Research Article

Hammerstein Box-Jenkins System Identification of the Cascaded Tanks Benchmark System

Ibrahim A. Aljamaan ⬡,[1] **Mujahed M. Al-Dhaifallah** ⬡,[2] **and David T. Westwick** ⬡[3]

[1]*Imam Abdulrahman Bin Faisal University, P. O Box 1982, Dammam, Saudi Arabia*
[2]*Systems Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran 31261, KSA, Saudi Arabia*
[3]*University of Calgary (U of C), 2500 University Dr NW, Calgary, Alberta T2N 1N4, Canada*

Correspondence should be addressed to Ibrahim A. Aljamaan; ibrjmn@gmail.com and Mujahed M. Al-Dhaifallah; mujahed@kfupm.edu.sa

A common process control application is the cascaded two-tank system, where the level is controlled in the second tank. A nonlinear system identification approach is presented in this work to predict the model structure parameters that minimize the difference between the estimated and measured data, using benchmark datasets. The general suggested structure consists of a static nonlinearity in cascade with a linear dynamic filter in addition to colored noise element. A one-step ahead prediction error-based technique is proposed to estimate the model. The model is identified using a separable least squares optimization, where only the parameters that appear nonlinearly in the output of the predictor are solved using a modified Levenberg–Marquardt iterative optimization approach, while the rest are fitted using simple least squares after each iteration. Finally, MATLAB simulation examples using benchmark data are included.

1. Introductions

Recently, system identification attracted the attention of many researchers and practitioners because of the difficulty in modeling many systems using physical modeling approaches. Good empirical mathematical models that represent a variety of practical applications in different engineering fields are required to address various needs [1]. These needs might be to understand and analyze the limitations of current systems, predict and simulate new experiments, or design new or modified controllers. The process of creating a mathematical model from measured data is called system identification [2, 3]. Unlike most machine learning methods, system identification gives more insight about the structure and dynamics of the system. A very common approach in system identifications is by using model structure methods, which is applied in this work. Unlike linear system identification methods that are well established and extensively applied in the literature,

nonlinear system identification needs more attention because our real-life applications are nonlinear. The uniqueness of each nonlinearity increases the complexity of nonlinear systems and their identification [4–11].

There are a few attempts to apply nonlinear system identification techniques to common applications in different engineering fields reported in the literature. In the area of chemical process control, Aljamaan et al. [12] identified the dynamics of the continuously stirred tank reactor (CSTR) in the presence of nonstationary disturbances, Zhu [13] identified a distillation column, and Birpoutsoukis et al. [14] estimated a Volterra series model of the cascaded two-tank system. Similarly, in biomedical engineering applications, Vlaar et al. [15] identified brain signals by electroencephalography (EEG) while Cescon et al. [16] identified models of the blood sugar level inside the human body.

This work is limited to the cascaded two-tank system, where different identification approaches have been

proposed in the literature. Unlike the nonlinear state-space method proposed by Relan et al. [1], Birpoutsoukis et al. [14] developed a nonparametric Volterra series-based approach. In this latter method, cancelling the undesired transient part played an important role in minimizing the error. Svensson and Schön [17] proposed a black-box nonlinear system identification state-space technique and applied it to cascaded tanks datasets. Different system identification approaches were proposed in the literature such as [18–32].

In this paper, a prediction error minimization (PEM) nonlinear system identification parametric approach, similar to the one proposed in [10], is applied to approximate the two-tank cascaded model, where the benchmark data provided by Schoukens et al. [26] were used. The main contribution is to use such general nonlinear model structure, Hammerstein Box-Jenkins model, and separable least squares (SLS) identification approach to predict the parameters of cascaded tanks system. The suggested model structure to fit this application is Hammerstein Box-Jenkins model, where a static nonlinearity is in series with a linear dynamic filter whose output is summed with an autoregressive moving average (ARMA) noise model. Separable least squares (SLS) techniques are suggested, where second-order gradient descent-based optimization approach is performed on the nonlinear parameters while ordinary least squares regression is used, after each iteration, to estimate the coefficients that appear linearly in the predictor.

The arrangement of the paper is as follows. After a physical description of the two-tank system and approximated model structure in Section 2, Section 3 includes the identification suggested approach. Then, simulation examples and results' discussion are illustrated. Finally, concluding remarks are provided in the conclusion.

2. Problem Definition

2.1. The Cascaded Water Tanks. The cascaded water tanks system is a benchmark system proposed in [26]. It comprises two water tanks, with free outlets as shown in Figure 1. The first tank is fed by a pump and drains into the second tank, which itself drains into the reservoir (see Figure 2). The input to the system is the voltage applied to the pump, while the output is the reading on a level sensor in the second tank. Note that if the upper tank overflows, some of the overflow will land in the second tank, while the rest will go directly into the reservoir. The source of nonlinearity is mainly at the input signal and combined with a small nonlinearity during the measurement [26].

2.2. The Proposed Model Structure. The Box-Jenkins model is, in some sense, ideal for the modeling and identification of linear time invariant (LTI) systems because many models can be regarded as special cases such as the AR, MA, ARMA, ARX, ARMAX, and OE models [33]. It includes separate filters for the system and noise models, which gives it a great deal of flexibility. Replacing the linear process filter in the Box-Jenkins model with a nonlinear model produces a much more flexible model, as it can model nonlinear behaviour in
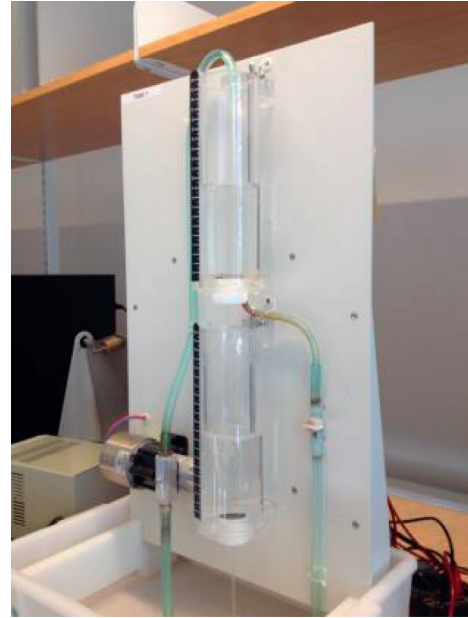

Figure 1: Cascaded tanks system with the pump voltage input and the water lower level tank output. The water is boosted from reservoir in the upper tank to the lower tank and then returns back to the reservoir [26].
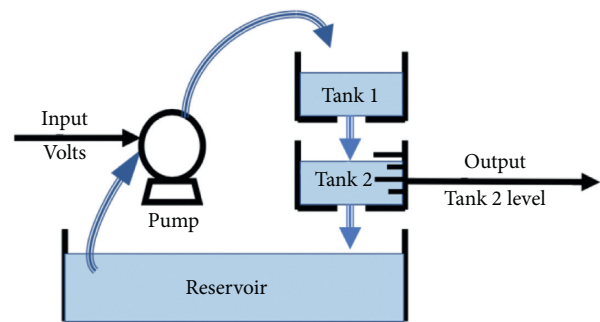

Figure 2: Schematic of the cascaded two-tank system, where the input and output are the pump voltage and water level of tank 2.

the plant but can still provide accurate predictions, even when the stochastic disturbance is highly autocorrelated. In this section, we consider the Hammerstein Box-Jenkins model, shown in Figure 3. It combines the linear dynamics of a Box-Jenkins model with the nonlinear structure of a Hammerstein model. The main reason for selecting Hammerstein is the previous knowledge of the differential balance equations for the two-tank system that imply the main nonlinearity is at the input which is in accordance with the Hammerstein model. Different system identification approaches are presented in [34–40].

The main identification target is to estimate the process and noise parameters and the expansion coefficients of the nonlinearity. An important factor in achieving an acceptable identification result is a good initial guess. This is because when any iterative optimization method is used, it is uncertain that the assessment function will converge to the
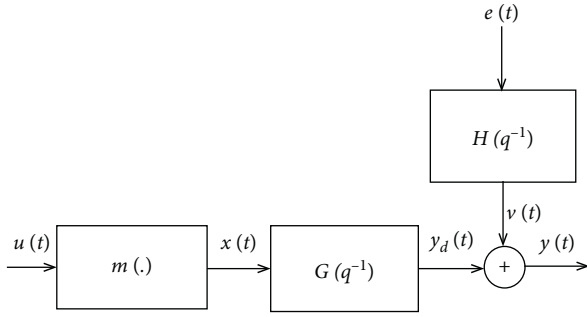
FIGURE 3: Hammerstein Box-Jenkins model.

global minimum unless the initial guess of the parameter vector is close enough. The model is formulated as

$$y(t) = y_d(t) + v(t), \qquad (1)$$

where $y_d(t)$ is the output of the deterministic part of the system and the noise part is

$$v(t) = H\big(q^{-1}\big)e(t) = \frac{B\big(q^{-1}\big)}{A\big(q^{-1}\big)}e(t). \qquad (2)$$

The deterministic process output side is computed by

$$y_d(t) = G\big(q^{-1}\big)x(t), \qquad (3)$$

where $x(t)$ is the output of the static nonlinearity:

$$x(t) = m(u(t)). \qquad (4)$$

In this paper, the memoryless nonlinearity $m(\cdot)$ will be approximated by a polynomial of degree M because of the efficiency and ability to apply LS to find the coefficients of the polynomial nonlinearity.

$$x(t) = \sum_{i=0}^{M} \gamma_i u^i(t). \qquad (5)$$

3. Preliminary

3.1. Prediction Error Method (PEM).
The prediction error method (PEM) searches for the parameter vector, θ, that reduces the difference between the measured output and the output prediction generated by the model. Thus, ε is the prediction error.

$$\begin{aligned} \varepsilon(t|t-1,\theta) &= y(t) - \hat{y}(t|t-1,\theta) \\ &= H^{-1}\big(z^{-1}\big)\big(y(t) - y_d(t,\theta)\big). \end{aligned} \qquad (6)$$

Note that the parameter vector, $\theta \in R^{n \times 1}$, consists of the elements of the process and noise filters and coefficients of the polynomial nonlinearity.

3.1.1. One-Step Ahead Prediction.
The most common option in PEM is the one-step ahead prediction, where the existing output, $y(t)$, at time t is approximated given all I/O measurements up to including $t-1$. The one-step ahead

predictor equation for the Hammerstein Box-Jenkins model is as follows:

$$\hat{y}(t|t-1,\theta) = y_d(t,\theta) + \hat{v}(t|t-1,\theta), \qquad (7)$$

where the noise model predictor, $\hat{v}(t|t-1,\theta)$, is solved by

$$\hat{v}(t|t-1,\theta) = \big(1 - H^{-1}\big(q^{-1}\big)\big)\big(y(t) - y_d(t,\theta)\big). \qquad (8)$$

By substituting (3) and (8) in (7):

$$\begin{aligned} \hat{y}(t|t-1,\theta) &= \big(1 - H^{-1}\big(q^{-1}\big)\big)y(t) + H^{-1}\big(q^{-1}\big) \\ &\quad \cdot \left(G\big(q^{-1}\big)\sum_{i=0}^{M}\gamma_i u^i(t)\right). \end{aligned} \qquad (9)$$

The main objective, (10), is to compute the parameter vector θ that minimizes the mean squared error (MSE) of the cost function, $V_N(\theta)$, in (11).

$$\hat{\theta} = \arg \min_{\theta \in DM} V_N(\theta), \qquad (10)$$

$$V_N = \frac{1}{N}\sum_{t=1}^{N}\varepsilon^2(t,\theta). \qquad (11)$$

3.2. Separable Least Squares.
In the SLS algorithm, the parameter vector, θ, is divided into two sets: a group of parameters that appears linearly in the prediction error, θ_ℓ, and the remaining parameters, θ_n [41]. The prediction error is generally a nonlinear function of θ_n; although in some cases, one or more elements of θ_n may appear bilinearly with elements of θ_ℓ. In these cases, the decision as to which parameter is "linear" or "nonlinear" is arbitrary. Thus, the parameter vector is subdivided as

$$\theta = [\theta_n | \theta_l]. \qquad (12)$$

For the Hammerstein model, the prediction error can be treated as linear in either the polynomial coefficients or the plant numerator. The output of process model, $y_d(t)$, and polynomial X is formulated as follows:

$$y_d(t) = X(\theta_n)\theta_\ell, \qquad (13)$$

$$X(\theta_n, t, :) = \big[\, G(\theta_n)u^0 \ \ G(\theta_n)u^1 \ \ G(\theta_n)u^2 \ \ \ldots \ \ G(\theta_n)u^M \,\big]. \qquad (14)$$

The parameter vector of the linear part, θ_ℓ, is estimated by doing simple LS.

$$\theta_l = \big(X^T(\theta_n)X(\theta_n)\big)^{-1}X^T(\theta_n)Y, \qquad (15)$$

where Y is the vector that contains all experimental results of $y(t)$. The nonlinear parameter vector, θ_n, is estimated using an iterative approach.

$$\theta_{n(k+1)} = \theta_{n(k)} + d_k. \qquad (16)$$

It is noticed from (16) that θ_ℓ relied on θ_n in finding the direction of the search. Similar to the parameter vector θ, the

Jacobian is divided into two parts too: J_ℓ and J_n that consist of the linear parameters and nonlinear coefficients, respectively:

$$J = [J_\ell | J_n], \tag{17}$$

and formulated in (18) and (19):

$$J_n(t, :) = \frac{\partial \widehat{y}(t|t-1, \theta)}{\partial \theta_n}, \tag{18}$$

$$J_l(t, :) = \frac{\partial \widehat{y}(t|t-1, \theta)}{\partial \theta_\ell}. \tag{19}$$

The Jacobian of the SLS J_s is computed by

$$J_s(t, :) = \frac{\partial \widehat{y}(\theta_n, \theta_\ell(\theta_n))}{\partial \theta_n}. \tag{20}$$

Because the predictor output relies on the nonlinear parameters that might result variations in the linear elements the projection of the linear part, P_ℓ, defined in (21), is required to calculate this variation.

$$P_\ell = J_\ell (J_\ell^T J_\ell)^{-1} J_\ell. \tag{21}$$

The complementary projection, Q_ℓ, is defined in the following equation:

$$Q_\ell = I - P_\ell. \tag{22}$$

Now J_s can be calculated by

$$J_s = Q_\ell J_n. \tag{23}$$

Once J_s has been computed, θ_n can be updated using any quasi-Newton optimization, and Levenberg–Marquardt is proposed in this work. The separable least squares approach is described in more detail by Aljamaan et al. [12].

The nonlinear vector parameters θ_n can be updated using any quasi-Newton optimization.

3.3. L-M Optimization. The Levenberg–Marquardt algorithm is a second-order iterative optimization-based technique, where the update applied to the parameter depends on the local gradient and the approximation of the error surface's local curvature [42, 43]. It is applied to find the nonlinear LS problems, where minimum is explored, but it is not guaranteed to be the global minimum. Also, it is used in software applications to fit best curve. The phase direction and length were measured together in the L-M iteration. The update is defined in (14).

Note that the distance and the direction of the kth step are chosen so that the cost decreases.

$$V_n[\theta_{(k+1)}] \leq V_n[\theta_{(k)}]. \tag{24}$$

The L-M method formulas are as follows:

$$\widehat{H} = -\frac{1}{N} J^T J + \mu_{(k)} I_M, \tag{25}$$

$$d_{(k)} = \left(-\frac{1}{N} J^T J + \mu_{(k)} I_M\right)^{-1} J^T \varepsilon = \widehat{H}^{-1} J^T \varepsilon, \tag{26}$$

where ε is the error and J is the Jacobian. $d_{(k)}$ is the updated step to control the step size and also responsible for speeding up of the convergence and ensures the stability, and $\widehat{H}$ is the approximate Hessian; the step size for the Hessian approximation is included. IM is the $M \times M$ identity Matrix.

3.4. Instrumental Variable (IV). The global minimum can be always achieved by using LS if the parameters of the model are linear. On the other hand, models that are nonlinear in their parameters require iterative optimization techniques to converge to a minimum. This minimum may not be the global minimum and hence the importance of a good initial guess. Moreover, predicting an excellent initial guess plays a significant factor to have a successful identification. The IV approach, used in this work, is a familiar option that yields unbiased approximated coefficients (this is explained in [2, 3]). An extended IV approach is provided in [44].

3.5. Hammerstein Box-Jenkins Model Identification Using SLS. In the following, a summary of the proposed SLS identification algorithm (Algorithm 1) is presented.

4. Experimental Results

MATLAB simulation examples using two datasets of 1024 points each, identification and validation data, were provided by Vlaar et al. [15]. These datasets were collected from the cascade tanks system shown in Figure 2. The input u (t) is a multisine signal, which is the voltage applied to pump the water from the reservoir to the upper tank with frequency range between 0 and 0.0144 Hz and sampling time Ts equal to 4 seconds. The water level of the lower tank is the output of the system, y (t), with signal to noise ratio, SNR, approximately equal to 40 dB. Input and output are shown in Figure 4. More information about the data is included in [15].

A Hammerstein Box-Jenkins model structure similar to the illustrated one in Figure 3 is applied to identify the approximated mathematical model. A third-order polynomial nonlinearity in cascade with second-order filter that represents the plant model is selected. Unlike the process model, the noise subsystem was of first-order ARMA filter.

The order was selected manually using the training data. On the other hand, Table 1 illustrates the MSE results using the validated dataset. MSE function was chosen as the assessment tool. The first experiment was applied by selecting a first-order process and noise model with zero-order nonlinearity, and it produces high MSE results. Then, the order of the nonlinearity increases and it is clear that the cost decreases in the 2nd and 3rd nonlinearity cases. Unlike in the 2nd and 3rd cases, there is very small difference in the cost between 3rd and 4th order nonlinearities. After that, the

```
(1) Find initial guess $\theta_{n(0)}$ using IV method.
(2) Compute $\hat{x}(t)$ using (5), and then fit $\theta_{\ell(0)}$ using LS.
(3) Find $J_\ell$ and $J_n$ as in (18) and (19).
(4) Calculate the updated step $d_k$, (26), and then update, $\theta_{n(k+1)}$, (16).
(5) Compute $\hat{x}(t)$, (5), using updated $\theta_{n(k+1)}$, and fit $\theta_{(\ell+1)}$.
(6) Using LS.
(7) Compute the error and cost function, $\varepsilon$ and $V_n$, respectively.
(8) If the cost $V_n$ decreases.
(9)         If $V_n$ is significant Then
(10)        Exit
(11)        else increase speed of convergence by setting the update step $\delta = \delta/\gamma$, then go back to step 3
(12)        End
(13) else the cost, $V_n$, increases Then
(14) Decrease the speed of convergence by setting: $\delta = \delta\gamma$ and go back to step 5
(15) end
```

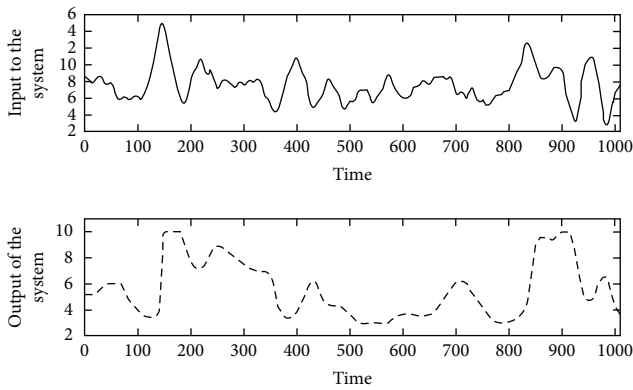ALGORITHM 1: Summary of the proposed identification algorithm.

FIGURE 4: Recorded input to of the cascaded tanks system, pump voltage, and output of the cascaded tanks system, the water level at the lower tank.

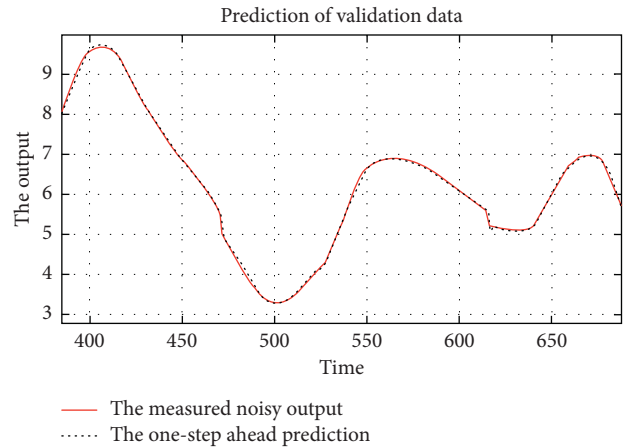The measured noisy output
······ The one-step ahead prediction

FIGURE 5: Zoomed figure of the exact and estimated output using validation data.

TABLE 1: The orders of the plant model vs. the nonlinearity and the resulted mean square error (MSE).

Linear plant order	Nonlinearity order	MSE
1	0	3.0031
1	1	0.0099
1	2	0.0093
1	3	0.0092
1	4	0.009
2	3	0.0073
3	3	0.0087

process linear dynamic order was increased, and it was noticed that using 2nd order plant model and 3rd order nonlinearity produces the minimum cost. Notice that a 1st order noise model was fixed for all the experiments.

The proposed parametric identification method based on Section 3 was applied. Figure 5 shows the difference between the simulated outputs, dotted black lines, that tracks the measured output represented in red solid line. Moreover, the cost function using iterative second order gradient descent method based is plotted in Figure 6, where the final
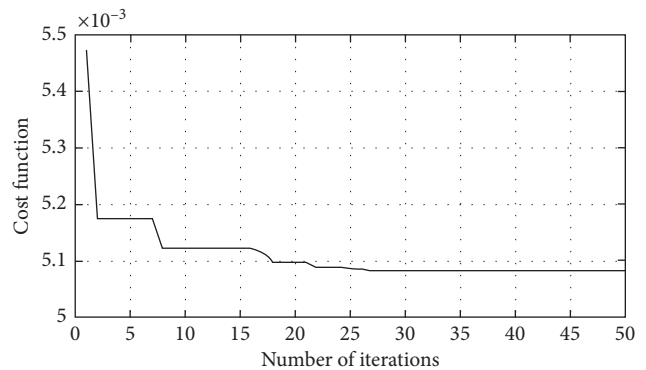

FIGURE 6: The cost convergence using MSE as the cost function method.

convergence was at iteration 25. The nonlinearity is illustrated in Figure 7. The mean square error, MSE, of the 2nd order linear dynamic subsystem, the 1st order noise filter, and 3rd order nonlinearity was the minimum with 0.0073.
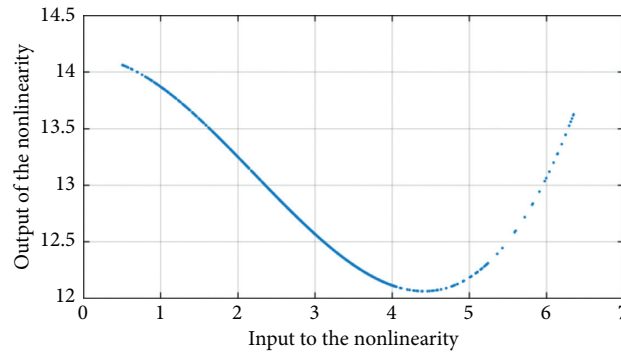
FIGURE 7: The nonlinearity of the system.

TABLE 2: A comparison of MSE results using different approaches from previous studies using the validated dataset.

Method	MSE (training)	MSE (validation)
PNLSS-I (polynomial nonlinear state space)	0.0324	0.4498
NLSS2	0.1165	0.3433
BLA (best linear approximation)	0.5471	0.7574
Volterra models (3rd degree) (transient removal)		0.5400
Volterra models (3rd degree) (no transient est.)		2.0800
Proposed HBJ	0.0045	0.0087

— True
--- RLS

(a)

— The measured noisy output
...... The predicted output (simulated data)

(b)

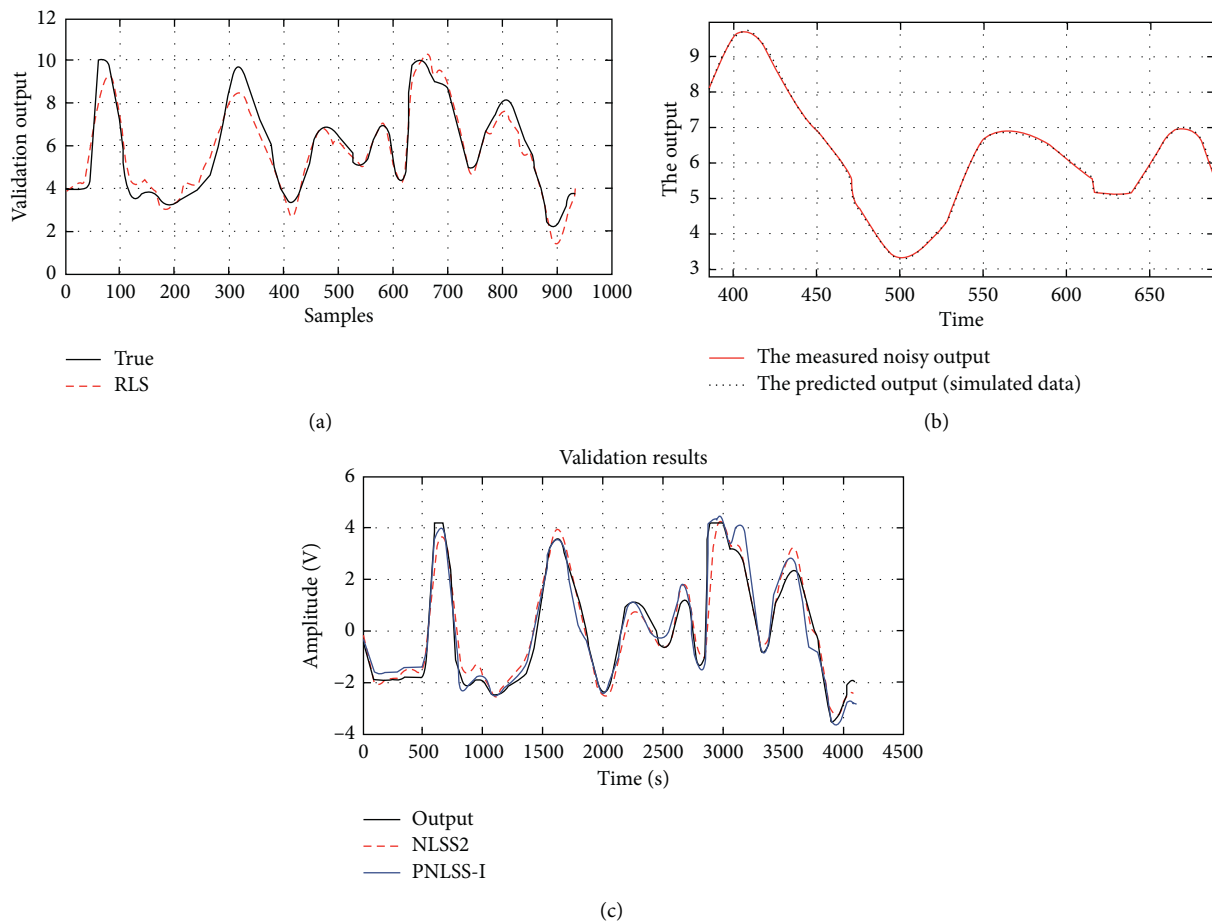— Output
--- NLSS2
— PNLSS-I

(c)

FIGURE 8: (a) Comparison of the measured and estimated output using RLS approach (Figure 4(c) in [14]). (b) The real and estimated output using the proposed method HBJ and (c) real output against the predicted using PNLSS-I and NLSS2 models using simulated data (Figure 3 in [1]).

A comparison between the results obtained by applying this proposed approach and two previous studies using the same benchmark data by Relan et al. and Birpoutsoukis et al. [1, 14] is given in Table 2 and Figure 8. It is clear that the proposed Hammerstein Box-Jenkins approach outperformed the other algorithms and has the least MSE value.

5. Conclusions

In this research, a nonlinear identification approach is applied to estimate a cascade tank system from data provided on the nonlinear system identification benchmark website as one of the challenging datasets, provided by Aljamaan[11]. The suggested model structure that fits this application is a Hammerstein Box-Jenkins model. This parametric method is prediction error based, where the linear dynamic process and noise subsystem elements are approximated using iterative Hessian method optimization technique; on the other hand, the polynomial nonlinearity coefficients are fitted using the simple LS method after each iteration because they appear linearly in the predictor. In order to prevent the ill-condition issue and increase the chance of convergence, an IV method is used for the initial guess. Different orders were tested to select the suitable one, and simulated data were applied to validate the results. The final results are promising for the successful identification of real two-tank systems.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

All authors collaborated and contributed equally to this study. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

References

[1] R. Relan, K. Tiels, A. Marconato, and J. Schoukens, "An unstructured flexible nonlinear model for the cascaded water-tanks benchmark," in *Proceeding of the 20th World Congress IFAC (The International Federation of Automatic Control) Conference*, pp. 454–459, Toulouse, France, July 2017.

[2] L. Ljung, *System Identification Theory for the User*, Prentice-Hall, Upper Saddle River, NJ, USA, 1999.

[3] T. Söderstrom and P. Stoica, *System Identification*, Prentice-Hall, Upper Saddle River, NJ, USA, 1989.

[4] R. Pintelon and J. Schoukens, *System Identification A Frequency Domain Approach*, Wiley-IEEE Press, Piscataway, NJ, USA, 2012.

[5] F. Giri and E. Bai, *Block-oriented Nonlinear System Identification*, Springer, Berlin, Germany, 2010.

[6] D. Westwick and R. Kearney, *Identification of Nonlinear Physiological System*, Wiley-IEEE Press, Piscataway, NJ, USA, 2003.

[7] J. Schoukens and L. Ljung, "Nonlinear system identification: a user-oriented road map," *IEEE Control Systems Magazine*, vol. 39, no. 6, pp. 28–99, 2019.

[8] O. Nelles, *Nonlinear System Identification*, Springer-Verlag, New York, NY, USA, 2001.

[9] K. J. Åström and P. Eykhoff, "System identification-a survey," *Automatica*, vol. 7, no. 2, pp. 123–162, 1971.

[10] A. K. Tangirala, *Principles of System Identification: Theory and Practice*, CRC Press, Boca Raton, Fl, USA, 2014.

[11] I. Aljamaan, "Nonlinear identification of two different closed-loop models using separable least squares methods," Master's Dissertation, University of Calgary, Calgary, Canada, 2011, http://library.ucalgary.ca/.

[12] I. Aljamaan, D. Westwick, and M. Foley, "Non-linear closed-loop identification of CSTR in the presence of a non-stationary disturbances," *IFAC-PapersOnLine*, vol. 48, no. 28, pp. 1029–1034, 2015.

[13] Y. Zhu, "Distillation column identification for control using wiener model," in *Proceedings of the American Control Conference*, vol. 5, pp. 3462–3466, San Diego, CA, USA, June 1999.

[14] G. Birpoutsoukis, P. Z. Csurcsia, and J. Schoukens, "Non-parametric Volterra series estimate of the cascaded water tanks using multidimensional regularization," in *Proceedings of the 20th World Congress IFAC (The International Federation of Automatic Control) Conference*, pp. 478–483, Toulouse, France, July 2017.

[15] M. P. Vlaar, G. Birpoutsoukis, J. Lataire et al., "Modeling the nonlinear cortical response in EEG evoked by wrist joint manipulation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering* in *Proceedings of the*, vol. 26, no. 1, pp. 205–215.

[16] M. Cescon, F. Ståhl, M. Olsson, and R. Johansson, "Subspace-based model identification of diabetic blood glucose dynamics," in *Proceedings of the 15th IFAC Symposium on System Identification*, pp. 233–238, Saint-Malo, France, March 2009.

[17] A. Svensson and T. B. Schön, "A flexible state-space model for learning nonlinear dynamical systems," *Automatica*, vol. 80, pp. 189–199, 2017.

[18] J. Belz, T. Münker, T. Heinz, G. Kampmann, and O. Nellez, "Automatic modeling with local model networks for benchmark processes," in *Proceedings of the 20th World Congress IFAC (The International Federation of Automatic Control) Conference*, pp. 472–477, Toulouse, France, July 2017.

[19] M. Brunot, A. Janot, and F. Carrillo, "Continuous-time nonlinear systems identification with output error method based on derivative-free optimisation," in *Proceedings of the The 20th World Congress IFAC (The International Federation of Automatic Control) Conference*, pp. 466–471, Toulouse, France, July 2017.

[20] M. Schoukens and J. Noël, "Three benchmarks addressing open challenges in nonlinear system identification," in *Proceedings of the the 20th World Congress IFAC (The*

International Federation of Automatic Control) Conference, pp. 448–453, Toulouse, France, July 2017.

[21] H. Zhou, C. Ibrahim, and W. Pan, "A sparse bayesian deep learning approach for identification of cascaded tanks benchmark," 2019, https://arxiv.org/abs/1911.06847.

[22] M. H. D. M. Ribeiro, V. C. Mariani, L. Manke, and L. S. Coelho, "Nonlinear system identification based on ensemble learning: the cascaded tanks system case," in *Proceedings of the 14th Brazilian Symposium on Intelligent Automation (SBAI)*, pp. 415–420, Ouro Preto, Brazil, October 2019.

[23] G. Giordano and J. Sjöberg, "Black-and white-box approaches for cascaded tanks benchmark system identification," *Mechanical Systems and Signal Processing*, vol. 108, pp. 387–397, 2018.

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

[25] J. G. Stoddard and J. S. Welsh, "Regularized basis function estimation of Volterra kernels for the cascaded tanks benchmark," *IFAC-PapersOnLine*, vol. 51, no. 15, pp. 413–418, 2018.

[26] M. Schoukens, P. Mattson, T. Wigren, and J.-P. Noël, "Cascaded tanks benchmark combining soft and hard nonlinearities," in *Proceedings of the Conference: ERNSI 2015*, Varberg, Sweden, September 2015, http://homepages.vub.ac.be/~7Emschouke/benchmark2016.html.

[27] T. Wigren and J. Schoukens, "Data for benchmarking in nonlinear system identification," Technical Reports, Department of Information Technology 2013-006, Uppsala, Sweden, 2013, http://www.it.uu.se/research/publications/reports/2013-006/SNLA80mVZipped.zip.

[28] B. Stenlund and A. Medvedev, "Level control of cascade coupled flotation tanks," *IFAC Proceedings Volumes*, vol. 33, no. 22, pp. 185–190.

[29] T. Wigren, "Recursive prediction error identification and scaling of non-linear state space models using a restricted black box parameterization," *Automatica*, vol. 42, no. 1, pp. 159–168, 2006.

[30] T. Wigren, "MATLAB software for recursive identification and scaling using a structured nonlinear black-box model—revision 2," Technical reports 2005-022, Department of Information Technology, Uppsala University, Uppsala, Sweden, 2005, http://www.it.uu.se/research/reports/..

[31] T. Wigren, *Input-output Data Sets for Development and Benchmarking in Nonlinear Identification*, Technical Reports, Department of Information Technology, Uppsala University, Uppsala, Sweden, 2010.

[32] K. Hamid and M. Fadali, "Nonlinear system identification using neural networks and trajectory-based optimization," 2018, https://arxiv.org/abs/1804.10346.

[33] G. Box and G. Jenkins, *Time Series Analysis Forecasting and Control*, Holden Day, San Francisco, CA, USA, 1970.

[34] I. A. D. Westwick and M. Foley, "Prediction error identification of Hammerstein models in the presence of ARIMA disturbances," in *Proceedings of the IEEE Conference on Control Applications (CCA)*, Juan Les Antibes, France, October 2014.

[35] J. Sjöberg and M. Viberg, "Separable nonlinear least squares minimization-possible improvements for neural net fitting," *IEEE Workshop on Neural Networks for Signal Processing*, vol. 7, pp. 345–354, 1997.

[36] E.-W. Bai and D. Li, "Convergence of the iterative Hammerstein system identification algorithm," *IEEE Transactions on Automatic Control*, vol. 49, no. 11, pp. 1929–1940, 2004.

[37] D. T. Westwick and R. E. Kearney, "Separable least squares identification of nonlinear Hammerstein models: application to stretch reflex dynamics," *Annals of Biomedical Engineering Society*, vol. 29, no. 8, pp. 707–718, 2001.

[38] E. J. Dempsey and D. T. Westwick, "Identification of Hammerstein models with cubic spline nonlinearities," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 2, pp. 237–245, 2004.

[39] K. Narendra and P. Gallman, "An iterative method for the identification of nonlinear systems using a Hammerstein model," *IEEE Transactions on Automatic Control*, vol. 11, no. 3, pp. 546–550, 1966.

[40] S. A. Billings and S. Y. Fakhouri, "Identification of systems containing linear dynamic and static nonlinear elements," *Automatica*, vol. 18, no. 1, pp. 15–26, 1982.

[41] A. Ruhe and P. Å. Wedin, "Algorithms for separable nonlinear least squares problems," *SIAM Review*, vol. 22, no. 3, pp. 318–337, 1980.

[42] A. Ranganathan, "The levenberg-marquardt algorithm," *Tutoral on LM Algorithm*, vol. 11, no. 1, pp. 101–110, 2004.

[43] J. Moré, "The Levenberg-Marquardt algorithm: implementation and theory," in *Numerical Analysis*, pp. 105–116, Springer, Berlin, Heidelberg, 1978.

[44] V. Laurain, M. Gilson, H. Garnier, and P. Young, "Re-fined instrumental variable method for identification of Hammerstein continuous-time box-jenkins models," in *Proceedings of the 47th IEEE Conference on Decision and Control*, pp. 1386–1391, Cancun, Mexico, December 2008.