Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Approximations of the NARMA model of Non-affine Plants

O. Adetona, S. Sathananthan, and L.H. Keel

Abstract— **When the NARMA model is used in adaptive control using neural networks, it often requires heavy computation due to its nonlinear dependence on the control input. To overcome this, two classes of approximate models to the NARMA model which are linear in the control input were introduced. They are known as NARMA-L1 and NARMA-L2. Though these approximations are useful in practical implementation of the controllers, their use is restricted to systems with small input magnitude. In this paper, we introduce two new classes of approximated models, referred as NARMA-L1B and NARMA-L2B, that relax the small input magnitude restriction. The proposed models are also linear in the control input and therefore suitable for control design. A simulation example is provided for illustration.**

I. INTRODUCTION

Concise recursive input-output maps for nonlinear systems have been presented (see [1], [2] and reference therein). Necessary and sufficient conditions for the existence of *global* recursive input-output maps have been reported [3]. It was shown that the future output of an $n^{\rm th}$ order plant is a nonlinear function of the last $2n$ inputs and last $2n$ outputs if the functions in the state equations are smooth, the system is state invertible, and an upper bound on the system order is known. Neural networks were subsequently proposed for identifying the input-output map from input-output data for the case where the structure of the state equations is unknown, and the states are inaccessible.

The NARMA model is an exact description of the input-output behavior of a finite dimensional nonlinear discrete-time plant in a neighborhood of an equilibrium point. It often leads to mathematically intractable nonlinear control equations and is therefore approximated by ARMA models for tractability. Though adequate for most applications, the ARMA model is only accurate for non-affine plants with small input magnitudes. To relax this restriction, NARMA-L1 and NARMA-L2 were recently introduced as approximations of NARMA models with inputs larger than permitted with ARMA models [4]. They are nonlinear with respect to past outputs but linear with respect to the current input and therefore suitable for control design. However, they are still restricted to small input *magnitudes*. In this paper, we further relax the input magnitude restriction and develop two new models (NARMA-L1B and NARMA-L2B). Like the NARMA-L1 and NARMA-L2, the proposed models are linear with respect to the control input and therefore suitable for control design. The proposed models eliminate the small input magnitude requirement and require small input *changes* rather than small input *magnitudes*.

This work was supported in part by NASA Grant NCC-5228

Authors are with Center of Excellence in Information Systems, Tennessee State University, Nashville, TN 37203-3401, USA

II. PRELIMINARIES

We consider an observable $n^{\rm th}$ order, non-affine, discrete-time, time-invariant system Σ with state equations

$$\begin{aligned} \Sigma: \quad \mathbf{x}(k+1) &= f\left[\mathbf{x}(k), u(k)\right] \\ y(k) &= h\left[\mathbf{x}(k)\right] \end{aligned} \tag{1}$$

where $\{\mathbf{x}(k)\}$, $\{u(k)\}$, $\{y(k)\}$ are discrete sequences at time instants k, subject to $\mathbf{x}(\cdot) \in \mathcal{X} \subseteq \Re^n$, $u(\cdot) \in \mathcal{U} \subseteq \Re$, $y(\cdot) \in \mathcal{Y} \subseteq \Re$, $f : \Re^n \times \Re \rightarrow \Re^n$, $h : \Re^n \rightarrow \Re$, and $f(\cdot,\cdot), h(\cdot) \in C^\infty$. As is typically done, we assume $\mathbf{0} = f\left[\mathbf{0}, 0\right]$, i.e. the origin is an equilibrium point of Σ. In eq. (1), $f(\cdot,\cdot)$ and $h(\cdot)$ are unknown functions, $f(\cdot,\cdot)$ is a nonlinear function of $\mathbf{x}(\cdot)$, $u(\cdot)$, and only $u(\cdot)$ and $y(\cdot)$ are accessible.

In a neighborhood $\mathbf{x}(\cdot) \in \Omega$ of an equilibrium state, the implicit function theorem can be used to derive a NARMA model [3], [4]

$$\begin{aligned} y(k+1) = F\,[&y(k), y(k-1), \cdots, y(k-n+1), \\ &u(k), u(k-1), \cdots, u(k-n+1)] \end{aligned} \tag{2}$$

for system Σ, where $F : \Re^n \times \Re^n \rightarrow \Re$ and $F(\cdot,\cdot) \in C^\infty$. In general, this input-output model is nonlinear in the control input and therefore not well suited to control system design.

Using eq. (2) for control design leads to mathematically intractable nonlinear difference equations. Unfortunately, techniques for solving such equations are limited to a few restricted families. A typical way of avoiding this problem is to approximate the NARMA model so that it becomes linear in the control input and then use the approximate model for control design. The resulting control design problem is linear for which a wide variety of general solution techniques are available.

We start our analysis by reviewing ARMA, NARMA, NARMA-L1 and NARMA-L2 models.

A. ARMA Models

The well known linear state space approximation of Σ is

$$\begin{aligned} \Sigma_L: \quad \mathbf{x}(k+1) &= A\mathbf{x}(k) + \mathbf{b}u(k) \\ y(k) &= \mathbf{c}\mathbf{x}(k) \end{aligned} \tag{3}$$

with $A \in \Re^{n\times n}$, $\mathbf{b} \in \Re^n$, and $\mathbf{c} \in \Re^n$ defined as:

$$A := \left.\frac{\partial f[\mathbf{x}, u]}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{0}, u=0}, \qquad \mathbf{b} := \left.\frac{\partial f[\mathbf{x}, u]}{\partial u}\right|_{\mathbf{x}=\mathbf{0}, u=0},$$

$$\mathbf{c} := \left.\frac{\partial h[\mathbf{x}]}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{0}}.$$

From eq. (3),

$$\begin{aligned} y(k) &= \mathbf{c}\mathbf{x}(k) \\ y(k+1) &= \mathbf{c}A\mathbf{x}(k) + \mathbf{c}\mathbf{b}u(k) \\ &\vdots \\ y(k+n-1) &= \mathbf{c}A^{n-1}\mathbf{x}(k) + \sum_{i=0}^{n-2} \mathbf{c}A^{i-1}\mathbf{b}u(k+i). \end{aligned}$$

Since Σ is observable, then Σ_L is observable and the matrix

$$\begin{bmatrix} \mathbf{c} \\ \mathbf{c}A \\ \vdots \\ \mathbf{c}A^{n-1} \end{bmatrix}$$

is nonsingular. This implies the state $\mathbf{x}(k)$ is a linear function of $y(k)$, $y(k-1)$, $\cdots$, $y(k-n+1)$, $u(k)$, $u(k-1)$, $\cdots$, $u(k-n+1)$. From eq. (3), $\mathbf{x}(k+n)$ is a linear function of $\mathbf{x}(k)$, $u(k)$, $u(k+1)$, $\cdots$, $u(k+n-1)$. From the facts above, the well known ARMA model of Σ_L is :

$$\Sigma_L: \quad y(k+1) = \sum_{i=0}^{n-1} \alpha_i \cdot y(k-i) + \sum_{j=0}^{n-1} \beta_j \cdot u(k-i) \quad (4)$$

where α_i and β_j and $i, j = 0, 1, \cdots, n-1$, are functions of A, $\mathbf{b}$ and $\mathbf{c}$.

Note that the linear approximation Σ_L is only applicable in a sufficiently small neighborhood of the origin. This is a consequence of the fact that the linearization is performed around $\mathbf{x} = \mathbf{0}$ and $u = 0$. Accordingly, the accuracy of the linear model (and the performance of controllers designed using it) will degrade as the operating region gets larger.

B. NARMA Model

The NARMA model is used when f and h are unknown and only the input u and output y are accessible. It is an exact input-output representation of the plant and can be derived from eq. (1) as follows. If $\frac{\partial Y_n(k)}{\partial \mathbf{x}(k)}$ is nonsingular at $\mathbf{x} = \mathbf{0}$, $U_{n-1} = \mathbf{0}$, the implicit function theorem can be used to show that ([3], pg. 33; [4], pg. 476),

$$\mathbf{x}(k) = \overline{g}\left[Y_n(k), U_{n-1}(k)\right] \quad (5)$$

in a neighborhood of the equilibrium state $\mathbf{x} = 0$, $U_{n-1} = 0$, where $Y_n(k)$ denotes the output sequence $y(k), y(k+1), \cdots, y(k+n-1)$, while $U_{n-1}(k)$ denotes the input sequence $u(k), u(k+1), \cdots, u(k+n-2)$, and $\overline{g}: \Re^n \times \Re^{n-1} \to \Re^n$ is a smooth function.

From ([4], pg. 476), eq. (5)

$$\begin{aligned} \mathbf{x}(k+n) &= g\left[Y_n(k), U_n(k)\right] \\ &= g\left[y(k), y(k+1), \cdots, y(k+n-1), \right. \\ &\qquad \left. u(k), u(k+1), \cdots, u(k+n-1)\right] \end{aligned}$$

where $g: \Re^{2n} \to \Re^n$ is a smooth function and $U_n(k)$ denotes the input sequence $[u(k), u(k+1), \cdots, u(k+n-1)]$.

From the expression above and noting that $y(k+n) = h[\mathbf{x}(k+n)]$, a NARMA model

$$\begin{aligned} y(k+1) = F\left[y(k), y(k-1), \cdots, y(k-n+1), \right. \\ \left. u(k), u(k-1), \cdots, u(k-n+1)\right] \end{aligned}$$

can be derived for a neighborhood $\mathbf{x}(k) \in \Omega$ of the equilibrium state. As shown below, the concept of relative degree can also be used to derive an alternative representation of the NARMA model. Here we define the term *relative degree* as in ([4], pg. 477):

Definition 1 If

$$\frac{\partial\left(h \circ f_e^k \circ f(\mathbf{x}, u)\right)}{\partial u} \begin{cases} = 0 & \text{when } 0 \le k \le d-2 \\ \neq 0 & \text{when } k = d-1 \end{cases}$$

in a neighborhood Ω of the equilibrium state, where f_e denotes the dynamics $f(\cdot,\cdot)$ and f_e^k is the k-times iterated composition of f, the dynamical system is said to have a *relative degree* d. Qualitatively, it implies that an input at instant k with the initial condition $\mathbf{x}(k) \in \Omega$ (i.e in a neighborhood of the equilibrium state) affects the output only d units of time later. Consequently, the input-output representation of the system Σ with a relative degree d can be written as

$$\begin{aligned} y(k+d) = \overline{F}\left[y(k), y(k-1), \cdots, y(k-n+1), \right. \\ \left. u(k), u(k-1), \cdots, u(k-n+1)\right]. \quad (6) \end{aligned}$$

The control design problem for a general function $\overline{F}$ is mathematically intractable because of the nonlinear relationship between $y(k+d)$ and $u(k)$. To make the problem tractable, several well known control design techniques have been developed assuming $\overline{F}$ is restricted to one of a few small classes (e.g linear or affine functions). This approach still leaves the control design problem for a general class of $\overline{F}$ unsolved. An alternative approach is to make approximations in the solution of the control design equations of eq. (6) and/or approximate $\overline{F}$ with an input-output map that is linear with respect to $u(k)$. As with system Σ_L, approximating $\overline{F}$ with a linearized model greatly simplifies the control design problem. For situations where the constraint on the input imposed by Σ_L is too severe, the NARMA-L1 and NARMA-L2 models were proposed in [4] as alternative models that can handle larger input values.

C. NARMA-L1 Model

This model was derived from the Taylor series expansion of $\overline{F}$ in eq. (6) (the NARMA model) in a compact domain $(Y, U) \in \mathcal{K}_Y \times \mathcal{K}_U$ where $\mathcal{K}_Y \times \mathcal{K}_U$ is a compact subset of $\Re^n \times \Re^n$. The expansion is performed around $[y(k), y(k-1), \cdots, y(k-n+1), u(k) = 0, u(k-1) = 0, u(k-2) = 0, \cdots, u(k-n+1) = 0]$. Accordingly, if $\| U_n(k-n+1) \|$ is sufficiently small, then ([4], pg. 477)

$$y(k+d) = f_0\left[y(k), y(k-1), \cdots, y(k-n+1)\right]$$

$$+\sum_{i=0}^{n-1} g_i\left[y(k), y(k-1), \cdots, y(k-n+1)\right] u(k-i) \quad (7)$$

where

$$f_0\left[y(k), y(k-1), \cdots, y(k-n+1)\right]$$
$$:= \overline{F}\left[y(k), y(k-1), \cdots, y(k-n+1), 0, 0, \cdots, 0\right],$$
$$g_i := \left.\frac{\partial\, \overline{F}}{\partial\, u(k-i)}\right|_{(y(k),\cdots,y(k-n+1),u(k)=0,\cdots,u(k-n+1)=0)}.$$

If the identification error is to be less than ϵ, $|u(k)|$ must be chosen less than δ for all k where $\delta \leq \sqrt{\frac{2\epsilon}{M_1 n}}$, and $M_1(Y, U)$ is the maximum matrix norm of the Hessian matrix $\frac{\partial}{\partial U}\left(\frac{\partial}{\partial U}\right)^T$ when evaluated over the compact domain $\mathcal{K}_Y \times \mathcal{K}_U$ ([4], pg. 478).

D. NARMA-L2 Model

This model was obtained from a Taylor series expansion of $\overline{F}$ around $[(y(k), y(k-1), \cdots, y(k-n+1), u(k) = 0, u(k-1), u(k-2), \cdots, u(k-n+1)]$. Accordingly, if $|u(k)|$ is sufficiently small then ([4], pg. 477)

$$\begin{aligned} y(k+d) = \overline{f}_0 &\left[y(k), y(k-1), \cdots, y(k-n+1),\right. \\ &\left. u(k-1), u(k-2), \cdots, u(k-n+1)\right] \\ + \overline{g}_0 &\left[y(k), y(k-1), \cdots, y(k-n+1),\right. \\ &\left. u(k-1), u(k-2), \cdots, u(k-n+1)\right] u(k), \end{aligned}$$

where

$$\overline{f}_0[y(k), \cdots, y(k-n+1), u(k-1), \cdots, u(k-n+1)]$$
$$:= \overline{F}\left[(y(k), y(k-1), \cdots, y(k-n+1), 0,\ u(k-1), u(k-2), \cdots, u(k-n+1)\right],$$
$$\overline{g}_i := \left.\frac{\partial \overline{F}}{\partial u(k)}\right|_{(y(k),\cdots,y(k-n+1),u(k)=0,u(k-1),\cdots,u(k-n+1))}.$$

Though the NARMA-L2 is theoretically different from the NARMA-L1 model, the region of validity of the two models may be the same in practical applications ([4], pg. 478). This situation occurs because the restriction of $|u(k)|$ to sufficiently small values imposed for NARMA-L2 implies $|u(k-1)|$, $|u(k-2)|$, $\cdots$, $|u(k-n+1)|$ must also be sufficiently small for $k = n, n+1, n+2, \cdots$. In effect, restrictions imposed for the NARMA-L2 model and its corresponding operating region may be equivalent to those of the NARMA-L1 model in practical applications.

III. MAIN RESULTS

Motivated by the NARMA-L1 and NARMA-L2 models, we propose NARMA-L1B and NARMA-L2B models. These models are not designed for sufficiently small input *magnitudes* like the NARMA-L1 and NARMA-L2 models. Instead, they require sufficiently small *changes* in the input. The NARMA-L1B and NARMA-L2B provide control system designers with a new pair of models in addition to the NARMA-L1 and NARMA-L2 model.

A. NARMA-L1B

This approximation of the NARMA model is obtained from the Taylors series expansion of $\overline{F}(\cdot, \cdot)$ around $[y(k), y(k-1), \cdots, y(k-n+1), u(k) = u(k-1), u(k-1) = u(k-2), \cdots, u(k-n+1) = u(k-n)]$. The assumption of $|u(k)|, |u(k-1)|, \cdots, |u(k-n+1)|$ sufficiently small (for all k) made for the NARMA-L1 model is replaced by the assumption that $|u(k) - u(k-1)|, |u(k-1) - u(k-2)|, \cdots, |u(k-n+1) - u(k-n)|$ is sufficiently small for the NARMA-L1B model. As demonstrated shortly, the NARMA-L1B (eq. (8)) can be used for operating regions that may be significantly larger than that of both the NARMA-L1 and NARMA-L2 models.

Theorem 1 *If δ is sufficiently small, the first two partial derivatives of $\overline{F}\left(Y_n(k-n+1), U_n(k-n+1)\right)$ with respect to $u(k-i)$, $i = 0, 1, 2, \cdots, n-1$, exist and are continuous, then*

$$\begin{aligned} y(k+d) = &\overline{F}[y(k), y(k-1), \cdots, y(k-n+1), \\ &u(k-1), u(k-2), \cdots, u(k-n)] \\ + &\sum_{i=0}^{n-1} p_i\left[y(k), y(k-1), \cdots, y(k-n+1),\right. \\ &u(k-1), u(k-2), \cdots, u(k-n)][u(k-i) - u(k-i-1)] \end{aligned} \quad (8)$$

where

$$\delta := \max\left\{|u(k) - u(k-1)|, |u(k-1) - u(k-2)|, \cdots, |u(k-n+1) - u(k-n)|\right\},$$

and

$$p_i := \left.\frac{\partial \overline{F}}{\partial u(k-i)}\right|_{u(k)=u(k-1),\cdots,u(k-n+1)=u(k-n)}.$$

Proof: For clarity, the notation $\overline{F}\left(Y_n(k-n+1), U_n(k-n+1)\right)$ will denote a scalar function $\overline{F} : \Re^{2n} \rightarrow \Re$, where $Y_n(k-n+1) := [y(k), y(k-1), \cdots, y(k-n+1)]$, and $U_n(k-n+1) := [u(k), u(k-1), \cdots, u(k-n+1)]$. Accordingly, the Taylor series expansion of $\overline{F}$ (from eq. (6)), around $u(k) = u(k-1)$, $u(k-1) = u(k-2)$, $\cdots$, $u(k-n+1) = u(k-n)$, is

$$\begin{aligned} y(k+d) = &\overline{F}\left[Y_n(k-n+1), U_n(k-n)\right] \\ &+ \sum_{i=0}^{n-1} p_i\left[Y_n(k-n+1), U_n(k-n)\right] \Delta u(k-i) \\ &+ R_1\left[Y_n(k-n+1), U_n(k-n+1)\right] \end{aligned} \quad (9)$$

where Δ is the difference operator,

$$U_n(k-n) := [u(k-1), u(k-2), \cdots, u(k-n)]$$

and the remainder term is

$$\begin{aligned}&R_1\left[Y_n(k-n+1),U_n(k-n+1)\right]\\ :=&\overline{F}\left[Y_n(k-n+1),U_n(k-n+1)\right]\\ &-\overline{F}\left[Y_n(k-n+1),U_n(k-n)\right]\\ &-\sum_{i=0}^{n-1}p_i\left[Y_n(k-n+1),U_n(k-n)\right]\Delta u(k-i).\end{aligned}$$

Using arguments similar to those in the Appendix of [4],

$$|R_1\left[Y_n(k-n+1),U_n(k-n+1)\right]|\leq N_1\left(K_y,K_u\right)\frac{n\delta^2}{2}$$

where

$$\begin{aligned}\delta:=\max\{&|u(k)-u(k-1)|\},|u(k-1)-u(k-2)|,\cdots\\ &\cdots,|u(k-n+1)-u(k-n)|\}\end{aligned}$$

and $N_1\left(K_y,K_u\right)$ is the maximum matrix norm of the Hessian matrix $\frac{\partial}{\partial U}\left(\frac{\partial\overline{F}[Y_n(k-n+1),U_n(k-n)]}{\partial U}\right)^T$ evaluated over the compact domain $\mathcal{K}_Y\times\mathcal{K}_U\subset\Re^n\times\Re^n$.

If δ is sufficiently small, the term $R_1(\cdot,\cdot)$ can be dropped from eq. (9) to get

$$\begin{aligned}y(k+d)=&\overline{F}\left[y(k),y(k-1),\cdots,y(k-n+1),\right.\\ &\left.u(k-1),u(k-2),\cdots,u(k-n)\right] \qquad (10)\\ +&\sum_{i=0}^{n-1}p_i\left[y(k),y(k-1),\cdots,y(k-n+1),u(k-1),\right.\\ &\left.u(k-2),\cdots,u(k-n)\right]\left[u(k-i)-u(k-i-1)\right].\end{aligned}$$

The term δ will be sufficiently small if $\delta\leq\sqrt{\frac{2\epsilon}{n\cdot N_1}}$, where ϵ is an upper bound on the approximation error $R_1(\cdot,\cdot)$. Note that a control system designer is free to select any value for ϵ and therefore compute δ in advance. The NARMA-L1B model can then be used to design (or constrain) controllers to generate $u(k)$ subject to $|u(k)-u(k-1)|\leq\delta$, for all $k\geq 0$. ■

B. NARMA-L2B

This approximation of the NARMA model is obtained from the Taylors series expansion of $\overline{F}(\cdot,\cdot)$ around $u(k)=u(k-1)$. The assumption of $|u(k)|$ sufficiently small (for all k) made for the NARMA-L2 model is replaced by the assumption that $|u(k)-u(k-1)|$ is sufficiently small for the NARMA-L2B model. As demonstrated shortly, the NARMA-L2B (eq. (11)) can be used for operating regions that may be significantly larger than that of both the NARMA-L1 and NARMA-L2 models.

Theorem 2 *If $|u(k)-u(k-1)|$ is sufficiently small, and the first two partial derivatives of $\overline{F}\left(Y_n(k-n+1),U_n(k-n+1)\right)$ with respect to $u(k)$ exist and are continuous, then*

$$\begin{aligned}y(k+d)=&\overline{F}\left[y(k),y(k-1),\cdots,y(k-n+1),\right.\\ &\left.u(k-1),u(k-1),u(k-2),\cdots,u(k-n+1)\right]\\ &+\overline{p}_0\left[y(k),y(k-1),\cdots,y(k-n+1),\right.\\ &\left.u(k-1),u(k-1),u(k-2),\cdots,u(k-n+1)\right]\\ &\cdot\left[u(k)-u(k-1)\right] \qquad (11)\end{aligned}$$

where

$$\overline{p}_0:=\left.\frac{\partial\overline{F}}{\partial u(k)}\right|_{u(k)=u(k-1)}. \qquad (12)$$

Proof: For clarity, the notation

$$\overline{F}\left(Y_n(k-n+1),u(k-1),U_{n-1}(k-n+1)\right)$$

will denote a scalar function $\overline{F}:\Re^{2n}\rightarrow\Re$, where

$$\begin{aligned}Y_n(k-n+1)&:=[y(k),y(k-1),\cdots,y(k-n+1)],\\ U_{n-1}(k-n+1)&:=[u(k-1),u(k-2),\cdots,u(k-n+1)].\end{aligned}$$

Accordingly, the Taylor series expansion of $\overline{F}$ (from eq. (6)) around $u(k)=u(k-1)$ is

$$\begin{aligned}y(k+d)=&\overline{F}\left[Y_n(k-n+1),u(k-1),U_{n-1}(k-n+1)\right]\\ &+\overline{p}_0\left[Y_n(k-n+1),U_n(k-n+1)\right]\left[u(k)-u(k-1)\right]\\ &+\overline{R}_1\left[Y_n(k-n+1),U_n(k-n+1)\right] \qquad (13)\end{aligned}$$

where the remainder term is

$$\begin{aligned}&\overline{R}_1\left[Y_n(k-n+1),U_n(k-n+1)\right]\\ :=&\overline{F}\left[Y_n(k-n+1),U_n(k-n+1)\right]\\ &-\overline{F}\left[Y_n(k-n+1),u(k-1),U_{n-1}(k-n+1)\right]\\ &-\overline{p}_0\left[Y_n(k-n+1),U_n(k-n+1)\right]\left[u(k)-u(k-1)\right].\end{aligned}$$

Using arguments similar to those in the Appendix of [4],

$$\begin{aligned}&\left|\overline{R}_1\left[Y_n(k-n+1),U_n(k-n+1)\right]\right|\\ &\qquad\leq\overline{N}_1\left(\overline{K}_y,\overline{K}_u\right)\frac{|\Delta u(k)|^2}{2} \qquad (14)\end{aligned}$$

where the term $\overline{N}_1\left(\overline{K}_y,\overline{K}_u\right)$ is the maximum value of $\frac{\partial^2\overline{F}(\cdot,\cdot)}{\partial\, u(k)^2}$ in the compact domain $\overline{K}_y\times\overline{K}_u\subset\Re^n\times\Re^n$.

For sufficiently small $|\Delta u(k)|$, the term $\overline{R}_1\left[Y_n(k-n+1),U_n(k-n+1)\right]$ can be dropped from eq. (13) to get

$$\begin{aligned}y(k+d)=&\overline{F}[y(k),y(k-1),\cdots,y(k-n+1), \qquad (15)\\ &u(k-1),u(k-1),u(k-2),\cdots,u(k-n+1)]\\ &+\overline{p}_0\left[Y_n(k-n+1),U_n(k-n+1)\right]\left[u(k)-u(k-1)\right].\end{aligned}$$

From eq. (14), the term $|\Delta u(k)|$ will be sufficiently small if $|\Delta(k)|\leq\sqrt{\frac{2\epsilon}{N_1}}$ for all $k\geq 0$, where ϵ is a desired upper bound on the estimation error $\overline{R}_1(\cdot,\cdot)$. ■

IV. AN EXAMPLE

The following simulation example is used to compare the approximation capabilities of the ARMA, NARMA-L1, NARMA-L2, NARMA-L1B and NARMA-L2B. All five approximation models will approximate the bounded-input-bounded-output non-affine discrete-time plant

$$\begin{aligned} y(k+1) &= \overline{F}[y(k), y(k-1), u(k), u(k-1)] \\ &= 0.2\cos[0.8y(k)+0.8y(k-1)] + 0.9 + 0.1y(k) \\ &+0.4\sin[0.8y(k)+0.8y(k-1)+2u(k)+u(k-1)] \\ &+0.1y(k-1) + 2\left[\frac{u(k)+[u(k)]^3+u(k-1)}{2+\cos[y(k)]}\right], \end{aligned} \quad (16)$$

with the following initial conditions:

$$y(-1) = y(0) = 1.38, u(-1) = u(0) = 0$$

and sampling time $t(k+1) - t(k) = 0.01$(Secs) for $k = 1, 2, \cdots, 400$. The approximate models (computed below) will be used as *series-parallel* ([3], pg. 35) models for the plant when the plant is subjected to input

$$u(k) = \theta\left[\sin\left(\frac{\pi k}{50}\right) + \sin\left(\frac{\pi k}{100}\right) + \sin\left(\frac{\pi k}{150}\right)\right], \quad \text{for} \ \ \theta \in \{0.01 \ \ 0.25 \ \ 0.5\}. \quad (17)$$

The ARMA model of the plant is:

$$\hat{y}(k+1) = \sum_{i=0}^{1}\alpha_i y(k-i) + \sum_{j=0}^{1}\beta_j u(k-j),$$

where

$$\alpha_0 = \left[\frac{\partial\overline{F}[\cdot,\cdot,\cdot,\cdot]}{\partial y(k)}\right]_{y(k)=y(k-1)=u(k)=u(k-1)=0} = 0.42,$$

$$\alpha_1 = \left[\frac{\partial\overline{F}[\cdot,\cdot,\cdot,\cdot]}{\partial y(k-1)}\right]_{y(k)=y(k-1)=u(k)=u(k-1)=0} = 0.42,$$

$$\beta_0 = \left[\frac{\partial\overline{F}[\cdot,\cdot,\cdot,\cdot]}{\partial u(k)}\right]_{y(k)=y(k-1)=u(k)=u(k-1)=0} = 1.8,$$

$$\beta_1 = \left[\frac{\partial\overline{F}[\cdot,\cdot,\cdot,\cdot]}{\partial u(k-1)}\right]_{y(k)=y(k-1)=u(k)=u(k-1)=0} = 1.4.$$

The NARMA-L1 model of the plant in the simulation example is

$$\hat{y}(k+1) = f_0[y(k), y(k-1)] + \sum_{i=0}^{1} g_i[y(k), y(k-1)]u(k-i),$$

where

$$\begin{aligned} f_0[\cdot,\cdot] &= 0.2\cos[0.8y(k)+0.8y(k-1)] + 0.9 + 0.1y(k) \\ &\quad +0.4\sin[0.8y(k)+0.8y(k-1)] + 0.1y(k-1), \\ g_0[\cdot,\cdot] &= 0.8\cos[0.8y(k)+0.8y(k-1)] + \frac{2}{2+\cos y(k)}, \\ g_1[\cdot,\cdot] &= 0.4\cos[0.8y(k)+0.8y(k-1)] + \frac{2}{2+\cos y(k)}. \end{aligned}$$

Similarly the NARMA-L2 Model is

$$\begin{aligned} \hat{y}(k+1) &= \overline{f}_0[y(k), y(k-1), u(k-1)] \\ &\quad +\overline{g}_0[y(k), y(k-1), u(k-1)]u(k), \end{aligned}$$

where

$$\begin{aligned} \overline{f}_0[\cdot,\cdot,\cdot] &= 0.2\cos[0.8y(k)+0.8y(k-1)] + 0.9 + 0.1y(k) \\ &\quad +0.4\sin[0.8y(k)+0.8y(k-1)+u(k-1)] \\ &\quad +0.1y(k-1) + \frac{2u(k-1)}{2+\cos y(k)}, \\ \overline{g}_0[\cdot,\cdot,\cdot] &= 0.8\cos[0.8y(k)+0.8y(k-1)+u(k-1)] \\ &\quad +\frac{2}{2+\cos y(k)}. \end{aligned}$$

The NARMA-L1B Model is

$$\begin{aligned} \hat{y}(k+1) &= \overline{F}[y(k), y(k-1), u(k-1), u(k-2)] \\ &\quad + \sum_{i=0}^{1} p_i[y(k), y(k-1), u(k-1), u(k-2)] \\ &\quad \cdot[u(k-i) - u(k-i-1)], \end{aligned}$$

where

$$\begin{aligned} &\overline{F}[\cdot,\cdot,\cdot,\cdot] = 0.2\cos[0.8y(k)+0.8y(k-1)] + 0.9 + 0.1y(k) \\ &+0.4\sin[0.8y(k)+0.8y(k-1)+2u(k-1)+u(k-2)] \\ &+0.1y(k-1) + 2\left[\frac{u(k-1)+[u(k-1)]^3+u(k-2)}{2+\cos y(k)}\right], \end{aligned}$$

$$\begin{aligned} p_0(\cdot,\cdot,\cdot,\cdot) &= 0.8\cos[0.8y(k)+0.8y(k-1)+2u(k-1) \\ &\quad +u(k-2)] + 2\left[\frac{1+3[u(k-1)]^2}{2+\cos y(k)}\right], \\ p_1(\cdot,\cdot,\cdot,\cdot) &= 0.4\cos[0.8y(k)+0.8y(k-1)+2u(k-1) \\ &\quad +u(k-2)] + \frac{2}{2+\cos y(k)}. \end{aligned}$$

The NARMA-L2B Model is

$$\begin{aligned} \hat{y}(k+1) &= \overline{F}[y(k), y(k-1), u(k-1), u(k-1)] \\ &\quad +p_0[y(k), y(k-1), u(k-1)][u(k) - u(k-1)] \end{aligned}$$

where

$$\begin{aligned} &\overline{F}[\cdot,\cdot,\cdot,\cdot] = 0.2\cos[0.8y(k)+0.8y(k-1)] + 0.9 + 0.1y(k) \\ &+0.4\sin[0.8y(k)+0.8y(k-1)+2u(k-1)+u(k-1)] \\ &+0.1y(k-1) + 2\left[\frac{u(k-1)+[u(k-1)]^3+u(k-1)}{2+\cos y(k)}\right], \\ &p_0(\cdot,\cdot,\cdot) = 0.8\cos[0.8y(k)+0.8y(k-1)+2u(k-1) \\ &+u(k-1)] + 2\left[\frac{1+3[u(k-1)]^2}{2+\cos y(k)}\right]. \end{aligned}$$

This simulation study was performed for time instants $k = 1, 2, \cdots, 400$ in a highly nonlinear region ($|u| \le 1.19$, $|y| \le 7.12$) of the plants input/output space. A plot of the input/output signals for the first four seconds (i.e. for $k = 1, 2, \cdots, 400$) of the plant (NARMA model) is shown in Figure 1. The highly nonlinear nature of the plant is evident from its response to the periodic input. In particular,

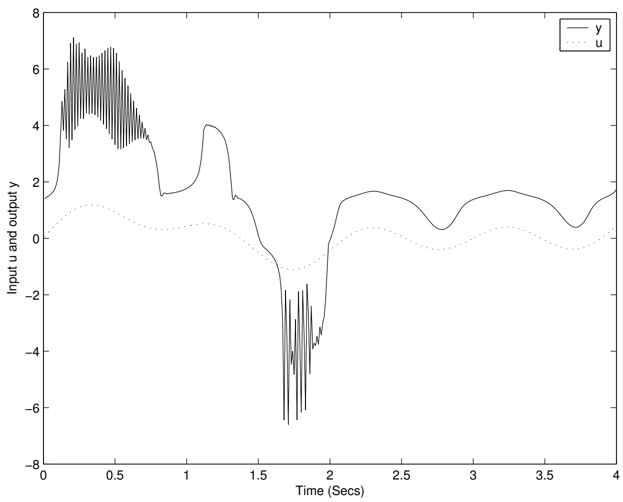

Fig. 1. Input (for $\theta = 0.5$) and Output Signals for NARMA Model.

TABLE I

INPUT AND OUTPUT DATA FROM THE NARMA MODEL

Model	θ	max $\lvert\Delta u(k)\rvert$	max $\lvert u(k)\rvert$	max $\lvert y(k)\rvert$
NARMA	0.01	0.001	0.02	1.40
NARMA	0.25	0.03	0.59	4.35
NARMA	0.50	0.06	1.19	7.12

the plant displays transient high frequency behavior in its response from 0.1 to 0.8 seconds and from 1.5 to 2 seconds.

Table I provides some general information on the input and output signal of the plant (NARMA model). From Table 1, the maximum change in the control signal $\max|u(k) - u(k-1)|$ and output magnitude $\max|y|$ increased with the input magnitude $\max|u|$. The table also shows that $\max|\Delta u(k)| := \max|u(k) - u(k-1)|$ was always much less than $\max|u(k)|$.

Table II is a comparison of the estimation errors (with respect to the NARMA model) of each approximate model evaluated over $k = 1, 2, \cdots, 400$. The term "Error" in Table II is defined as

$$\text{Error} := \frac{\max|y(k) - \hat{y}(k)|}{\max|y|} \times 100\%.$$

TABLE II

ESTIMATION ERRORS OF APPROXIMATE MODELS

MODEL	max $\lvert u(k)\rvert$	Error %	RMS Error
ARMA	0.02	2.07 $\times 10^{1}$%	2.18×10^{-1}
NARMA-L1	0.02	5.53 $\times 10^{-2}$ %	2.82×10^{-4}
NARMA-L2	0.02	2.42 $\times 10^{-2}$ %	1.25×10^{-4}
NARMA-L1B	0.02	1.36 $\times 10^{-4}$ %	5.29×10^{-7}
NARMA-L2B	0.02	6.09 $\times 10^{-5}$ %	2.37×10^{-7}
ARMA	0.59	39.952 %	0.7500
NARMA-L1	0.59	18.936%	0.1081
NARMA-L2	0.59	12.328%	0.0636
NARMA-L1B	0.59	0.024%	0.0003
NARMA-L2B	0.59	0.018%	0.0002
ARMA	1.19	79.07 %	1.574
NARMA-L1	1.19	53.74 %	0.699
NARMA-L2	1.19	40.57 %	0.726
NARMA-L1B	1.19	0.15 %	1.996×10^{-3}
NARMA-L2B	1.19	0.12 %	1.708×10^{-3}

The following observations are apparent from Table II. At low input levels (i.e $|u(k)| \leq 0.02$) all estimation models are accurate enough for some control applications. The errors (in percentage and RMS terms) of NARMA-L1B and NARMA-L2B were orders of magnitude less than those of NARMA-L1, NARMA-L2, and ARMA model for $\max|u| \leq 1.19$. As expected, the errors of all approximate models increased as the control signal magnitude increased. However, the errors of ARMA, NARMA-L1 and NARMA-L2 model increased at a much faster rate than those of NARMA-L1B and NARMA-L2B models. This occurred because we deliberately chose a control signal that kept $\max|\Delta u(k)|$ small at the values of $|u(k)|$ investigated. For such a case, the NARMA-L1B and NARMA-L2B worked better than others.

V. CONCLUDING REMARKS

Motivated by the NARMA-L1 and NARMA-L2 models of [4], we proposed NARMA-L1B and NARMA-L2B as alternatives suitable for control of non-affine nonlinear plants with larger input magnitudes. For input magnitudes that exceed those permitted by the ARMA, NARMA-L1 and NARMA-L2 models, the estimation error $e(k) = |y(k) - \hat{y}(k)|$ of the proposed models can be made arbitrarily small if $\max|\Delta u(k)|$ is sufficiently small. However, the "best" approximation model (linear in the control input) is application dependent because existing approximation models may provide better results if $\max|\Delta u(k)|$ takes on large values. Therefore, the proposed models complement existing models and do not replace them.

The NARMA-L1B and NARMA-L2B were developed assuming the NARMA model and input-output data are available. These models can be easily extended to the case where the NARMA model is not available but input-output data is available. For this case, the nonlinear functions of the NARMA-L1B and NARMA-L2B can be identified by neural networks using procedures similar to those described in [4].

REFERENCES

[1] I.J. Leontaris, and S.A. Billings, "Input-output parametric models for nonlinear systems part I: deterministic non-linear systems", *International Journal of Controls*, Vol. 41, No. 2, pp. 303-328, 1985.

[2] H. Wang, A.P Wang, M. Brown, and C.J. Harris, "One-to-one mapping and its application to neural networks based control system design", *International Journal of Systems Science*, Vol. 27, No. 2, pp. 161-170, 1996.

[3] A.U. Levin, and K.S. Narendra, "Control of nonlinear dynamical systems using neural networks - Part II: observability, identification, and control," *IEEE Transactions on Neural Networks*, Vol. 7, No. 1, pp. 30-42, 1996.

[4] K.S. Narendra and S.M. Mukhopadhyay, "Adaptive Control Using Neural Networks and Approximate Models", *Proceedings IEEE Transactions on Neural Networks*, Vol. 8, No. 3, pp. 475-485, May, 1997.