OZONE CONCENTRATION MODELING USING A FUZZY MODEL OVER THE BASSE NORMANDIE

E. ELAYAN * , F. GIRI, E. PIGEON and J-F. MASSIEU

* GREYC, UMR 6072 CNRS, Universite de Caen, ENSICAEN,
6, Bd Marechal Juin, 14050 Caen Cedex, FRANCE.
elamari.elayan@greyc.ensicaen.fr

Abstract: Takagi-Sugeno (TS) modeling formalism has been widely used to build up multi-model representations for nonlinear systems. In this paper, TS approach is applied to get a multi-model description of Ozone generation process in a specific geographical zone. This is done in two steps: firstly, an adequate structure of the desired multi-model is designed. This structure involves a set of local linear models (each one is valid for a certain range of operating conditions) and an interpolative mechanism that combines the outputs of the local models into a continuous global output. The second step consists in identifying the parameters of the local models using the parametric identification approach. The identification scheme developed is applied to model ozone generation in the Basse-Normandie region (France). The model thus obtained turned out to be satisfactory and currently used to build-up a predictor for this region.

Keywords: Nonlinear system identification; Takagi-Sugeno fuzzy modeling; Environmental data; Ozone modeling; Atmospheric pollution.

1. INTRODUCTION

One of the major issues in air pollution is the high ozone concentration of the troposphere in summer. About 10 % of ozone exists in the troposphere of the earth. High ozone concentration appears from May to September. Generally, the mechanism of Ozone concentration is formed by photochemical reaction and metrological variation. In this mechanism, nitric dioxide and hydrocarbon act as pollution materials, while solar radiation, wind speed, and temperature play the role of metrological materials. Nowadays, developing industries and increasing automobile are responsible for air pollution materials. The physio-chemical mechanism taking place are poorly understood, but it is clear that this process is multivariable, strongly nonlinear and time varying. Ozone process modeling has been dealt with following different approaches. Among these, the statistical and physical methods which account for the atmospheric chemistry of involved reactions, the emission of primary pollutants, as well as vertical and horizontal exchanges linked to movements of the atmosphere [5]. The obtained models are seldom resorted to because they need measurement that are seldom available (in air quality monitoring), also they are very complex and costly.
The dynamic system approach is a quite different alternative to Ozone process modeling. It consists in determining a black-box type model that relates the process inputs to the output. A deep comprehension of the atmospheric reactions is not needed. In [6], neural networks formalism is proposed to model this process. More simple and mathematically tractable models can be obtained with the fuzzy modeling approach developed by [4], [5]. The fundamental idea is to partition the input space into fuzzy areas and approximate, in each area, the studied process by a more simple local model (often a linear one). The global model is then obtained by interpolation between the different local models.
In this paper, a multi-model representation is devel-

oped based on the TS approach [4], for the ozone process in the region of Basse-Normandie (France). Such a process is multi-input single output (MISO) and highly nonlinear. The output is the Ozone concentration while the input vector includes the metrological variables (wind speed, solar radiation, pressure, temperature, ...) and the pollutants (NO2, NO,..). It is worth noting that most of these inputs are only measurable and can not be act on. Then, the input vector space is shared in zones. Consequently, each input zone is characterized by a membership function of the sigmoidal type. Furthermore, the zone process dynamics are represented, in each zone, by a local linear model whose parameters are still to be identified. Finally, the outputs of local models are combined to construct the output of the global model using an interpolation mechanism involving the membership functions. The model structure thus obtained is based upon to perform parameter estimation for the different local models. Since the number of unknown parameters grow exponentially with the number of inputs and the order of local models, it is clear that the model may becomes rapidly complex. Therefore, the different design choices, regarding the model structure, should be done bearing in mind the balance between model accuracy and complexity. For the considered studied region, choosing first order local models and 6 inputs reasonably selected, turned out to be a satisfactory compromise. A validation study with real data shows that the multi-model representation thus developed is a model gives satisfactory results.

This paper is organized as follows: in Section 2, a concise overview of fuzzy-based multi-model approach is described. In Section 3, parameter identification of the local models is performed and related issues are discussed. In Section 4, the above identification scheme is applied to Ozone process in the concerned geographical region. Section 5 is devoted to discuss some model improvement. In Section 6, a conclusion, perspective and references.

2. DEVELOPMENT OF MULTI-MODEL REPRESENTATIONS BASED ON (TS) FUZZY APPROACH

Consider a nonlinear system with n inputs and one output, let $u(k) = [u_1(k),\ u_2(k),\ \ldots,\ u_n(k)]$ be the vector of input variables and y(k) is the output variable, at each instant k. According to TS approach, the system is locally described by M local models. The way these models describe the system is formulated using M rules of the form:

$$R_i : \begin{cases} \textbf{If} \\ z_1(k-1)\ is\ A_1^{i_1}\ and\ \cdots\ and\ z_m(k-1)\ is\ A_m^{i_m}, \\ \textbf{then} \\ y^{(i)}(k) = \sum_{j=1}^{p} a^{(i)} y(k-j) + \sum_{l=1}^{n}\sum_{j=1}^{ql} b_{lj}^{(i)} u_l(k-d_l-j), \end{cases} \quad (1)$$

where i = 1,2, ... ,M and $(i_1,\ldots,i_m) \in \{1,2\} \times \ldots \times \{1,2\}$ m times. In (1), $\mathbf{z(k-1)} = [z_1(k-1),\ z_2(k-1)\ ,\ldots,\ z_m(k-1)]$ represents the antecedent variables, while A_r^i represents the i-th fuzzy domain related to the r-th input. Then, m represents the number of modalities (the number of fuzzy sets) defined over the variation domain of each antecedent variable.

The coefficients $a_j^{(i)} and\ b_{lj}^{(i)}$ $(i = 1,\ \ldots,\ M,\ l = 1,\ \ldots\ n,\ j = 1,\ \ldots\ q_l)$ are the parameters of the i-th local model, while $p, q_l, d_l)(l = 1,\ldots,n)$ define the structure of such a model. Note that the antecedent (premise) vector can include, in the general case, both consequence variable (especially) y and input variables **u** (at times prior to k). Auxiliary measured variables may also be included in **z**, provided that they contribute in capturing the process nonlinearity.

An estimation $\hat{y}$ of the process output is defined to be a weighted average of outputs of the local models i.e.

$$\hat{y}(k) = \frac{\sum_{i=1}^{M} \nu_i(k-1)\hat{y}^{(i)}(k)}{\sum_{i=1}^{M} \nu_i(k-1)} \quad (2)$$

with,

$$\nu_i(k-1) = \prod_{r=1}^{m} \mu_r^i(z_r(k-1)), \quad (3)$$

where $\nu_i(k-1)$, called the truth degree of the R_i, is defined as the product of the membership grades $\mu_r^i(z_r(k-1))$of the premise variables $z_r(k-1))$ to their corresponding modalities A_r^i (r = 1, ... m).

On the other hand the local models involved in (1) can be given the following regressive form:

$$\hat{y}^{(i)}(k) = \varphi^T(k-1)\theta_i \quad (4)$$

Substituting (4) in (2), yields:

$$\hat{y}(k) = \frac{\sum_{i=1}^{M} \nu_i(k-1)\varphi^T(k-1)\theta_i}{\sum_{i=1}^{M} \nu_i(k-1)} \quad (5)$$

where $\varphi(k-1) = (y(k-1)\ldots y(k-p)\ u_1(k-d_1-1)\ldots u_1(k-d_1-q_1)\ \ldots\ u_n(k-d_l-1)\ \ldots\ u_n(k-d_l-q_n))^T$ is the global regression vector and $\theta_i = (a_1^{(i)}\ldots a_p^{(i)}\ b_{11}^{(i)}\ \ldots b_{1q_1}^{(i)} b_{21}^{(i)}\ \ldots\ b_{2q_2}^{(i)}\ b_{n1}^{(i)}\ \ldots\ b_{nq_n}^{(i)})^T$is the parameter vector of the local model i.

Introducing the normalized truth degree of rules:

$$\omega_i(k-1) = \frac{\nu_i(k-1)}{\sum_{i=1}^{M} \nu_i(k-1)}$$

equation (5) becomes:

$$\hat{y}(k) = \sum_{i=1}^{M} \omega_i(k-1)\varphi^T(k-1)\theta_i \quad (6)$$

Now, let $\varepsilon(k)$ denote the equation error

$$\varepsilon(k) = y(k) - \hat{y}(k) \quad (7)$$

where y(k) is the measured output (ozone concentration). In the sequel, $\varepsilon(k)$ is supposed to undergo the following stochastic model:

$$\varepsilon(k) = \xi(k) + c_1\xi(k-1) + \quad \cdots \quad + c_n\xi(k-n) = C(q^{-1})\xi(k) \quad (8)$$

where, $\xi(k)$ is a white noise and (q^{-1}) is the delay operator. Then, the global model becomes:

$$y(k) = \sum_{i=1}^{M} \omega_i(k-1)\varphi^T(k-1)\theta_i + C(q^{-1})\xi(k) \quad (9)$$

Remarks: a) The above stochastic model of the output error $\varepsilon(k)$ is a moving average (MA) type. Other model structures (AR, ARMA) could be used [1].
b) Equation (9) is a multi-model representation of the considered nonlinear system. In this representation, the parameters c_j (j=1,2,...,n) and θ_i (i=1, 2, ... , M) and (c) are still to be identified. This is the subject of the next section.
c) With n premise variables, and m modalities per input(where the premise variables is equal to the input variables in this work), the number of local models is m^n.

3. A MULTI-MODEL IDENTIFICATION SCHEME FOR THE OZONE PROCESS

In this section, an identification scheme is designed to get a multi-model representation like (9) for the ozone process in the Basse-Normandie region. This will be done in three steps: the first one is devoted to determining (or making an appropriate choice) of the model antecedent. Model consequents structure identification is dealt with in the second step. Finally, model validation criteria will be proposed.

3.1 Model antecedents Choice:

Model antecedents determination amounts to selecting appropriate inputs, and membership functions as well as making an adequate partitioning of the input space

3.1.1. Selection of premise variable In most identification problems, the considered process is characterized by many input variables. The ozone process we are dealing with will involve the following inputs NO2, NO, Temperature, Humidity, Wind speed, Solar radiation. All these inputs are considered to be premise variables, ozone concentration is not. Consequently, $\mathbf{z(k)} = \mathbf{u(k)}$.

3.1.2. Choice of membership functions It is clear that the number and nature of membership functions play a certain role in achieving a high quality model. For a given type of functions, the more modalities the higher quality. But increasing the number of membership functions leads to a complex model. Then, the number of modalities should be a compromise between model quality and complexity. Since the number of inputs is high (n=6), it turned out to be impractical to take more than two modalities. Therefore, the most convenient membership functions are the following sigmoidal type

$$\mu_i^1(\tilde{z}_i(k)) = \mu_r^1 = \frac{1-\tan((\tilde{z}_i(k)-\bar{z}_i)/\sigma}{2} \quad (10)$$

$$\mu_i^2(\tilde{z}_i(k)) = \mu_r^2 = \frac{1+\tan((\tilde{z}_i(k)-\bar{z}_i)/\sigma}{2} \quad (11)$$

where:
$\tilde{z}_i(k)$ (i=1, 2, ..., m) is a normalized version of $z_i(k)$:

$$\tilde{z}_i(k) = \frac{z_i(k) - min_{\ k}(z_i(k))}{max_{\ k}(z_i(k)) - min_{\ k}(z_i(k))} \quad (12)$$

$\bar{z}_i$ is the mean value of the normalized variable $\tilde{z}$. The parameter σ characterizes the width of the membership function. Such a parameter plays a similar role as the standard deviation in a gaussian distribution.
The normalized variables ($\tilde{z}_i(k)$) vary between 0 and 1, and the data is distributed into the intervals $[0, \bar{z}_i]$ and $[\bar{z}_i, 1]$, where (i=1, 2, ..., m).
Notice that, due to normalization the structure of the membership functions satisfy, at each instant, the following relation:

$\mu_{r1} + \mu_{r2} = 1$

Also, it follows from (3) that the truth degree function $\nu_i(k)$, satisfies the following formula

$\sum_{i=1}^{M} \nu_i(k) = 1$

3.2 Identification of model consequents

Identification of model consequences includes determination of both the structure and the parameters for the local models.

3.2.1. Local models structure the structure of the local models should be chosen taking in account the balance between model accuracy and complexity. It follows from (1) that, there are $(\sum_{l=1}^{n} q_l + p) \times M$ parameters in the consequence part that have to be estimated. For example, if n=6 and m=2, p=1, $q_l = 1(\forall l)$ and $M = m^n = 64$, then the number of estimated parameters is 448 parameters. Such a large number of parameters may lead to a long time computation.
In the present study, we have decided to not limit the number of inputs. Therefore, the model structure parameters should be limited as follows: $q_l = 1$ *and* $m =$

2, the most important is to obtain a model that is the most accurate possible.

3.2.2. Parameter identification In this subsection, the focus is made on estimation of the parameter vector θ_i, (i=1, 2, ..., M) based on model structure (9). The latter is a (nonlinear) equation-error type parametrization. Let us consider the simple case where c=0, this can be given the following regression form:

$$y(k) = \Phi(k-1)^T \theta^* + \xi(k) \qquad (13)$$

with:
$\Phi(k-1) = [\omega_1(k-1)\phi^T(k-1) \quad \omega_2(k-1)\phi^T(k-1) \quad \ldots \quad \omega_M(k-1)\phi^T(k-1)]^T$

$\theta^* = [\theta_1^T \; \theta_2^T \; \ldots \; \theta_M^T]^T$
As θ^* comes linearly in (13), it can be estimated, without bias, using the standard least squares algorithm:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{F(k-1)\Phi(k-1)e(k,\hat{\theta}(k-1))}{1+\Phi(k-1)^T F(k-1)\Phi(k-1)} \qquad (14)$$

$$F(k) = F(k-1) - \frac{F(k-1)\Phi(k-1)\Phi(k-1)^T F(k-1)}{1+\Phi(k-1)^T F(k-1)\Phi(k-1)} \qquad (15)$$

Where $\hat{\theta}(0) = 0$, and $F(0) = 10^4.I$

$$e(k,x) = y(t) - \Phi(k)^T x \qquad (16)$$

Remarks :
a) The case where parameter c is not zero will be discussed latter (see section 5.2)
b) For the estimated parameter vector $\hat{\theta}(k)$ to be asymptotically unbiased, it is sufficient that $\Phi(k)$ is persistency exciting [1].

3.3 Model validation

Model validation can be performed using two criteria. The first is a prediction type criterion. It consists in comparing different model structures by the following quantity:

$$\bar{e} = \frac{1/N\sum_{k=1}^{N}|e(k,\hat{\theta})|}{1/N\sum_{k=1}^{N}|y(k)|} \qquad (17)$$

where N is the length of the available vector data and $\hat{\theta}$ denote the final estimated model $\hat{\theta}(N)$ obtained by (14). Note that $\bar{e}$ represents a relative mean value of the estimation error $e(k,\hat{\theta})$. The best model is the one that yields the smallest value of $\bar{e}$.
The second criterion consists in determining the autocorrelation $\varphi_e(\tau)$ of the estimation error $e(k,\hat{\theta})$. The ideal model is one which generate an uncorrelated sequence $e(k,\hat{\theta})$. Consequently, the ideal shape for $\varphi_e(\tau)$ is a Dirac like impulse.

4. APPLICATION OF THE IDENTIFICATION SCHEME TO THE OZONE PROCESS

The identification scheme developed in section 3, is applied to get a multi-model representation of the ozone process in the Basse-Normandie region. The real data has been provided by AIRCOM (Air de Calvados de lòrne et de la Manche). This is an association for supervision and study of the atmospheric pollution in the French town Caen.

4.1 Data description

The data we have used was provided by AIRCOM from different measurement sites (Herouville, Carpiquet, Ouisterham, Ranville, Vaucelles, Chemin vert, Ifs and Tour Leroy). All these sites are located in Caen, at the west of France, over the period from first May to thirty September, for the years 95-2001. Many metrological variables were measured and Table I illustrates the most important of them.
The sampling period is one hour and the number of samples input for each measured variable is 6720 (this is less than what it should be, i.e. 25000, because a part of the data is missing or impertinent.

Input variables	output variables
Temperature,C	Ozone,$\mu g/m^3$
Wind direction,degree	
Wind speed, m/s	
Sulfur dioxide SO_2, $\mu g/m^3$	
Carbon monoxide CO, $\mu g/m^3$	
Nitrogen oxide NO, $\mu g/m^3$	
Nitrogen dioxide NO_2, $\mu g/m^3$	
Pressure, hPa	
Relative humidity, %	
Solar radiation, W/m_2	

Table 1. measured variables

4.2 Variables selection

All input variables of Table I do not play a role of equal importance in the ozone process. Only those which play an essential role have to be retained.
The selection criterion we have used consisted in computing the correlation between each input variable and the output (ozone concentration). An example of such correlation is shown by (Fig 1). The input variables that showed the highest correlation with the output turned out to be the following: NO2, NO, Temperature, Wind speed, Humidity, and Solar radiation.

4.3 Premise parameter selection

It has already been mentioned (in section 3.1.1) that the number of inputs is 6 with 2 modalities per input. The parameters of the membership functions have

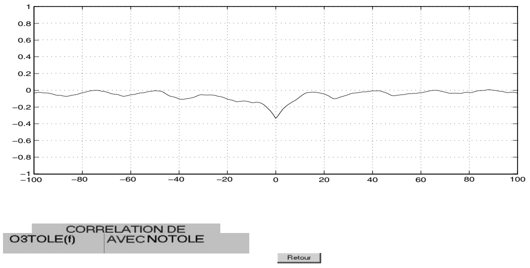

Fig. 1. Correlation between Ozone and NO

been fixed at the following values:
$\bar{z_1} = 0.0883, \quad \bar{z_2} = 0.2220, \quad \bar{z_3} = 0.4613,$
$\bar{z_1} = 0.2592, \quad \bar{z_5} = 0.6909, \quad \bar{z_1} = 0.2216,$
Where the value of $\bar{z}_i$ represents the mean value of the variable z_i; therefore The best distribution of the (normalized) values for each input is guaranteed with $\sigma = 0.3239$.

4.4 *Validation of the identified model*

Many models have been identified with different number of inputs using the proposed approach. The accuracy of models has been evaluated using the criteria of section 3.C. The best model was obtained with the six inputs that have been mentioned. Figure 2 compares the real output (ozone concentration) and the output of the estimated model. It is readily seen that the model tracks well the true process. This is made precisely by criterion (17) which gives $\bar{e} = 8.07\%$. The correlation

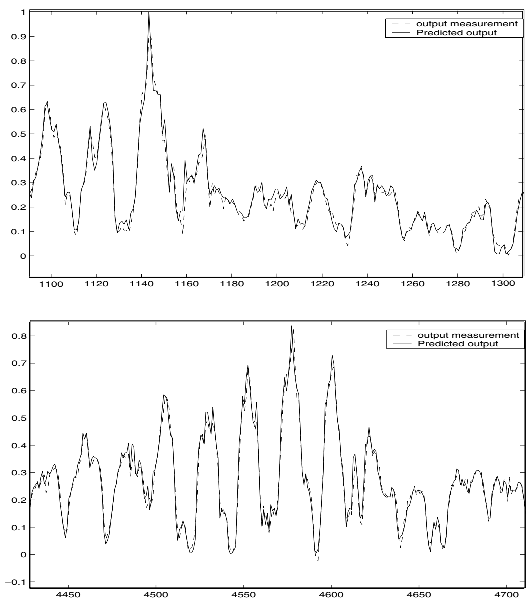

Fig. 2. Parts of the comparison of the multi-model output and the measured ozone

criterion (section 3.3) has confirmed the good quality of the estimated model. Actually, the autocorrelation shape of the estimated error $e(k, \hat{\theta})$ is close to that with white noise (see Figure 3).

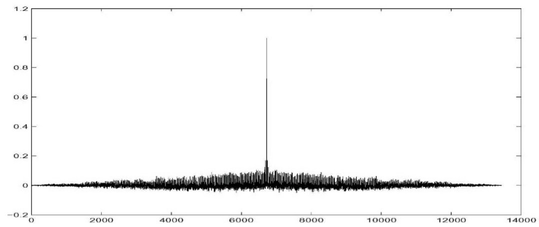

Fig. 3. The autocorrelation of the error $e(k, \hat{\theta})$

5. MODEL IMPROVEMENT

The results of the previous section may be improved in two ways. The first consists in increasing the number of modalities per each input. The second consists in not supposing that the parameter c to be zero.

5.1 *Increasing the number of modalities per input*

As we mentioned in section (3.1.2), two membership functions of the sigmoidal type have been used. The number of rules (or local models) is then 2^n where n is the number of inputs used in this model. A new experiment has been done using three modalities per input, a new membership function of the gaussian type has been added, therefore, the resulting membership functions are the following:

$$\mu_i^1(\tilde{z}_i(k)) = \mu_r^1 = \frac{1 - \tan((\tilde{z}_i(k) - \bar{z}_i)/\sigma}{2}$$
$$\mu_i^2(\tilde{z}_i(k)) = \mu_r^2 = \exp(-((\tilde{z} - \bar{z})/\sigma)^2)$$
$$\mu_i^3(\tilde{z}_i(k)) = \mu_r^3 = \frac{1 + \tan((\tilde{z}_i(k) - \bar{z}_i)/\sigma}{2}$$

where:
$\bar{z}$ is the normalized variable $\tilde{z}$
Model identification has been performed using the approach of section 3, with different number of inputs. For a given number of inputs the model obtained with three modalities turned out to be more precise than that corresponding to two modalities. But it was observed that with three modalities the new model becomes too complex for more than 4 inputs. For example, for 6 inputs, the model involves $3^6 = 729$ local models and $729 \times 5 = 3645$ unknown parameters (the order of these local models is a first order.
Therefore, the number of inputs has been limited to 4 (NO_2, Humidity, Wind speed, Temperature). The resulting model (with 3 modalities) turned out to be about more accurate than the model obtained in section 3 with 6 inputs and two modalities. Actually, the criterion (17) gives for the average estimation error the value $\bar{e} = 7.6\%$ which is slightly less than that obtained in section 3.

5.2 *The case where $\varepsilon(k) = \xi(k) + c_1\xi(k-1)$*

Ozone process identification is now dealt with based on model structure (9), supposing c_1 not to be zero. To this end, notice that equation (9) can be rewritten as follows [1]:

$$y(k)=\sum_{i=1}^{M}\omega_i(k-1)\varphi^T(k-1)\theta_i+c_1\delta(k-1,\Theta^*)+\xi(k) \tag{18}$$

where:

$$\delta(k,\Theta^*)=\frac{1}{1+cq^{-1}}[y(k)-\sum_{i=1}^{M}\omega_i(k-1)\varphi^T(k-1)\theta_i] \tag{19}$$

and,

$$\Theta^*=[\theta_1^T\ \theta_2^T\ ...\ \theta_M^T\ c]^T \tag{20}$$

It can be easily checked using (18) and (19) that $\delta(k,\Theta^*)=\xi(k)$, (for all k). Equation (18) can be given the following pseudo-linear regression form:

$$y(k)=\Phi_1(k-1,\Theta^*)^T\Theta^*+\xi(k) \tag{21}$$

with:

$$\Phi_1(k-1,\Theta^*)=[\Phi(k-1)^T\ \delta(k-1,\Theta^*)] \tag{22}$$

Then, parameter identification is performed with the prediction error algorithm:

$$\hat{\Theta}(k)=\hat{\Theta}(k-1)+\frac{F_1(k-1)\Phi_1(k-1,\hat{\Theta}(k-1))\delta(k,\hat{\Theta}(k-1))}{1+\Phi_1(k-1,\hat{\Theta}(k-1))^TF_1(k-1)\Phi_1(k-1,\hat{\Theta}(k-1))} \tag{23}$$

where, $F_1(k)$ is updated with similar law as (15), replacing F by F_1.
The resulting model turned out to be slightly progressed than the one obtained when supposing c = 0. (Fig 4) shows the autocorrelation of the estimated error $\delta(k,\hat{\Theta})$ where $\hat{\Theta}$ denote the final estimated model provided by (23). This shows that there is no significant improvement with respect to figure 3.

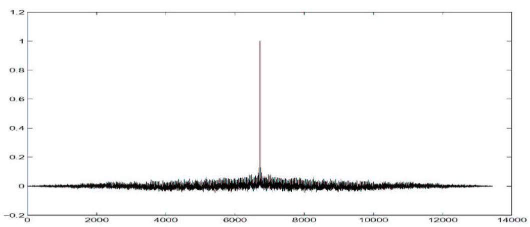

Fig. 4. The autocorrelation of the error $e(k,\hat{\theta})$

6. CONCLUSION

In this paper, a multi-model representation has been designed for the Ozone process using the Takagi-Sugeno fuzzy approach. Such a representation involves local models and rules that describes the relevance of each model (see (1)). The global (multi-model representation) is a weighted average of local models (9). This is achieved by using membership functions of the sigmoidal type (10)-(11). The effect of external distribution has been captured through an equation error of the (MA) type (7)-(8). The number of unknown parameters involved in (9) depends on the number of inputs (see Table I), the number of modalities per input and the order of local models. A quite accurate model for the Basse-Normandie region has been achieved considering six inputs, two modalities per input and first order local models.
The obtained model has been validated using the prediction criterion (17) and the correlation criteria (figure 3). The real data correspond to the measures assembled on the seven years 1995-2001. Then, the change of the structure of the model with four inputs and three modalities has slightly improved the accuracy of the model, where the complexity of both models is about the same. An involved modeling of external disturbances seems not to be crucial comparing figure 3 and figure 4. The results of the present work are currently used to build up a 24 hours predictor for the ozone process.

7. ACKNOWLEDGMENTS

The authors gratefully acknowledge the contribution of AIRCOM, who made available the atmospheric pollution data used in this paper.

REFERENCES

[1] L.Ljung, System identification:Theory of the user. Englewood Cliffs, NJ:Prentic-Hall.(1987)
[2] Sugeno, M. Kang, G.T. Structure identification of a fuzzy model. Fuzzy sets and systems, 28 (1988), 15-33.
[3] Takagi, T. Sugeno, M. . Fuzzy identifications of systems and its application to modelling and control. IEEE. Transactions on systems Man and Cybernetics, 15 (1985), 116-132.
[4] Gilles Mourot, Komi Gasso, Jos Ragot. Modelling of ozone concentrations using a Takagi-Sugeno model. Control Engineering Practice 7 (1999), 707-715
[5] Ioannis C.Ziamos, Dimitrios Melas, Christos S. Zerefos and Alkiviadis F.Bais. Forcasting peak pollutionlevels from meteorological variables. Atmospheric Environment 29(1995)3703-3711.
[6] M. Kolehmainen,H. Martikainen, J. Ruuskanen. Neural networks and periodec components used in air quality forcasting. Atmospheric Environment 35 (2001), 815-825.
[7] J.M. Davis and P. Speckman. A model for predicting maximum and 8h average ozone in Houston. Atmospheric Environment 33 (1999), 2487-2500