# A new approach for identification of MIMO non linear system with RKHS model

TAOUALI OKBA, ELAISSI ILYES, GARNA TAREK AND MESSAOUD HASSANI
Unité de Recherche ATSI
Ecole Nationale d'Ingénieurs de Monastir
Rue Ibn El Jazzar 5019 Monastir ; Tel : +(216) 73 500511, Fax : +(216) 73 500 514,
Tunisia
taoualiokba@yahoo.fr, ilyes.elaissi@yahoo.fr, tarek.garna@enim.rnu.tn,
hassani.messaoud@enim.rnu.tn

*Abstract:* - In this paper we propose a new approach for the modelling of the multi-variable systems (MIMO) on the Reproducing Kernel Hilbert Space (RKHS). The proposed approach considers the MIMO system as a set of MISO processes modelled in RKHS space. We propose also a comparative study of three identification kernel methods of nonlinear systems modelled in Reproducing Kernel Hilbert Space (RKHS), where the model output results from a linear combination of kernel functions. Theses methods are support vector machines (SVM), regularization networks (RN) and kernel Principal Component Analysis (KPCA). The performances of the proposed MIMO RKHS model and of each kernel method in terms of generalization ability and computing time were evaluated on numerical simulations.

*Key-Words:* - Identification, RKHS, SLT, MIMO model, SVM, RN, KPCA

## 1 Introduction

In supervised learning- or learning-from- examples-a machine is trained, instead of programmed, to perform a given task on a number of input-output pairs. According to this paradigm, training means choosing a function which best describes the relation between the inputs and the outputs. The central question of Statistical learning theory SLT [1], [2] is how well the chosen function generalizes, or how well it estimates the output for previously unseen inputs. The model provide is known a RKHS model (Reproducing Kernel Hilbert Space) as a linear combination of the kernels function forming the RKHS space [6], [7], [12], [13], [14] and [24].

The developed models are an attractive alternative to other modeling techniques based on Volterra series, neural networks,…. Indeed, the solution of optimization problem in space RKHS is a global minimum contrary to that provided by neural networks. The solution is obtained by solving a quadratic optimization problem by using the learning algorithms such as support vector machines (SVM) [1], regularization networks (RN) [2] and Kernel Analysis Principal Component KPCA [9]. These algorithms known as kernel methods construct RKHS models on the principle of structural risk minimization (SRM). The number of parameters of these models depends only on the number of observations and not on the structure model as in the conventional modeling approaches.

The paper is organized as follows. In section 2 we remind the presentation of the RKHS space. Section 3 is devoted to the modelling in RKHS. In the section 4 we propose the new approach of MIMO RKHS model. The SVM, RN and KPCA methods are presented in the section 5 and then tested to identify a benchmark in the section 6 [11].

## 2 Reproducing Kernel Hilbert Space

Let $E \subset \mathbb{R}^d$ an input space and $L^2(E)$ the Hilbert space of square integrable functions defined on $E$. Let $k : E \times E \to \mathbb{R}$ be a continuous positive definite kernel. It is
proved [15] that it exists a sequence of an orthonormal eigen functions $(\psi_1, \psi_2, ..., \psi_l)$ in $L^2(E)$ (where $l$ can be infinite) and a sequence of corresponding real positive eigenvalues $(\sigma_1, \sigma_2, ..., \sigma_l)$ so that the kernel $k$ is defined as:

$$k(x,\ t)=\sum_{j=1}^{l}\sigma_j\ \psi_j(x)\ \psi_j(t)\ \ ;\ \ x,\ t\in E \qquad (1)$$

Let $F_k \subset L^2(E)$ be a Hilbert space associated to the kernel $k$ and defined by:

$$F_k=\left\{ f\in L^2(E)/\ f=\sum_{i=1}^{l} w_i\varphi_i \ \ and\ \ \sum_{j=1}^{l}\frac{{w_j}^2}{\sigma_j}<+\infty \right\} \qquad (2)$$

Where $\varphi_i=\sqrt{\sigma_i}\,\psi_i$ ; $i=1,...,l$. The scalar product in the space $F_k$ is given by:

$$<f,\ g>_{F_k}=<\sum_{i=1}^{l} w_i\ \varphi_i,\ \sum_{j=1}^{l} z_j\ \varphi_j>_{F_k}=\ \sum_{i=1}^{l} w_i\ z_i \qquad (3)$$

The kernel $k$ is said to be a reproducing kernel of the Hilbert space $F_k$ if and only if the following conditions are satisfied.

$$\begin{cases} \forall\, x\in E,\ \ k(x,.)\in F_k \\ \forall\ x\in E \text{ and } \forall\ f\in F_k\ , <f(.),\ k(x,.)>_{F_k}=f(x) \end{cases} \qquad (4)$$

Where $k(x,.)$ means $k(x,\ x')\ \ \forall\ x'\in\ E$

$F_k$ is called reproducing kernel Hilbert space (RKHS) with kernel $k$ and dimension $l$. Moreover, for any RKHS, there exists only one positive definite kernel and vice versa [3].

Among the possible reproducing kernels, we mention the Radial Basis function (RBF) defined as:

$$k(x,\ t)=\exp\left(-\left\|x-t\right\|^2/2\mu^2\right)\ \ ;\ \forall\ x,\ t\ \in\ E \qquad (5)$$

with $\mu$ a fixed parameter.

Let's define the application $\Phi$ :

$$\Phi: E\to\mathbb{R}^l$$

$$x\mapsto\Phi(x)=\begin{pmatrix}\varphi_1(x)\\ .\\ .\\ .\\ \varphi_l(x)\end{pmatrix} \qquad (6)$$

Where $\varphi_i$ are given in (2). The kernel trick [8] is so that:

$$k(x,\ x')=\ <\Phi(x),\ \Phi(x')>\ \ x,\ x'\in E \qquad (7)$$

# 3 Modeling in RKHS

The use of RKHS models in identifying systems giving only the observations of their behaviour is of a great interest. Consider a set of observations $D=\left\{x^{(i)},y^{(i)}\right\}_{i=1}^{N}$, known as learning set, with $x^{(i)}\in\mathbb{R}^n$, $y^{(i)}\in\mathbb{R}$ are respectively the system input and output. According to the statistical learning theory (SLT) [1], the identification problem in the RKHS $F_k$ can be formulated as a minimization of the regularized empirical risk which consists in finding the function $f^*\in F_k$ such that:

$$\begin{aligned} f^* &= \sum_{j=1}^{l} w_j^*\varphi_j \\ &= \arg\min_{f\in F_k}\ \frac{1}{N}\ \sum_{i=1}^{N} V\left(y^{(i)},f\left(x^{(i)}\right)\right)^2+\lambda\left\|f\right\|_{F_k}^{\ 2} \end{aligned} \qquad (8)$$

Where $\lambda$ is a regularization parameter chosen in order to ensure a generalization ability to the solution $f^*$ and $V(.)$ is a cost function. According to the representer theorem [5], the solution of the problem (8) is a combination of $M$ kernel functions applied on the $M$ observations.

$$f^*(x)=\sum_{i=1}^{N} a_i\ k\left(x^{(i)},\ x\right) \qquad (9)$$

The number of parameters is equal to the observation number used in the learning phase and contained in the learning set.

# 4 MIMO RKHS model

In the MIMO case the process output is a $p$-dimensional vector, we consider the network of kernel functions illustrated by Figure 1

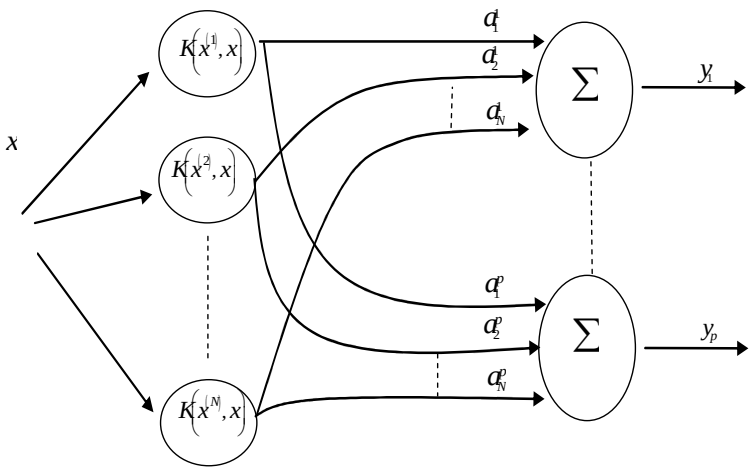


Fig1. Network of kernels functions for the MIMO modelling

The MIMO process is considered as a set of MISO processes modelled in RKHS space as above. To decrease the model complexity, all the MISO output are linear combinations of the same kernel components and with different parameters.

The output of the $q^{th}$ MISO model is:

$$y_q = \sum_{i=1}^{N} a_i^q K\left(x^{(i)}, x\right) \text{ for } q = 1, ..., p \tag{10}$$

$$y_q = A_q^T \; H(x) \quad q = 1, ..., p \tag{11}$$

Where:

$$H(x) = \left[H_1(x), ..., H_N(x)\right] \in \mathbb{R}^N \tag{12}$$

With

$$\begin{cases} H_i(x) = K\left(x^{(i)}, x\right), i = 1, ..., N \\ and \\ A_q = \left[a_1^q, ..., a_N^q\right]^{\mathrm{T}}, q = 1, ..., p \end{cases} \tag{13}$$

The output vector $Y_p$ is then given by:

$$Y_p = \begin{pmatrix} y_1 \\ . \\ . \\ y_p \end{pmatrix} = \begin{pmatrix} a_1^1 a_2^1 & . & . & . & a_N^1 \\ a_1^2 a_1^2 & . & . & . & a_1^2 \\ . & . & . & . & . \\ a_1^p a_2^p & . & . & . & a_N^p \end{pmatrix} \begin{pmatrix} K(x_1, x) \\ . \\ . \\ K(x_N, x) \end{pmatrix} \tag{14}$$

Or from (12) and (13)

$$Y_p = \begin{pmatrix} A_1^T \\ . \\ . \\ . \\ A_p^T \end{pmatrix} H(x) \tag{15}$$

# 5 Kernel methods

## 5.1 Support Vector machine (SVM)

We now discuss Support Vector Machines (SVM) [1], [2] in its regression form [25] that corresponds to the minimization of problem (8) and where the cost function $V$ is:

$$V = \left|y_i - f(x_i)\right|_{\xi} \tag{16}$$

Where $\xi$ width of tube

The coefficients $\{a_i\}$ in the relation (9) can be determined by solving a Quadratic Programming (QP) problem with linear constraints.

## 5.2 Regularization networks (RN)

In this case the cost function is:

$$V\left(y^{(i)}, f\left(x^{(i)}\right)\right) = \left(y^{(i)} - f\left(x^{(i)}\right)\right)^2 \tag{17}$$

and the solution provides the parameters $\{a_i\}$ which verify:

$$a_i = \sum_{j=1}^{N} \left(K + \lambda N I_N\right)_{i,j}^{-1} y^{(j)} \tag{18}$$

In a matrix form, the relation (18) is:

$$\left(K + \lambda N I_N\right) \mathrm{A} = Y \tag{19}$$

where: $K$ is the Gram matrix that satisfy:

$$\begin{aligned} &K_{ij} = k\left(x^{(i)}, x^{(j)}\right) \\ &A = \left(a_1 \; a_2 ... a_N\right)^T \text{ and } Y = \left(y^{(1)}, y^{(2)}, ..., y^{(N)}\right)^T \end{aligned} \tag{20}$$

### 5.3 Kernel Principal Component Analysis (KPCA)

As the dimension $l$ of the RKHS spaces is high (possibly infinite), the determination of the coefficients $w_j$ of the relation (2) is problematic. Using the Principal Component Analysis (PCA) [10] method to identify the RKHS model is very interesting. The KPCA method doesn't target the input space such as in the linear case but it treats the transformed of these data by the application $\Phi$ defined in (6) and satisfying (7).

The RKHS model of this system is given by the relation (9). The Gram matrix $K$ associated to the kernel $k$ is an $N$ - dimensional square matrix, so that:

$$K_{i,j} = k\left(x^{(i)}, x^{(j)}\right) \text{ for } i, j = 1, ..., N \qquad (21)$$

We assume that the transformed data $\left\{\Phi\left(x^{(i)}\right)\right\}_{i=1, ..., N} \in \mathbb{R}^l$ are centred [8], [9]. The approximation covariance matrix $C_\phi$ of the transformed data is symmetrical, $l$ - dimensional and it is written as following:

$$C_\phi = \frac{1}{N}\sum_{i=1}^{N}\Phi\left(x^{(i)}\right)\Phi\left(x^{(i)}\right)^T,\ C_\phi \in \mathbb{R}^{l\times l} \qquad (22)$$

Let $l'$ the number of the eigenvectors $\left\{\theta_j\right\}_{j=1}^{l'}$ of $C_\phi$ corresponding to the non zeros positive eigenvalues $\left\{\lambda_j\right\}_{j=1}^{l'}$. It is proved in [9] that the number $l'$ is less or equal to $N$. Due to the large size $l$ of $C_\phi$, the calculus of $\left\{\theta_j\right\}_{j=1}^{l'}$ can be difficult. The KPCA method shows that these $\left\{\theta_j\right\}_{j=1}^{l'}$ are related to the eigenvectors $\left\{\beta_j\right\}_{j=1}^{l'}$ of the gram matrix $K$ according to [9]:

$$\theta_j = \sum_{i=1}^{N}\beta_{j,i}\ \Phi\left(x^{(i)}\right)\ ,\ j = 1, ..., l' \qquad (23)$$

Where $\left(\beta_{j,i}\right)_{j=1, \dots, p}$ are the components of $\left\{\beta_j\right\}_{j=1}^{l'}$ associated to their nonzero eigenvalues $\mu_1 > ... > \mu_{l'}$ of the Gram matrix $K$

The principle of the KPCA method consists in organizing the eigenvectors $\left\{\beta_j\right\}_{j=1}^{l'}$ in the decreasing order of their corresponding eigenvalues $\left\{\mu_j\right\}_{j=1}^{l'}$. The principal components are the $p$ first vectors $\left\{\theta_j\right\}_{j=1}^{p}$ associated to the highest eigenvalues and are often sufficient to describe the structure of the data [8], [4]. The number $p$ satisfies the Inertia Percentage criterion IPC given by:

$$p^* = \arg\left(IPC \geq 99\right) \qquad (24)$$

Where

$$IPC = \frac{\sum_{i=1}^{p}\mu_i}{\sum_{i=1}^{N}\mu_i} * 100 \qquad (25)$$

The principal components $\theta_j$ $j = 1, ..., p$ are unit vectors if the corresponding vectors $\beta_j$ satisfy:

$$\left\langle \beta_j, \beta_j \right\rangle = \frac{1}{\mu_j} \qquad j = 1, ..., l' \qquad (26)$$

Let $F_{kpca}$ the space spanned by the $p$ principal components $\left\{\theta_j\right\}_{j=1}^{p}$. The projection $\tilde{\Phi}(x) \in \mathbb{R}^p$ of $\Phi(x) \in \mathbb{R}^l$ on $F_{kpca}$ is given by:

$$\tilde{\Phi}(x)_j = \langle \theta_j, \Phi(x) \rangle \qquad j = 1, ..., p \qquad (27)$$

and from (23), we have:

$$\begin{aligned}\tilde{\Phi}(x)_j &= \langle \sum_{i=1}^{N}\beta_{j,i}\Phi\left(x^{(i)}\right), \Phi(x)\rangle \\ &= \sum_{i=1}^{M}\beta_{j,i}\langle\Phi\left(x^{(i)}\right), \Phi(x)\rangle\end{aligned} \qquad (28)$$

and according to (7), we have

$$\tilde{\Phi}(x)_j = \sum_{i=1}^{N}\beta_{j,i}\ k\left(x^{(i)}, x\right) \qquad (29)$$

Similarly to the linear PCA [10], the model proposed by the Kernel PCA [4] is given by:

$$\tilde{y}_{kpca} = \sum_{j=1}^{p} w_j \tilde{\Phi}(x)_j \qquad (30)$$

Where $\tilde{y}_{kpca}$ and $\tilde{\Phi}(x)_j$ are respectively the model output and the $j^{th}$ component of $\tilde{\Phi}(x)$. It is interesting to mention that the number of the parameters $\left\{w_j\right\}_{j=1, ..., p}$ of the model (30) is less than $l$ given by the relation (8). To identify the model parameter $w_j$ we solve the following problem:

$$\min_{w\in\mathbb{R}^p}\frac{1}{2}\sum_{i=1}^{M}\left(y^{(i)}-w^T\tilde{\Phi}\left(x^{(i)}\right)\right)^2;\quad w=\left(w_1,...,w_p\right)^T \quad (31)$$

For any new input data $x^{(new)}\in E$, the corresponding KPCA model output is:

$$\tilde{y}_{kpca}^{(new)}=\sum_{j=1}^{p}w_j\,\tilde{\Phi}\left(x^{(new)}\right)_j \quad (32)$$

According to (29), the relation (32) is given by:

$$\tilde{y}_{kpca}^{(new)}=\sum_{j=1}^{p}w_j\sum_{i=1}^{N}\beta_{j,i}\,k\left(x^{(i)},x^{(new)}\right) \quad (33)$$

Thus:

$$\tilde{y}_{kpca}^{(new)}=\sum_{i=1}^{N}\sum_{j=1}^{p}w_j\beta_{j,i}\,k\left(x^{(i)},x^{(new)}\right) \quad (34)$$

Let $a_i=\sum_{j=1}^{p}w_j\,\beta_{j,i}$, relation (34) yields the model (9).

$$\tilde{y}_{kpca}^{(new)}=\sum_{i=1}^{N}a_i\,k\left(x^{(i)},x^{(new)}\right) \quad (35)$$

# 6 Application

To illustrate the efficiency of the proposed model we proceed to its validation on Chemical process: The Tennessee Eastman process.

## 6.1 Process description

The Tennessee Eastman (TE) process [19] is a highly non linear, non-minimum phase, and open-loop unstable chemical process consisting of a reactor/separator/recycle arrangement. This process produces two products G and H from four reactants A, C, D and E. Also a byproduct F is present in the process. The simultaneous, irreversible and exothermic gas-liquid reactions are:

$$\mathrm{A(g)}+C(g)+D(g)\rightarrow G(\mathrm{lig}),\text{Product 1}$$
$$\mathrm{A(g)}+C(g)+E(g)\rightarrow G(\mathrm{lig}),\text{Product 2}$$
$$\mathrm{A(g)}+E(g)\rightarrow F(\mathrm{lig}),\text{byroduct}$$
$$3D(g)\rightarrow 2F(\mathrm{lig}),\text{byroduct}$$

The process has 12 valves available for manipulation and 41 measurements available for monitoring or control. The detailed description of these variables, process disturbances and base case operating conditions, is given in [17]. The process flowsheet is presented in Figure 2.

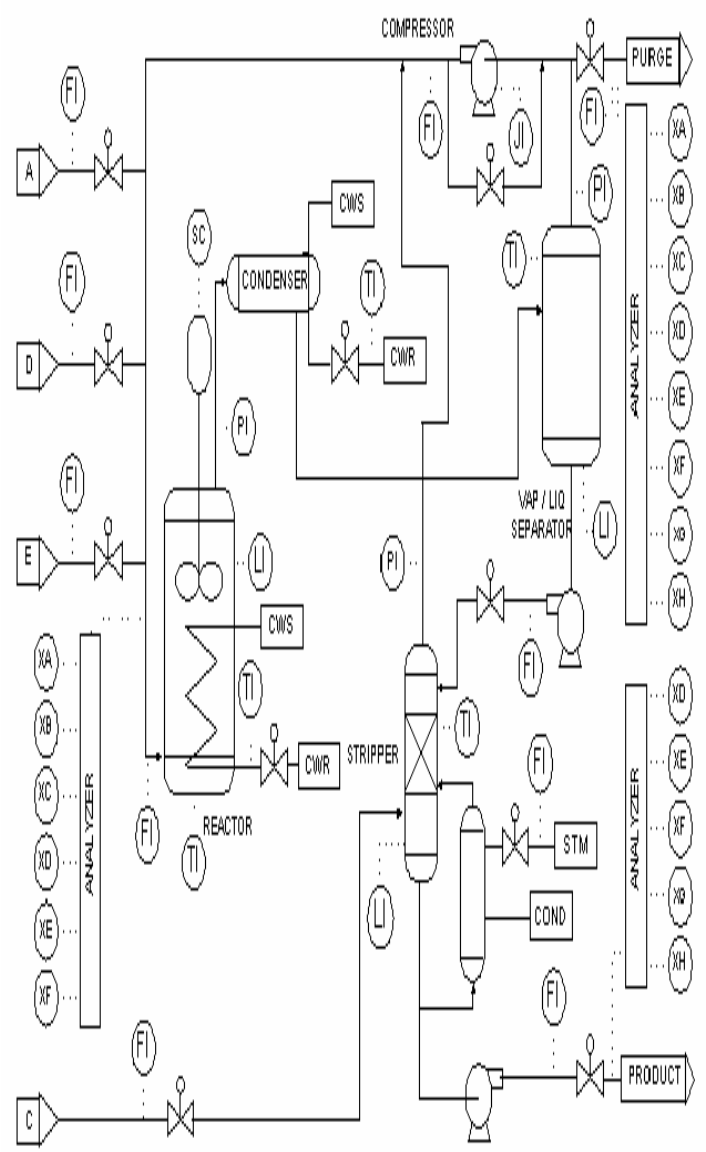


Fig.2 Tennessee Process

The modelling and the identification of the Tennessee Eastman process represent a challenge for the control community. It has been the subject of several studies [16], [18], [20] and [21], but most of them have tackled the process control without giving importance to modeling step. R. Sriniwas and Y. Arkun [22] have used input/ output process data to identify an autoregressive (AR) model parameters.

In our paper we intend to identify the parameters of the corresponding MIMO RKHS model of this process using the same technique of [23] for generating the data.

### 6.2 Data Extraction

The input/output data used to built the model were generated from the model of Tennessee implemented for the program Matlab © in the toolbox Simulink [19]. The process has 12 inputs and 41 outputs. According to the work of [22] the process was divided into two fields.

The first with a PID controller to maintain the process stability. The second field is devoted to the identification where only four inputs (reactor pressure, reactor level, D feed flow and E feed flow) are tuned and the others are maintained as suggested by mode 3 of Simulink model. Assuming the reactor outputs, we select the separator temperature and reactor liquid level.

### 6.3 Knowledge model of Tennessee process

In this section, we consider for the knowledge model of the Tennessee process that suggested by A. Mauricio Sales Cruz [19] as shown by Figure 3.

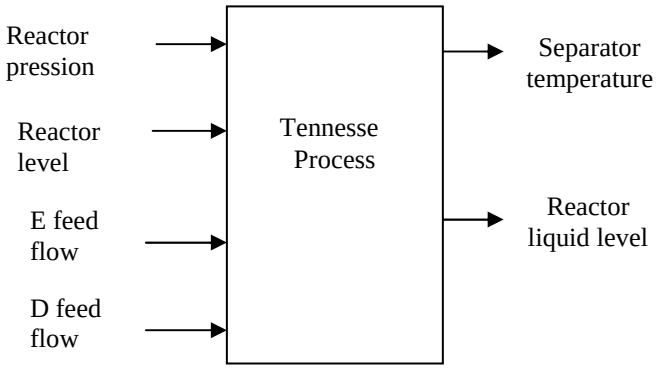


Fig. 3: Scheme input /output of Tennessee process

* Output equations:

The first output, separator temperature is given by:

$$T_S = \left(T_{CW,S,out} - T_{CW,S,in}\right) \quad (36)$$

Where:

$T_{CW,S,out}$ Cooling water outlet temperature in the separator

$T_{CW,S,\text{in}}$ Cooling water inlet temperature in the separator

The temperature $T_S$ is linked to the energy $Q_S$ removed for the separator by the differential equation:

$$\dot{Q}_S = m_{CW,S} c_{p,CW} \left(T_{CW,S,out} - T_{CW,S,in}\right) \quad (37)$$

Where:

$C_{p,CW}$ Specific heat capacity cooling water, kJ $kg^{-1}$ $K^{-1}$

$m_{CW,S}$ Cooling water flow rate separator, kg $h^{-1}$

For the second output, the reactor liquid level is given by:

$$V_{Lr} = \sum \frac{N_{i,r}}{\rho_i} \; ; \; i = D,E,F,G,H \quad (38)$$

Where:

$\rho_i$ Molar density of component i, mol $m^{-3}$

$N_{i,r}$ is the total molar holdup of the component $i$ in the reactor

### 6.4 Modelling in RKHS

To generate the data from simulink, the simulation step size was 0.0005 s and the data were collected every 0.02 s.
To build the RKHS model we use the Kernel ERBF (Exponentiel Radial Basis Function)

$$K(x,x') = \exp\left(-\frac{\|x - x'\|}{\sigma^2}\right) \quad (39)$$

Where
$\sigma = 15$, and $\| \; \|$ is the euclidean norm. The term of regularisation $\lambda = 1,002 \times 10^{-18}$
In the identification phase we use a training set of 1000 inputs/outputs and in the validation phase 800 new inputs/outputs are used to evaluate the performance of the resulting RKHS model.

### 6.5 Result

Figure 4 concerns the separator temperature. We plot the evolution of the model output and that given by the simulink during the learning phase. However, in Figure 5 the same outputs are plotted using validation data.

We notice a concordance between the RKHS model output and the process output in the learning phase and this concordance remains excellent in the validation phase. Indeed the normalised mean square error in validation is equal to 0.0012.

Figure 6 concerns the reactor liquid level where we plot the evolution of the output of the model and that given by the simulink during the learning phase and figure 7 the same outputs are plotted using validation data.

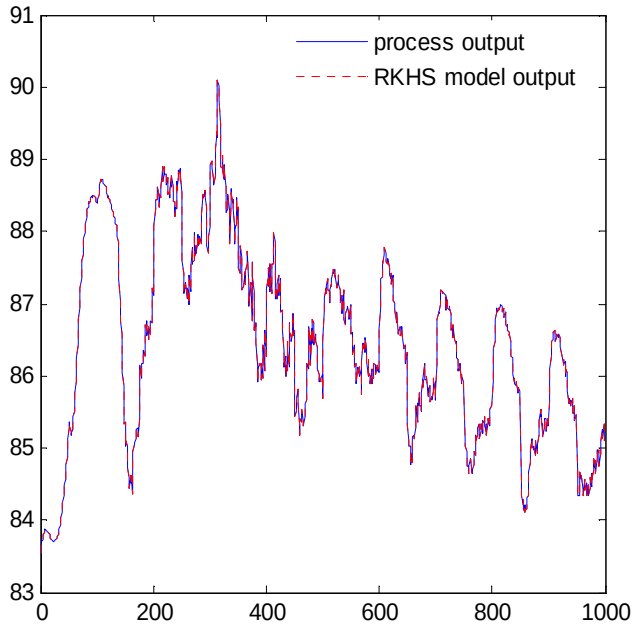


Fig .4. Identification Phase

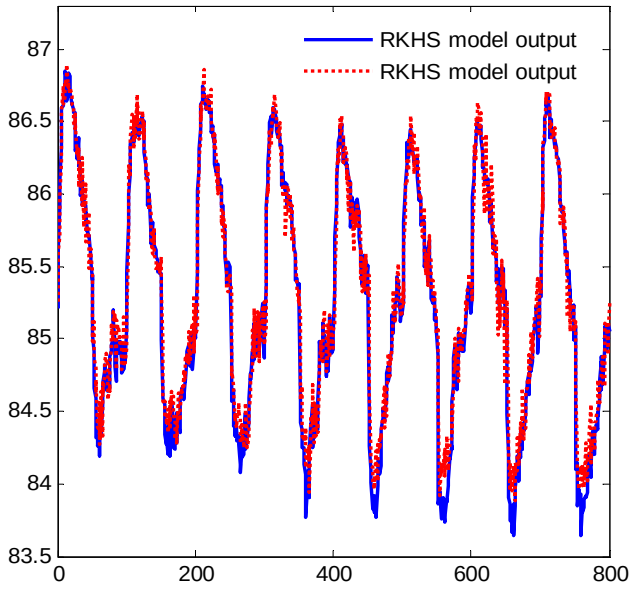


Fig .5. Validation Phase

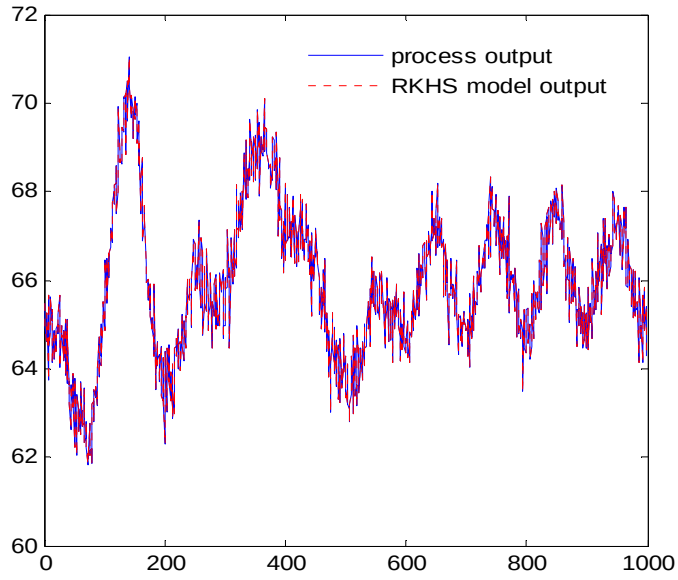


Fig 6. Identification Phase

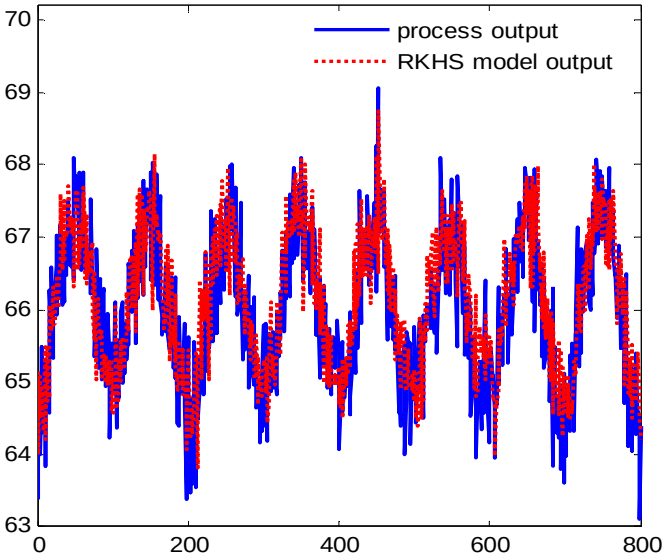


Fig .7. Validation Phase

On the test set, the validation phase confirm the good performances of the RKHS model: Indeed the similarity between the RKHS model output and the real output process is very important and proves that there is not an overfitting

## 7 Comparative study

To illustrate the efficiency of all methods, we proceed to their validation on nonlinear dynamic system used as a benchmark [11] and given by:

$$y(i)=\left(0.8-0.5\exp\left(-\left(y(i-1)\right)^2\right)\right)y(i-1)$$
$$-\left(0.3+0.9\exp\left(-\left(y(i-1)\right)^2\right)\right)y(i-2) \quad (40)$$
$$+0.1\sin\left(\pi.y(i-1)\right)+e(i)$$

Where:

$e(i)$: is a Gaussian noise with variance 0.4 and unit mean so that the Signal to Noise Ration (SNR) is equal to $5$. The input vector $x\in\mathbb{R}^d$ of RKHS model is:

$$x(k)=\left[y(k-1)\;\; y(k-2)\right]^T \quad (41)$$

To build the RKHS model we use the RBF kernel (Radial Basis Function) defined as:

$$k\left(x,x^{'}\right)=\exp\left(-\frac{\left\|x-x^{'}\right\|^2}{100}\right) \quad (42)$$

We have used 80 observations in the identification phase and 150 other observations in the validation one.

In Table 1 we present the performances of the three kernel method carried out in the same condition. The first performance is the generalization ability evaluated by the Normalized Means Square Error (NMSE) in the identification phase and the validation one and the second concerns the compute time. For the simulation the clock frequency is set to 3 GHz

From Table 1 we see that the compute time of both methods KPCA and RN are significantly less than the time needed by the SVM method (7.53 seconds). This is due to the fact that the SVM method solves an optimization problem with constraints.

However that the SVM method generalizes better than the other two methods RN and KPCA as the corresponding NMSE in learning and validation phase is significantly smaller than other techniques.

We also note that three principal components is enough to build an RKHS model from the system which leads to a less complex model.

Table I
Performances of kernel methods SVM, RN and KPCA

| | SVM | RN | KPCA |
|---|---|---|---|
| Compute time (seconds) | 7.53 s | 0.9204 s | 0.5460s |
| NMSE identification % | $8{,}88\ 10^{-11}$ | $9{,}5810^{-5}$ | $3\ 10^{-3}$ |
| NMSE validation % | $2{,}32\ 10^{-9}$ | $1{,}38\ 10^{-4}$ | $8.710^{-3}$ |

In figure 8 we draw the RKHS model outputs using SVM, RN and KPCA methods and the output process in the identification phase. We notice the concordance between the three model outputs and that of the process output.

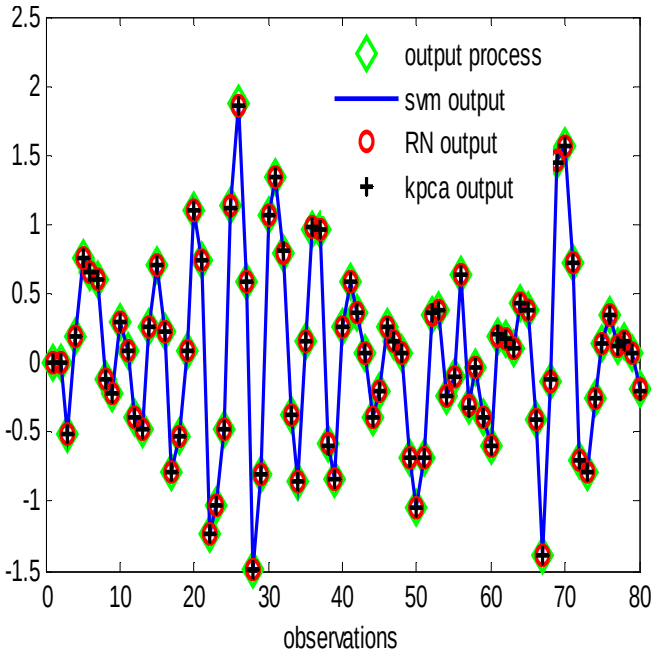


Fig.8. Identification Phase using SVM, RN and KPCA

In figure 9 we plot the RKHS model outputs using SVM, RN and KPCA methods and the output process in the validation phase. As previously we notice the concordance between all the outputs.

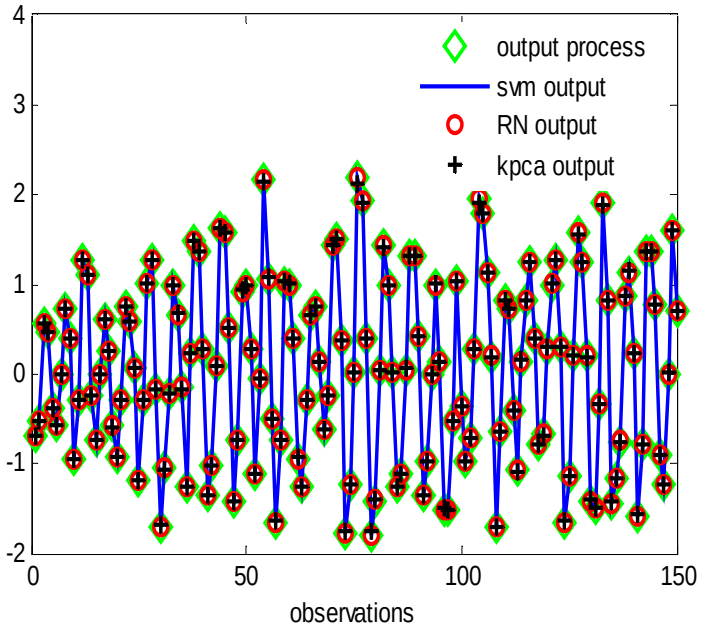


Fig. 9 Validation Phase using SVM, RN and KPCA

## 8 Conclusion

In this paper we have presented a new MIMO RKHS model for modeling a multivariate nonlinear system on RKHS space, also we have presented three kernel methods SVM, RN and KPCA to identify the nonlinear system in the RKHS space. A comparative study between the three methods has been achieved and shows the effectiveness of the SVM method in terms of generalization ability.

*References:*

[1] V. N Vapnik. *The Nature of Statistical Learning Theory*. Springer, New York, 1995

[2] Vapnik. V. N. *Statistical Learning Theory, Wiley, New York*, 1998

[3] Aronszajn N, Theory of reproducing Kernels. *Transactions of the American Mathematical Society*, Vol.68, pp. 337-404, 1950

[4] Ilyes. A. Modélisation, *Identification et Commande Prédictive des systèmes non linéaires par utilisations des espaces RKHS*, thèse de doctorat de l'université de Tunis, ENIT, 4-4-2009, Tunisie

[5] Wahba. G. *An introduction to model building with Reproducing Kernel Hilbert Spaces*, Technical report No 1020, Department of Statistics, University of Wisconsin-Madison, 2000

[6] Taouli. O, Saidi .N and Messaoud. H. Identification of Non linear MISO Process using RKHS and Volterra models. *WSEAS TRANSACTIONS on SYSTEM*. Vol 8, 2009, pp.723-732.

[7] Taouli. O, Aissi. I, Villa. N and Messaoud. H. Identification of Non Linear Multivariable Processes Modelled on Reproducing Kernel Hilbert Space: Application to Tennessee Process. *ICONS, 2nd IFAC International Conference on Intelligent Control Systems and Signal Processing*, Istanbul, Turquie 21-23 September 2009, pp.1-6.

[8] Scholkopf. B, Smola. A. *Learning with Kernel: Support Vector Machine Regularization*, Optimization and Beyond, MIT Press, 2001

[9] Scholkopf. B, Smola. A et Muller. K-R. Nonlinear component analysis as Kernel eigenvalue problem, *Neural computation*. v10, 1998, pp. 1299-1319.

[10] Jolliffe. I.T. *Principal components analysis, Springer-Verlag* , New York , 1986

[11] Chan . W. C, Chan. C. W, K. C. Cheung, and C. J. Harris. Modelling of nonlinear dynamical systems using support vector neural networks. *Engineering Applications of Artificial Intelligence*, 14, 2001, pp. 105-113.

[12] Ghate. V.N, Dudul. S.V. Induction Machine Fault Detection Using Support Vector Machine Based Classifier, *WSEAS TRANSACTIONS on SYSTEMS*, Volume 8, May 2009, pp. 591-603.

[13]Yang. M.H. Kernel Eigenfaces vs.kernel Firstfaces:Face recognition using kernel method. *In IEEE FGR*, Washington, DC, USA, 2002, pp. 215-220.

[14] Veropoulos. K, N. Cristianini, and Campbell C. The Application of Support Vector Machines to Medical Decision Support: A Case Study, *ACAI conference*, Chania, Greek Islands, 1999, pp 17-21.

[15] J. Mercer. Functions of positive and negative type and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society*, London, A 209, 1909, pp.415- 446.

[16] Banerjee, A and Arkun, Y. Control configuration design applied to the Tennessee Eastman plant-wide control problem. *Computers Chemical Engineering*, 19, 1995, pp.453-480.

[17] Downs, J.J and Vogel, E. A plant-wide industrial process control problem. *Computers and Chemical Engineering*, 17, 1993, pp.245-255.

[18] Mc Avoy, T.J and Ye, N. Base control for the Tennessee Eastman problem. *Computers and Chemical Engineering*. 18,1994, pp.383-413.

[19] Mauricio Sales Cruz, A. *Tennessee Eastman Plant –wide Industrial Process*. Technical report, CAPEC, Technical University of Denmark, 2004.

[20] Ricker, N.L and Lee, J.H. Nonlinear model predictive control of the Tennessee Eastman challenge process. *Computers and Chemical Engineering*, 19, 1995a, pp. 961-981.

[21] Ricker, N.L and Lee, J.H. Nonlinear modeling and state estimation for the Tennessee Eastman challenge process. *Computers and Chemical Engineering,* 19, 1995b, pp. 983-1005.

[22] Sriniwas, R and Arkun, Y. Control of the Tennessee Eastman process using input–output models. *Journal of Process Control*, 7, 1997, pp. 387-400.

[23] Ylöstalo, T and Hyötyniemi, H. Maps of Process Dynamics. *In Proceedings of EUROSIM'98 Simulation Congress, K. Juslin (ed.), Espoo, Finland: Federation of European Simulation Societies*, 2, 1998, pp. 225-229.

[24] G. Pajares, Support vector machines for shade identification in urban areas, *WSEAS Transactions on Information Science and Applications*, Vol. 2, No. 1, 2005, pp. 38-41

[25] Milena R. Petkovic, Milan R. Rapaic, Boris B. Jakovljevic, Electrical Energy Consumption Forecasting in Oil Refining Industry Using Support Vector Machines and Particle Swarm optimization, *WSEAS Transactions on Information Science and Applications*, Vol. 6, 2009, pp. 1761-1770.