Choice of complexity in constitutive modelling of fatigue mechanisms

Erland Johnson and Thomas Svensson

Choice of complexity in constitutive modelling of fatigue mechanisms

Abstract

The uncertainties in the virtual design stages of the product development chain are considered. The current computer capacity enables the processing of complex models to handle these steps. However, in the present work it is shown that the prediction capacity is not always favoured by an increased model complexity. In fact, it is observed that the prediction capacity deteriorates when the complexity exceeds a certain level. An example is shown and its implication for fatigue design are discussed.

Key words: fatigue, model complexity, prediction uncertainty

Erland Johnson

SP Swedish National Testing and Research Institute, Borås, Sweden

Tel. +46-33-165622, email: erland.johnson@sp.se

Thomas Svensson

Fraunhofer Chalmers Research Center for Industrial Mathematics, Göteborg, Sweden

Tel. +46-31-7724284, email: thomas.svensson@fcc.chalmers.se

**SP Sveriges Provnings- och
Forskningsinstitut**

SP Rapport 2005:15

ISBN 91-85303-46-1

ISSN 0284-5172

Borås 2005

**SP Swedish National Testing and
Research Institute**

SP Report 2005:15

Postal address:

Box 857,

SE-501 15 BORÅS, Sweden

Telephone: +46 33 16 50 00

Telex: 36252 Testing S

Telefax: +46 33 13 55 02

E-mail: info@sp.se

Contents

Abstract	2
Contents	3
1 Introduction	5
2 Method	7
3 Empirical modelling	8
4 A polynomial example	10
5 A general linear formulation	14
6 A fatigue example	15
Conclusion	16
References	17

1 Introduction

The product development process within the engineering- and vehicle industry typically involves the following activities:

1. Develop an idea for a product
2. Establish a specification of requirements for the product
3. Choose concepts for the different components of the product
4. Decompose the requirements to the component level
5. Establish design loads on the different components
6. Develop a design proposal for each of the components
7. Predict the properties of the different components
8. Optimize the properties of the different components
9. Develop and manufacture a prototype
10. Verify product properties and specification of requirements

Following the fourth step, a concept with requirements is established for each component in the product. In the following four steps (steps 5-8) the purpose is to virtually develop an optimum product, i.e. without producing a physical prototype (performed in step 9 above). The optimality is defined in each individual case from concurrent requirements on both cost (e.g. material consumption) and performance (e.g. structural strength or endurance).

Today, this virtual optimization is commonly performed with the finite element method (FEM). A finite element program essentially needs three types of input data, namely

1. Load history
2. Geometry (design proposal)
3. Material data (to be used in one of the predefined material models in the FE-program or, alternatively, in separate user-defined material models that are called from the program)

The FE-program uses this information to calculate the stresses and the displacements in the model. These values are then combined with fatigue data for the material to determine the fatigue life of the product. In practice the process is iterative where the engineer, when cost/performance requirements are not fulfilled, has to go back and perform modifications in geometric design, manufacturing technique, type of heat treatment or choice of material. Sometimes the engineer could reach an impasse and be forced to also reduce the requirements on, for instance, the load level. Modifications in the load or the geometry are connected to the input types 1 and 2 above while the remaining modifications are related to type 3, material data. This category involves changes on two different levels. When only modifications of the material are performed, for instance modifications in the manufacturing or heat treatment procedures, the material model could often be kept, while only the material parameters are changed. On the other hand, when the material is replaced, a change of material model might be necessary. Examples of material models are elasticity (with different sub classes regarding non-linearity and anisotropy), plasticity (with different subclasses regarding hardening, e.g. isotropic or kinematic hardening) or viscoelasticity. Examples of material parameters are Young's modulus, Poisson's ratio, modulus of hardening and relaxation time. Some of these parameters are present in more than one material model (e.g. Young's modulus), while some material parameters are specifically connected to a certain material model (e.g. modulus of hardening or relaxation time).

A critical and recurrent issue in the described iterative procedure is to judge if a certain design fulfils the requirements on performance and endurance or not. On which

foundation is this decision taken? A number is obtained from the calculation program but in practise an uncertainty (sometimes large) prevails on how to interpret the result. A consequence is that decisions in some companies quite frequently are postponed until a prototype has been manufactured. This is cost demanding for several reasons:

1. Resources (computer equipment and man power) have been put into calculation activities, which still are not sufficient for decision making.
2. The product development time increases which, in itself, corresponds to increased costs and also lost market time since the product cycles today become shorter and shorter.
3. Increased costs for manufacturing several prototypes since physical prototypes to a larger extent become involved in the iterative process described above.

Parallel to the increased need for decisions based on virtual simulations, the computers continuously become even more powerful. The computer capacity today enables a discretization of a geometry into a large number of finite elements with a very small discretization error. The computer capacity also implies that extremely advanced and complex material models could be used. The computers could today, in contrast to the situation only 10 years ago, perform simulations with advanced material models while maintaining moderate execution times. But, on the same time as discretization errors decrease and an increased model complexity has become possible to simulate, it is easy to loose the perspective and disregard the fact that other sources of error have not decreased to the same extent. These other sources of error have, instead, become the parameters controlling the accuracy in the results and therefore also become the bottleneck for reaching a rational decision. These sources of error are:

1. There is an uncertainty in the load representativity. The loads are chosen from a more or less guessed typical customer, or perhaps a worst customer. Irrespectively of how the loading is chosen, there is an uncertainty in to which degree the chosen load agrees with the actual service loads.
2. In spite of the use of computationally complex material models, critical model simplifications might be introduced. There is a risk that the model complexity is not focused on the most dominant effects for the performance requirements of interest.
3. There is an uncertainty in the material parameter values, which usually increases with increasing model complexity since it becomes too expensive (too many laboratory tests) to keep the accuracy in each material parameter for a more complex model compared to a model with fewer parameters. An increased level of complexity also reduces the application range of the model with a poorer prediction capability outside this range as a result.

Today there is a large difference in complexity between the models that are used in the industry and those models that are developed within academic research activities. The industrial models are phenomenological with few material parameters, which are calibrated against standard tests. As an example, for a steel material, a plasticity model with linear (kinematic) hardening is often used and its material parameters are calibrated against tensile tests and cyclic tests. For high cycle fatigue, the models are usually based on Wöhler curves, calibrated against high cycle fatigue tests on smooth specimens. In contrast to this, the research literature has plenty of much more advanced material models. This difference between industry and academia could partly be explained by the fact that more complex material models require an increased amount of testing to determine the (larger number of) material parameters. This counteracts the industrial requirement of a reduced number of tests.

Is the increase in available computer capacity used in the most efficient way today? It is not possible to give a definite answer to this question. However, one thing is clear: Discretization and material modelling could today be performed with such a high

accuracy in the computer, in contrast to the situation only a decade ago, such that, from a product development perspective, other bottle necks within virtual simulations have arisen. (It should, however, be mentioned that the use of complex material models requires a certain amount of interpretations, which still give rise to scattering in numerical results, *cf.* Bernauer & Brocks [1].) The problem today is most often not to reach sufficient accuracy in the FE-calculations but rather to specify the total uncertainty in the complete product development chain and, from this, choose a discretization and material model with appropriate complexity and accuracy. This specified uncertainty is different at different stages of the product development. During earlier stages, focus is on rapidity at the sacrifice of accuracy. It might, for instance be sufficient to identify areas with high stress without obtaining accurate information about the specific stress levels, while in the later stages a high accuracy is required, *e.g.* for verifications of fatigue life, and thereby more computationally heavy models are acceptable. In all stages there are, however, a need to, from a specific accuracy, choose an optimum discretization and material model. The influence of discretization on accuracy is today well known and depends, in practise only on the setting of the numerical parameters in the FE-program. The accuracy in a material model is, however, a more complex issue since it is based on test results. These involve scattering from the material itself as well as from the test procedure. The discussion points to the fact that when choosing a material model, it is necessary to simultaneously consider all sources of errors in all calculation stages within the product development chain, and, based on given resources for laboratory testing, identify the optimum material model.

The discussion leads to the following questions:

- How does the total uncertainty in the prediction of stress and fatigue life depend on the complexity (the number of parameters) in the chosen material model? Also, how complicated material models is it profitable to introduce in a certain situation?
- Is it a cost effective industrial strategy to keep to the simpler models or could testing methodology be developed (towards more similarity with the service life situation) such that current testing costs in combination with more advanced material models could increase the total prediction accuracy? Is the application area of the model reduced below critical levels through this procedure?
- Where should research efforts within virtual testing and simulations be directed in order to develop more efficient and better life predictions in the future?

A methodology to handle the overall complexity in product development in general and in fatigue life determination specifically is introduced and discussed in this report.

2 Method

The different sources of error when predicting the fatigue life could typically be grouped into the categories monotonic material behaviour, load history, geometry, FEM and cyclic material behaviour. These are illustrated in Figure 1. In order to understand the influence of the complexity of the material model on the total prediction inaccuracy, all sources of error must first be quantified. This cannot be done generally but a restriction to a component of a certain material, exposed to a certain load sequence must be done. By choosing a typically occurring industrial case with an appropriate parameterization of the material behaviour, load sequence and geometry, it is judged that comparatively general conclusions still could be drawn. To obtain quantitative values of the different errors, different sources and methods as, for instance, literature results, testing experience and calculation models, could be used. The total prediction error could then be expressed as a function of a variable number of parameters corresponding to a variable complexity in the

material model. A general approach to the problem of complexity in empirical models is outlined in the next section.

3 Empirical modelling

Modelling of physical phenomena are more or less always based on empirical observations. A mathematical formulation of such a model of a scalar measure of the phenomenon can be written

$$y = f(x_1, x_2, \dots, x_m),$$

where $x_1, x_2, \dots, x_m$ are different measurable variables that influence the phenomenon and f is an arbitrary mathematical formulation. By a Taylor expansion of the function around a nominal value $y_0 = \theta_0$ one can write

$$y \approx \theta_0 + \theta_1(x_1 - x_{1,0}) + \theta_2(x_2 - x_{2,0}) + \dots + \theta_m(x_m - x_{m,0}) + \theta_{m+1}(x_1 - x_{1,0})^2 + \dots, \quad (1)$$

where $x_{.,0}$ are the nominal values of the influentials, the parameters $\theta_1, \theta_2, \dots$ are proportional to different derivatives of the original function f and the approximation can be done as good as we need by adding terms of higher degree or limit the domain of application. This linear form of the function makes it somewhat easier to understand the problem of model complexity and we will use it here to demonstrate some fundamental problems in empirical modelling.

In case of a pure empirical function the partial derivatives are not known and the method for determining the function from observations is to estimate the parameters $\theta_0, \theta_1, \theta_2, \dots$. This is usually done by means of some least square method, i.e. by minimizing the squared errors between the observed values and the model fit. Such a fit can be done arbitrarily close to the observations by choosing more variables and in the limit one can obtain a perfect fit by choosing the same number of variables as the number of observations. However, in such a limiting case the modelling is quite useless, since no data reduction has been made. Further, no information is left to judge about the uncertainty in the model and consequently nothing is known about the quality of future predictions based on the model. This fact gives rise to the complexity problem in modelling: What is the optimal trade off between model complexity and prediction ability.

Figure 1. The different sources of error within the calculation of fatigue life time for a component.

4 A polynomial example

The following example from simple polynomial regression on one variable demonstrates the complexity problem: We have observed ten values $y_1, y_2, \dots, y_{10}$ depending on one variable x and want to find the function

$$y = f(x).$$

A Taylor expansion of the function is

$$y = \theta_0 + \theta_1(x - x_0) + \theta_2(x - x_0)^2 + \dots + \theta_{p-1}(x - x_0)^{p-1} + e, \quad (2)$$

where e is the error in the model, which represents both neglected x -terms and other unknown or neglected influences to the measure y . By assuming a random occurrence of such influences in the observations one can model the error term e as a random variable. This is the general statistical approach giving tools for estimating both confidence bands for the estimated parameters, and confidence- and prediction bands for the model.

The ten observations of y obtained for ten reference values of x gives the parameter estimates $\hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_{p-1}$ and we now want to decide how many parameters one should use in the model. Of course, the more parameters included the better fit one can get, but what about the possibilities for prediction? Using the statistical approach we can search for the number of parameters that gives the best prediction abilities. If the random variable is assumed to have a Gaussian distribution the following prediction limits will contain 95% of future measurements $\tilde{y}$:

$$\tilde{y}(x) = \hat{y}(x) \pm t_{0.025, n-p} \cdot s \cdot \sqrt{1 + g_{n,p}(x; x_{ref})}, \quad (3)$$

where $\hat{y}(x)$ is the estimated value based on the model (2) and the estimated parameters $\hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_{p-1}$, $t_{0.025, n-p}$ is the 2.5% quantile in the *student-t* distribution with $n-p$ degrees of freedom, s is the estimated standard deviation of the random variable e , $g_{n,p}$ is a function of the value of the influential variable x for the actual prediction situation and the values of the reference vector x_{ref} used in the estimation procedure. The theory behind this formula and an expression for the g function can be found in ordinary text books on linear regression.

The illustrations in figure 2 below show the results of using the model (2) with different polynomial degrees, corresponding to different number of parameters. Each figure shows the ten observations as dots, the fitted polynomial function as a line, and 95% prediction limits around the fitted function calculated using (3).

In the polynomial of zero degree the estimated function is simply $y = \hat{\theta}_0$ and the given prediction band is expected to contain 95% of new observations whose influential variable is chosen within the interval $0 \leq x \leq 1$. The first degree polynomial fit unexpectedly gives a wider prediction band. This is the result of that the reference measurements happened to give a very weak slope, and the improvement of the fit does not compensate enough for the more uncertain parameter estimates. The second degree polynomial does not any better, but the third degree shows clearly narrower prediction band and thereby a substantial improvement in prediction ability. The fourth degree polynomial gives a better fit to the observed values, but the prediction band now grows again, depending on the uncertainties in the parameter estimates, and finally the fifth degree performs even worse by means of prediction purposes. Further increased

complexity by choosing higher degrees are not shown here, but they give even wider prediction bands.

Figure 2. A polynomial function is fitted to experimental data and the corresponding 95% prediction bands are shown for increasing polynomial order.

One can conclude from these example figures that the best choice of model complexity seems to be the third degree polynomial, i.e. the model with four parameters, since this model gives the best prediction ability.

This simple polynomial example demonstrates the strength of statistical methods when choosing an optimal complexity for physical empirical models. However, it can be done more rigorous and rational by a formal criterion. The conclusion of optimal complexity in the example was based on a visual judgement about the area of the prediction bands in the figures. This area is a rough representation of the expected prediction variance and is an intuitive picture of one of the formal criteria for the optimal choice that has been presented in literature, namely the Breiman/Friedman criterion [2]: Minimize the estimated expected prediction variance $\hat{U}_{n,p}^2$:

$$\hat{U}_{n,p}^2 = s_{n,p}^2 \left(1 + \frac{p}{n-p-1} \right), \quad (4)$$

where $s_{n,p}$ is the estimated standard deviation for the random error in the empirical model, i.e. the error e in our polynomial example.

$$s_{n,p}^2 = \frac{1}{n-p} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (5)$$

where y_i is the i -th observed value and $\hat{y}_i$ is the predicted value. This simple criterion can be shown to be precisely the expected prediction variance in case of normally distributed variables in the function y . In the polynomial example this is not fulfilled, since nonlinear transformations of a variable cannot be normally distributed if the variable itself is normal. However, the criterion may still be a good approximate rule for decision, which will be seen in the example.

The estimated standard deviation $s_{n,p}$ in (5) depends highly on the complexity, since increasing the number of parameters p corresponds to smaller errors e , but it also depends on the number of reference observations n through its precision. On the other hand, the second term in the parenthesis increases with increasing complexity (i.e. increasing p) and thereby the criterion gives a trade off between model complexity and prediction uncertainty.

In the example the square roots of the prediction variances were estimated at:

p	1	2	3	4	5	6
$\hat{U}_{10,p}$	0.67	0.78	0.80	0.52	0.67	0.94

and the Breiman/Friedman criterion gives the same result as the visual inspection, the minimum is obtained for $p = 4$, i.e. the third degree polynomial.

Here, only increasing degrees of complete polynomial models have been tested. Of course, in a real case one should also try any subset of the different terms to find the optimal choice, but the limited choice still demonstrate the point: Optimal complexity can be found by means of statistical modelling of prediction uncertainty.

5 A general linear formulation

A generalisation of the statistical idea behind the example gives the following model for a general empirical relationship:

$$Y = E[Y] + \underbrace{\sum_{i=1}^p \theta_i (X_i - E[X_i]) + \sum_{j=p+1}^m \theta_j (X_j - E[X_j])}_{e} + \underbrace{\sum_{k=m+1}^{\infty} \theta_k (X_k - E[X_k])}_{\varepsilon}. \quad (6)$$

Here each influential variable is regarded as a random variable X with its expected value $E[X]$, which means that if the model is applied on the expected values for all X :s, then the result will be the expected value of Y , $E[Y]$. A certain variable X may be a function of some other variable which gives the possibility to include also non-linear terms into the model.

The sum in (6) is partitioned into three parts, the first one with p parameters is the model with optimal complexity, chosen by the criterion above, the second part contains the $m-p$ variables that are neglected in the model, because their influence is not large enough to add any useful information to the model. The third part contains an unknown number of variables that are not known or not measurable and which thereby represent the random contribution to the model.

A certain chosen model with p parameters can then be written:

$$Y = \theta_0 + \sum_{i=1}^p \theta_i (X_i - E[X_i]) + e, \quad (7)$$

where the error e may be modelled as a random variable if the model is used on a population of neglected influential variables $\{X_k; k = p+1, p+2, \dots\}$ (cf. eq. (6)). This means that if the outcome of the X -variables are randomly chosen according to an appropriate distribution, then the statistical procedure with confidence- and prediction intervals will be valid for each future use of the model on the same population.

The given abstract approach for treatment of empirical models gives insight in complexity problems and is helpful for the judgement of the validity of a certain model. However, in a specific situation the assumptions behind the approach are difficult to fulfil. The difficulties include:

- 1) A non-linear behaviour may violate the assumptions about Gaussian distributions of the X variables.
- 2) True representative populations of the X :s are not easy to establish, which makes the random choice of reference tests difficult. This is in particular true for the unknown influentials $\{X_k; k = m+1, m+2, \dots\}$.
- 3) The values of X :s included in the model may not be known completely in reference tests, but they are subjected to uncertainty.

These difficulties must be considered in each specific situation, and necessary approximate solutions be decided. The problem of non-linearities and non-Gaussian distributions of the variables can in a specific case be overcome by replacing the simple criterion (4) with simulated prediction variances.

6 A fatigue example

We will here apply the given approach to the problem of modelling high cycle fatigue life. Restriction is here made to uncertainty sources within the group *Cyclic material behaviour* in Figure 1. The concept is, however, equally applicable to all uncertainty sources shown in Figure 1 above. In order to diminish non-linearity problems we choose a log transformation and regard the logarithm of the fatigue life as a linear function of a number of influential variables:

$$\begin{array}{ll}
 \text{The Basquin model} & \ln N = \theta_0 + \theta_1 \ln \Delta S + \\
 \text{Neglected} & \theta_2 f_2(S_m) + \theta_3 f_3(seq) + \theta_4 f_4(T) + \theta_5 f_5(freq) + \quad (8) \\
 \text{Not measurable} & \left\{ \begin{array}{l} \theta_6 f_6(a_0) + \theta_7 f_7(C) + \theta_8 f_8(S_{op}) + \theta_9 f_9(G_{size}) + \\ \theta_{10} f_{10}(G_{conf}) + \theta_{11} f_{11}(G_{orient}) + \theta_{12} f_{12}(HV_{local}) + \\ \sum_{k=13}^{\infty} \theta_k (X_k - EX_k) \end{array} \right.
 \end{array}$$

A division of terms is introduced in accordance with eq. (6). In formula (8) the first line is the logarithm of the classical Basquin equation: $N = \alpha \cdot \Delta S^{-\beta}$ with $\theta_0 = \ln \alpha$ and $\theta_1 = -\beta$. The second line contains the influential variables mean stress, load sequence, temperature and loading frequency. The following two lines show examples of known influentials which usually are not measurable in the design stage, namely the initial crack length, the crack geometry, local grain properties such as size, configuration and orientation, and local hardness. Finally, the sum in the last line represents all influentials that are not known in the fatigue damage process. This choice of model complexity, neglecting all influentials except the load range, is common in industrial applications and in Svensson [3] this is roughly motivated by complexity arguments.

The presented formulation gives a special interpretation of systematic model errors in simplistic empirical modelling. When using the simple Basquin model one will avoid systematic model errors if the reference tests for the estimation of $\{\theta_1, \theta_2\}$ are performed on a random choice of neglected variable values, i.e. for instance using load mean values from the population of those load mean values that will appear in future applications of the model. If one succeeds to do such a choice, the reference test will give information about the variability around the model and, with an additional distribution assumption, confidence- and prediction limits can be calculated.

The resulting variability may be too large for an efficient design and then one must consider improvements of the model. This can be done in two different ways:

- 1) Increase the complexity in the model by including some of the neglected variables. This will *decrease the number* of variables in the error term e and thereby give possibilities for more precise predictions, but it demands more reference tests and more measurements of variables.
- 2) Narrow the application of the model to a subset of the population and put corresponding restrictions for the future use of the model. This will *decrease the variance* of the existing variables in the error term e and give more precise predictions.

A typical combination of these two ways is to use different Basquin equations for different classes of stress mean values, i.e. partition the population of mean values into a small number of subsets and estimate one pair of Basquin parameters $\{\theta_1, \theta_2\}$ for each

subset. This will result in larger overall complexity, since more parameters are used, without the need for estimating any parameter for the mean value influence.

Conclusion

A methodology has been introduced and exemplified for handling the total uncertainty during fatigue life determination. An optimum complexity could be chosen for the incorporated models in order to balance the use of measurement data between, on one hand, calibration of the model and on the other, improving its predictive capabilities.

The investigation sheds light upon the doubtful efficiency of investments in tools and modelling work for specific parts of the product development chain without looking at the complete chain of activities with all uncertainties, *cf.* figure 3 below.. The organisation of many companies with different specialists that minimise the uncertainty for their subactivities, respectively, is not optimal in this respect. Also the use of different program packages along the development chain that disguise the uncertainties within the individual steps and which are not analysed in the context of the complete life time calculation uncertainty are also questionable. Efficiency requires a comprehensive view of all the testing and calculation activities.

Figure 3. The influence of different sources of error on the fatigue life time determination.

References

- [1] Bernauer, G. and Brocks, W., Micro-mechanical modelling of ductile damage and tearing – results of a European numerical round robin, *Fatigue and Fracture of Engineering Materials and Structures*, Vol. 25, 2002, pp. 363-384.
- [2] Breiman, L. and Freedman, D., How many variables should be entered in a regression equation? *Journal of the American Statistical Association*, Vol. 78, No. 381, Theory and Methods Section, 1983, pp. 131-136.
- [3] Svensson, T., Complexity versus scatter in fatigue modelling, *Fatigue and Fracture of Engineering Materials and Structures*, Vol. 27, 2004, pp. 981-990.

SP Swedish National Testing and Research Institute develops and transfers technology for improving competitiveness and quality in industry, and for safety, conservation of resources and good environment in society as a whole. With Swedens widest and most sophisticated range of equipment and expertise for technical investigation, measurement, testing and certification, we perform research and development in close liaison with universities, institutes of technology and international partners.

SP is a EU-notified body and accredited test laboratory. Our headquarters are in Borås, in the west part of Sweden.

SP Building Technology and Mechanics
 SP REPORT 2005:15
 ISBN 91-7848-922-9
 ISSN 0284-5172

SP Swedish National Testing and Research Institute

Box 857
 SE-501 15 BORÅS, SWEDEN
 Telephone: + 46 33 16 50 00, Telefax: +46 33 13 55 02
 E-mail: info@sp.se, Internet: www.sp.se

A Member of

United Competence