Hindawi
Complexity
Volume 2024, Article ID 7741473, 13 pages
https://doi.org/10.1155/2024/7741473

Research Article

Behaviour Analysis of Modeling and Model Evaluating Methods in System Identification for a Multiprocess Station

A. Annie Steffy Beula [1] **Geno Peter** [2] **Albert Alexander Stonier** [3] **K. Ezhil Vignesh** [1] **and Vivekananda Ganji** [4]

[1] *Electrical and Electronics Engineering, Stella Mary's College of Engineering, Ganapathipuram, India*
[2] *CRISD, School of Engineering and Technology, University of Technology Sarawak, Sibu, Sarawak, Malaysia*
[3] *School of Electrical Engineering, Vellore Institute of Technology, Vellore, India*
[4] *Department of Electrical and Computer Engineering, Debre Tabor University, Debre Tabor, Ethiopia*

Correspondence should be addressed to Vivekananda Ganji; vivekganji@dtu.edu.et

Received 14 February 2024; Revised 2 April 2024; Accepted 3 May 2024; Published 20 May 2024

Academic Editor: Basil M. Al-Hadithi

Systems are designed to perform specific task by giving certain input which produces the required output in an orderly manner known as process. The input, output, and the state variables should be known that will help in interacting with the system. The relation between these variables can be brought out by building a model that resembles or expresses the original performance of the system. The parameters of the model are estimated using the least squares approximation, maximum likelihood, maximum log-likelihood, and Bayesian parameter estimation methods by utilizing the experimental data from the multiprocess station. The selected parameters are converted to nine different transfer function models that represent the given dynamic system. The models framed are analyzed by the criterion curve technique using seven criterion functions evaluating the fitness of the model. Order of the model is found from Hankel matrix representation methods such as singular value decomposition and determinant method. Response of the models is compared with the original response to choose the best fit model by calculating ISE standard. All the above methods are used to model the system without physical and theoretical laws which is known as system identification.

1. Introduction

Our day-to-day lives depend on a lot of dynamic systems. In order to improve the quality and behaviour of the system, the stability, controllability, and observability of the system should be analyzed or even a controller design might be required. The knowledge about the input and the output variables of the system is required so as to make the design better. The system must be modeled, which means the input and the output of the system should be related in the form of a transfer function. Modeling is done with the help of experimental data [1] collected from the experiments conducted in the system without physical laws, known as system identification.

The prior knowledge about the process and the response of the system [2] helps to conduct the experiments [3] in the system. Hence, the input, output, and state variables can be framed.

The least squares method is used, which approximates the experimental data with least prediction error into useful parameters. It can be used hierarchically and iteratively [4] for linear and nonlinear transfer function models. The objective function of the least squares method is based on prediction error [5]. The parameter that minimizes the objective function [2, 6] will have better fitness [4] and is chosen to form the system model. Genetic algorithm [7] or any such search algorithms can also be used to minimize the objective function. The statistical parameter estimation methods are also similar to the least squares approximation method in minimizing the cost function [8] or the objective function. One of the statistical methods is maximum likelihood method [9] that chooses the parameters with the maximum possibility to occur and also with the least prediction error. The maximum log-likelihood method [1, 9] is the same as maximum likelihood in addition to the logarithmic term, which eliminates the complexity of large numbers. The Bayesian's estimation is a priori estimator predicting the posterior parameters.

The estimated parameters are used to form the transfer function models. The models are formed by using the input variables, output variables, delay variables, error variables, and errorless or undisturbed output variables under different combinations [10, 11]. The Autoregressive eXogenous model (ARX) is the combination of input and output parameters. The Autoregressive Moving Average model (ARMA) is the combination of output and error parameters. The Autoregressive Moving Average eXogenous model (ARMAX) is the combination of output, input, and error parameters. The Auto Regressive Auto Regressive eXogenous model (ARARX) is the combination of output, input, and delay parameters. The Autoregressive Autoregressive Moving Average eXogenous model (ARARMAX) is the combination of output, input, error, and delay parameters [12]. The Output Error model (OE) is the combination of input and undisturbed output parameters. The Box–Jenkins model (BJ) is the combination of input, undisturbed output, error, and delay parameters. The Autoregressive Integration Moving Average model (ARIMA) is the integrated ARMA model. The Autoregressive Integration Moving Average eXogenous model (ARIMAX) is the integrated ARMAX model.

The best model is selected on certain factors based on complexity, performance, and accuracy. To find the best model, certain assumptions [13] should be made which leads to the complexity of the model. Therefore, the models are selected by using criterion functions [14] which depend upon the number of datasets, number of parameters in the model, loss function, and penalty functions. The criterion-based model selection overcomes [15] the drawback of prespecified order or any order assumption for the system. The fitness of the model is checked by comparing the predicted output and mean value of the output with the original output data. The fitness can also be checked by comparing the response [16] of the models with the original response of the system or by changing the signal to noise ratio [17] value each time. The best fit model is selected and subjected to various kinds of disturbances [18] for analyzing. The models are also given for analog to digital conversion [19] tests to check its reconstructability. The Levenberg–Marquardt algorithm [20] is used to check the fitness of the models after evaluating them using various parameter estimation methods.

Criteria are used to find the order of the system. The order of the system is found [21] instead of fixing the order limits that avoids the overestimation of orders. The model with least criterion value [22] is found from the pool of models with various orders. The selected model will have the order equal or close to the true order of the system. The order of the system can also be found [23] by the Hankel matrix representation. The Hankel matrix is the state-space model of the selected model [24]. The singular value decomposition [25] of the state model shows how far the experimental data are approximated into the parameters [26]. The determinant method is an order estimation method using Hankel matrix which eliminates the assumption of orders. Once the order is estimated, the model is reduced to the true order of the system by the use of order reduction technique. Thus, the model for the dynamic system could be evaluated from the experimental data and it is known as system identification.

2. System Identification

Modeling the given process station is the aim of identification and the model should resemble the true performance of the system. The order of the model should be the same order when found using the physical laws. The precision during the manufacture may vary due to ageing or by rough usage of instrument from time to time. Therefore, it is better to use the experimental data rather than using the theoretical equation in evaluating the model. Also, the controller could be designed for the current state of the system. System identification is creating a mathematical model for the dynamic systems from the experimental dataset. Figure 1 shows how to identify a system.

The concept for the input, output, and disturbance variables, as well as the method for measuring the system's variables, is determined. Based on this knowledge, the experimental setup is arranged for the collection of data.

The collected data are converted into useful parameters as the input to the model sets. The best model has to be selected from the model sets using the criterion functions. The order of the selected model is estimated. The model with best fitness is chosen to be the final model for the system.

3. Multiprocess Station

The multiprocess station [2] with tank and heater system is chosen for this work. The flow, level, and temperature of the process are analyzed. The water is collected in the reservoir tank of the process station. This water is pumped to the overhead tank of the process station, from which the water is supplied to the level tank and the heater system separately through a rotameter to control the flow rate. Figure 2 gives the hardware setup for the process station.

There is a front panel flow diagram to give connections for the required process. There are two separate switches for the power and heater on/off. The pump speed can be varied to vary the flow rate. There are two differential pressure transmitters or transducers where the pressure and the flow rate relation are converted to level in the level tank. There are two temperature transducers to measure the temperature of the heater. There are two input channels to give the analog to digitally converted input for data acquisition and an output channel which gets controller output. This controller output can be changed to manual mode in order to get an output without controller. There are indicators to show the output current, output pressure, air regulator output, and the output of current to pressure converter.

The pumping motor is turned on to fill the water in the tank. The water flows through the rotameter to measure the

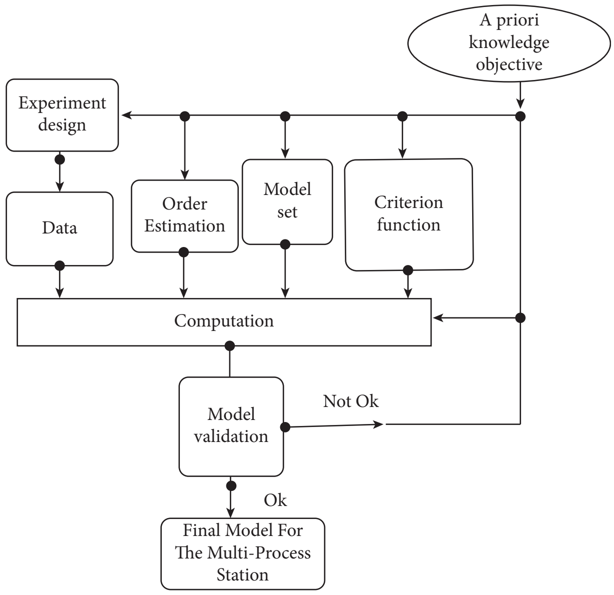

Figure 1: Block diagram of system identification.

flow rate. The flow rate can be adjusted by opening and closing the valve. The change in water level along with the change in flow rate can be observed from the dataset. Similarly, while working on with the heater system, the flow rate, time, and temperature change can be recorded. Data acquisition card or digital controller is inbuilt with analog to digital and digital to analog converters that link the process and the controller actions. The specifications of the components used in the multiprocess station are listed in Table 1.

4. Parameter Estimation

The experimental data have to be converted into useful parameters by the least squares approximation method which finds the parameters with least standard deviation. This method estimates a set of parameters depending upon the order assumed for the model. The input and the output data are arranged in a matrix with the order as the model order and the number of datasets collected. Therefore, the estimated parameters are completely dependent on the input and the output collected from the experimental setup. The input and the output are first represented as the linear difference equation as

$$\begin{aligned} y(t) + a_1 y(t-1) +, \ldots, + a_{n_a} y(t-n_a) \\ = b_1 u(t-1) +, \ldots, + b_{n_b} u(t-n_b), \end{aligned} \tag{1}$$

$\varphi(k) = \begin{bmatrix} y(k) \\ u(k) \end{bmatrix}$, where $k = 1, 2, \ldots, l$ is the length of the experimental data collected.

Figure 2: Hardware of the multiprocess station.

$$\theta = \left[b_0 \left(-a_1 \;\; b_1 \right) \cdots \left(-a_p \;\; b_p \right) \right]. \tag{2}$$

The parameters are estimated for various orders, and the parameter that gives the least prediction error or the parameter that minimizes the cost function, which is the function of prediction error, is chosen. The error could be predicted from the estimated output and the measured output values as in

$$e = y - \widehat{y}, \tag{3}$$

$$J = \frac{1}{N} \sum_{i=1}^{N} \left[e \;\; e^T \right]. \tag{4}$$

The objective function J (4) is formed from the error. It is also known as the cost function where N is the number of datasets considered.

There are a few statistical methods which are used to estimate the parameters for the models. The maximum likelihood method estimates the parameters that are more likely suitable for the given model that is with the least prediction error. As said prior, these methods depend on the cost function of the least squares method. The likelihood function method depends upon the type of distribution considered. The distribution considered is the normal distribution. The likelihood function is the probability density function for maximum likelihood; the maximum log-likelihood and Bayesian's estimation methods are represented in (8), (9), and (11), respectively.

Table 1: Specification of components used in process station.

S. no	Name of the component	Parameter	Specification
1	Rotameter	Range	(50–100) lph
2	Heater	Power	3 kW
3	Pump	RPM Voltage Discharge	6500 RPM 230 V AC/DC, 50 Hz 800 lph
4	Reservoir	Capacity	15 liters
5	RTD sensor	Type Length Tube material	Pt-100 60 mm SS 316
6	RTD transmitter	Temperature range Supply Output	(0–100)°C 24 V DC (4–20) mA DC

$$f(\theta, y_1, y_2, \cdots, y_N) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y_i-\theta)^2}{2\sigma^2}\right], \tag{5}$$

$$\log f(\theta, y_1, \cdots, y_N) = \log\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y_i-\theta)^2}{2\sigma^2}\right], \tag{6}$$

$$P(\mu|X) = \frac{P(X|\mu)P(\mu)}{P(X)} = \text{likelihood} \times \text{prior}. \tag{7}$$

The general form or the standard form of the estimation is $y = X\beta + E$, where X is the vector of probability density function of data with known standard deviation and mean. E is the vector of probability density function of data with zero mean and one standard deviation.

5. Transfer Function Models

The transfer function models are used to relate the parameters estimated from the least squares, maximum likelihood, and Bayesian parameter estimation methods as listed in Table 2. The transfer function models are framed using the five different parameter sets. They are A, B, C, D, and F representing the output, input, error, delay, and disturbance as vectors.

$$A(q) = \sum_{k=0}^{n_a} a_k q^{-k} = a_0 + a_1 q^{-1} +, \ldots, + a_{n_a} q^{-n_a}, \quad \text{with } \mathrm{a}_0 = 1, \tag{8}$$

$$B(q) = \sum_{k=0}^{n_b} b_k q^{-k} = +b_1 q^{-1} +, \ldots, + b_{n_b} q^{-n_b}, \quad \text{with } \mathrm{b}_0 = 0, \tag{9}$$

$$C(q) = \sum_{k=0}^{n_c} c_k q^{-k} = c_0 + c_1 q^{-1} +, \ldots, + c_{n_c} q^{-n_c}, \quad \text{with } \mathrm{c}_0 = 1, \tag{10}$$

$$D(q) = \sum_{k=0}^{n_d} d_k q^{-k} = d_0 + d_1 q^{-1} +, \ldots, + d_{n_d} q^{-n_d}, \quad \text{with } \mathrm{d}_0 = 1, \tag{11}$$

$$F(q) = \sum_{k=0}^{n_f} f_k q^{-k} = f_0 + f q^{-1} +, \ldots, + f_{n_f} q^{-n_f}, \quad \text{with } \mathrm{f}_0 = 1, \tag{12}$$

Table 2: Simplified combinations of transfer function models.

S. no	Transfer function model	Parameters of the model
1	ARX	A, B
2	ARMA	A, C (experimental output)
3	ARIMA	A, C (integration of exp. output)
4	ARMAX	A, B, C (exp. output)
5	ARIMAX	A, B, C (integration of exp. output)
6	ARARX	A, B, D
7	ARARMAX	A, B, C, D
8	Output error (OE)	B, F
9	Box–Jenkins (BJ)	B, F, C, D

where $A(q)$, $B(q)$, $C(q)$, $D(q)$, and $F(q)$ are the coefficients of the output, input, error, delay, and disturbance coefficients. The general representation of the OE model is represented as follows: OE model gives the relation between the input and the undisturbed output. The undisturbed output can be expressed as follows:

$$\hat{y} = \phi^T \theta, \tag{13}$$

$$\phi^T(t) = \left[u(t), \ldots, u(t-n_a)\hat{y}(t), \ldots, \hat{y}\left(t-n_f\right)\right]^T, \tag{14}$$

$$\theta = \left[b_1, \ldots, b_{na} f_1, \ldots, f_{nf}\right]^T. \tag{15}$$

The differential equation can be written as

$$\begin{aligned} &\hat{y}(t) + f_1\hat{y}(t-1) +, \ldots, + f_{nf}\hat{y}\left(t-n_f\right) \\ &\quad = b_1 u(t-1) + b_{nb} u\left(t-n_b\right) + f_0 e(t) +, \ldots, f_1 e(t-1) \\ &\qquad +, \ldots, + f_{nf} e\left(t-n_c\right). \end{aligned} \tag{16}$$

The above equation can be written as

$$y(t) = \left(\frac{B(q)}{F(q)}\right)u(t) + e(t). \tag{17}$$

The transfer functions for the OE model obtained from the MATLAB programming were run on a Dell Inspiron I5 processor and are given by

$$\begin{aligned} \left(\frac{B(q)}{F(q)}\right)u(t) &= \frac{0.01765s}{2.128e^{-005}s^2 + 2.625e^{-007}s + 1}u(t), \\ y(t) &= \frac{0.01765s}{2.128e^{-005}s^2 + 2.625e^{-007}s + 1}u(t) + e(t). \end{aligned} \tag{18}$$

6. Criterion Selection

The models are selected on the basis of number of datasets considered, number of parameters in the model, amount of estimation error in the model parameters, and the penalty which is directly proportional to the error. Akaike Information Criterion (AIC) has penalty higher than the Kullback–Leibler Information Criterion (KIC). Therefore, AIC produces least values when compared with the KIC. Bayesian Information Criterion (BIC) produces more penalty for the lower order systems, and as the order increases, the penalty decreases when compared with the AIC and KIC. Hannan–Quinn Information Criterion (HQC) gives fine-tuning of the data, that is, tells how far it approximates the large amount of data into a very few parameters. Minimum Description Length criterion (MDL) probabilistically predicts the parameters by compressing the long length of unseen data. Mallows' Cp Criterion (MCp) depends on the number of parameters in the model, sum of squares of errors, and the variance in the prediction of the parameters. Akaike's Final Prediction Error Criterion (AFPEC) depends on the same likelihood or loss functions as in AIC, BIC, and KIC, but the penalty increases as the ratio of the number of datasets taken into account to the total amount of data measured increases. The criterion values for statistical estimation for both the level and the temperature process are analyzed in Tables 3 and 4. It is found that the OE model has the best criterion value. Therefore, the selected model is the OE model.

7. Order Estimation from Criterion Curve

The order of the system could only be estimated using the knowledge model structure, and the prediction error of the estimated parameters could be estimated using the criterion functions. The AIC, BIC, and KIC criteria are used to estimate the order of the system. Figures 3 and 4 show the estimated order for the level and the temperature process.

Figures 4 and 4 represent the results of estimating the order of the system using the criterion curve method. Here AIC, KIC, and BIC are the criterion functions chosen. These criteria are objective functions framed using number of samples considered, error, standard deviation, and order of the system. The order which reduces the objective function is 2 and 1 for level and temperature process, respectively, and so it is chosen as the order of the system. This can be seen from Figures 3 and 4.

The order of the level process is fixed to be 2. Since, the least value of all the three criteria in Figure 3 is at order 2, while the order for the temperature process is fixed to be 1, the least value of all the three criteria in Figure 4 is at order 1.

Table 3: Criterion values for level process with 2^{nd} order transfer function models for statistical estimation.

2^{nd} order	Estimation method	ARX	ARMA	ARMAX	ARARX	ARARMAX	OE	BJ	ARIMA	ARIMAX
AIC	MLE	14.8101	14.8379	14.8934	14.8934	14.9768	8.3593	8.526	14.959	15.514
	Log MLE	14.8100	14.8378	14.8933	14.8933	14.9767	8.3516	8.5282	13.127	13.682
	Bayesian	14.8099	14.8377	14.8933	14.8933	14.9766	8.3608	8.5275	15.96	16.516
BIC	MLE	14.7193	14.729	14.7482	14.7482	14.7771	8.2686	8.3263	13.87	14.062
	Log MLE	14.7193	14.7289	14.7481	14.7481	14.777	8.2608	8.3286	12.038	12.23
	Bayesian	14.7192	14.7288	14.7481	14.7481	14.777	8.2701	8.3278	14.871	15.064
KIC	MLE	14.8795	14.9112	15.0045	15.0045	15.1295	8.4288	8.6788	15.792	16.625
	Log MLE	14.8795	14.9211	15.0045	15.0045	15.1295	8.421	8.681	13.96	14.793
	Bayesian	14.8794	14.9211	15.0044	15.0044	15.1294	8.4303	8.6803	16.794	17.627
HQC	MLE	10.5213	10.4907	10.4704	10.4704	10.4398	4.3647	4.4462	3.2872	3.2872
	Log MLE	10.5219	10.4913	10.471	10.471	10.4404	4.3646	4.4591	6.2916	6.2916
	Bayesian	10.5212	10.4907	10.4703	10.4703	10.4398	4.365	4.4455	4.5064	4.5064
AFPEC	MLE	−2.48*e*003	−2.48*e*003	−2.48*e*003	−2.48*e*003	−2.48*e*003	−3.928*e*003	−3.928*e*003	−3.9938	−3.9938
	Log MLE	−2.48*e*003	−2.48*e*003	−2.48*e*003	−2.48*e*003	−2.48*e*003	−3.937*e*003	−3.937*e*003	−3.3253	−3.3253
	Bayesian	−2.48*e*003	−2.48*e*003	−2.48*e*003	−2.48*e*003	−2.48*e*003	−3.934*e*003	−3.934*e*003	−4.4147	−4.4147
MCp	MLE	−66	−60.0002	−56.0002	−56.0002	−50.0002	−66.005	−50.005	−59.895	−55.8945
	Log MLE	−66	−60.0002	−56.0002	−56.0002	−50.0002	−66.0051	−50.0051	−59.972	−55.972
	Bayesian	−66	−60.0002	−56.0002	−56.0002	−50.0002	−66.005	−50.005	−59.948	−55.9481
MDL	MLE	6.11*e*007	6.11*e*007	6.11*e*007	6.11*e*007	6.11*e*007	96.445	9.6445*e*004	98.0513	98.0513
	Log MLE	6.11*e*007	6.11*e*007	6.11*e*007	6.11*e*007	6.11*e*007	96.443	9.6663*e*004	81.6374	81.6374
	Bayesian	6.11*e*007	6.11*e*007	6.11*e*007	6.11*e*007	6.11*e*007	96.591	9.6591*e*004	108.3828	108.3828

Table 4: Criterion values for temperature process with 1st order transfer function models OE, BJ, ARIMA, and ARIMAX for statistical estimation.

1st order	Estimation method	ARX	ARMA	ARMAX	ARARX	ARARMAX	OE	BJ	ARIMA	ARIMAX
AIC	MLE	8.8703	8.8981	8.9258	8.9258	8.9814	7.1228	9.2339	10.2344	10.2622
	Log MLE	8.9462	8.9740	9.0018	9.0018	9.0573	7.1664	9.2775	7.891	7.9188
	Bayesian	8.9462	8.9740	9.0018	9.0018	9.0573	7.1664	9.2775	7.9173	7.9451
BIC	MLE	8.7869	8.7869	8.7869	8.7869	8.7873	7.0386	9.0394	10.1233	10.1267
	Log MLE	8.8629	8.8629	8.8629	8.8629	8.8633	7.0809	9.0831	7.7799	7.7803
	Bayesian	8.8629	8.8629	8.8629	8.8629	8.8633	7.0809	9.0831	7.8062	7.8071
KIC	MLE	8.9119	8.9536	8.9953	8.9953	9.0786	7.1644	9.3311	10.29	10.3314
	Log MLE	8.9879	9.0295	9.0712	9.0712	9.1545	7.1081	9.3747	7.9465	7.9882
	Bayesian	8.9879	9.0295	9.0712	9.0712	9.1545	7.2081	9.3748	7.9728	8.0145
HQC	MLE	10.5213	10.4907	10.4704	10.4704	10.4398	7.3647	4.4462	3.2872	3.2872
	Log MLE	10.5219	10.4913	10.471	10.471	10.4404	4.3636	4.4591	6.2916	6.2916
	Bayesian	10.5212	10.4907	10.4703	10.4703	10.4398	4.364	4.4455	4.5064	4.5064
AFPEC	MLE	−6.73*e*003	−6.73*e*003	−6.73*e*003	−6.73*e*003	−6.73*e*003	−8.7*e*003	−8.7*e*003	−2.6*e*003	−2.6*e*003
	Log MLE	−7.26*e*003	−7.26*e*003	−7.26*e*003	−7.26*e*003	−7.26*e*003	−9.1*e*003	−9.1*e*003	−2.46*e*003	−2.46*e*003
	Bayesian	−7.26*e*003	−7.26*e*003	−7.26*e*003	−7.26*e*003	−7.26*e*003	−9.1*e*003	−9.1*e*003	−2.52*e*003	−2.52*e*003
MCp	MLE	−69.9976	−63.9976	−61.9976	−61.9976	−57.9976	−69.9996	−57.9963	−63.9993	−61.9993
	Log MLE	−69.9978	−63.9978	−61.9978	−61.9978	−57.9978	−69.9994	−57.9964	−63.9865	−61.9865
	Bayesian	−69.9979	−63.9979	−61.9979	−61.9979	−57.9979	−69.9997	−57.9971	−63.9869	−61.9869
MDL	MLE	6.548*e*003	6.548*e*003	6.548*e*003	6.548*e*003	6.548*e*003	8.43*e*002	8.43*e*002	2.5*e*004	2.5*e*004
	Log MLE	7.064*e*003	7.064*e*003	7.064*e*003	7.064*e*003	7.064*e*003	8.81*e*002	8.81*e*002	2.4*e*003	2.4*e*003
	Bayesian	7.064*e*003	7.064*e*003	7.064*e*003	7.064*e*003	7.064*e*003	8.81*e*003	8.81*e*003	2.46*e*003	2.46*e*003

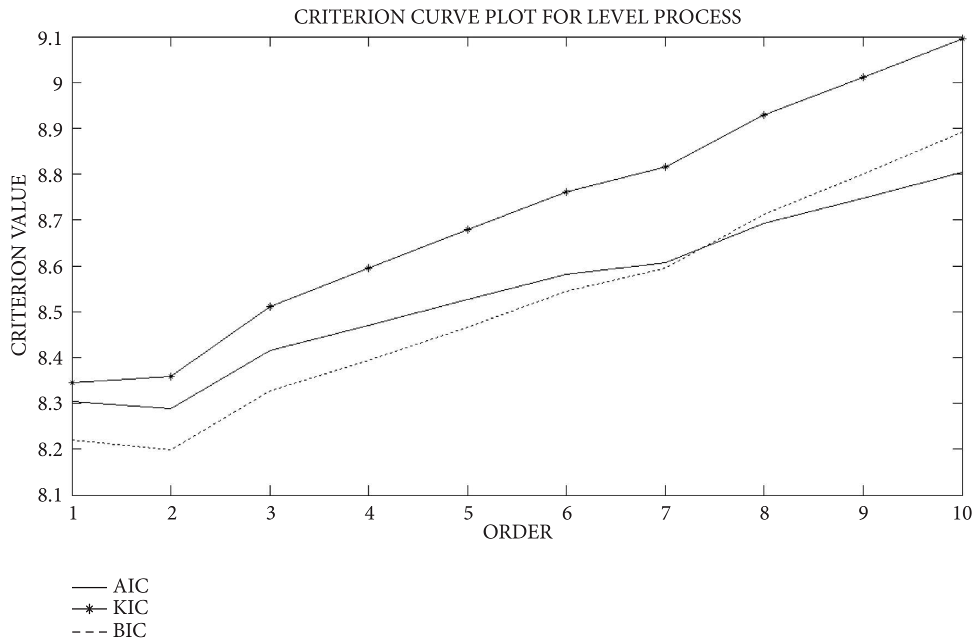

Figure 3: Criterion curve for level process.

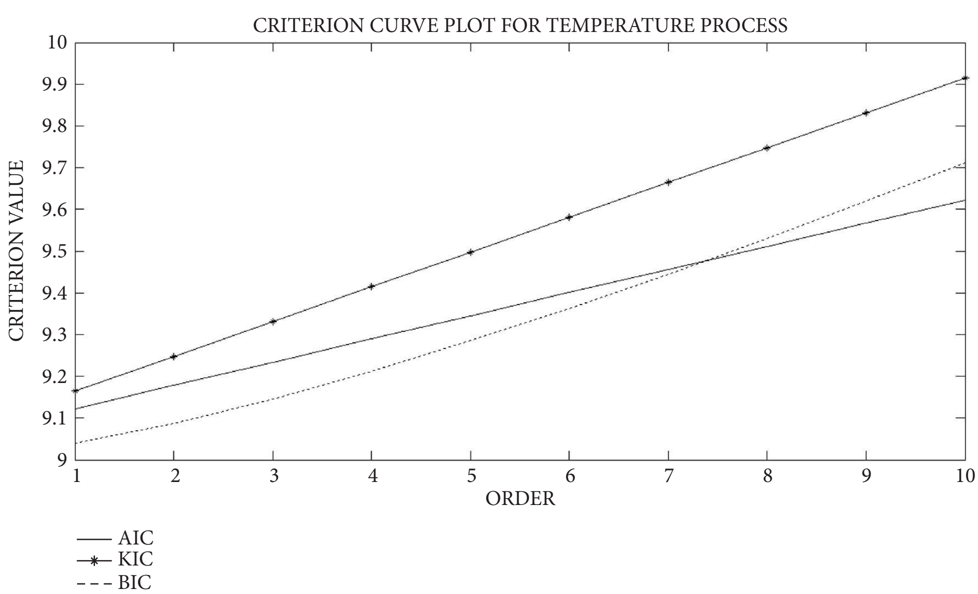

Figure 4: Criterion curve for temperature process.

8. Model Fitness

The fitness of the models is seen by comparing the response of the models formed with response of the true system from the experimental data. Each transfer function model formed from the parameters of the different estimation methods is compared with the original response of the system for both level and temperature processes. Finally, the OE model from Maximum Log-Likelihood estimation for both level and temperature process lies more accurately on the original response of the system shown in Figures 5 and 6.

9. Performance Parameters of Final OE Model

The performance of the selected OE model from the four different types of estimation is analyzed after giving a disturbance. The Integral of Squared Errors for the models is evaluated and listed in Table 4. It is seen that the model from the least squares method gives a better response and the least ISE value. So, the OE model estimated from the least squares method is chosen to be the final model for the multiprocess station. The responses of the level and temperature models after the disturbance are given in Figures 7 and 8, respectively.

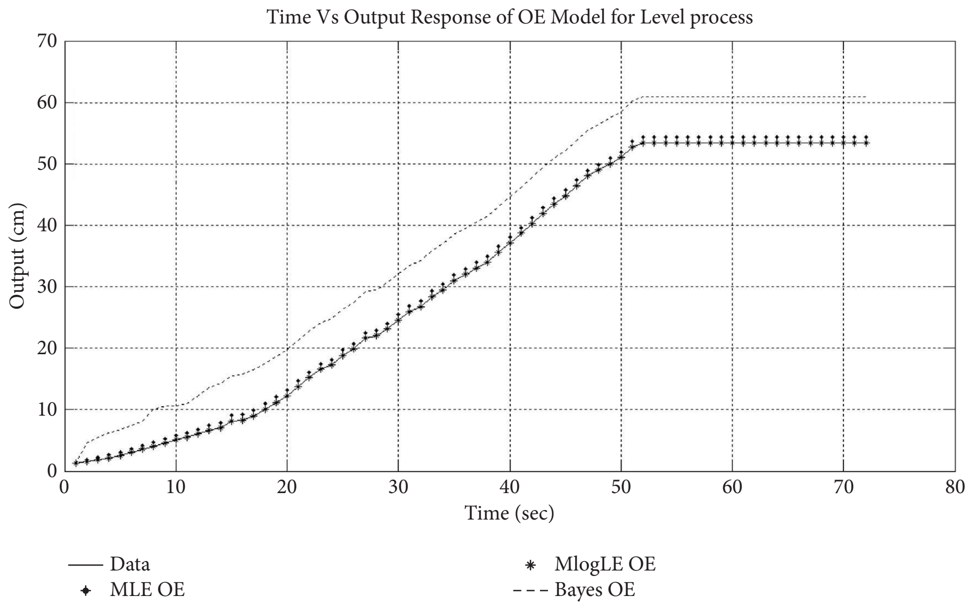

Figure 5: Output response of OE model for level process.

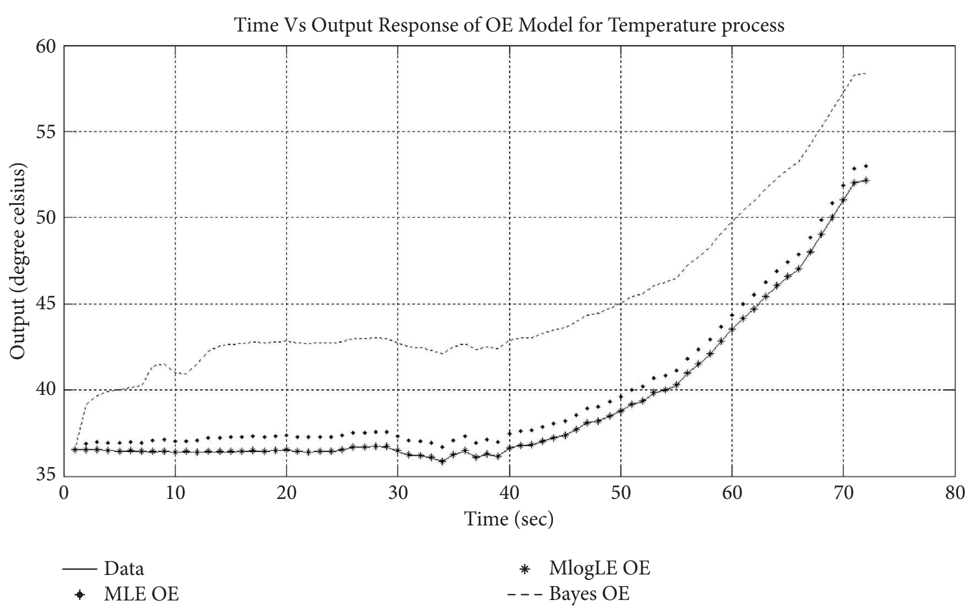

Figure 6: Output response of OE model for temperature process.

The transfer function for the level process is

$$G(s) = \frac{0.02255s^2 - 0.01418s + 0.01013}{s^2 + 0.41s + 0.009741}. \quad (19)$$

The transfer function for the temperature process is

$$G(s) = \frac{0.9257s + 12.99}{s + 21.35}. \quad (20)$$

The order of the system can also be found using the Hankel matrix representation which has been formed from the state-space model of the finalized model without any order fixation. The rank of the decomposed matrix in the singular value decomposition matrix will be equal to the true order of the system. The ratio of the eigenvalues reveals the order of the system [12]. The order of the least determinant value predicted for the possible least and maximum order of the Hankel matrix gives the order of the system.

Hankel matrix is formed using the state-space model of the selected OE transfer function. The controllability and observability matrix is determined from the state-space model. The product of observability and controllability matrix is the Hankel matrix.

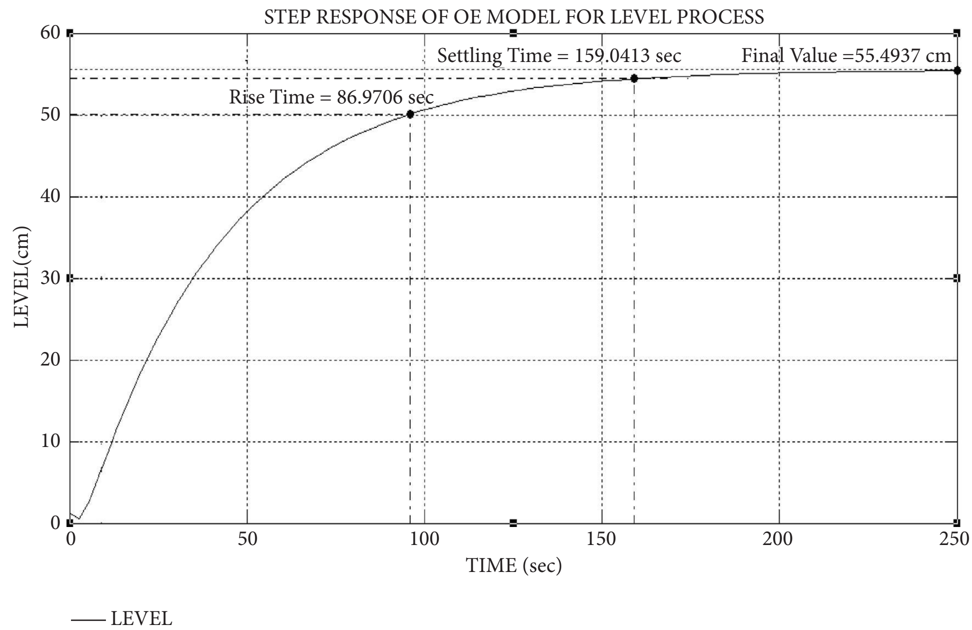

Figure 7: Step response of the OE model for level process.

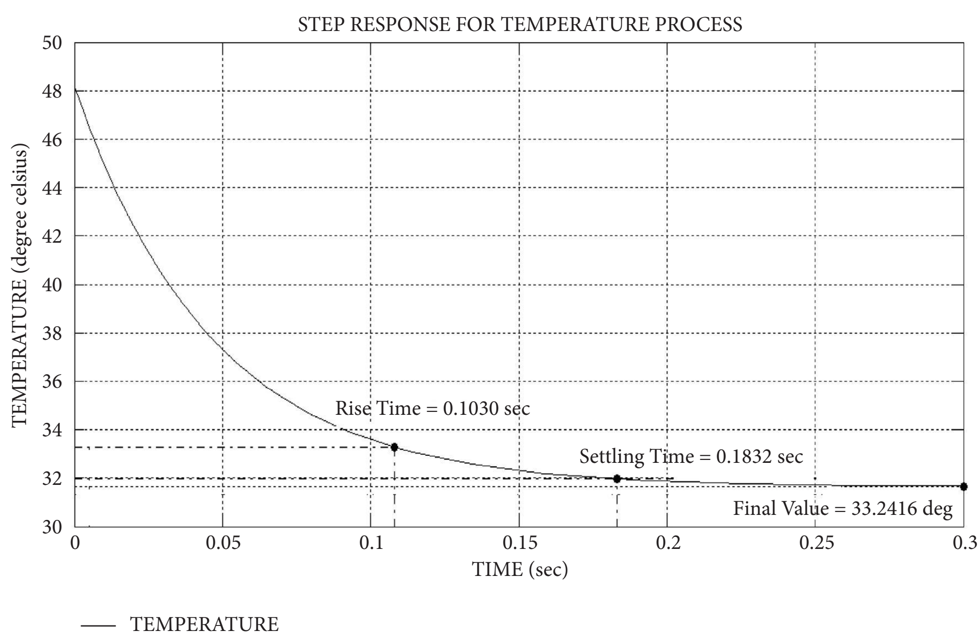

Figure 8: Step response of the OE model for temperature process.

$$H = \begin{bmatrix} 0.0446 & 1.412 & -1.9256 & 1.542 \\ 1.4122 & -1.9258 & 1.5424 & -0.6505 \\ -1.926 & 1.5426 & -0.6504 & -0.2635 \\ 1.5424 & -0.6501 & -0.2636 & 0.8316 \end{bmatrix}. \quad (21)$$

Then the determinants are found for the Hankel matrix. The order for which the determinant value is minimum or close to one is chosen. And the order less than one of the corresponding order is taken as the order of the system. Thus, the given multiprocess station is modeled using the system identification technique.

10. Results and Discussion

The experiment was conducted in the multiprocess station and the datasets for level and temperature process are collected using the data acquisition system. The datasets are transformed into parameter for the transfer function model by the approximation methods. The parameters are utilized in the transfer function models and the models are validated using the criterion function. The OE model has no error or disturbance coefficients in its structure while analyzing Tables 3 and 4, which is guessed to be suitable for the given multiprocess station. The order of the system must be fixed

Table 5: Performance analysis of the selected models.

S. no	Estimation type	Process	Rise time (sec)	Settling time (sec)	Magnitude of settling	ISE value
1	Least squares approximation	Level	86.9706	159.0413	55.4937	4.51010
		Temperature	0.1030	0.1832	33.2416	351.8775
2	Statistical methods	Level	34.5701	78.1639	42.3107	122.3081
		Temperature	2.67	7.49	23.42	824.8384

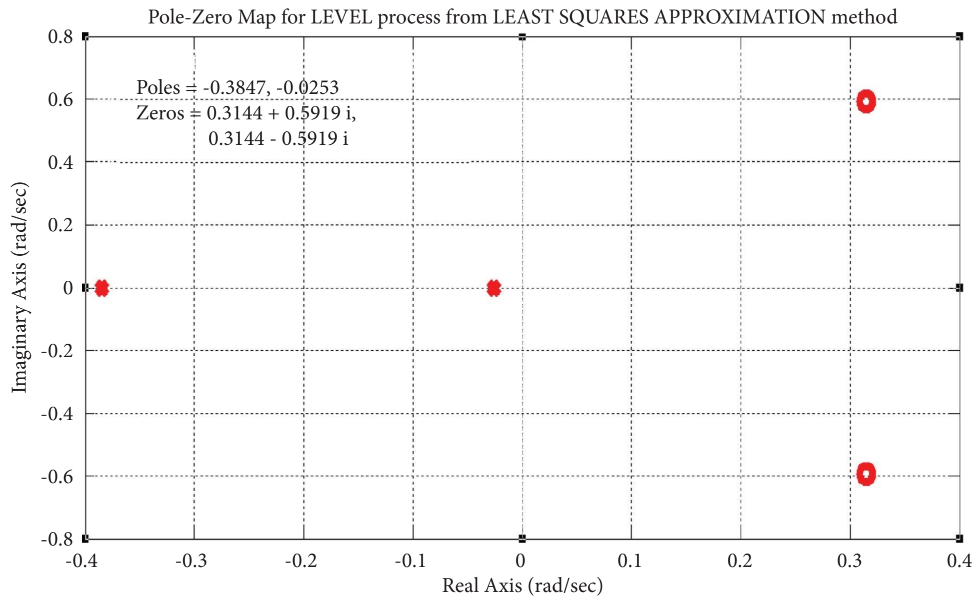

Figure 9: Pole zero location of level process.

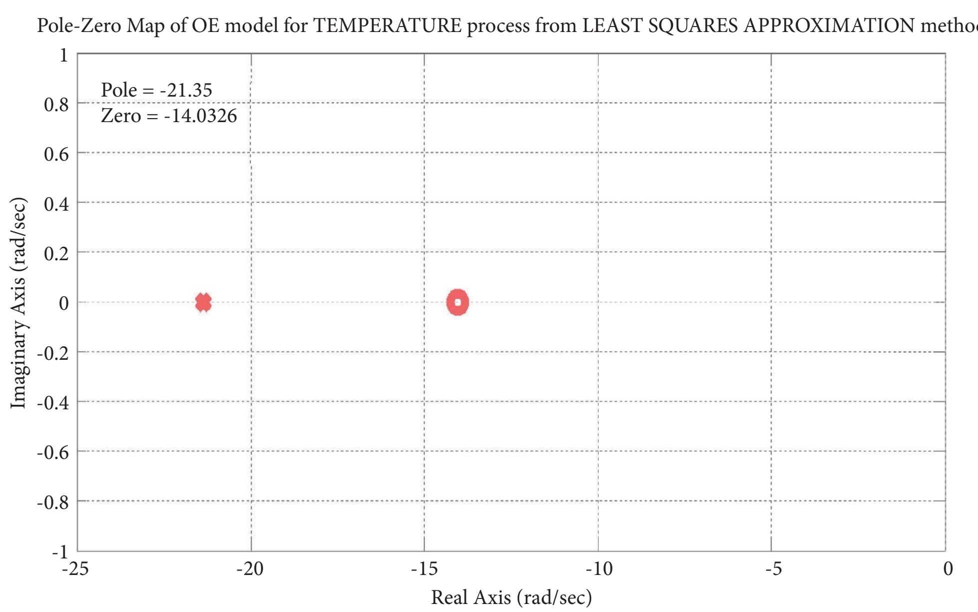

Figure 10: Pole zero location of temperature process.

since the order of the validated models is high. The criterion curve method is been adopted to estimate the order of the system. In analyzing Figures 3 and 4, the order for level and temperature process can be fixed to 2 and 1, respectively, as it matches the original experimental data curve. The OE model for both level and temperature process whose parameters are estimated from the Maximum Log-Likelihood method seems to be more accurate with the original response of the system while observing Figures 5 and 6. And so the corresponding transfer function is selected. The step response of the OE model for three parameter estimation methods is taken and it is found that the OE model using Least Squares Approximation method has better response and ISE value as shown in Table 5. Therefore, the model is finalized for the chosen multiprocess station. The stability of the model is then analyzed. The transfer function of the temperature process has all the poles and zeros on the left half of S-Plane and so it is stable as shown in Figures 9 and 10. The transfer function of the level process has all the poles on the left half of S-Plane and zeros are complex conjugate in the nonminimum phase.

11. Conclusion

The experimental data are collected using prior knowledge about the multiprocess station and the variables required for modeling the given tank and heater system. The data collected are converted into parameters for system modeling using the least squares approximation method, maximum likelihood method, maximum log-likelihood method, and Bayesian estimation method. The parameters estimated are applied to the transfer function models such as ARX, ARMA, ARMAX, ARARX, ARARMAX, OE, BJ, ARIMA, and ARIMAX models. The models are analyzed for the selection using the criterion functions such as AIC which gave the least value of 8.3516 where the parameters are predicted using the log-likelihood method. KIC gave the least value of 8.421 where the parameters are predicted using the log-likelihood method. Similarly, all other criteria such as BIC, HQC, AFPEC, MCp, and MDL gave least values for the parameters predicted using the log-likelihood method. The least values from various criteria are analyzed for various models. The OE model has the least error of −66.0051. The selected model is checked for its fitness by examining its response with the experimental data response. The model that is more accurate to the original data is formed from the parameters estimated using the maximum log-likelihood method. The true order of the model is evaluated using the criterion curve. The models are further examined with step response. The data approximated using least squares have least ISE standard value of 4.5101 for level process and 351.8775 for temperature process. Finally it is found that the OE model from the least squares algorithm has better performance. The general representation of the OE model does not include the disturbance parameters. The experimental data are collected by neglecting the disturbance parameters. Thus, the multiprocess station is modeled using the system identification method. The work will be extended in future by implementing the technique in real-time applications like boilers in plants.

Data Availability

The required data can be obtained from the corresponding author upon an e-mail request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

A. Annie Steffy Beula and Albert Alexander Stonier were responsible for conceptualization, methodology, and software. Geno Peter and K. Ezhil Vignesh were responsible for data curation and original draft preparation. Vivekananda Ganji was responsible for visualization and investigation.

References

[1] L. Ljung, *System Identification Theory for the User*, Linkoping University, Sweden, Europe, 1991.

[2] A. Rahrooh and S. Shepard, "Identification of nonlinear systems using NARMAX model," *Nonlinear Analysis: Theory, Methods and Applications*, vol. 71, no. 12, pp. 1198–1202, 2009.

[3] User Manual, *Multiprocess Station, VMPA-62*, Vi Microsystems Pvt. Ltd, Perungudi, India.

[4] M. Jafari, M. Salimifard, and M. Dehghani, "Identification of multivariable nonlinear systems in the presence of colored noises using iterative hierarchical least squares algorithm," *ISA Transactions*, vol. 53, no. 4, pp. 1243–1252, 2014.

[5] C. R. Rojas, R. Toth, and H. Hjalmarsson, "Sparse estimation of polynomial and rational dynamical models," *IEEE Transactions on Automatic Control*, vol. 59, no. 11, pp. 2962–2977, 2014.

[6] C. Wiegand, C. Hedayat, W. John, L. Radic-Weissenfeld, and U. Hilleringmann, "Nonlinear identification of complex systems using radial basis function networks and model order reduction," *2007 IEEE International Symposium on Electromagnetic Compatibility*, 2007.

[7] H.-Q. Cao, Li-S. Kang, T. Guo, Yu-P. Chen, and H. de Garis, "A two-level hybrid evolutionary algorithm for modeling one-dimensional dynamic systems by higher- order ODE models," *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, vol. 30, no. 2, pp. 351–357, 2000.

[8] W. Wang, F. Ding, and J. Dai, "Maximum likelihood least squares identification for systems with autoregressive moving average noise," *Applied Mathematical Modelling*, vol. 36, no. 5, pp. 1842–1853, 2012.

[9] V. Vandewalle, C. Biernacki, G. Celeux, and G. Govaert, "A predictive deviance criterion for selecting a generative model in semi-supervised classification," *Computational Statistics and Data Analysis*, vol. 64, pp. 220–236, 2013.

[10] Y. Zou and W. P. Heath, "The quantification of large SNR for MLE of ARARMAX models," *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, 2009.

[11] Y. Naung, A. Schagin, H. L. Oo, K. Z. Ye, and Z. M. Khaing, "Implementation of data driven control system of DC motor by using system identification process," in *Proceedings of the 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, pp. 1801–1804, Moscow and St. Petersburg, Russia, January 2018.

[12] X. Ping, X. Luan, S. Zhao, F. Ding, and F. Liu, "Parameters-transfer identification for dynamic systems and recursive form," *IEEE Signal Processing Letters*, vol. 30, pp. 1302–1306, 2023.

[13] E. Walter and L. Pronzato, *Identification Parametric from Experimental Data*, Springer, Great Britain, UK, 1997.

[14] G. Vallejo, E. Tuero-Herrero, J. C. Nunez, and P. Rosario, "Performance evaluation of recent information criteria for selecting multilevel models in Behavioral and Social Sciences," *International Journal of Clinical and Health Psychology*, vol. 14, no. 1, pp. 48–57, 2014.

[15] C.-M. Ting, S. B. Samdin, Sh-H. Salleh, M. H. Omar, and I. Kamarulafizam, "An expectation-maximization algorithm based kalman smoother approach for single-trial estimation of event-related potentials," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2012, pp. 6534–6538, 2012.

[16] G. Sparacino, C. Tombolato, and C. Cobelli, "Maximum-likelihood versus maximum-A-posteriori parameter estimation of physiological system models: the C- peptide impulse response case study," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 6, pp. 801–811, 2000.

[17] V. A. O. Alves, R. Juliani Correa de Godoy, and C. Garcia, "Searching theOptimal order for high order models- MIMO case," *2012 IEEE International Conference on Control Applications*, 2012.

[18] R. Pourramazan, S. Vaez-Zadeh, and H. Nourzadeh, "Power system MIMO identification for coordinated design of PSS and TCSC controller," *2007 IEEE Power Engineering Society General Meeting*, 2007.

[19] J. Saliga, I. Kollar, L. Michaeli, J. Busa, J. Liptak, and T. Virosztek, "A comparison of least squares and maximum likelihood methods using sine fitting in ADC testing," *Measurement*, vol. 46, no. 10, pp. 4362–4368, 2013.

[20] J. M. Blasco, E. Sanchis, V. Gonzalez et al., "Maximum likelihood estimation and non-linear least squares fitting implementation in FPGA devices for high resolution hodoscopy," *IEEE Transactions on Nuclear Science*, vol. 60, no. 5, pp. 3578–3584, 2013.

[21] R. T. Baillie, G. Kapetanios, and F. Papailias, "Modified information criteria and selection of long memory time series models," *Computational Statistics and Data Analysis*, vol. 76, pp. 116–131, 2014.

[22] M. Karimi, "Order selection criteria for vector autoregressive models," *Signal Processing*, vol. 91, no. 4, pp. 955–969, 2011.

[23] H. Yin, Z. Zhu, and F. Ding, "Model order determination using the Hankel matrix of impulse responses," *Applied Mathematics Letters*, vol. 24, no. 5, pp. 797–802, 2011.

[24] R. Schaefer and P. Hauptmann, "Ultrasonic density measurement of liquids – a novel method using a generalized singular value decomposition based system identification," *IEEE Trans. Ultrasonics*, 2006.

[25] R. S. Smith, "Frequency domain subspace identification using nuclear norm minimization and hankel matrix realizations," *IEEE Transactions on Automatic Control*, vol. 59, no. 11, pp. 2886–2896, 2014.

[26] A. Javed and M. I. Ahmad, "Projection-based model order reduction for biochemical systems," in *Proceedings of the 2019 International Conference on Applied and Engineering Mathematics (ICAEM)*, pp. 133–138, Taxila, Pakistan, August 2019.