Model Validation in Metrology

Alistair B Forbes

Centre for Mathematics and Scientific Computing
National Physical Laboratory

1. Abstract

Modelling is the basis on which we extract information from measurement. For the information to be reliable the model must be valid. The main output of the modelling process in metrology is usually a measurement result, i.e., an estimate of the value of a parameter, and a statement of its uncertainty. In validating a measurement result it is necessary to consider both the parameter estimate and its stated uncertainty. In this paper, we give a brief introduction to the modelling and model validation Best Practice Guides that have been produced by the NMS Software Support for Metrology Programme.

2. Discrete Modelling in Metrology

Modelling is the basis on which we extract information from measurement. This statement is illustrated by Figure 1. The upper graph shows a seemingly random scatter of data points while the lower graph shows the data along with the fitted model. It is only from knowledge of the model that we are able to characterise the behaviour of a harmonic oscillator from the data.

The main components of discrete modelling are:

A *functional model* consisting of:

> Problem variables representing all the quantities that are known or (to be) measured.
>
> Problem parameters representing the quantities that have to be determined from the measurement experiment. The problem parameters describe the possible behaviour of the system.
>
> The functional relationship between the variables and parameters.

A *statistical model for the measurement errors* consisting of:

> The error structure describing which variables are known accurately and which are subject to significant measurement error.
>
> The description of how the measurement errors are expected to behave.

An *estimator* describing a method of determining estimates of the problem parameters from the measurement data. Good estimators are unbiased (will find the true solution on average) and efficient (make good use of the data).

3. What is a valid model?

There are two aspects of model validity. The first is *internal consistency*. To the extent that a model consists of a set of mathematical statements, its validity can be checked for mathematical correctness. Typically, a model has a set of inputs described in terms of facts and assumptions about components of the physical system. It also has a set of outputs in terms of predicted behaviour of systems to which it applies. If the model is internally consistent then the outputs are valid so long as the inputs are valid.

The second aspect of validity is *external consistency* with prior information and/or experimental results. A model is valid if information extracted from it is not contradicted by other valid information. The validation status of a

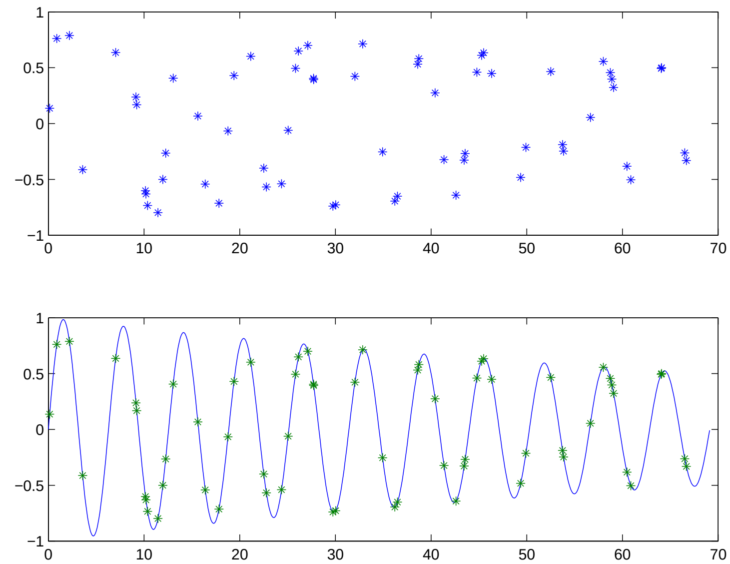

Figure 1: Data and fitted model for a harmonic oscillator. The data alone carries little information.

model is consolidated as it is shown to be consistent with more external information. Validation of an internally consistent model focuses on the extent to which the assumptions associated with the model apply, i.e., a validation of the applicability of the model.

The main output of the modelling process in metrology is usually a measurement result, i.e., an estimate of the value of a parameter, and a statement of its uncertainty. A valid model should produce a valid measurement result. A measurement result is valid if the estimate of the parameter is close to its true value (determined by a significantly more accurate measuring system, for example) relative to the stated uncertainty. In validating a measurement result it is necessary to consider both the parameter estimate *and* its stated uncertainty. For this reason, the model validation in metrology must be concerned not only with validation of the method by which parameter estimates are defined but also the method of uncertainty estimation.

Numerical simulation is an important tool in model validation. In particular, Monte Carlo simulation in which calculations are repeated for different inputs generated according to their statistical models is useful in determining the likely variation in the parameter estimates.

While the need for model validation seems self-evident, validation will generally require resources. It is therefore necessary to assess the risks associated with an invalid model and then design and implement validation responses that aim to control the risks to at acceptable level.

4. What Can Go Wrong

The functional model. Reasons for an invalid functional model include:

- Significant components of the model are either ignored or approximated inadequately.
- Incomplete understanding of the underlying science.
- Functional models are mathematical equations in which mistakes, typographical errors, etc., can arise.
- The model is used outside the area for which it was developed.

Validation responses include:

Conduct design review by an expert in the metrology field to examine the choice of model variables and check assumptions.

Perform numerical simulations to compare the behaviour of comprehensive models with simpler models.

Conduct design review by an expert in the metrology area to check the modelling of the underlying science.

Conduct design review by a modelling expert to check the mathematical derivation of equations.

Perform numerical simulations to check the effect of approximations, simplifications, linearisations, etc., on the model values (relative to the likely measurement error).

Evaluate the model at variable/parameter values for which the physical response is known accurately.

Perform numerical simulations to check the qualitative behaviour of the model against expected behaviour.

The statistical model. Reasons for an invalid statistical model include:

Measurements with different uncertainties are treated as having equal weight.

Variables subject to significant measurement error are treated as known exactly.

Correlated measurement errors are treated as independent. (The errors in quantities derived from a common set of measurements will be correlated.)

Validation responses include:

Conduct design review by an expert in the metrology field to examine the statistical model for the measurement data and check assumptions.

Conduct design review by modelling expert to check the statistical models for derived quantities.

Perform numerical simulation to check the effect of approximations, simplifications, linearisations, etc., associated with the statistical model.

Perform Monte Carlo simulations to check the variation in derived quantities against the predicted variation.

The estimator. Reasons for an invalid parameter estimation method include:

The computational problem to determine appropriate parameter estimates is replaced by a simpler one. For example, a nonlinear problem is replaced by linear one.

The estimator takes poor account of the statistical model of the measurement data. For example, the estimator may be appropriate for data (x_i, y_i) in which x_i is measured accurately but y_i is subject to significant measurement error but in fact it is variable x_i that is measured with error.

The estimator is statistically inefficient, ignoring a large component of the measurement information.

The estimator implements an algorithm that does not provide adequate estimates of the solution.

The uncertainties in the solution parameters are calculated according to an incorrect statistical model for the data.

Validation responses include:

Perform Monte Carlo simulations to examine the bias and variation of the solution estimates on datasets generated according to the statistical model.

Perform Monte Carlo simulations to compare the predicted variation of parameter estimates with the actual variation on datasets generated according to the statistical model.

Apply the estimator to datasets for which the estimates provided by an optimal estimator are known.

Compare the actual variation of parameter estimates on datasets generated according to the statistical model with the predicted variation for an optimal estimator.

Compare the actual statistical model with the statistical model for which the estimator is known to perform well.

5. Validation of the Model Solution

In addition to the validation approaches described above which examine how a model solution is arrived at, it is also possible to examine the validity of a solution itself. Validation measures include:

Examine the goodness of fit in terms of the size of the residual errors.

Compare of the size of the residual errors with the statistical model for the measurement errors.

Plot the residual errors to check for random/systematic behaviour.

Plot the root-mean-square residual error for a number of model fits to select an appropriate model (e.g., the polynomial of appropriate degree).

Calculate and check the uncertainty associated with the model predictions againsts requirements and/or expected behaviour.

6. Conclusions

An approximate functional model or approximate solution method will generally provide estimates of the model parameters that are not optimal. If the uncertainty estimates for this solution fully take into account the approximations, then the solution and its uncertainty provide valid information (but a better approach would provide better information). What gives rise to invalid results is when the uncertainty estimates do not reflect the approximations in the model. In these cases, it is quite possible to arrive at uncertainty estimates that are incorrect by a factor of 2, 5 or even more.

A comprehensive discussion on mathematical and statistical modelling and model validation is presented in the Software Support for Metrology Best Practice Guides Nos. 4, 6 and 10.

7. References

1. M. G. Cox, A. B. Forbes, and P. M. Harris. Software Support for Metrology Best Practice Guide 4: *Modelling Discrete Data*. National Physical Laboratory, Teddington, 2000.

2. M. G. Cox, M. P. Dainton, and P. M. Harris. Software Support for Metrology Best Practice Guide No. 6: *Uncertainty and Statistical Modelling*. National Physical Laboratory, Teddington, 2001.

3. R. M. Barker and A. B. Forbes. Software Support for Metrology Best Practice Guide No. 10: *Discrete Model Validation*. National Physical Laboratory, Teddington, March 2001.