

**17th Inc. 2008 Available online at www.sciencedirect.com** 





IFAC-PapersOnLine 48-28 (2015) 086–091

### Model Validation Criteria for System Identification in Time Domain Model Validation Criteria king in Time Domain I<br>References Model Validation Criteria Model Validation Criteria Model Validation Criteria

#### Hideo Muroi *<sup>∗</sup>* and Shuichi Adachi *<sup>∗</sup>* Hideo Muroi *<sup>∗</sup>* and Shuichi Adachi *<sup>∗</sup>* Hideo Muroi *<sup>∗</sup>* and Shuichi Adachi *<sup>∗</sup>* Hideo Muroi *<sup>∗</sup>* and Shuichi Adachi *<sup>∗</sup>*

*<sup>∗</sup> Faculty of Science and Technology, Keio University, Yokohama, 223-8522, Japan (e-mail: muroi@z8.keio.jp). ∗ Faculty of Science and Technology, Keio University, Yokohama,*<br>*∗ Faculty of Science and Technology, Keio University, Yokohama, 223-8522, Japan (e-mail: muroi@z8.keio.jp). 223-8522, Japan (e-mail: muroi@z8.keio.jp).*

system identification in the time domain. This criterion is given by the root mean squared error (RMSE) divided by the standard deviation of the measured signal. However, the fit ratio has some problems. For example, it can take negative values because it is not normalized, and it is easy to obtain a better value for a low-amplitude signal than for a high-amplitude signal. In this paper, we introduce some normalized criteria from the field of physical geography and consider criteria that resolve these paper, we introduce some normalized criteria from the field of physical geography and consider criteria that resolve these problems. We evaluated these criteria through two case studies. We found that the correlation coefficient was effective to evaluate the phase, and a criterion obtained from the triangle inequality was effective for evaluating the gain and phase. Abstract: The fit ratio is one of the most commonly used criteria to evaluate a result of Abstract: The fit ratio is one of the most commonly used criteria to evaluate a result of system identification in the time domain. This criterion is given by the root mean squared error (RMSE) divided by the standard deviation of the measured signal. However, the fit ratio has some problems. For example, it ca

© 2015, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved. from the triangle income.<br>Commonly the triangle for the second the second phase in the gain and phase. The second the second the gain an<br>integrating the second the second phase in the second phase in the second the second

*Keywords:* System identification, Validation, Time domain, Error criteria, Correlation coefficients coefficients coefficients coefficients *Keywords:* System identification, Validation, Time domain, Error criteria, Correlation Keywords: *Keywords:* System identification, Validation, Time domain, Error criteria, Correlation

# 1. INTRODUCTION 1. INTRODUCTION 1. INTRODUCTION 1. INTRODUCTION 1. INTRODUCTION

Recently, mathematical models have been used for the design of control systems even in the field of industrial applications. The mathematical models are often constructed by physical modeling or system identification. In this paper, we focus on system identification. Once the models are obtained, it is necessary to validate the some aspects of them. This step is called model validation. For example, Ljung (1998) noted that Ljung (1998) noted that Ljung (1998) noted that Recently, mathematical models have been used for the design of control systems even in the field of industrial ap-design of control systems even in the field of industrial ap-Recently, mathematical models have been used for the design of control systems even in the field of industrial ap-<br>plications. The mathematical models are often constructed by physical modeling or system identification. In this pa-plications. The mathematical models are often constructed by physical modeling or system identification. In this pa-<br>per, we focus on system identification. Once the models per, we focus on system fuentification. Once the models per, we focus on system identification. Once the models<br>are obtained, it is necessary to validate the some aspects<br>of them. This step is called model validation. For example

- simulation in the time domain, *•* simulation in the time domain, *•* simulation in the time domain, *•* simulation in the time domain,
- comparison of some methods in the frequency domain, and main, and main, and • simulation in the time domain,<br>• comparison of some methods in the frequency do-• comparison or some methods in the frequency do-<br>main and
- *•* pole and zero cancellation in the *s* (or *z*) domain main, and<br>
• pole and zero cancellation in the  $s$  (or  $z$ ) domain *•* pole and zero cancellation in the *s* (or *z*) domain

are often checked. Also, independence between the residuals and the past input is usually tested using the sample covariance. If the model is constructed by physical modeling, the feasibility of the physical parameters is often checked and is used to evaluate. checked and is used to evaluate. are often checked. Also, independence between the reside-between the reside-between the reside-between the reside-between the residence between the residence between the residence between the residence between the residen are often checked. Also, independence between the residare often checked. Also, independence between the resid-are often checked. Also, independence between the residuals and the past input is usually tested using the sample are often checked. Also, independence between the residuals and the past input is usually tested using the sample<br>covariance. If the model is constructed by physical modeling, the feasibility of the physical parameters is often eling, the feasibility of the physical parameters is often

Simulation in the time domain is the most useful and intuitive validation method for evaluating the goodness of the model. Mean squared error (MSE) or root mean squared error (RMSE) are often used for the comparison criteria. Because these criteria are based on a physical quantity, their orders can be significantly different, and they do not seem to be suitable for discussion in general terms about comparison of two waveforms.  $S_{\rm{max}}$  is the time domain is the most useful and  $\alpha$ Simulation in the time domain is the most useful and covariance. It the model is constructed by physical modeling, the feasibility of the physical parameters is often checked and is used to evaluate.<br>Simulation in the time domain is the most useful and intuitive validation m Simulation in the time domain is the most useful and intuitive validation method for evaluating the goodness Simulation in the time domain is the most useful and of the model. Mean squared error (MSE) or root mean intuitive validation method for evaluating the goodness squared error (RMSE) are often used for the comparison squared error (RMSE) are often used for the comparison squared error (RMSE) are often used for the comparison criteria. Because these criteria are based on a physical criteria. Because these criteria are based on a physical criteria. Because these criteria are based on a physical quantity, their orders can be significantly different, and quantity, their orders can be significantly different, and they do not seem to be suitable for discussion in general they do not seem to be suitable for discussion in general quantity, their orders can be significantly different, and of the model. Mean squared error (MSE) or root mean they do not seem to be suitable for discussion in general

Therefore, a criterion based on a dimensionless quantity that does not depend on physical quantity appears to be required for simulation. Some criteria based on dimensionless quantities (called dimensionless criteria in this  $\mu$  paper) have been proposed in a field of physical geography (especially hydrology and climatology) since the 1970s. First of all, Nash and Sutcliffe (1970), and Garrick et al. That of all, Nash and Sutchile (1970), and Garrick et al.<br>(1978) proposed the coefficient of efficiency  $E$  to evaluate Therefore, a criterion based on a dimensionless  $\mathcal{L}_\text{max}$ Therefore, a criterion based on a dimensionless quantity Therefore, a criterion based on a dimensionless quantity that does not depend on physical quantity appears to be Therefore, a criterion based on a dimensionless quantity required for simulation. Some criteria based on dimen-that does not depend on physical quantity appears to be sionless quantities (called dimensionless criteria in this sionless quantities (called dimensionless criteria in this<br>paper) have been proposed in a field of physical geography (especially hydrology and climatology) since the 1970s. (especially hydrology and climatology) since the 1970s. (especially hydrology and chinatology) since the 1970s.<br>First of all, Nash and Sutcliffe (1970), and Garrick et al. required for simulation. Some criteria based on dimenpaper) have been proposed in a field of physical geography the rainfall runoff model. The coefficient of efficiency is sometimes called the Nash-Sutcliffe model efficiency coefficient (NSE). This criterion is given by the MSE divided by the variance of the observed values, and it takes a maximum value when the predicted values are identical to the observed values. However, it takes negative values when the MSE is greater than the variance of the measured values. Therefore, Willmott (1981) and Willmott et al.  $(1985)$  proposed the index of agreement, which normalizes the MSE to a velue from 0 to 1. the MSE to a velue from 0 to 1. the MSE to a velue from 0 to 1. the rainfall runoff model. The coefficient of efficiency is the rainfall runoff model. The coefficient of efficiency is<br>sometimes called the Nash-Sutcliffe model efficiency coefsometimes called the Negle Cotton and definition of the subset sometimes can ed the Nash-Sutchine model efficient (NSE). This criterion is given by the MSE divided by the variance of the observed values, and it takes a by the variance of the observed values, and it takes a ficient (NSE). This criterion is given by the MSE divided maximum value when the predicted values are identical maximum value when the predicted values are identical by the variance of the observed values, and it takes a to the observed values. However, it takes negative values to the observed values. However, it takes negative values to the observed values. However, it takes negative values<br>when the MSE is greater than the variance of the measured when the MSE is greater than the variance of the measured<br>values. Therefore, Willmott (1981) and Willmott et al. values. Therefore, whillfull (1981) and whillfull et al.<br>(1985) proposed the index of agreement, which normalizes maximum value when the predicted values are identical values. Therefore, Willmott (1981) and Willmott et al.

In the field of system identification, the fit ratio (sometimes called the NRMSE fitness value), which is the RMSE divided by the standard deviation of the measured values, is often used for evaluation. This criterion is implemented<br>in System Identification Toolbox for MATLAB<sub>*®*</sub> is often used to evaluation. This criterion is implemented<br>in System Identification Toolbox for MATLAB® (Mathworks.co.jp, 2015). However, similarly to the coefficient of efficiency, this criterion takes values from *−∞* to 100, works.co.jp, 2015). However, similarly to the coencient<br>of efficiency, this criterion takes values from  $-\infty$  to 100,<br>making it unquitable for use as an intuitive exitence for or enteriency, this criterion takes values from  $-\infty$  to 100,<br>making it unsuitable for use as an intuitive criterion for evaluation. evaluation. evaluation. In the field of system identification, the fit ratio (some-In the field of system identification, the fit ratio (some-In this paper, we introduce some normalized criteria from In the field of system identification, the fit ratio (some-It is called the NRMSE fitness value), which is the RMSE<br>times called the NRMSE fitness value), which is the RMSE divided by the standard deviation of the measured values, divided by the standard deviation of the measured values, is often used for evaluation. This criterion is implemented is often used for evaluation. This criterion is implemented m system Identification Toolbox for MATLAB<sub>*W*</sub> (Math-<br>works.co.jp, 2015). However, similarly to the coefficient In the field of system identification, the fit ratio (some-

In this paper, we introduce some normalized criteria from the field of physical geography, and consider criteria that resolve the fit ratio's problem. Then we evaluate the criteria using some case studies. criteria using some case studies. In this paper, we introduce some normalized criteria from In this paper, we introduce some normalized criteria from In this paper, we introduce some normalized criteria from resolve the ft retic's problem. Then we evaluate the criteria using some case studies. the field of physical geography, and consider criteria thatresolve the fit ratio's problem. Then we evaluate the In this paper, we introduce some normalized criteria from<br>the field of physical geography, and consider criteria that<br>resolve the fit ratio's problem. Then we evaluate the<br>criteria using some case studies.<br>2. UNNORMALIZED

#### 2. UNNORMALIZED DIMENSIONLESS CRITERIA 2. UNNORMALIZED DIMENSIONLESS CRITERIA 2. UNNORMALIZED DIMENSIONLESS CRITERIA 2. UNNORMALIZED DIMENSIONLESS CRITERIA

We introduce two unnormalized criteria, that is, the coefficient of efficiency and the fit ratio. Then, we show the problems of these criteria. problems of these criteria. problems of these criteria. We introduce two unnormalized criteria, that is, the co-We introduce two unnormalized criteria, that is, the cowe introduce two unnormalized criteria, that is, the co-<br>efficient of efficiency and the fit ratio. Then, we show the efficient of efficiency and the fit ratio. Then, we show the

# *2.1 Coefficient of efficiency 2.1 Coefficient of efficiency 2.1 Coefficient of efficiency 2.1 Coefficient of efficiency*

The coefficient of efficiency (hereinafter denoted by *E*) is defined by defined by defined by defined by The coefficient of efficiency (hereinafter denoted by *E*) is The coeffic The coefficient of efficiency (hereinafter denoted by  $E$ ) is

2405-8963 © 2015, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved. Peer review under responsibility of International Federation of Automatic Control.<br> **Control.** 10.1016/j.ifacol.2015.12.105

$$
E = 1 - \frac{\sum_{k=1}^{N} (y(k) - \hat{y}(k))^2}{\sum_{k=1}^{N} (y(k) - \bar{y})^2},
$$
\n(1)

where  $y(k)$  and  $\hat{y}(k)$  are the measured output and the simulated output at time k, respectively,  $\overline{y}$  is the average value of *y*, and *N* is the number of samples. *E* takes a value of 1 when the MSEs of  $y$  and  $\hat{y}$  are equal to 0, and  $E$ takes a negative value when the MSE is greater than the variance of the measured output. Furthermore, *E* takes a value of *−∞* when the simulated output diverges. In this way, *E* does not intuitively appear to be associated with the similarity between  $y$  and  $\hat{y}$ .

However, *E* was recommended as a criterion for evaluating the rainfall runoff model by the ASCE Task Committee (1993) and Legates and McCabe (1999) because it has a meaningful interpretation. One interpretation is that if  $E \leq 0$ , the mean measured value is a better predictor than the simulated value. Additionally, Moriasi et al. (2007) claimed that if  $E > 0$ , the model can be viewed as having acceptable performance. Thus, we can conclude that

(1)  $E \leq 0$  : model is not acceptable,

(2)  $0 < E \leq 1$  : model is acceptable.

In this paper, we call this Moriasi's rule.

*2.2 Fit ratio*

The fit ratio (hereinafter denoted by FIT) is defined by

$$
FIT = \left\{ 1 - \frac{\sqrt{\sum_{k=1}^{N} (y(k) - \hat{y}(k))^2}}{\sqrt{\sum_{k=1}^{N} (y(k) - \bar{y})^2}} \right\} \times 100\% . (2)
$$

When  $y(k)$  is identical to  $\hat{y}(k)$  for all  $k = 1, ..., N$ . FIT becomes 100. Since in the second term the RMSE is normalized by the standard deviation of *y*, FIT can take negative values when the RMSE is greater than the standard deviation of *y*, similarly to *E*, i.e.,

$$
\sqrt{\sum_{k=1}^{N} (y(k) - \hat{y}(k))^2} > \sqrt{\sum_{k=1}^{N} (y(k) - \bar{y})^2}.
$$

In particular, if the RMSE tends to *∞*, FIT tends to *−∞*. This property of FIT appears to make it unsuitable for the evaluation of system identification.

## *2.3 Problems of these criteria*

In this section we show through an example that *E*, FIT, and Moriasi's rule are not appropriate for evaluating the result of system identification. As an illustrative example, let us consider the Linear Time-Invariant (LTI) system:

$$
y(k) = G(q)u(k),\tag{3}
$$

where  $u(k)$  and  $y(k)$  are the input and output at time k, respectively,  $G(q)$  denotes a discrete time transfer function, and *q* denotes a shift operator. As the identification input



Fig. 1. Waveforms in the example: gain shift case (black solid line:  $y(k)$ , Blue dashed line:  $\hat{y}_+(k)$ , Red dashed line:  $\hat{y}_-(k)$ 



Fig. 2. Waveforms in the example: phase shift case (black solid line:  $y(k)$ , blue dashed line:  $y_p(k)$ , Red dashed line:  $y_{p-}(k)$ )

*u*, two periods of a pseudo-random binary signal (PRBS) with a period of 1023 are used.

We assume that true system is given by

$$
G(q) = \frac{q^{-1} + 0.5q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}}.
$$

Now we assume that following four models are estimated:

$$
\hat{G}_{+}(q) = 2.01G(q), \quad \hat{G}_{-}(q) = 0.49G(q), \n\hat{G}_{p}(q) = G(q)q^{-3}, \quad \hat{G}_{p-}(q) = 0.49G(q)q^{-3},
$$

where the first two models  $G_+$  and  $G_-\$  are the same except for their gains, and the latter two models  $\hat{G}_p$  and  $\hat{G}_p$ <sup>−</sup> are the same except for their gains. We apply *u* to the models, and we obtain the following four simulated outputs:

$$
\hat{y}_{+}(k) = \hat{G}_{+}(q)u(k), \quad \hat{y}_{-}(k) = \hat{G}_{-}(q)u(k),
$$
  
\n $\hat{y}_{p}(k) = \hat{G}_{p}(q)u(k), \quad \hat{y}_{p-}(k) = \hat{G}_{p-}(q)u(k).$ 

Moreover, if  $\hat{G} = 0$ ,  $\hat{y}$  takes a value of 0 all the time. We call this output "zero output". We calculate *E* and FIT for four simulated outputs and the zero output.

Figure 1 shows the true output *y* (black solid line) and the simulated outputs  $\hat{y}_+$  and  $\hat{y}_-$  (blue and red dashed lines, respectively) for 0 to 200 samples. In the same way, Fig.2 shows the true output *y* and the simulated outputs  $\hat{y}_p$  and  $\hat{y}_{p-}$  for 0 to 200 samples. Table 1 shows the values of *E* and FIT and the acceptance results according to Moriasi's rule. *E* and FIT take negative values for  $\hat{y}_+$  having about twice the amplitude of *y*, but they take positive values for *y*<sup> $j$ </sup>− having about half the amplitude of *y*. Similarly, they take negative values for  $\hat{y}_p$  having three taps delay and

Download English Version:

# <https://daneshyari.com/en/article/711154>

Download Persian Version:

<https://daneshyari.com/article/711154>

[Daneshyari.com](https://daneshyari.com)