2022 IEEE 21st International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA) | 978-1-6654-8261-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/STA56120.2022.10019110

Predictive Control of nonlinear systems based on PWA model

LASSOUED Zeineb
Numerical Control of Industrial Processes Laboratory
National School of Engineers of Gabes,
University of Gabes 6029 Gabes, Tunisia
zeineb.lassoued1@gmail.com

ABDERRAHIM Kamel
Numerical Control of Industrial Processes Laboratory
National School of Engineers of Gabes,
University of Gabes 6029 Gabes, Tunisia
kamelabderrahim@yahoo.fr

***Abstract*—This work concerns the control of nonlinear systems. For this purpose we identify the nonlinear system using PieceWise affine models (PWA). This choice is due to their ability to represent nonlinear systems with precision. Then, we proceed to design the one-step-ahead predictive control using the identified models. The proposed strategy exploit the same principle of the multimodel adaptive control with switching. For this strategy, a regulator is synthesized for each submodel using one of linear systems control approaches. The obtained results of the proposed control approache are satisfying.**

***Index Terms*—PWA systems, Modeling, Clustering approach, One-Step-Ahead predictive Control.**

I. Introduction

Piecewise Affine PWA models represent the most important class of hybrid systems. They can represent many nonlinear processes such as embedded systems, chemical process, telecommunication networks, air traffic management systems, etc [1], [2], [3], [4]. These systems involve continuous and discrete dynamics interacting with each other. The continuous dynamic is the feature of any physical system although the discrete dynamic is obtained by switches, transitions, operating phases, etc.

The most successfully used control strategy for piecewise affine systems is the Hybrid Model Predictive Control (HMPC) [5]–[8]. It is a technique of optimal control methodology, since the control sequence is computed from minimization of a criterion to maintain the system output closest to the desired reference trajectory. The receding and the anticipatory action features distinguish it from other optimal control strategies [9]. Nevertheless, the HMPC approach has some limitaions. In fact, the arbitrary choice of the tuning parameters can affect the results because the PWA systems have arbitrary switch phenomenon. Indeed, switching from one sub-model to another need to redefine the tuning parameters to keep the desired performance. Some methods exist in the literature for the parameters adjustement of the generalized predictive control such as the parametric identification algorithms using fuzzy logic [10], the multi-objective optimization algorithms [11], the fuzzy supervisor [12].

In this paper, the proposed approach for PWARX systems control is inspired from the principles of the adaptive open loop control or switching multimodel control. This approach proceed to synthesis a controller to every sub-model ensuring the desired performances and then in developing a supervisor which allows selecting the best controller at every instant based on the minimization of a performance index. The precision of the control depends on the efficiency of the performance index. However, the synthesis parameters used by this index, such as the forgetting factors are set empirically. Consequently, they must be properly chosen in order to avoid instability of the system control. The use of the PWA models allows to overcome this problem since the suitable controller is selected automatically based on the actual regressor which designates the corresponding region and then generates the active sub-model. In this work, we suggest the synthesis of predictive control for PWA systems based on linear techniques. The organization of this paper is as follows. In section 2, the mathematical modeling and problem formulation are given. Section 3 presents the proposed predictive control strategy for PWA systems. In section 4, we present the results of the proposed control approach.

II. Mathematical modeling and problem formulation

A. Hybrid systems representation

Dynamic hybrid systems are heterogeneous dynamical systems which gather continuous and event phenomena. These systems offer a solution that is both simple and rigorous for the representation of complex systems since the hybrid behavior occurs in most industrial processes. Indeed, there exist numerous representation of these systems. Among them, we distinguish the PieceWise Auto-Regressive eXogenous models (PWARX) which are obtained by the decomposition of the regression domain into convex polyhedral regions, and the association of an affine model to each region. These models are considered as universal approximators. This property gives the possibility to exploit linear systems analysis techniques if the local models are linear. The input-output representation of a PWA system can be formulated as follows:

$$y(k) = \begin{cases} \theta_1^T \bar{\varphi}(k) + e(k) & if \quad \varphi(k) \in \quad H_1 \\ \vdots & \\ \theta_s^T \bar{\varphi}(k) + e(k) & if \quad \varphi(k) \in \quad H_s \end{cases} \tag{1}$$

$$\varphi(k) = \begin{bmatrix} y(k-1) \\ \vdots \\ y(k-n_a) \\ u(k-1) \\ \vdots \\ u(k-n_b) \end{bmatrix} \quad \theta_i = \begin{bmatrix} a_{i,1} \\ \vdots \\ a_{i,n_a} \\ b_{i,1} \\ \vdots \\ b_{i,n_b} \\ g_i \end{bmatrix} \tag{2}$$

$$\bar{\varphi} = \begin{bmatrix} \varphi^T & 1 \end{bmatrix}^T. \tag{3}$$

where

- $y(k) \in \mathbb{R}$, $u(k) \in \mathbb{R}$, $e(k) \in \mathbb{R}$, $s \in \mathbb{N}$ are respectively the output, the input, the additive noise and the number of sub-models.
- $\theta_i \in \mathbb{R}^{n_a+n_b+1}$ is the parameter vector of the i^{th} sub-model having n_a and n_b as orders.
- $a_{i,j}$ and $b_{i,j}$ represent the coefficients of the i^{th} sub-model while g_i represents an independent affine parameter of the i^{th} sub-model.
- $\varphi(k) \in \mathbb{R}^{n_a+n_b}$ is the regressor vector.
- $H_i \in \mathbb{R}^{n_a+n_b}$ is the polyhedral partition of the i^{th} sub-model. The polyhedral partitions $H_i, i = 1, .., s$ must verify the following assumptions:

$$\begin{cases} \bigcup_{i=1}^{s} H_i = H \\ H_i \bigcap H_j = \emptyset \quad \forall i \neq j \end{cases} \tag{4}$$

B. Hybrid systems identification

System identification is the construction of mathematical models from the input-output measurements of the system. This approach is the most used in the field of system control design because it ensures a good compromise between simplicity and precision of the model and is distinguished by its easy of implementation. Moreover, it is generally applicable to all physical systems. We can formulate the identification problem of a PWA system as follows:
Using the input-output measurements, estimate the number of submodels, the parameters vectors and the coefficients of the affine hyperplanes that define the regions.

It is easy to deduce that PWARX systems identification is one of the most difficult problems. For many approach, the orders of the sub-models must be defined in order to alleviate the complexity. Despite the consideration of these hypotheses, the subject still remains difficult because it requires the resolution of two problems which are the identification of the parameters of the sub-models and the estimation of the coefficients of the hyperplanes delimiting the regression domain. There exist a plenty of methods for PWARX systems identification: the algebraic method [13], the classification method [14], the Bayesian method [15], the bounded error method, and so forth. Only the classification approach is considered in this work because of its simplicity and its good performances in many real-time examples.

C. Identification of PWA models based on clustering approach

The main steps that caracterize the clustering based approach are: data classification, parameters vectors estimation and regions definition.

1) Constructing local data sets: Constructing the local set C_k to each pair of data $\{\varphi(k), y(k)\}_{k=1}^{N}$. Every set C_k containing the regressor vectors and a finite number of nearest neighbors [14].

The number of neighbors n_ρ is a random parameter. It has an important role in the algorithm. So, it must be properly chosen in order to ameliorate the identification result. For the obtained local sets $\{C_k\}_{k=1}^{N}$, we identify the parameters vectors $\{\theta_k\}_{k=1}^{N}$ using least square method or any standard linear regression techniques.

1 teta per ogni coppia di
dati (N) e poi li
raggruppo in s clusters

$$\theta_k = (\phi_k^T \phi_k)^{-1} \phi_k^T Y_k. \tag{5}$$

$$\begin{aligned} \phi_k &= \left[\bar{\varphi}(t_k^1) ... \bar{\varphi}(t_k^{n_\rho})\right]^T \\ Y_k &= \left[y(t_k^1) ... y(t_k^{n_\rho})\right]^T \end{aligned} \tag{6}$$

$\left\{t_k^1, \cdots, t_k^{n_\rho}\right\}$ are the indices of the elements of sets C_k.

2) Data clustering and parameters' estimation: The objective of this step is to classify the parameters vectors $\{\theta_k\}_{k=1}^{N}$ into s clusters and determine the sub-models parameters vectors $\{\theta_i\}_{i=1}^{s}$.

The efficiency of the classification technique has an important role in obtaining good estimation of sub-models parameters and hyperplane parameters. Several techniques of classification can be found in the literature. For example, in [14] the data classification step is achieved by the *k-means* algorithm which is sensitive to additive noise and it can not deal with outliers. Moreover, it assumes that the number of classes is known a priori. These problems lead to a degradation of the estimation quality of the parameters of the sub-models as well as of the coefficients of the hyperplanes. Therefore, we have proposed to replace the existing classification technique by other ones. Among them, we cite the DBSCAN approach (Density-Based Spatial Clustering of Applications with Noise) [16]. This algorithm allows assigning the data into distinct classes while being based on certain density conditions, *i.e.* the classes are considered as dense regions which are separated by low density regions. This method is also able to eliminate outliers during the partitioning process. In addition, it can determine the number of classes.
This algorithm was tested on simulated examples and on real measurements where it provided the best results by comparison with existing method like *k-means*. Indeed, the problems of convergence towards local minima and the divergence of the algorithm does not exist because we don't need any initialization. Moreover, it automatically generates the number of sub-models.

The main principle of the DBSCAN algorithm is that the neighborhood of the object of any cluster ϵ must hold in a minimum number *MinPts* of objects with a chosen radius. ϵ and $(MinPts)$ represent the input parameters of the algorithm. Therefore, they must be properly chosen in order to guarantee a good classification of the data.

Given ϵ and *MinPts* as input and a data set $S = \{\theta_k\}_{k=1}^{N}$, the ϵ-neighborhood of an object θ_k is:

$$N_\varepsilon(\theta_k) = \{\theta_j \in S;\ \|\theta_k - \theta_j\| \leq \varepsilon\}$$

Clusters are constructed by reviewing the ϵ-neighborhood of all the elements. So, a step of differentiation of points nature is necessary. Then, the point is considered as:

- *core*, when the cardinality of his neighborhood is higher or equal to MinPts.
- *border*, when the cardinality of his neighborhood is less than MinPts and it is within the neighborhood of any core point.
- *noise*, for the other points that are not core nor border.

The first cluster is formed by the first core point and its neighbors then all the other *core* points are evaluated. If the considered object isn't assigned to a cluster, another cluster will be created. In [16], the DBSCAN algorithm is more detailed.

3) Region reconstruction: The region reconstruction problem consists in determining the different partitions $\{H_i\}_{i=1}^{s}$. Since the polyhedral regions are defined by hyperplanes, estimating the regions amounts to separate the groups of points using linear classifiers (hyperplanes). Separating the points in H_i from H_j, $i \neq j$ with an hyperplane without errors is a fabulous task because the sets H_i and H_j have intersecting convex hulls. Therefore, we just have to find the hyperplane that minimizes some misclassification index. For the s sets H_i,...,H_s, two types of linear separation can be handled:

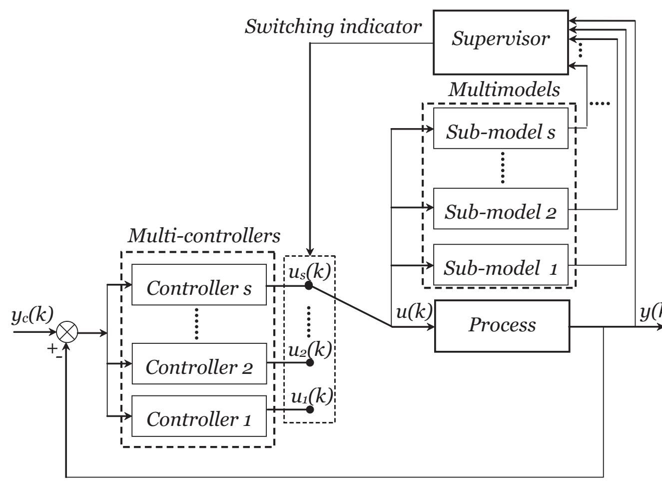
- Binary classification: for each pair (H_i, H_j), with $i \neq j$, a linear classifier is constructed.
- Multi-class classification: a piecewise linear classifier is determining having as object the discrimination of s classes.

The separation task can be accomplished by resorting to the support-vector machine (SVM) approach [17], [18].

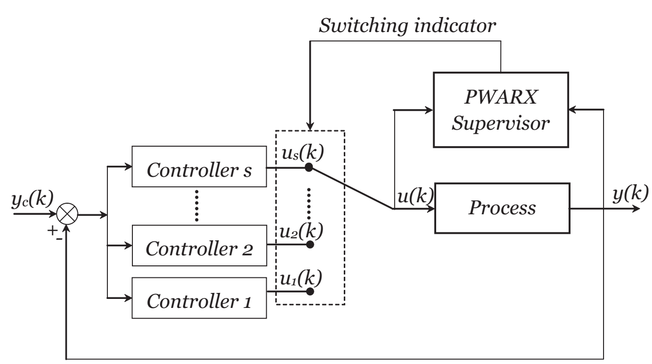
III. Proposed One-step-ahead predictive control strategy for PWA systems

After describing the nonlinear behaviour of the system with a PWA model based on the proposed clustering identification approach, we can proceed with the design of a system control law. The one-step predictive control is a convenient control approach which can emphasize the accuracy of the model. Indeed, in order to design the predictive control for PWARX systems, we propose a solution that exploit the principle of multimodel adaptive control with switching [5].
This approach is based on the estimation of the sub-models and the design of the supervisors that used to select the suitable sub-mode representing the process at every sampling time. After that, the suitable controller output is chosen.

For the multimodel adaptive control approach, the supervisor consists in solving a criterion based on the error of the output of the process and the outputs of the sub-models. To achieve this goal, many criteria can be found in the literature [19], [10].
However, the proposed supervisor uses the regression vector to decide about the active sub-model. The convenient controller is then determined by the SVM approach. Therefore, we can avoid the arbitrary choice of the ponderations of the existing approach.
Since the system is defined by a PWA model, the control structure is then simplified as shown in Figure 1(a) and 1(b).



(a) Proposed PWARX control strategy



(b) Multimodel adaptive control strategy

Fig. 1. Control design.

Results of linear system control can be applied to the PWA model. In fact, we propose the one-step-ahead predictive control. This method is broadly used in the field of industry as well as in the field of research. This control strategy addresses

system control issues that can be constrained while solving an optimization problem. The principle of this control strategy involves the knowledge of the mathematical model in order to anticipate the future behavior of the process.
The principle of the predictive control can be formulated as follows: "*use a model to predict the behavior of the system and determine the convenient controller by minimizing a performance criterion while respecting the constraints*".
The elements of the predictive control are therefore:

- a model representing the process for prediction;
- a performance criterion (cost function);
- constraints to impose on the state variables, inputs or outputs;
- an optimization algorithm generating the control law.

For each element, several options can be considered, which gives a multitude of algorithms of the predictive control.
In this work, we are interested in one-step predictive control. The choice of a prediction horizon of a single step is justified by the fact that the switching dynamics of the submodels is not known a priori and therefore switching from one submodel to another can occur in an unknown way.
The control law is calculated through the minimization of a quadratic criterion which penalizes the differences between the predicted outputs and the reference trajectory.
For each sampling time k, the available informations are the current output $y(k)$ and the previous inputs and outputs $y(k-1),...,y(k-n_a)$ and $u(k-1),...,u(k-n_b)$.
The main principle of the one-step predictive control is to find the control low $u(k)$ that coincides the output $y(k)$ with the reference trajectory $y_c(k)$. More precisely, this control is obtained by minimizing the following criterion:

$$J = \frac{1}{2}\left\|y(k+1) - y_c(k+1)\right\|_Q^2 + \frac{1}{2}\left\|u(k)\right\|_R^2 \qquad (7)$$

$Q \in \mathbb{R}^{m\times m}$ is a positive definite matrix and $R \in \mathbb{R}^{n\times n}$ is a semi-definite positive matrix and m and n are respectively the number of outputs and inputs.

PWA control problem formulation

The quadratic form of the criterion is given by:

$$\begin{array}{l} J = [y(k+1) - y_c(k+1)]^T Q [y(k+1) - \\ y_c(k+1)] + u(k)^T R u(k) \end{array} \qquad (8)$$

The output $y(k)$ of a PWA system is given by:

$$\left\{ \begin{array}{c} a_{1,1}y(k-1) + .. + a_{n_a,1}y(k-n_a) + \\ b_{1,1}u(k-1) + .. + b_{n_b,1}u(k-n_b) + g_1 \\ if \;\; \varphi(k) \in H_1 \\ \vdots \\ a_{1,s}y(k-1) + .. + a_{n_a,s}y(k-n_a) + \\ b_{1,s}u(k-1) + .. + b_{n_b,s}u(k-n_b) + g_s \\ if\varphi(k) \in H_s \end{array} \right. \qquad (9)$$

The system output is as follows:

$$\begin{aligned} y(k) &= \theta_{\sigma(k)}^T \bar{\varphi}(k) \\ &= a_{1,\sigma(k)}y(k-1) + \cdots + a_{n_a,\sigma(k)}y(k-n_a) \\ &+ b_{1,\sigma(k)}u(k-1) + \cdots + b_{n_b,\sigma(k)}u(k-n_b) \\ &+ g_{\sigma(k)} \end{aligned} \qquad (10)$$

where $a_{i,\sigma(k)}$, $b_{i,\sigma(k)}$ and $g_{\sigma(k)}$ represent the coefficients of the active sub-model and $\sigma(k)$ represents the considered sub-model.
The criterion of the one-step predictive control allowing a coincidence between the output $y(k+1)$ and the reference $y_c(k+1)$ for a PWARX system described by the relation (10) is formulated as follows:

$$J = Q.\Psi^2 + Ru^2(k) \qquad (11)$$

where

$$\begin{array}{l} \Psi = y_{ref}(k+1) - a_{1,\sigma(k)}y(k) - \cdots \\ -a_{n_a,\sigma(k)}y(k-n_a+1) - b_{1,\sigma(k)}u(k) - \cdots \\ -b_{n_b,\sigma(k)}u(k-n_b+1) - g_{\sigma(k)} \end{array}$$

An explicit solution of $u(k)$ can be obtained by minimising the criterion (11).

$$u(k) = \frac{Q.b_{1,\sigma(k)}.\Delta}{Q.b_{1,\sigma(k)}^2 + R} \qquad (12)$$

where

$$\begin{array}{l} \Delta = y_{ref}(k+1) - a_{1,\sigma(k)}y(k) - \cdots \\ -a_{n_a,\sigma(k)}y(k-n_a+1) - b_{2,\sigma(k)}u(k-1) - \cdots \\ -b_{n_b,\sigma(k)}u(k-n_b+1) - g_{\sigma(k)} \end{array}$$

A convenient choice of the weights Q and R has an impact on the stability of the synthesized controller.
In addition, the weight R can be considered as a tuning parameter allowing to obtain a balance between control magnitude and tracking accuracy.

IV. Simulated Example : Process of Level Control

We consider a level control system consisting of two connected tanks described in [12].

This system has to control the level h_2 in tank R_2 using the input flow q_e [20]. This process can be modeled by using the principles of Bernoulli and the conservation of mass as follows:

$$\frac{dh_1}{dt} = \frac{q_e}{S_1} - \frac{\alpha_1}{S_1}\sqrt{2g(h_1-h_2)} = f_1(h_1,h_2,q_e) \qquad (13)$$

$$\frac{dh_2}{dt} = \frac{\alpha_1}{S_2}\sqrt{2g(h_1-h_2)} - \frac{\alpha_2}{S_2}\sqrt{2gh_2} = f_2(h_1,h_2,q_e) \qquad (14)$$

where h_1 and h_2 are the levels in R_1 and R_2 respectively, S_1 and S_2 are the areas of R_1 and R_2, α_1 and α_2 are the effluent areas of R_1 and R_2, q_e is the sytem input, q_1 is the input of R_2, q_2 is the flow of th output while g represent the gravity constant.
The following parameters are used to simulate this process: $\alpha_1 = 0.002m^2$, $\alpha_2 = 0.002m^2$, $S_1 = 0.25m^2$, $S_2 = 0.1m^2$,

$l = 1.2m$, $g = 9.81m/s^2$ and $q_1 = 0.005m^3/mn$.
The identification of this process using the DBSCAN method id presented in our paper [12]. In fact, we have used a multi-sine sequence characterized by data length $N = 200$, sampling time $T_e = 2s$, input frequencies $Fe \in [0.001, 0.008]Hz$ and input amplitudes $A \in [0.001, 0.01]m^3/mn$.
The input-output measurements are depicted in Figure. 2.

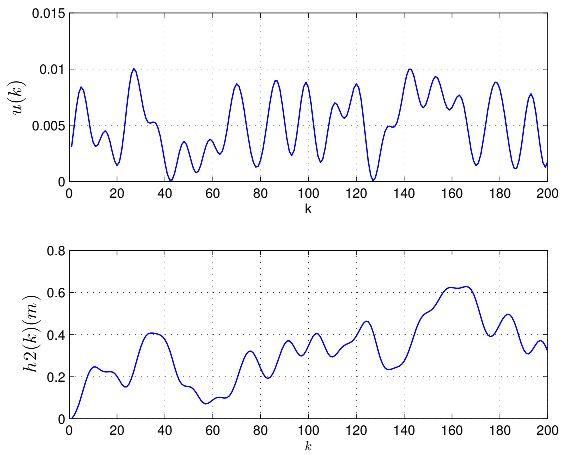


Fig. 2. Input-Output measurements.

Applying the DBSCAN identification method, we obtain the following results:

- The number of submodels s is equal to three.
- The parameter vectors $\theta_i = \{a_{i,1}, b_{i,1}, b_{i,2}\}_{i=1,2,3}$ are:

$$\theta_1 = \{0.7915, 2.2125, 5.2863\}$$
$$\theta_2 = \{0.8877, 0.1246, 7.5378\}$$
$$\theta_3 = \{0.9105, -0.7323, 8.4514\}$$

- The regions $\{H_i\}_{i=1,2,3}$ are as follows:

$$H_1 = \Big\{\varphi \in \mathbb{R}^3; \ [-4.989 - 0.105 - 0.103]^T \varphi(k) - 0.2528 \leq 0\Big\}$$

$$H_2 = \Big\{\varphi \in \mathbb{R}^3; \ [-4.5735 - 0.0207 - 0.0348]^T \varphi(k) + 0.1977 \leq 0 \ and \ [-4.989 - 0.105 - 0.103]^T \varphi(k) - 0.2528 > 0\Big\}$$

$$H_3 = \Big\{\varphi \in \mathbb{R}^3; \ [-4.5735 - 0.0207 - 0.0348]^T \varphi(k) + 0.1977 > 0\Big\}$$

We apply, firstly, the classical Hybrid Model Predictive Control (HMPC) method with constant tuning parameters ($Q_u = 9.9796 * 10^{-5}$ and $Q_y = 7.9601 * 10^{-7}$) [12].

The obtained results are presented in Figure. 3 which shows that the output and the reference trajectory don't have the same dynamics.

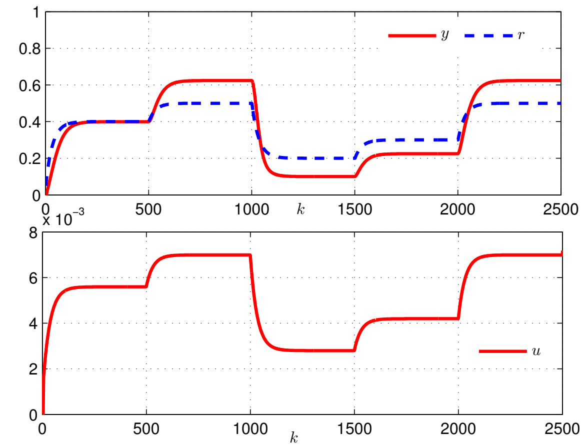


Fig. 3. HMPC control results: Reference trajectory and system output.

The application of the proposed control strategy with the same reference trajectory gives the results presented in Figure. 4 which presents the evolutions of the system's output, the reference, the error and the control signal.

We observe that the dynamics of the output and the reference trajectory are very close and then the tracking error is close to zero. Indeed, the proposed control strategy allow to obtain good closed-loop performance.

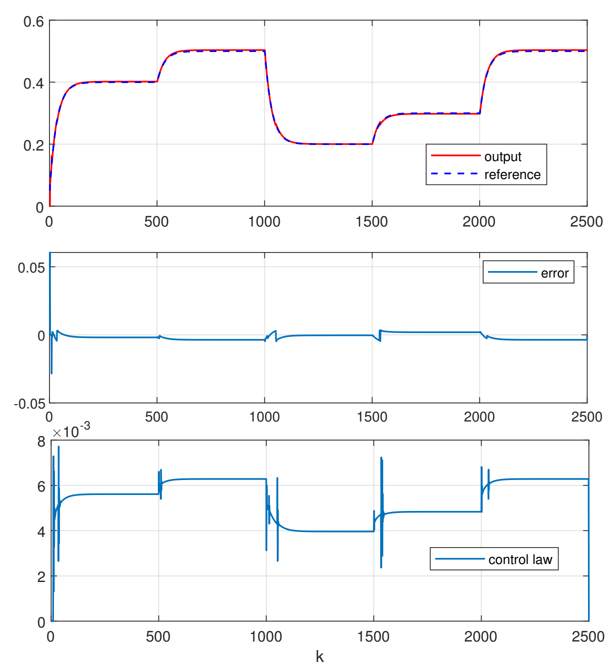


Fig. 4. Proposed Control results : Reference trajectory and system output.

Figure 5 presents the switching instances. We remark that if the reference changes value, the active sub-model is also changed.

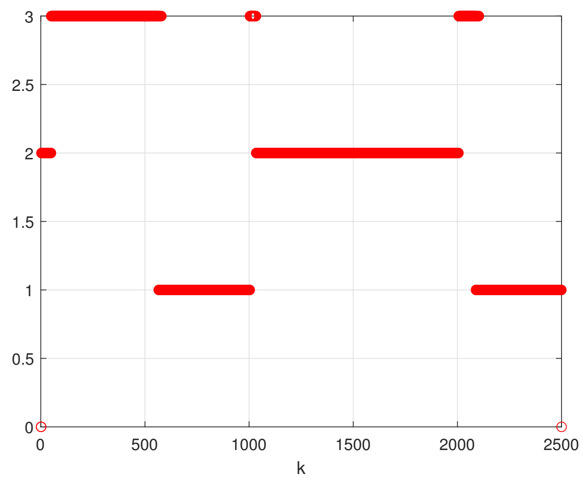


Fig. 5. Switching instances .

V. Conclusion

In this paper, the one-step predictive control is proposed to solve the problem of nonlinear system control. Indeed, the suggested strategy is inspired by the switching multimodel control.

The proposed approach develops a controller to every sub-model using one-step ahead predictive control. The selection of the best regulator for every sampling time is ensured in a systematic way using the active region mechanism. This represents an important advantage because the knowledge of this regulator in the case of the switching multimodel adaptive control is ensured by minimizing an explicit criterion.

However, it is important to point out that the implementation of these approaches for both identification and control is generally difficult in the case of a real system. These difficulties arise at several levels such as the determination of the number of sub-models, the choice of the structure of the sub-models, the choice of the synthesis parameters, the stability of the control system, etc. Indeed, the control of real time nonlinear systems is the core of our future contributions and therefore the problem is still open.

References

[1] Y. Du, F. Liu, J. Qiu, and M. Buss, "Online identification of piecewise affine systems using integral concurrent learning," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 10, pp. 4324–4336, 2021.

[2] P. Sindareh-Esfahani and J. Pieper, "Machine learning–based piecewise affine model of wind turbines during maximum power point tracking," *Wind Energy*, vol. 23, no. 2, pp. 404–422, 2020.

[3] B. Hadid, E. Duviella, and S. Lecoeuche, "Data-driven modeling for river flood forecasting based on a piecewise linear arx system identification," *Journal of Process Control*, vol. 86, pp. 44–56, 2020.

[4] Z. Lassoued and K. Abderrahim, "An experimental validation of a novel clustering approach to pwarx identification," *Engineering Applications of Artificial Intelligence, Elsevier*, vol. 28, pp. 201–209, 2014.

[5] I. D. Landau, R. Lozano, M. M'Saad, and A. Karimi, *Adaptive Control : Algorithms, Analysis and Applications*. Communications and Control Engineering, Springer, 2011.

[6] M. Sarailoo, B. Rezaie, and Z. Rahmani, "Fuzzy predictive control of three-tank system based on a modeling framework of hybrid systems," *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 228, no. 6, pp. 369–384, 2014.

[7] X. Sun, H. Zhang, Y. Cai, S. Wang, and L. Chen, "Hybrid modeling and predictive control of intelligent vehicle longitudinal velocity considering nonlinear tire dynamics," *Nonlinear Dynamics*, vol. 97, no. 2, pp. 1051–1066, 2019.

[8] M. Moness and A. M. Moustafa, "Hybrid modelling and predictive control of utility-scale variable-speed variable-pitch wind turbines," *Transactions of the Institute of Measurement and Control*, vol. 42, no. 9, pp. 1724–1739, 2020.

[9] F. Camacho, E, R. Ramirez, D, D. Limon, D. Munoz, and T. Alamo, "Model predictive control techniques for hybrid systems," *Annual Reviews in Control*, vol. 34, no. 1, pp. 21–31, 2010.

[10] L. Chen and K. S. Narendra, "Nonlinear adaptive control using neural networks and multiple models," *Automatica*, vol. 37, no. 8, pp. 1245–1255, 2001.

[11] A. Bemporad and D. L. Penab, "Multiobjective model predictive control," *Automatica*, vol. 45, no. -, pp. 2823–2830, 2009.

[12] O. Yahya, Z. Lassoued, and K. Abderrahim, "Predictive control based on fuzzy supervisor for pwarx hybrid model," *International Journal of Automation and Computing*, vol. 16, no. 5, pp. 683–695, 2019.

[13] F. Lauer, G. Bloch, and R. Vidal, "A continuous optimization framework for hybrid system identification," *Automatica*, vol. 47, no. 3, pp. 608–613, 2011.

[14] G. Ferrari-Trecate, M. Muselli, D. Liberati, and M. Morari, "A clustering technique for the identification of piecewise affine systems," *Hybrid Systems: Computation and Control*, pp. 218–231, 2001.

[15] O. Kocadağlı, "A novel hybrid learning algorithm for full bayesian approach of artificial neural networks," *Applied Soft Computing*, vol. 35, pp. 52–65, 2015.

[16] Z. Lassoued and K. Abderrahim, "Identification and control of nonlinear systems using piecewise auto-regressive exogenous models," *Transactions of the Institute of Measurement and Control*, vol. 41, no. 14, pp. 4050–4062, 2019.

[17] M. Awad and R. Khanna, "Support vector machines for classification," in *Efficient learning machines*. Springer, 2015, pp. 39–66.

[18] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, 2020.

[19] X. Wang, S. Li, W. Cai, H. Yue, X. Zhou, and T. Chai, "Multi-model direct adaptive decoupling control with application to the wind tunnel system," *ISA transactions, Elsevier*, vol. 44, no. 1, pp. 131–143, 2005.

[20] P. Mattsson, D. Zachariah, and P. Stoica, "Identification of cascade water tanks using a pwarx model,," *Mechanical Systems and Signal Processing*, vol. 106, pp. 40–48, 2018.