Finite element model correlation and calibration of historic masonry monuments: review

Sezer Atamturktur and Jeffrey A. Laman

SUMMARY

Building a reliable finite element (FE) model of historic masonry structures is a difficult undertaking due to the challenges in accurate representation of irregular geometry, complex material behaviour and complicated boundary conditions between structural masonry components. Model calibration refers to correcting the inherent deficiencies within the FE model by matching outputs to measured data. Model calibration has the potential to produce more reliable computer models and thus aid in the economical management and maintenance of contemporary as well as heritage structures. Researchers involved in the assessment of historic masonry monuments have devoted decades of consistent attention to FE model calibration. This paper reviews the recent developments in this topic with a focus on complex vaulted masonry monuments. Studies on simpler forms of masonry structures, such as masonry arch bridges or masonry towers, are also discussed since they lay the groundwork for studies on more complex structures. This paper identifies several remaining technical challenges as model calibration approaches gain wider recognition and usage in historic monument structural assessment. Copyright © 2010 John Wiley & Sons, Ltd.

1. INTRODUCTION

Growing interest in the preservation of historic structures has created a need for tools capable of reliably analyzing unreinforced masonry monuments. These structures are typically composed of complex interactions between arches, vaults and buttresses. For the structural appraisal of these monuments, finite element (FE) analysis tools have been increasingly utilized over the last three decades. Compared with the graphical or semi-graphical analysis methods proposed by Heyman (1966), FE analysis provides a generally applicable and straightforward approach - especially when investigating complex, three-dimensional interactions between structural components. The versatility of FE analyses in incorporating various constitutive laws, as well as practically all geometric configurations, has resulted in its widespread use by the masonry society (Figure 1).

In the early 1970s, FE methods were first seriously applied to the analysis of masonry monuments (Sachs, 1971). Questions regarding the adequacy of FE methods were immediately raised. Mark, Abel and O’Neill (1973) initiated one of the earliest efforts to confirm FE method adequacy in simulating masonry structure behaviour. Photoelastic tests on small-scale plastic samples of Gothic vaults were conducted to identify internal stress distributions under a gravity load (Figure 2). The measured stress resultants and support reactions were compared with the corresponding FE solutions. Later, this approach was extended to wind loading conditions (Mark, 1982). Today, the adequacy of the FE method for various structural analysis problems has been confirmed and widely accepted. However, the accuracy of an FE simulation is determined by the accuracy of its input parameters. The current challenge in FE analysis is in properly implementing physically substantiated input parameters in order to increase accuracy. For typical masonry structures, the most difficult input parameters to quantify are the boundary conditions between structural components and the material properties of masonry and mortar assembly. When developing FE models, particularly for historic masonry monu-
ments, there are numerous opportunities to enter imprecise input parameters that can result in unsuitable models and erroneous solutions.

Progress has been made in comparing FE solutions with measurements for various contemporary civil structures (e.g., bridges, frame buildings, towers, stadiums, etc.), a procedure commonly known as model correlation. As the need for structural assessment of historic buildings increases, the model correlation concept has been increasingly applied to the analysis of masonry structures such as bridges (e.g., arch bridges), towers (e.g., bell towers, minarets), buildings (e.g., residential, public), and monuments (e.g., churches, mosques, basilicas). Typically, when the FE solutions compare favourably with the corresponding measurements, the FE model is considered to be accurate. However, if the comparison yields a mismatch, the observed discrepancy is attributed to the deficiencies in the FE model due to either imprecise model parameters (e.g., imprecise values for the Young’s modulus of stone) or erroneous modelling decisions (e.g., improper boundary conditions).

Researchers in other fields developed methods to reduce these FE model deficiencies. Using measurements, one can fine-tune the appropriate model input parameters and improve the agreement between FE simulations and measured data, a process commonly known as model calibration. The concept of FE model calibration was first developed within the context of linear dynamics (Motterhead and Friswell, 1995). During the calibration process, imprecise parameters are adjusted until the FE
model output reproduces acceptable results as compared with the physical evidence. In this context, physical evidence that is relevant to the identified deficiencies in the FE model should be obtained. In the absence of a formal quantitative approach, the relevance of physical evidence to model deficiencies is typically decided based on engineering judgment. For example, a heat transfer experiment would be unlikely to provide relevant information for a model solving dynamics problems.

Because researchers involved in the assessment of historic masonry monuments have devoted decades to FE model calibration, what is now needed is a review of the extent to which these prior studies have been successful and an appraisal of the remaining technical challenges. This paper provides that overview with a focus on complex vaulted monumental masonry structures such as churches and cathedrals. Studies on simpler forms of masonry structures such as masonry arch bridges or masonry towers are also discussed as they lay the groundwork for studies on even more complex structures. Although there are several alternative analysis methods developed for the assessment of masonry structures, the scope herein is strictly confined to FE analysis.

The earliest efforts to correlate FE simulations with measurements, discussed in Section 2, date from the 1980s and range from qualitative visual comparisons based on field observations to quantitative comparisons of static tests. In the late 1980s, non-destructive, in situ dynamic testing gained popularity in the assessment of masonry structures. Section 3 highlights these experimental studies, and Section 4 details the use of dynamic experiments for FE model calibration. The studies discussed in Section 4 are deterministic in nature, meaning the model parameters are assumed to be known with certainty and repeated experiments are assumed to yield identical results. In Section 5, these assumptions are challenged with the implementation of stochastic calibration. It is expected that the thorough review of the FE model calibration provided here will encourage wider recognition and implementation of these techniques in the assessment of historic structures.

2. MODEL CORRELATION BASED ON STATIC OBSERVATIONS OR EXPERIMENTS

The governing philosophy of model correlation is the comparison of the FE simulations against in situ measurements. Several studies have achieved this comparison through visual, on-site observations or through static response measurements. Because model correlation is the necessary first-step for model calibration, the studies discussed in this section provide a valuable resource for demonstrating the advantages and disadvantages of visual and static methods.

2.1. Visual methods of correlation

The earliest model correlation effort applied to masonry monuments was the visual comparison of crack locations to analytical estimates of the tensile zones. Mark and Hutchinson (1986) compared available information on existing cracks of the Roman Pantheon against tension region predictions of several alternative FE models. Based on this comparison, the suitability of various modelling strategies was investigated (e.g., modelling of the hemispherical dome with and without the walls). The same approach was applied to the simplified FE model of a historic cathedral by Ricart-Nouel (1991). The simplified FE model was used to investigate the cathedral’s structural behaviour under earthquake excitation. From solutions of the correlated FE model, Ricart-Nouel (1991) derived several conclusions about the necessary structural modifications that included the isolation of cathedral superstructure from the foundations through base-isolators and the addition of a rigid core to resist the lateral forces. These visual inspections of existing cracks are concentrated at a few locations in a historic masonry building. Therefore, this approach is limited in its effectiveness and susceptible to significant error when differential support settlements or long-term creep are present in the structure.

2.2. Static methods of correlation

The problems associated with visual methods can be remedied by destructive and non-destructive static tests focusing on stress, strain, or deflection under controlled loading. This approach was most
widely and successfully used to investigate masonry arch bridge behaviour. While successful when applied to masonry bridges, the use of in situ strain or deflection measurements is impractical for monumental masonry structures due to the difficulty in sufficiently loading the structure to achieve a detectable response. Due to this technical infeasibility, several researchers have attempted to examine the static behaviour of such systems through scaled laboratory models. Also, as it is highly unlikely that destructive tests on a historic monument will be permitted for research purposes, scaled laboratory experiments provide a convenient alternative for investigations of failure regimes in these structures.

Vaulted or domed sections of historic monuments are among the most vulnerable structural forms to structural failure. As such, FE analysis is frequently adapted to simulate collapse mechanisms of vaults, which require the definition of a proper, damage criterion. To verify the accuracy of novel damage criteria incorporated into FE models, destructive tests on scaled vault specimens are frequently completed. One successful example was conducted by Creazza et al. (2001). In this study, the three-dimensional behaviour of a masonry barrel vault reinforced by FRP was simulated through an FE model. Next, a barrel vault scaled model was tested to failure under a vertical, quasi-static load located around the quarter span. The load was increased until the vault formed an unstable mechanism (Figure 3). The measured displacement and load at failure were then compared with the FE model predictions. The same study was repeated a year later on a ribbed, cross-vault, scaled masonry model (Creazza et al., 2002). Creazza et al. focused on the locations and magnitudes of maximum strain and deformation as well as on the characteristics of the collapse mechanism under slowly increasing static load.

Theodossopoulos et al. (2002, 2003) investigated the behaviour of masonry cross-vaults through static tests conducted on scaled wood models. These models represented an aisle vault of the partially collapsed Abbey Church of Holyrood in Edinburgh. First, the strains and displacements under gravity loading were experimentally recorded. Subsequently, the scaled model was tested to failure by progressive abutment movement in order to characterize the failure. A nonlinear FE model was built using a smeared crack approach along with a biaxial bending damage model. Although the FE solutions for deformation patterns agreed well with the scaled model, an order of magnitude difference was observed between strain measurements and predictions. Theodossopoulos et al. attributed lack of model predictivity to the averaging of strains (as well as deflections) in the mortar joints and

Figure 3. The failure mechanism of the barrel vault (Creazza et al., 2001, with permission).
masonry units. This study is a relevant example that demonstrates the difficulties in modelling material behaviour of inhomogeneous masonry and mortar assembly.

Obtaining physical evidence through scaled model tests enables examination of the behaviour under load regimes that usually cannot be tested on an actual structure, e.g. collapse mechanisms. Physical evidence obtained from destructive tests is the only way to assess the accuracy of the damage criterion incorporated in FE algorithms. However, problems occur if scaled models are used to make inferences about an existing structure. Scaled test models typically only represent a portion of the existing structure and thus exclude the actual elastic restraint exerted by adjacent elements or boundary conditions. Such a deficiency means that the alternative load paths within the structure would also be absent from the analysis. While analyzing existing structures, the problems associated with testing scaled models combined with the difficulties in conducting in situ static tests lead one to use in situ small amplitude dynamic tests. These studies are discussed in the next section.

3. STUDIES BASED ON DYNAMIC EXPERIMENTS

Dynamic characteristics of masonry monumental structures are significantly different than those of contemporary reinforced concrete and steel buildings. As a result of an experimental program conducted on a series of Gothic cathedrals, Atamturktur et al. (2009) investigated the practical issues related to the testing of monumental masonry structures. Atamturktur et al. emphasized that the connectivity of two masonry walls involve factors that depend on contact pressure, surface friction, and existing cracks, as well as the elastic behaviour of each stone unit and mortar joint. The interaction of these factors typically yields a rather flexible connection between structural components that causes local modes to be more pronounced relative to global modes. As a result, the structural component connectivity, load distribution, and dynamic response are affected by the amplitude and location of the excitation. This aspect has also been noted by Sortis et al. (2005) and Zembaty and Kowalski (2000). Also, complex vault geometry results in a large number of closely spaced modes with low participation factors (see Figure 4). Furthermore, high dissipative forces in a masonry assembly make identification of modes with low participation factors difficult.

Despite these challenges, both traditional and operational modal analysis techniques have been successfully applied to masonry structures. The practical and technical differences between these two non-destructive testing procedures were recently evaluated by Atamturktur, Fanning and Boothby (2007) through references to tests conducted on the choir fan vaults of the Washington National Cathedral. Impact hammer excitation was used during traditional modal analysis tests while excitation from the peal bells, carillon bells, an orchestra and chorus, an organ, and ambient noise were exploited separately for operational modal analyses. The acoustic input of the first four excitation types was

![Figure 4. The first three modes of the choir vaults of National Cathedral.](image-url)
measured by microphone and found to approximate a uniform acoustic excitation. Although modes with high participation factors were observed to be consistent for both test techniques, certain modes obtained by traditional modal analysis were not detected by operational modal analysis.

In past research, dynamic testing techniques have been applied to both scaled models in controlled laboratory conditions and existing structures in operational conditions. Because controlled laboratory experiments are largely immune to environmentally caused complications, higher quality measurements are typically obtained. Studies on scaled laboratory models are discussed in the first part of this section. However extending in situ dynamic testing beyond laboratory tests is still necessary to incorporate practical difficulties and gain a realistic view about the limits of dynamic testing. These studies, which address the difficulties of testing an existing structure, are discussed in the second part of this section.

3.1. Dynamic tests on scaled laboratory models

The early studies on this topic focused on scaled arch bridge models. Pretlove and Turner, in their pioneering 1985 study, reported a possible 5% reduction in certain natural frequencies of a beam-like structure upon the development of a severe crack. For a masonry arch bridge loaded monotonically to failure, Brown et al. (1995) reported up to a 10% reduction in natural frequencies between the onset of damage and the formation of first hinge (also see Pretlove and Ellick (1988), Pretlove and Ellick (1990), and Ellick and Brown (1994)). As the damage in the arch bridge was increased, these previous studies noted a progressive reduction in natural frequencies. This reduction is related to the load carrying capacity of the damaged arch bridge through a serviceability threshold criterion.

Armstrong et al. (1995a) also noted significant variations in natural frequencies in their study investigating the spandrel wall separation of arch bridges. Otherwise identical, two arch bridge models were built, both with and without spandrel wall separation and tested in the laboratory under impact hammer excitation. Aside from changes in frequencies, Armstrong et al. observed distortions in mode shapes as well as changes to the mode sequence. While, investigating the influence of additional mass on the bridge, the authors observed up to a 10% reduction in natural frequencies and as high as 400% deviations in damping ratios. Accurate identification of damping ratios, especially from masonry structures, is a very difficult task, and this study illustrates this challenge.

In the studies discussed above, modal parameters, such as natural frequencies and mode shapes, provided physically meaningful and low dimensional features for the comparison of two datasets. However, comparisons of frequency response functions (FRF) were also used as a convenient, higher dimensional alternative to modal parameters. For instance, Bensalem et al. (1998) investigated the natural frequencies and FRF amplitudes measured from scaled arch bridge models loaded monotonically to failure. The first natural frequency of the arch model consistently declined from 40 Hz (in its intact stage) to 8 Hz (in its unstable stage). With the development of first cracks, the FRF magnitudes decreased over the entire frequency range. However, as the damage level gradually increased, FRF magnitudes were observed to increase for lower frequencies and decrease for higher frequencies. In a later study, Bensalem et al. (1999) concluded that FRF amplitudes are the most reliable indicators of presence and size of voids in the arch bridge backfill (see Figure 5) (also see Bensalem et al., 1995 and Bensalem et al., 1997).

Another comparative feature is the structure dynamic stiffness that is the inverse of the displacement FRF amplitude at zero Hz. Armstrong et al. (1995b) repeated their previous study using measured dynamic stiffness of the arch bridge models and concluded that dynamic stiffness is sensitive to the spandrel wall separation. Although this approach is a convenient way to measure the stiffness of a structure, it may yield erroneous results if the FRF measurements are contaminated by noise. This often occurs due to the high noise floors of accelerometers at low frequencies.

Scaled masonry building models have also been a research subject of much interest. Vestrioni et al. (1996) completed experiments on a 1/5th scale masonry building under mechanical vibration excitation. First, the scaled model dynamic characteristics were investigated by inducing small amplitude vibrations. When the baseline modal parameters were obtained, forces with successively increasing amplitudes were applied to the scaled model to induce structural damage. In agreement with the findings of Armstrong et al. (1995b), a reduction in the dynamic stiffness due to structural
damage was observed through dynamic measurements. Ramos et al. (2005) also had success in establishing a relationship between damage level and natural frequencies during his study on a full-scale masonry building built of rubble stone. Increasing numbers of cracks were induced by shaking tests. Modal identification through operational modal analysis was performed at each damage state revealing consistently decreasing natural frequencies as the damage level increased. However, a direct relationship between the crack patterns and dynamic response was not evident (see Figure 6).

Ramos (2007) conducted a similar study on masonry arch and wall models built with clay bricks of low compressive strength and mortar with poor mechanical properties to represent the typical
material present in historic construction. Cracks were successively induced in the scaled models through controlled static tests. In between these tests, operational modal analysis was performed to identify the modal parameters. Consistent with earlier studies, Ramos (2007) also observed a significant reduction in natural frequencies as damage increased. Ramos was successful in accurately identifying the damping ratios and noted a gradual increase in damping ratios after damage. He also noted that the modal assurance criterion, that is the correlation between the mode shapes, generally remains unchanged before and after the damage.

3.2. Dynamic tests on existing structures

The studies discussed in this section inherently have more difficulties in accurately extracting comparative features. One of the sources of difficulty is environmental conditions. Brown et al. (1995) reported a maximum 3% change in the first three natural frequencies of masonry arch bridges due to variations in temperature and rainfall when temperatures below freezing were excluded. However, when temperatures below freezing are considered, the change in natural frequencies was observed to be as high as 12%. The authors explain this high variation as being due to changes in the infill soil dynamic behaviour of upon freezing. The effects of temperature and rainfall were later studied by Ramos (2007) for an ancient monastery. As a result of long-term monitoring, Ramos (2007) reported an average 6% variation in frequencies due to annual temperature fluctuations. Also, absorbed moisture increased the mass of the stone units and reduced the stiffness of the mortar joints. During his studies on a masonry clock tower, Ramos (2007) reported a 4% reduction in natural frequencies of a clock tower with the beginning of the rainy season. These two studies quantify the potential variability due to annual changes in environmental conditions.

Because the studies discussed in this section are not characterized by extensive test campaigns, their measurements are not sufficiently numerous for statistical evaluation and thus must be considered as a snapshot representation of reality. As an example, the dynamic behaviour of an old, masonry building was investigated by Genovese and Vestroni (1998). The acceleration response of the building was recorded due to small-amplitude, forced oscillations. The FRFs obtained at increasing excitation levels were compared with investigate the nonlinear characteristics of the masonry structure. Similar to their earlier study in the laboratory (Vestroni et al., 1996), Genovese and Vestroni (1998) observed a reduction in the stiffness of the structure under increased force levels. This conclusion, however, strictly depends on the excitation force levels. For low amplitude vibration, other studies have reached different conclusions; Zonta (2000) found the frequency response to be independent of vibration amplitudes. The experiment focused on a Roman amphitheatre that had been experiencing structural problems due to aging and material deterioration (Figure 7). Zonta observed linear behavior in the frequency response of the wing wall. Likewise, for small vibration amplitudes, Atamturktur et al. (2009) observed a linear response when testing the masonry vaults of National Cathedral in D.C.

Figure 7. Section Roman Amphitheatre (Zonta, 2000 with permission).
However, once the excitation levels were increased beyond 400 lbs, the resonant peaks of FRF exhibited reduced frequency and increased damping.

Variations in natural frequencies successfully detected damage in the study completed on 534 stone pinnacles of the Palace of Westminster in London by Ellis (1998). Ellis experimentally identified the natural frequencies of the pinnacles and compared them to each other. A total of five pinnacles with a significantly lower fundamental frequency were successfully identified as damaged pinnacles (Figure 8).

Vibration-based studies, just as they were used to detect damage, were also used to detect structural improvements after retrofit or strengthening projects. Turek et al. (2002) conducted ambient vibration analysis on a recently repaired historical church. The natural frequencies identified before and after retrofit were compared and an increase in the dynamic stiffness was observed. Increased dynamic stiffness after retrofit was also observed by Antonacci et al. (2001) in a similar study on a historic basilica and by Ramos (2007) on a historic masonry tower. These two studies extended their experimental campaign to the calibration stage and are discussed in Section 4.

In numerous studies, natural frequencies have proven themselves to be sensitive to a change in the structure, which makes natural frequencies suitable comparative features for calibration. Though mode shapes have exhibited distortion in certain studies, the success of these studies in relating mode shapes with damage has been rather limited. The FRFs contain a larger amount of information regarding the frequency content and the modal mass and damping of the structure. Therefore, FRFs exhibit varying sensitivity to the changes in the structure, depending on the severity and type of this change.

4. DETERMINISTIC MODEL CALIBRATION

As a result of the experimental studies, which were discussed in the previous section, significant improvements have been made in the dynamic identification of masonry structures. However, solely experimental methods are limited in their spatial resolution. As only a limited set of measurements can be taken from the structure, the spatial incompleteness poses a greater problem when monumental structures are concerned. The available dynamic measurements, however, can be used to calibrate FE simulations that permit prediction of a more complete set of structural responses.

Model calibration begins from the premise that the FE solutions to predict phenomena of interest (A, in Figure 9) can be improved by calibrating the appropriately selected input parameters according to the physical evidence provided by experimental measurements (B, in Figure 9). When the calibra-
Physical phenomena where estimates are available:
- spring constant estimates via FE model

Physical phenomena of interest where measurements are not available:
- tensile stresses, maximum deformations

Figure 9. Calibration of the imprecise input parameter of the numerical model by the use of comparative features.

Model calibration seeks to determine the most probable values for poorly known input parameters though comparison of FE solutions against in situ measurements. During this comparison, comparative features are a bridge between simulations and measurements. Like other inverse problems, in model calibration, ill conditioning is a potential problem if the quality or quantity of the comparative features is insufficient. The success of model calibration depends not only on selecting the right comparative features but also in calibrating the right parameters. The calibration parameters must be selected according to the combined effects of parameter uncertainty and parameter sensitivity. Parameter uncertainty can be determined from a family of specimen tests or by prior knowledge of the structure. The sensitivity of the model parameters can be determined by a sensitivity analysis that measures the changes in the model outcomes due to a unit change in the model input. After the comparative features and calibration parameters are identified, the model calibration is a matter of updating the calibration parameters based on the functional relationships between the measured and calculated comparative features.

The properties that can be calibrated are closely tied to the quantity, quality and type of comparative features. The studies discussed in this section invariably adopt comparative features based on linear dynamics, specifically, modal parameters. Because modal parameters are related to both structure mass and stiffness, no direct information about the ultimate load capacity or damage scenarios can be determined from these comparative features. To learn such information, comparative features must be derived from destructive testing in the form of maximum strain, stress, displacement, and collapse mechanism. Moreover, the load levels that may induce damage in a masonry structure would induce nonlinearities in the masonry structure. Therefore, if the ultimate use of the FE model is to investigate damage in masonry structures, the material properties, and perhaps some boundary conditions, must be built to represent nonlinearities. In turn, the model comparative features must contain relevant information about these nonlinearities. Although studies discussed in this section are invariably limited to linear FE models, they are the necessary first steps before extending the model calibration approach to nonlinear FE model with damage criterion. In sections 4.1 and 4.2, two calibration approaches are discussed. The first approach is manual calibration; a trial-and-error based approach calibrating selected parameter values according to engineering judgment and intuition. The second approach is automated calibration; performed by constructing a series of loops based on optimization procedures or Bayesian inference.
4.1. Manual deterministic calibration

Manual calibration, aided by engineering judgment, is an appealing and convenient approach for calibrating FE model parameters. However, by its nature, manual calibration cannot incorporate uncertainties. When there are sources of high uncertainty that challenge understanding of structural behaviour, manual calibration may compensate for errors caused by one factor by adjusting another. It is also likely that, through manual calibration, the hidden dependencies or correlations potentially present between input parameters will be revealed. If these dependencies or correlations are strong, this will raise the problem of the calibration of one parameter compensating for imprecision in another. Although eliminating these problems is fundamental for accurate manual calibration, many published papers have practically ignored these criteria.

Manual calibration of parameters can be justified when the initial model is a close representation of reality, which can be confirmed during test-analysis correlation. In this case, typically, after calibration, the parameters are only minimally adjusted and they, in effect, maintain their physical meaning. Manual calibration tends to be successful if deficiencies arising from imprecise parameters are independent and uncorrelated. The literature is replete with many successful examples of this type of calibration. The majority of these studies characterize the development of linearly elastic FE models and calibrate the poorly known Young’s modulus of the appropriately selected construction material using modal parameters as comparative features.

Antonacci et al. (2001) conducted transient dynamic tests to obtain natural frequencies and global mode shapes of a basilica. By comparing FE model predictions against the experimentally identified, first four natural frequencies and mode shapes, the ratio of the Young’s modulus of the upper and lower portions of an ancient basilica are adjusted. Antonacci et al. then used the calibrated FE model to investigate the structure static behaviour before and after the repair and strengthening. Arêde et al. (2002) completed a similar study on an ancient monastery church. The poorly known Young’s modulus of the surcharge infill was calibrated based upon the experimentally obtained first four natural frequencies and mode shapes. The calibrated model, which delivered better agreement with the experiments, was used to assess the seismic vulnerability of the structure. Multi-tiered masonry temples in Nepal were the subject of a similar study by Jaishi et al. (2003) in which three temples were tested with output-only techniques. The first three experimental bending modes in each orthogonal direction were identified for each temple and then paired with the numerical modes calculated by the corresponding FE models. The identified frequencies were used to manually calibrate the uncertain material properties of the mud-brick walls of the temple. For some temples, Young’s modulus magnitudes were significantly reduced to achieve better agreement.

Erdogmus (2004) applied such an approach to the vaults of Gothic cathedrals. Erdogmus identified the first axis-symmetric mode of the choir vaults of a 20th-century cathedral built based on medieval construction techniques. This one mode was used as a reference to manually adjust the FE model boundary condition and material properties. The calibrated FE model was then used as a baseline for the development of FE models of two other complex vaulted historic churches (Figure 10). Atamturktur (2006) and, in a later interpretation, Atamurktur and Boothby (2007) completed a complementary study on two masonry tile domes. These studies obtained high-quality test data and identified ten clear mode shapes. These studies, using non-destructive and destructive techniques to identify the material properties of the tile and mortar, confined calibration parameters to boundary conditions. Upon completion of manual calibration, FE model predictions compared favourably with the first eight experimentally identified natural frequencies and mode shapes (Figure 11). Bayraktar et al. (2008) were the first to investigate the dynamic characteristics of an old masonry minaret. Here, the minaret’s modal parameters, identified through output-only modal analysis techniques, were used to calibrate the material properties and boundary conditions of the model. Through such calibrations, they were ultimately able to achieve favourable agreement between the test and analysis. The studies discussed in this section show the success of modal parameter based FE model calibration for a wide range of structures. This approach itself is a proven applicable and versatile means of fine-tuning the typically poorly known material properties of masonry structures.

Júlio et al. (2008) applied a similar manual calibration procedure to a clock tower adjacent to a faculty building at the University of Coimbra in Portugal. The tower, of rubble-stone construction.
with coarse stone masonry at the corners, exhibited degradation of joints, cracking of stone blocks, and biological colonization. The tests were conducted using operational modal analysis techniques. The restraints imposed on the tower by the adjacent building walls and slabs, as well as the soil structure interaction, were uncertain. Julio et al. had to make several assumptions regarding these interactions to represent them with idealized boundary conditions commonly found in FE tools. These established boundary conditions were then adjusted through trial-and-error until the authors found an acceptable agreement between simulations and measurements for the first five mode shapes. Julio et al. acknowledged that, without a survey of geometry and material of the surrounding structural components, it was not possible to validate the accuracy of the final boundary conditions. Once the boundary conditions were determined, the material properties of the tower walls were tuned to achieve a better agreement of frequencies. The calibrated model was then used to draw inferences about the structural integrity of the tower. The approach used by Júlio et al. uncoupled the calibration of boundary conditions and material properties where the boundary conditions are calibrated based on mode shapes and the material properties are calibrated based on natural frequencies. The same approach has been applied to a half-scale Guastavino dome-scaled model by Erdogmus (2008). However, the uncoupling of the boundary condition and material property calibration must be applied with caution.
when there are multiple material types in the FE model. As illustrated by Antonacci et al. (2001), mode shapes are sensitive to the relative ratios of the material property values. Therefore, if these ratios are erroneous while adjusting the boundary conditions using mode shapes, incorrect boundary conditions may be obtained that would ultimately result in imprecise material properties. This issue can be resolved by calibrating all of the poorly known parameters simultaneously as in Gentile and Saisi (2007) and Atamturktur (2009). These studies are discussed in the next section.

4.2. Automated deterministic calibration

The masonry tower study by Júlio et al. (2008) was similar to Gentile and Saisi (2007) in that both studies represented poorly known structural interaction of the respective towers with the walls of an adjacent structure. While Júlio et al. used idealized boundary conditions, such as hinged or fixed connections to represent this interaction, Gentile and Saisi used linear springs with a spring constant requiring calibration. The tower studied by Gentile and Saisi also showed signs of partial damage in the form of extensive vertical cracks that altered the wall stiffness. In addition to the spring constants representing the interaction between the tower and the adjacent structure, Gentile and Saisi (2007) selected the poorly known Young’s modulus values of the defective walls as calibration parameters. Based on the extent of the damage, the exterior walls were defined in six distinct regions with independent material properties. Instead of uncoupling the calibration of boundary conditions and material properties, Gentile and Saisi calibrated all of the parameters simultaneously using modal parameters obtained through ambient vibration testing (Figure 12). Calibration was completed by both the Inverse-Eigen-Sensitivity (IES) method and Douglas-Reid (DR) method. Through calibration, the difference between each of the first five theoretical and experimental natural frequencies was minimized. The findings obtained from both methods consistently yielded lower Young’s modulus values in the damaged regions when compared with the undamaged regions, thus supporting the potential of vibration-based model calibration methods to deliver useful information on the damaged state of masonry structures.

Using the manual calibration method discussed in Section 3, parameters are calibrated, often the Young’s modulus of masonry, which has a global effect on the structural response. However, with the aid of automated calibration algorithms, it is possible to effectively calibrate the individual properties of each FE. Such an approach also enables the representation of local variability of material properties and localized cracks, an important uncertainty source for masonry structures. Aoki et al. (2005); and in their refined dynamic model (Aoki et al., 2008), present results of material property variability and localized cracks through dynamic identification where FE modelling and an FE model calibration campaign was applied to a brick chimney. The first three modes in two orthogonal directions were identified by Autoregressive Moving Average and Eigensystem Realization Algorithm techniques. The chimney FE model was built with 20-node, isotropic, solid elements assuming a fixed

![Figure 12. The masonry tower and first five vibration modes identified from ambient testing, (Gentile and Saisi, 2007, with permission).](image-url)
support at the base. A correction factor is assigned for both the mass and stiffness of each FE. Using Inverse-Eigen-Sensitivity (IES), the elemental mass and stiffness correction factors were sought. To mitigate the inevitable incompleteness of the measurements, a weighting function was applied to eliminate FEs with low sensitivity. Aoki et al. determined that elemental stiffness was reduced at the base of the chimney due to the chimney/soil interaction, and was increased at the chimney corners due to the iron angles.

Ramos (2007) conducted tests on a masonry clock tower that exhibited substantial material degradation, biological growth and loss of material due to lack of proper maintenance. Similar to Turek et al. (2002) and Antonacci et al. (2001), Ramos (2007) investigated the dynamic behaviour of the tower before and after a strengthening rehabilitation. Ramos (2007) observed that the tower vibrated at higher natural frequencies after the retrofit while the damping coefficients were observed to be lower. It must be emphasized that the study achieved a remarkable correlation of the first five theoretical and experimental mode pairs. This is perhaps due to the fact that the test structure was a standalone tower without uncertain interactions with adjacent structures. The first five natural frequencies and mode shape vectors were used to calibrate the FE model by the nonlinear least square method. Ramos also obtained significantly lower Young’s modulus results for the walls where damage was dominant.

5. STOCHASTIC MODEL CALIBRATION

The studies discussed in the previous section, implementing either manual or automated calibration algorithms, assume that both FE simulations and in situ dynamic measurements are deterministic. The purpose of these studies was to establish a direct connection between the analytically and experimentally derived comparative features. However, FE models contain uncertainty in their parameters and experiments contain uncertainty in their measurements. In this section, an overview will be presented of studies intended to reach a statistical correlation between the simulations and measurements by formulating the FE model input parameters and FE model output response probabilistically.

Antonacci et al. (2000); and in a later, refined version Sortis et al. (2005), presented a study on a two-story stone masonry structure. These studies collected vibration measurements from the structure due to low-amplitude vibratory forces induced by shakers placed at four different locations. The modal parameters extracted from these measurements were used to calibrate individual exterior wall segment Young’s modulus values for the corresponding FE model, based on a nonlinear output error approach. Both the input parameters and the output error were treated as random variables with a normal distribution. The optimal parameter values, which yield maximum posterior distributions for the input parameters and minimum for the nonlinear objective function, were sought. The calibration parameters, selected based on their significance according to the Fisher matrix, were Young’s modulus of four exterior wall sections. The discrepancies between the experimental and analytical modal parameters, upon calibration, were within the measured frequency variations obtained by exciting four different locations.

Theoretically, if all the modes of a structure are identified experimentally, the optimization-based calibration can be expected to calibrate all the input model parameters. However, because the available resources limit the available experimental information to a limited number, optimization-based calibration tends to be ill-conditioned. One approach for circumventing this problem is to eliminate those insensitive calibration parameters, thus reducing the overall number of parameters. The study below describes such a procedure.

De Stefano (2007) conducted dynamic experiments on a baroque chapel masonry dome using four different excitation sources: ambient, hammer, dropped object and wind turbulence caused by a helicopter. As a result, the first six modes were identified. The structure was divided into a number of substructures in which the material of each substructure was assumed to be homogenous. To represent the interaction of the chapel with the neighbouring buildings, elastic springs were added to the model, and the most influential parameters were selected based on a sensitivity analysis. The initial probability distributions of the selected calibration parameters, also known as a priori distribution, were defined as uniform. The test-analysis disagreement for the first five natural frequencies (the cost func-
tion) was minimized through a Probabilistic Global Search algorithm. The algorithm explores the domain defined by the calibration parameters, generating multiple alternative models to be run. Among these models, the algorithm selects those that show reasonable agreement with the measurements. Next, the calibration parameters of the models that passed the first elimination are perturbed consecutively, and at each iteration, the probability distribution of these calibration parameters is updated. The last step of this multi-model approach is clustering the final set of models that fit the minimum error requirements. From this approach, the clustered five alternative models, which only differ from each other for the values of the calibration parameters, were selected. The study incorporates uncertainty in the calibration parameters by treating them probabilistically and ultimately offers alternative FE models among which an analyst can make the selection based upon engineering judgment. In this study, engineering judgment is incorporated with the sophisticated and computationally intensive calibration procedure remarkably well.

Although these two earlier studies are extraordinary in both success and sophistication, neither account for the experimental variability, unlike Atamturktur’s (2009) masonry cathedral study. Here, the author integrated large amounts of experimental and computational information collected from testing the choir vaults of the National Cathedral, DC. Measurement uncertainty was assessed from 48 replicated experiments and FE model uncertainty was incorporated through experimental design. A total of 100 computer experiments, which explored the model parameter domain, were run by perturbing model parameters. From both the measurements and numerical analysis comparative features, the natural frequencies and mode shape vectors were extracted probabilistically as mean and variance statistics. Using a Phenomenon Identification and Ranking Table, the uncertain parameters deemed candidates for calibration were ranked based on the sensitivity of test-analysis comparative features. Atamturktur used Bayesian inference to compound the prior knowledge about the calibration parameters together with experimental observations collected from vibration testing. Prior probability distribution incorporated expert judgment while the variance of measured features accounted for the experimental uncertainty. Bayesian inference was then applied and resulted in updated knowledge of the calibration parameters in the form of a posterior probability distribution (Figure 13). It must be emphasized that the Bayesian approach to calibration differs from optimization based calibration techniques. The concept of calibration in the Bayesian approach is handled with successive propaga-

![Diagram](image)

Figure 13. The masonry tower and first five vibration modes identified from ambient testing.
tions of input and output uncertainty and evaluated as the characterization of the probability distributions of parameters.

6. DISCUSSION AND CONCLUSIONS

This literature review reveals the common need for analysts to find FE solution supporting evidence. Mark’s photoelastic studies of plastic scaled models (Mark 1982), Fanning’s scaled masonry bridges and recent applications of in situ dynamic tests convey one common message (Fanning et al., 2001): until the model is validated with physical evidence, numerical predictions must be treated with due caution. These studies illustrate that testing of existing buildings yields very useful information about its response characteristics. Additionally, experimental measurements integrated with computerized FE tools makes it possible to gain a thorough understanding of the structural behaviour. Although experimental measurements are always incomplete in the sense of their spatial resolution, they play a critical role in model calibration that ultimately yields mathematical representation of the global structural behaviour. However, there are several issues that remain to be addressed.

Manual calibration studies—that is, the calibration of uncertain parameters to improve the agreement between calculations and measurements by trial and error—have a benefit, as they incorporate engineering judgment and expert opinion into the calibration process which keeps the calibration from converging to an unrealistic model. However, because calibration parameters are treated as deterministic, they have limitations in incorporating the uncertainties into the calibration. Manual calibration studies become more successful in the absence of a structural configuration difficult to interpret by engineering judgment (e.g., the complex interaction between two masonry components). In the presence of several sources of uncertainty, manual calibration of material properties will likely compensate for the errors introduced by an inappropriate boundary condition.

While automated calibration studies, commonly based on optimization techniques, can be stochastic and can incorporate uncertainties, they are not typically seen as incorporating expert opinion that can be very successfully achieved through manual calibration approaches. One alternative is the use of Bayesian methods to characterize calibration parameters that can account for both uncertainty and expert opinion.

In the calibration of masonry system FE models, uncertainties originate from many different sources. Therefore, calibration must be stochastic and account for uncertainties in both the experimental measurements and the model definition. Because the tasks required for stochastic model calibration require extensive resources and expertise, they are rarely undertaken. However, the stochastic approach is necessary to bring calibration of analytical models into the analytical mainstream.

Most masonry structures exhibit a very complex, inelastic, and nonlinear dynamic behaviour, making the experimentation and comparative feature selection very difficult. Features with strong linearity assumptions, such as modal parameters, tend to smear the effects of inelastic nonlinear behaviour, degrading the quality of the calibration studies. Features that are not deeply rooted in strong linearity assumptions, unlike modal parameters and frequency response functions and their derivatives, warrant particular attention.

For FE models, which are built to model the ultimate load capacity of historic structures, the most likely calibration parameters are the nonlinear material properties, the coefficients of the damage criterion as well as the parameters of the nonlinear connections. Therefore, comparative features for these studies must contain information about the behaviour of the structure at damage levels. Commonly employed non-destructive vibration testing would not provide relevant information for these FE models.

To gain validity and to quantify the accuracy of an FE model, an independent set of experimentally derived information, other than that used in the calibration, is necessary. Until this step is completed, there can be no justification of the FE solutions, and the term ‘validation’ should not be used (Trucano et al., 2007). This requirement increases the already high demands on resources; however, it is important to distinguish between a calibrated model and a validated model.

The credibility of a calibrated model is increased by the increasing amounts of experimental information accurately reproduced by the calibrated model. In the current state of the art, an objective
criterion to determine the predictive maturity of a numerical model and thus the completion of a calibration exercise is not available; there is a need to develop a measure of sufficiency for experimental information and an indicator of completion for the calibration exercise.

Subcomponent testing, commonly implemented in mechanical and nuclear engineering applications, also warrants attention. Subcomponent testing evaluates components individually (such as piers or walls) and later integrates the calibrated subcomponents into the global structure model. Subcomponent testing can also supply information about local phenomena that may not have a strong effect on the global response of the structure. It must be noted that calibrating a model with global dynamic response measurements of the structure can only be used for phenomena that have a substantial influence on the global behaviour of the structure.

Although the immediate benefits of model calibration are not as obvious in structural engineering as they are in fields where prototyping and mass production are common, the determination of modelling strategies learned through model calibration can ultimately serve the structural engineering community with an improved accuracy with numerical modelling. The extent of research efforts in model calibration is presented herein with a specific emphasis on historic masonry structures. Calibrated FE models will enable a better understanding of historic monument behaviour, and ultimately enable successful repair and retrofit schemes. As a result of this review, it is apparent that the ever-increasing popularity of FE model calibration will result in the routine application of model calibration to a diverse group of masonry structures.

REFERENCES


