# Identification of a Gas Furnace Process Using WSDP Model

Nguyen-Vu Truong and Nguyen Tran Hiep

***Abstract*— This paper presents the identification of a gas furnace process using the so called wavelet based State Dependent Parameter (WSDP) models. Here, each of the process' dominant dynamics are characterized by a state dependent parameter in form of wavelet which is able to provide valuable insight, and effective means to measure the system's degree of nonlinearity. To illustrate the advantages of the proposed modeling approach, the identified model is benchmarked with a few welknown reported results, i.e. fuzzy logic, generalised kernel and linear ARX models.**

## I. INTRODUCTION

The gas furnance process data set has been used for a long time as one of a common benchmark for artificial intelligence based modelling approach, i.e. fuzzy logic, generalised kernel models, etc.. This dataset consists of 296 data samples at $Ts = 9s$. The input for this system is Methane input into gas furnace: cu. ft/ min, while the output is carbon dioxide output concentration from gas furnace-% of output gas. In the open literature, several techniques have been reported to approach the modelling of such a process, including (1) Box and Jenkins model [4] in which linear models were assumed, (2) a number of fuzzy logic based approach (i.e. [6]-[8] ), generalised kernel models [16] assuming a certain degree of nonlinear dynamics exist in this process. However, these assumptions are rather misleading since they did not justify the measure and location of the process' nonlinearity / linearity and their performance were often measured by Mean-Squared-Error (MSE).

Our recent publications (i.e. [1]-[3]) have presented approaches to the identification of nonlinear systems using wavelet based *State Dependent Parameter* (SDP) models. This model structure expresses the nonlinear system as a set of the linear regressive output/input terms (states) multiplied by associated *State Dependent Parameters* to characterize the nonlinearities. These state dependencies, in the first step, are non-parametrically estimated using a SDP algorithm based on recursive fixed interval smoothing (i.e. [9]-[15],etc). The shapes of the SDP relationships (as defined by the plots of the parameters against the state variables) indicate and visualize the nature of the most significant nonlinearities within the dynamic SDP model. They are then, in the second step, parameterized in a compact manner via wavelet series expansion by employing appropriate types of wavelet basis functions that are selected corresponding to the features of the SDP relationships. This formulates the wavelet based SDP model (WSDP).

Looking back to this identification problem, this paper presents the use of WSDP models to model this gas furnace process. Here, in the first step, the nonparameterically estimated SDP provides valuable insight, effective measures as well as location of the most dominant nonlinearities. Based on this measure, the model's structure of the process will be identified, combining both nonlinear and linear terms. The final parameterized model will be then obtained using Orthogonal decomposition and PRESS statistics forward regression (i.e. [1]-[3]) which delivers a parsimonious model representation of this benchmark systems, that is the WSDP model excellently characterizes the system's dynamic behaviour in a compactly generalizable manner (12 terms) with a second order model insteading of upto fourth order models as in the previously reported results.

This paper is organized as follows. Background information about WSDP models is provided in Section 2. The simulation and benchmark results are reported in Section 3. Finally, Section 4 concludes this paper.

## II. WAVELET BASED SDP MODEL

A wide range of nonlinear systems can be represented by a nonlinear autoregressive model with exogenous inputs (NARX), as described below:

$$y(k) = f\{\chi(k)\} = f\{y(k-1), ..., y(k-n_y), u(k), ..., u(k-n_u)\} + e(k) \tag{1}$$

where $f\{.\}$ is a nonlinear function (mapping); $u(k)$ and $y(k)$ are, respectively, the sampled input-output sequences; while $\{n_u, n_y\}$ refer to the maximum number of lagged inputs and outputs. Finally, $e(k)$ refers to the noise variable, assumed initially to be a zero-mean, white noise process that is uncorrelated with the input $u(k)$ and its past values.

It is assumed that the above system (1) can be represented by the following *State Dependent Parameter* (SDP) model:

$$y(k) = \sum_{q=1}^{n_y} f_q\{y(k-q)\}y(k-q) + \sum_{q=0}^{n_u} g_q\{u(k-q)\}u(k-q) + e(k) \tag{2}$$

Here, the parameters $f_q\{.\}$, $g_q(.\}$ are the State Dependent Parameters (SDPs) which are respectively functions of the state variables defined by the input and output variables and their past sampled values.

Nguyen-Vu Truong is with the Intelligent Robotics and Automation Laboratory (IRAL), Institute of Applied Mechanics and Informatics, Vietnam Academy of Science and Technology, Vietnam. Email: nvtruong@irobotics.ac.vn / truong.nguyen.vu@gmail.com. Nguyen Tran Hiep is with Department of Automatic Control, Le Qui Don Technical University, Hanoi, Vietnam

978-1-4799-0572-0/13/$31.00 ©2013 IEEE

At this point, the nonlinear system identification problem is equivalent to the problem of estimating and parameterizing the system's nonlinearities characterized by the respective SDPs. More specifically, the nonlinearities carried by these SDPs are first non-parametrically estimated as discussed in the previous sections. In the transformed space (sorted space), the SDP's non-parametric estimate is smoother and less rapid variation. As a result, in order to obtain the compactness in the SDP's parameterization, linear wavelet functional approximation is utilized, in which the respective SDP is represented by a set of scaled and translated wavelet basis functions in combination with a linear function. As proposed and discussed in Section 2.3, this functional approximation scheme is, particularly, advantageous in approximating functions with slow variation features like SDP relationships.

These non-parametric estimates are then used as *a priori* knowledge for the selection of wavelet basis functions and the associated scaling parameters (*i.e.* finest and coarsest scales $i_{\min}$, $i_{\max}$) to be used to effectively parameterize the system's nonlinearities characterized by the respective SDP relationships.

Using the linear wavelet functional approximation as proposed in Section 2.3, the respective SDPs can be generally approximated as follows:

$$f_q\{y(k-q)\} = \sum_{i=i_{\min}}^{i_{\max}} \sum_{j\in L_i} d_{fq,i,j}\Psi_{i,j}[y(k-q)] + a_{fq}[y(k-q)] + b_{fq} \quad (3)$$

$$g_q\{u(k-q)\} = \sum_{i=i_{\min}}^{i_{\max}} \sum_{j\in L_i} d_{gq,i,j}\Psi_{i,j}[u(k-q)] + a_{gq}[u(k-q)] + b_{gq} \quad (4)$$

$$\Psi_{i,j}(x) = \Psi(2^{-i}x - j) \quad (5)$$

in which, $\Psi(x)$, $i_{\min}$,$i_{\max}$ are selected corresponding to the features of the SDP relationships.

Substitute (3) and (4) into (2), we obtain a *wavelet based SDP model* (WSDP) as follows.

$$\begin{aligned} y(k) = &\left\{\sum_{q=1}^{n_y}\left[\sum_{i=i_{\min}}^{i_{\max}} \sum_{j\in L_i} d_{fq,i,j}\Psi_{i,j}[y(k-q)] + a_{fq}[y(k-q)] + b_{fq}\right]\right\}\times \\ &\times y(k-q)+ \\ &\left\{\sum_{q=0}^{n_u}\left[\sum_{i=i_{\min}}^{i_{\max}} \sum_{j\in L_i} d_{gq,i,j}\Psi_{i,j}[u(k-q)] + a_{gq}[u(k-q)] + b_{gq}\right]\right\}\times \\ &\times u(k-q) + e(k) \end{aligned} \quad (6)$$

In this WSDP model, the parameters are the coefficients of the respective wavelet/linear functions, i.e. $d_{fq,i,j}, a_{fq}, b_{fq}$ and $d_{gq,i,j}, a_{gq}, b_{gq}$. With given information of the basis functions (wavelets/linear function) and $\{n_y, n_u\}$ as well as the associated scaling parameters $i_{\min}$ and $i_{\max}$, the next task here is to formulate (6) as an estimation problem of a *linear-in-the-parameter* regression equation, starting from the inner-most summation ($j$) to the outer-most summation ($q$).

Let

$$\begin{cases} \phi_{fq,L_i}(k) = \left[\Psi_{i,j i\min}\{y(k-q)\}, ..., \Psi_{i,j i\max}\{y(k-q)\}\right] \\ A_{fq,L_i} = [d_{fq,i,j_{i\min}}, ..., d_{fq,i,j_{i\max}}]^T \end{cases} \quad (7)$$

$$\begin{cases} \phi_{gq,L_i}(k) = \left[\Psi_{i,j i\min}\{u(k-q)\}, ..., \Psi_{i,j i\max}\{u(k-q)\}\right] \\ A_{gq,L_i} = [d_{gq,i,j_{i\min}}, ..., d_{gq,i,j_{i\max}}]^T \end{cases} \quad (8)$$

where $j_{i\min}$ and $j_{i\max}$ are the lower and upper limits of $L_i$.

By being defined as in (7) and (8), $\phi_{fq,L_i}(k)$ and $\phi_{gq,L_i}(k)$ are functions of $y(k-q)$ and $u(k-q)$ respectively.

Then, (6) can be re-arranged into the following form:

$$\begin{aligned} y(k) = &\sum_{q=1}^{n_y}\left[\sum_{i=i_{\min}}^{i_{\max}} \phi_{fq,L_i}(k)A_{fq,L_i} + a_{fq}[y(k-q)] + b_{fq}\right] y(k-q) \\ &+ \sum_{q=0}^{n_u}\left[\sum_{i=i_{\min}}^{i_{\max}} \phi_{gq,L_i}(k)A_{gq,L_i} + a_{gq}[u(k-q)] + b_{gq}\right] u(k-q) \\ &+ e(k) \end{aligned} \quad (9)$$

Now define,

$$\begin{cases} \Gamma_{fq}(k) = \left[\phi_{fq,L_{i\min}}(k), ..., \phi_{fq,L_{i\max}}(k), y(k-q), 1\right] y(k-q) \\ A_{fq} = \left[A^T_{fq,L_{i\min}}, ..., A^T_{fq,L_{i\max}}, a_{fq}, b_{fq}\right]^T \end{cases} \quad (10)$$

$$\begin{cases} \Gamma_{gq}(k) = \left[\phi_{gq,L_{i\min}}(k), ..., \phi_{gq,L_{i\max}}(k), u(k-q), 1\right] u(k-q) \\ A_{gq} = \left[A^T_{gq,L_{i\min}}, ..., A^T_{gq,L_{i\max}}, a_{gq}, b_{gq}\right]^T \end{cases} \quad (11)$$

We obtain,

$$y(k) = \sum_{q=1}^{n_y}\Gamma_{fq}(k)A_{fq} + \sum_{q=0}^{n_u}\Gamma_{gq}(k)A_{gq} + e(k) \quad (12)$$

Here, $A_{fq}$ and $A_{gq}$ are the parameter vectors. Again, as defined in (10) and (11), $\Gamma_{fq}(k)$ and $\Gamma_{gq}(k)$ are functions of $y(k-q)$ and $u(k-q)$ respectively.

To integrate (12) with measured input and output data, we assume that $y(0), y(1), ..., y(N-1)$ and $u(0), u(1), ..., u(N-1)$ are available.

With

$$Y = [y(0), ..., y(N-1)]^T \quad (13)$$

$$U = [u(0), ..., u(N-1)]^T \quad (14)$$

$$\Xi = [e(0), ..., e(N-1)]^T \quad (15)$$

Equation (12) is written into the matrix form as below

$$Y = \sum_{q=1}^{n_y}\Gamma_{fq}A_{fq} + \sum_{q=0}^{n_u}\Gamma_{gq}A_{gq} + \Xi \quad (16)$$

where,

$$\begin{aligned} \Gamma_{fq} &= \left[\Gamma^T_{fq}(0), ..., \Gamma^T_{fq}(N-1)\right]^T \\ &= \begin{bmatrix} \phi_{fq,L_{i\min}}(0)y(-q) & \dots & \phi_{fq,L_{i\min}}(N-1)y(N-1-q) \\ \vdots & \vdots & \vdots \\ \phi_{fq,L_{i\max}}(0)y(-q) & \dots & \phi_{fq,L_{i\max}}(N-1)y(N-1-q) \\ y(-q)^2 & \dots & y(N-1-q)^2 \\ y(-q) & \dots & y(N-1-q) \end{bmatrix}^T \end{aligned} \quad (17)$$

$$\Gamma_{gq} = \left[\Gamma^T_{gq}(0),...,\Gamma^T_{gq}(N-1)\right]^T = \begin{bmatrix} \phi_{gq,L_{i\min}}(0)u(-q) & \dots & \phi_{gq,L_{i\min}}(N-1)u(N-1-q) \\ \vdots & \vdots & \vdots \\ \phi_{gq,L_{i\max}}(0)u(-q) & \dots & \phi_{gq,L_{i\max}}(N-1)u(N-1-q) \\ u(-q)^2 & \dots & u(N-1-q)^2 \\ u(-q) & \dots & u(N-1-q) \end{bmatrix}^T \tag{18}$$

Let us define

$$\begin{Bmatrix} A_f = \left[A^T_{f1},...,A^T_{fn_y}\right]^T \\ A_g = \left[A^T_{g0},...,A^T_{gn_u}\right]^T \end{Bmatrix} \tag{19}$$

$$\begin{Bmatrix} \Gamma_f = [\Gamma_{f1},...,\Gamma_{fn_y}] \\ \Gamma_g = [\Gamma_{g0},...,\Gamma_{gn_u}] \end{Bmatrix} \tag{20}$$

Here, $A_f$ and $A_g$ are the parameter matrices while $\Gamma_f$ and $\Gamma_g$ are the data matrices.

Substitute (19) and (20) into (16), we obtain

$$Y = \Gamma_f A_f + \Gamma_g A_g + \Xi \tag{21}$$

As a result, (6) can be written in the following matrix form which is a standard least squares parameter estimation:

$$Y = P\theta + \Xi \tag{22}$$

where,

$$\begin{Bmatrix} P = \left[\Gamma_f, \Gamma_g\right] \\ \theta = [A^T_f, A^T_g]^T \end{Bmatrix} \tag{23}$$

## III. Model Structure Identification and PRESS Statistics

The model as described in (22) is often over-parameterized, since it consists of all the possible combinations of regression terms as derived from the selected finest and coarsest scales. With these redundancies, the data matrix is often *numerically ill-conditioned*, leading to a number of disadvantages in both the computation as well as efficiency associated with the parameter estimation. An approach to overcome this is to use *Orthogonal Decomposition (OD)* algorithm (as described in Section 2.4.1). In the proposed approach, it is incorporated into the model structure selection algorithm to enable the algorithm to automatically eliminate any associated *ill-conditioning* problems.

The principle of a model structure determination algorithm lies on the selection of a final model structure which is *simple* but *adequate* to explain the essentials of the underlying system's dynamics. The key here is to justify the significance of each terms within the original over-parameterized model based on a criterion, and determine which term is necessary to be included into the final model.

A well known approach to this problem for a linear-in-the-parameter model is to use the *Predicted Residual Sums of Squares* (PRESS) statistic and forward regression. Here, these methods are used to detect the most significant terms for the optimized SDP parameterization, as discussed in the following.

For this model (22), as described in Section 2.4, its associated *PRESS* value is calculated as

$$PRESS(m) = \sum_{k=0}^{N-1} \xi^2_{-k}(k|m) \tag{24}$$

This value measures the predictive capabilities of the estimated model, thus accesses the model's structure selection problem. If the addition of a term into the model yields a positive increment of the *PRESS* value, it means that term is decreasingly significant for the parameterization, and *vice versa*. Consequently, the *PRESS* statistic can be used as criteria to detect the significance of each term in the model. It is, in fact, the incremental value of *PRESS* resulted from excluding a term from the model that reflects its contribution in the model. If we let $PRESS^{-i}_m(m-1)$ be the value calculated from a $(m-1)$-term model by excluding the $i^{th}$ term from the original $m$-term model, then $\Delta PRESS_i = PRESS^{-i}_m(m-1) - PRESS(m)$ can be used as criterion for the *new term selection algorithm.*

Note that traditional approaches that use the *PRESS* statistic in term selection are based on the so called '*growing model*' concept . This normally starts from an initial small term subset and gradually adds in new terms based on the reduction in the *PRESS* value caused from adding these terms. But, in this case, how can we ensure that the selected initial subset contains the most significant terms? Otherwise, it might easily lead to model over-parameterization.

In order to avoid the possibility of such over-parameterization, in our new algorithm, we first detect the significance and contribution of each term in the model, based on the incremental value $\Delta PRESS_i = PRESS^{-i}_m(m-1) - PRESS(m)$. In this way, the maximum $\Delta PRESS_i$ signifies the most significant term, while its minimum reflects the least significant term. Based on this, in the next step, forward regression is employed to select the system's model structure. By doing so, we can ensure that the algorithm initializes with the initial subset being the most significant term. It then starts to grow to include the subsequent significant terms in a forward regression manner, until a specified performance is achieved. To be more specific, the forward regression-based *PRESS* term selection algorithm is described below.

### A. The PRESS Term Selection Algorithm

For the ease of representation, let us denote $\phi_i$ be the $(i+1)^{th}$ column of $\Phi$: $\phi_i = \Phi(:, i+1)$, and $P^{(-i)}$ denotes the matrix which is resulted from excluding the $i^{th}$ column from the original matrix $P$.

1) Initialize $\Phi = P$, $[N, m] = size(P)$
2) Orthogonal Decomposition
   a) $[N, m_1] = size(\Phi)$. Initialize $\omega_0 = \phi_0$, $g_0 = \frac{\omega_0^T Y}{\omega_0^T \omega_0}$.

b) For $1 \le i \le m_1 - 1$, compute

$$\alpha_{j,i} = \frac{\omega_j^T \phi_i}{\omega_j^T \omega_j}, \; j = 0, 1, ..., i-1$$
$$\omega_i = \phi_i - \sum_{j=0}^{i-1} \alpha_{j,i} \omega_j$$
$$g_i = \frac{\omega_i^T Y}{\omega_i^T \omega_i}$$

3) PRESS computation

$$\xi_{-k}(k) = \frac{y(k) - \sum_{i=0}^{m_1-1} \omega_i(k) g_i}{1 - \sum_{i=0}^{m_1-1} \frac{\omega_i(k)^2}{\|\omega_i\|^2}}$$
$$PRESS = \sum_{k=0}^{N-1} \xi_{-k}^2(k)$$

4) $PRESS(m) = PRESS$. For $1 \le i_1 \le m$,
   a) Set $\Phi = P^{(-i_1)}$. Repeat steps 2 and 3.
   b) $PRESS_m^{-i_1}(m-1) = PRESS$. Calculate

$$\Delta PRESS_{i_1} = PRESS_m^{-i_1}(m-1) - PRESS(m)$$

5) Based on the largest $\Delta PRESS_{i_1}$ value, select the most significant term to be added to the regressor matrix.
6) Solve for the intermediate parameter estimate in a least squares manner.
7) Calculate the approximation accuracy, and compare it to the desired value:
   - If satisfactory performance is achieved, stop the algorithm;
   - Otherwise, add extra terms into the regressor matrix based on the next largest $\Delta PRESS_{i_1}$ values, and repeat from step 6 to 7.

### B. Nonlinear System Parameter Estimation

Even though the model parameter estimate can be obtained as a by-product of the above described model structure selection algorithm, to facilitate the understanding and support the presentation of the subsequent chapters, in the following, we formulate a standard least squares parameter estimation framework.

Upon completing the above model structure selection procedure, the optimized functional structures for all the SDPs are revealed. They are defined in the following manner:

$$f_q(x) = \sum_{j=0}^{n_{f_q}} a_{fq,j} l_{fq,j}(x)$$
$$g_q(x) = \sum_{j=0}^{n_{g_q}} a_{gq,j} l_{gq,j}(x) \quad (25)$$

where, $L_{f_q} = \{l_{fq,0}, ..., l_{fq,n_{fq}}\}$, $L_{g_q} = \left\{l_{gq,0}, ..., l_{gq,n_{g_q}}\right\}$ are, respectively, the optimized sets of wavelet functions and/or the linear function $(ax+b)$ used for parameterization of $f_q(x)$ and $g_q(x)$. Substituting (25) into (2) yields

$$y(k) = \sum_{q=1}^{n_y} \sum_{j=0}^{n_{f_q}} [a_{fq,j} l_{fq,j}\{y(k-q)\}] y(k-q) + \sum_{q=0}^{n_u} \sum_{j=0}^{n_{g_q}} [a_{gq,j} l_{gq,j}\{u(k-q)\}] u(k-q) + e(k) \quad (26)$$

Let us define

$$\theta_{fq} = [a_{fq,0}, ..., a_{fq,n_{f_q}}]^T$$
$$\theta_{gq} = [a_{gq,0}, ..., a_{gq,n_{g_q}}]^T$$
$$L_{k,fq} = [l_{fq,0}\{y(k-q)\}, ..., l_{fq,n_{fq}}\{y(k-q)\}] y(k-q)$$
$$L_{k,gq} = [l_{gq,0}\{u(k-q)\}, ..., l_{gq,n_{g_q}}\{u(k-q)\}] u(k-q) \quad (27)$$

Equation (26) can be rewritten in to the following form:

$$y(k) = \sum_{q=1}^{n_y} L_{k,fq} \theta_{fq} + \sum_{q=0}^{n_u} L_{k,gq} \theta_{gq} + e(k) \quad (28)$$

Let

$$\theta = \left[\theta_{f1}^T, ..., \theta_{fn_y}^T, \theta_{g0}^T, ..., \theta_{gn_u}^T\right]^T$$
$$L_k = [L_{kf1}, ..., L_{kfn_y}, L_{kg0}, ..., L_{kgn_u}]^T \quad (29)$$

By substituting (29) into (28), we obtain:

$$y(k) = L_k^T \theta + e(k) \quad (30)$$

Assume

$$Y = [y(0), ..., y(N-1)]^T$$
$$U = [u(0), ..., u(N-1)]^T$$

to be the measurement output-input data.

Equation (30) is written into the following matrix form:

$$Y = L\theta + \Xi \quad (31)$$

in which,

$$L = [L_0, ..., L_{N-1}]^T$$
$$\Xi = [e(0), ..., e(N-1)]^T \quad (32)$$

Via Orthogonal Decomposition, $L$ is orthogonally decomposed into 2 components: $m_L \times m_L$ upper triangular matrix $T$, and $N \times m_L$ matrix $W$ with orthogonal columns of $\omega_i$,

$$L = WT \quad (33)$$

Doing so, the estimate $\hat{\theta}$ of the parameter vector $\theta$ is obtained in an Orthogonal Least Squares manner, *i.e.*

$$\hat{\theta} = T^{-1} \hat{G} \quad (34)$$

in which,

$$\hat{G} = diag[\frac{1}{\omega_0^T \omega_0}, ..., \frac{1}{\omega_i^T \omega_i}, ..., \frac{1}{\omega_{m_L-1}^T \omega_{m_L-1}}] \left[W^T Y\right] \quad (35)$$

This linear least squares estimate will have optimal statistical properties if $e(k)$ is a zero-mean, normally distributed, white noise process, independent of the input signal $u(k)$. However, depending on the nature of the data and the SDP model, these assumptions may not be applicable. In this situation, other estimation solutions are necessary: for instance, *Instrumental Variable* (IV); nonlinear least squares based on a response error function; or maximum likelihood estimation based on prediction errors, *etc.* In particular, the standard and optimal IV approaches, which have proven so useful in the linear model estimation context, are very robust in practical application, can be developed for use in IV estimation of the parameters in this model setting.

### C. Identification Procedure

The overall nonlinear system identification using the proposed approach can be summarized into the following steps:

1) *Determining the SDP model's order.* This includes the selection of the initial values[1] of $n_y$ and $n_u$.
2) *Non-parametrically estimating the associated SDP parameters.*
3) *SDP's optimized parameterization structure selection.* This involves the following steps:
   a) Based on the features of the respective SDP non-parametric estimate, determine an appropriate wavelet function and the associated scaling parameters (*i.e.* finest and coarsest scaling parameters) to be used for the parameterization.
   b) Using the *PRESS* based selection algorithm, determine an optimized functional structure used for the respective SDP's parameterization.
4) *Final parametric optimization.*
   a) Substitute the optimized SDP functional structures into the original SDP model.
   b) Using the measured data, estimate the associated parameters via an Orthogonal Least Squares algorithm.
5) *Model validation.*
   - If the identified values of $n_y$ and $n_u$ as selected in step 1 provide a satisfactory performance over the considered data, terminate the procedure.
   - Otherwise, increase the model's order, *i.e.* $n_y = n_y + 1$ and/or $n_u = n_u + 1$, and repeat steps 2 through 5.

## IV. THE CHOISE OF WAVELET BASIS FUNCTION

It is obvious that the choice of the wavelet basis function in the linear wavelet functional approximation may well be different for each SDP. In order to simplify the general procedure, therefore, in this chapter, we use a simple form of the radial mother wavelet called the *Mexican Hat Wavelet* as described in (36) which is compactly supported in $(-4,4)$. Of course, there may be cases where the SDP relationships are more volatile and rich in frequency features than allowed by this mother wavelet and then a more complicated basis function may be necessary ( for example, the Morlet wavelet or other well known wavelet forms). The Mexican hat wavelet takes the following form:

$$\Psi(x) = \begin{Bmatrix} (1-x^2)e^{-0.5x^2} & \text{if } x \in (-4,4) \\ 0 & \text{otherwise} \end{Bmatrix} \quad (36)$$

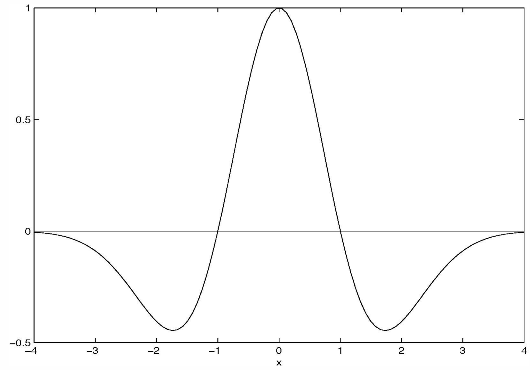

Fig. 1. Compactly supported Mexican hat wavelet function.

## V. RESULTS

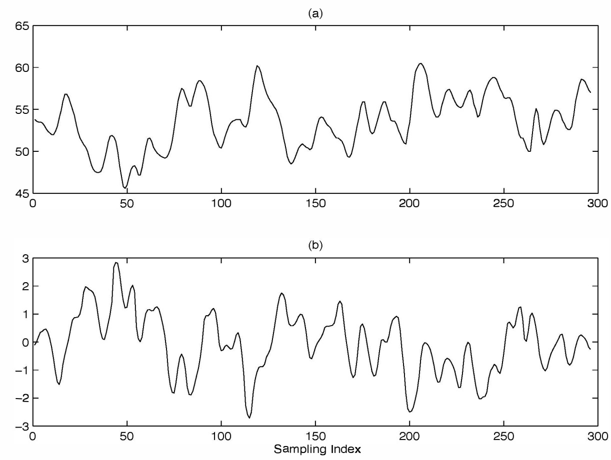


Fig. 2. (a) output, (b) input.

Consider input-output data from a gas furnace process, sampled at $Ts = 9s$. The input for this system is Methane input into gas furnace: cu. ft/ min, while the output is carbon dioxide output concentration from gas furnace-% of output gas. This 296-input-output data set (Figure 2) is separated into: estimation data set consisting of 200 data points, and the remaining 96 data points for model testing. For the convenience of implementation,the recorded output data was standardized: $y_d(k) = \{y(k) - mean(y)\}/std(y)$, and the standardized output sequence is still designated as $y(k)$.

[1]Which normally start with lower values.

Using discrete model for this system, the SDP model structure is identified as below

$$y(k) = f_1\{y(k-1)\}y(k-1) + f_2\{y(k-2)\}y(k-2) + g_0\{u(k)\}u(k) + g_2\{u(k-2)\}u(k-2) \quad (37)$$

The non-parametrically estimated dependencies for this model are shown in Figure 5 where all SDPs are modelled as Integrated Random Walk (IRW) processes. Carrying out the similar process as the above, the final parametric model is found to be

$$\begin{aligned} y(k) = & \begin{bmatrix} -0.0336\Psi_{2,-1}(x) - 0.1314\Psi_{2,0}(x) \\ +1.4561 \end{bmatrix}_{y(k-1)} y(k-1) \\ & + \begin{bmatrix} 0.7572\Psi_{2,-1}(x) + 0.45281\Psi_{2,0}(x) \\ +1.0231\Psi_{3,1}(x) + 2.2873 \end{bmatrix}_{y(k-2)} y(k-2) \\ & + \begin{bmatrix} 0.0627\Psi_{2,0}(x) + 0.0110\Psi_{2,-1}(x) \\ +0.0408\Psi_{2,-2}(x) - 0.0083 \end{bmatrix}_{u(k)} u(k) \\ & - 0.1837u(k-2) \end{aligned} \quad (38)$$

where,

$$\Psi_{i,j}(x) = \Psi(2^{-i}x - j) \text{ and } \Psi(x) = (1-x^2)e^{-0.5x^2}$$

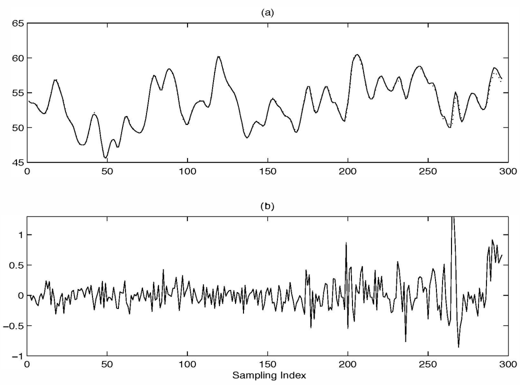


Fig. 3. (a) Comparison between the actual output (solid) and model (38) prediction (dot-dot) which are very well overlapped, and (b) the associated residual.

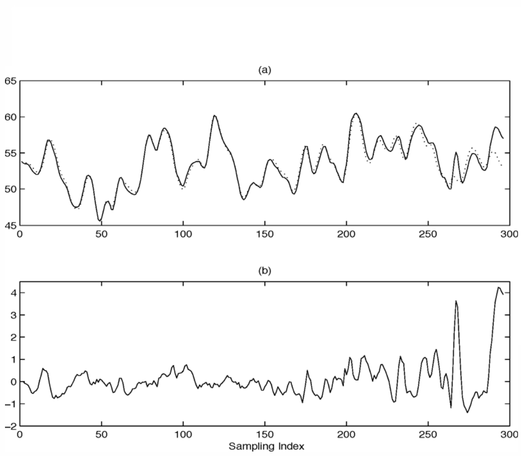


Fig. 4. (a) Comparison between the actual output (solid) and model iterative output of (38) (dot-dot), and (b) their difference

Figure 3 compares the model prediction (which is recovered to the original amplitude by de-standardization) to the actual output over the whole data. The model's iterative (simulated) output[2] is shown in Figure 4 in comparison to the actual output signal, implying that the 12-term identified model excellently characterizes the system's dynamic behaviour. In fact, this application example was also studied in [16] (Example 2), where generalized Kernel model is used to model the system. Nevertheless, in comparison to the work in the present study, the proposed approach may be more advantageous over the generalized kernel model [16] for this particular example, in the sense that the system is well represented with a smaller number of terms (12 versus 21, or 43% term's reduction), smaller order (2 versus 3) and using less training data (200 versus 296) while producing much better result ($MSE = 0.0285$ versus 0.053482 of the generalized kernel model [16] over the considered data set). In comparison with other approaches, i.e. one of the best results of fuzzy based systems as reported in [6], the MSE is 0.068, third order, linear models at fourth order delivers MSE at 0.061 [5]. This, again, confirms the efficiency of the proposed approach.

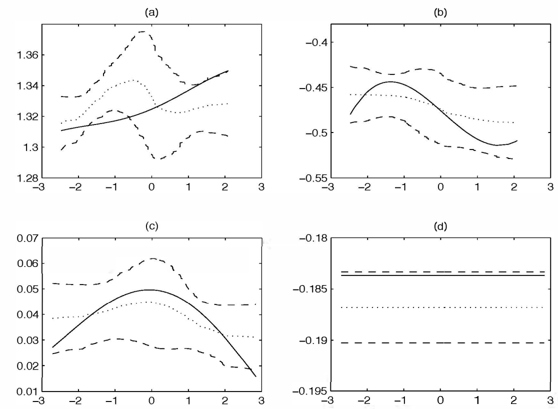


Fig. 5. Non-parametrically estimated (dot-dot), and actual SDPs (solid): (a) $f_1\{y(k-1)\}$ versus $y(k-1)$ (b) $f_2\{y(k-2)\}$ versus $y(k-2)$ (c) $g_0\{u(k)\}$ versus $u(k)$, (d) $g_2\{u(k-2)\}$ versus $u(k-2)$ , and standard error bound (dash).

## VI. Conclusions

Benchmark is vital to the validation of various identification techniques. In this paper, WSDP model is applied in the study of a commonly used benchmark: the gas furnace process data set. It has been demonstrated that WSDP models are able to provide a effective analytical insights, efficient measures of the nature and location of the nonlinear dynamics within this systems.

## References

[1] N.V. Truong, L.Wang and P.C. Young, " Nonlinear system modeling using non-parametric identification and linear wavelet estimation of SDP models", *Proc. 45th IEEE Conf. on Decision and Control*, San Diego, US, pp. 2523-2528, 2006.

[2] *i.e.* the output obtained by generating the deterministic model output from the model input alone, without any reference to the output measurements

[2] N.V. Truong, L. Wang and P.C. Young, "Nonlinear system modeling using non-parametric identification and linear wavelet estimation of SDP models", *International Journal of Control*, Vol. 80, No. 5, pp. 774-788, 2007.

[3] N.V. Truong, L. Wang and J.M. Huang, "Nonlinear modeling of a magnetic bearing using SDP model and linear wavelet parameterization", *Proc. 2007 American Control Conference*, New York, USA, pp. 2254-2259, 2007.

[4] G. E. P. Box and G.M. Jenkins, Time Series Analysis: Forecasting and Control, *2nd ed. San Francisco: Holden-Day*, 1970.

[5] W. Farag and A. Tawfik, " On Fuzzy Model Identification and the Gass Furnace Data", in Proc. IASTED International Conf. on Intelligent Systems and Control, Hawaii, USA, 2000, pp. 1-5.

[6] M. Sugeno and K. Tanaka, "Successive Identification of a Fuzzy Model and its Application to Prediction of a Complex System", *Fuzzy Sets Syst.*, 42, 1991, 315-334.

[7] M. Sugeno, and T. Yasukawa, "A Fuzzy-Logic-Based Approach to Qualitative Modeling", I*EEE Trans. on Fuzzy Systems*, 1(1), 1993, 7-31.

[8] L. Wang and R. Langari, "Complex Systems Modeling via Fuzzy Logic", *IEEE Trans. on Syst., Man, and Cybernetics*, 26(1), 1996, 100-106.

[9] P.C. Young, "Time variable and state dependent modelling of nonstationary and nonlinear time series", in T.S. Rao (ed), *Developments in Time Series, volume in honour of Maurice Priestley*, Chapman and Hall: London, 374-413 , 1993.

[10] P.C. Young, "Data-based mechanistic modelling of environmental, ecological, economic and engineering systems", *Journal of Environmental Modelling and Software*, Vol. 13, pp. 105-122, 1998.

[11] P.C. Young, "Data-based mechanistic modelling of engineering systems", *Journal in Vibration and Control*, Vol. 4, pp. 5-28, 1998.

[12] P.C. Young, "Top-down and data-based mechanistic modelling of rainfall-flow dynamics at the catchment scale", *Hydrological Processes*, **17**, 2195-2217, 2003.

[13] P.C. Young, "Stochastic, dynamic modelling and signal processing: time variable and state dependent parameter estimation". In Fitzgerald, W. J., Walden, A., Smith, R., and Young, P. C., (Eds.), *Nonlinear and Nonstationary Signal Processing*, Cambridge University Press: Cambridge, 74-114, 2000.

[14] P.C. Young, "The identification and estimation of nonlinear stochastic systems", In *Nonlinear Dynamics and Statistics* (A.I. Mees ed.), Birkhauser: Boston, 127-166, 2001.

[15] P.C. Young, P. McKenna and J. Bruun "Identification of nonlinear stochastic systems by state dependent parameter estimation", *International Journal of Control*, Vol. 74, pp. 1837-1857, 2001.

[16] S. Chen, X. Hong, C.J. Harris, and X. Wang, "Identification of nonlinear system using generalized kernel models", *IEEE Trans. on Control Systems Technology*, Vol. 13, No. 3, pp. 401-411, 2005.