Model Validation and Model Error Modeling

Lennart Ljung

Department of Electrical Engineering
Linköping University, S-581 83 Linköping, Sweden
WWW: http://www.control.isy.liu.se
Email: ljung@isy.liu.se

March 9, 1999

Report no.: LiTH-ISY-R-2125
For the Åström Symposium on Control, Lund, Sweden, August 1999

Model Validation and Model Error Modeling

Lennart Ljung

Division of Automatic Control
Department of Electrical Engineering
Linköping University, S-581 83 Linköping, Sweden
E-mail: ljung@isy.liu.se
URL: http://www.control.isy.liu.se/

March 9, 1999

Dedicated With Admiration to Karl Johan Åström
For Well Tuned Initial Conditions

Abstract

To validate an estimated model and to have a good understanding of its reliability is a central aspect of System Identification. This contribution discusses these aspects in the light of *model error models* that are explicit descriptions of the model error. A model error model is implicitly present in most model validation methods, so the concept is more of a representation form than a set of new techniques. Traditional model validation is essentially a test of whether the confidence region of the model error model contains the zero model. However, the model error model allows a better visualization of the possible deficiencies of the nominal model. Based on such information, the nominal model may very well be accepted even if the model error model does not contain the zero model. Conversely, it will be illustrated that the model error model may give good reason – because of if its more precise infomation –to reject a nominal model, that has passed a conventional model validation test.

1 Introduction

Much of the renewed interest in system identification actually concerns *model validation*. Approaches with unknown but bounded disturbances, stochastic embedding, control oriented model validation, H_∞-identification, etc., typically have the ambition to provide more valid, and reliable, models for, say, control design.

We shall in this contribution discuss what we term *model error models* in this light. There are five aspects of this concept that will be stressed:

1. Model error models allow an alternative interpretation of standard model validation (residual analysis) tests
2. Model error models allow better visualization of the result of residual analysis for dynamical systems

3. Model error models allow safe use of nominal models that themselves have been falsified

4. Model error models allow a combination of simple nominal models evaluated within a potentially (much) more complex set of candidate system descriptions

5. Model error models put a finger on a perhaps neglected aspect of experiment design: Experiment design for *model invalidation*.

The first two aspects are demagogical in nature, and technically trivial.

The third one could prove quite useful in model applications like control design. The model error model takes care of both unmodeled dynamics and uncertainty, and as long as the robust control design can live with these, there is no need to burden the nominal model with more complexity. The nature of this aspect is that the nominal model and the error model together constitute an unfalsified model, even if this is never made explicit.

The fourth aspect is a deviation from basic scientific principles, like Occam's razor, to accept the simplest possible model that has not been falsified by experimental data. Therfore it clearly needs more discussion and a debate on this would be welcome. The reason for aspect four can be illustrated as follows: Suppose the intended model application is control design. The model builder shall deliver a set of possible system descriptions, based on measured input-output data. The control designer constructs a regulator that gives acceptable behavior when applied to all these models. The delivered model set should preferably be small, at the same time as the model builder should feel confident (at some level) that when the control design will give good behavior when tested on the real system. The ongoing discussion on identification for control, control oriented model validation, etc., indeed concerns the problem how to achieve this.

Now, the classical statistical approach to deliver such model sets is of the following kind:

> Under the hypothesis that the true system can be described as a third order linear system – *and I have not been convinced by the data that this is not the case* – it can with 99.9 % confidence be found in the delivered set.

Notice the double standards applied in this statement! The confidence level can be given at an impressive level, essentially matching the "hard bounds" cherished by the robust control community. On the other hand, the message also hinges upon the statement in italics, where we have demanded convincing evidence (perhaps at a 95 or 99 % level) to accept a more complex reality. There is still a substantial possibility – or risk – that are more complex model is required to describe the system, so delivering just the nominal third order model with its confidence region does not give full security for the model builder that the control design task will succeed. To deal with this in a formal way would require sophisticated decision theory. We shall in this paper suggest how model error models can handle it in a more informal way. Among other things, we can deal with unmodeled dynamics in terms of non-linearities rather than just being neglected linear dynamics.

The fifth aspect, experiment design for model invalidation, concerns issues in experiment design that are not so often stressed, but are closely linked with item four above. Experiment design usually deals with the question of maximizing the accuracy of certain model aspects. This might be in conflict with another aspect of the design, namely to display unknown or unexpected sides of the system properties. This is what helps constructing a powerful model error model.

The paper is organized so that a brief summary of essential issues in System Identification is given in Section 2. Aspects of model quality in connection with control design are treated in Section 3, while typical approaches to model validation are outlined in Section 4. The essential split into model error and disturbance is commented upon in Section 5, and two interpretations of standard model validation techniques are given in Sections 6 and 7. The four first aspects on model error modeling, listed above, are dealt with in Section 8, and the fifth one is further commented upon in Section 9.

2 System Identification For Control

"Identification for Control" is a term that has been coined rather recently, and it has been accompanied by quite a substantial literature. See, e.g., the special issues [12] and [25], as well as one or more special sessions on this topic at each CDC and ACC in the 90's.

However, it must be pointed out that the control application aspect of System Identification has been present ever since the birth of the subject. The fundamental paper [2] lines up all the relevant concepts of estimation of parameters in dynamical systems (although applied just to ARMAX models in that paper), with the clear intention to use the estimated model for control. This is also stressed in [3].

The traditional statistical framework of [2] set the stage for the major part of the subsequent development of System Identification. This framework suggests that you estimate models of increasing complexity until a validation test does not falsify the model. Part of such a test is typically a whiteness test of the model residuals. Under the assumption that the residuals indeed are white, the statistical uncertainty (the distribution) of the parameter estimates can be computed using standard procedures. If the hypothesis of whiteness is not rejected, it is reasonable to accept (i.e. not reject) not only the model but also the uncertainty measure of its parameters.

This traditional framework has been questioned in other parts of the control community, most notably in the early 90's, e.g., [23], in connection with robust control design. The criticism has three main aspect:

1. *The framework with disturbances described as stationary white noise is unrealistic.* "Real-life" disturbances are more complex. This leads to the "worst-case", "unknown-but-bounded" and "set-membership" approaches, see, e.g., [5], [17], [16], [27], and [26].

2. *Traditional System Identification does not consider dynamic uncertainties* (e.g. "... the most popular system identification methods assume that all uncertainty is in the form of additive noise"; quoted from the abstract of[23].) This has led to an error model with explicit disturbances and dynamic errors, suggested and discussed in, e.g., [23], [9], [20], and [22].

3. *The delivered nominal, estimated model always contains a dynamic model error (bias error) in addition to the statistical uncertainty of the parame-*

ters (variance error). A number of techniques have been developed to take this fact into account, among them the *stochastic embedding* approach, See [18] for an excellent survey of such approaches.

Let us make a few comments on these items:

1. The point is well taken, since indeed actual disturbances may be non-stationary, have "deterministic" components, etc. However it is important to point out (see [13]) that as long as the disturbances $w(t)$ have the property that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} w(t)u(t-\tau) = 0 \tag{1}$$

   the (asymptotic) model is not affected by the actual form of $w(t)$. We could phrase (1) as "the disturbance is uncorrelated with the input" but no probabilistic interpretation needs to be made for this. It is simply a statement about the two sequences w and u. It is true that "worst-case" disturbances (when Nature is allowed to select the disturbance after It has seen the input) will typically not be subject to (1), but it is questionable if such a signal is a "disturbance" and not a model error, see e.g., [8].

2. This statement, I think, is partly a misunderstanding. The traditional approach, as we saw, takes the model's confidence intervals seriously, and really delivers a set of possible dynamic models to the user. Perhaps it is item 3 that is the target of the criticism.

3. This is an issue that we shall discuss further in this contribution. We may note that that it has not been made clear whether proper uncertainty regions of a properly validated model, based on properly informative data, do not include both the bias and variance errors.

Experiment design in System Identification is focused on selecting input properties and includes the possibility of letting the input be generated (partly) as output feedback. Also this subject has seen a renewed interest in connection with identification for control. The idea then is to let the experiment excite the control-relevant dynamics of the plant, possibly in conjuction with a data prefilter to enhance the model fit in certain frequency ranges. Since it might not be known a priori what the "control-relevant dynamics" is, iterative schemes might be required, e.g. [7], [1]. The close connection to adaptive control, [4] then also becomes obvious.

3 Model Quality, Occam's Razor, and Control Design

We now assume that the identified model is to be used for control design. We can then picture the interplay between identification and control as a game between the model builder (MB) and the control designer (CD):

1. The MB delivers a set of models to the CD. The rationale he is using to compute this set is immaterial; be it worst case noise models, classical confidence regions or whatever.

2. The CD will now have to construct a regulator that gives acceptable closed loop behavior for all models in the set.

3. If the designed regulator performs well for all models in the set, but fails when applied to the true system, the MB looses face.

4. The CD may also find his task to be impossible (maybe he looses face then). He then turns back to the MB and asks for a smaller set in certain respects ("I cannot have such a big uncertainty around 5-10 rad/sec"). The MB may or may not be able accommodate this request without collecting more data. If more data are required, it's the task of the MB to design the experiment that delivers the new model set.

In this perspective *Model Quality* is a combination and a compromise between

- Having *small* model sets so that task number 2) may be more easily solved
- Having *large* model sets so as to increase the chance that the real life test of the design will succeed.

The MB can play it safe by delivering large, conservative sets to the CD. His pride should however forbid that, and make him deliver model sets with higher quality according the the definition just made.

By the way, it would be interesting to see more extensive tests of the model quality in this sense, using the many different approaches to model estimation: Will the traditional approach with properly validated models handle also model errors and "bad" disturbance behavior (subject to (1), though)? Do the worst case models have bad quality, being too conservative? And so on. This is the proof of the pudding.

The traditional statistical approach to estimation follows Occam's razor: Use the simplest possible description that is not in conflict with known facts. We should thus try and order the model structures of interest in increasing complexity. For dynamical systems this could typically be first linear models in increasing order and then adding non-linearities of different kinds. Prior knowledge, as well as desired control design techniques, will play an important role in this ordering. We would then select that model, together with is estimated uncertainty region, that is the first one, within the chosen ordering, not to be falsified by the validation tests.

Model quality for control design as defined above does not fully comply with this line of thought. Suppose that we test for a non-linearity by estimating some parameters in a non-linear structure. If the confidence region for these parameters contain zero, we should reject the hypothesis of a non-linear system, according to Occam's Razor, since no convincing evidence of this more complicated structure has been given. On the other hand, if the MB wants to play it safe in the game and be honest to the CD, he might very well include a possible non-linearity: "I cannot tell for sure that there are no non-linear effects, but in any case, they should not the larger than so-and-so."

A similar situation is at hand when the estimation is based on data of limited information value (poorly exciting, or a short record). For example, if the input consists of two sinusoids, and there are no harmonics in the output, a second order linear model will always pass the validation tests. According to Occam's razor, this model is also what should be delivered to the user. However, for control design it is clear that such a model must be delivered with a disclaimer about it properties at other frequencies than those excited. This is actually a case of principal interest, since the unfalsified second order model will itself come with uncertainty regions in the frequency domain that do not reveal the lack of information outside the excited area. Only when the model order is increased, this will be clear.

4 Two Typical Model Validation Tests

We are now in the situation that we are given a *nominal model* $\hat{G}$ along with a validation data set

$$Z^N = \{u(1), y(1), \ldots, u(N), y(N)\} \tag{2}$$

y and u being the output and the input of the system. We would like to devise a test by which we may *falsify* the model using the data, that is to say that it is not possible, reasonable or acceptable to assume that the validation data have been generated by the nominal model. If the model is not falsified, we say that the model has *passed the model validation test.*

Now what tests are feasible? It is natural to evaluate a model by its capability to reproduce the input-output behavior on new data sets. We thus compute the *residuals* ε from the nominal model $\hat{G}$ as

$$\varepsilon(t) = y(t) - \hat{G}(q)u(t) \tag{3}$$

(The nominal model need not all all be linear, but we use the above notation for simplicity. We may also include all possible prefiltering by proper preprocessing of Z^N.)

4.1 Classical residual correlation analysis

One of the most basic tests, [6], is to compute the correlation between the regressors, in our case the past inputs, and the residuals:

$$\hat{r}^N(\tau) = \frac{1}{N}\sum_{t=1}^{N} u(t-\tau)\varepsilon(t) \tag{4}$$

It is customary to plot these estimates as a function of τ and compare with their standard deviations to check if they are significantly different from zero. If not, we have not traced any significant influence of u in ε, so we cannot say that the model $\hat{G}$ has not picked up all the influence of u on y. (Note the double negation: we are not saying that "$\hat{G}$ has picked up all ..."). It is convenient to form

$$\varphi(t) = \begin{bmatrix} u(t-1) & \ldots & u(t-M) \end{bmatrix}^T \tag{5}$$

$$h_N^M = \begin{bmatrix} \hat{r}^N(1) \\ \vdots \\ \hat{r}^N(M) \end{bmatrix} = \frac{1}{N}\sum_{t=1}^{N} \varphi(t)\varepsilon(t) \tag{6}$$

Under the assumption that ε is white noise with variance λ, h has a normal distribution with zero mean and variance λ/NR_N, where

$$R_N = \frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t) \tag{7}$$

so

$$\xi_N^M = \frac{1}{\lambda N}\|\sum_{t=1}^{N}\varphi(t)\varepsilon(t)\|_{R_N^{-1}} \tag{8}$$

will in this case have a χ^2 distribution, and the familiar χ^2-test

$$\xi_N^M < \alpha \tag{9}$$

is based on this. Note that other kind of dependences can be tested quite analogously by letting $\varphi(t)$ be other, nonlinear, functions of past inputs:

$$\varphi(t) = f(u^t) \tag{10}$$

4.2 Control oriented model validation

The philosophy listed under item 2 in Section 2 proposes the following relationship (in a somewhat simplified version)

$$\varepsilon(t) = \Delta(q)u(t) + w(t); \qquad \|\Delta\|_\infty \le \alpha_1; \quad \|w\|_2 \le \alpha_2 \tag{11}$$

and the nominal model $\hat{G}$ "passes" the test (and is delivered together with α_1 and α_2) if there exists a linear system Δ with norm less than α_1 and a signal w with L_2-norm less than α_2 that solves (11) for the given ε, u. No further requirement is put on w. See, e.g., [9], [23], [24], [19].

As a special case, when $\alpha_1 = 0$, we obtain the validation test for the "unknown-but-MSE-bounded" approach, and if $\|w\|_2$ is changed to $\|w\|_\infty$ we have the more standard unknown-but-bounded (worst case, set membership) model characterization.

One might ask where the thresholds α come from in the two cases. In a sense, this has to rely upon prior information about the noise source. The first test, (9) is "self-contained", in the sense that it corresponds to a hypothesis test that ε is white noise, and then α corresponds to a certain confidence level for the test

5 A Fundamental Split of the Residuals

It is very useful to consider two sources for the model residual ε: One source that originates from the input $u(t)$ and one that doesn't. With the assumption that these two sources are additive we could write (*The model error model*)

$$\varepsilon(t) = \tilde{f}(u^t) + w(t) \tag{12}$$

The assumption that the contribution from w is additive is non-trivial and restrictive, but we shall not be concerned about that now. Note that the distinction between the contributions to ε is fundamental and has nothing to do with any probabilistic framework. We have not said anything about $w(t)$, except that it would not change, if we changed the input $u(t)$. We refer to (12) as the separation of the model residuals into *Model Error* and *Disturbances*.

We may further specialize the the case of a linear model error Δ:

$$\tilde{f}(u^t) = \Delta(q)u(t) \tag{13}$$

as in (11). Note though, that in (11), there is no assumption about w, other than it has bounded norm.

Describing w as "the part of ε that wouldn't change if we changed the input" is not a scientifically precise statement. The traditional way of specifying this is to introduce a probabilistic framework and require u and w to be *independent* in the mathematically well defined sense of the word. In practical work, one will however be content with tests of the kind (1). The standard test (9) is clearly devised to test if $\tilde{f}$ in (12) is zero.

For control use of the model it is important that w in (12) has this independence property if $\hat{G}$ and $\tilde{f}$ are delivered to the user along with uncertainty bounds, and some measure of a bound on w. Without this property, there is nothing that would prevent us from, say, interpreting the model error

$$\varepsilon(t) = 0.4u(t), \qquad \text{with } |u(t)| \leq 1$$

as a disturbance error

$$\varepsilon(t) = w(t), \quad |w(t)| \leq 0.4$$

This could clearly have a devastating effect on the control design. In other words, the independence paradigm eliminates the built-in ambiguity in (12).

6 Model Validation as Set Membership Identification

The model validation tests can also be seen as *set membership identification methods* in the sense that we may ask, for the given data set Z^N, which models within a certain class would pass the test. This set of *unfalsified models* would be the result of the validation process, and could be delivered to the user. The interpretation would be that any model in this set could have generated the data, and that thus a control design must give reasonable behavior for all models in this set. Let us now further discuss what sets are defined in the different cases.

To more clearly display the basic ideas we shall here work with models of FIR structure, i.e., we ask which models of the kind

$$\hat{G}(q)u(t) = \sum_{t=1}^{N} g_k u(t-k) = \theta^T \varphi(t) \tag{14}$$

will pass the test. φ is defined by (5). The validation measures above will then be given the argument θ as in $\varepsilon(t, \theta)$ and $\xi_N^M(\theta)$ to emphasize the dependence. See also [15].

6.1 Uncorrelated residuals and inputs

Suppose now we use the standard correlation test (9). Let $\hat{\theta}_N$ be the standard LS estimate using the validation data. Simple calculations give

$$\begin{aligned}\frac{1}{N}\sum_{t=1}^{N}\varepsilon(t,\theta)\varphi(t) &= \frac{1}{N}\sum_{t=1}^{N}(\varepsilon(t,\theta)-\varepsilon(t,\hat{\theta}_N))\varphi(t)\\ &= -\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)(\theta-\hat{\theta}_N)\\ &= -R_N(\theta-\hat{\theta}_N)\end{aligned}\tag{15}$$

where the first step follows from the definition of the LS estimate. We then find that

$$\begin{aligned}\xi_N^M(\theta) &= (\theta-\hat{\theta}_N)^T R_N R_N^{-1} R_N(\theta-\hat{\theta}_N)\\ &= (\theta-\hat{\theta}_N)^T R_N(\theta-\hat{\theta}_N)\end{aligned}\tag{16}$$

Inserting this in $\xi_N^M < \alpha$ gives that the set of non-falsified models is given by

$$(\theta-\hat{\theta}_N)^T R_N(\theta-\hat{\theta}_N) < \alpha \tag{17}$$

Note the connection between this result and traditional *confidence ellipsoids*. In a probabilistic setting, the covariance matrix of the LS estimate $\hat{\theta}_N$ is proportional to R_N^{-1} (see e.g. [14]). This means that (17) describes those models θ that are within a standard confidence area from the LSE. The level of confidence depends on α.

We note also, in passing, that if nominal model $\hat{G}$ also is a FIR model of the same order, it will pass the test if and only if it belongs to the confidence region of the estimate based on the validation data. The test (9) does not make explicit use of the estimate $\hat{\theta}_N$, but the interpretation is intuitively appealing and can be applied in more generality: A reasonable validation test is to split the data record, estimate separate models of the same structure on each part, and accept the model structure if the two models lie in each others' confidence regions. For this to be efficient, the input properties should be different in the two data parts.

6.2 Control oriented model validation

The model validation test (11) has been suggested for control oriented model validation. In this context is has also been customary to compute the set of unfalsified models, parameterized by α_1 and α_2. This is quite a formidable computational task, but results in a curve in the α_1-α_2 plane, below which the set of unfalsified models is empty, [10]. See figure 1. The shaded area corresponds to "possible" model descriptions, but it is normally interesting to consider just the models on the boundary.

One of the end-points is easy to deal with: If $\alpha_1 = 0$, all errors are to be explained from the disturbance w. For the L_2-norm $\|w\|_2$ it is readily shown, e.g. [15], that

$$\frac{1}{N}\sum_{t=1}^{N}\varepsilon^2(t,\theta) \le \alpha_2 \tag{18}$$

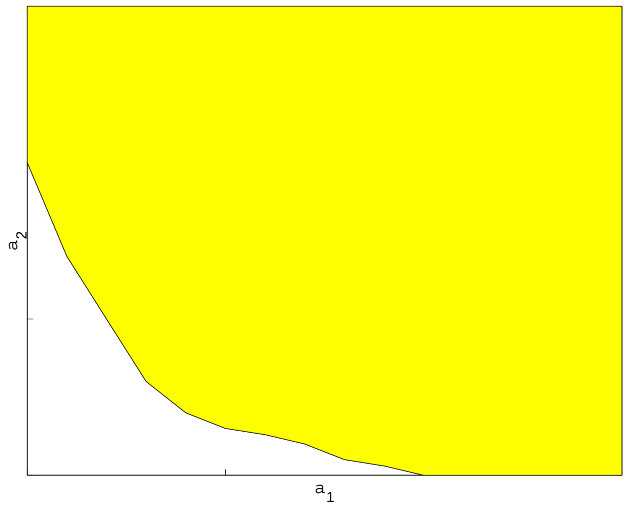

Figure 1: Shaded area: Models that pass the test that they can explain data with a model error less than α_1 and an additive disturbance less than α_2.

if and only if

$$(\theta - \hat{\theta}_N)^T R_N (\theta - \hat{\theta}_N) \leq \alpha_2 - \frac{1}{N} \sum_{t=1}^{N} \varepsilon^2(t, \hat{\theta}_N) \tag{19}$$

This shows that the α_2-axis is crossed at α_2=LSE-fit (=estimated noise variance) and the lowermost model is the LS model computed for the validation data.

Similarly, with an ∞-norm on w in (11), the models on the α_2-axis in figure 1 correspond to the traditional *unknown-but-bounded* set of models, e.g., [5], [17], [16], [26].

7 Model Validation As Model Error Modeling

It is immediate that the classical test (9) also can be interpreted in in terms of a *model error model*

$$\varepsilon(t) = \varphi^T(t)\eta + w(t) \tag{20}$$

(compare (10), (12)).

The LS estimate of η is given by

$$\hat{\eta}_N = R_N^{-1} \frac{1}{N} \sum_{t=1}^{N} \varphi(t)\varepsilon(t) \tag{21}$$

with covariance matrix $\hat{\lambda} R_N^{-1}$, where $\hat{\lambda}$ is an estimate of the variance of w. This means that a standard χ^2 test whether the true η is zero, has the form

$$\hat{\eta}^T (\hat{\lambda} R_N^{-1})^{-1} \hat{\eta} < \alpha' \tag{22}$$

or (cf (8))

$$[\sum_{t=1}^{N}\varphi(t)\varepsilon(t)]^T R_N^{-1}\frac{1}{\hat{\lambda}}R_N R_N^{-1}[\sum_{t=1}^{N}\varphi(t)\varepsilon(t)] = \frac{N}{\hat{\lambda}}\xi_N^M < \alpha' \tag{23}$$

That is, the standard χ^2-test can equivalently be described as testing whether the model error model (20) gives an estimate that is significantly different from zero. Yet another interpretation is that the test corresponds to checking whether the improvement in model fit (in terms of a prediction error criterion) is significant when the extra model (20) is appended. This gives the link to the classical model order hypothesis tests, [2].

As long as we just perform a yes/no test on the model residuals it is immaterial how we phrase the interpretation of the test. However, in case we have some opinion that certain model errors are more serious than others, it may be useful to think of the χ^2-test (9) as a statement of the character of the model error. This is what we turn to now.

8 Direct Model Error Modeling

8.1 FIR Model Error Models: Visualizing the Result of Residual Analysis.

The classical residual analysis test for dynamical systems, included in most software packages, is obtained by (4)-(9). Normally the result is also visualized by plotting $\hat{r}^N(\tau)$ as a function of τ.

It is clear that this test corresponds to a FIR model error model

$$\varepsilon(t) = \sum_{k=1}^{M} b(k)u(t-k) + w(t) \tag{24}$$

With this said, it is also clear that, at least from a control user's point of view it would be more natural to visualize the result in terms of this model's properties, like its impulse response or its frequency response.

We illustrate this in a few plots. We have simulated a system (same as in Example 8.5 in [14]) and estimated a second order ARX model. The model with its 99% confidence region (very thin) is shown in Figure 2. The model is also shown together with the true system, which reveals that the true system is not at all to be found within the confidence region. The reason is that the model order is too low, in combination with a fairly poor excitation at high frequencies.

The result of conventional residual analysis is shown in Figure 3. The cross correlations for 20 lags are computed. Visual inspection of this plot shows that the model is falsified by the data, but the character of the deficiencies is not clear. We interpret the calculations of the test as a corresponding 20:th order FIR model, and display its impulse response and frequency response in Figure 4.

Here, we see more clearly the character of the model errors. In particular the frequency function plot makes it clear that there are significant errors around 0.1 – 1 rad/s.

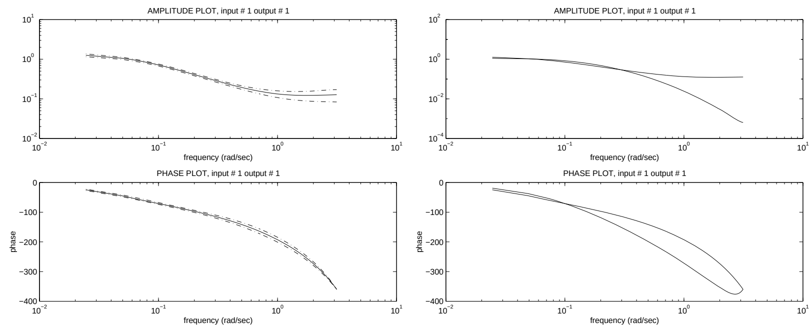

Figure 2: Left:Bode plot of the second order ARX model with 99% confidence intervals. Right: The true system together with the model.

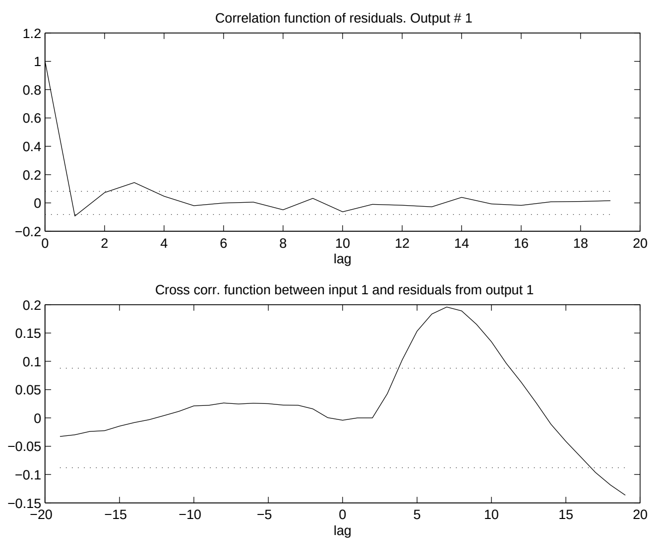

Figure 3: Residual analysis of the second order ARX model with 99% confidence intervals. Upper plot shows the auto-correlation of the residuals and the lower plot shows the cross-correlations

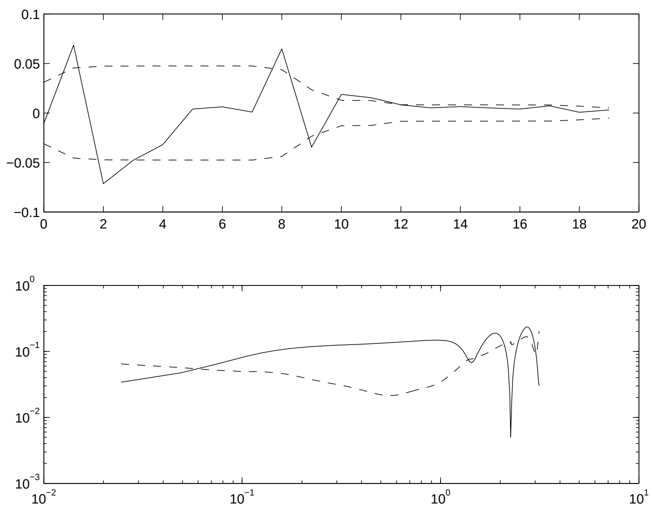

Figure 4: Above: The impulse response of the 20th order FIR error model. Below: Its frequency response Bode plot(amplitude only). Dotted lines correspond to 99% confidence limits around zero, i.e. anything outside these is a significant deviation from zero. In the lower plot the threshold is of course one-sided

To display this even more clearly, we propose the visualization according to Figure 5. The model error model is represented in the frequency domain, with its uncertainty regions around itself shaded. At the top figure the model error model and its uncertainty region is added to the nominal model. (The addition is of course applied to the complex-valued frequency functions.) This is shown as a shaded region, and would correspond to the model set to be delivered to the user. The nominal model is shown as a dashed line. In this case we also include the true system. We see that the delivered model set contains an accurate description of the system.

In the linear case, plots of the kind of Figure 5 will be our preferred way of presenting the nominal model, and its sidekick, the model error model. They work together to provide a suitable representation of the information in the collected data.

8.2 General Linear Black Box Model Error Models.

The model error model concept gives us more freedom in investigating the residuals than the classical residual correlation test. A more general linear model, like, e.g., the Box-Jenkins model

$$\varepsilon(t) = \frac{b_0 + b_1 q^{-1} + \ldots b_n q^{-n}}{1 + f_1 q^{-1} + \ldots + f_n q^{-n}} u(t) + \frac{1 + c_1 q^{-1} + \ldots + c_m q^{-m}}{1 + d_1 q^{-1} + \ldots + d_m q^{-m}} w(t) \qquad (25)$$

could improve the estimate of the error model, since a more sophisticated model of the disturbance is used.

In connection with this, it should be noted that if the nominal model contains

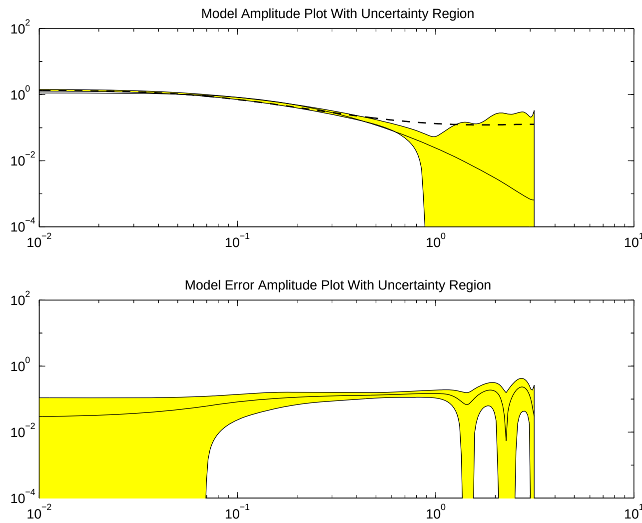

Figure 5: Upper plot: Nominal second order ARX model (dashed line) as well as nominal model + model error model with uncertainty region. The true system is marked with a solid line. (The nominal model plus model error model is not marked as a separate curve; just the corresponding uncertainty region.) Lower plot: The model error model (20th order FIR model) and its uncertainty region.

a noise model

$$y(t) = \hat{G}(q)u(t) + \hat{H}(q)e(t) \tag{26}$$

it is natural to build the error model

$$\varepsilon(t) = y(t) - \hat{G}(q)u(t)$$
$$\varepsilon(t) = \tilde{G}(q)u(t) + w(t)$$

from the prefiltered data

$$\varepsilon_F(t) = \hat{H}^{-1}(q)\varepsilon(t), \quad u_F(t) = \hat{H}^{-1}(q)u(t) \tag{27}$$
$$\varepsilon_F(t) = \tilde{G}(q)u_F(t) + w(t) \tag{28}$$

Instead of parametric linear models, we may apply spectral analysis to try and extract any linear influence of u on ε, [11]. In any case, there is a close relationship between the Blackman-Tukey spectral analysis estimate of this transfer function and the one obtained by a FIR-model.

8.3 Non-linear Model Error Models

In the literature, most model error discussions as well as the *identification-for-control* approaches are dealt with in a setting where "the true system" is a high order linear model, and the models are of lower order. That brings us to error models of the kind (11). In practical use, it is of course more common that the model errors are ignored non-linearities rather than unmodeled linear dynamics. From a model error perspective, this simply means that we should test non-linear models:

$$\varepsilon(t) = \tilde{f}(u^t) + w(t) \tag{29}$$

In the absence of specific, suspected non-linearities, it is reasonable to test non-linear black boxes like neural network NNFIR model, cf [21]:

$$\varepsilon(t) = g(u(t), u(t-1), \ldots, u(t-M+1), \eta) + w(t) \tag{30}$$

The number of lagged inputs can be chosen relatively small here, like $M = 5$ or so. To appreciate the size of any estimated non-linearity (in particular for control applications) it is natural to use the sup-norm

$$\|g(u(t), u(t-1), \ldots, u(t-M+1), \hat{\eta})\|_\infty = \sup_{u_1, \ldots, u_M} \frac{|g(u_1, u_2, \ldots, u_M, \hat{\eta})|^2}{u_1^2 + \ldots + u_M^2} \tag{31}$$

Then also determine the worst cases value of this norm in a properly chosen confidence regions for the estimate $\hat{\eta}$:

$$\|g\| = \sum_{\hat{\eta} \in \Theta} \|g(u(t), u(t-1), \ldots, u(t-M+1), \hat{\eta})\|_\infty \tag{32}$$

8.4 How to Use the Model Error Model

The model error modeling approach to model validation and model set delivery can be summarized as follows

1. Select beforehand a model error model structure that is versatile enough to handle a variety of model errors, and possibly also adjusted for suspected problems in the system at hand, and for errors that would be especially damaging for the intended model application. More about this in the next subsection.

2. Estimate nominal models in preferred order of increasing complexity. Determine the corresponding model error model with its confidence regions.

3. If the model error model contains the zero model, the nominal model has (essentially) passed a traditional model validation test. Then deliver the model *with the uncertainty region given by the model error model.* This is in conflict with the classical use of model validation and in conflict with Occam's razor. The reason why not to use the nominal model's own confidence region is explained below.

4. Even if the model error model is significantly different from zero, we may choose to stay with the nominal model, and take it plus the model error model and its uncertainty region as the estimated model set. The reasons for doing this may be that the errors (with regard to their frequency function, or to the norm (32)), even if statistically significant, are deemed harmless for the intended (control) application.

5. It is not the intention to treat the nominal model plus the model error model as a new and better nominal model. In this case one should reestimate a more complex nominal model.

Let us comment specifically on item 3. The point is to use the more complex model error model's uncertainty region, even when the simpler nominal model has not been falsified. That this might be wise is illustrated in Figures 10, 11 and 12. To push the issue, think of system identification using an input consisting of two sinusoids. Under this input – provided the output contains no harmonics – it is impossible to invalidate a second order linear model. Its own uncertainty (which is computed under the assumption that the true system indeed is of second order) cannot be used as a realistic description of what possible systems may have generated the data. The model error model thus also acts like a safeguard for poorly informative data.

8.5 How to Choose the Model Error Model

The remaining question now is: How shall we choose the model error model. Unfortunately, this section is bound to be a disappointment, since there will be no unique, scientifically sound way of selecting the error model structure. This is a reflection of the well know dilemma that we can never *verify* models and hypotheses, only *falsify* them.

This means that the structure of the model error model must be chosen on *ad hoc* grounds, based on experience and also on prior information about the system, and the intended application. We may list a few items to consider:

- The model error model structure must be so rich that the estimated model error model itself should not be falsified from data
- It should be considerably richer than the nominal model.
- It should reflect suspected or possible properties of the system
- ... in particular those that could be damaging for the intended model application

As a default structure, in case no specific information is at hand we may suggest

$$\varepsilon(t) = \sum_{t=0}^{20} b(k)u(t-k) + g_{\text{NN}}(u(t), u(t-1), \ldots, u(t-5), \eta) \qquad (33)$$
$$+ \frac{1 + c_1 q^{-1} + \ldots c_4 q^{-4}}{1 + d_1 q^{-1} + \ldots + d_4 q^{-4}} w(t)$$

where g_{NN} as a Neural Network black-box non-linear model. We stress that the error model should be built in two steps, so that first the linear part is extracted, before the Neural Network model is applied. This will remove the ambiguity of how to split the first two terms.

8.6 Example

Let us illustrate the discussion with the following example. We simulate the fourth order system

$$y(t) = \frac{q^{-1} + 0.5q^{-2}}{1 - 2.2q^{-1} + 2.42q^{-2} - 1.87q^{-3} + 0.7225q^{-4}} u(t) + w(t) \qquad (34)$$

The input u is a PRBS signal with clock period 5. To protect ourselves from the criticism of naive use of stationary stochastic processes as disturbance models we pick w to be the signal registered at the Charles F. Richter Seismological Laboratory (east-west accelerations) during the Santa Cruz Mountain Earthquake in 1989. (This signal has been made available by The MathWorks, Inc.) The signal is hardly a realization of stationary stochastic process, but it seems reasonable to assume that is is "independent" of the PRBS input in our simulation. The level has been adjusted so that the output SNR in the data is about 1. Figure 6 shows the noise-free output as well as the signal w. We estimate a second

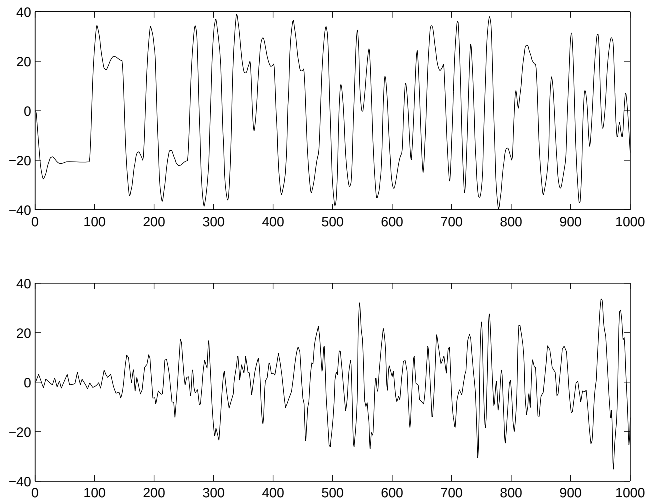

Figure 6: Upper plot: the noise-free part of the output in (34). Lower plot: The noise term w (actually an earthquake signal)

order Output Error model from the data. The traditional way of presenting the model is shown in Figures 7 and 8. We see that although the residuals seem to pass the correlation test (lower part of Figure 8), the confidence region of the frequency function of the second order model does not contain the true system. Figure 9 shows the Model Error way of presenting the result of the residual analysis. The model error model used here is a Box-Jenkins model with 20 FIR parameters and 5 parameters each in the numerator and denominator. Clearly, the information from the model error model is accurate.

Now, the fact that Figure 7 gives bad information, is not so serious as one may think at first sight. The residual analysis test is not passed, since the whiteness test of the residual fails, and hence the confidence regions in the lower plot of Figure 8 are not reliable.

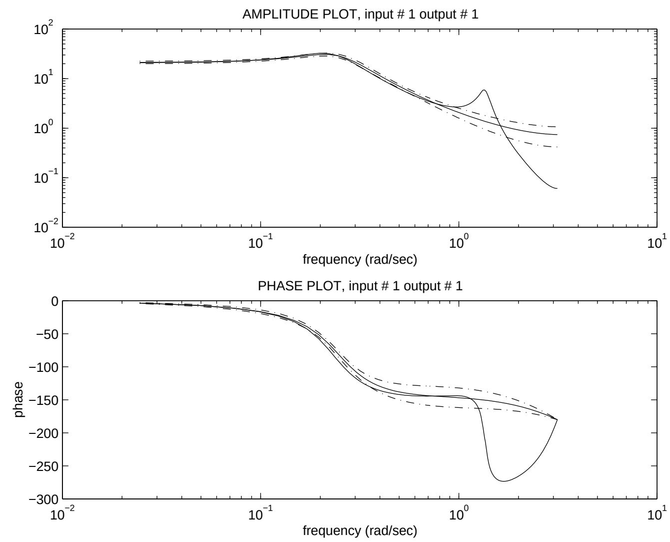

Figure 7: Bode plot of the second order output error model with 99% confidence intervals. The true system is also shown.

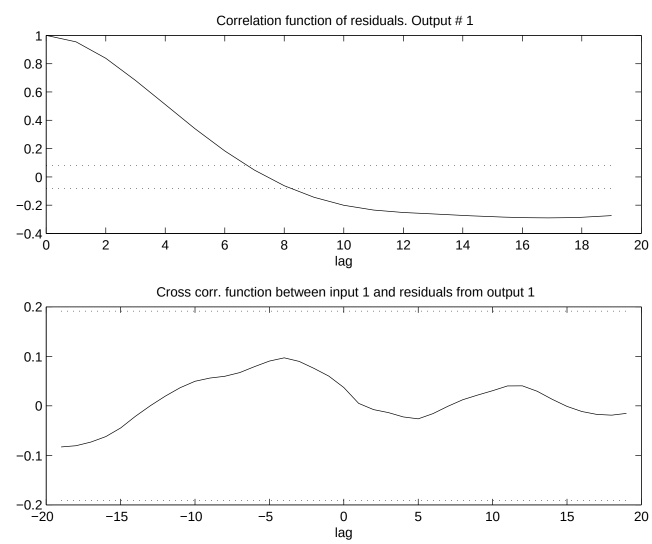

Figure 8: Residual analysis of the second order output error model with 99% confidence intervals. Upper plot shows the auto-correlation of the residuals and the lower plot shows the cross-correlations

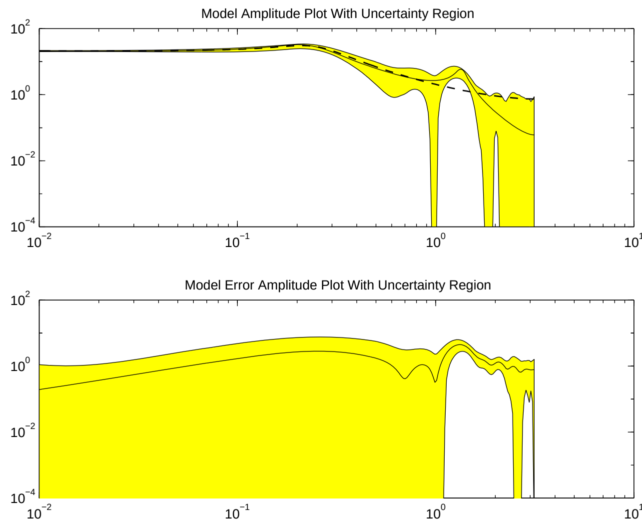

Figure 9: Upper plot: Nominal second order OE model (dashed line) as well as nominal model + model error model with uncertainty region. The true system is marked with a solid line. Lower plot: The model error model (Box-Jenkins type) and its uncertainty region.

We therefore try a second order Box-Jenkins model instead. The corresponding plots are given in Figures 10-12. We see here that with a more reliable noise prefilter, the correlation between residuals and past inputs is on the verge of being significant, but the model is not rejected by a χ^2 test at any reasonable level. It is quite difficult to see from the lower plot of Figure 11 if there are any deficiancies in the nominal model. The model error model in the lower plot of Figure 12 displays this information in a much more intuitive way, and we see also that the true system lies within the model set delivered in the upper plot. Since we have a noise model also in the Box-Jenkins estimate, we have used that, as suggested in (27) to build the error model based on the model error and input prefiltered by the inverse noise model.

There is one more important comment to be made:

- Even though the lower plot of Figure 12 shows that the nominal second order model is falsified, the qualitative information may still tell us that we may safe work with the simple nominal second order model, as long as the control design does not rely critically on the behavior above 0.5 rad/sec.

9 Experiment Design for Model Invalidation

The topic of experiment design for system identification has typically focused on how to obtain models of optimal accuracy within certain model structures. However, one should note that from this respect, optimal inputs could prove to be very bad in other respects. For example, an optimal input for identifying a second order linear system could very well consist of two sinusoids. With such

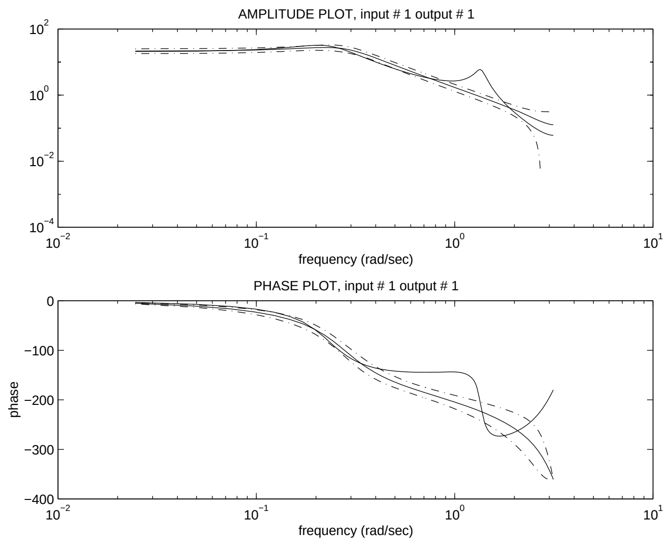

Figure 10: Bode plot of the second order Box-Jenkins model with 99% confidence intervals. The true system is also shown.

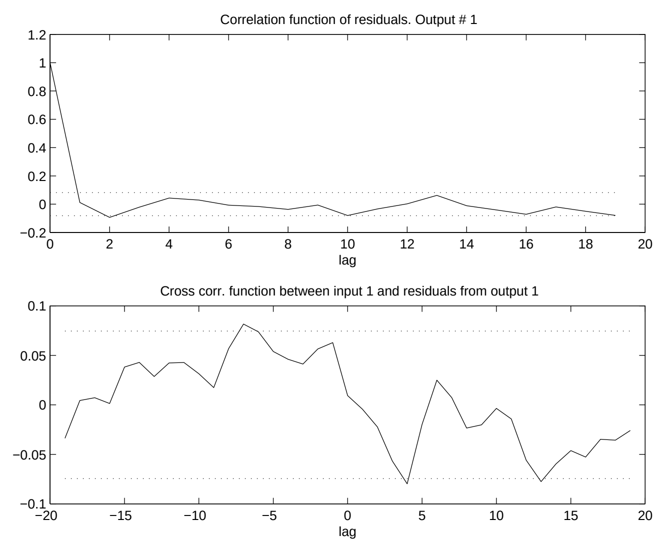

Figure 11: Residual analysis of the second order Box-Jenkins model with 99% confidence intervals. Upper plot shows the auto-correlation of the residuals and the lower plot shows the cross-correlations

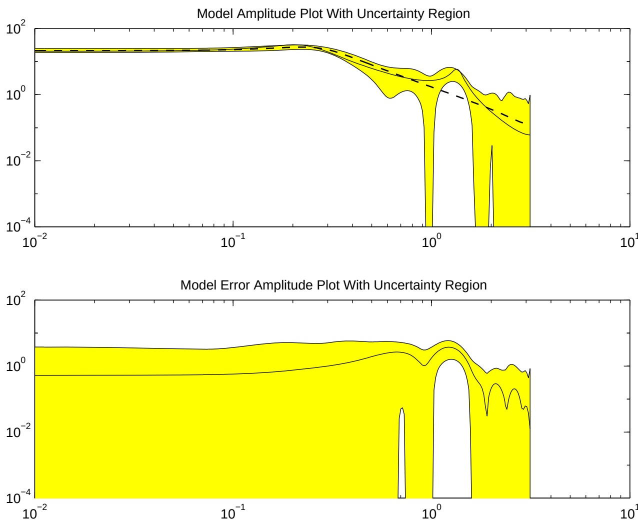

Figure 12: Upper plot: Nominal second order Box-Jenkins model (dashed line) as well as nominal model + model error model with uncertainty region. The true system is marked with a solid line. Lower plot: The model error model (Box-Jenkins type) and its uncertainty region.

an input it is impossible to falsify a second order linear model, so the optimal design really relies upon the correctness of prior information.

Similarly, an optimal input with optimal crest factor for a linear system is a binary signal. Such an input is incapable of finding a static non-linearity at the input of the system (Hammerstein non-linearity).

We would therefore need to develop *experiment design for model invalidation*. In the current framework it corresponds to *experiment design for powerful model error models*. This is not the place to go deeper into this issue, but let us point to one useful possibility: *periodic inputs*. A periodic input regularly takes us back to the same point in the regressor space, and thereby allows us to distinguish disturbance contributions from model errors. That is, we will get a certain help in the fundamental, but difficult split (12).

Assume that the model error model $\tilde{f}$ in (12) is stable and the input u is periodic with period P, and that transients can be neglected. Then the split in (12) is defined as separating ε into its, with period P, periodic part and the rest. (Note that this is a non-probabilistic way of catching the concept of independence between u and w.) We could obtain

$$\hat{\tilde{f}}(t) = \frac{1}{M}\sum_{k=1}^{M} \varepsilon(t + kP) \tag{35}$$

(M being the number of periods) as a simple estimate of the error model, and correspondingly an estimate of the size of the disturbances.

Periodic inputs have some additional advantages (as well as a few disadvantages), but from a model validation perspective, the quick model error model (35) is the most important one. This is a topic that should be further pursued.

10 Conclusions

We have pointed to a few ways of using an explicit model error model as a tool in model validation and for determining what set of possible system descriptions shall be delivered to the user. We have stressed that the model error model is more of a conceptual issue than a technical one: Many of the involved calculations are closely related to standard ones performed in model validation. The examples have illustrated that the explicit model error model gives more freedom in picking up both errors and uncertainties that may be hidden in traditional approaches, and is well suited, e.g., for control design.

References

[1] K. J. Åström. Matching criteria for control and identification. In *Proc. ECC 1993*, pages 248–251, Groningen, The Netherlands, 1993.

[2] K. J. Åström and T. Bohlin. Numerical identification of linear dynamic systems from normal operating records. In *IFAC Symposium on Self-Adaptive Systems*, Teddington, England, 1965.

[3] K. J. Åström and P. Eykhoff. System identification – a survey. *Automatica*, 7:123–162, 1971.

[4] K.J. Åström and B. Wittenmark. *Adaptive Control.* Addison-Wesley, Reading, MA, second edition, 1995.

[5] J. Deller. Set membership identification in digital signal processing. *Acoust. Speech and Signal Process. Mag.*, 6(4):4–20, 1990.

[6] N.R. Draper and H. Smith. *Applied Regression Analysis, 2nd ed.* Wiley, New York, 1981.

[7] Michel Gevers. Towards a joint design of identification and control? In H L Trentelman and J C Willems, editors, *Essays on control: Perspectives in the theory and its applications*, ECC '93 Groningen, 1993.

[8] Håkan Hjalmarsson. *Aspects in Incomplete Modeling in System Identification.* PhD thesis, Linköping University, Linköping, Sweden, 1993. Linköping Studies in Science and Technology, Dissertations No. 298.

[9] R. Kosut, M. K. Lau, and S. P. Boyd. Set-membership identification of systems with parametric and nonparametric uncertainty. *IEEE Trans. Automatic Control*, AC-37:929–941, 1992.

[10] R. L. Kosut. Iterative adaptive robust control via uncertainty model unfalsification. In *Proc. 13th IFAC Congress*, pages 91–96, San Francisco, July 1996. International federation of Automatic Control.

[11] R.L. Kosut. Adaptive calibration: An approach to uncertainty modeling and on-line robust control design. In *Proc. 25th IEEE Conference on Decision an Control*, volume 1, pages 455–461, Athens, Greece, December 1986.

[12] R.L. Kosut, G. C. Goodwin, and M. P. Polis (Eds). *Special Issue on System Identification for Robust Control Design, IEEE Trans. Automatic Control, Vol 37.* 1992.

[13] L. Ljung. A nonprobabilistic framework for signal spectra. In *Proc. 24th IEEE Conf. on decision and Control*, pages 1056–1060, Fort Laudredale, Fla, 1985.

[14] L. Ljung. *System Identification - Theory for the User.* Prentice-Hall, Upper Saddle River, N.J., 2nd edition, 1999.

[15] L. Ljung and H. Hjalmarsson. System identification through the eyes of model validation. In *Proc. Third European Control Conference*, volume 3, Rome, Italy, Sep 1995.

[16] P. M. Mäkilä, J. R. Partington, and T.K. Gustafsson. Worst-case control-relevant identification. *Automatica*, 31(12):1799–1819, 1995.

[17] M. Milanese and A. Vicino. Optimal estimation theory for dynamic systems with set membership uncertainty: An overview. *Automatica*, 27:997–1009, 1991.

[18] B. Ninness and G. C. Goodwin. Estimation of model quality. *Automatica*, 31(12):1771–1797, 1995.

[19] K. Poolla, P. P. Khargonekar, A. Tikku, J. Krause, and K.Nagpal. A time-domain approach to model validation. *IEEE Trans. on Automatic Control*, AC-39:951–059, 1994.

[20] S. Rangan and K. Poolla. Time-domain validation for sample-data uncertainty models. *IEEE Trans. Automatic Control*, AC-41:980–991, 1996.

[21] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.Y. Glorennec, H. Hjalmarsson, and A. Juditsky. Nonlinear black-box modeling in system identification: A unified overview. *Automatica*, 31(12):1691–1724, 1995.

[22] R. E. Skelton. Model error concepts in control design. *Int. J. Control*, 49:1725–1753, 1989.

[23] R. Smith and J. C. Doyle. Model validation: a connection between robust control and identification. *IEEE Trans. Automatic Control*, AC-37:942–952, 1992.

[24] R. Smith and G.E. Dullerud. Continuous-time control model validation using finite experimental data. *IEEE Trans. Automatic Control*, AC-41:1094–1105, 1996.

[25] T. Söderström and K. J. Åström, editors. *Special Issue on Trends in System Identification.* Automatica, Vol 31, No 12, 1995.

[26] D. Tse, M. Dahleh, and J. Tsitsiklis. Optimal asymptotic identification under bounded disturbances. *IEEE Trans Automatic Control*, AC-38:1176–1190, 1993.

[27] B. Wahlberg and L. Ljung. Hard frequency-domain model error bounds from least-squares like identification techniques. *IEEE Trans. on Automatic Control*, pages 900–912, 1992.