XXXVI Ibero-Latin American Congress on Computational Methods in Engineering

CILAMCE

Rio de Janeiro, 22-25 Nov 2015

AN APPROACH OF INFORMATION CRITERIA FOR MODEL SELECTION IN STATE SPACE

Jean Pierre López Vargas

Paulo David Battaglin

Gilmar Barreto

jean_88@dsif.fee.unicamp.br

paulodav@dsif.fee.unicamp.br

gbarreto@dsif.fee.unicamp.br

Laboratory of Control and Intelligent Systems (LCSI).

Department of Semiconductors, Instruments and Photonics, DSIF.

Faculty of Electrical and Computer Engineering- FEEC

University of Campinas, UNICAMP, Campinas, SP

Av. Albert Einstein-400, 13083-852, Brasil

Abstract. *In this article, we show a form to evaluate mathematical models represented in state space through the Akaike Information Criterion (AIC). We use the maximum likelihood method for estimation the set of optimal parameters of the model and the recursive Kalman Filter algorithm for estimating the state of same model. Thus we estimate the statistical quality of models by the Criterion of Akaike. This article proposes to clarify and make it simpler to understand the AIC in state space, which will be shown through an example.*

Keywords: *Akaike Information Criterion, Maximum likelihood method, Recursive Kalman Filter, Time series, State Space models.*

1 INTRODUCTION

The selection of models is important and of interest to study several areas. In the control area one way to represent the data of an observed system is the modeling in state space, although the criteria of this kind of modeling are underused. The objective of this selection is to know the *statistic quality* of the set of models proposed to represent the real. From the studies of

(Akaike, 1973) which proposes the Akaike Information Criterion (AIC), develop other known criteria such as the Bayesian Information Criterion (BIC), Schwarz Information Criterion (SIC) and Hannan and Quinn(HQ), we see no one criteria is focused on modeling in the state space. In the work of J.Cavanaught arising from his doctoral thesis(Cavanaugh, 1993), studies for modeling in state space uses different information criteria. These works show how to use the Akaike through the maximum likelihood method for the estimated parameters, and then use the log-likelihood function. This function contains innovations and covariance matrices of state space model, according to the equation (20). The development of AIC for modeling in state space with large samples is one of the goals of this technique, which we will illustrate through a didactic example.

2 Modeling in State Space

The state space mathematical models represent a real system and these models are expressed by two equations that describe the inputs, outputs and disturbance of a dynamic system. The first equation (1a) represent the states of the model and the second (1b) is the equation of the outputs or observations. Generally this model is written as follows (Barreto, 2002):

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}_t\boldsymbol{x}_t + \boldsymbol{B}_t\boldsymbol{u}_t + \boldsymbol{w}_t \tag{1a}$$

$$\boldsymbol{y}_t = \boldsymbol{C}_t\boldsymbol{x}_t + \boldsymbol{D}_t\boldsymbol{u}_t + \boldsymbol{v}_t \tag{1b}$$

for $t = 1, \ldots, n$, where $\boldsymbol{u}_t \in \boldsymbol{R}^m$, $\boldsymbol{y}_t \in \boldsymbol{R}^l$ and $\boldsymbol{x}_t \in \boldsymbol{R}^n$ which represent the input vector, output vector and state vector respectively in t-th time instant. The matrices $\boldsymbol{A}_t \in \boldsymbol{R}^{n\times n}$, $\boldsymbol{B}_t \in \boldsymbol{R}^{n\times m}$, $\boldsymbol{C}_t \in \boldsymbol{R}^{l\times n}$, $\boldsymbol{D}_t \in \boldsymbol{R}^{l\times m}$, are respectively the transition matrix of the system states, the bonding matrix of the inputs to the system, the connection matrix of the systems states to the outputs, and the matrix which binds the inputs to the outputs. These matrix have elements invariant in time. Measurements of system noise are represented by the vector $\boldsymbol{w}_t \in \boldsymbol{R}^n$ and the noises in the measurements on system outputs are represented by the vector $\boldsymbol{v}_t \in \boldsymbol{R}^l$.

One of the model variations in state space is the system without inputs, in which case we will have a time series of this type

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}_t\boldsymbol{x}_t + \boldsymbol{w}_t \tag{2a}$$

$$\boldsymbol{y}_t = \boldsymbol{C}_t\boldsymbol{x}_t + \boldsymbol{v}_t \tag{2b}$$

We will work with this type of model in state space in the form of time series based on the reference (Cavanaugh, 1993).

A relation of the error covariance matrix and the state transition matrix can be made through of the equation (3), when $\boldsymbol{x}_t$ is a stationary process with zero mean, whereas $\boldsymbol{x}_0$, $\boldsymbol{w}_t$ and $\boldsymbol{v}_t$ are mutually uncorrelated. The initial state is $\boldsymbol{x}_0$. From this point we will simplify the notation of matrices $\boldsymbol{A}_t$,$\boldsymbol{B}_t$,$\boldsymbol{C}_t$ and $\boldsymbol{D}_t$ for $\boldsymbol{A}$,$\boldsymbol{B}$,$\boldsymbol{C}$ and $\boldsymbol{D}$.

Thus we have the following equations:

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{w}_t \\ E[(\boldsymbol{x}_{t+1})(\boldsymbol{x}_{t+1})^T] &= E[(\boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{w}_t)(\boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{w}_t)^T] \\ &= E[\boldsymbol{A}\boldsymbol{x}_t\boldsymbol{x}_t^T\boldsymbol{A}^T] + E[\boldsymbol{A}\boldsymbol{x}_t\boldsymbol{w}_t^T] + E[\boldsymbol{w}_t^T\boldsymbol{x}_t^T\boldsymbol{A}^T] + E[\boldsymbol{w}_t\boldsymbol{w}_t^T] \\ &= \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^T + \boldsymbol{Q} \end{aligned} \tag{3}$$

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{C}\boldsymbol{x}_t + \boldsymbol{v}_t \\ E[(\boldsymbol{y}_t)(\boldsymbol{y}_t)^T] &= E[(\boldsymbol{C}\boldsymbol{x}_t + \boldsymbol{v}_t)(\boldsymbol{C}\boldsymbol{x}_t + \boldsymbol{v}_t)^T] \\ &= E[\boldsymbol{C}\boldsymbol{x}_t\boldsymbol{x}_t^T\boldsymbol{C}^T] + E[\boldsymbol{A}\boldsymbol{x}_t\boldsymbol{v}_t^T] + E[\boldsymbol{v}_t^T\boldsymbol{x}_t^T\boldsymbol{C}^T] + E[\boldsymbol{v}_t\boldsymbol{v}_t^T] \\ &= \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^T + \boldsymbol{R} \end{aligned} \tag{4}$$

where the error covariance matrices are $E[\boldsymbol{w}_t\boldsymbol{w}_t^T] = \boldsymbol{Q}$, $E[\boldsymbol{v}_t\boldsymbol{v}_t^T] = \boldsymbol{R}$. Another topic developed in this work is the relation of the innovations in the Kalman filter algorithm with the likelihood function (Cavanaugh, 1993)(Schweeppe, 1965). Representation of the innovation $\boldsymbol{e}_t$ of the Kalman filter((Brown & Hwang, 2012)) is defined from the equation (2b) an it is shown by equation (5)

$$\boldsymbol{e}_t(\boldsymbol{\theta}) = \boldsymbol{y}_t - \boldsymbol{C}\boldsymbol{x}_t^{t-1}(\boldsymbol{\theta}) \tag{5}$$

wherein $\boldsymbol{x}_t^{t-1}(\boldsymbol{\theta})$ is the expectation of $\boldsymbol{x}_t$ given the observed by the $t-1$ time according to equation (6)

$$\boldsymbol{x}_t^{t-1}(\boldsymbol{\theta}) = E\{\boldsymbol{x}_t|\boldsymbol{Y}_{t-1}\} \tag{6}$$

where $\boldsymbol{\theta}$ represents the system parameters and $\boldsymbol{Y}_{t-1} = [\boldsymbol{y}_1, ..., \boldsymbol{y}_{t-1}]$ represents the observed data until the $t-1$ time. We define the error estimate as the difference between the state vector observable and the estimated according to equation (7).

$$(\boldsymbol{\varepsilon}_t = \boldsymbol{x}_t - \boldsymbol{x}_t^{t-1}(\boldsymbol{\theta})) \tag{7}$$

The system state covariance matrix at the t time can be write as

$$\boldsymbol{P}_t^{t-1}(\boldsymbol{\theta}) = E[\varepsilon_t\varepsilon_t^T|\boldsymbol{Y}_{t-1}] = E[(\boldsymbol{x}_t - \boldsymbol{x}_t^{t-1}(\boldsymbol{\theta}))(\boldsymbol{x}_t - \boldsymbol{x}_t^{t-1}(\boldsymbol{\theta})^T)|\boldsymbol{Y}_{t-1}] \tag{8}$$

If we do not considerer the system noise $\boldsymbol{w}_t$, which has zero mean and is not correlated with the noise $\boldsymbol{w}_{t-1}$, we have the system state

$$\boldsymbol{x}_t^{t-1}(\boldsymbol{\theta}) = \boldsymbol{A}\boldsymbol{x}_{t-1}^{t-1} \tag{9}$$

In the $t+1$ time we have for the system state

$$\boldsymbol{x}_{t+1}^{t}(\boldsymbol{\theta}) = \boldsymbol{A}\boldsymbol{x}_t^t \tag{10}$$

The error estimate in the $t+1$ time is provided by

$$\begin{aligned}
\boldsymbol{\varepsilon}_{t+1} &= \boldsymbol{x}_{t+1} - \boldsymbol{x}_{t+1}^{t}(\boldsymbol{\theta}) \\
\boldsymbol{\varepsilon}_{t+1} &= (\boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{w}_t) - \boldsymbol{A}\boldsymbol{x}_t^{t} \\
\boldsymbol{\varepsilon}_{t+1} &= \boldsymbol{A}(\boldsymbol{x}_t - \boldsymbol{x}_t^{t}(\boldsymbol{\theta})) + \boldsymbol{w}_t \\
\boldsymbol{\varepsilon}_{t+1} &= \boldsymbol{A}\boldsymbol{\varepsilon}_t + \boldsymbol{w}_t
\end{aligned} \tag{11}$$

The error covariance matrix in the $t+1$ is computed by

$$\begin{aligned}
\boldsymbol{P}_{t+1} &= E[(\boldsymbol{\varepsilon}_{t+1})(\boldsymbol{\varepsilon}_{t+1})^T] \\
\boldsymbol{P}_{t+1} &= E[(\boldsymbol{A}\boldsymbol{\varepsilon}_t + \boldsymbol{w}_t)(\boldsymbol{A}\boldsymbol{\varepsilon}_t + \boldsymbol{w}_t)^T] \\
\boldsymbol{P}_{t+1} &= \boldsymbol{A}E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t^T]\boldsymbol{A}^T + E[\boldsymbol{w}_t\boldsymbol{w}_t^T]
\end{aligned}$$

As the error covariance matrix in the t time is $\boldsymbol{P}_t$ and the system noise covariance matrix of $\boldsymbol{w}_t$ is $\boldsymbol{Q}$, then we have for the covariances matrices in $t+1$ time

$$\boldsymbol{P}_{t+1} = \boldsymbol{A}\boldsymbol{P}_t\boldsymbol{A}^T + \boldsymbol{Q} \tag{12}$$

Otherwise, the innovations covariance matrix can be achieved by developing the equation (2b) and considering the equation (6)

$$\begin{aligned}
\boldsymbol{e}_t(\boldsymbol{\theta}) &= \boldsymbol{y}_t - \boldsymbol{C}\boldsymbol{x}_t^{t-1}(\boldsymbol{\theta}) \\
\boldsymbol{e}_t(\boldsymbol{\theta}) &= (\boldsymbol{C}\boldsymbol{x}_t + \boldsymbol{v}_t) - \boldsymbol{C}\boldsymbol{x}_t^{t-1}(\boldsymbol{\theta}) \\
\boldsymbol{e}_t(\boldsymbol{\theta}) &= \boldsymbol{C}(\boldsymbol{x}_t - \boldsymbol{x}_t^{t-1}(\boldsymbol{\theta})) - \boldsymbol{v}_t \\
\boldsymbol{e}_t(\boldsymbol{\theta}) &= \boldsymbol{C}\boldsymbol{\varepsilon}_t + \boldsymbol{v}_t
\end{aligned} \tag{13}$$

The measurement covariance matrix $\boldsymbol{R}$ and the system state covariance matrix $\boldsymbol{P}$ are related by the equations (14) and (15) below:

$$\begin{aligned}
E[\boldsymbol{e}_t(\boldsymbol{\theta})\boldsymbol{e}_t^T(\boldsymbol{\theta})] &= E[(\boldsymbol{C}\boldsymbol{\varepsilon}_t + \boldsymbol{v}_t)(\boldsymbol{C}\boldsymbol{\varepsilon}_t + \boldsymbol{v}_t)^T] \\
E[\boldsymbol{e}_t(\boldsymbol{\theta})\boldsymbol{e}_t^T(\boldsymbol{\theta})] &= \boldsymbol{C}E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t^T]\boldsymbol{C}^T + E[\boldsymbol{v}_t\boldsymbol{v}_t^T]
\end{aligned} \tag{14}$$

We defined that the innovations covariance matrix in t time as $\boldsymbol{\Sigma}_t(\boldsymbol{\theta})$

$$\boldsymbol{\Sigma}_t(\theta) = \boldsymbol{C}\boldsymbol{P}_t^{t-1}(\boldsymbol{\theta})\boldsymbol{C}^T + \boldsymbol{R} \tag{15}$$

3 Evaluation of mathematical model in state space through AIC criterion

3.1 Maximum Likelihood Method

The value that maximizes the likelihood function is also the value that maximizes the function provided by the log-likelihood function, called log-likelihood function. This function is more suitable for computations of the model parameters and allows that mathematical models can be compared additively. The log-likelihood function of vector $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_p)$ for a parametric model specified $f(x|\boldsymbol{\theta})(\boldsymbol{\theta} \in \Theta \subset R^p)$ is given by

$$\ell(\boldsymbol{\theta}) = \sum_{\alpha=1}^{n} \log f(x_\alpha|\boldsymbol{\theta}) \tag{16}$$

The maximum likelihood method seeks to maximize $\boldsymbol{\theta} \in \Theta$ of the estimator $\ell(\boldsymbol{\theta})$ such that it satisfies following

$$\ell(\hat{\boldsymbol{\theta}}) = \max_{\boldsymbol{\theta}\in\Theta} \ell(\boldsymbol{\theta}) \tag{17}$$

wherein $\hat{\boldsymbol{\theta}}$ is called maximum likelihood estimator.

If the data used in the estimation should be explicit in the notation, then the maximum likelihood estimator is denoted by $\hat{\boldsymbol{\theta}}(\mathbf{x}_n)$. The model $f(x|\hat{\boldsymbol{\theta}})$ determinated by $\hat{\boldsymbol{\theta}}$ is called maximum likelihood method and the term $\ell(\hat{\boldsymbol{\theta}}) = \sum_{\alpha=1}^{n} \log f(x_\alpha|\hat{\boldsymbol{\theta}})$ is called maximum log-likelihood. If the likelihood function is continuously differentiable, the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ is provided by (Konishi & Kitagawa, 2008)

$$\frac{\partial\ell(\boldsymbol{\theta})}{\partial\theta_i} = 0, \quad i = 1, 2, ..., p \quad and \quad \frac{\partial\ell(\boldsymbol{\theta})}{\boldsymbol{\theta}} = \mathbf{0} \tag{18}$$

being θ_i elements of the vector $\boldsymbol{\theta}$.

3.2 Akaike Information Criterion for State Space Models

Most problem in the statistical inference can be considered as problems related to statistical modeling. They are formulated as comparisons between different statistical models. We consider to use the AIC to statistical inference problems such as verification distributions, order selection of regression models and statistical determination of the model . This information criterion is defined by,

$$AIC = -2(\text{maximum log-likelihood}) - 2(\text{numbers of the free parameters}) \tag{19}$$

The numbers of free parameters p refers to the vector dimensions $\boldsymbol{\theta}$ contained in the specified model as $f(x|\boldsymbol{\theta})$. The AIC is an evaluation criterion for the models whose parameters are estimated by the maximum likelihood method and indicates that the bias of maximum log-likelihood function is approximately equal to the “number of free parameters of the model”. The bias is obtained under the assumption that the true model is contained in the parametric model specified $f(x|\boldsymbol{\theta})$(Konishi & Kitagawa, 2008).
According (Cavanaugh, 1993) Log-likelihood function applied to state space models can be written as

$$L(\boldsymbol{\theta}|Y_n) \propto -\frac{1}{2}\sum_{t=1}^{n} ln|\Sigma_t(\boldsymbol{\theta})| - \frac{1}{2}\sum_{t=1}^{n} \boldsymbol{e}_t^T(\boldsymbol{\theta})\Sigma^{-1}(\boldsymbol{\theta})\boldsymbol{e}_t(\boldsymbol{\theta}) \tag{20}$$

Using the maximum likelihood method, the parameters can be obtained as follows

$$\frac{\partial L(\boldsymbol{\theta}|Y_n)}{\partial \theta_i} = -\frac{1}{2}\sum_{t=1}^{n}\{tr\{[\Sigma_t^{-1}(\boldsymbol{\theta})\frac{\partial \Sigma_t(\boldsymbol{\theta})}{\partial \theta_i}](\boldsymbol{I} - \Sigma_t^{-1}(\boldsymbol{\theta})\boldsymbol{e}_t(\boldsymbol{\theta})\boldsymbol{e}_t^T(\boldsymbol{\theta}))\} \\ - \frac{\partial \boldsymbol{e}_t^T(\boldsymbol{\theta})}{\partial \theta_i}\Sigma_t^{-1}(\boldsymbol{\theta})\boldsymbol{e}_t(\boldsymbol{\theta})\} \tag{21}$$

The next figure 1 shows a block diagram intuitively, which is a way to evaluate a state space model by AIC.

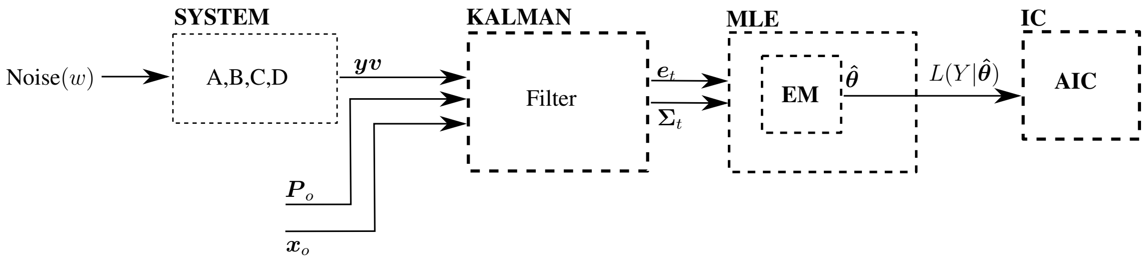

Figure 1: Blocks diagram for evaluate the AIC in State Space.

4 Example

The purpose of this example is to make it easier to understand the concepts presented in this paper and its implementation, considering the Akaike Information Criterion applied in a state space model.

If the idea is to evaluate the different order models with AIC, the steps to follow are:

1. Select the order models that you think it is the real system.
2. You can initialize the kalman filter to obtain the covariance matrices $\boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{P}, \boldsymbol{\Sigma}$ and the vectors $\boldsymbol{x}_t$, $\boldsymbol{e}_t$.
3. Get the parameters unknown (Shumway & Stoffer, 2010) by EM algorithm: $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots \boldsymbol{\theta}_p)$,
4. With the parameters estimated, obtained the Log-likelihood: equation(20)
5. Derivate the log-likelihood function with respect of each parameters estimated: equation(21)
6. Get the Optimum parameters that minimize the log-likelihood function.
7. Replace in the AIC formula: equation(19)
8. Repeat these steps to the other models and compare the AIC.
9. The model with the smallest AIC is the best suitable to the real system between sets of models for comparison.

If we know a system with the parameters:

$$A = \begin{pmatrix} 1.1269 & -0.4940 & 0.1129 \\ 1.0000 & 0 & 0 \\ 0 & 1.0000 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} -0.3832 \\ 0.5919 \\ 0.5191 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \qquad D = \begin{pmatrix} 0 \end{pmatrix}$$

We can simulate the output of this system on Matlab using a random input. For a sample of 1000 observation, we can use the Matlab function *estimate* to obtain all of parameters unknown and the AIC and then we can compare numerically and graphically the output in the system.

If we consider the ***order 2***, the parameters estimated are

$$\hat{A} = \begin{pmatrix} -0.0175 & 1.0397 \\ 0.3184 & 0.0696 \end{pmatrix} \qquad \hat{B} = \begin{pmatrix} 0.2352 \\ 1.4536 \end{pmatrix}$$

$$\hat{C} = \begin{pmatrix} 0.4273 & 0.9505 \end{pmatrix} \qquad \hat{D} = \begin{pmatrix} -0.0202 \end{pmatrix}$$

Thus the output generated from these optimal parameters for this order is as shown in figure 2. The x-axis for better viewing was limited.

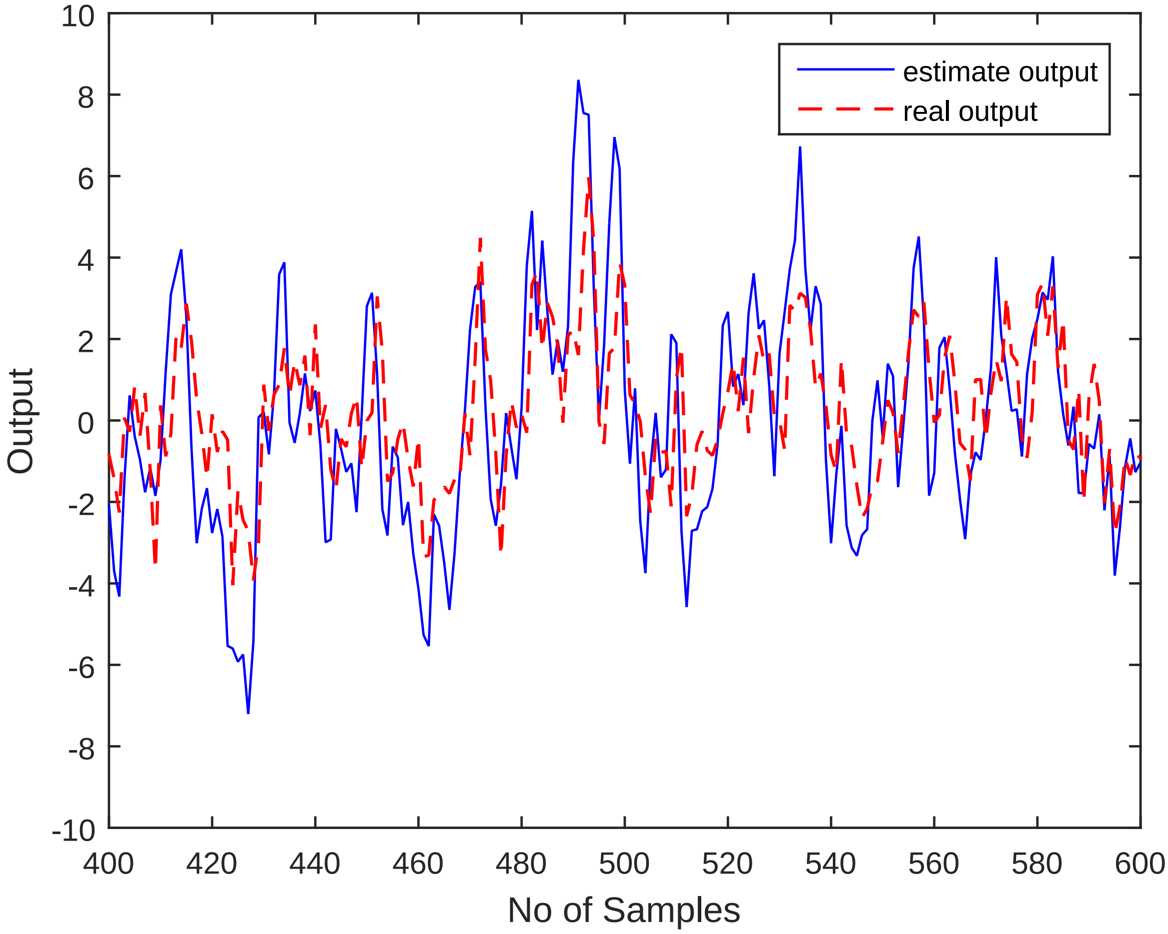

Figure 2: Output model of order 2 with estimated parameters

If we consider the ***order 3***, the parameters estimated are

$$\hat{A} = \begin{pmatrix} 0.1425 & 0.3851 & 0.0870 \\ 0.4739 & 0.1550 & -0.1848 \\ 0.2555 & 0.8432 & 0.8009 \end{pmatrix} \qquad \hat{B} = \begin{pmatrix} 1.2310 \\ 1.8070 \\ 0.4445 \end{pmatrix}$$

$$\hat{C} = \begin{pmatrix} 0.8005 & 0.1055 & 0.0333 \end{pmatrix} \qquad \hat{D} = \begin{pmatrix} 0.7441 \end{pmatrix}$$

Thus the output generated from these optimal parameters is as shown in figure 3. The x-axis for better viewing was limited.

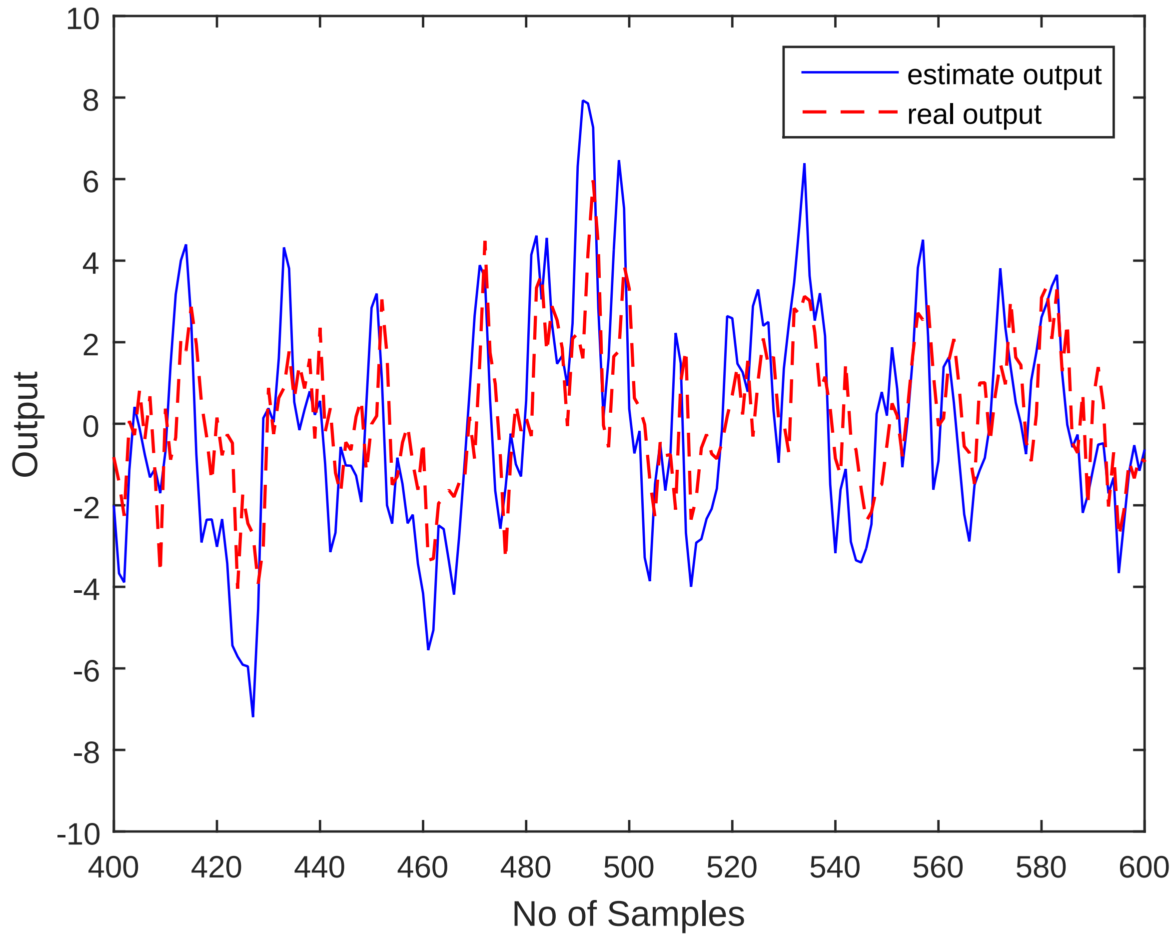

Figure 3: Output model of order 3 with estimated parameters

If we consider the ***order 4***, the parameters estimated are

$$\hat{A} = \begin{pmatrix} 0.0812 & -0.4306 & 0.4653 & 0.4672 \\ 0.7681 & 0.1003 & 0.0605 & -0.1916 \\ 0.6157 & -0.4962 & 0.3373 & -0.0893 \\ 0.4320 & 0.7663 & 0.8354 & -0.3496 \end{pmatrix} \qquad \hat{B} = \begin{pmatrix} 0.8599 \\ 0.3921 \\ 0.8654 \\ -0.0925 \end{pmatrix}$$

$$\hat{C} = \begin{pmatrix} 0.8058 & 0.2636 & 0.0507 & 0.3562 \end{pmatrix} \qquad \hat{D} = \begin{pmatrix} 0.9645 \end{pmatrix}$$

Thus the output generated from these optimal parameters is as shown in figure 4. The x-axis for better viewing was limited.

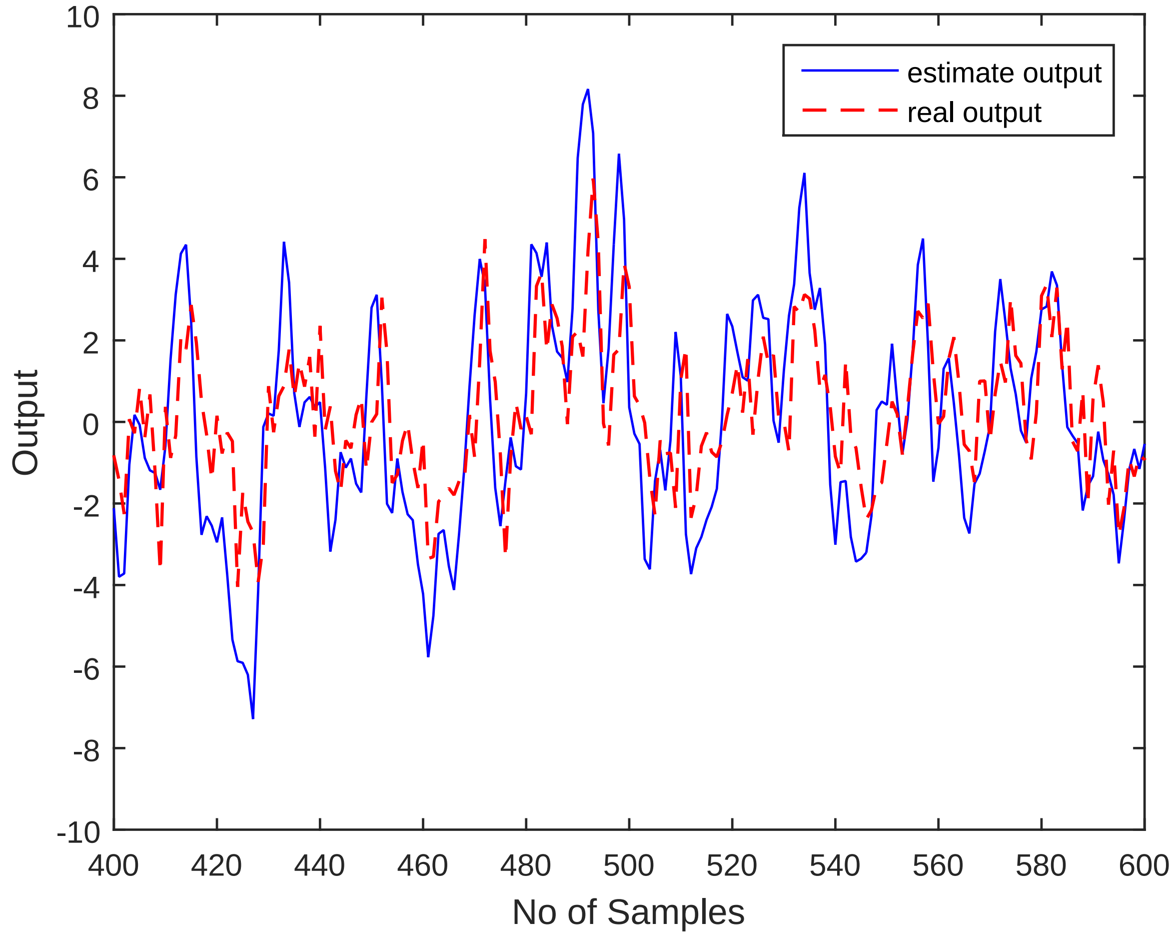

Figure 4: Output model of order 4 with estimated parameters

If we consider the ***order 5***, the parameters estimated are

$$\hat{A} = \begin{pmatrix} -0.4471 & 0.0099 & 0.8376 & 0.3876 & -0.9261 \\ 0.0875 & 0.2629 & -0.1276 & 0.1947 & 0.7233 \\ -0.1021 & 0.2034 & 0.2862 & 0.5369 & 0.0828 \\ 0.2770 & -0.5580 & 0.3741 & -0.1295 & 0.1082 \\ 0.7405 & 0.3794 & 0.3395 & -0.2135 & 0.1294 \end{pmatrix} \qquad \hat{B} = \begin{pmatrix} 0.4745 \\ 0.8428 \\ 0.6072 \\ 0.6189 \\ 0.1378 \end{pmatrix}$$

$$\hat{C} = \begin{pmatrix} 0.5128 & 0.6278 & 0.3279 & 0.4905 & 0.3448 \end{pmatrix} \qquad \hat{D} = \begin{pmatrix} 0.5298 \end{pmatrix}$$

Thus the output generated from these optimal parameters is as shown in figure 5. The x-axis for better viewing was limited.

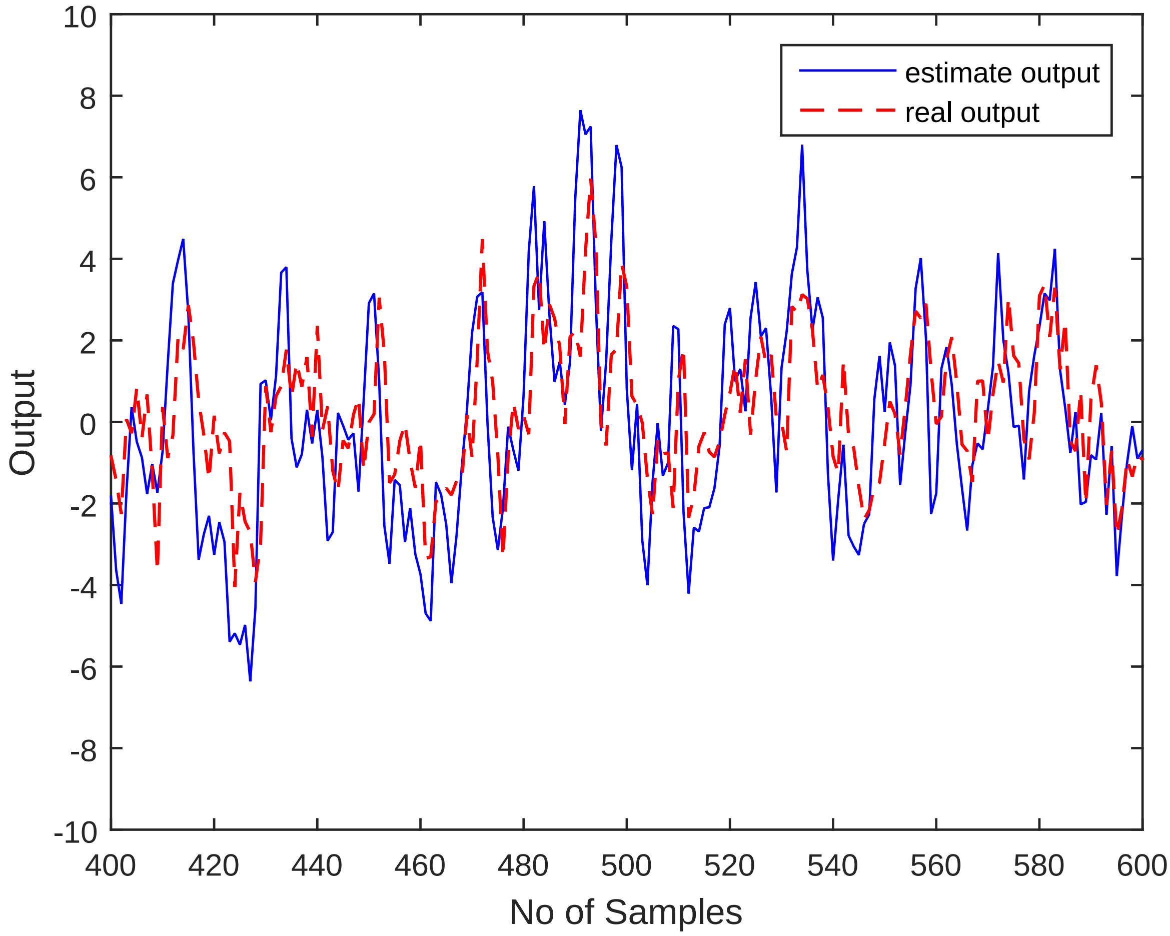

Figure 5: Output model of order 5 with estimated parameters

In the following table we can see all the Log-Likelihood, AIC and BIC measures obtained by the function "*estimate*" of Matlab software. This set of proposed models are compared to the real system, so choose the model with a smaller value of the chosen criteria.

Table 1: Measures of the criteria between the set of proposed models

	Log-Likelihood	AIC	BIC
Order 2	-1814.83	3647.65	3691.83
Order 3	-1813.76	3659.51	3738.05
Order 4	-1813.1	3676.19	3798.91
Order 5	-1809.41	3690.83	3867.54

5 Conclusions

Using the studies of (Bengtsson & Cavanaugh, 2006, Cavanaugh, 1993, Hurvich & Tsai, 1989, Mader et al., 2014), we built a mini guide through a simple didactic example, listing the steps reach the value of AIC. We can say the Akaike can be used to reliably assess the quality of state-space models, considering that need large amounts of data sampled in relation to the system order. In this paper we managed to simplify a comprehensible way to apply the AIC for a set of models in state space as well as compare it with the actual system. With the help of software can implement and get these criteria a simple method.

REFERENCE

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. *2nd International Symposium on Information Theory (Petrov, B. N. and Csaki, F., eds.)*, vol. , pp. 267–281.

Barreto, G., 2002. *Modelagem Computacional Distribuída e Paralela de Sistemas e de Séries Temporais Multivariáveis no Espao de Estado*. PhD thesis, Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

Bengtsson, T., & Cavanaugh, J. E., 2006. An improved akaike information criterion for state-space model selection. *Computational Statistics & Data Analysis*, vol. 50, n. 10, pp. 2635–2654.

Brown, R. G., & Hwang, P. Y. C., 2012. *Introduction to Random Signal and Applied Kalman Filtering with Matlab Exercises*. Wiley.

Cavanaugh, J. E., 1993. *Small-Sample Model Selection in the General State-Space Setting*. PhD thesis, University of California Davis.

Hurvich, C. M., & Tsai, C.-L., 1989. Regression and time series model selection in small samples. *Biometrika*, vol. 76, n. 2, pp. 297–307.

Konishi, S., & Kitagawa, G., 2008. *Information Criteria and Statistical Modeling*. Springer.

Mader, W., Linke, Y., Mader, M., Sommerlade, L., Timmer, J., & Schelter, B., 2014. A numerically efficient implementation of the expectation maximization algorithm for state space models. *Applied Mathematics and Computation*, vol. 241, pp. 222–232.

Schweeppe, F. C., 1965. Evaluation of likelihood functions for gaussian signals. *Information Theory, IEEE Transactions*, vol. 11, n. 1, pp. 61–70.

Shumway, R. H., & Stoffer, D. S., 2010. *Time series analysis and its applications: with R examples*. Springer Science & Business Media.