Modelling of atmospheric pollutants in petrochemical refineries

Nicola Barbalace and Letterio Giannetto

Abstract

In the present work, the model identification of industrial process was considered, in particular a model for the prediction of atmospheric pollutant output (NOx) from a chimney of a refinery was developed to substitute line analysers in case of breakdown or anomalies. The model will permit the analysis of anomalous behaviour which can occur in the plant. In fact, with this aim, models for realising a "soft sensor" have been implemented, which supply "on-line" monitoring of the levels of atmospheric pollution (NOx, nitrogen oxide) in the chimneys of a refinery. The best results, in terms of quality indices and reliability, were obtained through non linear models. Thanks to the potential of neural networks, these results represent a valid alternative to on line measurements during maintenance of the analysers. After simulation and comparisons, it is shown that two static neural models, one for each chimney, can be considered completely satisfactory, above all if compared with both the linear model currently in use in refineries and the linear model obtained by the Squared Minimum Method. Copyright ©2004.

Mathematics Subject Classification: 92B20, 93B05.
Key words: neural networks, model approximation, instability, non-linear models, identification.

1 Description of the problem.

The inherent security problems of petrochemical refineries and the pollution produced by them are more frequently becoming the subject of research and studies. In fact, the best procedures for the optimisation for the running and control of a plant are currently being defined. In particular, in the field of refining petrol and its derivatives, a working adjustment of the qualitative specifications imposed by the law in relation to atmospheric pollution from plants, has become a need and an obligation.

In the production processes carried out inside a refinery, there are numerous factors which produce an undesirable and harmful impact on the environment. Legislation establishes that all emission sources of hydrocarbon, compound and harmful

Proceedings of The 3-rd International Colloquium "Mathematics in Engineering and Numerical Physics" October 7-9 , 2004, Bucharest, Romania, pp. 23-31.

and /or foul smelling vapours must be accurately examined and tested. Therefore, it is necessary to continually monitor polluting emissions from chimneys into the atmosphere. Furthermore, all necessary means to guarantee that the concentrations of these emissions are always contained within the lowest levels established by the law, must be foreseen. In studies on methodologies suitable for reducing emissions into the atmosphere, it has emerged from accurate studies, that to improve the inherent problems of pollution, it is necessary to centralise the smoke discharge from secondary plants. In fact, in the petrol plant, the smoke discharge from secondary plants is centralised in one chimney called side A with an overall capacity for all discharges of all consumptions.

The emissions from fumes burnt inside the chimney are monitored by hardware sensors: ”online analysers”. These allow a continual monitoring of pollution levels by carrying out measurements with time of sampling equal to one minute. All the measured samples are checked and given a validity index. On all valid samples a hourly average is effected and stored in a database; this is necessary since all checks and running conditions of the chimney are developed using hourly average; therefore it is opportune, as dictated by precise legislation, that such data has, step by step, a reliable value. So that the pollution level, measured by analysers, expelled by the chimneys, supplied by monthly averages, can be considered reliable, it is necessary that the number of measurements is sufficiently high and in particular, it must not fall below 80% of the conditions of ideal functioning. Often, however, due to breakdowns or anomalies in the analysers, the data is not reliable and it is for this reason that the technology of soft sensor was chosen in the refinery.

A linear model obtained through empirical calculations is used in the refinery as an instrument for predicting output values of the chimneys based on selected and tested inputs and coefficients. The simulated values obtained from the model, currently in use in the refinery, are, from a comparison reported below, unsatisfactory, in particular if compared with models implemented in the present article. The typologies considered for the implementation of the models, for the realisation of ”soft sensors” for the monitoring of pollution levels, are: linear (least squares technique) and non linear (implemented by neural networks, [2]). From a comparison of the models implemented, the most representative was determined, based on precise indices of quality. The sensors were designed on the basis of input/ output data contained in the database.

2 Theoretical calculation of NOx emissions.

In the refinery, a theoretical calculation was implemented to compensate for eventual breakdowns of the instrumentation placed on the chimneys, suitable for surveying NOx emissions in the atmosphere. The fuels burnt in the refinery furnaces were considered in this calculation. There are two types of fuels: Fuel Gas and Fuel Oil. Fuel oil has a greater polluting power in terms of NOx and the calculation was made by multiplying the capacity of burnt fuel in the furnaces (whose fumes are conveyed to that chimney) by appropriate emission factors. To calculate the contribution of each furnace, a value of NOx concentration, ”typical of the burner” is referred to, from which an emission factor which allows the quantity of burnt fuel to be correlated with kilogram’s of NOx products, is obtained. The development of a software sensor is, in general, realised off-line through the following phases: analysis and pre-elaboration of historical data of the process, definition of the model structure, training and finally, validation.

Fig.1 depicts the diagram of the flow of information necessary for the development of a virtual software sensor.

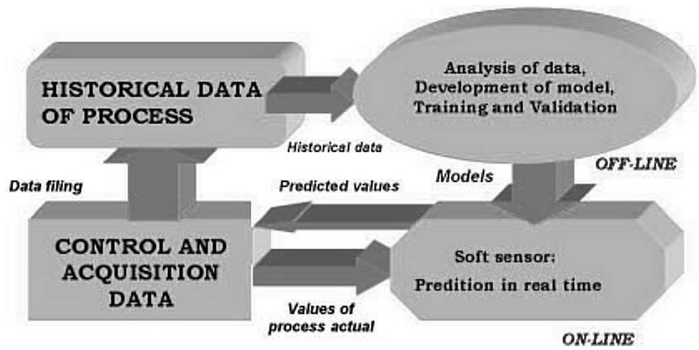

Figure 1: Developement of a sensor.

3 Identification techniques.

When physical modelling of the system is not possible the attention must be focused on strategies which directly use the measurements of the signals produced by the system. These are defined by the term "system identification". When the structure of the system to be modelled is not known previously, one attempts to characterise the output of the system to a given instant (k) as an expansion of the input/output data through a non linear static function S (NARMAX model) [3]:

$$y(k) = S[y(k), u(k)]\,. \tag{3.1}$$

4 Identification through neural networks.

A useful approach in approximating the S function of 3.1 is represented by neural networks ([4], [5], [6]). These models are composed of many non linear computational elements operating in parallel and arranged in patterns reminiscent of biological neural nets. Computational elements or nodes are connected via weights that are typically adapted during use to improve performance. The weight of networks constitutes the parameters to be identified, and the learning algorithm, which minimises the quadratic error between the output of the model and the real output of the process, assumes the role of an identification algorithm. The inputs of the network correspond to the input/output samples of the system, while the unknown S function is determined repeatedly through the learning algorithm as a combination of sigmoidal function.

The data furnished by the experts of refinery are listed in table 4; the number of inputs is 22; the considered entries are those, given by the experts, relative to the theoretical calculation factor NOx.

DESCRIPTION	UNIT	Indicative Contribution
COURSE FUEL GAS	kg/h	23,50%
VENT GAS	kg/h	
FUEL OIL A F-101	kg/h	
Return FUEL OIL	kg/h	
O2 zone conv. 06F101 mountain	%	
O2 zone conv. 06F101 sea	%	
F-101 FUEL-GAS	kg/h	1,20%
FUEL GAS F101	kg/h	0,38%
VENT GAS OF BURNER F201	kg/h	
VENT GAS OF BURNER F201	kg/h	
OXYGENSMOKES F201	%	
FUEL GAS A F101	kg/h	4,20%
% O2 FUMI F101	%O2	
FUEL GAS A LIMITE BATT.	kg/h	0,05%
METHANE L.B.A IMP. 1200	kg/h	1,12%
F.G. A HEAT CELL.	kg/h	3,20%
F.G. A SOAK. CELL	kg/h	
O2 FUMES H CELL. F101	%	
O2 FUMES S CELL. F101	%	
F.G. TO F301	kg/h	1,15%
VENT GAS from D311 to F301	T/h	
ANAL. %O2 FUMES F301	%	
GAS OF COMBUSTION C.1	kg/h	21,82%
OIL OF COMBUSTION C.1	kg/h	
OXYGEN FUMES C.1	%	
GAS OF COMBUSTION C.2	kg/h	21,20%
OLIO OF COMBUSTION C.2	kg/h	
OXYGEN FUMES C.2	%	
GAS OF COMBUSTION C.3	Kg/h	20,77%
OIL OF COMBUSTION C.3	Kg/h	
OXYGEN SMOKES C.3	%	

Table 1 Inputs

The theoretical calculation has been determined on the base of experimental evaluations and represents the coefficient of multiplication of the variable in the theoretical linear model used in the refinery. The considered output is the concentration of NOx in the fumes.

The model used in refinery is a linear model obtained using as coefficients the values obtained with "the theoretical calculation factor" reported in the column FN in table 4.

To be able to evaluate if the error signal, that represents the difference between the simulated output of the implemented model and the real output of the process, is of a random type, it was necessary to determine some parameters. These parameters are the autocorrelation function of the error and the probability density function. The "best conditions" present themselves in the case in which: the average of the error tends towards zero, the autocorrelation function assumes a trend as similar as possible to a white noise, the correlation coefficient tends towards one, the simulated output of the model as similar as possible towards the real output of the process.

4.1 Linear models.

The model used in the refinery is a linear model obtained by using ai as a coefficient, the values of the "Theoretical Calculation Factor" reported in the FN column in table

Description	UNIT	Name	FN
NOX COMP. Time average	mg/Nm3	OUT	
Fuel Gas	Kg/h	IN.1	0.004823
Vent Gas	kg/h	IN.2	0.004823
Fuel Oil	kg/h	IN.3	0.00729
Fuel Oil	kg/h	IN.4	0.00729
Fuel Gas	Kg/h	IN.5	0.00195
Fuel Gas	Kg/h	IN.6	0.00373
Fuel Gas	Kg/h	IN.7	0.002365
Methane	Kg/h	IN. 8	0.002365
Fuel Gas	Kg/h	IN.9	0.00278
Vent Gas	Kg/h	IN.10	0.00278
Fuel Gas	Kg/h	IN.11	0.002087
Vent Gas	Kg/h	IN.12	0.002087
GAS OF COMBUSTION C. 1	kg/h	IN. 13	0.00403
GAS OF COMBUSTION C. 2	kg/h	IN. 14	0.00403
GAS OF COMBUSTION C. 3	kg/h	IN. 15	0.00403
OIL OF COMBUSTION C. 1	kg/h	IN. 16	0.00448
OLIO OF COMBUSTION C. 2	kg/h	IN. 17	0.00448
OLIO OF COMBUSTION C. 3	kg/h	IN. 18	0.00448

Table 2 Grouping of inputs in base to FN and output

4. The output of the model is represented by the following expression:

$$\text{out-linear} = in_1 \cdot a_1 + in_2 \cdot a_2 + \cdots + in_{18} \cdot a_{18}\,. \tag{4.1}$$

A comparison between the real output and that of the linear model obtained on the entire set of data is reported in Fig.4.

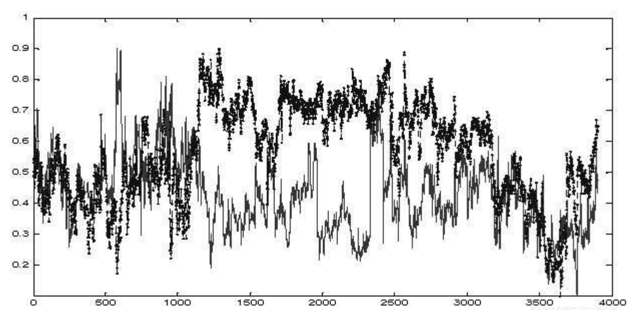

Figure 2: Real output (continuous) - Linear model output (dotted).

It is possible to see how the output of the linear model differs, especially in some areas, from the real output, of the process; therefore, the model doesn't sufficiently identify the real dynamics of the system. Furthermore, the analysed linear model, as seen in Fig.2 doesn't have a conservative behaviour in that it tends to under estimate the effective value of pollutants in the chimney output.

A first approach implemented to resolve these problem is that of "Least Squares Method" (LMS), it allows to calculate the coefficients that minimize the quadratic error between real output and model output. Using LMS, after numerous tests on implemented models, it emerged that the coefficients that minimise most the quadratic error average permit, therefore, a better approximation of the real dynamics of the process and are reported in table 4.1:

Variable	Coefficient	Variable	Coefficient
h1	-0. 0262	*h10*	0.0266
h2	-0.0007	*h11*	0.2609
h3	-0.0145	*h12*	-0.4203
h4	0.0053	*h13*	-0.0078
h5	-0.0304	*h14*	-0.0014
h6	-0.0223	*h15*	0.0423
h7	-0.0303	*h16*	-0.0025
h8	0.0930	*h17*	-0.0018
h9	-0.2904	*h18*	0.0061

Table 3 Coefficients obtained with LMS

For the models implemented with this method, considering the entire set of data (3900 samples), only the best, in terms of performance, was reported; reported in the following graph are: the comparison between the real output and the output relative to the coefficients determined by LMS (Fig.3).

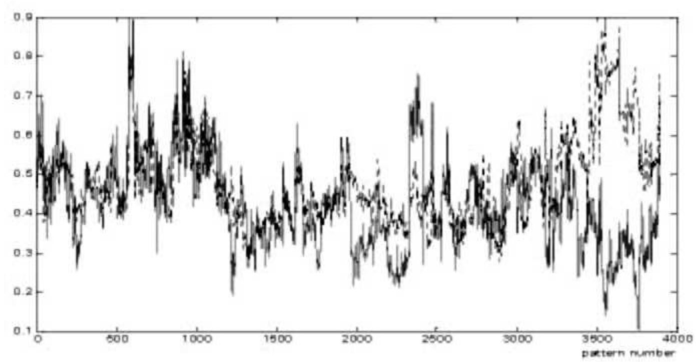

Figure 3: Real output (continuous) - Linear model output M.M.Q. (dotted).

As we can see, a first comparison can be made to determine which of the analysed linear models is better in terms of reliability and quality indices. Fig.4 reports the comparison between the outputs: real (plotted in continuous) of the NOx emission process from the A side chimney; the linear model 1 (light and dotted) currently in use in the refinery and obtained by considering the coefficients relative to the Theoretical Calculation; the linear model 2 (dotted) considering the coefficients obtained with LMS.

5 Neural models.

Three series of models, using neural networks, were realised and more precisely: the first considering ten inputs those with a higher degree of influence on the output;

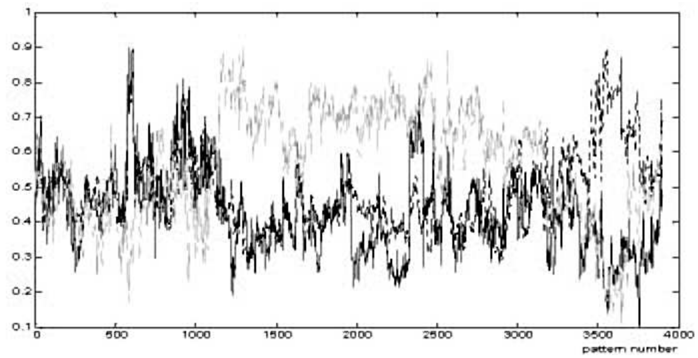

Figure 4: Output: model 1 (light and dotted) - real (continuous) - model 2 (dotted and marked).

the second, using all the inputs (Table 4) and finally, using nine inputs and, that is, considering as inputs the sum of the variable with the same influence index. The best results, in terms of reliability of the model and quality indices, were obtained by considering as inputs all the variables of the process taken separately, considering therefore, also those with a low influence index.

Fig.5 reported the results of the neural model (18 inputs with 9 neurons in the hidden layer) which present the best performances and thus most closely approach the real dynamics of the process. A comparison of the obtained results has consequently

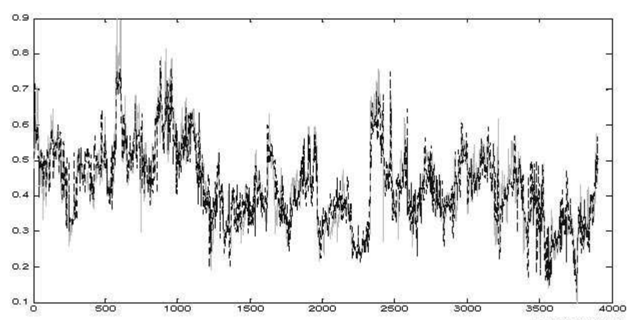

Figure 5: Simulated output (dotted) - Real output (light thickness).

shown that a neural model can be considered fully satisfactory, especially if compared with a linear model. In fact, in the following figures we report the comparison between the outputs obtained by applying the best, in terms of performance, neural network (18 inputs and 9 neurons in the hidden layer) and the best, in terms of performance, linear model (least squared method) on the entire data set; in Fig.6, the real output of the process is plotted in continuous, the linear model output in light thickness and the neural model output dotted.

In table 5 there are the indexes of quality of the linear model gotten with the L.M.S. and of the neural model.

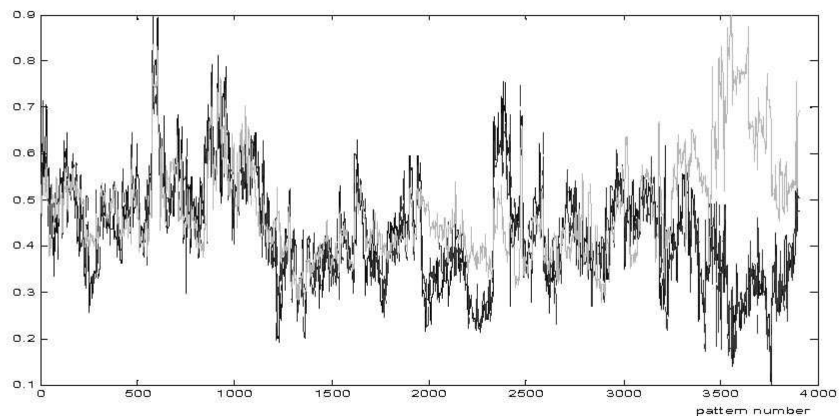

Figure 6: Output: neural model (dotted) - real (continuous) - linear model (light thickess).

	Average of the signal error	Correlation coefficient
Neural Net	5.3360e-004	0.9130
Linear model	-0. 0506	0.2323

Table 4 Indexes of quality

As we can see all the indices and the graphs give complete evidence in favour of the neural model; in fact, the average of the error is lower, the mutual coefficient of correlation between the real output and the simulated one is higher in the neural model and finally, the dynamics of the process is most similar estimated by the neural model. The results encourage the use of neural architecture in a vast range of applications in the refinery. However, it has been observed that the success or less in determining a model depends strongly on the choice and the elaboration of the input sizes and an accurate analysis of the process to be modelled.

6 Conclusion.

In the present work, the problem of identification in an industrial process was confronted and consequently, the determination of a model for the control and prediction of atmospheric pollutant output (NOx) from one chimney. It has therefore, determined a model that best lends itself to the problems of the process regarding atmospheric polluting emissions in a petrochemical plant, from both chimneys, for the development of software sensors to substitute line analysers in case of breakdown or anomalies.

The best results, in terms of quality indices and reliability, were obtained through non linear models. Thanks to the potential of neural networks, these results have demonstrated that today, they represent a clearly valid alternative to the classic analysers used in industry to monitor and store data.

In fact, after simulation and comparisons, it is shown that two static neural models, one for each chimney, can be considered completely satisfactory, above all if compared with both the linear model currently in use in refineries and the linear model obtained by the Squared Minimum Method. These results encourage the use of neural architecture in a vast range of applications in the automation of a refinery. The software sensors, realised for both chimneys, therefore represent a new type of approach which has numerous advantages compared to the various traditional sensors. Furthermore, software sensors, in comparison with traditional hardware sensors, don't need any maintenance. Consequently, it is possible to place the software sensors next to the measuring instrumentation in the refinery, to test the measurements of hardware sensors and which they can effectively temporarily replace in case they malfunction.

References

[1] H.A. Henson and D.E. Seborg, *Non linear process control*, Prentice Hall, 1991.

[2] J.M. Zurada, *Introduction to artificial neural systems*, St.Paul, MN.

[3] S. Chen, S.A. Billings, *Representations of non-linear system: the NARMAX model model"*, Int. J. Control, 43 (1989), 1013-1032.

[4] P. Arena, L. Fortuna, G. Nunnari, M.G. Xibilia, A. Gallo, *Air pollution prediction via neural networks*, Proc. 7-th IMACS-IFAC Symp. LSS'95, London, 1995.

[5] S.R. Chu, R. Shoureshi, M. Tenorio, *Neural networks for system identification"*, IEEE Control System Magazine, April 1990, 31-34.

[6] S. Chen, S.A. Billings, P.M. Grant, *Non linear system identification using neural networks*, Int. J. Control, 51, 60 (1990), 1191-1214.

Nicola Barbalace
Department of Mathematics
University of Messina, Sicily - Italy
e-mail: nbarbalace@ingegneria.unime.it

Letterio Giannetto
Department of Statistics
University of Messina, Sicily - Italy
e-mail: lgiannetto@unime.it