CHEMICAL ENGINEERING TRANSACTIONS

VOL. 76, 2019

A publication of AIDIC

The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Petar S. Varbanov, Timothy G. Walmsley, Jiří J. Klemeš, Panos Seferlis

**ISBN** 978-88-95608-73-0; **ISSN** 2283-9216

DOI: 10.3303/CET1976111

# Improved Intelligent Identification of Hammerstein-Wiener Systems by Particle Swarm Optimization and K-Means Clustering

Zhu Wang*, Haoran An, Xionglin Luo

Department of Automation, China University of Petroleum, Beijing 102249, PR China
zwang@cup.edu.cn

This paper considers the robust identification of Hammerstein-Wiener systems in the presence of Gaussian or non-Gaussian noises. An improved intelligent identification scheme is exploited by combining particle swarm optimization (PSO) and K-means clustering. The proposed scheme has strong ability to keep the balance between exploration and exploitation. Its procedure is about "global particle swarm optimization search — K-means clustering — local particle swarm optimization search". The proposed scheme can identify the parameters of the general Hammerstein-Wiener system with dead zone and saturation characteristics, and obtain a more accurate model for the actual production process. Relative to other improved particle swarm optimization methods, the accuracy of parameter estimation is improved by nearly 53 % at data length L=2000. In particular, the method can better model nonlinear dynamics and facilitate the precise implementation of control in chemical production.

## 1. Introduction

Nonlinear systems widely exist in industrial processes. In order to achieve accurate modelling for nonlinear systems, the block-oriented nonlinear models that separate system statics and dynamics are employed (Cao and Luo, 2014). A Hammerstein-Wiener model (Figure 1) can give good descriptions of the nonlinearities of both actuators and sensors. The identification of Hammerstein-Wiener models with polynomial or dead-zone (Wang et al., 2017) can be found in the literature. Hammerstein-Wiener models can describe some processes completely and explicitly. e.g., the ionospheric dynamic process and the polymerase chain reaction process.

The key of accurate modelling of block-oriented nonlinear systems is to obtain global optimal parameter estimations under known and complex structures (Schoukens and Tiels, 2017). Based on measurement data of systems, lots of techniques have been developed for nonlinear parameter estimation. These include iterative methods (Behl et al., 2019), subspace-based techniques. In recent contributions, the conclusive prediction error (PE) framework is adopted to solve the identification of general Hammerstein-Wiener models (Rasouli et al., 2015), and the maximum likelihood (ML) framework is employed to further allow for a disturbance before the final Wiener nonlinearity (Wills et al., 2013). In pre-existing literature, the Gaussian-based colored or white noise assumptions are necessary. In engineering areas, the measurement distributions are non-Gaussian because they contain outliers. Since the swarm intelligence belongs to the nature-inspired optimization framework (Gotmare et al., 2015), its relevant techniques have powerful global search abilities that can be applied into the identification of block-oriented nonlinear systems under Gaussian or non-Gaussian noises. The nature-inspired optimization can make up for the weaknesses of PE and ML effectively. Compared with the simple PSO, this paper uses an adaptive inertia weight (Maleki et al., 2015) for PSO and the K-means clustering (Tang et al., 2017) to avoid the search being trapped in local optima. The whole intelligent scheme is about "global PSO search — K-means clustering — local PSO search". Meanwhile, through the appropriate algorithmic settings, a good balance between exploration and exploitation can be harvested, and the global optimal solution can be found.

The identification results by the proposed scheme are compared with the results obtained by the CPSO (Chen et al., 2015) and NPSO (Jin et al., 2013) intelligent search algorithms. The proposed scheme can estimate more

Paper Received: 05/03/2019; Revised: 25/04/2019; Accepted: 25/04/2019
Please cite this article as: Wang Z., An H., Luo X.-L., 2019, Improved Intelligent Identification of Hammerstein-Wiener Systems by Particle Swarm Optimization and K-Means Clustering, Chemical Engineering Transactions, 76, 661-666 DOI:10.3303/CET1976111

accurate parameters to improve the quality of modelling. This is a follow-up to better predict, control, fault diagnosis to lay the foundation. Applying the scheme can increase production efficiency, save energy and reduce pollution.

## 2. Problem formulation

### 2.1 Description of model structure

Consider the nonlinear dynamic Hammerstein-Wiener system shown in Figure 1, where $u_t$ and $y_t$ are input and output respectively, $e_t$ is zero-mean stationary stochastic noise, and $v_t$ and $x_t$ represent internal variables. Note that the distribution of $e_t$ can be either Gaussian or non-Gaussian. From Figure 1, the Hammerstein-Wiener system is formed by a linear block $G(z^{-1})$ and two memoryless nonlinearities $\{f_H, f_W\}$. Among them, the linear block $G(z^{-1})$ is depicted by a discrete transfer function:

$$G(z^{-1}) = \frac{\beta_1 z^{-1} + \cdots + \beta_m z^{-m}}{1 + \alpha_1 z^{-1} + \cdots + \alpha_n z^{-n}}, \tag{1}$$

where $m$ and $n$ are model orders. The memoryless nonlinearities $f_H(\cdot, \boldsymbol{\rho})$ and $f_W(\cdot, \boldsymbol{\gamma})$ are quite general, and the only restriction is that the derivatives $\frac{\partial}{\partial \boldsymbol{\rho}} f_H(\cdot, \boldsymbol{\rho})$ and $\frac{\partial}{\partial \boldsymbol{\gamma}} f_W(\cdot, \boldsymbol{\gamma})$ with respect to their parameter vectors $\boldsymbol{\rho}$ and $\boldsymbol{\gamma}$ exist. Thus, the system model in the presence of stochastic noise is expressed as:

$$y_t = f_W(x_t, \boldsymbol{\gamma}) + e_t, \tag{2}$$

$$x_t = G(z^{-1}) \cdot v_t, \tag{3}$$

$$v_t = f_H(u_t, \boldsymbol{\rho}). \tag{4}$$

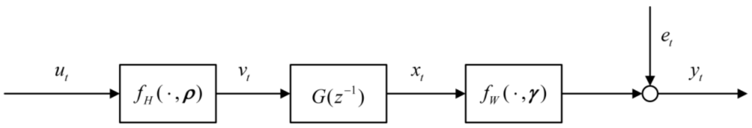


*Figure 1: The Hammerstein-Wiener model structure*

Let $\boldsymbol{\alpha}^{\mathrm{T}} = [\alpha_1, \ldots, \alpha_n]$ and $\boldsymbol{\beta}^{\mathrm{T}} = [\beta_1, \ldots, \beta_m]$. From Eq(1) to Eq(4), it is easy to see that the system structure is parameterized by the parameter vector $\boldsymbol{\theta} = \left[\boldsymbol{\rho}^{\mathrm{T}}, \boldsymbol{\alpha}^{\mathrm{T}}, \boldsymbol{\beta}^{\mathrm{T}}, \boldsymbol{\gamma}^{\mathrm{T}}\right]^{\mathrm{T}}$. Since the internal variables $v_t$ and $x_t$ are unavailable, the objective of the paper is to exploit an improved intelligent identification scheme to obtain the parameter vector $\boldsymbol{\theta}$.

### 2.2 Description of nonlinear blocks

In order to give the reasonable descriptions of memoryless nonlinear blocks $f_H(\cdot, \boldsymbol{\rho})$ and $f_W(\cdot, \boldsymbol{\gamma})$ that satisfy the condition of derivative existence, we discuss following fundamental cases:
The common nonlinear blocks contain saturation nonlinearity, dead-zone nonlinearity and piecewise linearities. Referring to the work in (Wills et al., 2013), the piecewise linear modeling method is adopted for the description of non-smooth nonlinearities. Assuming that the nonlinearity $f_H(\cdot, \boldsymbol{\rho})$ is non-smooth, it can be described by an initial linear term together with a number $l_k$ of “hinge” functions $h_k(\cdot, \boldsymbol{\rho})$:

$$f_H(u_t, \boldsymbol{\rho}) = \rho_{0,0} + \rho_{0,1} \cdot u_t + \sum_{k=1}^{l_k} h_k(u_t, \boldsymbol{\rho}) \tag{5}$$

Then, the hinges are built by

$$h_k(u_t, \boldsymbol{\rho}) = \begin{cases} \rho_{k,0} + \rho_{k,1} \cdot u_t, & u_t > -\rho_{k,0} / \rho_{k,1} \\ 0, & otherwise \end{cases} \tag{6}$$

Here, the parameter vector $\boldsymbol{\rho}$ is shown as:

$$\boldsymbol{\rho}^{\mathrm{T}}=\left[\rho_{0,0},\,\rho_{0,1},\,\rho_{1,0},\,\rho_{1,1},\,\ldots,\,\rho_{l_k,0},\,\rho_{l_k,1}\right] \tag{7}$$

It is shown that Eq(5) to Eq(7) are used for the description of non-smooth nonlinearities. For example, a dead-zone or a saturation can be described by a linear base together with two hinges. That is, $l_k$ is equal to 2, and the vector $\boldsymbol{\rho}^{\mathrm{T}}$ contains a number of six parameters.

## 3. Improved intelligent identification scheme

### 3.1 Cluster analysis for elite swarm

In this section, the nonlinear identification problem is framed as swarm intelligence-based optimization problem. Specifically, the positions of particles in PSO algorithm correspond to possible parameter estimation vectors of Hammerstein-Wiener systems. The following mean-squared-error (MSE) index is derived as the fitness function of PSO search:

$$MSE=(1/L)\cdot\sum_{k=1}^{L}\left(y_t(k)-\hat{y}_t(k)\right)^2 \tag{8}$$

where $L$ is the number of measurement samples, and estimated output $\hat{y}_t(k)$ is obtained by estimated model. The smaller the fitness values are, the better the search quality is. It should be noted that the estimated model is formed by the identification result of model parameters. According to the fitness function, the top $N$/5 particles are reserved as candidate solutions. For K-means, determine the suitable value of cluster number $k_s$ according to the silhouette coefficient (Jia and Deng, 2018) and divide candidate solutions into $k_s$ categories. Next, Define $E_i^{(r)}=\left[e_{i1}^{(r)},\ldots,e_{id}^{(r)}\right]^{\mathrm{T}}$ as the $I$-th particle in the $r$-th cluster, and $N_r$ as the number of elite particles in the $r$-th cluster (i.e. $N_1+\cdots+N_{k_s}=N/5$ ). According to different clusters, establish boundaries for multiple local search spaces:

$$\begin{aligned}&e_{\min j}^{(r)}=\min\left\{e_{ij}^{(r)}|i=1,\ldots,N_r\right\},\ e_{\max j}^{(r)}=\max\left\{e_{ij}^{(r)}|i=1,\ldots,N_r\right\},\ j=1,2,\ldots,d,\\&E_{\min}^{(r)}=\left[e_{\min 1}^{(r)},\ldots,e_{\min d}^{(r)}\right],\ E_{\max}^{(r)}=\left[e_{\max 1}^{(r)},\ldots,e_{\max d}^{(r)}\right],\ r=1,2,\ldots,k_s,\end{aligned} \tag{9}$$

where $E_{\min}^{(r)}$ and $E_{\max}^{(r)}$ are the lower and upper bounds in the $r$-th local search space. Finally, for the exploitation of the promising local spaces, randomly generate $N$/10 new particles within each local space.

Because the parameter search space is large (Nezamivand Chegini et al., 2018), the lack of prior knowledge of model parameters restricts the search quality and identification accuracy. In view of the above situation, the elite swarm is taken for analysis instead of the best particle.

From the above procedure of cluster analysis, since the relevant minimum and maximum values are extracted for the $j$-th part of particle boundary, the Eq(9) is able to ensure the relatively large local search boundaries. There may exist some overlaps between the local search spaces. The focus of the problem is to ensure the population diversity in the top $N$/5 particles. The effective approach to solve the problem involves two aspects: (a) the parameter search space for global PSO should be set large ensures the diversity of population and contributes to the global search of whole space; (b) set the iteration number of global PSO to a relatively small number. This is because the smaller the iteration number is, the bigger the particle differences are. However, if the iteration number is chosen too small, the cluster analysis results start to deteriorate rapidly. Here, Define $P_g^{(r)}$ as the best position that the swarm has reached so far in the $r$-th local space, the different iteration numbers are tested by the varying fitness values $MSE_g$ of best particle $P_g$ during the iteration calculation.

### 3.2 Intelligent identification scheme by PSO and clustering

By combining the PSO search and the K-means clustering, an improved intelligent identification scheme is proposed to harvest robust parameter estimation in the presence of stochastic noise. In short, the whole intelligent identification scheme is about “global PSO search — K-means clustering — local PSO search”. The detailed intelligent identification scheme is provided as Figure 2.

There are two points to note in the intelligent identification scheme by PSO and clustering: (a) For keeping balance between exploitation and exploration in PSO, an adaptive inertia weight is proposed as:

$$w = w_{\max} - k/k_{\max} \cdot (w_{\max} - w_{\min}) + r_3/\sqrt{k} \tag{10}$$

where $k$ represents the current iteration, $k_{\max}$ is the maximum iteration number, $r_3$ is the random number in the range of [0, 1], $w_{\min}$ and $w_{\max}$ are the lower and upper bounds respectively.(b) When the particle flies out of boundaries, this particle can be reset as:

$$x_{ij}^{(r)}(k) = e_{\min j}^{(r)} + r_4 \cdot (e_{\max j}^{(r)} - e_{\min j}^{(r)}),\quad x_{ij}^{(r)} > e_{\max j}^{(r)} \ \ or \ \ x_{ij}^{(r)} < e_{\min j}^{(r)}, \tag{11}$$

where $r_4$ is a random number in the range of [0, 1]. The modified boundary processing method can improve the search efficiency significantly.

---

Intelligent identification scheme by PSO and clustering

---

Begin

1. Initialization:
   - 1.1 Initialize the population number $N$, the initial search space and the particle velocities and positions.
   - 1.2 Evaluate the particle fitness values.
   - 1.3 Initialize the best historical position of the $i$-th particle $P_i$, $i = 1, \ldots, N$.
2. Global PSO search:
   - 2.1 Update velocities and positions from, and compute fitness values.
   - 2.2 Check if there are updates for $P_i$, $i = 1, \ldots, N$.
   - 2.3 If a stop criterion is reached, rank $P_i$, $i = 1, \ldots, N$; otherwise, go to 2.1.
   - 2.4 Output the top $N/5$ particles as candidate solutions (elite swarm).
3. K-means clustering for elite swarm:
   - 3.1 Determine the suitable value of $k_S$ according to the silhouette coefficient.
   - 3.2 Use K-means clustering to divide elite swarm into $k_S$ categories.
   - 3.3 From Eq(9), establish search boundaries for multiple local search spaces.
4. Local PSO search:
   - 4.1 Randomly generate $N/10$ new particles within each local space.
   - 4.2 Evaluate particle fitness values and initialize best positions in local spaces.
   - 4.3 Update velocities and positions, and compute fitness values.
   - 4.4 Check if there are updates for $\{P_g^{(r)} \mid r = 1, 2, \ldots, k_s\}$.
   - 4.5 If a stop criterion is reached, output multiple best positions achieved so far; otherwise, go to 4.2.
   - 4.6 Choose an optimal solution from $\{P_g^{(r)} \mid r = 1, 2, \ldots, k_s\}$.

End

---

*Figure 2: The detailed intelligent identification scheme*

## 4. Numerical simulation

In this section, simulation example is conducted to evaluate the proposed intelligent identification scheme for general Hammerstein-Wiener models. By setting MSE index of Eq(8) as the fitness function, the identification results are compared with the results obtained by the CPSO and NPSO intelligent search algorithms. Specially, for all the PSO searches, the parameters are set to: $c_1 = 1.5, c_2 = 1.5, w_{\min} = 0.7, w_{\max} = 1$.

In the example, a Hammerstein-Wiener model with non-smooth nonlinearities is considered. The linear block (Wills et al., 2013) is in the form of:

$$G(z^{-1}) = \frac{z^{-1} + 0.1z^{-2} - 0.49z^{-3} + 0.01z^{-4}}{1 + 0.3676z^{-1} + 0.88746z^{-2} + 0.52406z^{-3} + 0.55497z^{-4}}. \tag{12}$$

The true Hammerstein nonlinear block and the true Wiener nonlinear block are dead-zone function and saturation function respectively that can be expressed as:

$$f_{H\text{true}}(u_t) = \begin{cases} u_t + 0.9, & u_t < -0.9 \\ 0, & -0.9 \le u_t \le 0.9, \\ u_t - 0.9, & u_t > 0.9 \end{cases} \quad f_{W\text{true}}(x_t) = \begin{cases} 3, & x_t > 3 \\ x_t, & -3 \le x_t \le 3 \\ -3, & x_t < -3 \end{cases}$$

For the identification models $\{f_H(\cdot, \rho),\ f_W(\cdot, \gamma)\}$ of the above nonlinear blocks, the piecewise linear description shown in Eq(5) to Eq(7) is adopted. The description is composed of a linear base and two "hinge" functions. It is easy to see that there are 6 parameters in vector $\boldsymbol{\rho}$ as well as in vector $\boldsymbol{\gamma}$ . Thus, the vector $\boldsymbol{\theta} = \left[\boldsymbol{\rho}^{\mathrm{T}}, \boldsymbol{\alpha}^{\mathrm{T}}, \boldsymbol{\beta}^{\mathrm{T}}, \boldsymbol{\gamma}^{\mathrm{T}}\right]^{\mathrm{T}}$ contains twenty parameters that need to be identified. Besides, the input $u_t$ is an independent persistent excitation signal with zero-mean and variances $\sigma_u^2 = 2^2$ . The ambient additive noise $e_t$ follows the Student's $t$-distribution:

$$e_t \sim t(0, 0.3^2, 4), \tag{13}$$

where the location parameter is 0, the scale parameter is $0.3^2$, and the freedom degree is 4. From Eq(13), it is known that stochastic noise $e_t$ has properties of both zero-mean and heavy-tailed distribution.
For the above model structure, the proposed intelligent identification scheme is conducted by PSO search and K-means clustering. Specifically, the initial parameter space is set to [-2, 2]. The inertia weights for both global PSO and local PSO take the form of Eq(10). After several search tests with the population number $N = 600$ , it is known that a suitable iteration number of global PSO is 10 because the change of $MSE_g$ is starting to be unobvious after $k \geq 10$ .The varying process of $MSE_g$ is shown in Figure 3. Further, by observing different silhouette coefficients, the optimal category number $k_s$ of elite swarm is set to 6. After K-means clustering, the local PSO is implemented, and its iteration number is set to 500. Under these algorithmic settings, the global optimal solution can be obtained. The compared CPSO and NPSO search algorithms are conducted to harvest their results, and the relevant algorithmic settings remain the same as the settings of global PSO of proposed scheme.
For the identification purpose, the first collected input-output samples are used to identify the model, while the next samples are employed as a validation set to ensure the predictive capability. The MSE indicator Eq(8) was used to evaluate the accuracy of the estimated model. For NPSO, CPSO and the proposed scheme, we run these three identification schemes ten times. Table 1 shows the mean values and standard deviations of final fitness indexes. From Table 1, it is concluded that the proposed intelligent identification scheme can provide more accurate identification results than the other two mainstream search algorithms. For the predictive capability of proposed scheme, the validating and estimated outputs are compared in Figure 4.

*Table 1: The final fitness indexes with different data lengths*

| Data length L | 100 | 300 | 1000 | 2000 | 3000 |
|---|---|---|---|---|---|
| Final MSE for NPSO | 1.338±0.248 | 1.094±0.332 | 0.901±0.298 | 1.207±0.221 | 0.904±0.210 |
| Final MSE for CPSO | 1.030±0.361 | 1.009±0.289 | 0.867±0.229 | 1.396±0.301 | 0.814±0.193 |
| Final MSE for our scheme | 0.774±0.153 | 0.616±0.112 | 0.746±0.091 | 0.653±0.074 | 0.600±0.051 |

Besides, the population number $N$ has an impact on the accuracy of proposed scheme. In order to analyze this impact, for the different $N$, the global optimal fitness values under $L = 300$ are presented in Table 2. From Table 2, it is easy to see that the fitness value becomes better with the population number $N$ increasing.

*Table 2: The final fitness indexes with different population numbers s*

| Population number N | 300 | 400 | 500 | 600 | 700 |
|---|---|---|---|---|---|
| Index | 1.1176 | 0.9157 | 0.9338 | 0.7010 | 0.4954 |

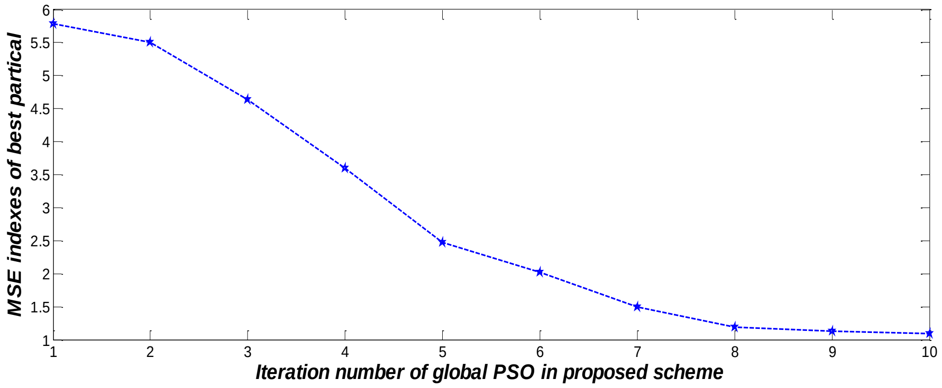


*Figure 3: The values of $MSE_g$ along time for global PSO in the proposed scheme*

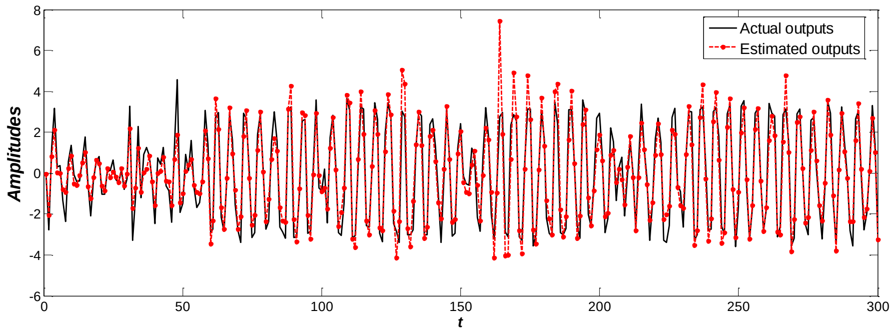


*Figure 4: The predicted output quality of proposed intelligent identification scheme*

## 5. Conclusions

A swarm intelligence-based identification scheme can make up for the weaknesses of classical PE and ML frameworks, and can give the global optimal parameter estimation from different perspectives. Especially for non-Gaussian noises, the accuracy of the parameter estimation is increased by 23 % to 53 % respectively at different data length relative to NPSO and CPSO. An increase in the length of the data and an increase in the population number will cause the MSE value to decrease. The best final MSE value is 0.4954. Thus, with careful design of algorithmic details, the comprehensive intelligent scheme achieves high identification accuracy and suppresses pulse disturbance. Besides, the future research work can extend this identification scheme into the identification of other nonlinear systems, e.g., Wiener systems, Hammerstein OEMA systems, Wiener-Hammerstein systems. By using its clustering and search of this identification scheme, the future research work will focus on the identification of large-scale nonlinear systems.


### Acknowledgments

This work is supported by National Natural Science Foundation of China (No. 61703434) and Science Foundation of China University of Petroleum, Beijing (No. 2462016YJRC028).


### References


Behl R., Sarría Í., González R., Magreñán Á.A., 2019, Highly efficient family of iterative methods for solving nonlinear models, Journal of Computational and Applied Mathematics, 346, 110-132.

Cao P.F., Luo X.L., 2014, Modeling for soft sensor systems and parameters updating online, Journal of Process Control, 24, 975-990.

Chen J.H., Yau H.T., Lu J.H., 2015, Chaos embedded particle swarm optimization algorithm-based solar optimal ReflexTM frequency charge, Journal of Applied Research and Technology, 13, 321-327.

Gotmare A., Patidar R., George N.V., 2015, Nonlinear system identification using a cuckoo search optimized adaptive Hammerstein model, Expert Systems with Applications, 42, 2538-2546.

Jia H., Deng L.H., 2018, Water flooding flowing area identification for oil reservoirs based on the method of streamline clustering artificial intelligence, Petroleum Exploration and Development, 45, 328-335.

Jin Q.B., Hao F., Wang Q., 2013, A multivariable IMC-PID method for non-square large time delay systems using NPSO algorithm, Journal of Process Control, 23, 649-663.

Maleki A., Ameri M., Keynia F., 2015, Scrutiny of multifarious particle swarm optimization for finding the optimal size of a PV/wind/battery hybrid system, Renewable Energy, 80, 552-563.

Nezamivand Chegini S., Bagheri A., Najafi F., 2018, PSOSCALF: A new hybrid PSO based on Sine Cosine Algorithm and Levy flight for solving optimization problems, Applied Soft Computing, 73, 697-726.

Rasouli M., Westwick D., Rosehart W., 2015, Quasiconvexity analysis of the Hammerstein model, Automatica, 50, 277-281.

Schoukens M., Tiels K., 2017, Identification of block-oriented nonlinear systems starting from linear approximations: A survey, Automatica, 85, 272-292.

Tang J.L., Wang D., Zhang Z.G., He L.J., Xin J., Xu Y., 2017, Weed identification based on K-means feature learning combined with convolutional neural network, Computers and Electronics in Agriculture, 135, 63-70.

Wills A., Schön T.B., Ljung L., Ninness B., 2013, Identification of Hammerstein–Wiener models, Automatica, 49, 70-81.

Wang Z., Wang Y., Ji Z., 2017, A novel two-stage estimation algorithm for nonlinear Hammerstein–Wiener systems from noisy input and output data, Journal of the Franklin Institute, 354, 1937-1944.