learning because intermediate neurons not directly connected to output neurons have very small weight changes and learn very slowly.

Although the relationship between the network performance and its hidden layer size is not well understood, a principle can be used as a guide: the principle of generalization versus conver-

Table 1
Saint Cyprien WWTP process variables

Variables (unit)	Description
Rain (mm/day)	Rain
$Q_{influent}$ (m^3/day)	Influent flow
$SS_{influent}$ (mg/l)	Suspended solids in the influent
$COD_{influent}$ (mg/l)	Chemical oxygen demand in the influent
DO (mg/l)	Dissolved oxygen in the first aerated basin
Q_{air} (m^3/day)	Influent air flow in the first aerated basin
$TKN_{influent}$ (mg/l)	Total Kjeldahl nitrogen in the influent
$NH_{4\,influent}$ (mg/l)	Ammonia in the influent
$NO_{2\,influent}$ (mg/l)	Nitrite in the influent
$NO_{3\,influent}$ (mg/l)	Nitrate in the influent
$TP_{influent}$ (mg/l)	Total phosphorus in the influent
$Q_{effluent}$ (m^3/day)	Effluent flow
$SS_{effluent}$ (mg/l)	Suspended solids in the effluent
$COD_{effluent}$ (mg/l)	Chemical oxygen demand in the effluent
$TKN_{effluent}$ (mg/l)	Total Kjeldahl nitrogen in the effluent
$NH_{4\,effluent}$ (mg/l)	Ammonia in the effluent
$NO_{2\,effluent}$ (mg/l)	Nitrite in the effluent
$NO_{3\,effluent}$ (mg/l)	Nitrate in the effluent
$TP_{effluent}$ (mg/l)	Total phosphorus in the effluent

gence [13]. Thus, the two following user-definable parameters, because of their great influence on the learning phase efficiency, were optimized [5]:

- The number of iterations completed during the training phase.
- The number of neurons placed in the network hidden layer.

The MLP and many others neural networks learn using an algorithm called "backpropagation". With this algorithm, input data are repeatedly presented to the neural network. With each presentation the neural network output is compared to the desired output (the target) and an error is computed. This error is then fed back (backpropagated) to the network and used to adjust the weights such that the error decreases with each iteration and the neural model gets closer and closer to producing the desired
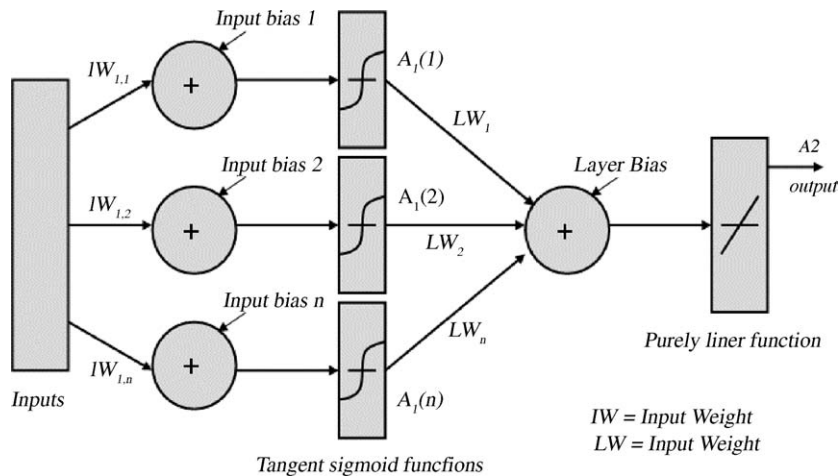
$$A2 = tanh(LW1xA1(1)+LW2xA1(2)+ \ldots +LWnxA1(n)+LB)$$
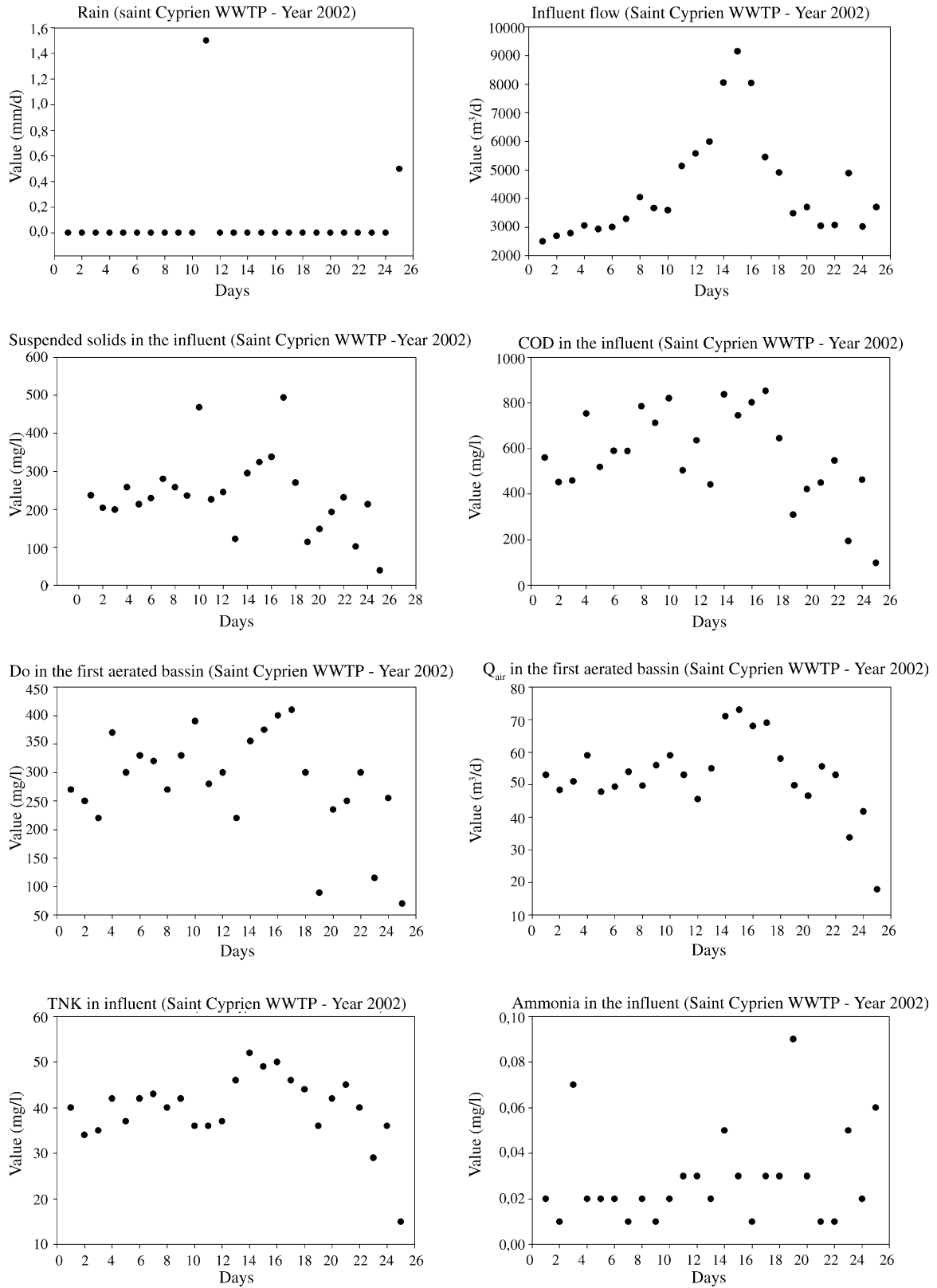
Fig. 3. MLP neural network structure.

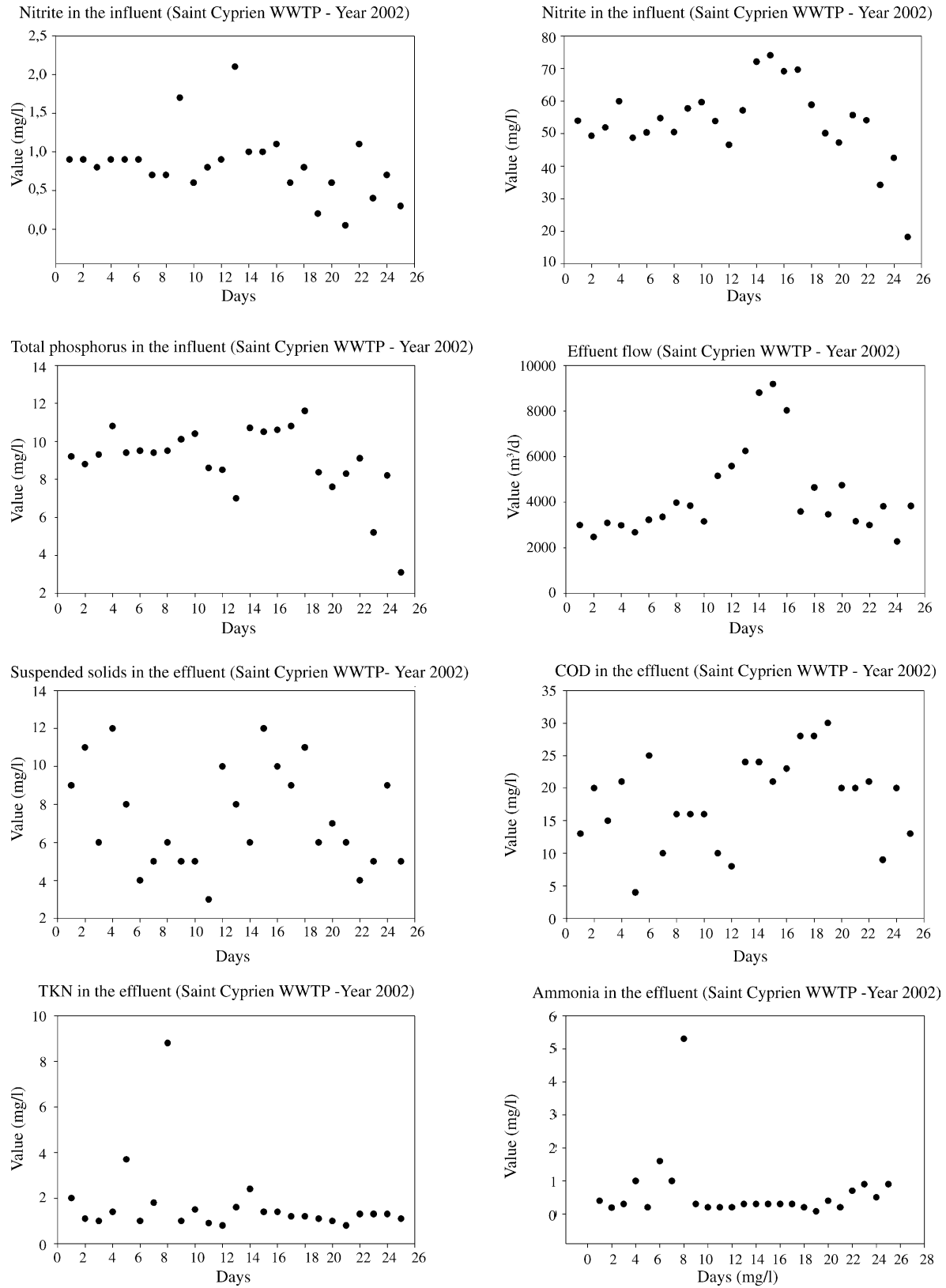Fig. 4. Saint Cyprien WWTP process variables.
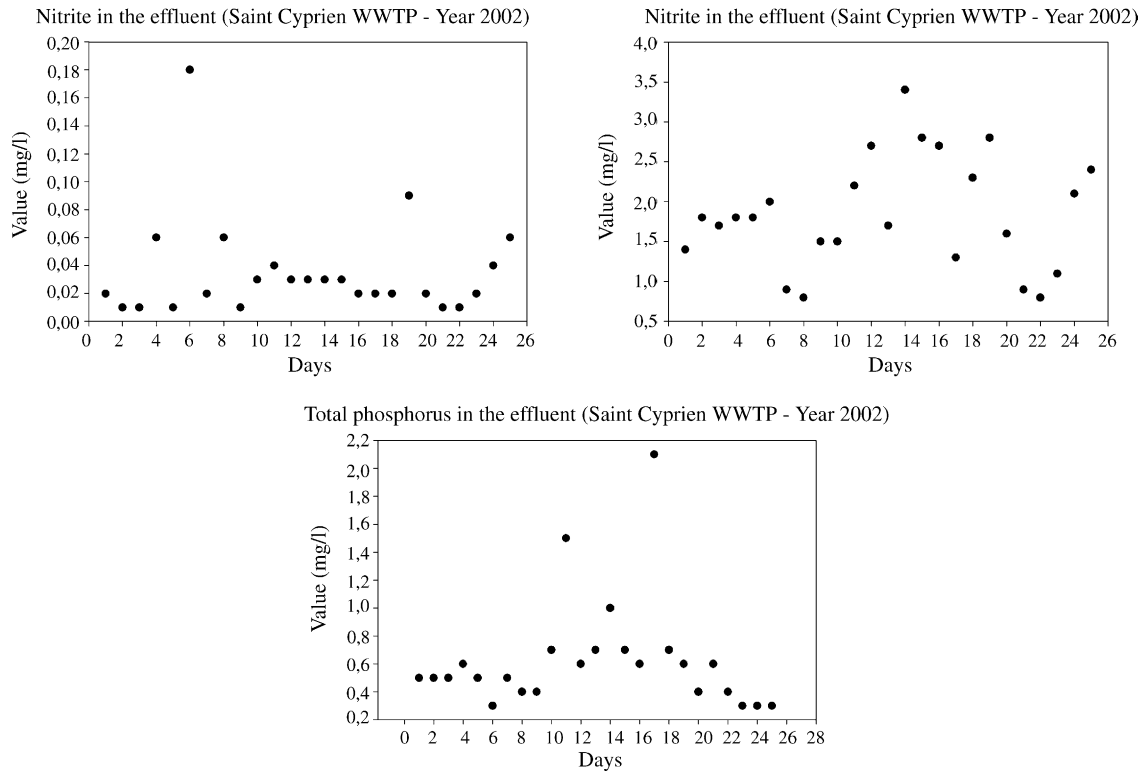
Fig. 4. (*Continued*).

Fig. 4. (*Continued*).

output. The minimization algorithm uses an iterative process and various weight values are explored with the aim of minimizing the quadratic criterion on the sum of squared error made during the learning phase [14].

3.4. Levenberg–Marquardt (LM) algorithm

Several training methods were used, but the Levenberg–Marquardt algorithm proved to be the fastest and the most robust
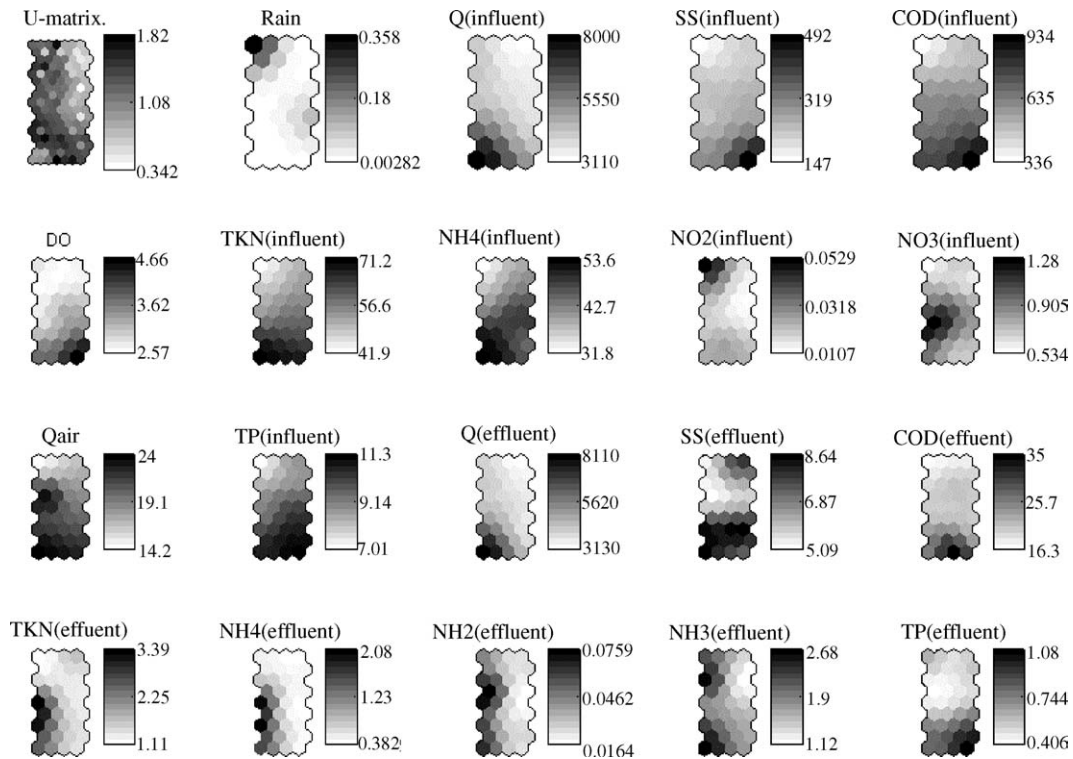

Fig. 5. Component planes for the process variables and U-matrix.

[15]. It is particularly adapted for networks of moderate size and has memory reduction feature for use when the training set is large [16].

Like the quasi-Newton methods, the Levenberg–Marquardt algorithm was designed to approach second-order training speed without having to compute the Hessian matrix.

When the performance function has the form of a sum of squares, then the Hessian matrix can be approximated as:

$$H = J^\mathrm{T} J \qquad (5)$$

The gradient can be computed as:

$$g = J^\mathrm{T} e \qquad (6)$$

where J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases and e is a vector of network errors. The Jacobian matrix can be computed through a standard backpropagation technique that is much less complex than computing the Hessian matrix. The Levenberg–Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton-like update:

$$x_{k+1} = x_k - [J^\mathrm{T} J + \mu I]^{-1} J^\mathrm{T} e \qquad (7)$$

When the scalar μ is zero, this is just Newton's method, using the approximate Hessian matrix. When μ is large, this becomes gradient descent with a small step size. Newton's method is faster and more accurate near an error minimum, so the aim is to shift towards Newton's method as quickly as possible. Thus, μ is decreased after each successful step and is increased only when a tentative step would increase the performance function. In this way, the performance function will always be reduced at each iteration of the algorithm [17].

4. Results and discussions

In the present section, both results of the Saint Cyprien WWTP study using Kohonen's self-organizing maps and COD, NH$_4$ and SS estimation using a multi-level perceptron are described.

4.1. Results of the Saint Cyprien WWTP study using KSOMs

Raw data obtained from the Saint Cyprien WWTP consist of many process variables (Table 1) and describe the plant operation during years 2000, 2001, 2002 and 2003. Fig. 4 presents the data set for all the process variables and year 2002. These variables are only available as daily average data and are measured in different units having different magnitudes. Then, if raw data are fed into the KSOM neural network, variables having a larger magnitude are given unequal importance due to the nature of the weight update procedure. Standardizing the dynamic range of each variable in the input vector is required and will improve the model performance.

In this work, all the simulations are done using MATLAB$^\mathrm{TM}$ and a Self-Organizing Maps Toolbox (http://www.cis.hut.fi/

projects/somtoolbox/). The most significant features of this toolbox are its visualization capabilities [9]. Two different kinds of data presentation tools for the KSOM were used: (1) the component planes and scatter plots for the visualization of the relationships among the process variables and (2) distance matrices and K-means clustering algorithm for the visualization of clusters. The most widely used distance matrix is the unified distance matrix (U-matrix). In this method, a matrix of distances between the weight vectors of map units and their neighbors is calculated. The relative distances between map nodes on the map can be seen. Fig. 5 presents the component planes for the data set. Each hexagon represents one map node and its color tells the value of the component in that node. Nodes representing high values are in black and nodes representing low values are in white. These component planes highlight interesting correlations between variables and allow to choose the multi-level perceptron inputs used during the second phase of this work to estimate the state of the Saint Cyprien activated sludge process.

COD, NH$_4$ and SS are physicochemical parameters, very difficult to measure on-line and characterizing the studied biological process of organic mater and nitrogen removal [18].

Suspended solids are a significant visual quality factor for water. COD represents the oxygen quantity necessary for
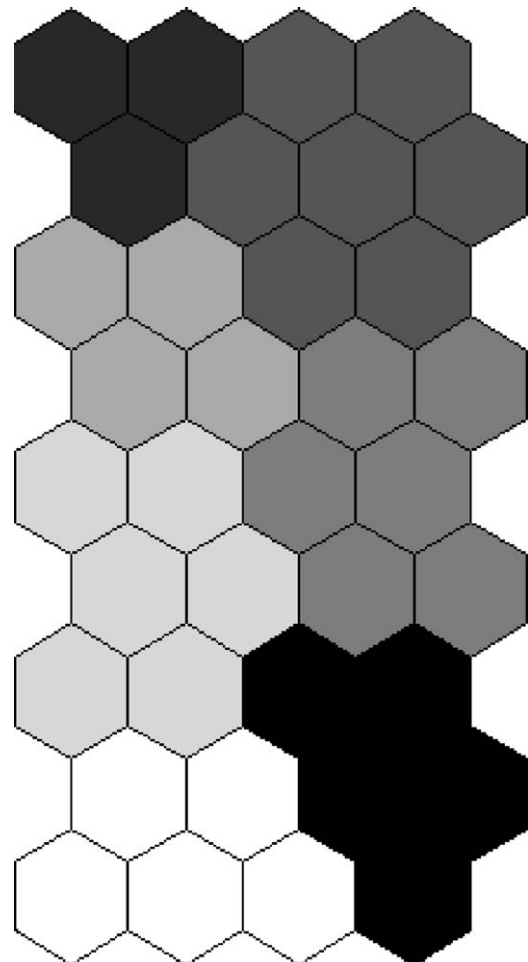

Fig. 6. The seven clusters (operation classes) characterizing the Saint Cyprien WWTP process.

chemically eliminating the organic matter. NH_4 presence in water is a human or industrial contamination sign. The knowledge of these parameters in entry and exit of the plant allows estimating the process efficiency. They were used as targets of the estimation phase.

By comparing the component planes, DO, Q_{in} and Q_{air} seem to be positively correlated with COD (influent), NH_4 (influent) and SS (influent). Similar black and white areas on the various maps allow to highlight these correlations. All these variables present high values in the lower part of their component plane and weak values in the upper part. According to these results and that DO, Q_{in} and Q_{air} are easily on-line measured parameters at the Saint Cyprien WWTP, these three variables were used as MLP inputs to estimate COD (influent), NH_4 (influent) and SS (influent). The estimation of COD (effluent), NH_4 (effluent) and SS (effluent) is identically done but using DO, Q_{in}, Q_{air}, COD (influent), NH_4 (influent) and SS (influent) as MLP inputs.

4.2. Analysis of the cluster structure

One important technique often used in knowledge extraction for process diagnosis is a cluster analysis that generates the classification automatically from the raw process data [19]. Initially, the cluster structure of the Kohonen self-organizing maps is investigated by the visual inspection of the U-matrix. High values (dark shades) on the U-matrix corresponds to a large distance between neighboring nodes, and thus indicate cluster borders on the maps. Lower values (white shade) describe neighboring nodes that are close to each other in the input space.

The U-matrix being difficult to interpret, the K-means clustering algorithm was used to find the optimal number of clusters by partitioning the maps. The Davies-Bouldin clustering index shows that there are seven clusters i.e. seven operation classes characterizing the Saint Cyprien pollution removal process (Fig. 6). The K-means clustering algorithm allows to
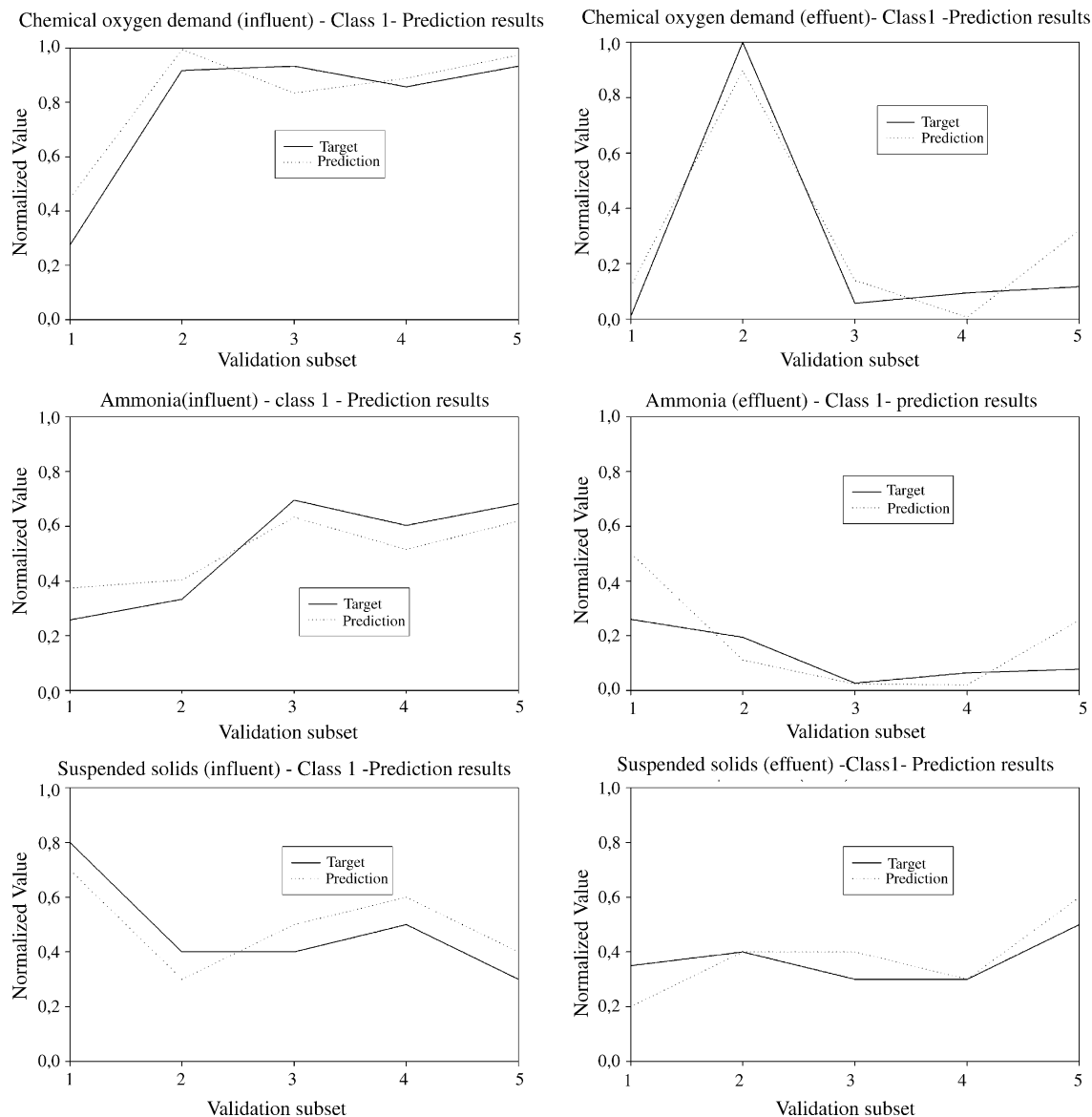

Fig. 7. Estimation results for COD, NH_4 and SS (influent and effluent).

Table 2
Average relative errors for COD, NH$_4$ and SS estimation

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
COD (influent) (%)	9.8	10.5	9.2	11.4	12.3	9.7	13.7
NH$_4$ (influent) (%)	12.8	13.4	12.4	11.9	13.4	14.1	15.2
SS (influent) (%)	14.1	10.9	12.1	13.6	13.3	11.7	14.2
COD (effluent) (%)	13.3	14.9	14.8	15.4	13.1	13.6	14.1
NH$_4$ (effluent) (%)	14	13.8	13.2	14.9	14.2	14.3	16.3
SS (effluent) (%)	15.9	16.2	14.9	15.6	15.1	14.4	17.4

highlight significant differences in operation and to carry out a fractional training for the estimation phase using the MLP. The network learning phase can be conducted separately for each class and is so more precise and effective [5].

The seven operation classes can be described as follows:

- Classes 1 and 2 describe the plant operation during the summer season. The four parallel tanks are used during this period according to the quantity of water to be treated and the pollution to be removed. Class 1 is mainly composed of August days and class 2 is composed of June and July days.
- Classes 3–7 describe the plant operation during the remainder of the year with only two of the parallel tanks in action. Differences between these classes depend on the period of the year and the quantity of water to be treated.

4.3. Estimation results for COD, NH$_4$ and SS

This section presents the results obtained for the COD, NH$_4$ and SS estimation before the pollution removal treatment (influent) and after this pollution removal treatment (effluent), using the seven operation classes defined by K-means clustering. Only graphs for class 1 are presented (Fig. 7) but Table 2 summarizes all the obtained estimation results.

As previously mentioned, DO, Q_{in} and Q_{air} were used as MLP inputs to estimate COD (influent), NH$_4$ (influent) and SS (influent). The estimation of COD (effluent), NH$_4$ (effluent) and SS (effluent) is identically done but using DO, Q_{in}, Q_{air}, COD (influent), NH$_4$ (influent) and SS (influent) as MLP inputs. For each class, 80% of the data set were reserved for the MLP learning phase and 20% for the validation. This distribution proved to be a good choice and is the most frequently used. Fig. 7 presents the estimation results using the validation subset.

According to various tests, for the 7 operation classes, 15 neurons were placed in the MLP hidden layer and 15 iterations were carried out during the learning phase [5].

For each WWTP operation highlighted by K-means clustering, the estimation results can be considered as satisfactory. The obtained information is sufficiently reliable and precise to be exploitable by operators in charge of the Saint Cyprien WWTP monitoring and maintenance.

Indeed, COD, NH$_4$ and SS on-line knowledge at plant entry (influent) and exit (effluent) allows estimating the pollution level before and after biological treatment. It makes it possible to know how the system is evolving and if the pollution removal

process efficiency corresponds (or not) to the fixed objectives for the plant. This information completes the process expert knowledge and available on-line sensors measurements.

As previously mentioned, Table 2 summarizes the obtained results for all the operation classes highlighted by K-means clustering. Average relative errors are satisfactory, ranging between 9.2% (influent COD for class 3) and 17.4% (effluent SS for class 7), according to the strongly non-linear nature of the studied process. However, these results show that the effluent parameters estimation seems to be globally more difficult than the influent one. The biological process efficiency has an important influence on the effluent parameters estimation.

5. Conclusion

This paper presents a tool based on Kohonen's self-organizing maps and a multi-level perceptron, to obtain a COD, NH$_4$ and SS on-line estimation for the Saint Cyprien WWTP influent and effluent. These parameters are not on-line easily measurable and are generally evaluated by laboratory analyses which can be relatively long. Knowledge of these parameters, at WWTP entry and exit, is fundamental, and makes it possible to estimate the correct working of the organic pollution removal process. Their estimation by neural networks was carried out using real data obtained from the Saint Cyprien WWTP.

A preprocessing phase using Kohonen's self-organizing maps completed the neuronal estimation phase in order to improve the results obtained from real data characterizing a strongly non-linear process. Kohonen's self-organizing maps are used to highlight, by means of their visualization capabilities, interesting correlations between variables and significant differences in plant operation (operation classes). These correlations allow choosing suitable MLP inputs to carry out the COD, NH$_4$ and SS estimation. The highlighted operation classes allow to carry out a fractional MLP training which facilitates the phenomena comprehension by the neural network.

The obtained results are satisfactory. The information provided by the estimation procedure is sufficiently reliable and precise to be exploitable by operators in charge of the plant monitoring and maintenance. It allows understanding how the system is evolving. The whole procedure (Kohonen's self-organizing maps and multi-level perceptron) uses tools which proved to be efficient and complementary. Kohonen's self-organizing maps neural network provides useful information by selecting interesting combinations of process variables and extracting knowledge

contained in multi-dimensional data. The multi-level perceptron neural network proved to be an effective estimation and prediction tool.

References

[1] M.M. Hamed, M.G. Khalafallah, E.A. Hassanien, Prediction of wastewater treatment plant performance using artificial neural networks, Environ. Modell. Software 19 (10) (2004) 919–928.

[2] M. Henze, C.P.L. Grady Jr., W. Gujer, G.V.R. Marais, T. Matsuto, Activated Sludge Model No. 1, IAWQ Scientific and Technical Report No. 1, IAWQ, London, 1986.

[3] S. Grieu, A. Traoré, M. Polit, Fault detection in a wastewater treatment plant, in: Eighth IEEE International Conference on Emerging Technologies and Factory Automation, Antibes – Juan les pins, France, 2001, pp. 399–402.

[4] A.K. Jain, R.C. Dubes, Algorithms for clustering data, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[5] S. Grieu, A. Traoré, M. Polit, J. Colprim, Prediction of parameters characterizing the state of a pollution removal process, Eng. Appl. Artif. Intell. 18 (2005) 559–573.

[6] J. Hertz, A. Krogh, R.G. Palmer, Introduction to the Theory of Neural Computation, Computation and Neural Systems Series, Addison-Wesley, New York, NY, 1991.

[7] M. Moller, Efficient training of feed-forward neural networks, Ph.D. Thesis, Computer Science Department, Aarhus University, Denmark, 1993.

[8] W. Eckenfelder, J. Musterman, Activated Sludge Treatment of Industrial Wastewater, Technomic, 1995.

[9] Y.-S.T. Hong, M.R. Rosen, R. Bhamidimarri, Analysis of a municipal wastewater treatment plant using a neural network-based pattern analysis, Water Res. 37 (2003) 1608–1618.

[10] R. Rosenblatt, The perceptron: a perceiving and recognizing automaton, Project PARA, Report 85-460-1, Cornell Aeronautical Lab., 1957.

[11] W.S. McCullogh, W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys. 5 (1943) 115–133.

[12] B. DasGupta, G. Schnitger, The power of approximating: a comparison of activation functions, in: Advances in Neural Networks Information Processing Systems, vol. 5, San Mateo, CA, 1993, pp. 615–622.

[13] C. Schittenkopft, G. Deco, W. Brauer, Two strategies to avoid overfitting in feedforward networks, Neural Networks 10 (1997) 505–516.

[14] K. Hornik, M. Stinchcombe, H. White, Multi-layer feedforward networks are universal approximation, Neural Networks 2 (1989) 359–366.

[15] C. Charalambous, Conjugate gradient algorithm for efficient training of artificial neural networks, IEEE Proc. 139 (1992) 301–310.

[16] M.T. Hagan, M. Menhaj, Training feedforward networks with the Marquardt algorithm, IEEE Trans. Neural Networks 5 (1994) 989–993.

[17] H. Demuth, M. Beale, Neural Network Toolbox, for Use with MATLAB^{TM} User's Guide Version 4, The MathWorks Inc., 2000.

[18] S. Grieu, A. Traoré, M. Polit, J. Colprim, Neural networks for estimating wastewater BOD5 and ammonia from simple on-line operational data, in: IEEE CESA'2003, Lille, France, 2003.

[19] A.K. Jain, M.N. Murty, P.J. Flynn, Data clustering: a review, ACM Computing Surveys 31 (1999) 264–323.