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

A novel hybrid neural network for modeling rare-earth extraction process

Wenjun Jia * **Tianyou Chai** *,** **Wen Yu** ***

* *Key Laboratory of Process Industry Automation (Northeastern University), Ministry of Education, Shenyang, 110006, China (e-mail: dr.wjjia@gmail.com).*
** *Research Center of Automation, Northeastern University, Shenyang, 110006, China.*
*** *Departamento de Control Automatico, CINVESTAV-IPN, A.P. 14-740, Av.IPN 2508, Mexico D.F., 07360, Mexico*

Abstract: Although neural networks are universal approximators, they are black-box models and it is difficult to obtain a suitable neural structure for a special unknown nonlinear system. Since many physical systems can be modeled by mechanistic models, we use these information to modify the normal neural identifiers and propose a novel neural modeling approach in this paper. This model includes a linear model which is linearized from the mechanistic model and a multilayer neural network. By Lyapunov stability approach, we prove that this hybrid neuro is stable in present of parameter and structure uncertainties. Then we apply these theory results on rare-earth extraction process. The results of application show that this new method can be used as soft-sensor of complex nonlinear systems.

1. INTRODUCTION

Element component content (ECC) is an important quality index for the rare-earth extraction process. There exit strong coupling, nonlinear, large time delay between ECC and solvent flow-rate, material liquid flow-rate and hydrochloric acid flow-rate, see Chai et al. [2004]. ECC also varies with the disturbances of solvent saponification degree and feed-in compositions etc. And it is hard to be measured online. Mjalli et al. [2005] proposed that this relation can not be described by precise model.

In order to describe the extraction process, many modeling methods are proposed. The population balance equation model is developed to describe the hydrodynamics and mass transfer of extraction process, see Tsouris et al. [1994] and Weinstein et al. [1998]. And the pulsed-flow model has also been used to predict the operating conditions and performance of the extractive separation of the rare earth metals by Wichterlova et al. [1999]. Recently a rigorous model for dynamic simulation of extraction process employs an improved detailed stage-wise mixing stage with back mixing and it takes into account the variation in hydrodynamics, mass transfer, and physical properties, see Mjalli et al. [2005].

Generally the exact mathematical model for rare-earth extraction process is too complex to be handled analytically. A common method is to use linear model, but the modeling error is relative big. Results show that neural network technique seems to be very effective to model a broad category of complex nonlinear systems when we do not have complete model information. A recent modeling of rare-earth extraction process by neural network is introduced by Giles et al. [1996]. But the lack of a process-based internal structure is a liability for the neural network when faced with sparse, noisy data. Insufficient data hamper the accuracy of a neural network because the network relies completely on the data when inducing process behavior. While it is a feasible approach by combining prior knowledge with neural networks to estimate the ECC online. It has more priorities comparing with prior knowledge based model or neural networks used along the papers of Thompson et al. [1994] and Oliveria et al. [2004]. The inclusion of prior knowledge is to improve the neural network estimations when trained on sparse and noisy process data. Prior knowledge enters the hybrid model as a linear model. The linear model controls the extrapolation of the hybrid in the regions of input space that lack training data. The neural network compensates for inaccuracy in the prior model.

Neuro modeling approach uses the nice features of neural networks, but the lack of mathematical model for the plant makes it hard to obtain theoretical results on stable learning. It is very important to assure the stability of neuro modeling in theory before we use them in some real applications. Discrete-time neural networks are more convenient for real applications. Two types stability for discrete-time neural networks were studied. The stability of neural networks can be found in the papers of Feng et al. [1999] and Suykens et al. [1997]. The stability of learning algorithms was discussed by Jin et al. [1999] and Polycarpou et al. [1992]. Polycarpou et al. [1992] assumed neural networks could represent nonlinear systems exactly, and concluded that backpropagation-type algorithm guar-

⋆ This work was supported by the State Key Program of National Natural Science of China (No. 60534010), the National Fundamental Research Program of China (No. 2002CB312201), the Funds for Creative Research Groups of China (No. 60521003), the 111 project(B08015), and the National High-tech Program(2006AA040307).

anteed exact convergence. It is well known that normal identification algorithms are stable for ideal plants, see Ioannou et al. [1996]. In the presence of disturbances or unmodeled dynamics, these adaptive procedures can go to instability easily. Several robust modification techniques were proposed by Ioannou et al. [1996]. The weight adjusting algorithms of neural networks is a type of parameters identification, the normal gradient algorithm is stable when neural network model can match the nonlinear plant exactly, see Polycarpou et al. [1992].

In this paper, a hybrid neuro is proposed by combining a linear model with a multilayer neural network. And Lyapunov stability approach is applied to obtain stable learning laws for the ECC estimator for rare-earth extraction process. An application study of a La, Ce, Pr, Nd tetracomponent for rare-earth extraction process is conducted to illustrate the hybrid neuro modeling approach.

2. MATHEMATICAL MODELS OF RARE-EARTH EXTRACTION PROCESS

In order to model the rare-earth extraction process by input/output data, we first study mathematical models of extraction process. In the process of rare earth extraction, a single mixer-settler cannot achieve an effective separation due to the small separation coefficients between different elements. As such, some mixer-settlers should be connected in cascade. A typical rare-earth extraction process is illustrated in Fig.1, which involves n extraction stages and m washing stages, where the material to be extracted in aqueous phase gets in touch with the organic phase continuously. As a result, the purified products in aqueous and organic phase are exported from the 1st stage and the $(n+m)$th stage respectively.

The blank organic phase flow rate V_S is added to the 1st stage and flows from the left to the right, whereas the material liquid flow rate V_F containing the elements to be separated is added to the nth stage and flows from the right to the left. The hydrochloric acid flow rate V_W is added to the $(n+m)$th stage, flowing against the organic phase and converging with the material liquid at the nth stage. In the extraction section, the distribution rates between different elements in aqueous and organic phase make many easy-extracted elements and partial hard-extracted elements enter into organic phase. In the scrubbing section, the hard-extracted elements can be washed down into aqueous phase much more than the easy-extracted elements under certain conditions. Although easy-extracted elements are partially washed down into aqueous phase in the scrubbing section, the aqueous phase from the scrubbing section enters into the extraction section along with the material liquid, and the washed down elements is extracted again. The extraction section can ensure the recovery ratio of easy-extracted elements and make more of them into organic phase, whilst the scrubbing section is utilized to ensure the quality of products and wash more hard-extracted elements down to purify the easy-extracted elements.

The rare-earth cascade extraction process is a typical multistage process, where the concentrations of the elements in aqueous and organic phases at every stage can be regarded as Stephanopoulos [1984].

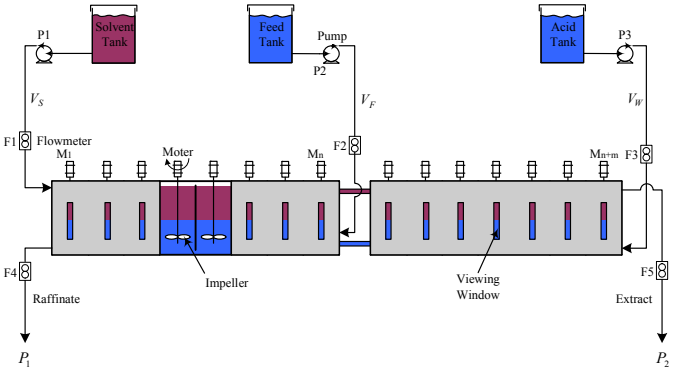

Fig. 1. Schematic diagram of rare-earth extraction process

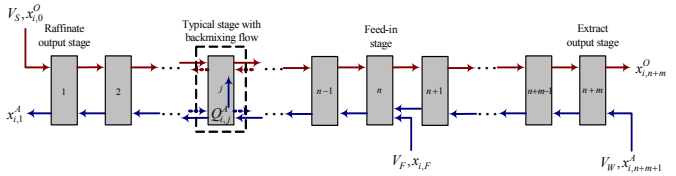

Fig. 2. Schematic diagram of the stagewise extraction with backmixing flow

$$\frac{dn_i}{dt} = \sum_{inlets} c_{i,k}F_k - \sum_{outlets} c_{i,l}F_l - r_i V \qquad (1)$$

where n_i is the molar accumulation of component i in mixer- settlers, $c_{i,k}$ and $c_{i,l}$ are the molar concentrations of the ith component in the kth inlet stream and the lth outlet stream respectively, F_k and F_l are the volumetric flow rates of the kth inlet stream and the lth outlet stream respectively, r_i is the mass-transfer velocity, V is the apparatus volume.

The above rare-earth extraction process is stage-wise with back mixing flow, it can be simply described as Fig.2. The entrainment between aqueous and organic phases should be considered, and elements transfer ratio is used to describe the mass transfer between the two phase. According to the mass balance equation (1), a dynamic model with back mixing flow for rare-earth extraction process is established. Firstly, the basic Assumptions is necessary.

(1) Flow non-idealities are handled by incorporating backflow streams opposite to the direction of the main flow streams. The values of these streams are expressed as fractions of the main flow streams using fractional symbol α_c and α_d.

(2) Mass transfer coefficient K is calculated for each stage as function of physical properties, operational parameters and stage design specifications.

(3) Equilibrium between phases at each stage is expressed as a distribution coefficient $D = \frac{y}{x}$. Its value is calculated for each stage from experimental data as a function of solute concentration in the refined phase.

(4) Hydrodynamics within stages is expressed as a fractional volume hold-up ε_i and calculated for each stage. The hold-up is as a function of phase flow ratio.

(5) The physical properties of the two phases are considered as variables throughout the mixer-settlers and are calculated for each stage as functions of concentration, mixer-settler geometry, and operational parameters.

(6) In order to approximate the damping and delaying action of the phase separation volumes (single phase) located between the interfaces and the conductor ends, a form of delay must be added to the theoretical model. This is attained by considering the volume between the interface and the sampling tube as comprising a perfectly mixed, single-phase stage without mass transfer.

In Fig.2, the holdup volume and of aqueous phase and organic phase at each stage are constants, while the flow rates of organic phase $V_{S,j}$, the material liquid $V_{F,j}$ and the hydrochloric acid $V_{W,j}$ are assumed to be invariant instantaneously. By step-wise extraction model with back mixing flow and using equation (1), the high-order nonlinear dynamic model of the ith component in the jth stage is obtained as Mjalli et al. [2005].

$$h_{A,j}\frac{dx^A_{i,j}}{dt} = V_{A,j+1}(1+\alpha_{A,j})x^A_{i,j+1} + V_{A,j-1}\alpha_{A,j}x^A_{i,j-1} - V_{A,j}(1+2\alpha_{A,j})x^A_{i,j} - Q^A_{i,j} \tag{2}$$

$$h_{O,j}\frac{dx^O_{i,j}}{dt} = V_{O,j-1}(1+\alpha_{O,j})x^O_{i,j-1} + V_{O,j+1}\alpha_{O,j}x^O_{i,j+1} - V_{O,j}(1+2\alpha_{O,j})x^O_{i,j} + Q^A_{i,j} \tag{3}$$

where $h_{A,j} = V(1-\varepsilon_j)$, $h_{O,j} = V\varepsilon_j$ are holdup volume of the aqueous and organic phase in the jth stage, $x^A_{i,j}$, $x^O_{i,j}$ are concentrations of the ith component at the jth stage, and the balance equation between $x^A_{i,j}$ and $x^O_{i,j}$ is $x^O_{i,j} = D_{i,j}\cdot x^A_{i,j}$, $Q^A_{i,j}$ represents the mass transfer between aqueous and organic phase, it can be described as Srivastava et al. [2000].

$$Q^A_{i,j} = V_c\frac{dx^A_{i,j}}{dt} = K_{oc,j}\alpha_j V(x^A_{i,j} - \overline{x}^A_{i,j}) \tag{4}$$

where V_c is the volume of continuous phase , V is the volume of mixer-settler, $K_{oc,j}$ is the total mass transfer coefficients, $\overline{x}^A_{i,j}$ is the aqueous balance concentration of the ith element in the jth stage.

At stage 1, n and $n+m$, the boundary equations are expressed as

$$h_{A,1}\frac{dx^A_{i,1}}{dt} = (V_{F,2}+V_{W,2})(1+\alpha_{A,1})x^A_{i,2} - (V_{F,1}+V_{W,1})\times(1+\alpha_{A,1})x^A_{i,1} - Q^A_{i,1} \tag{5}$$

$$h_{O,1}\frac{dx^O_{i,1}}{dt} = V_S x^O_{i,0} + V_{S,2}\alpha_{O,1}x^O_{i,2} - V_{S,1}(1+\alpha_{O,1})x^O_{i,1} + Q^A_{i,1} \tag{6}$$

$$h_{A,n}\frac{dx^A_{i,n}}{dt} = V_F x_{i,F} + V_{W,n+1}(1+\alpha_{A,n})x^A_{i,n+1} - (V_{F,n} + V_{W,n})(1+2\alpha_{A,n})x^A_{i,n} + (V_{F,n-1}+V_{W,n-1})\alpha_{A,n}x^A_{i,n-1} - Q^A_{i,n} \tag{7}$$

$$h_{O,n}\frac{dx^O_{i,n}}{dt} = V_{S,n-1}(1+\alpha_{O,n})x^O_{i,n-1} + V_{S,n+1}\alpha_{O,n}x^O_{i,n+1} - V_{S,n}(1+2\alpha_{O,n})x^O_{i,n} + Q^A_{i,n} \tag{8}$$

$$h_{A,n+m}\frac{dx^A_{i,n+m}}{dt} = V_W x^A_{i,n+m+1} + V_{W,n+m-1}\alpha_{A,n+m} \times x^A_{i,n+m-1} - V_{W,n+m}(1+\alpha_{A,n+m})x^A_{i,n+m} + Q^A_{i,n+m} \tag{9}$$

$$h_{O,n+m}\frac{dx^O_{i,n+m}}{dt} = V_{S,n+m-1}(1+\alpha_{O,n+m})x^O_{i,n+m-1} - V_{S,n+m}(1+\alpha_{O,n+m})x^O_{i,n+m} + Q^A_{i,n+m} \tag{10}$$

where $x^O_{i,0}$ and $x^A_{i,n+m+1}$ are boundary concentrations.

The extraction purpose is to improve the component contents of hard- and easy-extracted elements in refined and extract. The hard-extracted component content P_1 and easy-extracted component content P_2 is defined as

$$P_1 = 100\times\sum_{i=1}^{q}x^A_{i,1}/\sum_{i=1}^{p}x^A_{i,1} \tag{11}$$

$$P_2 = 100\times\sum_{i=q+1}^{p}x^O_{i,n+m}/\sum_{i=1}^{p}x^O_{i,n+m} \tag{12}$$

where $q(q<p)$ is the number of hard-extracted elements, and $p-q$ is the number of easy-extracted elements.

3. MODELING OF RARE-EARTH EXTRACTION PROCESS VIA HYBRID NEURO

The mathematical models discussed above work only under some special conditions, espacially some kinetics parameters can not be measured, and the calculation results are ideal. In real application, we have only input/output data. In this case a linear model and a neural network are used to modeling the whole extraction process. The linear model is for the linear part of the step-wise model (2) or (3) and the static neural network can be used to identify the nonlinear part (4). A new stable learning algorithm will be presented for static hybrid neuro modeling.

The properties of ECC can be written in the following form

$$y(k) = \Phi[X(k)] \tag{13}$$

where $X(k) = [u(k), u(k-1), u(k-2), \cdots]^T$, $y(k)$ is the element component content (ECC), $\Phi(\cdot)$ is an unknown nonlinear function representing the rare-earth extraction properties, $u(k)$ is measurable vector inputs or disturbances. We consider a linear model and a multilayer neural network to model the extraction properties.

The mechanistic model is linearized as an ARMAX model

$$\begin{aligned}\widehat{y}_l(k) = &-A_1\widehat{y}(k-1) - \cdots - A_{n_A}\widehat{y}(k-n_A)\\ &+B_0u(k-d) + \cdots + B_{n_B}u(k-d-n_B)\end{aligned} \tag{14}$$

where $u = [u_1(k), u_2(k), u_3(k), \cdots]^T$, d is the time delay. The above model (14) can be rewritten as a vector form

$$\widehat{y}_l(k) = \theta_k X_l(k) \tag{15}$$

where $\theta_k = [-A_1\cdots A_{n_A}, B_0\cdots B_{n_B}]$, $X_l(k) = \left[\widehat{y}_l^T(k-1)\right.$ $\left.\cdots\widehat{y}_l^T(k-n_A), u^T(k-d)\cdots u^T(k-d-n_B)\right]^T$

The linear model is liable for the process-based internal structure. Then we use a neural network to compensate the modeling error of the linear mechanistic model (15)

$$\widehat{y}_n(k) = W_k \phi [V_k X_n(k)] \tag{16}$$

where the scalar output $\widehat{y}_n(k)$ and vector input $X_n(k) \in R^{n\times 1}$ is defined in (13), the weights in output layer are $W_k \in R^{1\times m}$, the weights in hidden layer are $V_k \in R^{m\times n}$, ϕ is $m-$dimension vector function. The typical presentation of the element $\phi_i(.)$ is sigmoid function. According to the Stone-Weierstrass theorem Cybenko [1989], this general nonlinear smooth function (13) can be written as

$$y(k) - \widehat{y}_l = W^* \phi [V^* X_n(k)] - \mu(k) \tag{17}$$

where V^* and W^* are sets of unknown weights which may minimize the modeling error $\mu(k)$. Or

$$y(k) = \theta^* X_l(k) + W^* \phi [V^* X_n(k)] - \mu(k)$$

where θ^* is the ideal value of θ_k for the linear mechanistic model (15).

First we fix the hidden weight we V^0, let us define the identification error as

$$e(k) = (\widehat{y}_n(k) + \widehat{y}_l) - y(k) \tag{18}$$

So the error dynamic is

$$e(k) = \widetilde{\theta}_k X_l(k) + \widetilde{W}_k \phi \left[V^0 X_n(k)\right] + \mu(k) \tag{19}$$

where $\widetilde{W}_k = W_k - W^*$, $\widetilde{\theta}_k = \theta_k - \theta^*$.

In this paper we are only interested in open-loop identification, we can assume that the plant (13) is bounded-input and bounded-output stable, *i.e.*, $y(k)$ and $u(k)$ in (13) are bounded. Since $X_n(k) = [u(k), u(k-1), u(k-2), \cdots]^T$, $X_n(k)$ is bounded. By the boundedness of the active function ϕ, we assume that $\mu(k)$ in (19) is bounded.

3.1 Parameter Uncertainty

Let us first assume that exact linear mechanistic model and neural network model of the plant are available (without unmodeled dynamics $\mu(k) = 0$), i.e., there exist weights θ^* and W^* such that the nonlinear system (13) is completely described by these two models.

The following theorem gives a new stable learning algorithm for the hybrid neuro identification.

Theorem 1. If we use the linear mechanistic model (15) and the neural network (16) to model the nonlinear system (13), and there only exits parameter uncertainty, the following gradient updating law can make identification error be asymptotic stable.

$$\begin{aligned} \theta_{k+1} &= \theta_k - \eta_k e(k) X_l(k) \\ W_{k+1} &= W_k - \eta_k e(k) \phi \left[V^0 X_n(k)\right] \end{aligned} \tag{20}$$

where $0 < \eta_k = \dfrac{\eta_0}{1 + \|X_l\|^2 + \|\phi\|^2}$, $0 < \eta_0 \le 2$.

Proof. Lyapunov function is selected as

$$V_k = \left\|\widetilde{\theta}_k\right\|^2 + \left\|\widetilde{W}_k\right\|^2 \tag{21}$$

From the updating law, we can obtain

$$\begin{aligned} \widetilde{\theta}_{k+1} &= \widetilde{\theta}_k - \eta_k e(k) X_l(k), \\ \widetilde{W}_{k+1} &= \widetilde{W}_k - \eta_k e(k) \phi \left[V^0 X_n(k)\right] \end{aligned} \tag{22}$$

From (19) we know $e(k) = \widetilde{\theta}_k X_l(k) + \widetilde{W}_k \phi \left[V^0 X_n(k)\right]$,hence

$$\begin{aligned} V_{k+1} - V_k &= \left\|\widetilde{\theta}_{k+1}\right\|^2 + \left\|\widetilde{W}_{k+1}\right\|^2 - \left\|\widetilde{\theta}_k\right\|^2 - \left\|\widetilde{W}_k\right\|^2 \\ &= \eta_k^2 e(k)^2 \|X_l(k)\|^2 - 2\eta_k tr\left\{\widetilde{\theta}_k X_l(k) e(k)\right\} \\ &+ \eta_k^2 e(k)^2 \|\phi\|^2 - 2\eta_k tr\left\{\widetilde{W}_k \phi e(k)\right\} \\ &= \eta_k^2 \left(\|X_1(k)\|^2 + \|\phi\|^2\right) e(k)^2 \\ &- 2\eta_k tr\left\{\left(\widetilde{\theta}_k X_1(k) + \widetilde{W}_k \phi\right) e(k)\right\} \\ &= \eta_k^2 \left(\|X_1(k)\|^2 + \|\phi\|^2\right) e(k)^2 - 2\eta_k e(k)^2 \\ &= -\eta_k \left(2 - \eta_0 \frac{\|X_1\|^2 + \|\phi\|^2}{1 + \|X_1\|^2 + \|\phi\|^2}\right) e(k)^2 \end{aligned} \tag{23}$$

When $0 < \eta_0 \le 2$, $V_{k+1} - V_k \le 0$. From

$$V_{k+1} - V_1 = \sum_{i=1}^{k} -\eta_i \left(2 - \eta_0 \frac{\|X_l\|^2 + \|\phi\|^2}{1 + \|X_l\|^2 + \|\phi\|^2}\right) e(i)^2$$

and the boundedness of $\eta_k \left(2 - \eta_0 \dfrac{\|X_l\|^2 + \|\phi\|^2}{1 + \|X_l\|^2 + \|\phi\|^2}\right)$,

$\lim_{k\to\infty} V_k$ and V_1 are bounded, we can obtain

$$\lim_{k\to\infty} ||e(k)|| \to 0 \tag{24}$$

■

3.2 Unmodeled Dynamics Present

If the linear mechanistic model and neural network model can never follow the nonlinear system (13), the modelling error is defined as (19). The following dead-zone modification gives a stable backpropagation-like algorithm for training the neural network and the linear mode.

Theorem 2. If we use the linear mechanistic model (15) and the neural network (16) to model the nonlinear system (13), the following dead-zone gradient updating law can make identification error bounded

$$\begin{aligned} \theta_{k+1} &= \theta_k - \eta_k e(k) X_l(k) \\ W_{k+1} &= W_k - \eta_k e(k) \phi \left[V^0 X_n(k)\right] \end{aligned} \tag{25}$$

where

$$\eta_k = \frac{s_k \eta_0}{1 + \|X_l\|^2 + \|\phi\|^2},$$

$$s_k = \begin{cases} 1 & e(k)^2 \ge \dfrac{\overline{\mu}}{1 - \eta_0} \\ 0 & otherwise \end{cases}$$

$0 < \eta_0 \le 1$, $\overline{\mu}$ is an upper bound of $\mu(k)^2$. The average of the identification error satisfies

$$J = \limsup_{T\to\infty} \frac{1}{T} \sum_{k=1}^{T} e^2(k) \le \frac{\overline{\mu}}{1 - \eta_0} \tag{26}$$

Proof. Lyapunov function is selected as

$$V_k = \left\|\widetilde{\theta}_k\right\|^2 + \left\|\widetilde{W}_k\right\|^2 \tag{27}$$

Following the same process as *Theorem 1*, but $e(k) = \widetilde{\theta}_k X_l(k) + \widetilde{W}_k \phi \left[V^0 X_n(k)\right] + \mu(k)$

$$\begin{aligned} V_{k+1}-V_k &= \eta_k^2\left(\|X_l(k)\|^2+\|\phi\|^2\right)e(k)^2 \\ &-2\eta_k tr\left\{\left(\widetilde{\theta}_k X_l(k)+\widetilde{W}_k\phi\right)e(k)\right\} \\ &\le \eta_k^2\left(\|X_l(k)\|^2+\|\phi\|^2\right)e(k)^2-2\eta_k e(k)^2 \\ &+2\eta_k\|\mu(k)\|\|e(k)\| \\ &\le \eta_k^2\left(\|X_l(k)\|^2+\|\phi\|^2\right)e(k)^2-\eta_k e(k)^2 \\ &+\eta_k\mu(k)^2 \\ &= -\eta_k(1-\eta_0)e(k)^2+\eta_k\overline{\mu} \\ &= -\eta_k\left[(1-\eta_0)e(k)^2-\overline{\mu}\right] \end{aligned} \tag{28}$$

with $1\ge\eta_0>0$, when $(1-\eta_0)e(k)^2-\overline{\mu}\ge 0, V_{k+1}-V_k\le 0$, then θ_k and W_k are bounded. Therefore $e(k)$ is bounded. If $(1-\eta_0)e(k)^2-\overline{\mu}<0$, then, $\theta_{k+1}=\theta_k$, $W_{k+1}=W_k$,so θ_k and W_k are also bounded, and $e(k)^2<\frac{\overline{\mu}}{1-\eta_0}<\infty$ is also bounded.

(28) can be rewritten as

$$\Delta V_k\le -\eta_k(1-\eta_0)e^2(k)+\eta_k\overline{\mu} \tag{29}$$

Summarizing (29) from 1 up to T, and by using $V_T>0$ and V_1 is a constant, we obtain

$$\begin{aligned} V_T-V_1 &\le -\eta_k(1-\eta_0)\sum_{k=1}^{T}e^2(k)+T\eta_k\overline{\mu} \\ \eta_k(1-\eta_0)\sum_{k=1}^{T}e^2(k) &\le V_1-V_T+T\eta_k\overline{\mu} \\ &\le V_1+T\eta_k\overline{\mu} \end{aligned} \tag{30}$$

(26) is established. ■

Remark 1. (25) is the gradient descent algorithm, which the normalizing learning rate η_k is time-varying in order to assure the identification process is stable. This learning law is simpler to use, because we do not need to care about how to select a better learning rate to assure both fast convergence and stability. No any previous information is required.

Remark 2. If we also want to train the hidden weights V_k, with the similar approach as in Yu et al. [2003], the following backpropagation-like algorithm can make identification error $e(k)$ bounded

$$\begin{aligned} \theta_{k+1} &= \theta_k-\eta_k e(k)X_l(k) \\ W_{k+1} &= W_k-\eta_k e(k)\phi[V_kX_n(k)] \\ V_{k+1} &= V_k-\eta_k e(k)\phi' W_k^T X_n^T(k) \end{aligned} \tag{31}$$

where

$$\eta_k=\frac{s_k\eta_0}{1+\|X_l\|^2+\|\phi\|^2+\left\|\phi' VW_kX_n^T\right\|^2},$$
$$s_k=\begin{cases}1 & e(k)^2\ge\dfrac{\overline{\mu}}{1-\eta_0}\\ 0 & otherwise\end{cases}$$

where $0<\eta_0\le 1$.

4. APPLICATION STUDY

The hybrid neuro is used to estimate element component content (ECC) of hard–extracted component B(La, Ce) or easy-extracted component A(Pr, Nd) for a La, Ce, Pr, Nd tetra-component extraction production line in Jiangxi Province, South China. The extraction process under consideration is illustrated as Fig.1. The organic, material liquid and hydrochloric acid flow rates are added to the extraction process at the 1st, 20th and 48th stage. At the same time the hard–extracted component and easy-extracted component are exported from the 1st and 48th stage respectively. We mainly use the proposed hybrid neuro to estimate ECC of component B, and the similar result can also be drawn on component A by the similar modeling procedure.

The data vector of the linear model $X_l(k)=[\widehat{y}_l(k-d), u_1(k-2),u_1(k-d-1),u_2(k-n-1),u_2(k-n-d), u_3(k-n-m-1),u_3(k-n-m-d),u_4(k-n-1), u_5(k-n-d)]^T$, and data vector of the multilayer neural network $X_n(k)=[u_1(k),u_2(k),u_3(k),u_4(k),u_5(k)]^T$. The $\widehat{y}_l$ is ECC of component B or component A via linear model, $u_i(i=1,\cdots,5)$ are measurable inputs, which are solvent flow rate V_S, material liquid flow rate V_F, hydrochloric acid flow rate V_W, and measurable disturbances, which are material compositon $x_{A,F}$ and $x_{B,F}$. The time delay $d=48, n=25, m=23$. We use u, y data to train following hybrid neuro model

$$\widehat{y}=\theta_k X_l(k)+W_k\phi[V_kX_n(k)] \tag{32}$$

where $V_k\in R^{9\times 5}, W_k\in R^{1\times 9}$,the initial conditions for the elements of V^{0T}, W_k, V_k are random number in $[0,1]$. The learning algorithm is (31), $e(k)=\widehat{y}(k)-y(k)$, and learning rate

$$\eta_k=\frac{s_k}{1+\|X_l\|^2+\|\phi\|^2+\left\|\phi' VW_kX_n^T\right\|^2},$$
$$s_k=\begin{cases}1 & e(k)^2\ge 1\\ 0 & otherwise\end{cases}$$

where $\phi(\cdot)=\tanh(x)=\frac{e^x-e^{-x}}{e^x+e^{-x}}, \phi'(\cdot)=\mathrm{sech}(x)=\frac{2}{e^x+e^{-x}}$.

We use the following mean squared error for finite time to calculate the average modeling error

$$J(N)=\frac{1}{2N}\sum_{k=1}^{N}[\widehat{y}(k)-y(k)]^2 \tag{33}$$

where N is finite time.

The first 170 groups of data sampled at 24-hour intervals are used to train the hybrid neuro. The other 50 groups of data are used to verify the hybrid neuro. In the training phase, the average modeling error is shown in Fig.3. Modeling errors depend on the complexity of the particular model selected and how close it is to the actual plant. In this application the modeling error is bigger in verifying phase than training phase. The worse results due to the hybrid neuro cannot match the plant exactly. From the point of identification, it is because the parameters are not very close to their ideal value. We should mention that structure the of multilayer neural network also influences modeling error, but does not destroy stability of identification process.

The verified result is shown in Fig.4. *Theorem 2* gives a necessary condition of η for stable learning, $0<\eta\le 1$. The hybrid neuro identification discussed in this paper is online, we do not study the convergence of the weight, we care about the identification error $e(k)$ shown in Fig.5. Under the same verifying conditions, the root mean squared error of hybrid neuro and multilayer neuro are 0.0052403 and 0.0058344 respectively. In training phase and verifying phase, the hybrid neuro simultaneity represents the superi-

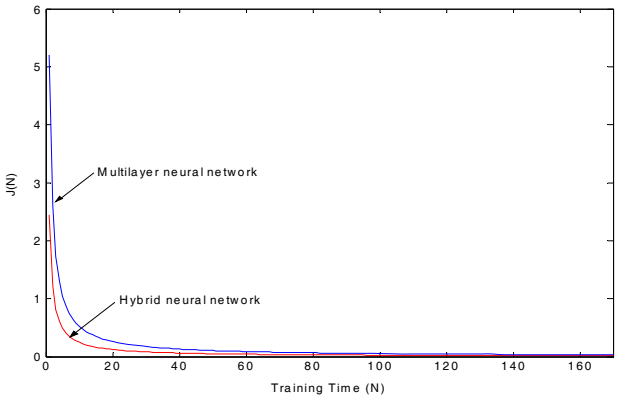

Fig. 3. Modeling error of hybrid neuro and multilayer neuro.

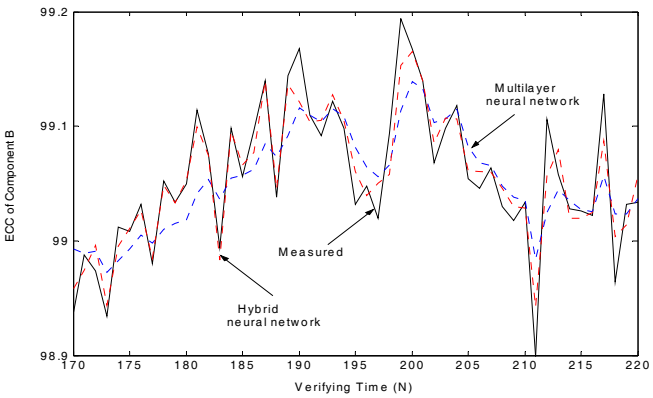

Fig. 4. Estimation of element component content (ECC) via neural network.

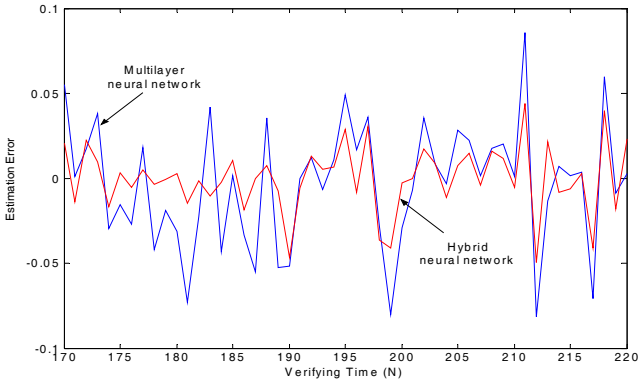

Fig. 5. Error comparison between hybrid neuro and multilayer neuro.

ority to traditional multilayer neural network. It is mainly attributed to the introducetion of process-based internal structure.

5. CONCLUSION

The main contributions of this paper are: 1) A novel mechanism based neural network is proposed to realize grey-box identification. 2) The stability of this modeling method is proven in present of parameter and structure uncertainties. 3) This method is successfully applied on a rare-earth extraction process. Comparing with measured data from an extraction site, the validity and precision of the hybrid neuro is verified.

REFERENCES

T. Chai, H. Yang. Situation and developing trend of rare-earth countercurrent extraction process control. *Journal of Rare Earths*, 22(4): 427-433, 2004.

G. Cybenko. Approximation by superposition of sigmoidal activation function. *Math.Control, Sig Syst*, 2: 303-314, 1989.

Z. Feng, and A. N. Michel. Robustness analysis of a class of discrete-time systems with applications to Neural Networks. *Proc. of American Control Conference*, 3479-3483, San Deigo, 1999.

A. E. Giles. Modelling of rare earth solvent extraction with artificial neural nets. *Hydrometallurgy*, 43(1-3): 241-255, 1996.

P. A. Ioannou, and J. Sun, *Robust Adaptive Control*, Prentice-Hall, Inc, Upper Saddle River: NJ, 1996

L. Jin, and M. M. Gupta. Stable dynamic backpropagation learning in recurrent neural networks. *IEEE Trans. Neural Networks*, 10(6): 1321-1334, 1999.

Mjalli Farouq S, Abdel-Jabbar Nabil M, Fletcher John P. Modeling, simulation and control of a scheibel liquid–liquid contactor Part 1. Dynamic analysis and system identification. *Chemical Engineering and Processing*, 44(5): 543-555, 2005.

R. Oliveira. Combining first principles modelling and artificial neural networks: a general framework. *Computers and Chemical Engineering*, 28: 755-766, 2004.

M. M. Polycarpou, and P. A. Ioannou. Learning and convergence analysis of neural-type structured networks, *IEEE Trans. Neural Networks*, 3(1): 39-50, 1992.

P. Srivastava, O. Hahr, R. Buchholz, R. M. Worden. enhancement of mass transfer using colloidal liquid aphrons: measurement of mass transfer coefficients in liquid-liquid extraction. *Biotechnology and Bioengineering*, 70(5): 525-532, 2000.

G. Stephanopoulos. *Chemical Process Control: an introduction to theory and practice*. Englewood Cliffs, New Jersey, Prentice Hall, 1984.

J. A. K. Suykens, J. Vandewalle, and B. De Moor. NLq Theory: checking and imposing stability of recurrent neural networks for nonlinear modelling. *IEEE Transactions on Signal Processing (special issue on neural networks for signal processing)*, 45(11): 2682-2691, 1997.

M. L. Thompson, M. A. Kramer. Modeling chemical process using prior knowledge and neural networks. *AIChE Journal*, 40(8): 1328-1340, 1994.

C. Tsouris, V. I. Kirou, L. L. Tavlarides. Drop size distribution and holdup profiles in a multistage extraction column. *AIChE Journal*, 40(3): 407-418, 1994.

O. Weinstein, R. Semiat, and D. R. Lewin. Modeling, simulation and control of liquid–liquid extraction columns. *Chemical Engineering Science*, 53(2): 325-339, 1998.

J. Wichterlova, V. Rod. Dynamics behaviour of the mixer-settler cascade. Extractive separation of the rare earths. *Chemical Engineering Science*, 54(18): 4041-4051, 1999.

W. Yu, X. Li. Discrete-time neuro identification without robust modification. *IEE Proceedings - Control Theory and Applications*, 150(3): 311-316, 2003.