Proceedings of PVP 2006 / ICPVT-11
2006 ASME Pressure Vessels and Piping/ICPVT-11 Conference
July 23-27, 2006, Vancouver, BC, Canada

PVP2006-ICPVT11-93631

PARTICLE SWARM OPTIMIZATION (PSO) FUZZY SYSTEMS AND NARMAX APPROACHES TRADE-OFF APPLIED TO THERMAL-VACUUM CHAMBER IDENTIFICATION

Ernesto Araujo*
Integration and Testing Laboratory (LIT)
Instituto Nacional de Pesquisas Espaciais (INPE)
São José Campos, SP, 12.227-010, Brazil
ernesto@lit.inpe.br

Ubiratan S. Freitas
Elbert E. N. Macau
Computer Sci. and Applied Math. Associated Lab - LAC
Instituto Nacional de Pesquisas Espaciais (INPE)
São José Campos, SP, 12.227-010, Brazil
ubiratan@lac.inpe.br
elbert@lac.inpe.br

Leandro S. Coelho
Automation and Systems Laboratory (LAS-PPGEPS)
Pontifícia Universidade Católica do Paraná (PUC-PR)
Curitiba, PR, Brazil
leandro.coelho@pucpr.br

Luis A. Aguirre
Department of Electronic Engineering(DELT)
Universidade Federal de Minas Gerais (UFMG)
Belo Horizonte, MG,31270-901, Brazil
aguirre@cpdee.ufmg.br

ABSTRACT

Two nonlinear identification methods are employed in this paper in an experimental comparative approach to generate dynamical models for a thermal-vacuum system. Used for space environment emulation and satellite qualification, a thermal-vacuum chamber presents highly nonlinear and time-delay characteristics. While, in the first nonlinear identification approach, Particle Swarm Optimization (PSO) derive a Takagi-Sugeno fuzzy model, the second one was based on NARMAX polynomial identification technique. PSO is a stochastic global optimization technique that uses a population of particles, where the position and velocity of each particle represent a solution to the problem. It is employed as an auxiliary mechanism for finding out a T-S fuzzy model. The NARMAX polynomial identification technique uses a criterion called Error Reduction Ratio (ERR) computed by employing an orthogonal least squares method whose terms are selected in a forward-regression manner. Results indicate that both methods are feasible solutions for eliciting models from the available data.

INTRODUCTION

Modelling of nonlinear dynamical processes from operating data is fundamental to diverse engineering problems. A model is any sort of abstract description that captures useful relevant features able to represent a system. Finding out a model for representing dynamical behaviour is of particular importance when dealing with thermal-vacuum chambers used for satellite qualification. In doing so, the model can be employed to estimate future thermo-dynamical behaviour. This approach, then, may be applied to support operators to decide what is the best control action for conducting a thermal-vacuum qualification testing. Additional advantages of identifying a model are, for instance, the ability to detect loss of vacuum, presence of unknown heat sources or sinks, training of thermal-vacuum operators, development of a supervisor decision-support system for helping to con-

*Address all correspondence to this author.

trol the whole operation, checking the instantaneous operation or even the operator's behaviour or performance, and, ultimately design an automatic control for the whole system [1, 2].

A question that arises is which modelling approach would be more suitable for modelling this process. An alternative for modelling this nonlinear and time delay system is to employ Takagi-Sugeno (T-S) fuzzy model since this methodology exhibits both flexibility to represent high nonlinearity and robustness to uncertainty in data. This method was used previously for describing the dynamics of the thermal-vacuum system when employing a hybrid intelligent approach [2, 3]. Particle Swarm Optimization (PSO) was employed as an auxiliary mechanism for finding out T-S fuzzy models. Working in synergy, PSO allows to determine the premise space partition and to obtain membership functions and thus to extract the number of IF-THEN fuzzy rules as well to determine the statements in the consequent of the rules. Advantages of using fuzzy modelling include its simple structure to describe nonlinear systems as well as its ability to represent human being behaviours. Since human beings handle the thermal-vacuum system, this fuzzy approach seems a natural method for both designing a control system and/or modelling the dynamical system.

Despite the suitability of this method to this particular application, another question that comes up is concerning the performance of this approach in comparison to at least another method found in the literature. In this paper, NARMAX (Nonlinear Auto-Regressive Moving Average with eXogenous inputs) polynomial model in conjunction with Error Reduction Rate criterion (ERR) [4] is employed as an option to deal with nonlinear system identification. This representation describes the output of a discrete-time system as a multivariate polynomial function of previous output values and input values.

The objective concerning this paper is twofold. This paper aims to compare the fuzzy T-S/PSO method with the NARMAX polynomial identification technique as well as to explore the limits to where the previous hybrid fuzzy T-S approach may be employed in comparison to the other approach.

METHODOLOGIES EMPLOYED IN THERMAL-VACUUM MODELLING

In this paper two methodologies for identification were employed to elicit thermal-vacuum models, one of the field of artificial intelligence and the other one from the field of stochastic and dynamical analysis.

Identification of T-S/PSO Fuzzy model.

Takagi-Sugeno models consist of IF-THEN rules in which IF statements define the premise part that is featured as linguistic terms meanwhile the THEN functions constitute the consequent part characterized as linear polynomial terms:

$$R^{(j)}: \text{ IF } (z_1 \text{ IS } A_1^j) \text{ AND } \ldots \text{ AND } (z_m \text{ IS } A_m^j) \\ \text{THEN } (y = b_0^j + b_1^j x_1^j + b_{q_j}^j x_{q_j}^j), \tag{1}$$

where $z = [z_1, \ldots, z_m]^T$ is the input vector of the premise p, and $A_i^j,\ i = 1, \ldots, m$ are linguistic terms (labels) of fuzzy sets. The fuzzy sets pertaining to a rule form fuzzy regions (clusters) within the input space, A_1^j x A_2^j x ... x A_m^j. The element $x = [x_1^j, \ldots, x_{q_j}^j]^T$ represents the input vector to the consequent part of R^j that comprises q_j terms; $y_i = y_j(x^j)$ denotes the j-th rule output which is a linear polynomial of the consequent input terms u_1^j; and $b = [b_0^j, b_1^j, \ldots, b_{q_j}^j]^T$ are the polynomial coefficients that form the consequent parameter set.

Given the input vectors z and $x^j, j = 1, \ldots, M$, the final output of the fuzzy system is inferred by taking the weighted average of the local outputs $y_j(x^j)$

$$y = \sum_{j=1}^{M} v_j(\underline{z}) \cdot y_j(\underline{x}^j) \tag{2}$$

where M denotes the number of rules and $v_j(z)$ is the normalized firing strength of $R(^j)$, which is defined as

$$v_j(\underline{z}) = \frac{\mu_j(\underline{z})}{\sum_{j=1}^{M} \mu_j(\underline{z})} \tag{3}$$

and

$$\mu_j(\underline{z}) = \mu_{A_1^j}(z_1) \cdot \mu_{A_2^j}(z_2) \cdots \mu_{A_m^j}(z_m). \tag{4}$$

Linguistic labels A_i^j may be, for instance, associated with Gaussian membership functions,

$$\mu_{A_1^j}(z_i) = \exp\left[-\frac{1}{2}\frac{(z_i - m_{ij})^2}{\sigma_{ij}^2}\right] \tag{5}$$

where m_{ij} and s_{ij} are, respectively, the centres (mean value) and the spreads (standard deviations) of the Gaussian function that defines the core and the support of membership functions. The objective of the optimization process consists of determining (tuning) these parameters when using measured input-output data so that a performance measure based on the output errors is minimized:

$$\min_{\theta} \sum_{k=1}^{N} \|\widehat{y}(k+1) - y(k+1)\| \tag{6}$$

where $\widehat{y}(k+1)$ is the estimate output used for computing the square error when compared with the actual output, $y(k+1)$. This activity corresponds to the parameter-learning task and, consequently, the parameter estimation process. Nevertheless, the identification problem in T-S fuzzy modelling involves not only parameter estimation but structure selection as well. The structure selection, in turn, consists of determining the premise space partition and extracting the number of rules and determining the structure of the output elements (equations). In this paper, the structure selection of T-S system is carried out by PSO for premise part optimization while the consequent part optimization is realized by batch least mean squares method.

Particle Swarm Optimization (PSO) was developed by Kennedy and Eberhart in 1995 [5, 6]. In this population-based swarm algorithm individuals learn primarily form the successes of their neighbors. The position of each particle in PSO is a potential solution in the solution space. Each particle has an associated random velocity and moves through the problem space. At each step, each particle changes its velocity flying toward its *pbest* and *gbest* locations (global version of PSO). The term *pbest* is associated with the best solution (fitness) each particle has achieved so far and it corresponds to the coordinates in the problem space. The element *gbest* is the overall best value. The past best position and the best overall position of the group are employed to minimize (maximize) the solution. Acceleration is weighted by random terms, with separate random numbers being generated for acceleration toward *pbest* and *gbest* locations. The main steps of the global version of PSO [7] are:

1. Initialize a population (array) of particles with random positions and velocities in the n-dimensional problem space using uniform probability distribution function.
2. For each particle, evaluate its fitness value.
3. Compare each particle's fitness with the particle's *pbest*. If current value is better than *pbest*, then set *pbest* value equal to the current value and the *pbest* location equal to the current location in n-dimensional space.
4. Compare the fitness with the population's overall previous best. If current value is better than *gbest*, then reset *gbest* to the current particle's array index and value.
5. Change the velocity, v_i, and position of the particle, x_i, according to equations (7) and (8):

$$\begin{aligned} v_i(t+1) &= w \cdot v_i(t) + c_1 \cdot ud() \cdot (p_i(t) - x_i(t)) \\ &\quad + c_2 \cdot Ud() \cdot (p_g(t) - x_i(t)) \end{aligned} \tag{7}$$

$$x_i(t+1) = x_i(t) + \Delta t \cdot v_i(t+1) \tag{8}$$

6. Return to step 2 until a stop criterion is met, usually a sufficiently good fitness or a maximum number of iterations (generations).

The vector $x_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^T$ stands for the position and $v_i = [v_{i1}, v_{i2}, \ldots, v_{in}]^T$ for the velocity of the i-th particle, while $p_i = [p_{i1}, p_{i2}, \ldots, p_{in}]^T$ represents the best previous position of the i-th particle (the position giving the best fitness value). The first part in equation (7) is the momentum of the particle. The second part is related to the "cognition", i.e., it represents the independent behaviour of the particle itself. Equation (8) represents the updated position, according to its previous position and speed, considering $\Delta t = 1$. The inertia weight, w, represents the degree of the momentum of the particles and it is responsible for dynamically adjusting the velocity of the particles [8]. This parameter is accountable for balancing between local and global search. There are distinct mechanisms for chosing it. In this paper, the linear decreasing inertia function is employed since it reduces the influence of past velocities during the optimization process. The index g represents the best particle among all the particles in the group according to the fitness criteria defined latter. Variables $ud(\cdot)$ and $Ud(\cdot)$ are two random functions in the range $[0, 1]$. Positive constants c_1 and c_2 are denominated, respectively, cognitive and social components. They are selected by the user and are usually set around 2. These are the acceleration constants, responsible for varying the particle speed towards *pbest* and *gbest*. The velocity of the i-th particles on each dimension is clamped to a maximum velocity V_{max}. and it is used to determine the resolution with which the regions around the current solutions are searched.

Among a population of potential solutions to a problem, every particle has a fitness value for expressing appropriate optimization result. The function representing this quality measure employs the position of the particle, x, that is calculated after each iteration. In this paper, the fitness criteria chosen for evaluating the relationship between the real output and the estimate output during the optimization process was the *Pearson multiple correlation coefficient*. This coefficient gives the rate between the variability of two measures (variables) in which one is described by the variability of the other. This Coefficient represents the harmonic mean of R of training and validation phases of T-S fuzzy model conducted by $R^2_{harmonic}$ as given by:

$$R^2_{harmonic} = \frac{2}{\dfrac{1}{R^2_{training} + \varepsilon} + \dfrac{1}{R^2_{validation} + \varepsilon}}, \tag{9}$$

where:

$$\begin{aligned} R^2_{training} &= 1 - \frac{\sum_{k=1}^{0.5Na} [y(k) - \widehat{y}(k)]^2}{\sum_{k=1}^{0.5Na} [y(k) - \overline{y}]^2} \\ R^2_{validation} &= 1 - \frac{\sum_{k=0.5Na+1}^{Na} [y(k) - \widehat{y}(k)]^2}{\sum_{k=0.5Na+1}^{Na} [y(k) - \overline{y}]^2} \end{aligned} \tag{10}$$

are, respectively, the R^2-training (estimation) and R^2-validation phases of the model; ε is the small tolerance value (10^{-16}), Na is the total number of samples evaluated, and $\overline{y}$ is the system real output. When $R(\cdot)^2$ is close to unit, $R(\cdot)^2 = 1.0$, a sufficient accurate model for the measured data of the system is found. A R^2 between 0.9 and 1.0 is suitable for applications in identification and model-based control [9].

NARMAX polynomial identification technique.

The NARMAX (Nonlinear Auto-Regressive Moving Average with eXogenous inputs) models [4] are nonlinear discrete-time dynamical systems. The models used here are of polynomial type. The general NARMAX polynomial model is represented by

$$y(k) = \sum_{i=1}^{p} \theta_i x_i(k) + \xi(k) \,, \tag{11}$$

where $y(k)$ is the model output at the discrete instant k, $\xi(k)$ is the one-step prediction error or residue, $x_i(k)$ is the i-th *term*, θ_i is its *coefficient*. The number of terms is p. The terms are given by

$$\begin{aligned} x_i(k) = {} & y(k-1)^{\alpha_{i1}} y(k-2)^{\alpha_{i2}} \cdots y(k-n_y)^{\alpha_{in_y}} \\ & \cdot u(k-1)^{\beta_{i1}} u(k-2)^{\beta_{i2}} \cdots u(k-n_u)^{\beta_{in_u}} \\ & \cdot e(k-1)^{\gamma_{i1}} e(k-2)^{\gamma_{i2}} \cdots e(k-n_e)^{\gamma_{in_e}} \,, \end{aligned} \tag{12}$$

where $u(k)$ is the model input at k, $e(k)$ is a noise term responsible for modelling the stochastic part of the data, $(\alpha_{ij}, \beta_{ij}, \gamma_{ij}) \in \mathbb{N}^3$ are integer exponents, and n_y, n_u and n_e are the maximum lag in y, u and e respectively. The quantity

$$d_i = \sum_{j=1}^{n_y} \alpha_{ij} + \sum_{j=1}^{n_u} \beta_{ij} + \sum_{j=1}^{n_e} \gamma_{ij} \tag{13}$$

is defined as the *degree* of term i.

The problem of identifying a NARMAX model given by (11) can be split into two parts. The first one consists of finding *which* terms x_i of an arbitrary candidate set of terms C should be included in the model. Such choice must be done in agreement with the particular system one is trying identify. This problem is called *structure selection*. The second part consists in finding suitable coefficients θ_i such that the model can correctly represent the system being identified. This problem is called *parameter estimation*.

The structure selection problem is specially important in nonlinear systems identification. A nonlinear model has more flexibility in its structure then a linear one. Also, the consequences of a wrong structure are usually more severe in a nonlinear model than in a linear one [10]. Therefore, the structure selection problem must be addressed carefully. In this work, we use a criterion called Error Reduction Ratio (ERR) [4, 11, 12]. The ERR is defined as the fraction of the variance of the output signal $y(k)$ that can be explained by a particular term of the model. The ERR is computed using an orthogonal least squares method and the terms are selected in a forward-regression manner, step by step. Assume that N samples were taken. Then, equation (11) can be written once at each sample, thus resulting in the matrix equation

$$Y = X\Theta + \Xi \,, \tag{14}$$

where $Y = [y(1), y(2), \ldots, y(N)]^T$ is the vector of outputs, $X = [\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_p]$ with $\underline{x}_i = [x_i(1), x_i(2), \ldots, x_i(N)]^T$ is the regressor matrix, $\Theta = [\theta_1, \theta_2, \ldots, \theta_p]^T$ is the parameter vector, and $\Xi = [\xi(1), \xi(2), \ldots, \xi(N)]^T$ is the residue vector.

If X has full column rank, it can be decomposed as $X = MA$, where A is a $p \times p$ unit upper triangular matrix and M is an $N \times p$ matrix whose columns, $\underline{m}_i$, $i = 1, \ldots, p$, are pairwise orthogonal. Then, (14) can be rewritten as

$$Y = (XA^{-1})(A\Theta) + \Xi = MG + \Xi \,, \tag{15}$$

where $G = A\Theta = [g_1, g_2, \ldots, g_p]^T$ is an auxiliary parameter vector that can be computed by

$$g_i = \frac{Y^T \underline{m}_i}{\underline{m}_i^T \underline{m}_i} \quad i = 1, 2, \ldots, p \,. \tag{16}$$

The ERR of the i-th orthogonal term $\underline{m}_i$ is then defined as

$$\mathrm{ERR}_i = \frac{g_i \underline{m}_i^T \underline{m}_i}{Y^T Y} = \frac{(Y^T \underline{m}_i)2}{(Y^T Y)(\underline{m}_i^T \underline{m}_i)} \,. \tag{17}$$

The procedure for constructing a parsimonious model using the ERR as a term selection criterion is the following.

1. Start with an empty model and a set of candidate terms C. One way to form such set is generating all possible terms up to some chosen maxima degree, n_y, n_u and n_e.
2. For each term x_i in C, compute its corresponding orthogonal regressor by forming an orthogonal basis with the other terms already in the model. Compute the ERR of this orthogonal regressor.
3. Chose the term of largest ERR computed in the previous step. Include it in the model and remove it from C.

4. Repeat steps 2. and 3. up to a stop criterion, such as a desired residue variance or a maximum number of terms. At the end, there will be a set of terms in the model that parsimoniously represent the data.

The orthogonal decomposition produced can then be used to compute the model coefficients, i. e. perform the parameter estimation part of the identification problem.

There is, however, one unsolved issue: the noise modelling. In equation (11), in order to compute the terms that involve e, one would need to know the noise that corrupted the system, which is usually unavailable. This is solved by an iterative estimator known as the *Extended Least Squares*. In brief, this approach starts the identification with no terms that involve e. Afterwards, a model is obtained and the residue vector Ξ is computed. The residue given by $\xi(k)$ at the instant k is then employed as estimate for the noise, $e(k)$, and another model, which includes noise terms, is computed. Finally, a new residue vector is obtained and used as a new estimate for $e(k)$. This procedure is repeated until the convergence of the parameters is accomplished.

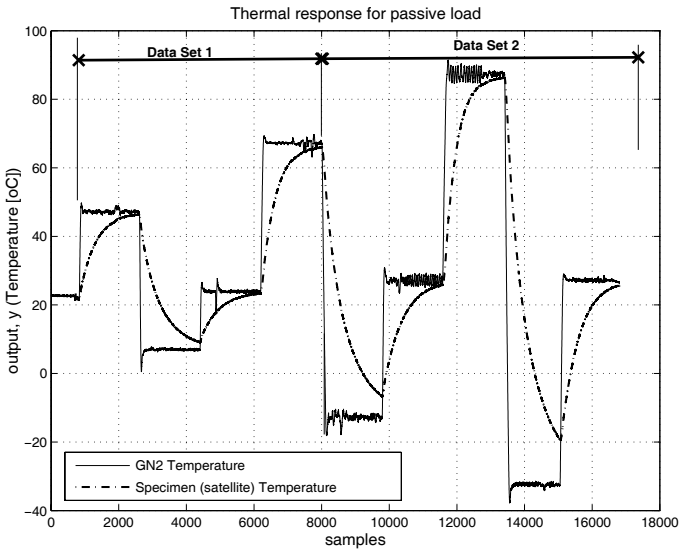


Figure 1. Temperature on the satellite and in the gas used for nonlinear identification.

PROBLEM FORMULATION

Thermal-vacuum systems consist of a chamber, a shroud (set of pipes) which heats or cools off the environment, and some devices and auxiliary equipment able to reproduce the conditions of expected post-launch space environments [13, 14].

Since pressure inside thermal-vacuum chamber is low, there is no convective heat transfer. Furthermore, temperature gradient inside the payload may be considered negligible if there is fast heat conduction inside the payload. Because radiation is the main source of heat transfer between the payload and shroud, thermal vacuum chambers are inherently nonlinear. This heat transfer depends nonlinearly on absolute temperature, T^4 [1, 15]. In a glimpse, this nonlinear behaviour may be confirmed by real-world industrial dynamical response corresponding to the thermal-vacuum system with passive load into it (Fig. 1). The solid line represents the temperature of the gas inside the shroud and corresponds to a typical temperature set up for thermal-vacuum environmental simulation for sattelite qualification. The dashed line, in turn, is the temperature on the satellite and represents its thermal response when submitted to step values of temperature.

The schematic and simplified diagram that depicts the operational characteristic of the thermal-vacuum chamber and the identification block is presented in Fig. 2. This paper focus on the nonlinear identification for modelling the relationship between the temperature on the satellite (output) and the controlled temperature of the gas inside shroud (input) which is used to change the temperature in the interior of the chamber.

Both temperature measured on the satellite and in the gas of the shroud employed to elicit the nonlinear models shown in Fig. 1 are split in two data subsets. The first data set (data set 1) was collected in one day while the second one (data set 2) was acquired a couple of days latter yielding relative independence between them. Such independence allows the use of different data sets for validation and identification.

SIMULATION RESULTS

The two identification approaches were performed using the data shown in Fig. 1. Instead of eliciting models based on purely sampled data (black-box approach), a hybrid semi-mechanistic approach (grey-box approach) is employed. In doing so, the temperature data was converted from degrees Celsius to Kelvin since radiation is the main source of heat transfer and the nonlinear characteristics of this process are naturally represented in absolute (Kelvin) temperatures. Afterwards, the resulting data was normalized, that is, they were divided by its maximum absolute value so all the data was in the range $(0, 1]$. Finally, data was also decimated by taking one sample each four.

The data set 1, in Fig. 1, consists of 1900 points and it was used in the parameter estimation procedure (training) while the entire set, comprising 3500 points, (data set 1 plus data set 2) was used for validation. This approach was applied in both identification methods.

The methodology of comparison includes that both Takagi-Sugeno Fuzzy models and NARMAX polynomial models were computed by using 3 inputs, $y(k-1)$, $y(k-2)$ and $u(k-1)$, i.e., it was considered here a second order model. The validation activity is carried out in two experiments: one-step ahead simulation and free simulation.

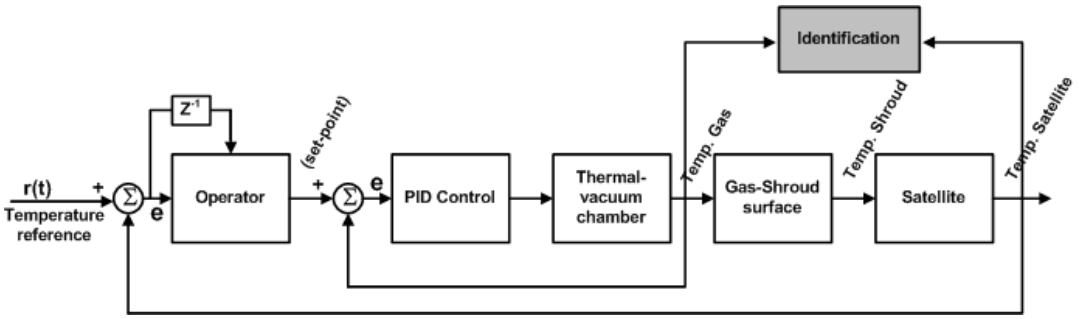


Figure 2. Simplified block diagram for thermal-vacuum chamber and identification process.

The number of rules and membership functions in the T-S/PSO approach were set as 3. Thus, PSO needs to deal with a vector whose elements are 9 centers and 3 spreads of a Gaussian membership function. The initial position of particles is randomly generated through an uniform distribution. The sufficient number of particles for this application was set as 5. Social and cognitive components were equally fixed as 2.05 while inertia weight was chosen to be linear decreasing departing from 0.729 and reaching 0.4 as its final value. It is worth mentioning that only 6 iterations were enough for accomplishing final T-S fuzzy models, that is, a very low computational cost.

The candidate term set C used in the NARMAX polynomial approach was formed by taking *all* terms up to degree 4, n_y up to 2 (i.e., up to second order) and $n_u = 1$. The noise terms employed were linear and up to $n_e = 5$. Models with total number of terms varying from 1 to 15, plus 5 noise terms, were computed. From those, the best one was chosen.

The one-step ahead simulation of the best T-S/PSO model and of the best NARMAX polynomial model along with the original system response is shown in Fig. 3(a). The output for the obtained models and the original system (specimen) are so close that they can hardly be distinguished. The one-step ahead prediction error for this approach is shown in Fig. 3(b).

The free simulation of both these models and the system's output are shown in Fig. 4(a). The free simulation prediction error is depicted in Fig. 4(b). As expected, the error of both models is smaller on the data used for training than on the second part of data.

CONCLUSION

In this paper, two nonlinear identification methods were employed in eliciting models for representing thermal-vacuum chambers used to qualify space devices, and satellites. While Takagi-Sugeno fuzzy modelling based on Particle Swarm Optimization (PSO) is a technique that uses a nonlinear-parameter representation (linear in the consequent of the rule but nonlinear for the whole process) the other approach, NARMAX polynomial modelling with ERR criterion, employs a linear-parameter representation. Results indicate that both approaches were able to generate satisfactory nonlinear models for one-step ahead and free simulation forecasting. When dealing with one-step ahead prediction the proposed methods yields scarcely distinguishable outputs. However, when free simulation was achieved, results show that the fuzzy model generates an error greater than the one obtained by NARMAX polynomial approach.

The performance difference of the two approaches seems to lie in the nonlinear characteristics of T-S/PSO representation when compared with the NARMAX polynomial one. Being nonlinear in the parameters, the former needs a nonlinear optimization technique for estimating its parameters and this makes finding a good model more difficult, which requires an extensive search for a T-S/PSO model at least as good as the NARMAX model. This will be done in a future work.

REFERENCES

[1] Jen, C.-L., and Tilwick, L., 2000. "On-line, self-learning, predictive toll for determining payload thermal response". In Proceedings of Space Simulation, pp. 193–200.

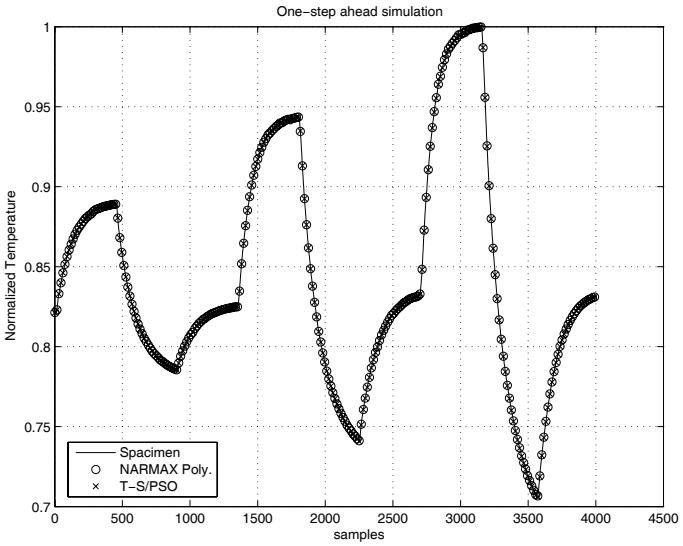
[2] Araujo, E., and Coelho, L. S., 2006. *Systems Engineering using Swarm Particle Optimisation*. ISE Book Series on System Engineering. Nova Science Publishers, New York, ch. Piecewise Reference-driven Takagi-Sugeno Fuzzy Modelling based on Particle Swarm Optimization (PSO), pp. 85–112. N. Nedjah, L. M. Mourelle (Eds.).

[3] Marinke, R., Araujo, E., Coelho, L. S., and Matiko, I., 2005. "Particle swarm optimization (PSO) applied to fuzzy modelling in a thermal-vacuum system". In International Conference on Hybrid Intelligent Systems, pp. 67–72.

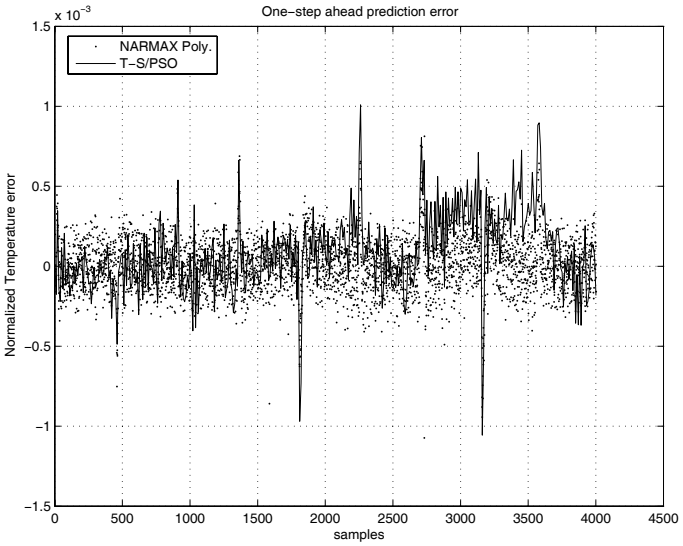
[4] Billings, S. A., Chen, S., and Korenberg, M. J., 1989. "Identification of MIMO nonlinear systems using a forward-regression orthogonal estimator". *Int. J. Control,* **49**(6), pp. 2157–2189.

[5] Kennedy, J., and Eberhart, R., 1995. "Particle swarm optimization". In IEEE International Conference on Neural Networks, pp. 1942–1948.

[6] Eberhart, R., and Kennedy, J., 1995. "A new optimizer us-

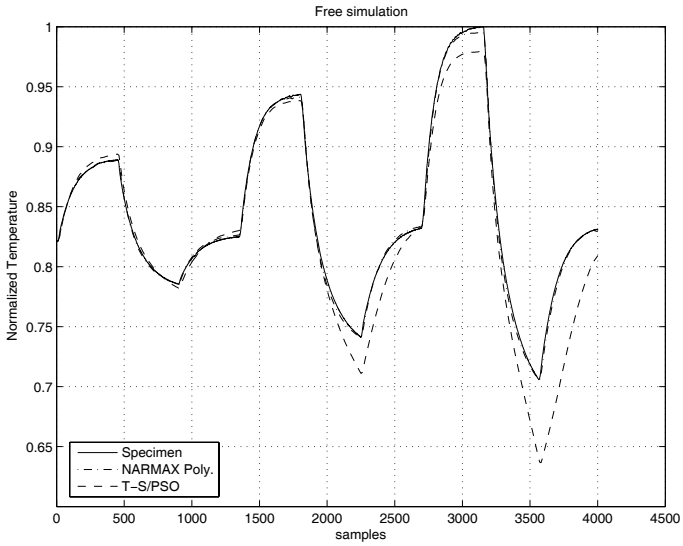


(a) Measured response and output estimation for one-step ahead simulation.

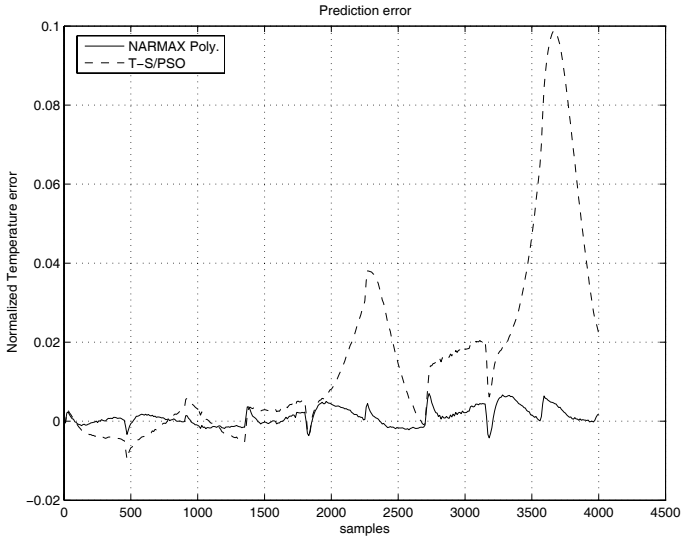


(b) One-step ahead simulation error.

Figure 3. One-step ahead simulation.



(a) Measured response and output estimation for free simulation.



(b) Free simulation error.

Figure 4. Free simulation.

ing particle swarm theory". In Proceedings of the Sixth International Symposium on Micro Machine and Human Science, pp. 39–43.

[7] Krohling, R. A., Hoffmann, F., and Coelho, L. S., 2004. "Co-evolutionary particle swarm optimization for min-max problems using gaussian distribution". In Proceedings of the Congress on Evolutionary Computation, pp. 959–964.

[8] Shi, Y., and Eberhart, R. C., 1998. "Parameter selection in particle swarm optimizer". In Proceedings of Annual Conference on Evolutionary Programming, V. W. Porto, N. Saravan, D. Waagen, , and A. E. Eiben, eds., Springer-Verlag, pp. 591–601.

[9] Schaible, B., Xie, H., and Lee, Y. C., 1997. "Fuzzy logic models for ranking process effects". *IEEE Transactions on Fuzzy Systems,* **5**(4), pp. 545–556.

[10] Aguirre, L. A., and Billings, S. A., 1995. "Improved structure selection for nonlinear models based on term clustering". *Int. J. Control,* **62**(3), pp. 569–587.

[11] Chen, S., Billings, S. A., and Luo, W., 1989. "Orthogonal least squares methods and their application to nonlinear system identification". *Int. J. Control,* **50**(5), pp. 1873–1896.

[12] Billings, S. A., and Wei, H.-L., 2005. "A new class of wavelet networks for nonlinear system identification". *IEEE Trans. Neural Networks,* **16**(4), pp. 862–874.

[13] HIGH VACUUM SYSTEMS INC. (HVSI), 1997. *Operations and Maintenance Manual - 250 dm3 Thermal Vacuum System with Thermally Conditioned Platen*, October.

[14] Garner, J. T., 1996. *Satellite control - a comprehensive approach*. John Wiley & Sons Ltd. and Praxis Publishing Ltd., Winchester.

[15] Gilmore, D. G., 1994. *Satellite thermal control handbook*. The Aerospace Corporation Press, California.