Inverse Problems in Structural Dynamics

M.I. Friswell

Abstract
The forward problem simulates the response of a given system, and is routinely used for performance prediction and optimisation. In contrast, the inverse problem takes response data and deduces information concerning the underlying model. In structural dynamics potential applications include model updating and structural health monitoring. This paper gives an overview of inverse problems, which is formulated as an optimisation problem. The similarities and differences with design optimisation are outlined, and the application of model updating within a design optimisation environment is discussed.

1 Introduction
The goal of inverse problems such as model updating or structural health monitoring is to improve a mathematical model of an existing structure using measurements performed on this structure. Inverse problems may be considered as optimisation problems, although a characteristic feature of inverse problems is that they may be ill-posed. Model updating will be considered first, as most health monitoring schemes are based on similar procedures. In model updating, two models with compatible responses need to be available, namely the mathematical model to be improved, usually obtained by the finite element method, and the experimental model which serves as a reference for updating, for example natural frequencies and mode shapes. These two models have to be specified for identical environments, such as boundary conditions, loads, operating history, etc. In structural dynamics the most common measured outputs are the modal properties, although frequency response functions may also be used. The parameterisation of the mathematical model may take a variety of forms, and examples will be outlined later in the paper.

It is instructive to compare and contrast inverse methods and design optimisation. In one sense they are very similar, in that they both consist of objectives and constraints related to the structure’s response, geometry and material properties, and an optimisation is performed. However there are subtle differences in the form of the objectives and the constraints, and this leads to significant differences in approach. In particular the objective of model updating is to correct uncertainties in the model, and therefore must include some physical insight into the potential modelling errors and an understanding of the measurement errors. Structural health monitoring takes this a stage further with the assumption that damage only changes a small number of parameters. These considerations are not present in design optimisation, which is able to allow changes to more parameters, with perhaps a larger number of constraints. The optimisation methods commonly used in inverse problems may be
considered quite basic, and often comprise gradient based searches. The reason for this is a practical one related to the application; if the cost function is very complex with many local minima, then small changes in the measured data, or the weights applied to the optimisation, may yield very different model parameters. Such sensitivity reduces the robustness of the model, and a better approach is to obtain more information, either through further measurements or improved regularisation.

2 Model updating methods


2.1 The direct methods

The objectives of model updating have changed subtly over the last 30 years, with many of the early methods being classified as direct. These methods update complete structural matrices, so that the updated matrices are those closest to the initial analytical matrices, but reproduce the measured data. Baruch [4] described these methods as reference basis methods, since one of three quantities (the measured modal data, and the analytical mass and stiffness matrices) is assumed to be exact, or the reference, and the other two are updated. Friswell et al. [12] extended the direct methods to update both the stiffness and the viscous damping matrices based on measured complex modal data. The direct methods have a long history, but two issues have restricted their use in practice. Usually measurements contain errors, and forcing the model updating procedure to reproduce the measured modal data causes these errors to be propagated to the parameters. Even worse, the stiffness matrix element values are dominated by the high frequency modes, whereas the low frequency modes are measured. Secondly, all the elements in the structural matrices are changed, and physically meaningful parameters cannot be chosen. The matrices are not guaranteed to be positive definitive, and connectivity of the original model is not enforced (although later work by Kabé [20] attempted to solve this problem). The combination of changing a large number of parameters to reproduce a small number of response measurements leads to an ill-conditioned estimation problem that is usually regularised using minimum norm type solutions, thus spreading localised modelling errors through the structure.

2.2 Sensitivity methods

The most popular methods for model updating minimise the errors in the response in a weighted least squares sense, by changing a pre-selected set of physical parameters. The optimum solution is usually obtained using sensitivity methods, and for well chosen parameters a simple linear approximation to the error surface is sufficient. The modal data is often used, where the sensitivity information is relatively easy to calculate. Frequency domain data may be used directly as the output error approach where the errors in the predicted frequency response functions are minimised, although this leads to a highly non-linear optimisation problem, with many local minima that have small radii of convergence. The alternative is to minimise the errors in the equations of motion, which are often linear in the unknown parameters, although the errors in the response prediction are not minimised directly.

The sensitivity methods allow a wide choice of parameters to be updated and both the measured data and the initial analytical parameter estimates may be weighted. This ability to weight the different data sets gives the method its power and versatility, but requires engineering insight to provide the correct weights. The basic approach is little changed from the key papers of Collins et al. [7,8], where the weighting of the natural frequencies were based on the statistical properties of the measurements. This approach will be considered in more detail later.

2.3 Error localisation

The great power of the sensitivity based updating methods is that they allow a wide choice of parameters to update. In a typical structure the number of potential parameters is huge, and to ensure a well-conditioned estimation problem the number of parameters identified should be relatively small. Only those areas of the model that are uncertain should be updated; the updated model will be physically unreasonable if areas that can be modelled accurately are updated. But how are these areas determined? Clearly engineering insight is a great
help, and engineers often know from experience that areas such as joints are likely to contain modelling errors. Automatic methods may be used. For example, if eigenvalues and eigenvectors are measured then the error in the eigenvalue equation may be used (Lallement and Piranda, [21]). This is an approach that localises the errors on a degree of freedom basis, and thus gives areas of the model that should be investigated further. Other methods localise errors on a element basis using energy (Link and Santiago, [22]). An alternative approach is subset selection (Friswell et al., [14]) where the optimum subset of a large number of candidate parameters is chosen, that is best able to fit the measured data. It is usually difficult to rely totally on automatic methods, and although they can help, the final decision should be taken based on a detailed understanding of the dynamics of the structure.

The other issue that should be addressed is the sensitivity of the response to the chosen parameters. Clearly the parameters should be sensitive, otherwise the updated parameter estimates are likely to be poor. Sometimes changing the boundary conditions of the test can make a big difference to the sensitivities, if for example a joint in the modified experimental arrangement is under strain when originally it was not. Low sensitivities may imply that the information content of the measurements is insufficient to estimate the parameters. However the converse is not true. Just because a parameter is sensitive does not mean it should be included in the model. If the parameter is likely to be estimated accurately in the initial model, then there is no reason to update it.

### 3 Sensitivity methods in model updating

Sensitivity based methods allow a wide choice of physically meaningful parameters and these advantages has led to their widespread use in model updating. The approach is very general and relies on minimising a penalty function, which usually consists of the error between the measured quantities and the corresponding predictions from the model. Parameters are then chosen that are assumed uncertain, and these are usually estimated by approximating the penalty function using a truncated Taylor series and iterating to obtain a converged solution. If there are sufficient measurements and a restricted set of parameters then the identification may be well-conditioned. Often some form of regularisation must be applied, and this is considered in detail later. Other optimisation methods may be used, such as quadratic programming, simulated annealing or genetic algorithms, but these are not considered further in this paper. Problems will also arise if an incorrect or incomplete set of parameters is chosen, or even worse, if the structure of the model is wrong.

Friswell and Mottershead [13] discussed sensitivity based methods in detail. The approach minimises the difference between modal quantities (usually natural frequencies and less often mode shapes) of the measured data and model predictions. This problem may be expressed as the minimization of $J$, where

$$J(\theta) = \|z_m - z(\theta)\|^2 = \varepsilon^T \varepsilon$$  \hspace{1cm} (1)

and $\varepsilon = z_m - z(\theta)$. Here $z_m$ and $z(\theta)$ are the measured and computed modal vectors, $\theta$ is a vector of all unknown parameters, and $\varepsilon$ is the modal residual vector. The modal vectors may consist of both natural frequencies and mode shapes, although often mode shapes are only used to pair individual modes. If mode shapes are included then they must be carefully normalised, the sensor locations must be carefully matched to the finite element degrees of freedom and weighting should be applied to Eq. (1). Frequency response functions may also be used, although a model of damping is required, and the penalty function is often a very complicated function of the parameters with many local minima, making the optimisation very difficult. The modal residual in Eq. (1) is a non-linear function of the parameters and the minimization is solved using a truncated linear Taylor series and iteration. Thus the Taylor series is

$$z_m = z_j + S_j \delta \theta_j + \text{higher order terms}$$  \hspace{1cm} (2)

where

$$z_j = z(\theta_j), \quad S_j = S(\theta_j), \quad \delta \theta_j = \theta_m - \theta_j.$$
The matrix \( S_j \) consists of the first derivatives of the modal quantities with respect to the model parameters, index \( j \) denotes the \( j \)th iteration and \( \theta_m \) is the parameter vector that gives the measured outputs. Standard methods exist to calculate the modal derivatives required (Friswell and Mottershead, [13]). By neglecting higher order terms in Eq. (2), an iterative scheme may be derived, using the linear approximation,

\[
\delta z_j = S_j \delta \theta_j
\]

where \( \delta z_m = z_m - z_j \) and \( \delta \theta_j = \theta_{j+1} - \theta_j \).

Often, for damage location studies, only the residual and sensitivity matrix for the initial model are used. Avoiding iteration reduces the computation required, particularly where multiple parameter sets have to be estimated. However, particularly if the damage is severe, there is a risk that the wrong location is identified.

As indicated above, one of the problems with sensitivity methods is the need for a parametric model of the damage. Mottershead et al. [24] proposed an approach where the system was constrained so that unknown stiffnesses are replaced with rigid connections. The constraint is not imposed physically but the behaviour inferred from the unconstrained measurements. The best fit between the measured and predicted data is obtained when the damage is located in the substructure that is made rigid.

### 3.1 Regularisation

The treatment of ill-conditioned, noisy systems of equations is a problem central to finite element model updating (Ahmadian et al. [3], Friswell et al. [15], Titurus and Friswell [30]). Such equations often arise in the correction of finite element models by using vibration measurements. The emphasis is usually to reduce the number of parameters to ensure that the equations are well conditioned. Even so, the proper treatment of the noisy data is important. Furthermore this reduction in the number of parameters is sometimes not possible or desirable.

In model updating the relationship between the measured output is generally non-linear. In this case the problem is linearised using a Taylor series expansion and iteration performed until convergence. At each iteration a set of linear equations given by \( Ax = b \) has to be solved for the unknown vector \( x \), related to the parameters. When the coefficient matrix is close to being rank deficient then small levels of noise may lead to a large deviation in the estimated parameters from its ‘exact’ value. The solution is said to be unstable and the estimation problem is ill-conditioned. A different problem occurs when the number of parameters exceeds the number of measurements, so that the estimation equation is under-determined and there are an infinite number of solutions. The Moore-Penrose pseudo-inverse provides the solution of minimum norm, as does singular value decomposition (SVD). For the over-determined case, the SVD will again result in the minimum norm solution. This is a form of regularisation which has been widely applied in the model updating community. Unfortunately minimum norm solutions rarely lead to physically meaningful updated parameters.

One solution to the problem of ill-conditioning is to select only a subset of the parameters for updating, and this is considered further in the section on health monitoring.

#### 3.1.1 Side constraints

Model updating often leads to an ill-conditioned parameter estimation problem, and an effective form of regularisation is to place constraints on the parameters. This constraint could take the form of minimising the deviation between the parameters of the updated and the initial model, or alternatively the differences between parameters could be minimised. For example, in a frame structure a number of ‘T’ joints may exist that are nominally identical. Due to manufacturing tolerances the parameters of these joints will be slightly different, although these differences should be small. Therefore a side constraint is placed on the parameters, so that both the residual and the differences between nominally identical parameters are minimised. Thus the penalty function is

\[
J(\theta) = \|Ax - b\|^2 + \lambda^2 \|Cx - d\|^2
\]
where $A$ and $b$ represent the linearised estimation problem, $C$ and $d$ represent the constraint, and $x$ is the vector of unknown parameters. The regularisation parameter, $\lambda$, is chosen to give a suitable balance between the residual and the side constraint. The constraints should be chosen to satisfy Morozov’s complementation condition, so that the regularised problem is well-conditioned.

### 3.1.2 The singular value decomposition

The singular value decomposition (SVD) of the coefficient matrix $A$ may be written in the form,

$$A = U \Sigma V^T = \sum_{i=1}^{m} \sigma_i u_i v_i^T$$

(5)

where the singular values, $\sigma_i$, are arranged in descending order, and $u_i$ and $v_i$ are the left and right singular vectors. The solution to the linear equation $Ax = b$ is then

$$x = \sum_{i=1}^{m} \frac{u_i^T b}{\sigma_i} v_i.$$  

(6)

Thus the components of $A$ corresponding to the low singular values only have a small contribution to $A$ but a large contribution to the estimated parameters. The elements of these singular vectors (corresponding to the low singular values) are also generally highly oscillatory. Equation (6) shows the noise will be amplified when $\sigma_i < u_i^T b$, and this may be used to decide where to truncate the singular values. If $A$ does not contain noise then the singular values will decay to zero whereas the $u_i^T b$ terms will decay to the noise level. Ahmadian et al. [3] considered this approach in more detail.

The standard SVD is incapable of taking account of the side constraint, as this requires the generalised SVD. Space does not permit a full explanation of the generalised SVD, and the reader is referred to Friswell et al. [15] and Hansen [19] for more complete detail of the decomposition. The regularisation parameter, $\lambda$, dampens the effect of the lower singular values (lower than about $\lambda$) and thus smoothes the solution.

### 3.1.3 ‘L’ curves and coss validation

Thus far the value of the regularisation parameter, $\lambda$, has not been determined. In model updating some idea of the measurement noise may be available, but the estimation of the noise (really errors) in the model is very difficult in any quantitative way. The ‘L’ curve approach plots the norm of the side constraint, $\|Cx - d\|$, against the norm of the residual, $\|Ax - b\|$, obtained by minimising the penalty function, Eq. (4), for different values of $\lambda$. Hansen [18] showed that the norm of the side constraint is a monotonically decreasing function of the norm of the residual. He pointed out that for a reasonable signal-to-noise ratio and the satisfaction of the Picard condition, the curve is approximately vertical for $\lambda < \lambda_{\text{opt}}$, and soon becomes a horizontal line when $\lambda > \lambda_{\text{opt}}$, with a corner near the optimal regularisation parameter $\lambda_{\text{opt}}$. The curve is called the ‘L’-curve because of this behaviour. The optimum value of the regularisation parameter, $\lambda_{\text{opt}}$, corresponds to the point with maximum curvature at the corner of the log-log plot of the ‘L’-curve. This point represents a balance between confidence in the measurements and the analyst’s intuition.

The idea of cross-validation is to maximise the predictability of the model by choice of the regularisation parameter $\lambda$. A predictability test can be arranged by omitting one data point, $b_k$, at a time and determining the best parameter estimate using the other data points, by minimising Eq. (4). Then for each of the estimates, the missing data is predicted and the value of $\lambda$ that on average predicts the $b_k$ best is found, in the sense of minimising the cross-validation function.
\[ V_0(\lambda) = \frac{1}{n} \sum_{k=1}^{n} \left( b_k - \tilde{b}_k(\lambda) \right)^2 \]  

where \( \tilde{b}_k(\lambda) \) is the estimate of \( b_k \) obtained from the remaining data. This is the method of cross-validation. The calculations may be conveniently computed via the generalised SVD. Ahmadian et al. [3] gave the background and further details.

4 Parameterisation of the finite element model

Parameterisation is a key issue in finite element model updating. It is important that the chosen parameters should be able to clarify the ambiguity of the model, and in that case it is necessary for the model output to be sensitive to the parameters. Usually elements in the mass and stiffness matrices perform very poorly as candidate parameters, and this is one reason why the direct methods of model updating are not favoured. One reason for this poor performance is that the stiffness matrix element values are dominated by the high frequency modes, whereas the low frequency modes are measured. Element parameters, such as the flexural rigidity of a beam element, may be used provided there is some justification as to why the element properties should be in error. Mottershead et al. [25] used geometric parameters, such as beam offsets, for the updating of mechanical joints and boundary conditions. Gladwell and Ahmadian [17] and Ahmadian et al. [1] demonstrated the effectiveness of parameterising the modes at the element level, and used both geometric parameters and element-modal parameters (i.e. the so-called generic element method) to update mechanical joints. The following subsections will concentrate on the modelling of joints, since these are often the most difficult areas of a structure to model.

4.1 Physical and geometric parameters

There are a number of physical parameters of a joint that could be updated. A beam with a flange welded as a ‘T’ joint, shown in Figure 1, will be taken as an example. The beam part was of length 0.4 m and cross-section 70 mm × 12 mm. The flange area was 110 mm × 70 mm and the thickness of the flange was 6 mm. Pairs of bolt holes, of diameter 12 mm and 40 mm apart, were drilled 25 mm from the edge of the longer part of the flange and 15 mm from the edge of the longer part of the flange, as shown in Figure 1. Only vibration in a single plane was considered. The resonances of the structure are lightly damped and well separated, making natural frequency identification and mode shape pairing straightforward.

The beam structure was modelled with cubic beam elements for the transverse motion and linear bar extension elements. The nodes possess axial and transverse translation degrees of freedom together with a rotation in the same plane. The beam was represented by eight elements, each of the flanges was represented by five elements, and nodes were located to coincide with the bolt holes. The shaded area is considered rigid and is enforced by a constraint matrix linking nodes a, b and c. Only the degrees of freedom corresponding to node c are independent and included in the model. The mass and inertia of the rigid area were lumped at node c.

One approach to updating this joint is to alter the beam stiffness of the elements closest to the joint. Although this often gives good results, the model error is not localised at the joint, but spread through the updated elements. Flexibility may be introduced into the rigid area by using discrete translational and rotary springs between nodes a, b and c. Mottershead et al. [25] showed that for typical joints the structure’s response is insensitive to the stiffness of these discrete springs, and such insensitivity causes great problems for the updating algorithms. A powerful alternative is to update geometric parameters, for example the offset of nodes b and c from node a, denoted by \( \alpha \) and \( \beta \) in Figure 1. The offset parameters have a physical meaning with regard to stiffness updating: the shaded (rigid) region in Figure 1 can be considered to expand or contract depending upon whether the offset dimensions are extended or reduced by updating. The offset dimensions are assumed to only affect the stiffness matrix and the mass matrix is unaffected. A third parameter, the variation in the thickness of the beam and flange was used to allow for a global shift in all the modelled natural frequencies. The beam was tested twice: under free-free conditions and clamped at the flanges. Updating was carried out using both sets of natural frequency data, using a sensitivity based approach. Mode shape data was measured but only the natural frequencies were used for updating. The mode shape data was used to check the pairing of the experimental and analytical modes, although since the natural frequencies were well separated (and the structure basically one
dimensional) this task is straightforward. The spatial incompleteness of the measured is not an issue here because
the number of parameters has been reduced using physical reasoning, resulting in an over-determined
identification problem. Table 1 shows the measured, and the initial and updated analytical natural frequencies,
and Table 2 shows the initial and updated parameter values. The natural frequencies are much improved after
updating. The beam thickness only changes by about 3%, which is within the tolerance that the thickness of the
beams was measured.

<table>
<thead>
<tr>
<th>Natural Frequencies (Hz)</th>
<th>Free-Free</th>
<th>Clamped-Free</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured</td>
<td>324 823 1243 1975 3022 3898</td>
<td>56 354 986 1523</td>
</tr>
<tr>
<td>Initial</td>
<td>318 813 1212 1940 2976 3833</td>
<td>55 349 972 1504</td>
</tr>
<tr>
<td>Updated</td>
<td>325 827 1235 1978 3023 3897</td>
<td>56 356 989 1525</td>
</tr>
</tbody>
</table>

Table 1: Natural frequencies for the welded joint in the short beam

| Initial               | 6.0 | 3.0 | -3.2 |
| Updated - Short Beam  | 6.4 | 3.0 |      |

Table 2: Updated parameters for the welded joint in the beam

4.2 Generic parameters

Gladwell and Ahmadian [17] and Ahmadian et al. [1] introduced the generic element approach. The method
depends on the eigenvalue decomposition of stiffness and mass matrices at the element level, or substructure
masses and stiffnesses typically at a joint. The joint would then be updated by adjusting a set of parameters
related to its own eigenvalues and eigenvectors. It would be possible, for example, to update parameters related
to the bending behaviour in a particular mode of a joint whilst retaining the original stiffness of the other modes.
Model correction using submodel coefficients or physical parameters (such as Young’s modulus or the thickness
of a beam) can be restrictive and may lead to converged models whose physical interpretation does not match the
real structure. The generic element approach is equivalent to modifying the coefficients in the element shape-
function equations but not the order of the shape functions. Generic elements are based on the element (or
substructure) free-free modes but other co-ordinate systems are possible, and might have advantages for
particular updating problems. Often only the eigenvalues of the substructure strain modes are changed, while
keeping the mode shapes unaltered. However other generic parameters have a meaning in terms of the
interaction between the physical modes, which is especially important if substructures are related through
constraints. However it should be noted that the updated generic parameters cannot be interpreted in terms of
physical parameters of the substructure.
As an example, consider the in-plane vibration of a ‘T’ joint in a frame structure. There are 4 nodes, each with 3 degrees of freedom, and the substructure has 4 elements. Thus the joint has 12 degrees of freedom, with 3 rigid body modes and 9 strain modes. Figure 2 shows the strain eigenvectors of the substructure, and show local deformation of the joint. Titurus et al. [31] updated the model of a frame structure using generic elements.

![Substructure eigenvectors for a ‘T’ joint](image)

**Eigenvector 1** **Eigenvector 2** **Eigenvector 3**
**Eigenvector 4** **Eigenvector 5** **Eigenvector 6**
**Eigenvector 7** **Eigenvector 8** **Eigenvector 9**

Figure 2: Substructure eigenvectors for a ‘T’ joint

### 4.3 Equivalent models

Occasionally part of the structure is so ill-defined that no finite element model can be constructed with confidence. Common examples are welded joints in frames and in structures such as automobile bodies. For example the rubber seal which provides the connection between a vehicle window and the car-body structure has a complicated cross-sectional shape into which the window glass and the steel sheet are pressed to form the joint. Furthermore it is important to model the seal accurately because vibration of the window has a strong influence on the acoustics of the passenger compartment. In such cases there seems to be no reasonable alternative to direct parameter estimation. Ahmadian et al. [2] performed tests on a seal with a very stiff foundation that was assumed to be rigid in the model. By using the various physical constraints and the symmetry of the element, the model of the seal was derived from measurements on the rigid foundation.

Spot welds are used extensively in the automotive industry to join panels to make automotive bodies. Extensive efforts are made to accurately simulate these bodies using finite element modelling, and these methods are used to optimise the designs. The most commonly used finite element models of spot welds are the CWELD and ACM2 models (Palmonella et al. [27,28]), and are shown schematically in Figure 3. These models share the great advantage of not constraining the mesh near to the spot welds, and thus no remeshing of the panels to be joined is required. The CWELD model is available in the NASTRAN element library, and is easier to implement than the ACM2 model. Palmonella et al. [28] gave guidelines on their implementation, and in particular considered the mesh size of the welded plates close to a spot weld. The group of elements involved in the definition of the spot weld model is called a patch, and often the area and shape of these patches are very different for different spot welds. Even for a single spot weld, the upper and lower patches may have very different dimensions. The patch area significantly influences the stiffness of the structure, and means that the areas surrounding the spot welds must be meshed with care. Another common practice is match the brick dimensions to the spot weld diameter so that the brick is inscribed on the circle representing the spot weld. In reality the spot weld diameter doesn’t influence the structure’s stiffness, whereas the brick dimensions do
influence the stiffness of the model. Thus changing the brick size to match the spot weld diameter is not acceptable. Palmonella et al. [28] updated models of spot welds in benchmark structures to give guidelines on suitable parameters for patch geometry.

Figure 3: Schematics of the patches associated with the CWELD (left) and ACM2 (right) spot weld models

5 Model updating in the design process

The purpose of model updating is to improve model quality. In the design process, prototypes are manufactured and the experimental response of these structures may be compared to predictions from the finite element model. Clearly changing parameters in a model to make the predicted response match the measured response is not sufficient in this case. The updated model must have physical meaning so that the effect of modifications or design changes to the structure are accurately predicted. In this case one should be very careful about only changing those parts of the structure that are uncertain, and these may be identified by automated methods or physical reasoning. Even so the number of potential parameters is huge and the amount of measured data is limited. It must be realised that model updating is not a panacea, and too much should not be asked of it. Regularisation can help to produce a well conditioned estimation problem, but if the information content of the data is insufficient to produce unique parameter estimates, then some physical meaning will be lost.

The trend in finite element modelling is to generate models with more and more elements and degrees of freedom. But are these models any more accurate? Certainly introducing more elements enables the geometry of the structure to be approximated more closely, and this will produce more accurate estimates of the response of complicated geometric components. However the modelling of joints is still a unsolved problem, and most of the problems with modelling an automotive body-in-white are due to problems in modelling the joints. Here finer models, or higher order elements, do not help very much, as the problem is the error in modelling the physics of the joint. Indeed, if the low frequency dynamics are important then a very fine mesh does seem excessive. The alternative approach, which is central to the model updating philosophy, is to have an equivalent model of the joint, whose properties are updated using measured data.

The difficulties in modelling joints are compounded further by variability in the dimensions and material properties within the joint, with the result that nominally identical components can have significantly different dynamic responses. It might be thought that model updating does not have a solution to this problem, when actually model updating is the solution. It is vital, and it will become increasingly common, that the variability within a structure is propagated through the model, to predict the variability in the response. Considerable work has been performed on the techniques to model and propagate uncertainty, but relatively little has been done to determine the statistical properties of the model parameters (Fonseca et al. [10], Mottershead et al. [26]). Model updating can help by identifying the parameters of a large number of components, from which their statistical properties may be estimated. Once again this relies on the estimated parameters being physically meaningful.

The major problem in model updating is the relatively low information content of the measured data. There are two possible solutions to this. The first is to increase the information content of the data. Perhaps the best solution to this is to test the structure in different configurations, so that the areas of model uncertainty are
stressed in different ways. The alternative is to reduce the number of parameters in the model. Estimating all of the uncertain parameters relating to the joints in a body-in-white will inevitably lead to a large number of uncertain parameters, and therefore an ill-conditioned identification problem. If the objective is to improve the modelling of the joints, then a better approach is to isolate the joint of interest, and use a simplified test geometry for the structure. Of course other problems then arise, and in particular the joint should be forced in a similar way to that which occurs in operation. Furthermore, the boundary conditions of this structure with the isolated joint (or the added components) should be capable of being modelled accurately, so that the parameters of interest can be estimated uniquely and with physical meaning. This leads to the notion of a library of joint models, that includes details of the variability in the parameters. These joint models may then be used with confidence in full models of the body-in-white or other structures.

6 Structural health monitoring

Inverse methods may also be used for monitoring the health of a structure (Doebling et al. [9], Friswell [11]). The four stages of damage estimation, first given by Rytter [29], are now well established as detection, location, quantification and prognosis. Detection is readily performed by pattern recognition methods or novelty detection. The key issue for inverse methods is location, which is equivalent to error localisation in model updating. Once the damage is located, it may be parameterised with a limited set of parameters and quantification, in terms of the local change in stiffness, is readily estimated. Prognosis requires that the underlying damage mechanism is determined, which may be possible using inverse methods using hypothesis testing among several candidate mechanisms. However, once the damage mechanism is determined, the associated model is available for prognosis, and this is a great advantage of model based inverse methods.

A frequent problem that arises in model-based vibration-based damage detection, whether parametric or non-parametric, is the need for a very accurate mathematical model, so that it correctly captures the actual structural dynamic behaviour in some predetermined frequency range. Often in structural health monitoring the changes in the measured quantities caused by structural damage are smaller than those observed between the healthy (i.e. undamaged) structure and the mathematical model. Consequently, it becomes almost impossible to discern between inadequate modelling and actual changes due to damage. There are two alternative approaches to this problem. The first is to update the healthy model so that the correlation between the model and the measured data is improved. This approach requires that the errors that remain after updating are smaller in magnitude than the changes due to the damage. Furthermore the changes to the model should be physically meaningful, so that the updating process corrects actual model errors, and doesn't merely reproduce the measured data. The second approach is based on the use of (relative) differences between data measured on healthy and potentially damaged structure. In this case, assuming that the only changes in the structure are due to damage, the problem may be reduced to finding those parameters that reproduce the measured changes.

Damage usually reduces the local stiffness of a structure. Thus equivalent models, including generic elements, that represent the stiffness of an element or groups of elements may be used for health monitoring. More bespoke models are available for cracks, for composite structures and for distributed damage; Friswell [11] gave a summary of these parameterisations.

6.1 Regularisation using subset selection

The advantages of sensitivity type model updating methods have been highlighted in this paper. However there are significant differences in the application of these methods in model updating and damage location, which necessitates different methods of regularisation. In both cases the number of potential parameters is very large and the estimation process is likely to be ill-conditioned unless the physical understanding can be used to introduce extra information. In model updating, the number of parameters may be reduced by only including those parameters that are likely to be in error. Thus if a frame structure is updated, the beams are likely to be modelled accurately but the joints are more difficult to model. It would therefore be sensible to concentrate the uncertain parameters to those associated with the joints. Even so, a large number of potential parameters may be generated, the measurements may still be reproduced and the parameters are unlikely to be identified uniquely. Using subset selection, where only the optimum subset of the parameters are used for the estimation, has been used for model updating and also for damage location. The reasoning for health monitoring applications is based on the premise that only a limited number of locations are likely to be damaged.
Parameter subset selection is a technique that selects the best subset of parameters from a candidate set, utilising some application dependent cost function that provides a measure of goodness of each subset. Often, these techniques only obtain a sub-optimal estimate of the best subsets in some sense due to the excessive computational burden posed by the original problem. These techniques are firmly rooted in statistics and related fields, although recently applications in structural mechanics have appeared. Friswell et al. [16] gave an overview of subset selection and also proposed the use of this technique for damage detection. They suggested an approach based on forward parameter subset selection, which is especially suited to local damage, and applied the method to a simulated cantilever beam example with physical parameters corresponding to either element or node properties. Different selection and iteration strategies were evaluated, and the case where multiple measurement sets are available was handled by computing the principal angles between two vector subspaces. Titurus et al. [32] used generic elements and subset selection to locate damage in a frame structure.

In damage location statistical methods and performance measures have been used that work on a similar principle (Cawley et al. [5,6]). Only a limited number of sites are assumed to be damaged, and the model updated based on the reduced number of parameters. This process is repeated for all possible combinations of damage site, and possibly even damage mechanism. The results from all the updated models are compared and the one that best matches the measured data is chosen.

The major problem with both subset selection and the statistical type approach, is that many smaller model updating exercises have to be performed. To optimally derive the best set of parameters, or the best damage location, requires the evaluation of many subsets of parameters. With a large number of parameters evaluating all subsets of even 2 or 3 parameters can become daunting. Thus sub-optimal methods must be used to derive good, but not necessarily the best, subsets of the parameters. In the forward approach parameters are chosen one at a time, and the parameters selected previously are retained. However there is no guarantee that the optimal subset will be found. The number of candidate damage locations may be controlled based on the expected reduction in the residual. The addition of a parameter to a previously selected subset inevitably reduces the residual terms, and thus there is a trade off between the number of parameters selected and the magnitude of the residual. Often only a single damage location will be required, in which case the optimal parameter may be determined, or a reasonable number of parameter subsets (say between 3 and 20) are selected for more detailed study.

7 Conclusions
This paper has given an overview of inverse models in structural dynamics. Inverse problems are essentially optimisation problems where the objective function and the constraints have been carefully chosen. The key issues are the choice of measured responses, the choice of parameters to identify and the approach to regularisation to cope with the inherent ill-conditioning.

References