RELIABILITY OF PARAMETRIC VARIANCE ESTIMATES FOR IDENTIFIED TRANSFER FUNCTIONS

J. Schoukens (*), P.M.J. Van den Hof (**), R. Pintelon (*)

(*): Vrije Universiteit Brussel, dep. ELEC, Pleinlaan 2, B1050 Brussels, Belgium
(**): Delft Center for Systems and Control, Delft University of Technology, NL-2628 CD Delft, The Netherlands
email: Johan.Schoukens@vub.ac.be

Abstract: Classical system identification methods result in an identified parametric plant and noise model. Due to modelling errors, the estimated variance of the plant model can fail completely to give a reasonable idea of the amplitude of the model errors while this is not detected at all in the validation test. A robust method is presented to validate the parametric variance estimate, and to indicate the reliable frequency bands where the estimated variance can be safely used as an indication of the remaining model errors. Outside these bands the model is unreliable (neither validated or invalidated).

Keywords**:** system identification, model validation, variance estimates.

1. Introduction

Consider the discrete time system

$$y(t) = G_0(q)u_0(t) + H_0(q)e(t) = y_0(t) + v(t). \quad (1)$$

Starting from the observed input and output a parametric plant and noise model $\hat{G}(q, \theta), \hat{H}(q, \theta)$ is identified, and eventually also the standard deviation $\hat{\sigma}_G(q, \theta)$ is obtained, using one of the classical identification schemes (Ljung, 1999; Söderström and Stoica). Once these models have passed the validation step (e.g. cross-correlation test input-residue; autocorrelation test residue) it is very tempting to believe that these models are close to the real underlying system characteristics (Douma *et al.*, 2005). Quite a lot of results are available on the study of the plant model errors $\hat{G}(q, \theta) - G_0(q)$, but not much is said about the reliability of the variance estimate $\hat{\sigma}^2_G(q, \theta)$. Simple examples show that these have to be used with extreme care. $\hat{\sigma}^2_G(q, \theta)$ describes very well the variance of the estimated plant model $\hat{G}(q, \theta)$, but it should not be used as a bound on possible bias errors. $E\{\hat{G}(q, \theta)\} - G_0(q)$ can be much larger than $\hat{\sigma}_G(q, \theta)$, even for fully validated models as is shown in the next example (see Fig. 1).

A 4th order system is identified using a 4th order model (see Section 4 for more details). Also the estimated standard deviation $\hat{\sigma}_G(q, \theta)$ is plotted. From the figure it is seen that the 2nd resonance is completely

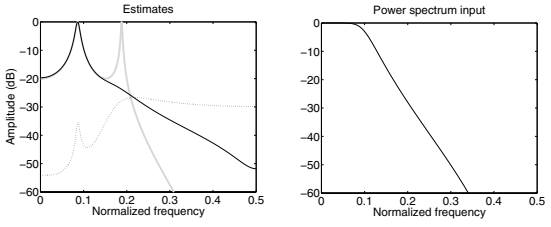


Fig. 1: Left: G_0 (gray line), $\hat{G}$ (black line), and $\hat{\sigma}_G$ (broken line);
Right: Power spectrum of the input

missed, and this is not indicated at all by the variance estimate. This poor model passes the validation tests because the second resonance was not well excited as can be seen from the right plot in Fig. 1, and so there is almost no evidence in the data for its presence. Beyond 0.1 Hz, the system is badly excited, so all results shown above this frequency are mainly extrapolations of the models (plant- and variance-model) obtained below 0.1 Hz. It is clear that such extrapolation results should not be relied on. In a frequency band without excitation, no conclusions should be made. In a well excited band, the identification results are reliable. However, in practice there are many situations where the problem is not that clear cut. In between both extremes, there is a grey zone where it is not clear at all how much the user can rely on the estimates.

The aim of this paper is to provide a simple and robust criterion that indicates the reliability of the variance estimate $\hat{\sigma}^2_G(q, \theta)$ as a function of the frequency, and next to indicate in what frequency bands the model is unreliable (it might be good, it might be bad).

As such this approach is different from the 'model error modelling' approach (Reinelt *et al.*, 2002) where an aditional more flexible model is identified between the input and the residuals. The original plant model is then considered to be validated in those frequency regions where this second error model equals zero within its own uncertainty bounds. There are two weak points in this method when considered from the point of view of this paper. Firstly the construction of the uncertainty bounds around the error model relies again on an identified noise model; secondly, such an approach will still not distinguish between well excited frequency regions and frequency bands where the extrapolation effect dominates due to a lack of excitation. The first remark is removed in the nonparametric model error modelling approach (Stenman and Tjarnstrom, 2000), where also a nonparametric uncertainty bound is extracted based on a local polynomial modelling of the smoothed empirical transfer function estimate ETFE (Ljung, 1999). The variance is then directly computed from the residuals.

Another alternative approach is described in Hakvoort and Van den Hof (1997). In this case uncertainty bounds are generated that include the effects of unmodelled dynamics, but in order to do so an exponentially decreasing bound on the impulse response is needed. The noise characterization is done using an analysis of the residuals of a nominal model on a validation set.

The outline of the paper consist of 3 parts. First the system setup is described. Next the method to detect 'local' extrapolation problems is explained. This leads to a simple criterium to inform the user about frequency regions where the model variance estimate and hence also the plant model is not reliable. Eventually the method is illustrated on a simulation.

2. Setup

A single input, single output system as described in (1) is considered. $e(t)$ is a sequence of independent random variables with zero mean values and variances λ.

The input and the output are observed:

$$u_0(t), y(t), \; t = 1, \ldots, N. \tag{2}$$

Later on, a non parametric noise model will be estimated. The method that will be applied restricts the input to be random with a dependency length that is (significantly) shorter than the length of the experiment.

Assumption 1. The excitation $u_0(t)$ is a random excitation with a dependency length that is significantly shorter than the experiment length.

Some calculations will be done in the frequency domain. The discrete Fourier transform (DFT) $U_0(l), Y(l)$ of the input/output signal (Brigham, 1974) will be used:

$$X(l) = \frac{1}{N}\sum_{k=0}^{N-1} x(k) e^{-j\frac{2\pi}{N}kl}. \tag{3}$$

where the frequency index l corresponds to the frequency $f_l = 2\pi l/N$. Due to the central limit theorem, the noise on the DFT of the output will be asymptotically normal and independently distributed (Pintelon and Schoukens, 2001).

3. Method

The basic idea to estimate the reliability of $\hat{\sigma}^2_G(q, \theta)$ is to measure the 'local' information on G. To do so four steps are needed.

i) A sliding rectangular frequency window is considered. Only the data in this window (local data) are considered; ii) On these local data, a zero order model $\hat{G}_{\mathrm{w}}$ is estimated; iii) A non parametric noise analysis is made to estimate the variance σ^2_Y as a function of the frequency; iiii) The variance $\sigma^2_{\hat{G}_{\mathrm{w}}}$ is calculated.

If the modelled variance $\hat{\sigma}^2_G(q, \theta)$ is below $\sigma^2_{\hat{G}_{\mathrm{w}}}$, then there is not enough local information to support the parametric variance. In the reverse situation $\hat{\sigma}^2_G(q, \theta)$ is reliable, and will give a good idea of the errors on $\hat{G}(q, \theta)$ at that frequency. Each of these steps will be first discussed in detail, next the choice of the window width will be considered.

3.1 'Local' information using a sliding window

In Fig. 2, a system is measured using an excitation with power spectrum S_{uu}. The dashed lines define a frequency window centred around f_c with width B. The 'local' information is then the spectral information that is cut out by this window.

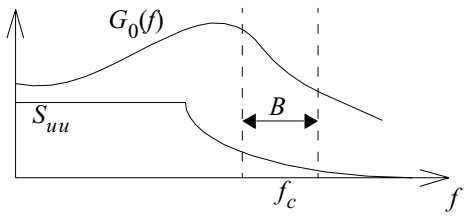


Fig. 2: The concept of 'local' information

Definition 2. Local spectral window $w(f_c)$
The spectral window $w(f_c)$ is 1 if $|f-f_c| \le B/2$ and zero elsewhere.

Definition 3. Local information
Consider the system (1) with data (2). The information $U(k)w(f_c)$ and $Y(k)w(f_c)$ that is available in the local spectral window is called the local information.

3.2 'Local' transfer function model

We are looking for the smallest uncertainty that can be obtained from the local information for the system G_0 at frequency f_c. This is given by the uncertainty of a zero order model estimate $G_{\mathrm{w}}(f_c)$. The least squares estimate of the latter one is:

$$\hat{G}_{\mathrm{w}}(f_c) = \frac{\sum Y(k)\bar{U}(k)}{\sum U(k)\bar{U}(k)}, \tag{4}$$

where the sum is taken over all frequencies f_k belonging to the window $w(f_c)$.

The variance on this estimate is

$$\sigma^2_{\hat{G}_{\mathrm{w}}}(f_c) = \frac{\sum \sigma^2_Y(k)|U(k)|^2}{\sum |U(k)|^4}, \tag{5}$$

where $\sigma^2_Y(k)$ is the variance on the Fourier spectra $Y(k)$.

Remarks:

i) It is clear that $\hat{G}_{\mathrm{w}}(f_c)$ will be a biased estimate of $G_0(f_c)$ if the window width B is large compared to the variations of G_0 around f_c. However, this does not disturb us because we are not interested in $G_{\mathrm{w}}(f_c)$, we only want to use the variance $\sigma^2_{\hat{G}_{\mathrm{w}}}(f_c)$. Increasing the model order will also increase the variance. As such (5) is a lower bound for it.

ii) If B is small enough, the variance $\sigma^2_Y(k)$ will be almost constant. At that moment the estimate (5) becomes the maximum Likelihood estimate.

3.3 Nonparametric noise analysis

In order to calculate (5), the variance $\sigma^2_Y(k)$ should be known. The variance estimate obtained from the identified noise model $H(q, \theta)$ can not be used here, because it can be strongly influenced by model errors. For that reason a non parametric method is needed that does not critically depend on the identified plant model. In Schoukens et. al. (2004) it is shown that a non parametric noise variance estimate can be obtained directly from the discrete Fourier transforms $U(k)$, $Y(k)$ under Assumption 1. It is given by:

$$\hat{\sigma}^2_Y(f_c) = \frac{d_{YY}}{d_{UU}}, \text{ with} \tag{6}$$

$$\begin{aligned} d_{YY} = \sum |&(Y(k+2)-Y(k+1))(U(k+1)-U(k)) \\ &-(Y(k+1)-Y(k))(U(k+2)-U(k+1))|^2 \end{aligned} \tag{7}$$

$$\begin{aligned} d_{UU} = \sum &|U_0(k)-U_0(k+1)|^2 \\ &+|U_0(k+1)-U_0(k+2)|^2, \\ &+|U_0(k+2)-U_0(k)|^2 \end{aligned} \tag{8}$$

where all the sums run over the frequencies in the window $w(f_c)$. The error on this estimate drops as an $O(N^{-1})$.

3.4 Validation of $\hat{\sigma}^2_G(q, \theta)$

Once the local minimum uncertainty bound (5) is available, it becomes quite easy to validate the parametric uncertainty estimate $\hat{\sigma}^2_G(q, \theta)$. If this value is significantly smaller than (5), it is unreliable because it heavily relates on the 'extrapolated' plant model. There is not enough local information available to invalidate the model during the validation step at the level of the estimated uncertainty. Model errors will be difficult to detect in that frequency band with the available data. Only if $\hat{\sigma}^2_G(q, \theta)$ is in the same order of magnitude or larger than (5), it can be reliably used. The following normalized reliability indicator for the variance is defined (alternative definitions are possible):

$$\Gamma(f_c) = \sqrt{\frac{\hat{\sigma}^2_G(f_c, \theta)}{\hat{\sigma}^2_G(f_c, \theta) + \sigma^2_{\hat{G}_{\mathrm{w}}}(f_c)}}, \tag{9}$$

where $\hat{\sigma}^2_G(f_c, \theta)$ indicates the modelled variance evaluated at frequency f_c. It is clear from the definition that $\Gamma(f_c)$ is always between 0 and 1. If $\Gamma(f_c) \approx \sqrt{0.5}$, the modelled variance is reliable. Remark that there is some arbitrary decision to be made here.

3.5 Tuning the bandwidth of the 'local' information

The only user choice to be made when calculating $\Gamma(f_c)$ is the bandwidth B of the spectral window. Choosing B large will decrease $\sigma^2_{\hat{G}_{\mathrm{w}}}(f_c)$ so that smaller uncertainty levels can be used. However, at the

same time the risk of missing the impact of sharp resonances is growing. Making B small increases the spectral resolution at a cost of decreasing the reliability of the variance estimate.

The choice of B sets also the spectral resolution of the model errors that can be traced. Local model errors (for example: missing a sharp resonance peak) should have a large amplitude (larger than the local uncertainty $\sigma^2_{\hat{G}_W}(f_c)$) in order to be detected. Small model errors should be present in a wide frequency band in order to be detectable. $\sigma^2_{\hat{G}_W}(f_c)$ is a natural measure to tune the conflicting choice between resolution and sensitivity.

4. Example

The method is illustrated on the example of the introduction. A fourth order system

$$G_0(z) = \frac{\sum_{k=0}^{n_b} b_k z^{-k}}{\sum_{k=0}^{n_a} a_k z^{-k}} \tag{10}$$

$$\text{with } a = \begin{bmatrix} 1 \\ -2.4097 \\ 3.1648 \\ -2.3326 \\ 0.9215 \end{bmatrix}, \; b = \begin{bmatrix} 2.1503\times10^{-3} \\ 8.6014\times10^{-3} \\ 1.2902\times10^{-2} \\ 8.6014\times10^{-3} \\ 2.1503\times10^{-3} \end{bmatrix} \tag{11}$$

is measured in $N = 2048$ points.

The system is excited by filtered white noise (cut-off frequency at $0.1 f_s$), the filter coefficients are

$$a = \begin{bmatrix} 1 \\ -2.3695 \\ 2.3140 \\ -1.0547 \\ 0.18738 \end{bmatrix}, \; b = \begin{bmatrix} 4.8243\times10^{-3} \\ 1.9297\times10^{-2} \\ 2.8946\times10^{-2} \\ 1.9297\times10^{-2} \\ 4.8243\times10^{-3} \end{bmatrix} \tag{12}$$

The output is disturbed with white noise. A typical realization of the input and output spectrum is shown in figure 3. A fourth order parametric plant model $\hat{G}(q, \theta)$ is identified together with its uncertainty $\hat{\sigma}_G(q, \theta)$ using an output error identification scheme. During the identification two typical situations appeared. Depending upon the actual realizations of the input and the noise, either both resonances were captured, or only one resonance is found. Both results are shown in Figure 4. The estimated transfer function $\hat{G}(q, \theta)$, and the estimated model uncertainties $\hat{\sigma}_G(q, \theta)$ are shown on this plot. Observe that the estimated uncertainty varies strongly from the first fit to the second one. In figure 5 the results of the whiteness

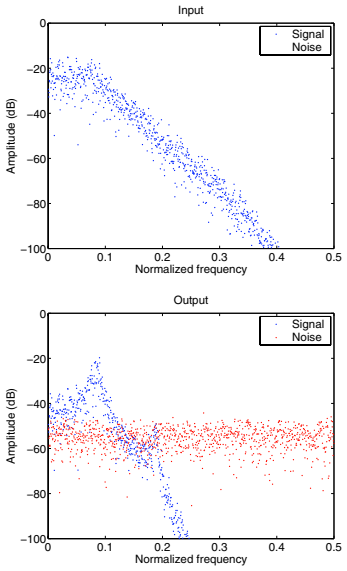


Fig. 3: Top: Amplitude spectrum of a realization of the input.
Bottom: Amplitude spectrum of a realization of the output (blue) and the disturbing noise (red)

test of the residuals, and the crosscorrelation test of the residuals with the input are shown. It can be seen that both models pass these validation tests. There is no indication at all that something is going wrong. In figure 6 the estimated uncertainty $\hat{\sigma}_G(q, \theta)$ is shown, together with $\Gamma(f_c)$ for three different resolutions as indicated by the bars. On the left figure it can be clearly seen that in the low frequency band a reliable estimate of $\hat{\sigma}_G(q, \theta)$ is obtained that also gives a good indication about the magnitude of the errors $G - G_0$. At the higher frequencies, where the excitation level sharply drops as shown in figure 1, it can be seen that $\hat{\sigma}_G(q, \theta)$ completely fails to indicate the level of the estimation error. This corresponds to a low value of $\Gamma(f_c)$ so that the user gets a warning that the results are not reliable in that frequency region. In the right figure, the situation is different. The variance gives in this case a good description of the remaining errors, also at the second resonance, as is indicated by the high value of $\Gamma(f_c)$, pointing to a high reliability.

5. Conclusions

In this paper a method is presented that allows a simple validation of the estimated uncertainty bounds and this with a minimum of user interaction. Those frequency regions where not enough local information is available to justify the calculated uncertainty bound are automatically detected. In these bands the model is unreliable (neither validated or invalidated). This protects the user against over optimistic conclusions.

6. Acknowledgement

This work was supported by the Flemish government (GOA-ILiNos), the FWO (onderzoeksgemeenschap ICCoS) and the Belgian government as a part of the

Belgian program on Interuniversity Poles of Attraction (IUAP V/22).

7. References

Brigham E.O., *The Fast Fourier Transform*. Prentice Hall, New Jersey, USA, 1974.

Douma, S.G., X. Bombois, P.M.J. Van den Hof (2005). Validity of the standard cross correlation test for model structure validation. Proc. 16th IFAC World Congress, Prague, paper We-M13-TO/6.

Hakvoort, R.G. and P.M.J. Van den Hof (1997). Identification of probabilistic system uncertainty regions by explicit evaluation of bias and variance errors. IEEE Transactions on Automatic Control, 42(100), 1516-1528.

Ljung, L. (1999). *System Identification - Theory for the user*. Prentice-Hall, Upper Saddle River, N.J., 2nd edition.

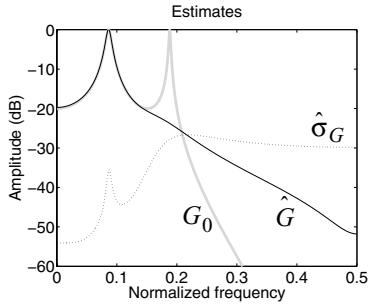
Pintelon R. and J. Schoukens (2001). *System Identification: A frequency domain approach*. IEEE Press, Piscataway.

Reinelt W., A. Garulli, L. Lung (2002). Comparing different approaches to model error modeling in robust identification. *Automatica*, 38, pp. 787-903.

Schoukens J., R. Pintelon, and Y. Rolain (2004). Box-Jenkins alike identification using non parametric noise models. *Automatica*, Vol. 40, pp. 2083-2089.

Stenman A. and F. Tjarnstrom (2000). A nonparametric approach to model error modelling. In *Proceedings of the IFAC symposium on system identification SYSID*, Santa Barbara, CA, USA, pp. WeMD1-1.

Söderström, T. and P. Stoica (1989). *System Identification*. Prentice-Hall, Englewood Cliffs.



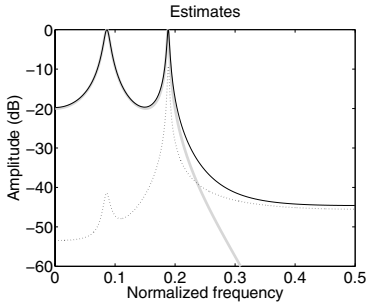
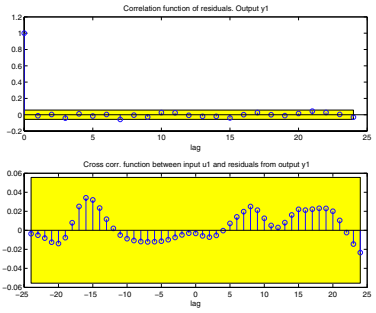


Fig. 4: Estimated plant model (full line) and $\hat{\sigma}_G$ (broken line). Gray line: G_0



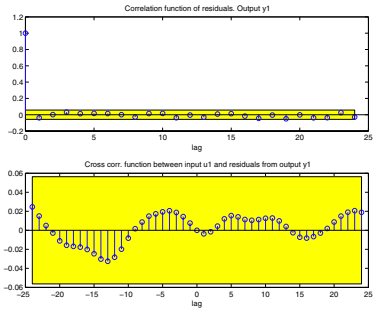
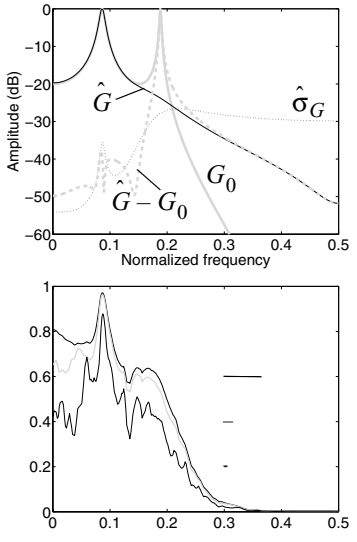


Fig. 5: Validation test for the estimated models. Top: cross-correlation of the residuals, bottom: cross-correlation between input and residuals.



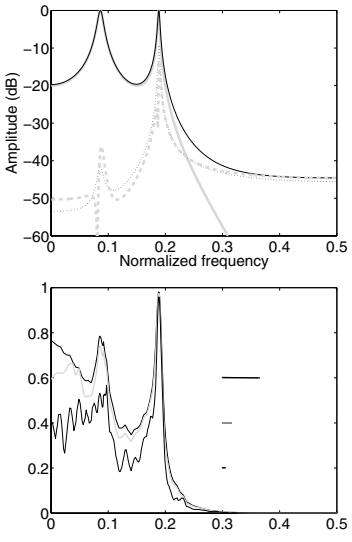


Fig. 6: Top: Gray: $\hat{G}_0(q)$, full black line: $\hat{G}(q, \theta)$, broken gray line: the error $\hat{G} - G_0$, broken black line: $\hat{\sigma}_G(q, \theta)$.

Bottom: Reliability of the variance for different resolutions (Large, medium, small) as indicated by the bars.