Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

A new model validity monitoring method for nonlinear recursive identification

Li Feng Zhang*. Quan Min Zhu, Ashely Longden*****

Faculty of Computing, Engineering and Mathematical Sciences
University of the West of England, Frenchay Campus
Coldharbour Lane, Bristol, BS16 1QY, UK
lifeng.zhang@hotmail.co.uk, **quan.zhu@uwe.ac.uk, * ashley.longden@uwe.ac.uk*

Abstract: In the present study, a new correlation test based model validity monitoring procedure is proposed to online check the quality of nonlinear recursively identified models. The new method provides a simple but effective diagnose of nonlinear recursive models by detecting if the residuals are reduced to uncorrelated noise sequences. In the monitoring procedure, the correlation functions are periodically computed with a specified frequency and a constant data window. The computational time and sensitivity of the correlation tests can be easily modified by adjusting the testing time interval, data length, and maximum lag. A simulated case study is employed to demonstrate the effectiveness and efficiency of the new method.

1. INTRODUCTION

Recursive identification deals with the problem of building equivalent models of unknown systems when the systems are in operation. It has been widely applied in many fields such as adaptive control and adaptive signal processing. In many cases, the recursive models may be applied online in parallel with the recursive identification. Hence, real time model validity detection is important to check the quality of the models during the identification procedure. Furthermore, it is sometimes difficult to effectively check the goodness of the models by only comparing the observed and predicted data sequences since, in practice, the data are always measured in unknown noisy environments. Therefore, efficient and effective model validity monitoring methods need to be developed.

In the last three decades, several correlation test based model validation methods have been developed to detect the quality of a wide range of linear and nonlinear models (Bohlin, 1971, 1978, Billings and Voon, 1983, 1986, Söderström and Stoica,1990, Billings and Zhu 1994, 1995, Mao and Billings 2000). Based on a common assumption that the system can be adequately described by a model, all these methods detect the validity of the identified model by checking if the residuals are reduced to a white noise sequence and independent to the delayed inputs and outputs. In 2007, two sets of first order correlation functions named combined omni-directional auto-correlation function (ODACF) and combined omni-directional cross-correlation function (ODCCF) have been proposed to detect linear and nonlinear associations between variables. The new tests provide a more effective and comprehensive correlation detection by separately investigating the associations between both the absolute values and signs of the analysed variables. Then, they have been used to construct a set of new nonlinear model validity tests (Zhang *et al*., 2007, Zhu *et al*., 2007). However, all these methods are designed for validating the identified models after the structures and parameters of the models are determined. It is obviously that they cannot be directly used to validate recursive models since in recursive identification the parameters are sequentially estimated.

To overcome this problem, a new online correlation test procedure is proposed in this study to monitor the validity of nonlinear models during the real-time estimation procedures. In the new method, the combined ODACF and ODCCF tests between residuals, inputs and outputs are repeatedly computed with a specified frequency and only the latest observed data with a specified data length is involved in the computation. In other words, the correlation functions are computed periodically with a constant moving data window. It is believed that the new method can provide simple and effective online validity detection to a wide range of recursive models including intelligent models and adaptive noise cancellers.

This study is organised as follows. In section 2, a brief introduction to recursive identification and correlation test based model validation is presented. In section 3, a new combined ODACF and ODCCF based model validity monitoring method is proposed. A simulated case study, then, is employed in section 4 to demonstrate the new method. Finally, in section 5 conclusions are drawn to summarise the study.

2. RECURSIVE IDENTIFICATION AND CORRELATION TESTS BASED MODEL VALIDATION

Initially, the basic principles of recursive identification and correlation tests based model validation are introduced.

2.1 Recursive Identification

Consider a generalised single input and single output (SISO) nonlinear model.

$$y(t) = \hat{y}(t) + \varepsilon(t) = \hat{f}(y(t-1),...,y(t-n_y), u(t-1),...,u(t-n_u), \varepsilon(t-1),...,\varepsilon(t-n_\varepsilon)) + \varepsilon(t) \quad (1)$$

where $\hat{f}(\cdot)$ denotes an identified nonlinear model. $\hat{y}(t)$, $y(t)$, $u(t)$, and $\varepsilon(t)$ respectively denote the predicted outputs, measured outputs, inputs, and residuals. A typical parametric expression of (1) is Nonlinear Auto-Regressive Moving Average with eXogenous input (NARMAX) model (Leontaritis and Billings 1985) formulated as

$$y(t) = \boldsymbol{\varphi}^T(t)\boldsymbol{\theta} + \varepsilon(t) = \sum_{i=1}^{r} \theta_i \varphi_i(t) + \varepsilon(t) \quad (2)$$

where $\boldsymbol{\theta}$ denotes the parameter vector, and $\boldsymbol{\varphi}(t)$ denotes the vector valued linear or nonlinear terms.

To estimate the parameters in (2), many offline parameter estimation methods have been developed. Nevertheless, the computational complexity of these methods increases rapidly with the amount of data. Hence, they are not suitable for real-time system identification where a new model will be estimated within each sampling instant by exploiting the information contained in the new collected data. In many cases, for the sake of making real-time decisions, it is necessary to identify the models online when the systems are in operation. To cope with this problem, many recursive identification methods have been developed, such as the least mean square (LMS) method, the recursive least square (RLS) method, the recursive instrumental variables (RIV) method, and the recursive prediction error method (RPEM). At time instant *t*, these methods estimate the new parameters based on the estimated parameter values at the previous sampling instant *t*-1 and the new observed data samples at *t*. Hence the computational complexity of these methods will not increase with the amount of data samples. A typical recursive identification method can be mathematically described as follows. Consider the general form of a nonlinear polynomial model expressed as

$$\hat{y}(t) = \boldsymbol{\varphi}^T(t)\boldsymbol{\theta}(t-1) \quad (3)$$

where $\boldsymbol{\theta}(t)$ denotes the vector valued estimated parameters at time instant *t*. $\boldsymbol{\theta}(t)$ can be derived as follows.

$$\boldsymbol{\theta}(t) = \boldsymbol{\theta}(t-1) + \boldsymbol{\gamma}(t)(y(t) - \hat{y}(t)) \quad (4)$$

where $y(t) - \hat{y}(t)$ denotes the prediction error at *t*. $\boldsymbol{\gamma}(t)$ is defined as the correction function vector that determines how the current prediction error affects the update of the estimated parameter vector.

2.2 Correlation tests based offline model validation

Model validation is the final step of any system identification procedure to check if the identified model is an adequate approximation of the underlying system (Ljung, 1999, Wigren, 2003). It is widely accepted that if a model is valid the residuals should be reduced to a white noise sequence with zero mean, finite variance, and independent to the delayed inputs, outputs. Auto-correlation function (ACF) and cross-correlation function (CCF) have been widely applied in linear model validation to check if the residuals are uncorrelated to the delayed residuals, inputs, and outputs (Bohlin, 1971, 9178, Söderström and Stoica,1990). To validate nonlinear models, several higher-order correlation tests based approaches have been proposed (Billings and Voon, 1983, 1986, Billings and Zhu 1994, 1995, Mao and Billings 2000).

Recently, two sets of first order correlation functions named combined omni-directional auto-correlation functions (ODACFs) and combined omni-directional cross-correlation functions (ODCCFs) have been proposed to detect nonlinear associations between variables. (Zhang *et al*., 2007, Zhu *et al*., 2007). Consider that $r_{u\varepsilon}(\tau)$ denotes the general form of CCF between $u(t)$ and $\varepsilon(t)$ with a time delay element of τ. ODCCF ($R_{u\varepsilon}(\tau)$), which includes four first-order correlation tests, can be simply formulated as

$$R_{u\varepsilon}(\tau) = \lfloor r_{\alpha'\beta'}(\tau), r_{\alpha'\varepsilon'}(\tau), r_{u'\varepsilon'}(\tau), r_{u'\beta'}(\tau) \rfloor \quad (5)$$

where $\alpha(t) = |u(t)'| = |u(t) - \bar{u}|$, $\beta(t) = |\varepsilon(t)'| = |\varepsilon(t) - \bar{\varepsilon}|$, and prime ’ denotes that the mean value has been removed from the corresponding data sequence. Then, the results obtained from using ODCCFs are combined together to constitute a more condensed correlation function named combined ODCCF ($\rho_{u\varepsilon}(\tau)$) derived as follows.

$$\begin{cases} if\ \ |\max(R_{u\varepsilon}(\tau))| > |\min(R_{u\varepsilon}(\tau))| \\ \quad \rho_{u\varepsilon}(\tau) = \max(R_{u\varepsilon}(\tau)) \\ else \\ \quad \rho_{u\varepsilon}(\tau) = \min(R_{u\varepsilon}(\tau)) \end{cases} \quad (6)$$

For the special case that $u(t) = \varepsilon(t)$, (5) and (6) are called ODACFs and combined ODACF respectively.

Based on these correlation functions, a set of nonlinear model validity tests have been proposed in the study of Zhu *et al.* (2007) to detect the omitted nonlinear terms in the residuals. For a valid model, the correlation tests can be derived as follows

Combined ODACF validation of residuals

$$\begin{cases} \rho_{\varepsilon\varepsilon}(\tau) = 1, \tau = 0 \\ \rho_{\varepsilon\varepsilon}(\tau) = 0, otherwise \end{cases} \quad (7)$$

Combined ODCCF validation between inputs and residuals:

$$\rho_{\varepsilon u}(\tau) = 0, \forall\, \tau \quad (8)$$

Combined ODCCF validation between outputs and residuals:

$$\begin{cases} 0 < |\rho_{\varepsilon y}(\tau)| < 1, \tau = 0 \\ \rho_{\varepsilon y}(\tau) = 0, otherwise \end{cases} \quad (9)$$

Compared to the previous methods, combined ODACF and ODCCF tests provide an enhanced nonlinear correlation detection power and a more condensed correlation illustration.

These methods, however, are not applicable in recursive identification since they all operate on the whole collected data set. The computational cost of the correlation tests will increase rapidly with the amount of the collected data. Particularly, as $t \to \infty$ the total data length will become extremely large.

3. A NEW MODEL VALIDITY MONITORING PROCEDURE

In this study, a new model validity monitoring method is proposed based on the combined ODACF and ODCCF tests introduced above. In the method, the correlation functions are repeatedly computed in parallel with the recursive identification. For reducing the computational cost without losing effectiveness, they are not computed at all t but computed periodically. To enhance the tracking capability and sensitivity, only the latest parts (a data window with specified length of N_r) of the data sequences are used to compute the correlation functions. It makes the computational time of each step constant, and avoids a step by step increasing computational cost. The online correlation test procedure is presented as follows.

1. Determine the time interval t_r between two steps of correlation computations, maximum lag τ_r, and maximum data length N_r for the correlation tests.

   Remark 1: The values of t_r, τ_r, and N_r are chosen by the user. Generally, τ_r needs to be selected lager than the order of the model. N_r musts be selected larger than t_r. Furthermore, the choice of N_r will influence the strict of the validity test that a larger N_r will result in a smaller confidence interval and a more strict validation. All these values have to be determined carefully to ensure that each computation of correlation functions can be finished within the time interval t_r.

2. For $t \le N_r$, compute the combined ODACF and ODCCF defined in (7) to (9) at time instant $t = t_r, 2t_r, 3t_r, ..., kt_r$ ($N_r \ge kt_r > N_r - t_r$) by using the testing data sequences $\boldsymbol{\varepsilon}_{r,(t)}$, $\mathbf{y}_{r,(t)}$ and $\mathbf{u}_{r,(t)}$ given as

$$\begin{cases} \boldsymbol{\varepsilon}_{r,(t)} = [\varepsilon(0),...,\varepsilon(t)]^{\mathrm{T}} \\ \mathbf{y}_{r,(t)} = [y(0),...,y(t)]^{\mathrm{T}} \\ \mathbf{u}_{r,(t)} = [u(0),...,u(t)]^{\mathrm{T}} \end{cases} \tag{10}$$

   Remark 2: It should be noticed that $\rho_{\varepsilon\varepsilon}(0)$ and $\rho_{y\varepsilon}(0)$ are not computed since theoretically they are nonzero numbers and unnecessary to the validity detection.

   Then, the correlation tests for a recursively identified model at t are derived as follows.

$$\begin{cases} \rho_{\varepsilon_r\varepsilon_r,(t)}(\tau) = 0, \quad 0 < \tau \le \tau_r, \\ \rho_{u_r\varepsilon_r,(t)}(\tau) = 0, \quad 0 \le \tau \le \tau_r \\ \rho_{y_r\varepsilon_r,(t)}(\tau) = 0, \quad 0 < \tau \le \tau_r \end{cases} \tag{11}$$

   To provide a better illustration and reduce the number of correlation plots, the results obtained from (11) are combined together as

$$\begin{cases} \Theta_{\varepsilon\varepsilon}(t) = \rho_{\varepsilon_r\varepsilon_r,(t)}(\tau), \\ \left|\rho_{\varepsilon_r\varepsilon_r,(t)}(\tau)\right| = \max(\left|\rho_{\varepsilon_r\varepsilon_r,(t)}(1)\right|,...,\left|\rho_{\varepsilon_r\varepsilon_r,(t)}(\tau_r)\right|) \\ \Theta_{u\varepsilon}(t) = \rho_{u_r\varepsilon_r,(t)}(\tau), \\ \left|\rho_{u_r\varepsilon_r,(t)}(\tau)\right| = \max(\left|\rho_{u_r\varepsilon_r,(t)}(0)\right|,...,\left|\rho_{u_r\varepsilon_r,(t)}(\tau_r)\right|) \\ \Theta_{y\varepsilon}(t) = \rho_{y_r\varepsilon_r,(t)}(\tau), \\ \left|\rho_{y_r\varepsilon_r,(t)}(\tau)\right| = \max(\left|\rho_{y_r\varepsilon_r,(t)}(1)\right|,...,\left|\rho_{y_r\varepsilon_r,(t)}(\tau_r)\right|) \end{cases} \tag{12}$$

3. For $t > N_r$, compute the correlation functions at time instant $t = (k+1)t_r, (k+2)t_r, ..., Kt_r$ ($N \ge Kt_r > N - t_r$) with a fixed data length of N_r.

   The testing data sequences $\boldsymbol{\varepsilon}_{r,(t)}$, $\mathbf{y}_{r,(t)}$ and $\mathbf{u}_{r,(t)}$ are defined as

$$\begin{cases} \boldsymbol{\varepsilon}_{r,(t)} = [\varepsilon(t-N_r+1), \varepsilon(t-N_r+2),...,\varepsilon(t)]^T \\ \mathbf{y}_{r,(t)} = [y(t-N_r+1), y(t-N_r+2),...,y(t)]^T \\ \mathbf{u}_{r,(t)} = [u(t-N_r+1), u(t-N_r+2),...,u(t)]^T \end{cases} \tag{13}$$

   The computational procedure of the correlation functions in this stage is same as that in stage 2.

Then, $\Theta_{\varepsilon\varepsilon}(t)$, $\Theta_{u\varepsilon}(t)$, and $\Theta_{y\varepsilon}(t)$ are used to online detect the quality of the recursive models. Theoretically, if a model is recursively identified adequately at t, the correlation functions can be derived as

$$\begin{cases} \Theta_{\varepsilon\varepsilon}(t) = 0 \\ \Theta_{u\varepsilon}(t) = 0 \\ \Theta_{y\varepsilon}(t) = 0 \end{cases} \tag{14}$$

According to the central limit theorem, the estimates of $\Theta_{\varepsilon\varepsilon}(t)$, $\Theta_{u\varepsilon}(t)$ and $\Theta_{y\varepsilon}(t)$ are asymptotically normal with zero mean and finite variance (Bowker and Lieberman, 1972). In practice, a model can be considered as valid when all the correlation functions lie within the 95% confidence interval. The confidence intervals for the online correlation tests are derived as follows.

1. For $t \le N_r$, the 95% confidence intervals are computed as $\pm 1.95/\sqrt{t}$ where $t = t_r, 2t_r, 3t_r, ..., kt_r$.

2. For $t > N_r$, the 95% confidence intervals are computed as $\pm 1.95/\sqrt{N_r}$

Remark 3: At the beginning of the recursive identification, $\Theta_{\varepsilon\varepsilon}(t)$, $\Theta_{u\varepsilon}(t)$, and $\Theta_{y\varepsilon}(t)$ should lie significantly outside the confidence interval since the model is unfitted. If the model structure and parameter estimation algorithm are selected correctly, the correlation functions will decrease as the model is recursively fitted. Finally, they will converge to the confidence interval while a valid model is achieved.

Remark 4: It should be noticed that all $\Theta(t)$ indicate the worst correlation values (with the highest amplitude) of $\rho(1),\ldots,\rho(\tau_r)$ at time instant *t*. For large τ_r there is always one or more correlation functions lie around the confidence interval. Consequently, a $\Theta(t)$, which closes to the confidence limits, is also acceptable.

4. SIMULATION STUDY

4.1 Simulation and Identification of a Hammerstein System

In this section, a simulated Hammerstein system (Kung and Shinh, 1986) was employed to demonstrate the method. Consider a Hammerstein system expressed as follows

$$\begin{cases} z(t) = 0.9z(t-1) - 0.1z(t-2) + 0.08u(t-1) \\ \quad - 0.13u(t-2)u(t-1) + 0.05u^2(t-2) \\ y(t) = z(t) + e(t) \end{cases} \tag{15}$$

where $u(t)$ was selected as a uniformly distributed input sequence with zero mean and amplitude from -1 to 1. The additive white noise $e(t)$ was selected as a normally distributed data sequence with zero mean and variance of 1×10^{-3}. All the data sequences were generated with length of 3000. Fig. 1 shows the data sequences for system (15).

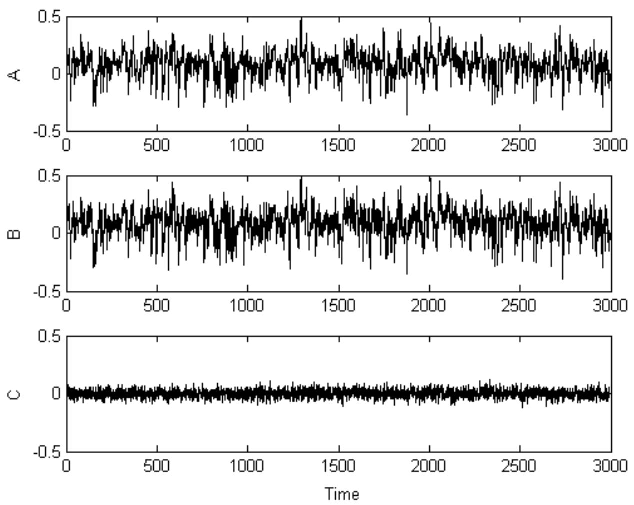

Fig. 1. The original data sequences for system (15): A. noise free outputs; B. measured outputs; C. additive noise.

Consider three candidate models formulated as follows.

$$\begin{cases} \hat{y}_1(t) = a_1 y(t-1) + a_2 y(t-2) + a_3 y(t-3) + a_4 y(t-4) \\ \quad + a_5 y(t-5) + b_1 u(t-1) + b_2 u(t-2) + b_3 u(t-3) \\ \quad + b_4 u(t-4) + b_5 u(t-5) \\ \varepsilon_1(t) = y(t) - \hat{y}_1(t) \end{cases} \tag{16}$$

$$\begin{cases} \hat{y}_2(t) = a_1 y(t-1) + a_2 y(t-2) + b_1 u(t-1) \\ \quad + b_1 u^2(t-2) + b_3 u(t-1)u(t-2) \\ \varepsilon_2(t) = y(t) - \hat{y}_2(t) \end{cases} \tag{17}$$

$$\begin{cases} \hat{y}_3(t) = a_1 y(t-1) + a_2 y(t-2) + a_3 y(t-3) + a_4 y(t-4) \\ \quad + a_5 y(t-5) + b_1 u(t-1) + b_2 u(t-2) + b_3 u(t-3) \\ \quad + b_4 u(t-4) + b_5 u^2(t-2) + b_6 u^2(t-3) + b_7 u^2(t-4) \\ \quad + b_8 u^2(t-5) + b_9 u(t-1)u(t-2) + b_{10} u(t-2)u(t-3) \\ \quad + b_{11} u(t-3)u(t-4) + b_{12} u(t-4)u(t-5) \\ \varepsilon_3(t) = y(t) - \hat{y}_3(t) \end{cases} \tag{18}$$

In this study, RLS was used to estimate the parameters associated in the models.

Moreover, to compare the performance of the models, an error signal $\xi(t)$, which is the difference between noise free outputs and predicted outputs, is computed as follows.

$$\xi(t) = z(t) - \hat{y}(t) \tag{19}$$

It should be notice that, in practice, it is impossible to obtain $\xi(t)$ since $z(t)$ is always unknown. In simulation studies, $z(t)$ is defined in advance so that $\xi(t)$ can be obtained to indicate the performance of the models. In this study, it is used only for providing visual illustrations and comparisons.

Figs. 2 to 4 depict the predicted data sequences for the three models.

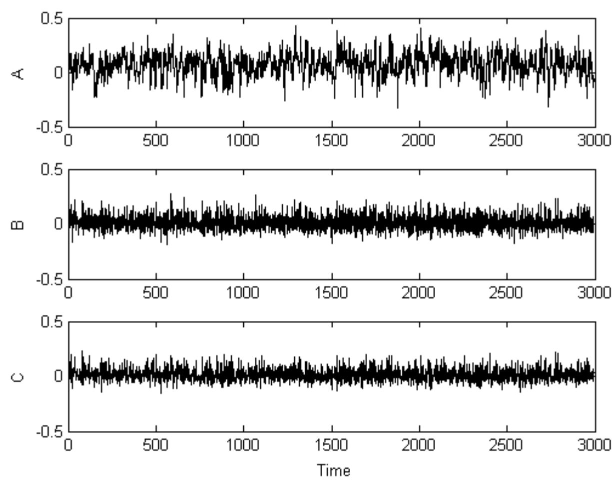

Fig. 2. The predicted signals for model (16): A. $\hat{y}_1(t)$; B. $\varepsilon_1(t)$; C. $\xi_1(t)$

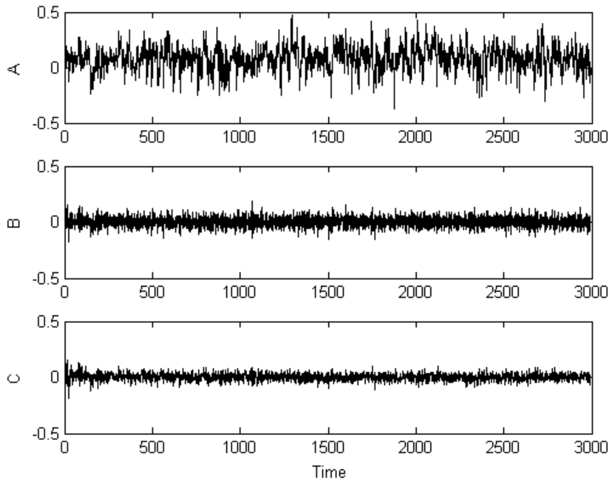

Fig. 3. The predicted signals for model (17): A. $\hat{y}_2(t)$; B. $\varepsilon_2(t)$; C. $\xi_2(t)$

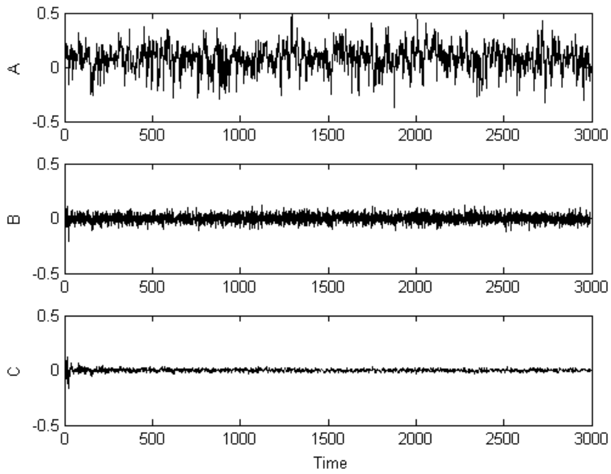

Fig.4. The predicted signals for model (18): A. $\hat{y}_3(t)$; B. $\varepsilon_3(t)$; C. $\xi_3(t)$

As shown in Figs. 2 and 3, both $\xi_1(t)$ and $\xi_2(t)$ did not converge to an acceptable level that (16) and (17) are inadequate. Fig. 4 shows that $\xi_3(t)$ was reduced to a low level that (18) is an adequate approximation of (15). In addition, the figures clearly suggest that it is difficult to diagnose the models by only comparing the measured outputs, predicted outputs, and residuals.

4.2 Validity Monitoring and Discussions

The new online correlation tests were used to monitor the validity of the three candidate models in parallel with the recursive identification. The parameters of the correlation tests were selected as $t_r = 50$, $\tau_r = 10$, and $N_r = 300$. Figs 5 to 7 show the validity monitoring results for the three models. Then, the results for each model are discussed below.

(16) is a linear model that the nonlinear terms of inputs are omitted. Fig 5 clearly shows that $\Theta_{u\varepsilon}(t)$ lies significantly outside the confidence interval at all the time instants that the residuals still correlated to the delayed inputs.

(17) is a nonlinear model with the same form as (15). Since there is an additive noise exists in the measured outputs which are used as important regressors in the identified models, the order of the model should be selected larger than the order of the underlying system to remove the effect of the additive noise in the measured outputs. Therefore, (17) cannot be used to adequately approximate (15). Fig. 6 shows that $\Theta_{\varepsilon\varepsilon}(t)$ lies significantly outside the confidence interval that the residuals were not reduced to a white noise sequence.

(18) is a proper nonlinear model for approximating (15). As shown in Fig. 4, $\xi_3(t)$ was reduced to an acceptable level. Fig. 7 shows that all the correlation functions converged to the confidence interval after a period of recursive identification that the third identified model (18) is valid.

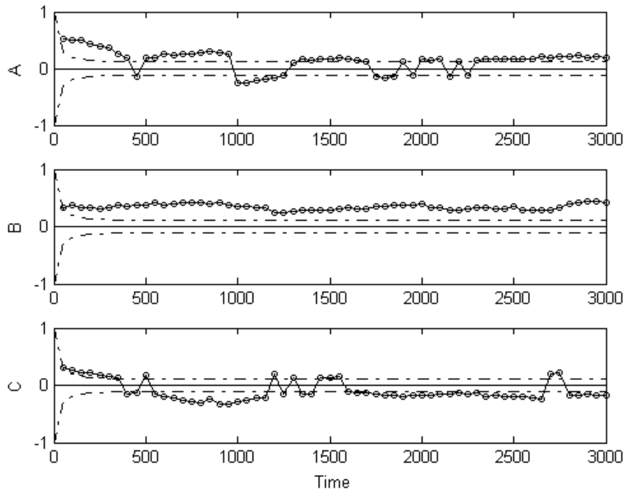

Fig. 5. Validity monitoring results for (16): A. $\Theta_{\varepsilon\varepsilon}(t)$; B. $\Theta_{u\varepsilon}(t)$; C. $\Theta_{y\varepsilon}(t)$.

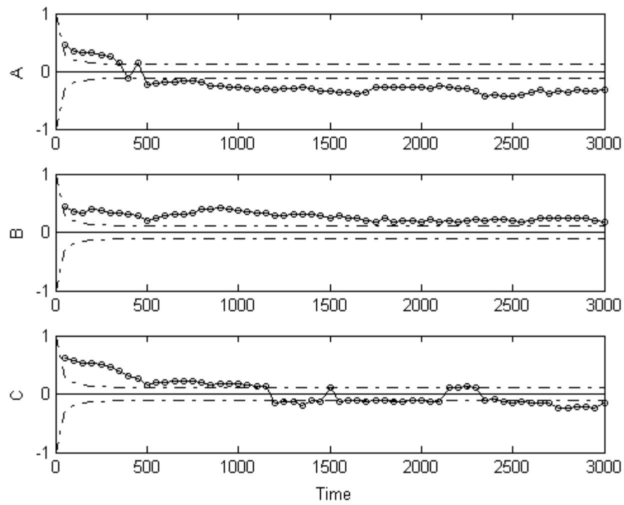

Fig. 6 Validity monitoring results for (17): A. $\Theta_{\varepsilon\varepsilon}(t)$; B. $\Theta_{u\varepsilon}(t)$; C. $\Theta_{y\varepsilon}(t)$.

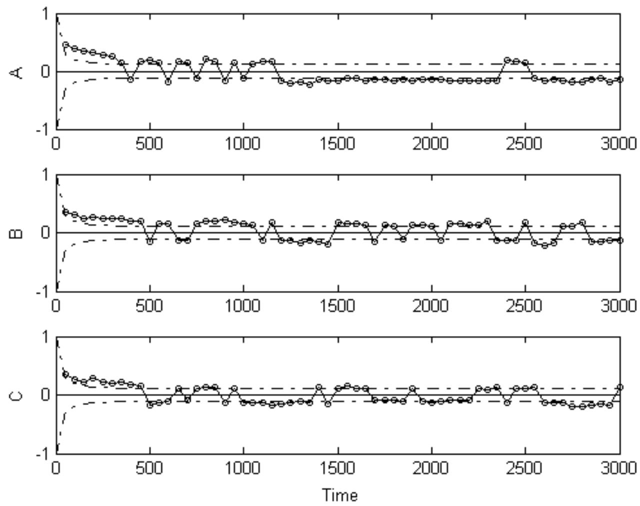

Fig. 7. Validity monitoring results for (17): A. $\Theta_{\varepsilon\varepsilon}(t)$; B. $\Theta_{u\varepsilon}(t)$; C. $\Theta_{y\varepsilon}(t)$.

5. CONCLUSIONS

In this study, a new online correlation test procedure is proposed to monitoring the validity of nonlinear models during the recursive identification procedure. Simulation study has been presented to demonstrate the new tests. The advantages of the new method can be summarized as follows.

1. The new online correlation tests can be used to clearly and precisely detect whether and when a recursive model is adequately identified since they are computed by normalized correlation functions.
2. Different requirements of the computational time and sensitivity can be easily satisfied by modifying the values of t_r, τ_r and N_r.
3. The correlation tests are computed by using the lastest collected data so that they can provide an immediate and sensitive indication of the performance change of the identified model.

It is believed that the new method should be applicable to much wider class of nonlinear recursive models including intelligence models and adaptive nose canceller, which are under investigation currently and will be reported in following publications.

REFERENCES

Billings, S.A. and Voon, W. S.F. (1983). Structure detection and model validity tests in the identification of nonlinear system, *Proceeding of the institution of electronic engineers*, Pt D, **130**, 193-199.

Billings, S.A. and Voon, W.S.F. (1986). Correlation based model validity tests for nonlinear models, *International Journal of Control*, **44**, 235-244.

Billings, S.A. and Zhu, Q.M.. (1994). Nonlinear model validation using correlation tests, *International Journal of Control*, **60**, 1107-1120.

Billings, S.A. and Zhu, Q.M. (1995). Model validation tests for multivariable nonlinear models including neural networks, *International Journal of Control*, **62**, 749-766.

Bohlin, T. (1971). On the problem of ambiguities in maximum likelihood identification. *Automatica*, **7**, 199-200.

Bohlin, T. (1978). Maximum power validation of models without higher order fitting. *Automatica*, **14**, 173-146.

Bowker, A.H. and Lieberman, G.J. (1972). *Engineering* Statistics, Prentice Hall.

Kung, F.C. and Shinh, D.H. (1986). Analysis and identification of Hammerstein model nonlinear systems using block-pulse function expansion, International Journal of Control, 43, 139-147.

Leontaritis, I.J. and Billings, S.A. (1985). Input-output parametric models for nonlinear systems, Part I and part Π, International Journal of Control, 41, 303-328, 329-344, 1985.

Ljung, L. (1999). System Identification Theory for the User 2nd Edition, Prentice Hall, Upper Saddle River, NJ.

Mao, K.Z. and Billings, S.A. (2000). Multi-directional model validity tests for non-linear system identification, International Journal of Control, 73, 132-143.

Söderströrm, T. and Stoica, P. (1990). On covariance function tests used in system identification, Automatica, 26, 125-133.

Wigren, T. (2003). User choices and model validation in system identification using nonlinear Wiener models, *Proceedings of 13th IFAC Symposium on System Identification*, Rotterdam, Netherlands, August 27-29, 2003, 863-868.

Zhang, L.F., Zhu, Q.M., and Longden, A. (2007). A set of novel correlation tests for nonlinear system variables, *International Journal of System Science*, **38**, 1, 47-60.

Zhu, Q.M., Zhang, L.F., and Longden, A. (2007). Development of omni-directional correlation functions for nonlinear model validation, *Automatica*, **43**, 9, 1519-1531.