FUZZY MODELING OF MEASUREMENT DATA IN WATER QUALITY ASSESSMENT

C. Temneanu[1], O. Postolache[2], P. Girão[2], J. Pereira[3]

[1]TU "Gh. Asachi", Faculty of Electrical Engineering, D. Mangeron 53, Iasi, Romania
[2]IT/IST, Av. Rovisco Pais, 1049-001 Lisbon, Portugal
[2]ESTS/IPS, Rua do Vale de Chaves, Estefanilha, 2914-508 Setúbal, Portugal

Abstract − The paper discusses a possible solution for water quality assessment based on a virtual system and on intelligent processing structures. Referring to the intelligent processing, a practical approach concerning the fuzzy based solutions in data sensor correction is presented. Several types of fuzzy models are designed, tested and implemented in order to perform the on-line temperature correction of the acquired information from pH, turbidity and conductivity sensors. As the fuzzy architectures used are mentioned the Mamdani and TS by fuzzy clustering techniques. Elements related the virtual system hardware and LabVIEW *temperature fuzzy corrector" are also presented.

Keywords: fuzzy theory, data sensor correction, virtual instruments.

1. INTRODUCTION

To characterise water quality (chemical, physical, and biological characteristics), different parameters classified in different groups are used. One of the most important groups connected to the water quality of rivers, lakes and oceans includes the temperature, the pH, the dissolved oxygen, the turbidity and the conductivity of the water. The measurement of these parameters requires the development of multi-transducer measuring systems. Even if pH, temperature, conductivity and turbidity transducers are commercially available, their integration in a high flexible and accurate system continues to be a very important research field.

Different data sensor processing methods to deal with measurement influence factors are reported in the literature [1][2]. One important data sensor processing solution is intelligent data processing based on the fuzzy set theory. Fuzzy set theory has found applications in major fields such as signal processing, identification, optimisation and control, with a considerable attention not only from the scientific community but also from industry. Many systems are not amenable to conventional modelling approaches due to the lack of precise, formal knowledge about the system, due to strongly non-linear behaviour, due to the high degree of uncertainty or due to time varying characteristics. Fuzzy logic systems are non-linear, so they can be used to work as universal function approximators [3]. In these conditions, different fuzzy blocks were designed and implemented in a virtual measurement system in order to perform an on-line correction of the temperature-undesired effects in the pH, turbidity and conductivity measurement. Experimental results prove the increase of the multi-parameter system 's accuracy at the expenses of a low increase of computational time related with fuzzy data processing.

2. SISTEM PRESENTATION

To acquire the pH, turbidity (TU), conductivity (C) and temperature (T) information an automatic measurement system, was designed tested and implemented. The main hardware components of the system are: the sensors (pH ISI-11, turbidity sensor -WQ770, conductivity - ISI Model OLS50- and temperature ON400), the acquisition unit (PCI-MIO-16E-4) and the control and processing unit expressed by a PC (Fig. 1).
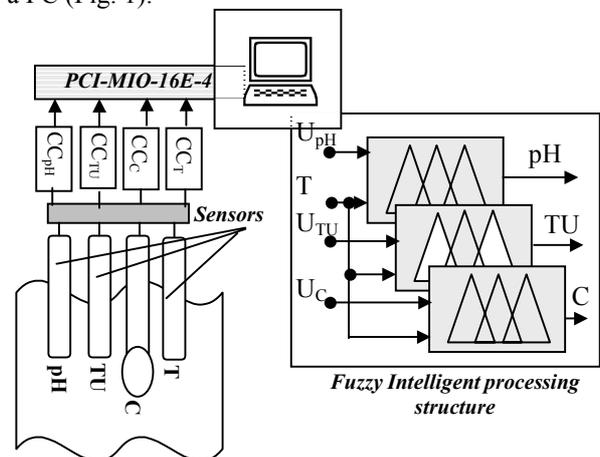

Fig. 1. Water quality monitoring system

Considering the output characteristics of the used sensors, additional conditioning circuits were designed and implemented in order to obtain more accurate voltage values U_{pH}, U_{TU}, U_C and U_T that inform on the monitored quantities.

In the pH case, the electrode polarisation due to current is avoided using a special operational amplifier (MAX406). Additional amplification is performed using an instrumentation amplifier (AD524).

Referring to the turbidity sensor that is characterised by a 4-20mA-output signal, the conditioning circuit is a current-to-voltage converter that outputs a voltage U_{TU} that depends on turbidity.

The conductivity sensor is of the inductive type and requires a special conditioning circuit materialised by a OLM 223 transmitter followed by a current-to-voltage converter.

To obtain U_T values, a resistance to voltage converter is connected to the temperature sensor (NTC thermistor). The temperature information is used both to access water temperature and to increase pH, TU and C characteristics accuracies, which are strongly dependent on temperature. The pH, TU and C corrections are performed using an intelligent processing structure based on fuzzy logic system designed in MATLAB. For on-line processing of the acquired U_{pH}, U_{TU}, U_C and U_T voltages the designed fuzzy system is implemented as a component of the pH, TU, C and T virtual instrument developed in LabVIEW.

3. FUZZY MODELING

It is already well known that the performance of a fuzzy model depends on its parameters: the fuzzification method used, the inference rules that can vary to infinite, the chosen defuzzification technique. In these conditions, most researchers agree with the fact that it is difficult to formulate a mathematical frame with general rules for design fuzzy models [4]. However, it is important to illustrate a few aspects regarding fuzzy model's tune for the reasons that make them viable.

The fuzzy structures taken into account for modelling conductivity, turbidity and pH-sensors are Mamdani and Takagi-Sugeno by fuzzy clustering techniques. Models' performances are evaluated and compared not only between this two kinds of architecture, but also from the point of view of the following structure parameters: membership functions (number, shape and distribution), inference engine and defuzzification method (Mamdani); number of clusters, fuzziness parameter and membership functions generation method (Takagi-Sugeno).

Because they can represent a large class of non-linear systems in a natural linguistic manner, linguistic fuzzy models (Mamdani) represent always the first choice in modelling. Because clustering techniques are among the unsupervised (learning) methods, fuzzy clustering is chosen to approximate the non-linear system (the sensor) by a collection of local linear models [5] of Takagi-Sugeno type in the product space of the regressors and the regressand.

Different scalar parameters have been used to assess the goodness of the obtained models. In the present application, the study was performed illustrating the model sensitiveness in terms of the following performance criteria:
- emax - maximum error;
- rms - root-mean-squared error between measured signal and modelled signal;
- VAF - variance accounted for performance index which is 100% for two equal signals;
- no – number of flops for data processing;
- t - processing time in model test phase.

Data used both for modelling and testing the models was acquired in the water quality monitoring system calibration phase. Thus, the following standards were used: different pH buffer solutions (1÷10 pH units), for PH; formazine standard solutions with values included in 0-100 NTU interval, for turbidity; reference standards with the values included in the 100uS/cm-100mS/cm, for conductivity.

In Fig. 2 are presented the data sets used for training ('o') and testing ('*') the models for pH, conductivity and turbidity, respectively.
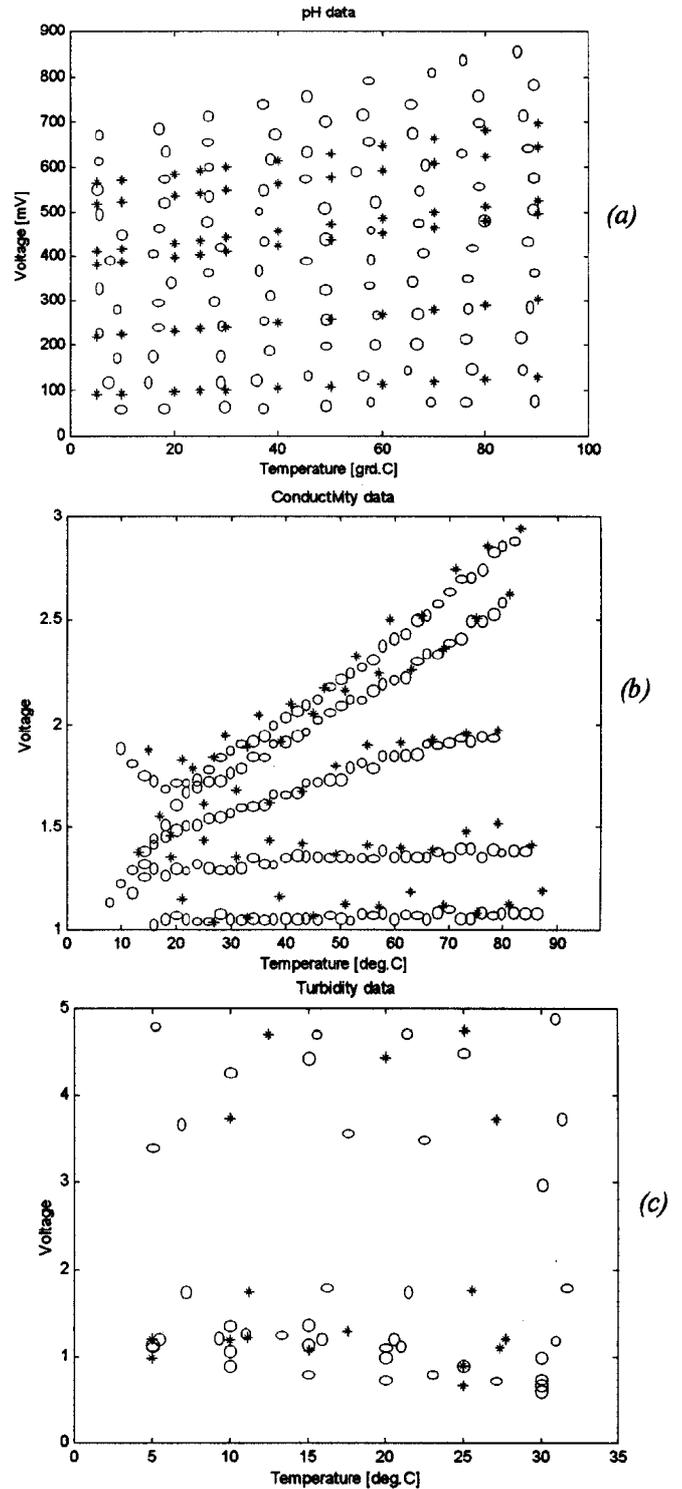

Fig. 2. The data sets for training ('o') and testing ('*') the models

4. RESULTS

4.1. pH fuzzy sensor models

In a Mamdani approach, results in terms of the maximum error are illustrated in Table I and Table II. Different models with a min-max inference mechanism, centroid defuzzification and different membership functions varying in number, shape and distribution were built and tested (Table I). Further, the obtained structure with the best result was tested for different methods of inference and defuzzification (Table II).

TABLE I. The pH-fuzzy Mamdani type models results for different fuzzifications

Description	5 gauss mf, uniform distributed	7 triangular mf, uniform distributed	7 gauss mf, uniform distributed	14 mf, logarithmic distribution for voltage	7 mf for temperature, 11 mf for voltage and pH
emax	19,75	6,85	7,76	6,99	3,97

TABLE II. The pH-fuzzy Mamdani type models results for different inference and defuzzification methods

Description	min-max, centroid	min-max, bisector	min-max, middle of maximum	max-dot, centroid
emax	3,97	8,45	11,15	5,72

From Table II it is clear that other inference and defuzzification methods do not improve the quality of the selected model and, generally, this kind of models does not offer satisfying results in this modelling case.

For the Takagi-Sugeno approach, the dimension of the data collection available for training (see Fig 2(*a*)) allowed us to design models by means of fuzzy clustering with a number of clusters up to 9. For each case, the fuzziness parameter in a range of 1,2÷2,6 was taken into account for evaluation. In Fig.3 are depicted the maximum error for each generated model with product-space membership functions and projected membership functions (*b*).

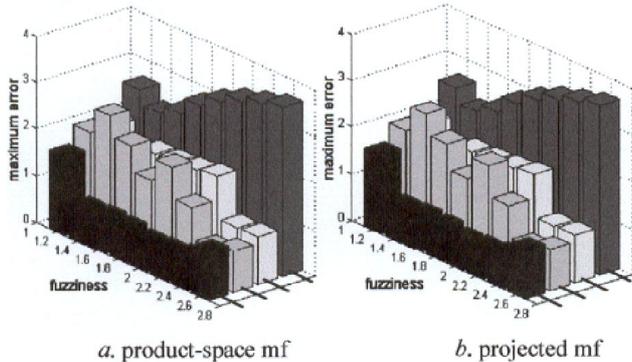*a.* product-space mf   *b.* projected mf

Fig. 3. Maximum error for pH-fuzzy clustering models

Another evaluation criterion in terms of *variance accounted for* (VAF) was computed to assess the goodness of the obtained models. In Fig. 4 is represented this parameter named here fulfillment degree with the meaning of how closed are the model and the original data.

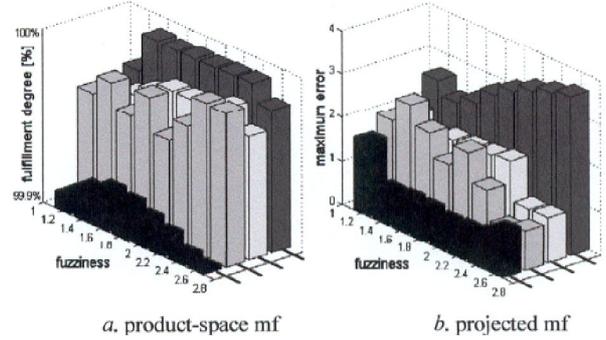*a.* product-space mf   *b.* projected mf

Fig. 4. VAF for pH-fuzzy clustering models

An additional measure of quality was computed, the rms. Table III shows the best solutions indicated by every performance criterion computed.

TABLE III. The best results in terms of emax, VAF and rms

Description	clusters=7 fuzziness=2 prod.-space mf	clusters=7 fuzziness=1,4 prod.-space mf	clusters=9 fuzziness=2,6 projected mf
emax	0,6663	0,7555	0,6916
VAF	99,9922	99,9928	99,9905
rms	0,0252	0,0242	0,0235

The big difference between this two approaches, Mamdani and fuzzy clustering, is due to the shape of the membership functions of the fuzzy clustering model, illustrated in Fig. 5, a shape which is hard to put into a linguistic fuzzy model.
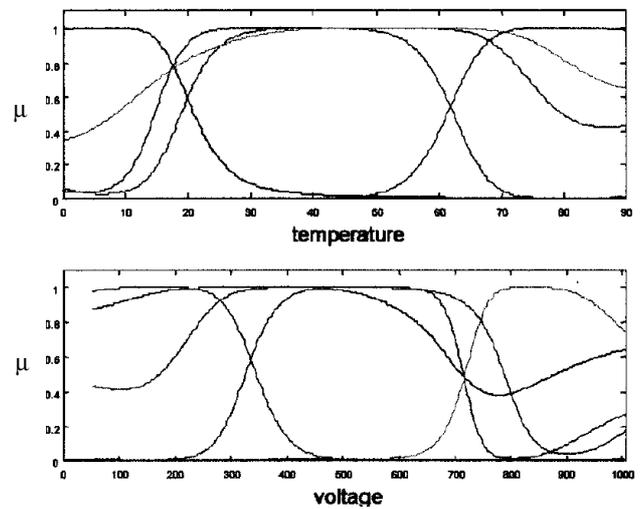

Fig. 5. Fuzzy clustering membership functions: 5 clusters and a fuzziness parameter of 1,8

4.2. Conductivity fuzzy sensor models

The lack of the goodness of the Mamdani models led us to continue only in the fuzzy clustering manner. So, the numerical data set available for modelling in the conductivity case was handled in the same way as in the case presented before (pH).

Fig. 6, which represents the maximum error for the obtained models, denotes a suitable solution with projected membership functions instead of product-space membership functions.
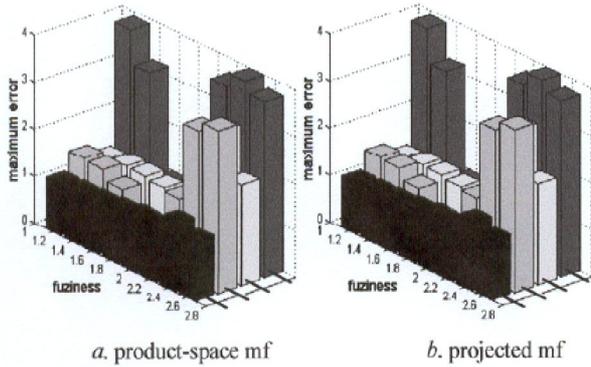

a. product-space mf   b. projected mf

Fig. 6. Maximum error in percentage for conductivity-fuzzy clustering models

Fig. 7 shows that from the VAF point of view, both types of membership functions produce good results.

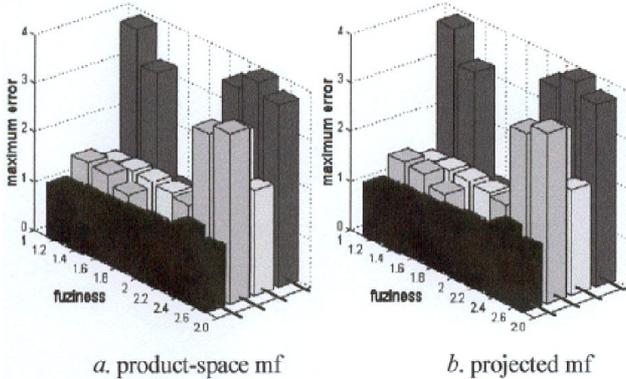a. product-space mf   b. projected mf

Fig. 7. VAF for conductivity-fuzzy clustering models

The above-mentioned suitability brings up another aspect: the rms measure points to the same best model as the emax's and VAF's ones (Table IV).

TABLE IV. Best results in terms of emax, VAF and rms

Description	clusters=9 fuzziness=2,4 projected mf	clusters=9 fuzziness=2,6 projected mf
emax	<u>0,3246</u>	0,3772
VAF	99,9992	<u>99,9993</u>
Rms	<u>0,2758</u>	<u>0,2758</u>

4.3. Turbidity fuzzy sensor models

In this case, the size of the numerical data set available for modelling (Fig. 2(c)) does not allow a number of clusters higher than 3. Notwithstanding, the results in terms of the maximum error are acceptable (Fig. 8).
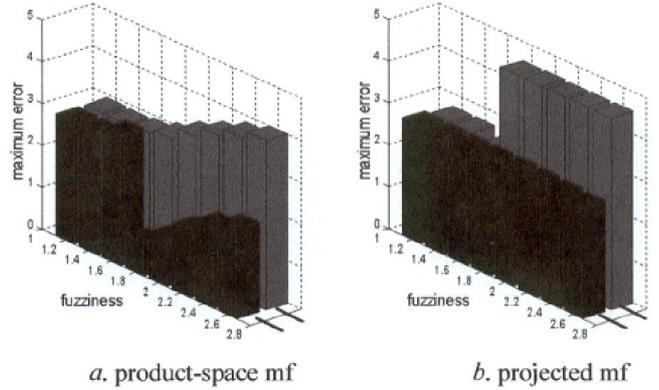

a. product-space mf   b. projected mf

Fig. 8. Maximum error in percentage for turbidity-fuzzy clustering models

Results in terms of the VAF (Fig. 9) underline the above remark if we are agree with the fact that, in that evaluation criterion, the second digit after decimal point is insignificant, a widely accepted principle within the modelling experts community.
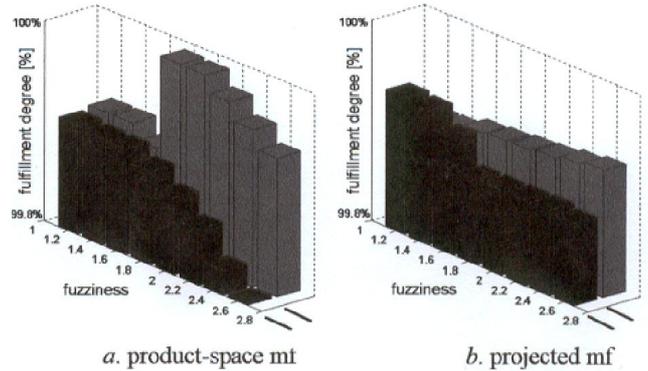

a. product-space mf   b. projected mf

Fig. 9. VAF for turbidity-fuzzy clustering models

4.4. Some practical aspects

Fig. 10 shows the number of floating point operations (flops) and the time used for processing a data input for a fuzzy clustering model. Both are strongly related with the number of clusters and clustering algorithm used in modelling.
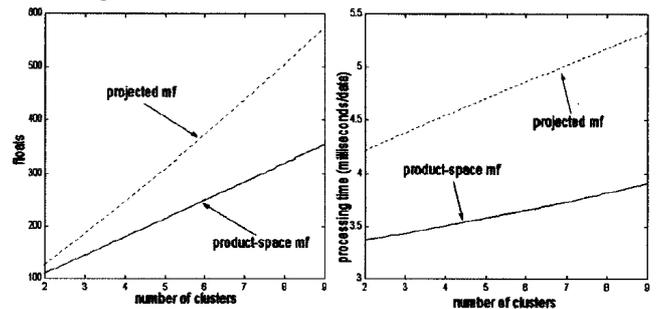

Fig. 10. Number of flops and processing time for a data input

5. CONCLUSIONS

The article proposes an intelligent virtual multi-transducer system that includes several fuzzy processing structures with application in water quality monitoring. The aim of the intelligent processing is to increase the pH, TU and C measurement accuracies performing on-line temperature compensation. It was emphasized in this paper that the tuning process of the fuzzy models must take into account the influences of many parameters onto model performances. Among these ones, the most important is the structure, clustering proving to be a very effective technique for dealing with large sets of data with an accurate prediction. When fuzzy clustering is applied to generate fuzzy models from data, the number of clusters is the most important parameter that must be chosen: Assuming that the clusters should represent local linear models of the function being approximated, they provide a correct local description of the sensor. This feature distinguishes fuzzy models generated by product-space clustering from the majority of other local approximation methods, like radial basis functions or splines. An important aspect could be that the overlap of the adjacent membership functions already gives an indication of the local character of the non-linearity. The results obtained allows to concluding that even with a small size data set, models with satisfactory quality can be obtained.

REFERENCES

[1] J.C.Patra, G. Panda, and R. Baliarsingh, "Artificial Neural Network based non-linearity estimation of pressure sensors", IEEE Trans. Instr. Meas., Vol. 43, pp.874-881, December 1994.

[2] O. Postolache, P. Girão, M. Pereira, "Neural Network in Automatic Measurement System: State of Art and New Research Trends", Proc. IEEE - IJCNN'2001, Washington DC, vol.3-4, pp.2310-2315, July 2001.

[3] G. Mauris, E. Benoit, L. Foulloy, "Fuzzy symbolic sensors: From concepts to applications", *Measurement*, vol. 12, pp. 357,1994.

[4] C. Temneanu, C. Lazar, "Linguistic Fuzzy Controllers. A Study of Some Influence Factors", XII-Th International Symposium on Electrical Apparatus and Technologies *SIELA 2001*, vol. 1, pp. 321, 2001.

[5] R. Babuška, *Fuzzy modelling for control*, USA: Kluwer Academic Publishers, 1998.