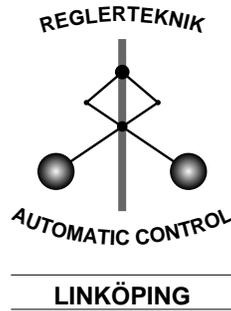


# Comments on Model Validation as Set Membership Identification

Lennart Ljung

Department of Electrical Engineering  
Linköping University, S-581 83 Linköping, Sweden  
WWW: <http://www.control.isy.liu.se>  
Email: [youremail@isy.liu.se](mailto:youremail@isy.liu.se)

March 9, 1999



Report no.: LiTH-ISY-R-2122

For “Robustness in Identification and Control”, Springer Verlag  
Conference in Siena, Italy, July 1998

Technical reports from the Automatic Control group in Linköping are available by anonymous ftp at the address [ftp.control.isy.liu.se](ftp://ftp.control.isy.liu.se). This report is contained in the compressed postscript file `2122.ps.Z`.

# Comments on Model Validation as Set Membership Identification

Lennart Ljung

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

March 9, 1999

## Abstract

We review four basic model validation techniques, one that relates to the “unknown-but-bounded” disturbance assumption, one that has been recently suggested in the “Identification-for-robust-control” context and two more classical statistical tests. By defining the set of models that would pass the chosen model validation test, we may interpret each of these as a set membership identification method. The consequences of such a viewpoint are discussed, and we focus on the important, but perhaps controversial concept of “independence” to make further selections of models within the thus defined sets.

## 1 Introduction

Model validation has always played a major role in System Identification, as a basic instrument for model structure selection and as the last “quality control” station before a model is delivered to the user.

The issues around model validation have shown renewed interest in connection with the recent discussions on “identification-for-control” or “control-oriented-validation”.

In this contribution we shall focus on some classical model validation criteria as well as on some more recently suggested ones. We shall in particular consider the validation process as a screening process that defines a set

of models that are not falsified by the test. We shall show the connections between these sets and common identification methods and we shall also discuss if there is any need for further choices within the set of un-falsified models.

We place ourselves in the following situation. A model is given. Let it be denoted by  $\hat{G}$  (more specific notation will follow later). We are also given a data set  $Z^N$  consisting of measured input-output data from a system. We do not know, or do not care, how the model was estimated, or constructed or given. We might not even know if the data set was used to construct the model.

Our problem is to figure out if the model  $\hat{G}$  is any good at describing the measured data, and whether it can be used for robust control design.

A natural start is to consider the model's simulated response to the measured input signal. Let that simulated output be denoted by  $\hat{y}$ . We would then compare this model output with the actual measured output and contemplate how good the fit is. This is indeed common practice, and is perhaps the most useful, pragmatic way to gain confidence in (or reject) a model. This will be the starting point of our discussion.

Some notation is defined in Section 2, while Section 3 discusses some relevant statistics around the measured and simulated outputs. Note that "statistics" here means some bulk, numerical descriptions of the fit; this has nothing to do with probability theory.

Section 4 deals with split of the residuals and a model error, while Section 5 lists the tests we are considering. The set membership aspects of model validation are outlined in Section 6 while a concluding discussion is given in Section 7.

## 2 Some Notation

We shall use the following notation. The input will be denoted by  $u(t)$  and the output by  $y(t)$ . The data record thus is

$$Z^N = \{y(1), u(1), \dots, y(N), u(N)\} \quad (1)$$

The given model  $\hat{G}$  will be assumed to be linear, and a function of the shift operator  $q$  in the usual way:  $\hat{G}(q)$ . The simulated output will thus be

$$\hat{y}(t) = \hat{G}(q)u(t) \quad (2)$$

It may be that the model contains a noise assumption, typically in the form of an additive noise or disturbance  $v(t)$  with certain properties. It would

then be assumed that the actual output is generated as

$$y_m(t) = \hat{G}(q)u(t) + v(t) \quad (3)$$

(We append a subscript  $m$  to stress the difference with the measured output.) The model could contain some "prejudice" about the properties of  $v(t)$ , but this is not at all essential to our discussion. A typical, conventional assumption would be that  $v(t)$  is generated from a white noise source through a linear filter:

$$v(t) = \hat{H}(q)e(t) \quad (4)$$

Most of the model validation test are based on simply the difference between the simulated and measured output:

$$\varepsilon(t) = y(t) - \hat{y}(t) = y(t) - \hat{G}(q)u(t) \quad (5)$$

Sometimes prefiltered model errors are studied:

$$\varepsilon(t) = L(q)(y(t) - \hat{y}(t)) = L(q)(y(t) - \hat{G}(q)u(t)) \quad (6)$$

For example, if the model comes with a noise model (4), then a common choice of prefilter is  $L(q) = \hat{H}^{-1}(q)$ , since this would make  $\varepsilon(t)$  equal to the model's *prediction errors*. The prefilter is however not at all essential to our discussion, and we shall cover the situation (6) by allowing the data set (1) be prefiltered.

In any case we shall call  $\varepsilon(t)$  the *Model Residuals* ("model leftovers").

### 3 Some Statistics Around the Residuals

Typical model validation tests amount to computing the model residuals and giving some statistics about them. Note that this as such has nothing to do with probability theory. (It is another matter that *statistical model validation* often is complemented with probability theory and model assumptions to make probabilistic statements based on the residual statistics. See, e.g., [2]. We shall not do that in this contribution.)

The following statistics for the model residuals are often used:

- The maximal absolute value of the residuals

$$M_N^\varepsilon = \max_{1 \leq t \leq N} |\varepsilon(t)| \quad (7)$$

- Mean, Variance and Mean Square of the residuals

$$m_N^\varepsilon = \frac{1}{N} \sum_{t=1}^N \varepsilon(t) \quad (8)$$

$$V_N^\varepsilon = \frac{1}{N} \sum_{t=1}^N (\varepsilon(t) - m_N^\varepsilon)^2 \quad (9)$$

$$S_N^\varepsilon = \frac{1}{N} \sum_{t=1}^N \varepsilon(t)^2 = (m_N^\varepsilon)^2 + V_N^\varepsilon \quad (10)$$

- Correlation between residuals and past inputs.

Let

$$\varphi(t) = [u(t-1), u(t-1), \dots, u(t-M)]^T \quad (11)$$

and

$$R_N = \frac{1}{N} \sum_{t=1}^N \varphi(t)\varphi(t)^T \quad (12)$$

Now form the following scalar measure of the correlation between past inputs (i.e. the vector  $\varphi$ ) and the residuals:

$$\xi_N^M = \frac{1}{N} \left\| \sum_{t=1}^N \varphi(t)\varepsilon(t) \right\|_{R_N^{-1}} \quad (13)$$

Now, if we were prepared to introduce assumptions about the true system (the measured data  $Z^N$ ), we could use the above statistical measures to make statements about the relationship between the model and the true system, typically using a probabilistic framework.

## Using Induction

If we do not introduce any explicit assumptions about the true system, what is then the value of the statistics (7)-(13)? Well, we are essentially left only with *induction*. That is to say, we take the measures as indications of how the model will behave also in the future: "Here is a model. On past data it has never produced a model error larger than 0.5. This indicates that in future data and future applications the error will also be below that value." This type of induction has a strong intuitive appeal.

In essence, this is the step that motivates the "unknown-but-bounded" approach. Then a model or a set of models is sought that allows the preceding statement with the smallest possible bound, or perhaps a physically reasonable bound.

Note, however, that the induction step is not at all tied to the unknown-but-bounded approach. Suppose we instead select the measure  $S_N^\varepsilon$  as our primary statistics for describing the model error size. Then the Least Squares (Maximum Likelihood/Prediction Error) identification method emerges as a way to come up with a model that allows the "strongest" possible statement about past behavior.

How reliable is the induction step? It is clear that some sort of invariance assumption is behind all induction. Here the statistics (13) plays a major role.

## 4 A Fundamental Split of the Residuals

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

$$\varepsilon(t) = \Delta(q)u(t) + w(t) \tag{14}$$

Note that the distinction between the contributions to  $\varepsilon$  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 (14) as the separation of the model residuals into *Model Error* and *Disturbances*.

The division (14) shows one weakness with induction for measures like  $M_N^\varepsilon$  and  $S_N^\varepsilon$  going from one data set to another. The implicit invariance assumption would require the input to be the same (or at least similar) in the two sets, unless we would have indications that  $\Delta$  is of insignificant size. The purpose of the statistics  $\xi_N^M$  in (13) is exactly to assess the size of  $\Delta$ . We shall see this clearly in Section 5. (One might add that more sophisticated statistics will be required to assess more complicated contributions from  $u$  to  $\varepsilon$ .)

In any case, it is clear that the induction about the size of the model residuals from one data set to another is much more reasonable if the statistics  $\xi_N^M$  has given a small value ("small" must be evaluated in comparison with  $S_N^\varepsilon$  in (10)).

## 5 Model Validation Tests

We are now in the situation that we are given a *nominal model*  $\hat{G}$  along with a validation data set  $Z^N$ . 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 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? Let us list some typical choices, based on the discussion above. In all cases we first compute the residuals  $\varepsilon$  from the nominal model as in (5).

1. Test if  $M_N^\varepsilon < \alpha$  (cf (7)) This corresponds to an assumption that the output noise is amplitude limited, and a model is valid if it does not break this assumption.
2. Test if  $S_N^\varepsilon < \alpha$  (cf (10)) This test is perhaps not common, but is based on an assumption that the variance of the output noise is known, and if the residuals show a significantly larger value, the model is rejected
3. Test if  $\xi_N^M < \alpha$  (cf(13)). This is the standard, “classical” residual analysis test, see e.g. [2].
4. Test if  $\exists \Delta, \|\Delta\|_\infty < \alpha_1$  such that  $\max_t |\varepsilon(t) - \Delta u(t)| < \alpha_2$  (cf (14)). This is, in simplified summary, the model validation test proposed in the “identification-for-robust-control” community, see, e.g., [3], [10], [11].
5. Estimate  $\Delta$  in (14), and let  $\hat{G}$  be unfalsified if the estimate  $\hat{\Delta}$  is not significantly different from zero.

In all these cases, one might ask where the threshold  $\alpha$  comes from. In a sense this has to rely upon prior information about the noise source, and we shall later discuss this issue. Only test number 3 is “self-contained”, in the sense that it corresponds to a hypothesis test that  $\varepsilon$  is white noise, and then the hypothesis to be tested also comes with a natural estimate of the size of the noise.

Let us also comment on test number 5. An estimate  $\hat{\Delta}$  can be viewed as a “model error model”, cf [4], [5], but there is a very intimate relationship to test number 3:

If  $\Delta$  is parameterized as a FIR model, its impulse response coefficients are estimated as

$$\hat{\eta}_N = R_N^{-1} \frac{1}{N} \sum_{t=1}^N \varphi(t) \varepsilon(t) \quad (15)$$

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

$$\hat{\Delta}^T (\hat{\lambda} R_N^{-1})^{-1} \hat{\Delta} < \alpha' \quad (16)$$

or (cf (13))

$$\left[ \sum_{t=1}^N \varphi(t) \varepsilon(t) \right]^T R_N^{-1} \frac{1}{\hat{\lambda}} R_N R_N^{-1} \left[ \sum_{t=1}^N \varphi(t) \varepsilon(t) \right] = \frac{N}{\hat{\lambda}} \xi_N^M < \alpha' \quad (17)$$

That is, test number 5 is equivalent to test number 3 for FIR model error models  $\Delta$ .

## 6 Model Validation as Set Membership Identification

Each of the five model validation test 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

$$\hat{G}(q)u(t) = \sum_{k=1}^N g_k u(t-k) = \theta^T \varphi(t) \quad (18)$$

will pass the test.  $\varphi$  is defined by (11). The validation measures above will then be given the argument  $\theta$  as in  $\varepsilon(t, \theta)$  and  $M_N^\varepsilon(\theta)$  to emphasize the dependence. See also [7].

## 6.1 Limited residual amplitude

The first test  $S_N^\varepsilon < \alpha$  gives the standard set membership approach to system identification. All models that produce an output error less than  $\alpha$  are computed, and for linear regression model structures, this problem can be solved by linear programming or bounding ellipsoids. See among many references, e.g. [1], [9], [8]

## 6.2 Limited MSE of the residuals

Suppose now we use test number 2. Let us define the LS-estimate of  $\theta$  for the validation data as  $\hat{\theta}_N$ . Then simple manipulations give that

$$\varepsilon(t, \theta) = -\varphi^T(t)(\theta - \hat{\theta}_N) + \varepsilon(t, \hat{\theta}_N)$$

and hence

$$\begin{aligned} \frac{1}{N} \sum_{t=1}^N \varepsilon^2(t, \theta) &= \frac{1}{N} \sum_{t=1}^N \varepsilon^2(t, \hat{\theta}_N) - \frac{2}{N} \sum_{t=1}^N \varepsilon(t, \hat{\theta}_N) \varphi^T(t) (\theta - \hat{\theta}_N) \\ &\quad + (\theta - \hat{\theta}_N)^T \frac{1}{N} \sum_{t=1}^N \varphi(t) \varphi^T(t) (\theta - \hat{\theta}_N) \\ &= \frac{1}{N} \sum_{t=1}^N \varepsilon^2(t, \hat{\theta}_N) + (\theta - \hat{\theta}_N)^T R_N (\theta - \hat{\theta}_N) \end{aligned}$$

where the second equality follows from the fact that

$$\frac{1}{N} \sum_{t=1}^N \varepsilon(t, \hat{\theta}_N) \varphi(t) = 0 \tag{19}$$

and  $R_N$  is defined as in (12).

This shows that the validation test will pass for exactly those models  $\theta$  for which

$$(\theta - \hat{\theta}_N)^T R_N (\theta - \hat{\theta}_N) \leq \alpha - S_N^\varepsilon(\hat{\theta}_N) \tag{20}$$

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  $R_N$  (see e.g. [6]). This means that (20) describes those models  $\theta$  that are within a standard confidence area from the LSE. The level of confidence depends on  $\alpha$ .

### 6.3 Uncorrelated residuals and inputs

Suppose now we use test number 3. We have

$$\begin{aligned}
\frac{1}{N} \sum_{t=1}^N \varepsilon(t, \theta) \varphi(t) &= \frac{1}{N} \sum_{t=1}^N (\varepsilon(t, \theta) - \varepsilon(t, \hat{\theta}_N)) \varphi(t) \\
&= -\frac{1}{N} \sum_{t=1}^N \varphi(t) \varphi^T(t) (\theta - \hat{\theta}_N) \\
&= -R_N (\theta - \hat{\theta}_N)
\end{aligned} \tag{21}$$

where the first step follows from(19). We then find that

$$\begin{aligned}
\xi_N^M(\theta) &= (\theta - \hat{\theta}_N)^T R_N R_N^{-1} R_N (\theta - \hat{\theta}_N) \\
&= (\theta - \hat{\theta}_N)^T R_N (\theta - \hat{\theta}_N)
\end{aligned} \tag{22}$$

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

$$(\theta - \hat{\theta}_N)^T R_N (\theta - \hat{\theta}_N) \leq \alpha \tag{23}$$

Here, again  $\hat{\theta}_N$  is the LSE for the validation data.

From these results we conclude that, for FIR models, the tests 2 and 3 are very closely related criteria. Furthermore, the results provide an alternative interpretation of probabilistic confidence regions. They are regions in the parameter space where, simultaneously, the *sample* cross correlation,  $\xi_N^M$ , and the Mean Square of the Model Residuals,  $S_N^\varepsilon$ , are small.

### 6.4 Control oriented model validation

Model validation test 4 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  $\alpha_1$  and  $\alpha_2$ . This is quite a formidable computational task, but results in a curve in the  $\alpha_1$ - $\alpha_2$  plane, below which the set of unfalsified models is empty. See figure 1. The shaded area corresponds to “possible” model descriptions, but it is normally interesting to consider just the models on the boundary.

## 7 Discussion

In a control oriented perspective, the reason for model validation is to find out if the nominal model can be used for reliable, robust control design.

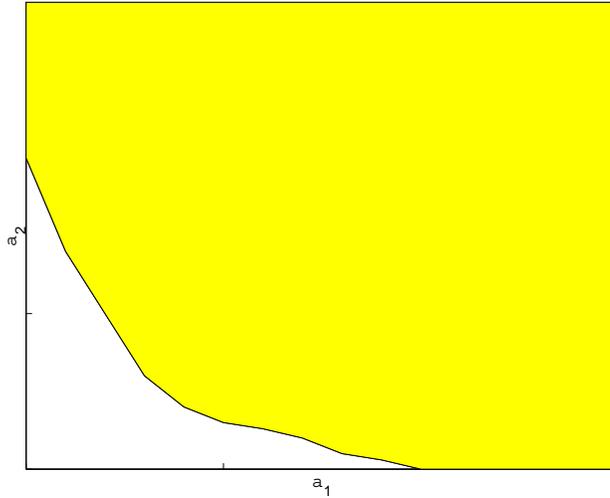


Figure 1: Shaded area: Models that pass the test that they can explain data with a model error less than  $\alpha_1$  and an additive disturbance less than  $\alpha_2$ .

From this perspective, the set membership aspect of model validation can be interpreted so that all models that pass the test should also give an acceptable closed loop system together with the constructed regulator. This indeed is closely related to the classical use of confidence intervals and uncertainty in connection with decision making.

A crucial question for this to hold is however that the validation process in some sense is successful in distinguishing the error model  $\Delta$  from the “noise”  $w$ . As a trivial example, let  $\varepsilon(t) = u(t)$ . If the input in the validation data happened to be less than 0.7, then, using a “deterministic setting” as in the tests 1) and 4) a model  $\Delta = 0$ ,  $|w(t)| < 0.7$  would not be falsified. Such a model clearly would not work for control design, however. It is really not up to the designer to pick any model in the shaded area for this reason: there is more information in the data. The notion of “independence” (between  $u$  and  $w$ ) is unavoidable, but not so easy to deal with if a probabilistic setting is rejected.

One useful possibility to get over this dilemma, in a non-probabilistic setting, is to use periodic inputs. That allow a distinction of the two terms in (14): the first one  $\Delta(q)u(t)$  is periodic with the same period as the input, while the disturbance  $w$  is not. This allows an independent, and non-parametric estimate of the size of the model error  $\|\delta\|$  and of the amplitude

or MSE of the disturbance  $w$ . Note that the periodicity is just a way to pinpoint “independence”:  $u$  and  $w$  are independent if they are not both periodic. Note also that in the pessimistic “worst-case” identification setup this would not help: the worst case disturbance is likely to be one that is periodic with the same period as the input.

With this we are back at yet another interpretation of the basic test 3): an interpretation in the frequency domain of this measure is that we are checking if the different harmonic components of  $u$  and  $w$  are systematically in phase with each other.

## References

- [1] J.R. Deller. Set membership identification in digital signal processing. *IEEE ASSP Magazine*, 4:4–20, 1989.
- [2] N.R. Draper and H. Smith. *Applied Regression Analysis, 2nd ed.* Wiley, New York, 1981.
- [3] 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.
- [4] L. Ljung. Identification, model validation and control. In *Plenary Presentation at the 36th IEEE Conference on Decision and Control*, San Diego, Dec 1997.
- [5] L. Ljung. Model validation and model error models. In *Symposium, in Honor of Karl Johan Åström, To appear*. Lund, Sweden, August 1999.
- [6] L. Ljung. *System Identification - Theory for the User*. Prentice-Hall, Upper Saddle River, N.J., 2nd edition, 1999.
- [7] 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.
- [8] P. M. Mäkilä, J. R. Partington, and T.K. Gustafsson. Worst-case control-relevant identification. *Automatica*, 31(12):1799–1819, 1995.
- [9] B. Ninness and G. C. Goodwin. Estimation of model quality. *Automatica*, 31(12):1771–1797, 1995.

- [10] R. Smith and J. C. Doyle. Model validation: a connection between robust control and identification. *IEEE Trans. Automatic Control*, AC-37:942–952, 1992.
- [11] R. Smith and G.E. Dullerud. Continuous-time control model validation using finite experimental data. *IEEE Trans. Automatic Control*, AC-41:1094–1105, 1996.