Model Order Selection Criterion with application to Physical Systems

Rajamani Doraiswami, Lahouari Cheded and Haris M. Khalid

Abstract— **In this paper, it is shown that poles of the discrete-time equivalent of continuous system will lie in the right-half of the complex plane if the sampling rate is chosen to be more than twice the Nyquist rate. This new criterion allows for a quick and reliable separation between the system's poles and any extraneous poles emanating from a variety of artifacts such as high frequency noise and nonlinearities in the system. The system is identified for different model orders and only those for which the poles lie in the right-half plane are chosen. Then using a conventional scheme such as Akaike Information Criterion (AIC), the correct order is chosen from the selected model set. The proposed model order selection criterion has been evaluated on a physical system.**

Index Terms—**Identification, model-order selection, Akaike information criterion, model-based fault diagnosis, model-based control design**

I. Introduction

In recent years there has been a strong emphasis placed on model evaluation criteria. This has been recognized as one of the important areas in model identification with applications to model based condition monitoring, fault diagnosis and controller design [1-7]. It consists of choosing a criterion and using it to select the best approximating model among a class of competing models for a given data set.

In identifying a model of a physical system, the structure of the model may not be identical to that of the mathematical model derived from the physical laws due to various factors including the presence of noise and fast dynamics [1].

For mathematical tractability, the well known criteria based on statistical decision theory when applied to physical system may require simplified assumptions such as long and uncorrelated data record, linearized model and Gaussian probability distribution function (PDF) of the residuals. In view of these assumptions, these criteria may not always give the correct model order. Generally the estimated order may be large due to the presence artifacts arising from noise, nonlinearities, and pole-zero cancellation effects.

A two-stage identification scheme is proposed in [12] is employed. First a high-order is selected so as to capture both the system dynamics and any artifacts (from noise or other sources). Then in the second stage lower order models are derived from the estimated high order model using a frequency-weighted estimation scheme.

The proposed model order selection scheme consists of selecting only the set of models, which are identified using the two-stage scheme for which all the poles are in the right-half plane. The rest of the identified models are not selected as they have extraneous poles. The extraneous poles are located either entirely in the left-half plane or some in the right-half and the rest in the left half plane. The correct model order is determined by evaluating the elected set of models, whose poles are entirely in the right-half plane, using a conventional model structure selection criterion such as the AIC.

A. Classical Information Theoretic Approach[8]

The well known the Maximum Likelihood (*ML*) approach although is the best for parameter estimation for a selected model order, it is unsatisfactory for model order selection. The optimal order, $\hat{n}$ will be maximum possible order since $\min_{\theta}\{L(y,\theta)\}$ is monotonically decreasing function of the order *n*. In other words it will over fit the data by fitting the noisy data (instead of the noise-free data) arbitrarily close to the estimated data as the order is increased. To overcome this problem a number of information theory based approaches has been proposed. They include the well known Akaike information criterion (AIC), Bayes Information criterion (BIC) and the minimum description length (MDL) criterion which address the model order selection problem in a general setup, by for example penalizing not only the loss function but also the model order in the case of AIC. The well known Information Criterion (IC) such as the AIC and BIC is based on including a penalty term to the *ML* criteria so that the optimal order is a trade-off between over fitting the data and minimizing the model complexity. It takes the general form given by:

$$IC = \min_{\theta}\left(-\log f(y\,|\,\theta)\right) + \nu M \qquad (1)$$

where $\log f(y\,|\,\theta)$ is the log-likelihood function of y given the parameter θ, M is the number of parameters , ν is a user-defined penalty usually takes values in the interval $\nu \in [2\ 6]$.

In addition to the above there are other criteria such as Akaike's Final Prediction Error (FPE), minimum description length (MDL) given by:

$$FPE = \left(-\log(f(y\,|\,\theta)\right)\left[N+M\right]/\left[N-M\right] \quad (2)$$

$$MDL = \left(1+\log(N)M\right)\left(-\log(f(y\,|\,\theta)\right) \quad (3)$$

Rajamani Doraiswami is with the Department of Electrical & Computer Engineering, University of New Brunswick, Fredericton, Canada, P. O. Box 4400, e-mail: dorai@unb.ca

Lahouari Cheded is with the Systems Engineering Department, King Fahd University of Petroleum and Minerals, P. O. Box 116, Dhahran 31261, Saudi Arabia, e-mail: cheded@kfupm.edu.sa.

Haris M. Khalid is with the Systems Engineering Department, King Fahd University of Petroleum and Minerals, P. O. Box 8283, Dhahran 31261, Saudi Arabia, e-mail: mhariskh@kfupm.edu.sa.

The optimal order, $\hat{n}$, is selected by minimizing the information criterion:

$$\hat{n} = arg\min_{n}\left\{\min_{\theta}\left[g\left(\log f(y\mid\theta)\right)\right]\right\} \quad (4)$$

In practice it is difficult to know *a priori* or estimate *a posteriori* the PDF of the measurement from the data, and hence the log-likelihood functions. In view of this, a direct approach is proposed.

B. Direct Approach to model order selection

Model of a continuous-time system: Most physical systems are inherently continuous in time, and its discrete-time equivalent model is employed for identification. The model in the frequency domain is given by:

$$y_0(s) = G_p^c(s)u(s) \quad (5)$$

$$y(s) = y_0(s) + v_p(s) \quad (6)$$

where $y(s)$ is the output, $y_0(s)$ is the noise-free output, $u(s)$ is the input, $v_p(s)$ is the error term, $G_p^c(s)$ is the transfer function of the system is s is the Laplace transform variable. The error v_p is in general a stochastic process modeled as an output of a linear system and driven by a zero mean white noise:

$$v_p(s) = G_v^c(s)v(s) \quad (7)$$

Assumptions

- The transfer function $G_p^c(s)$ is asymptotically stable.
- The transfer function $G_v^c(s)$ is minimal phase.
- The poles of the noise model $G_v^c(s)$ and the system model $G_p^c(s)$ are located at different frequency regions of the complex s-plane. For example, the system $G_p^c(s)$ is a low pass system transfer function and $G_v^c(s)$ is a band pass noise transfer function as shown in the figure 1.

The frequency region where the poles of $G_p^c(s)$ are located is indicated in the red while the region where the poles of $G_v^c(s)$ are located is indicated in the green.

The power spectral density of the noise-free output y_0 denoted $P_{y_0y_0}(\omega)$, and the power spectral density of the noise v_p denoted $P_{v_pv_p}(\omega)$ are given by:

$$P_{y_0y_0}(\omega) = \left|G_p(\omega)\right|^2 P_{uu}(\omega) \quad (8)$$

$$P_{v_pv_p}(\omega) = \left|G_v(\omega)\right|^2 \sigma^2 \quad (9)$$

Let the frequency regions of the s-plane where the poles of $G_p^c(s)$ and $G_v^c(s)$ are located, which are indicated in the figure by the red and the green colored regions in the s-plane in the above figure, be denoted by Ω and Ω_v respectively. The spectral densities $P_{y_0y_0}(\omega)$ and $P_{v_pv_p}(\omega)$ will be dominant in Ω and Ω_v respectively, and will flat in the rest of the s-plane regions. The ratio of the power spectral densities is measure of the Signal to Noise Ratio (SNR)

$$SNR = P_{y_0y_0}(\omega) / P_{vv}(\omega) \quad (10)$$

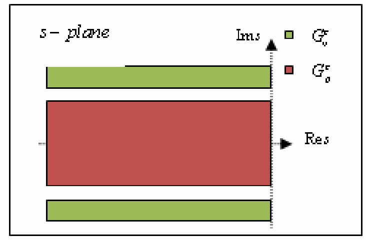

Figure 1. Low-Pass and Band-Pass transfer system

C. Discrete-time equivalent model of a system

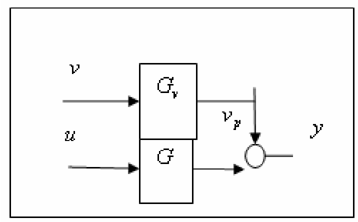

Figure 2. Model of a discrete-time system

A discrete-time equivalent of the continuous time model of the physical system takes form (See Figure 2):

$$y_0(z) = G_p(z)u(z) \quad (11)$$

$$y(z) = y_0(z) + v_p(z) \quad (12)$$

where $y(z)$ is the output, $y_0(z)$ is the noise-free output, $u(z)$ is the input, $v_p(z)$ is the error term, $G_p(z)$ is discrete-time equivalent of the continuous-time transfer function $G_p^c(s)$, z is the z-transform variable. The transfer function $G_p(z)$ takes the form:

$$G_p(z) = \frac{N_p(z)}{D_p(z)}$$

where $N_p(z)$ and $D_p(z)$ are the numerator and denominator polynomials. The error v_p is in general a stochastic process modeled as an output of a linear system and driven by a zero mean white noise process v, with variance σ^2, that is v_p is a *colored noise*,

$$v_p(z) = G_v(z)v(z) \quad where \quad G_v(z) = \frac{N_v(z)}{D_v(z)} \quad (13)$$

It is assumed that the noise transfer function $G_v(z)$ is minimal phase.

D. Selection of model order in identification: ideal noise-free case

The selection of model order is crucial step in identifying an accurate model of a system. An order less than the true value will not capture the system behavior, while a model higher than the true value will capture the noise and other artifacts.

In order to obtain an insight on the role of the selected model order on the identified model it is assumed that the error term $v(z)=0$ is zero. That is the system model is assumed to:

$$A(z)y(z)=B(z)u(z)$$

Let the true orders of $G_p(z)$, $G_v(z)$ and the selected order be n_p , n_v and n_s respectively. Thanks to the pole zero cancellation, the system transfer function $G(z)$ will be the same for all choices of the model order as long as the order is chosen to be greater than or equal to the true order.

$$G(z)=\frac{B(z)}{A(z)}=G_p(z)=\frac{N_p(z)}{D_p(z)} \quad for\, all \;\; n_s \geq n_p \quad (14)$$

The numerator $B(z)$ and the denominator $A(z)$ can be expressed in terms of the 'true' numerator $N_p(z)$ and 'true' denominator $D_p(z)$ as follows:

$$B(z)=N_p(z)P_{art}(z) \quad for\, all \;\; n_s \geq n_p \qquad (15)$$

$$A(z)=D_p(z)P_{art}(z) \quad for\, all \;\; n_s \geq n_p \qquad (16)$$

where $P_{art}(z)$ is some polynomial which is an artifact due to noise and other sources.

In the ideal noise-free case, as long as the order n_s is selected to be greater than or equal to the true order n_p , the resulting transfer function will contain the true model of the system. If on the other hand, the model order n_s is selected to be less than the true order n_p ,the resulting system will not contain the true model. If the selected order is equal to the true order $n_s=n_p$ the identified model will be equal to true model without any artifacts:

$$G(z)=G_p(z), \;\; B(z)=N_p(z), \;\; A(z)=D_p(z) \quad (17)$$

The question of whether the identified model is equal to the true model is addressed in the next section. The sampling period is chosen such that the poles of the system transfer function is located on the right-half while that of the noise is located on the left-half of the z-plane. This information is employed to determine whether the identified is the true model. The question as to whether iterating the model order period till the true model is identified is answered in a later section by extensive simulation.

E. Poles of the dicrete-time equivalent model

The poles of the discrete-time model, i.e. λ_d are related to those of the continuous-time model, i.e. λ_c , by:

$$\lambda_d=e^{\lambda_c T_s} \qquad (18)$$

where $T_s=1/f_s$ is the sampling period and f_s is the corresponding sampling frequency.

<u>Proposition</u> : If the sampling frequency is chosen in the range $2f_c \leq f_s < 4f_c$, then the complex-conjugate poles of the equivalent discrete system will all lie in the right-half of the z-plane, whereas the real ones will all lie on the positive real line, i.e.

$$\lambda_d \in \begin{cases} Z^+ & if \;\lambda_c \;\; is\, complex \\ R^+ & if \;\lambda_c \;\; is\, real \end{cases}$$

where $f_c=\omega_c/2\pi$ is defined in terms of $\lambda_c=\alpha_c+j\omega_c$

<u>Proof:</u> The poles of the discrete-time model, i.e. λ_d , are related to those of the continuous-time model, i.e. $\lambda_c=\alpha_c+j\omega_c$, by :

$$\lambda_d=e^{\alpha_c T_s}\left[\cos 2\pi\left(\frac{f_c}{f_s}\right)+j\sin 2\pi\left(\frac{f_c}{f_s}\right)\right] (19)$$

From the above equation, we deduce the following:

$$\lambda_d \in Z^+ \Leftrightarrow \cos\left(2\pi f_c/f_s\right)e^{\alpha_c T_s} \geq 0 \Leftrightarrow -\frac{1}{4} \leq \frac{f_c}{f_s} \leq \frac{1}{4} \qquad (20)$$

That is

$$\lambda_d \in Z^+ \;\; for \;\; all\, \lambda_c \Leftrightarrow -\frac{1}{4} \leq \frac{f_c}{f_s} \leq \frac{1}{4} \qquad (21)$$

This shows that the discrete-time poles lie in the right half of the Z-plane if the sampling rate (f_s) is more than twice the Nyquist rate ($2f_c$).

Let us now extend the result to the case where the system output is corrupted by colored noise. Let $BW_s=\left\{f_{\min}^s \quad f_{\max}^s\right\}$ be the bandwidth (the dominant spectral frequency interval) of the system transfer function $G_p^c(s)$ and $BW_v=\left\{f_{\min}^v \quad f_{\max}^v\right\}$ be the bandwidth of the noise transfer function $G_v^c(s)$.

From the proposition we deduce the following:

- The poles of the system will be in the right-half of the z-plane if the sampling rate f_s larger than four times the maximum frequency content of the system, $f_{\max}^s$, $\lambda_d \in Z^+ \;\; iff -\frac{1}{4} \leq \frac{f_{\max}^s}{f_s} \leq \frac{1}{4}$
- However, in order that the noise poles are located in a region different from that of the system poles, the sampling rate f_s must be smaller than four times the minimum frequency content of the noise, $f_{\min}^v$: $\lambda_d^v \in Z^- \;\; iff -\frac{1}{4} \geq \frac{f_{\min}^v}{f_s} \geq \frac{1}{4}$
- Thus to ensure that the system poles are located in the right-half plane and the noise poles in the left-half plane, the sampling rate, f_s must be larger than four times the maximum frequency $f_{\max}^s$ of the system, and less than four times the minimum

frequency of the noise, $f^{v}_{\min}$, $4f^{s}_{\max} \le f_s \le 4f^{v}_{\min}$

Example: Consider an example of low pass system whose output is corrupted by a colored noise:

$$y(z) = G_p(z)u(z) + v_p(z) \qquad (22)$$

Where $G_p(z) = \dfrac{z^{-2}}{1-0.8z^{-1}+0.64z^{-2}}$ and

$$v_p(z) = \left(\frac{1+z^{-1}}{1+0.8z^{-1}+0.64z^{-2}}\right)v(z)$$

The bandwidth of the system is $BW_s = \left[0.1Hz \quad 0.2\,Hz\right]$. The system has poles on the right-half of the z-plane, 0.4 ± j0.6928 while noise poles are located on the left-half plane −0.4 ± j0.6928. The noise v_p is a high frequency colored noise with spectral peak at 0.35 and has a bandwidth, $BW_v = \left[0.3Hz \quad 0.4\,Hz\right]$. It has no poles in the right half of the z-plane, and its spectrum in the is flat over the system bandwidth. The sampling frequency is selected to be $f_s = 1Hz$ to meet the requirement $4f^{s}_{\max} \le f_s \le 4f^{v}_{\min}$ that is $4(0.2) \le f_s \le 4(0.3)$.

Figures 3(a) and 3(b) below show the pole locations and the power spectral densities. The power spectral density of the system dominates that of the noise inside the system bandwidth.

$P_{y_0y_0}(\omega) >> P_{v_pv_p}(\omega) \quad \omega \in \Omega$ where Ω is the bandwidth of $G_p(\omega)$. The power spectral density of the noise is flat in Ω, while dominating outside Ω.

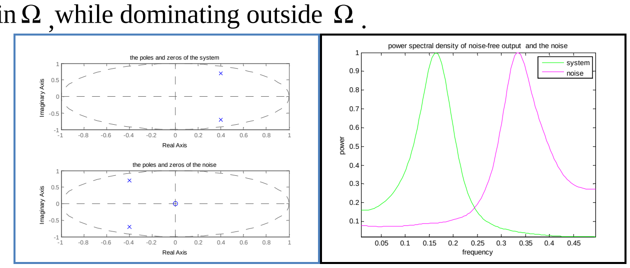

Figure 3. Figure 3(a) on the left shows the poles of the system and the noise and Figure 3(b) on the right shows the magnitude squared response of the plant and the power spectral density of the noise.

II. Identification Scheme

There are a number of identification scheme to overcome bias in the estimated model parameters due to colored noise. One approach is to preprocess the input-output data so as to attenuate the noise by cross correlating the data with a signal which is uncorrelated with the noise but is highly correlate with the input data. Example includes an instrument variable scheme. The other approach is to identify a high order model which captures both the system model and the noise artifacts, and separate the system model using some *a priori* information. In [12] a frequency weighted estimation scheme is employed to separate the system model from the high order model by weighting the spectral region where the model of the system dominates that of the noise.

The two-stage scheme proposed in [12] is employed here in. A very high order model is identified from which set of low order models covering all orders are estimated.

III. Model Order Selection

The proposed model order selection scheme consists of selecting only the set of models, which are identified using the two-stage scheme for which all the poles are in the right-half plane. The rest of the identified models are not selected as they have extraneous poles. The extraneous poles are located either entirely in the left-half plane or some in the right-half and the rest in the left half plane. The correct model order is determined by evaluating the elected set of models, whose poles are entirely in the right-half plane, using a conventional model structure selection criterion such as the AIC.

IV. Evaluation on Simulated Systems

The proposed model order selection scheme is first evaluated on a simulated system to gain some insight on the various factors that affect the model order selection such as the effects of the noise artifacts. The two-stage identification scheme proposed in [12] is employed here.

<u>Case 1:</u> Low-pass system with SNR= 1

$$y(z) = \left(\frac{z^{-2}}{1-0.7895z^{-1}+0.3769z^{-2}}\right)u(z) + \left(\frac{1}{1+1.79z^{-1}+0.8z^{-3}}\right)v(z)$$

where y is the output, u is the input and v is the measurement noise. The poles of the system and poles of the noise are located respectively at $0.1688 \pm j0.6144$ and $-0.8957 \pm j0.0882$

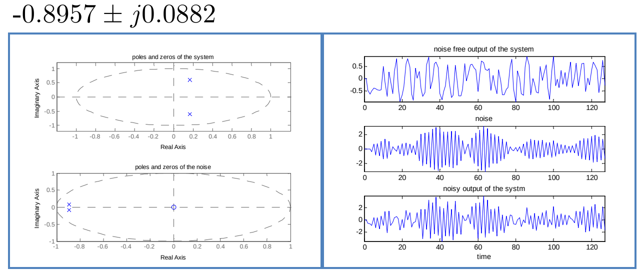

Figure 4. Figure on the left shows the poles of ths system and the noise. The figure on the right shows the noise-free output $y_0(k)$, the noise $v(k)$ and the noisy output $y(k)$. The SNR is 0.09.

The system is identified for different choices of the model order ranging from 2 to 6 (See Figure 4).

From Fig. 5, it can be seen that the poles of the discrete-time equivalent are located in the right-half of the z-plane when the selected model orders are 1 and 2. The model order value of 2 was selected as it yields the lowest AIC. For orders greater than 2, the poles located in the right half of the z-plane are extraneous ones due to noise and other artifacts. The extraneous poles are cancelled by the system zeros. In this case, the results obtained from the loss function, AIC and the proposed scheme all indicate correctly the model order. The selected order is 2 as orders higher than 2 will lead to poles in the left-half of the z-plane. In this special case, the estimated poles are, $0.3913 \pm j0.4616$ while the true poles are $0.3929 \pm j0.4620$.

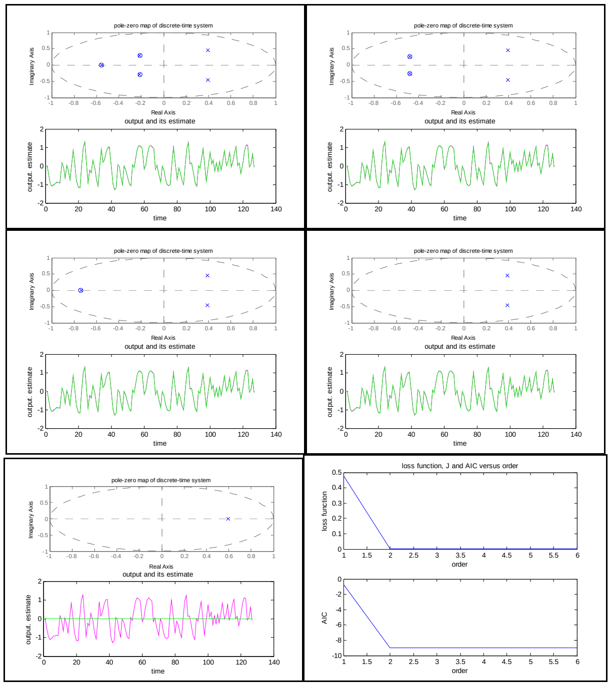

Figure 5. Pole-zero maps, the output and its estimate for selected orders 1 to 5. The last pair of graphs shows the loss functions and the AIC value.

The poles of the discrete-time equivalent are located in the right-half of the z-plane when the selected model orders are 1 and 2. The model order value of 2 was again selected because it yields the lowest AIC. For orders greater than 2, the poles located in the right half of the z-plane are extraneous ones due to noise and other artifacts.

Note: The AIC and the loss functions did not indicate the correct model order in low SNR cases**.**

V. EVALUATION ON SIMULATED SYSTEMS

The identification of a practical process is a challenging task as its model is generally stochastic, complex and nonlinear. For practical purposes, the identified model needs to be simple and linear for the design of a controller, a condition monitoring or a fault diagnosis system. In our practical evaluation of the proposed scheme, we consider the following physical process:

- A benchmark laboratory-scale process control system

An evaluation of the proposed scheme for fault diagnosis was performed on a benchmark laboratory-scale process control system using National Instruments LABVIEW as shown below in Fig 7.

The process control system under evaluation is a lab-scale two-tank system consisting of a dc motor, a pump, two tanks and a Proportional and Integral (PI) controller. The cascade connection of a dc motor and a pump relating the input to the motor, *u,* and the flow, Q_i, is a first-order time-delay system expressed by:

$$\dot{Q}_i = -a_m Q_i + b_m \phi(u) \quad (29)$$

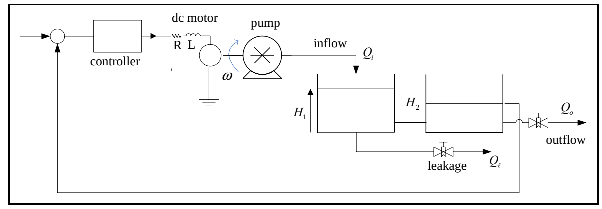

Figure 6. Process control system: A Lab-scale two-tank system

where a_m and b_m are the parameters of the motor-pump system and $\phi(u)$ is a dead-band and saturation type of nonlinearity. The PI controller is given by:

$$\begin{aligned} \dot{x}_3 &= e = r - h_2 \\ u &= k_p e + k_I x_3 \end{aligned} \quad (30)$$

where k_p and k_I are the proportional and integral gains, respectively, and r is the reference input. With the inclusion of the leakage, the liquid level system is modeled by:

$$\begin{aligned} A_1 \frac{dH_1}{dt} &= Q_i - C_{12}\varphi(H_1 - H_2) - C_\ell \varphi(H_1) \\ A_2 \frac{dH_2}{dt} &= C_{12}\varphi(H_1 - H_2) - C_0 \varphi(H_2) \end{aligned} \quad (31)$$

where $\varphi(.) = sign(.)\sqrt{2g(.)}$, $Q_\ell = C_\ell \varphi(H_1)$ is the leakage flow rate, $Q_0 = C_0\varphi(H_2)$ is the output flow rate, H_1 is the height of the liquid in tank 1, H_2 is the height of the liquid in tank 2, A_1 and A_2 are the cross-sectional areas of the 2 tanks, g=980 $cm/\sec^2$ is the gravitational constant, C_{12} and C_o are the discharge coefficients of the inter-tank and output valves, respectively. The linearized model of the entire system formed by the motor, pump, and the tanks is given by:

$$\dot{x} = Ax + Br \qquad y = Cx \quad (32)$$

$$x = \begin{bmatrix} h_1 \\ h_2 \\ x_3 \\ q_i \end{bmatrix}, \quad A = \begin{bmatrix} -a_1 - \alpha & a_1 & 0 & b_1 \\ a_2 & -a_2 - \beta & 0 & 0 \\ -1 & 0 & 0 & 0 \\ -b_m k_p & 0 & b_m k_I & -a_m \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & 0 & 1 & b_m k_p \end{bmatrix}^T, \quad C = [1 \quad 0 \quad 0 \quad 0]$$

Where q_i, q_l, q_0, h_1 and h_2 are respectively the increments in Q_i, Q_l, Q_0, H^0_1 and H^0_2 whereas α_1, α_2, α and β are parameters associated with linearization, α is associated with leakage, $q_l = \alpha h_1$, and β is the output flow rate,

$q_o = \beta h_2$.

Various orders for the model of the fluid system ranging from 1 to 6 were initially selected, and for each order, the system model was identified using a least-squares method. The following quantities were computed:

- Poles of the identified model
- The loss function, J given by

$$J = \frac{1}{N}\sum_{k=0}^{N-1}\left[y(k) - \hat{y}(k)\right]^2 \quad (33)$$

where $\hat{y}$ is the estimate of the system output, y, N is the number of data samples.

- AIC measure

TABLE I. POLES OF FOR DIFFERENT SELECTED ORDER

order 1	order 2	order 3	order 4	order 5
0.9850	0.9847	0.9847	0.9845	0.9845
	0.0712	0.0340 ± j0.5910	0.0915	0.3301 ± j0.5549
			-0.0189 ± j0.6197	-0.3058 ± j0.5938

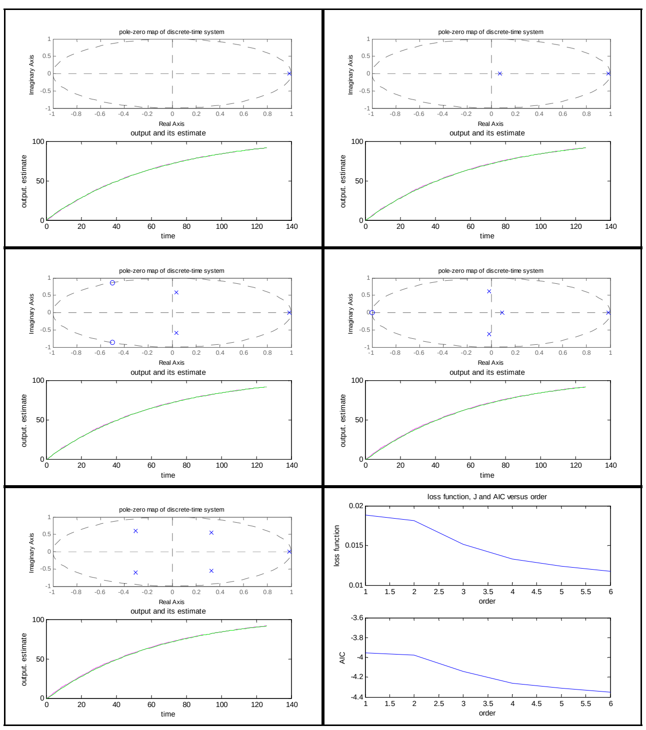

Figure 7. Pole-zero maps for various selected model orders ranging from 1 to 5. The output and its estimates are also shown. The last pair of graphs shows J and AIC

With reference to Table I below, it can be easily seen that for the selected model orders 1 to 3, all the poles are on the right half plane. The vital question that arises now is how to select the appropriate model order out of these three order values. The physical interpretation of that is that the identification algorithm cannot capture the fast actuator (dc-motor) dynamics. Table 1 lists the pole locations for the selected orders 1,2,..5. The poles located on the right-half plane are highlighted.

VI. CONCLUSION

The model order selection criteria complements the methods proposed in the literature. The method is direct, useful and simple and provides an additional check on the selected model order. However, it may fail in the case of low signal to noise ratio, and when the system poles are located close to the left half of the z-plane and the noise poles are located close to the right-half plane. The estimated system poles may drift to the left-half plane in the presence of noise. Similarly, the estimated poles of the noise may drift to the right-half plane system. If the estimated poles are located in the right-half plane they are the poles of system. However, it cannot ensure that all the poles located in the right-half plane are system poles and all the poles in the left-half plane are not system poles. The evaluation of the proposed model order selection criterion on the two tank system depicts the overall picture of the scheme.

ACKNOWLEDGEMENTS

Authors would like to acknowledge the support of NSERC, Universities of New Brunswick and KFUPM. The contribution of Mohammed Shahab of KFUPM is appreciated.

REFERENCES

[1] R. Doraiswami, C.P.Diduch and Jiong Tang, "A Diagnostic Model For Identifying Parametric Faults" , IFAC World congress Seoul, Korea, 2008 (also to be published in IEEE Trans. on Control Systems Technology, 2009)

[2] S.X. Ding, " Model-based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools" Springer-Verlag 2008

[3] Silvio Simani, Cesare Fantuzzi and Ronald J Patton, Model-based Fault Diagnosis using Identification Techniques, Advances in Industrial Control, Springer Verlag, 2003.

[4] Patton, R.J. Paul M. Frank, and Robert N. Clark, Issues in Fault Diagnosis for Dynamic Systems, Springer-Varlag, 2000.

[5] Chen, J. and Patton, Robust Model-based Fault Diagnosis for Dynamic Systems, Kluwer Academic Publishers, 1999.

[6] Janos J. Gertler Fault Detection and Diagnosis in Engineering Systems, Marcel Dekker, Inc, 1998.

[7] R. Isermann, " Fault diagnosis of Machines via parameter estimation and knowledge processing", Automatica, Vol. 29, No.4, pp. 825-825, 1993.

[8] Nils Lid Hjort, " Model Selection and Model Averaging", Cambridge Series in Statistical and Probabilistic Mathematics, 2008

[9] Ljung, L.: 'System identification: theory for the user' (Prentice-Hall, 1999)

[10] R. Palaniappan, P.Raveendran, Shogo Nishida and Naoki Saiwaki, "Autoregressive Spectral Analysis and Model Order Selection Criteria for EEG Signals" 0-7803-6355-WOO, 2000 IEEE.

[11] Broersen, P. M. T., and de Waele, S. Empirical time series analysis and maximum likelihood estimation. In IEEE Benelux Signal Processing Symposium (2000), pp. 1–4.

[12] R.Doraiswami, "A two-stage identification with application to control, feature extraction, and spectral estimation" IEEE Proceedings Control Theory & Applications, Vol. 152, No.4, July 2005.