ARMAV MODEL TECHNIQUE FOR SYSTEM IDENTIFICATION AND DAMAGE DETECTION

J.B. Bodeux, J.C. Golinval

Université de Liège, LTAS-Vibrations et Identification des Structures,
Chemin des chevreuils, I (Bât. 52), B-4000 Liège, Belgium

SUMMARY: This paper presents an application of ARMAV models in the fields of system identification and damage detection. It is shown how to estimate the modal parameters as well as their uncertainties only on the basis of output measurements, simply assuming that the excitation is a zero-mean stationary Gaussian white noise. The uncertainties estimation can be used for damage detection. Knowing the modal parameters and their uncertainties, it is possible to assess whether changes of modal parameters are caused by e.g. a damage or simply by estimation inaccuracies. The identification and damage detection method is illustrated on the “Steel-Quake” benchmark proposed in the framework of COST Action F3 “Structural Dynamics”. This structure is used at the Joint Research Centre in Ispra (Italy) to test steel buildings performance during earthquakes. The obtained results indicate the effectiveness of the method in estimating modal characteristics and their uncertainties.

KEYWORDS: Prediction Error Method, criterion function, Gauss-Newton algorithm, model order selection, white noise process, covariance matrix, modal parameters uncertainties.

INTRODUCTION

ARMAV models have been applied for the analysis of linear and time invariant systems under ambient excitation. These models only use time series obtained from the output signal of the system. The technique can be used directly to analyse data obtained from the free response or from the forced response due to uncorrelated random excitation [1]. The identification method known as the Prediction Error Method allows to find the model parameters in a non-linear, iterative way (Gauss-Newton algorithm) [2]. The method also incorporates model order selection via Akaike’s Final Prediction Error and Akaike’s Information Theoretic Criteria and structural mode distinction and extraction by use of stability plots. Finally, the model may be validated by examination of statistical tests on the prediction errors. Besides estimating the ARMAV model parameters, the method can also provide an estimate of the covariance matrix of these parameters. On the basis of this covariance matrix, it is then possible to estimate the uncertainties on the modal parameters [3]. The determination of these uncertainties is very relevant for structural monitoring based on vibration measurements. In this way, if the uncertainties of the estimated modal parameters can be computed, it becomes possible to establish a probabilistic confidence in the existence of a damage.
Vibrating structures

The dynamic behaviour of an ambient excited multi-DOF’s linear system can be represented in the state space by the usual form:

\[
\dot{z}(t) = F z(t) + B w(t)
\]  

with

\[
z(t) = \begin{bmatrix} \dot{x} \\ x \end{bmatrix}, \quad w(t) = \begin{bmatrix} f(t) \\ 0 \end{bmatrix}, \quad F = \begin{bmatrix} -M^{-1}C & -M^{-1}K \\ I & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} M^{-1} & 0 \\ 0 & 0 \end{bmatrix}
\]  

where \(M\), \(C\) and \(K\) are the mass, damping and stiffness matrices respectively. \(f(t)\) represents the ambient excitation vector and \(x(t)\) the displacement vector.

The poles of the system are the eigenvalues of the matrix \(F\). For low damped structures, the poles are described by pairs of complex conjugated values:

\[
\mu_r = -\zeta_i \omega_n \pm j \omega_n \sqrt{1 - \zeta_i^2}, \quad r = 1, 2, \ldots, d \text{ (number of degrees of freedom)}
\]  

where \(\omega_n\) and \(\zeta_i\) are the natural frequency and damping ratio of the \(i\)th mode.

ARMAV (Auto Regressive Moving Average Vector) model

For multivariate time series, described by a \(m\)-dimensional vector \(x[n]\), the parametric ARMAV(p,q) model is described by the following matrix equation [1] :

\[
x[n] = \sum_{k=1}^{p} a_k x[n-k] + u[n] + \sum_{k=1}^{q} b_k u[n-k] \quad \text{with} \quad \begin{cases} p = \text{AR order}, \\
q = \text{MA order.} \end{cases}
\]  

\(x[n]\) is the observed vibration vector at discrete time \(t_n = n \Delta t \) (\(\Delta t = \text{sampling interval}\)) and \(u[n]\) is a zero-mean stationary white noise process. \(a_k\) and \(b_k\) are \((m, m)\) matrices of AR (Auto-Regressive) and MA (Moving-Average) coefficients. The AR part describes the system dynamics while the MA part is related to the external noise as well as to the white noise excitation.

In the state space, the ARMAV model can be expressed in the form:

\[
x[n] = a x[n-1] + W[n]
\]  

with

\[
x[n] = \begin{bmatrix} x[n]^T \\ x[n-1]^T \ldots x[n-p+1]^T \end{bmatrix}^T
\]  

The dimension of vector \(x[n]\) is \((p.m, 1)\) and \(a\) is a \((p.m, p.m)\) matrix containing the different coefficients of the auto-regressive part of model. \(W[n]\) includes the MA terms. The ARMAV
model (4) can be applied to data obtained from the free response or from the forced response due to uncorrelated random signals [1].

Considering the covariance principle, it may be shown that an ARMAV(2s, 2s-1) model is the identical discrete model of a \(m\)-variate continuous system with \(d\) degrees of freedom, the number of channels \(m\) being equal to \(d/s\) [4], [5], [6]. If sampled response is affected by measurement noise, the adequate model changes in general to an ARMAV(2s, 2s) model [7].

**Modal parameter estimation**

Considering equation (3) and the analogy between equations (1) and (5), the modal parameters of the system can be extracted from the eigenvalues \(\tau_r\) of the AR matrix \(a\) as in [1] :

\[
\omega_r = \frac{\ln(\tau_r)}{\Delta t} \quad ; \quad \zeta_r = -\frac{\text{Real}(\ln(\tau_r))}{\ln(\tau_r)} \quad ; \quad r = 1,2,..,m.p
\]  

(7)

The mode-shape vectors can also be deduced from the eigenvectors of matrix \(a\).

The number of discrete time ARMAV model eigenvalues is in general larger or different from the number of eigenvalues corresponding to system (1). Therefore, only a subset of the discrete eigenvalues will represent structural modes. The distinction between physical and non-physical modes is performed by use of stability diagrams for increasing AR model order.

**MODEL PARAMETER ESTIMATION**

Let us note by \(\theta\) the model parameters to be determined i.e.

\[
\theta = [a_1 \ a_2 \ ... \ a_p \ b_1 \ b_2 \ ... \ b_q]^T
\]  

(8)

All systems are in principle stochastic, which means that the output \(x[n]\) at time \(t_n\) cannot be determined exactly from data available at time \(t_{n-1}\). Let us define \(\hat{x}[n \mid n-1, \theta]\), the predicted response at time \(t_n\) based on parameters \(\theta\) and on the available data for \(t \leq t_{n-1}\) [2] :

\[
\hat{x}[n \mid n-1, \theta] = \theta^T \varphi[n]
\]  

(9)

where \(\varphi[n]\) is the regression vector defined as

\[
\varphi[n] = \begin{bmatrix}
x^T[n-1] \ ... \ x^T[n-p] \ \epsilon^T[n-1 \mid \theta] \ ... \ \epsilon^T[n-q \mid \theta]
\end{bmatrix}^T
\]  

(10)

Thus the model parameters \(\theta\) are selected so that the prediction error defined as

\[
\varepsilon[n \mid \theta] = x[n] - \hat{x}[n \mid n-1, \theta]
\]  

(11)

becomes as small as possible. For this purpose, the quadratic criterion function \(V_N(\theta)\) that measures the “size” of \(\varepsilon[n \mid \theta]\) is formed :
The estimate \( \hat{\theta}_N \) based on \( N \) samples is then defined by minimisation of the criterion function

\[
\hat{\theta}_N = \arg \min_{\theta} V_N(\theta)
\]

(13)

where \( \arg \min \) means “the minimising argument of the function”. This way of estimating \( \theta \) is called the Prediction Error Identification Method (PEM) [2]. The predictor (9) is non-linear, since the prediction errors themselves depend on the parameters \( \theta \). So the function \( V_N(\theta) \) cannot be minimised by analytical methods. This implies that an iterative numerical minimisation of the function \( V_N(\theta) \) has to be applied. If \( \theta^{(i)} \) represents the \( v \)-dimensional column vector of model parameters at iteration \( (i) \), the iterative method is represented by:

\[
\hat{\theta}^{(i+1)}_N = \hat{\theta}^{(i)}_N + \alpha^{(i)}_N f^{(i)}
\]

with \( \text{dim}(\theta^{(i)}) = v = [p + q]m^2 \)

(14)

where \( f^{(i)} \) is a search direction based on information about \( V_N(\theta) \) acquired at previous iterations, and \( \alpha^{(i)}_N \) is a positive value determined so that an appropriate decrease in the value of \( V_N(\theta) \) is obtained. Here, the Gauss-Newton Method [2] is chosen where \( f^{(i)} \) is defined as

\[
f^{(i)} = -\left[H_N(\hat{\theta}^{(i)}_N)\right]^{-1} V_N'(\hat{\theta}^{(i)}_N)
\]

(15)

The Hessian matrix \( H_N(\hat{\theta}^{(i)}_N) \) and the gradient of the criterion function \( V_N'(\hat{\theta}^{(i)}_N) \) are defined as

\[
H_N(\hat{\theta}^{(i)}_N) = \frac{1}{N} \left[ \sum_{n=1}^{N} \psi[n, \hat{\theta}^{(i)}_N] \psi^T[n, \hat{\theta}^{(i)}_N] \right] \quad ; \quad V_N'(\hat{\theta}^{(i)}_N) = -\frac{1}{N} \left[ \sum_{n=1}^{N} \psi[n, \hat{\theta}^{(i)}_N] \epsilon[n \mid \hat{\theta}^{(i)}_N] \right]
\]

(16)

where

\[
\psi[n, \hat{\theta}^{(i)}_N] = \frac{\partial \hat{x}[n \mid n-1, \hat{\theta}^{(i)}_N]}{\partial \theta}
\]

(17)

is the gradient of the predictor (9), i.e. the derivative of (9) with respect to each of the ARMAV model parameters. The step size \( \alpha^{(i)}_N \) is chosen so that \( V_N(\hat{\theta}^{(i+1)}_N) < V_N(\hat{\theta}^{(i)}_N) \).

The Hessian matrix (16) may be singular or close to singular. This is the case, for example, if the model is over-parameterised or the data not informative enough. Then some numerical problems may arise in (15). One common way to overcome this problem is the Levenberg-Marquardt procedure which consists to use the following approximation of the Hessian matrix [2]:

\[
V_N(\theta) = \frac{1}{N} \sum_{n=1}^{N} \frac{1}{2} \left\| \epsilon[n \mid \theta] \right\|^2
\]

(12)
To start the iterative procedure, a first estimation of $\hat{\theta}_N$ is needed. For this reason, a high-order ARV model is first applied to the response $x[n]$. The prediction error $e[n]$ of this model is then used as external input in an ARX (Auto-Regressive eXtra-input) model. The estimated parameters of this model becomes then the initial estimate [1], [2].

MODEL ORDER DETERMINATION AND MODEL VALIDATION

Model order determination

The identification strategy consists of successive fittings of ARMAV models. ARMAV models of different orders are subsequently examined, and a final candidate model is selected. Model comparison and selection is based upon Akaike’s Final Prediction Error (FPE) and Akaike’s Information Theoretic Criteria (AIC) [7]. According to this principle, the order of an ARMAV model is selected to be the integer which minimises the criteria

$$FPE = V_N(\theta) \frac{1 + wN}{1 - wN} \quad AIC = N \log(V_N(\theta)) + 2v$$

where $N$ is the number of samples and $v$ is the total number of estimated parameters. These criteria include a penalty for badness of fit and for too high order models.

Model validation

Model validation is the final step of the system identification. There are many different ways to check the validity of the ARMAV model. Here, statistical tests of the prediction errors are examined. If the estimated model contains the true system then the prediction errors should be a white noise sequence. This can be investigated by plotting the correlation functions of these errors [2], [7].

ESTIMATION OF MODAL PARAMETER UNCERTAINTIES

It can be established that the PEM method is an asymptotic unbiased and efficient method for Gaussian distributed prediction errors. Let us define $P(\hat{\theta}_N)$ as the model parameter covariance matrix of the difference between the true parameters $\theta_0$ and estimated parameters $\hat{\theta}_N$ as $N$ tends to infinity. The covariance matrix of unbiased estimate $\hat{\theta}_N$ is evaluated, as in [2], [7], from

$$\hat{P}(\hat{\theta}_N) = \hat{E}[(\theta_0 - \hat{\theta}_N)(\theta_0 - \hat{\theta}_N)^T] = \left[ \sum_{n=1}^{N} \psi[n, \hat{\theta}_N] \Lambda^{-1}_N \psi^T[n, \hat{\theta}_N] \right]^{-1}$$

where the sampled covariance matrix of the prediction errors is given by

$$\Lambda_N = \frac{1}{N} \sum_{n=1}^{N} e[n \mid \hat{\theta}_N] e^T[n \mid \hat{\theta}_N]$$
Considering the auto-regressive part of estimated covariance matrix, it is possible to determine modal parameter uncertainties [3].

THE “STEEL-QUAKE” EXAMPLE

The method has been validated using experimental data proposed as benchmark in the framework of the European COST Action F3 “Structural Dynamics”. The “Steel-Quake” structure is used at the Joint Research Centre in Ispra (Italy) to test the performance of steel buildings during earthquakes [8]. The different tests performed correspond to the baseline undamaged and damaged states of the building.

Description and testing of the structure

The structure corresponds to a two-floor frame as depicted in Fig. 1. The main dimensions are 8 m × (4×2) m × 3 m. In the background, it can be observed the reaction wall which supports the 4 pistons (not present in the picture) used to deform the structure (on each side, on each stage) and to induce damage in the x-direction. Note that braces have been added in the plane parallel to the wall to reduce risk of collapse in that direction.

Four excitation points were tested using impact hammer. Their locations are shown in Fig. 2 (points A(x), A(y), B(-z), C(-z); x, y and z indicating the three directions). Eight to ten hammer impacts were recorded for each test. The sensor configuration is the same for all the four tests (Fig. 2). The sampling frequency was 128 Hz and for each channel 3200 data points were captured.

![Fig. 1: View of the Steel-Quake structure.](image1.png)

![Fig. 2: Sensor and excitation configuration.](image2.png)

Structural identification

Pre-processing

The analysis is concentrated on the frequencies below 25 Hz. Therefore a filter with a cut-off frequency of 32 Hz was applied and the data were 2 times decimated resulting in a new sampling rate of 64 Hz and 1600 points per channel.

Model order selection and structural mode distinction
The first step of the identification procedure is the determination of the ARMAV model order. The FPE and AIC criteria lead, for each excitation, to 15-dimensional ARMAV(4,4) candidate models. These models are characterised by an over-determination of the number of poles (the number of model poles is higher than the number of structural poles). The distinction between structural and spurious modes is performed by use of stability diagrams. As an example, the vibration data obtained from the “A(x) excitation” are considered hereafter. Fig. 3.a illustrates the evolution of the FPE criterion applied to different models while the stabilisation diagram is plotted in Fig. 3.b. The sum of the spectra of the 15 measured responses is also plotted in order to observe the localisation of stable modes. The stability of the mode-shape vectors may also be evaluated by computation of the Modal Assurance Criterion (MAC) between estimated mode-shapes obtained from two successive models. The chosen stabilisation criteria are 1% for frequencies, 5% for damping ratios and 1% for MAC values.

![STABILITY DIAGRAM](image)

Fig. 3 : a) FPE criterion and b) stability diagram applied to the “A(x) excitation” data.

**Model validation**

In order to validate the model, the prediction errors are checked. If the ARMAV(4,4) model is adequate, it can be assumed that the true system is contained in the estimated model so that the prediction errors should be a white noise sequence. This is investigated by plotting the correlation functions of these errors with their confidence intervals \[2\]. Fig. 4 gives the result for one channel in the case of data obtained from the “A(x) excitation”. It indicates that the auto-correlation function remains, for the most part, within the confidence interval, except at zero lag. Therefore, the prediction error is close to white noise. The same conclusion has been found for all channels, and therefore the model may be considered as validated.

![Auto-correlation function of prediction error](image)

Fig. 4 : Auto-correlation function of prediction error (horizontal lines indicate 95% confidence levels)
Identification

The identification procedure is applied for each excitation point and repeated for every hammer impact. The mean values and uncertainties for modal parameters are listed in Table 1 and the identified mode-shape vectors of the undamaged structure are shown in Fig. 5.

<table>
<thead>
<tr>
<th>Mode-shape</th>
<th>Undamaged state</th>
<th>Damaged state</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( f ) (Hz)</td>
<td>( \delta f ) (Hz)</td>
</tr>
<tr>
<td>Bending 1X</td>
<td>3.128</td>
<td>0.006</td>
</tr>
<tr>
<td>Bending 1Y</td>
<td>3.928</td>
<td>0.004</td>
</tr>
<tr>
<td>Torsion 1</td>
<td>6.129</td>
<td>0.003</td>
</tr>
<tr>
<td>Bending 2Y</td>
<td>9.687</td>
<td>0.005</td>
</tr>
<tr>
<td>Bending 2X</td>
<td>10.819</td>
<td>0.004</td>
</tr>
<tr>
<td>2nd slab bend. 1</td>
<td>12.271</td>
<td>0.016</td>
</tr>
<tr>
<td>1st slab bend. 1</td>
<td>13.053</td>
<td>0.013</td>
</tr>
<tr>
<td>2nd slab tors. 1</td>
<td>17.694</td>
<td>0.017</td>
</tr>
<tr>
<td>1st slab tors. 1</td>
<td>19.037</td>
<td>0.027</td>
</tr>
<tr>
<td>Torsion 2</td>
<td>21.415</td>
<td>0.018</td>
</tr>
</tbody>
</table>

Bending X

3.128 Hz; \( \delta v = 1.62 \% \)

10.819 Hz; \( \delta v = 0.41 \% \)

Bending Y

3.928 Hz; \( \delta v = 1.33 \% \)

9.687 Hz; \( \delta v = 0.97 \% \)

Torsion

6.129 Hz; \( \delta v = 0.49 \% \)

21.415 Hz; \( \delta v = 1.21 \% \)

Slab bending

12.271 Hz; \( \delta v = 1.92 \% \)

13.053 Hz; \( \delta v = 2.45 \% \)

Slab torsion

17.694 Hz; \( \delta v = 1.54 \% \)

19.037 Hz; \( \delta v = 2.52 \% \)

Fig. 5: Identified mode-shapes and their uncertainties (\( \delta v \)) for the undamaged state (\( \delta v \) represents the mean on components of the mode-shape).
The Modal Assurance Criterion (MAC) was used to compare mode-shapes of the undamaged and damaged structure. The MAC values are represented in Fig. 6 along with the frequency shifts from the undamaged to the damaged state.

From Table 1, it can be observed that the uncertainties on the estimated natural frequencies are very small compared to the uncertainties on estimated damping ratios. The estimates of the damping ratios being not enough accurate, the estimates of frequencies are used for damage detection. The frequencies which present the most significant changes will be used as damage indicators. Fig. 6 suggests that the first and the last five frequencies can be used for damage detection.

Fig. 6 : Comparison of frequencies and mode-shapes in the undamaged and damaged states.

In this particular example, the detected change of some natural frequencies is very significant and so, it is not difficult to find the presence of a damage. In order to illustrate the detection method using ARMAV models, let us analyse frequencies exhibiting small changes. The statistical approach for damage detection is based on confidence intervals obtained from the standard deviation of estimated natural frequencies to detect the damage [7]. It will be assumed that a damage has been detected if the confidence interval of the estimate frequency of a mode is non-overlapping with the 99% confidence interval of the frequency of the same mode in the undamaged state. The estimated natural frequencies of modes 2, 3 and 4 with their 99% confidence intervals are plotted in Fig. 7.

Fig. 7 : Estimated natural frequencies of modes 2, 3 and 4. The estimated 99% confidence intervals are represented by dotted lines.
In Fig. 7, it is seen that the confidence intervals of the three frequencies corresponding to modes 2, 3 and 4 in the damaged state do not overlap the undamaged state confidence intervals, so that damage can even be detected using these frequencies. In the case of overlapping confidence intervals, it should be possible to give a probability to the presence of damage. If frequencies are assumed to be independent distributed variables and that a negative change in frequency indicates a damage caused by structural change, the probability of negative change in frequency can be estimated with the unit normal distribution function [9].

CONCLUSIONS

In this paper, the application of ARMAV models in system identification and damage detection has been presented. The estimation of ARMAV model parameters has been carried out by the Prediction Error Method. The damage detection method is based on the evaluation of modal parameter uncertainties and on the use of statistical tools like confidence intervals and normal distribution of random variables. However, the present investigation is limited to damage detection; the problem of damage localization was not examined. In the Prediction Error Method, the criterion function is minimised using non-linear optimisation. If the application involves many response channels, the iterative updating of the model parameters may require many computations and be time-consuming. Therefore, it will be interesting to investigate another multivariate time series model: the Stochastic Subspace System Identification Method which does not involve any non-linear computations. This method and the Subspace-Based Fault Detection Algorithm are developed in [10].

ACKNOWLEDGEMENT

This text presents research results of the Belgian programme on Inter-University Poles of Attraction initiated by the Belgian state, Prime Minister’s office, Science Policy Programming. The scientific responsibility is assumed by its authors.

REFERENCES

