Modeling for gas flow measurement consumed by a boiler. Towards a low-cost sensor for energy efficiency.

Baya HADID[1], Régis OUVRARD[1], Laurent LE BRUSQUET [2], Thierry POINOT[1], Erik ETIEN[1], Frédéric SICARD[3]

[1]Laboratoire d'Informatique et d'Automatique pour les Systèmes – EA6315 (LIAS)
Université de Poitiers – ENSIP
2, rue Pierre Brousse, B.P. 633
86022 Poitiers Cedex, France.
baya.hadid@univ-poitiers.fr, regis.ouvrard@univ-poitiers.fr,
erik.etien@univ-poitiers.fr, thierry.poinot@univ-poitiers.fr

[2] Supélec Sciences des Systèmes – EA4454 (E3S)
3, rue Joliot Curie
91192 Gif-sur-Yvette, France
Laurent.Lebrusquet@supelec.fr

[3] EDF - R&D, EPI groupe E25
Avenue des Renardières Ecuelles
77818 Moret sur Loing
frederic.sicard@edf.fr

Abstract—**This paper deals with the modeling of the gas flow supplied to a boiler in order to implement a soft sensor. This study is a part of ANR CHIC project which has an aim to minimize the measuring chain cost in the energy efficiency improvement programs. This implementation requires the estimation of a mathematical model that expresses the flow rate from the control signal of the solenoid valve and the gas pressure and temperature measurements. Two types of models are studied : LPV (Linear Parameter Varying) model with pressure and temperature as scheduling variables and a non-parametric model based on Gaussian processes.**

Keywords—**Soft sensors, Identification, Gaussian process modeling, LPV model, Flow measurement**

I. INTRODUCTION

The concept of energy efficiency is becoming more and more important in the context of high energy demand. The international standard ISO 50001 represents the desire for saving energy. This standard is based on a preliminary energy audit and implementation of systems for measuring and monitoring to ensure that the objectives are achieved.

In the industrial sector, each investment is made in relation to the expected benefits. The cost of a program to improve energy efficiency must be offset by the gained benefits. Sometimes a project, though promising, is rejected on the basis of the amount of initial capital costs, the implementation requiring a production stop. To foster the acceptance of improved energy efficiency programs, production stops must be kept to a minimum and costs of measures need to be low.

It is in this context that the ANR CHIC project (CHaînes de mesures Innovantes à bas Coût) was born. The objective is to develop and to test low-cost sensors to monitor and to analyze the energy consumption of the major fluids used in industrial sites (electricity, gas, compressed air). The studied sensors in the ANR CHIC project should allow monitoring of consumption and drift detection consumption. EDF R&D, the initiators of this project gave the objective of achieving a measurement accuracy of about 5%. The project is to develop new sensors (both physical and "soft") at low cost in the following areas : current sensors, voltage sensors, power sensors, gas flowmeters. The work presented in this paper only concerns the study of gas flow measurement.

The objective of the study presented in this paper is the modeling of a boiler with the aim of developing a "soft" sensor. The concept of soft sensor is to combine measures available or easily achievable and mathematical models which link the measured quantities and the quantities to be determined. This concept is used in various fields and especially in chemical processes [4], [6], [10], or biological processes [3], [5], [8], [12]. The implementation of the soft sensor is based on a simulation, an observer or an inverse method and modeling is a key point for the measurement quality. Modeling can be based on physical principles or empirical approach, or a combination of both.

For the sake of economy, it is desirable that the sensor software can easily be installed. The development of a physical model dedicated to a plant is excluded because it would induce a too high cost of development. For this purpose, it is proposed to build black-box behavioral models. In the case of the gas flow measurement, the dynamic behavior of the signal to be modeled is very fast, consequently, the construction of static models is sufficient regarding the objectives of energy monitoring.

The study focuses on installation of gas boiler with a power of $750kW$, located on the *Renardières* site of EDF R&D near Paris, France. Two approaches are explored. The first one consists of a parametric model and the parameters depend on the pressure and temperature, i.e. an LPV (*Linear Parameter Varying*) model is estimated [7], [13]. The second approach is to estimate a nonparametric model [1].

Developed models allow representing the mass flow of gas in a boiler from the gas pressure, the gas temperature, and the solenoid valve control signal.

II. Boiler and experiments protocol

A schematic representation of the boiler installation and its instrumentation is shown in the figure 1.

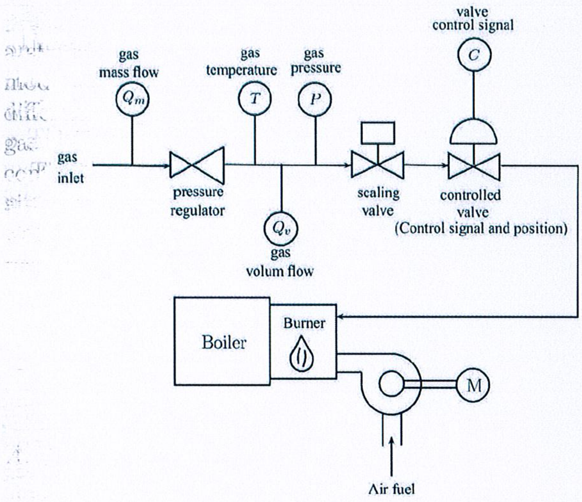

Fig. 1. Schematic view of the boiler

The experiments realized on this installation consist of a stepwise increase of the electrical control signal C of a modulating gas plug valve. These tests are repeated for different values of pressure p and temperature T of the gas. The pressure is an input data which can be practically controlled by a pressure regulator. The gas having a long air routing system, its temperature is thus influenced by weather conditions. These experiments are the same ones realized in operation with the aim of calibration of the proposed models.

Figure 2 shows a typical experiment. Figure 3 lists the operating points used in this study.

III. Modeling

A. Parametric modeling

In this modeling, the primary idea is to consider the mass flow Q_m as an output and C as an exogenous input. However, as can be seen in Figure 4, the pressure p and the temperature T also influence the flow value and a simple law only based on the control signal can not provide a good estimation of the flow. Therefore, we propose an LPV model with one scheduling variable p or two scheduling variables p and T.

The LPV model, like the other models described in this paper, is static.

The estimated LPV model is obtained by a local approach [7], [13] which consists of :

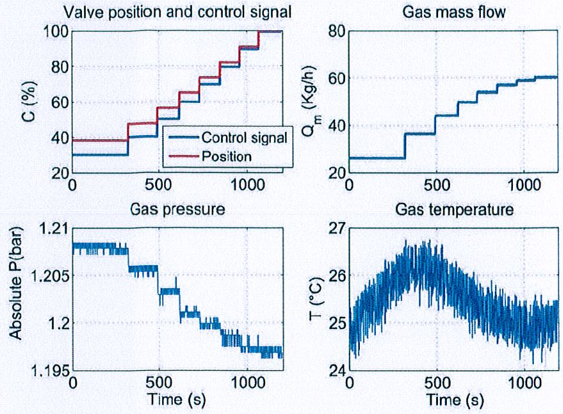

Fig. 2. Typical test at 200 mbar (effective pressure)

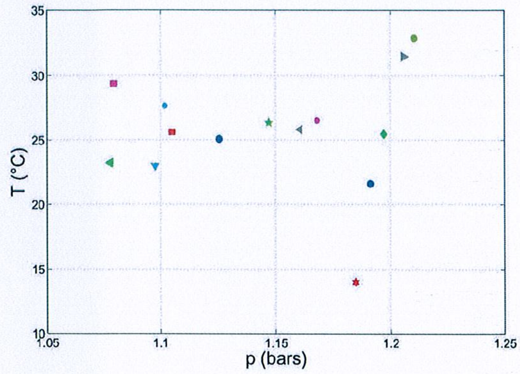

Fig. 3. Used operating points defined by constant values of the gas pressure and temperature

- estimating local models for different operating points of the scheduling variables,
- and calculate the global LPV model by a local models interpolation.

A.1 LPV model with one scheduling variable

Local models are estimated at different operating points defined by a constant pressure and temperature. With regard to the evolution of the gas flow Q_m depending on the control signal C, the chosen model is presented in the following polynomial form :

$$Q_m(t) = \theta_1\, C(t)^2 + \theta_2\, C(t) + \theta_3 \tag{1}$$

The global LPV model as a function of pressure p is determined from the local models. A fixed pressure value is considered throughout the experiment, and is equal to the average of the level corresponding to the highest control signal value. The choice of a fixed value is justified because the pressure varies slightly around a value set by the user via the pressure regulator. It justifies again to consider the pressure as a scheduling variable to fit to different installations. Thus, the parameters θ_1, θ_2 and θ_3 variations depending on the pressure are represented by the following polynomials :

$$\left\{\begin{array}{l} \theta_1 = \alpha_1\, p^2 + \alpha_2\, p + \alpha_3 \\ \theta_2 = \beta_1\, p^2 + \beta_2\, p + \beta_3 \\ \theta_3 = \delta_1\, p^3 + \delta_2\, p^2 + \delta_3\, p + \delta_4 \end{array}\right. \tag{2}$$

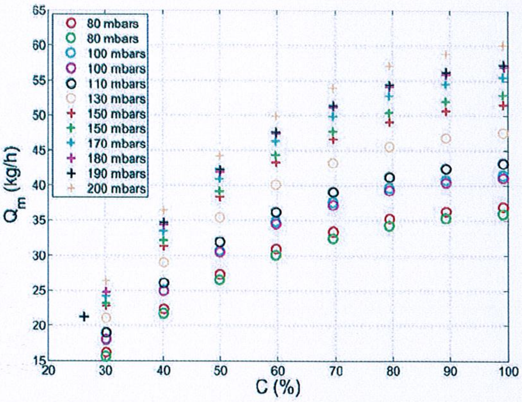

Fig. 4. Flow–control characteristic shape for different gas pressures – measured data

The global LPV model becomes :

$$
\begin{aligned}
Q_m(t) = \quad & (\alpha_1 p^2 + \alpha_2 p + \alpha_3)\ C(t)^2 \\
& + (\beta_1 p^2 + \beta_2 p + \beta_3)\ C(t) \\
& + \delta_1 p^3 + \delta_2 p^2 + \delta_3 p + \delta_4
\end{aligned}
\qquad (3)
$$

A.2 LPV model with two scheduling variables

The considered local models are the same as those given by the equation 1. The global LPV model is still obtained by interpolating the evolution of θ_1, θ_2 and θ_3. The only differences are :

- the average test pressure p is replaced by the instantaneous pressure $p(t)$;
- instantaneous temperature $T(t)$ is also taken into account.

By taking into account the instantaneous measurements and by adding temperature, it is hoped that more accurate estimates than those provided by the first model will be obtained. Each coefficient θ_1, θ_2 and θ_3 is modeled by a polynomial $p(t)$ and $T(t)$ (see section IV-A.2).

B. Non-parametric modeling

Modeling using Gaussian processes is also considered. It is a non-parametric approximation method that aims to build an approximation $\hat{f}$ of the function $Q_m = f(C, p, T)$ from n observations $Q_{m_i} = f(C_i, p_i, T_i)$,$1 \leq i \leq n$ (observations may contain measurement errors), and from *a priori* about the speed variations of the searched function. To simplify notations, note $x_i = (C_i, p_i, T_i)$.

The *a priori* is expressed assuming that the searched function is the realization of a regular random process, in practice a Gaussian process determined by its mean and covariance function. The mean is here taken equal to zero to reflect the absence of *a priori* about a possible tendency of $f(x)$. The covariance function is chosen from a set of parameterized covariance functions family (also called kernels) whose parameters are estimated using the maximum likelihood criterion. We considered that the process was stationary and we chose to model its covariance by a Matérn covariance [11]. This family of covariance was chosen both for its ability to represent a wide range of processes, because its parameters are easily interpretable, and also because it avoids potential numerical problems.

To express the constraint that the searched function $\hat{f}(x)$ is close to the n observations, we search among all the Gaussian process realizations, those that explain the observed points : it is the principle of the modeling with Gaussian process that consists of conditioning of the process law with respect to the observations. The conditioned process is actually a new process with a law, including both the *a priori* (regularity, process variation speed) and the information provided by the observation of the process at some points, can be calculated. This principle is shown in the figure 5 on an example where the variable x is a scalar.

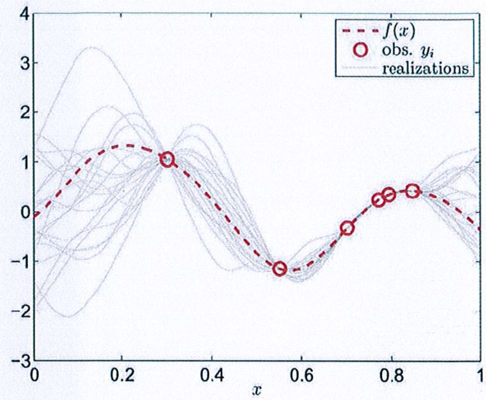

Fig. 5. Example of randomly generated realizations from a Gaussian process with Matérn covariance of parameters ($A = 1, h = 1, \nu = 2$) conditioned to a set of $n = 6$ observed values (assumed here free-noise case)

The estimate $\hat{f}(x)$ commonly used to estimate the function $f(x)$ at one point x is the mean of the process conditioned at this point. The covariance function of the conditioned process enable also to calculate confidence intervals for the function $f(x)$. Figure 6 includes the data of figure 5 (same function $f(x)$, the same abscissa x_i and same observed values) and gives the estimate $\hat{f}(x)$ and the associated confidence intervals.

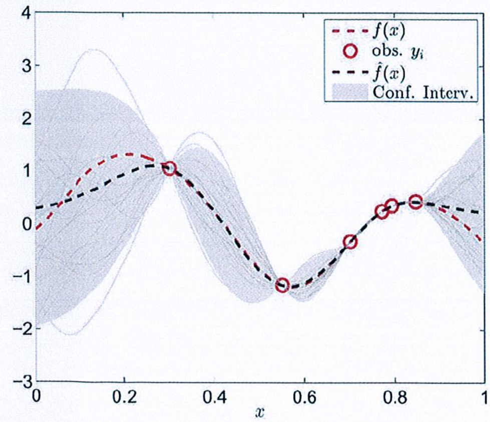

Fig. 6. Illustrative example of the Gaussian process modeling – Red dotted lines : function to estimate, red circles : observations, black dotted line : estimation, shaded area : confidence intervals 95 %

IV. Experimental results

A. Parametric modeling

A.1 LPV model with one scheduling variable

Figure 7 lists the local models defined by (1), estimated with all 14 available experiments. Figures 8 and 9 show the parameters of the different local models based on the operating point of the test. The polynomials defined by (2) allow a good approximation of the estimated values of the parameters θ_1, θ_2 and θ_3, depending on the pressure as it can be seen in Figure 8. Figure 9 shows that it is more difficult to define a mathematical law that fits these points. Initially, it is proposed to use a global LPV model only function of the control signal and the pressure.

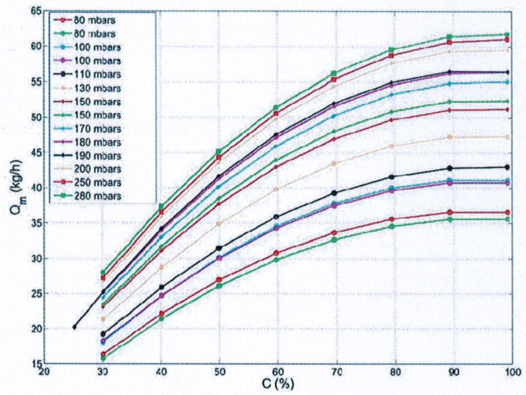

Fig. 7. Local models $Q_m(t) = \theta_1\, C(t)^2 + \theta_2\, C(t) + \theta_3$ for different pressure values p

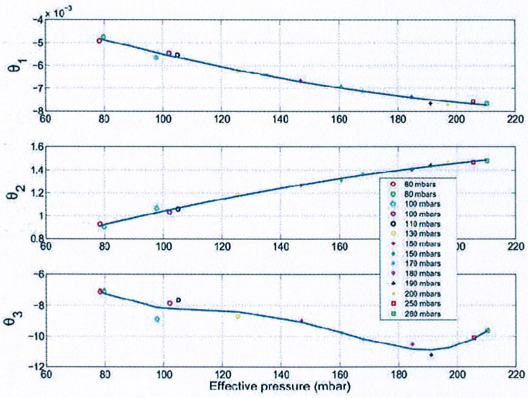

Fig. 8. Evolution of the local parameters models with respect to the pressure p (markers) and the polynomial models (2) (solid line)

The global LPV model can then be used directly as a soft sensor ; for a measured control signal C and a pressure setting, we simply apply the equation (3) to estimate flow gas. The results of models simulation for an experiment are shown in Figure 10 and compared to the real data. The maximum relative error (stepwise averaged) is shown in the figure 11 for all experiment set. Maximum relative error equal to 3.92 % is obtained for a 80 $mbar$ pressure.

A cross-validation was performed to verify the behavior of the virtual sensor for all experimental conditions poten-

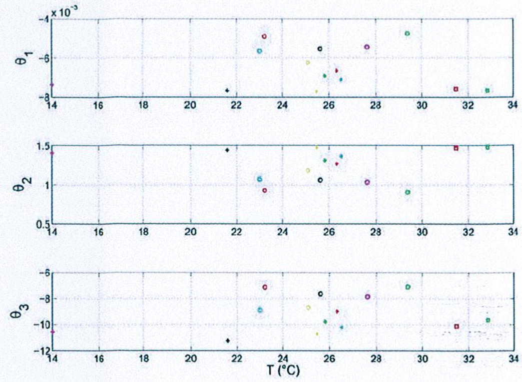

Fig. 9. Evolution of the local parameters models with respect to the temperature T

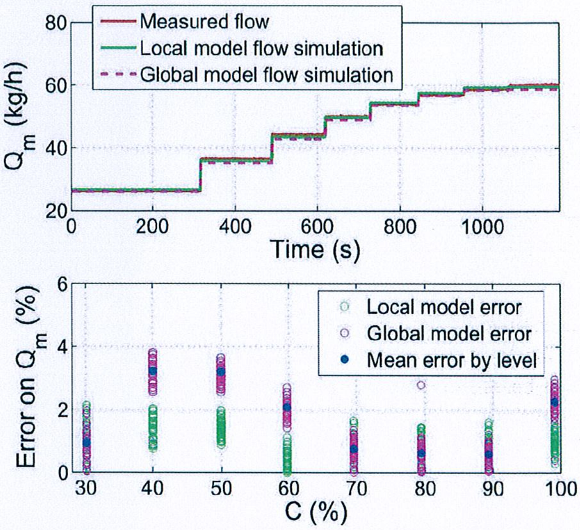

Fig. 10. Local model and one scheduling input global model simulations with 200 $mbar$ experiment data

tially faced in the operating phase. The number of experiments is relatively low (14 experiments) ; we chose to use a Leave–One–Out approach [9]. It consists to use 13 of the 14 tests for identification and one for validation, and repeat this operation so that each test is used as a validation.

The results are shown in Figure 12. The relative errors on the model simulation, estimated using 14 experiments are represented by crosses. The relative errors of the estimated models using 13 experiments and simulated on the validation test are represented by circles. A higher error is noted in validation. Nevertheless, it remains less than 5 % as shown in this figure.

A.2 LPV models with two scheduling variables

The global model is now based on the valve control signal $C(t)$ and the instantaneous measurements of pressure $p(t)$ and temperature $T(t)$. The model is given by :

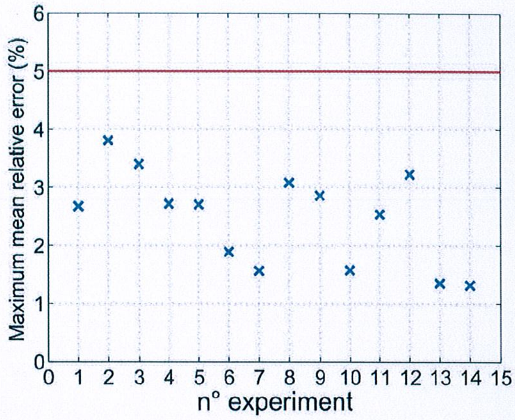

Fig. 11. Maximum relative errors (stepwise averaged) for 14 tests – LPV model with one scheduling variable

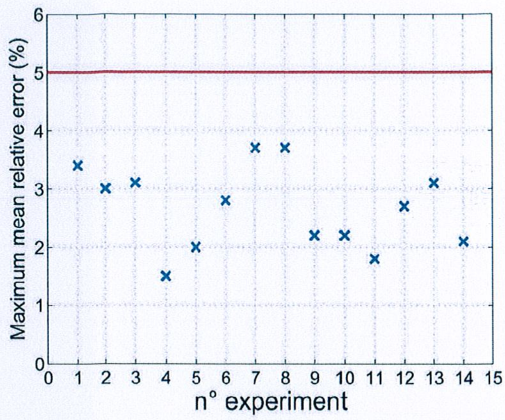

Fig. 13. Maximum relative errors (stepwise averaged) on validation data – LPV model with two scheduling variables

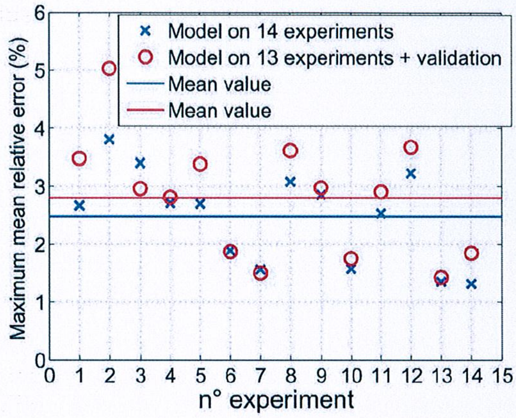

Fig. 12. Comparison between the estimated model using 14 experiments and the models from cross-validation – LPV model with one scheduling variable

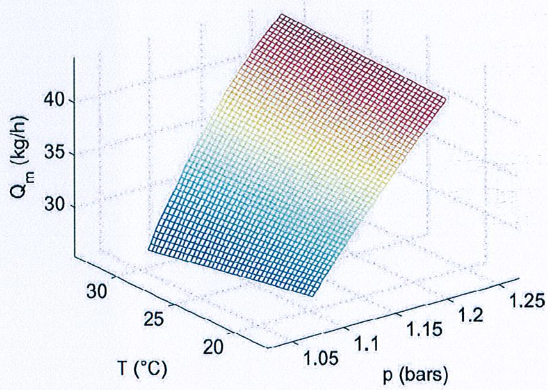

Fig. 14. Simulation of the Q_m estimated model defined by (4) for a fixed control signal at 50 %

$$
\begin{aligned}
Q_m &= \sum_{i=0}^{\mathrm{degP}} \sum_{j=0}^{\mathrm{degT}} \alpha_i\, p^i(t)\, T^j(t)\, C(t)^2 \\
&+ \sum_{i=0}^{\mathrm{degP}} \sum_{j=0}^{\mathrm{degT}} \beta_i\, p^i(t)\, T^j(t)\, C(t) \\
&+ \sum_{i=0}^{\mathrm{degP}} \sum_{j=0}^{\mathrm{degT}} \gamma_i\, p^i(t)\, T^j(t)
\end{aligned} \tag{4}
$$

where degP and degT represent the polynomials degrees of $p(t)$ and $T(t)$.

After several tests, the best results are obtained for degP = 2 and degT = 1. The stepwise maximum relative error in cross-validation are given in Figure 13. The maximum relative error is equal to 3.7 %, i.e. lower than those of the first model. However, it has a higher complexity. Figure 14 presents the simulation of the LPV model obtained for a value of $C = 50\,\%$ and varying pressures and temperatures. We can note that the influence of pressure on the flow variations is higher than the temperature, which may justify the use of a model taking into account only of the pressure.

B. Nonparametric modeling

The implementation of this method on the 14 experiments realized on the boiler was performed using the Matlab toolbox STK (*Small Toolbox for Kriging*) [2]. As for the other simulations, Figure 15 gives the maximum stepwise relative errors obtained by cross-validation. We can note that the errors are lower than the parametric model errors. In addition, the non-parametric model provides reasonable errors without having to specify structure models.

Figure 16 shows the results obtained by cross-validation tests on 4 from the 14 tests. As suggested by the results shown in Figure 15, the predictions are close to the real values. The simulation of the obtained model at variable pressures and temperatures for a 50 % control signal, provides similar results to those presented in Figure 14 .

C. Soft sensor implementation on the industrial boiler

The LPV model with one scheduling variable was experimented on site owing to its simplicity. The model was directly implemented on the PLC with C# programming language. The pressure setpoint has been tuned and fixed to the value read on the manometer during the pressure regulator setting. Figure 17 shows the flow measurement with the soft sensor. As we can see, the results are acceptable with a maximum mean relative error of 3.5 % for the LPV model with one scheduling variable.

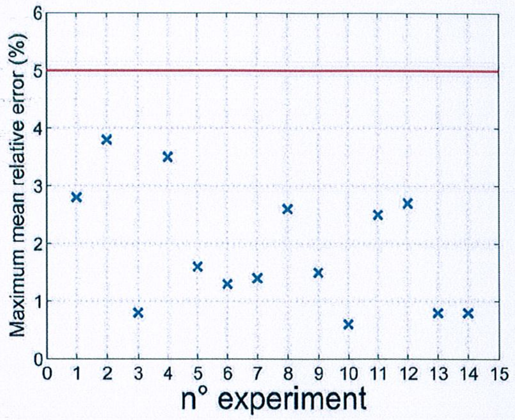

Fig. 15. Maximum relative errors (stepwise averaged) obtained in cross-validation – non-parametric model

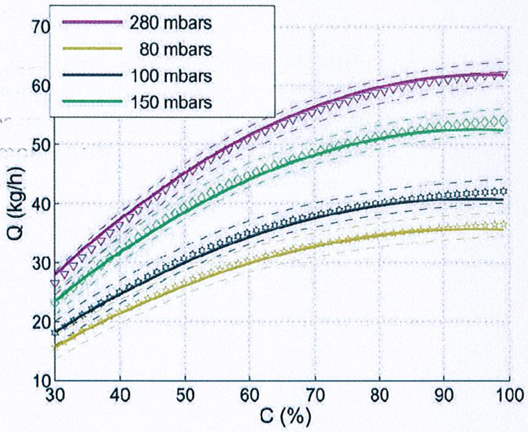

Fig. 16. Cross-validation results for $Q_m = f(C, p, T)$ models with respect to C – Solid lines : identified polynomials ; markers : predictions ; dotted lines : associated confidence intervals

V. Conclusion

In this paper, the modeling of an industrial boiler was investigated towards a consumed gas flow measurement. Three static models were estimated : two LPV parametric models and a nonparametric model. A cross-validation showed that the simulation of these models gives a flow measurement error lower than 5 %. This value corresponds to the fixed objectives of low-cost sensors implementation that allow consumption monitoring and detection of possibles drifts.

Models degraded uses should also be considered. While the online temperature measurement could be considered low-cost, this is not the case of the pressure. But, in practice, the boiler engineer tunes the pressure with the pressure regulator and measures it with manometer.

Finally, the genericity of models to different installations and other kinds of valves, should be studied. It is then necessary to define a model parameters calibration methodology which is the least intrusive.

Acknowledgment

This work has been supported by the French Research National Agency (ANR) through *Efficacité énergétique et réduction des émissions de CO_2 dans les systèmes industriels* program (project CHIC n° ANR-10-EESI-02).

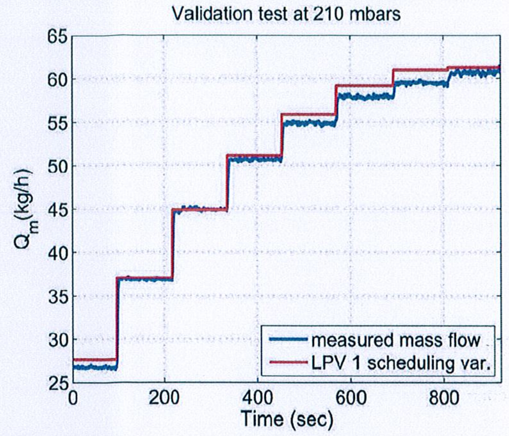

Fig. 17. Experimental simulation of the soft sensor based on the LPV model with one scheduling variable

References

[1] Arkoun O., Estimation non paramétrique pour les modèles autorégressifs. Thèse de doctorat de l'Université de Rouen, novembre 2009.

[2] Bect J. et Vasquez E., A small (Matlab/GNU Octave) Toolbox for Kriging. http ://sourceforge.net/projects/kriging/. Supelec, 2011.

[3] Bogaerts P. et Vande Wouwer A., Software sensors for bioprocesses, ISA Transactions 42, 547-558, 2003.

[4] Cecil D. et Kozlowska M., Software sensors are a real alternative to true sensors. Environmental Modelling & Software 25, 622-625, 2010.

[5] Chéruy A., Software sensors in bioprocess engineering. Journal of Biotechnology 52, 193-199, 1997.

[6] Dochain D., State and parameter estimation in chemical and biochemical processes : a tutorial. J. Proc. Control 13, 801-818, 2003.

[7] Dos Santos P. L., Perdicoúlis T. A., Novara C., Ramos J. et Rivera D., Linear Parameter-Varying System Identification : new developments and trends Advanced Series in Electrical and Computer Engineering. World Scientific, 2011.

[8] James S.C., Legge R.L. et Budman H., On-line estimation in bioreactors : a review. Rev. Chem. Eng. 16 (4), 311-340, 2000.

[9] Kohavi R., A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence 2 (12), 1137-1143, 1995.

[10] Queinnec I. et Spérandio M., Simultaneous estimation of nitrification/denitrification kinetics and influent nitrogen load using orp and do dynamics. ECC, Cambridge, UK, 2003.

[11] Rasmussen C.E. et Williams K.I., Gaussian Processes for Machine Learning. The MIT Press, 2006.

[12] Sotomayor O. A. Z., Won Park S. et Garcia C., Software sensor for on-line estimation of the microbial activity in activated sludge systems. ISA Transactions 41, 127-143, 2002.

[13] Tóth R., Identification and Modeling of Linear Parameter-Varying Systems Springer Verlag. Lecture Notes in Control and Information Sciences 403, 2010.