

University of West Bohemia
Faculty of Applied Sciences
Department of Mathematics



MASTER THESIS

**CORRELATION CRITERIA
FOR ACOUSTIC AND DYNAMIC DATA**

Plzeň 2018

Eva Brabcová

Sperrvermerk

Die vorliegende Diplomarbeit mit dem Titel Correlation criteria for acoustic and dynamic data enthält schützenswerte Informationen der ZF Friedrichshafen AG.

Die Veröffentlichung, Vervielfältigung -auch in digitaler Form-, Weitergabe -auch auszugsweise- sowie das Einstellen dieser Abschlussarbeit in eine Hochschulbibliothek oder öffentlich zugängliche Bibliothek sowie jede weitere öffentliche Zugänglichmachung sind vor Ablauf der unten angegebenen Sperrfrist ohne schriftliche Zustimmung der ZF Friedrichshafen AG nicht gestattet. Ausgenommen von diesem Zustimmungserfordernis ist eine Verwendung der Abschlussarbeit, die aufgrund der jeweils anwendbaren Prüfungs- oder Studienordnung zwingend erforderlich ist.

Die Sperrfrist gilt mit Abgabe der Abschlussarbeit an der Hochschule 3 Jahre.

Der Abgabetermin an der Hochschule für diese Arbeit ist am 18. 5. 2018.

.....

Ort, Datum

.....

Unterschrift Student/-in

ZF interner Hinweis: Schutzklasse der Vertraulichkeit (siehe Richtlinie des ZF-Konzerns KR97/05): **internal**

Declaration

I hereby declare that this Master thesis is my own work and that I used only the cited sources.

Plzeň, July 9th 2018

.....
Eva Brabcová

Acknowledgements

I would like to thank my supervisor, Mgr. Michal Friesl, Ph.D., for his expert guidance, patience and his time he dedicated to me.

I also want to thank my external supervisor, Ing. Josef Otta, Ph.D., for motivating leadership during my internship and all his advices.

I sincerely thank both of them for their willingness to conduct this Master thesis.

Abstract

This thesis is devoted to statistical methods for comparing data from acoustic and dynamic measurements and simulations. Firstly, the modal analysis and statistical methods are introduced, then chosen comparative criteria generalized. In the case of two data sets comparison, we primarily focus on acquirement and inclusion of a mass matrix and extensions for capturing differences in the data. The main results concern the case of multiple data sets, where the basic assurance criteria are generalized through regression models for comparison and differences detection for more input data. Demonstrations of the standard and improved criteria applied on real data are also provided. Within the thesis, an application, written in Python 3.6, for ZF Friedrichshafen AG usage was created as well and the documentation for the application is attached.

Keywords: modal analysis, dynamic measurements, Modal Assurance Criterion and generalizations, linear regression, statistical tests

Abstrakt

Práce je zaměřena na statistické metody pro porovnání dat z akustických a dynamických měření a simulací. Nejprve jsou představeny modální analýza a statistické metody, poté zobecněna vybraná srovnávací kritéria. V případě porovnání dvou datových sad se především věnujeme stanovení a zahrnutí matice tuhosti a rozšířením zachycujících rozdílů v datech. Hlavní výsledky se týkají případu vícero dat, kde jsou základní kritéria zobecněna pomocí regresních modelů pro porovnání a detekci rozdílů pro více vstupních dat. Součástí práce jsou také ukázky kritérií aplikovaných na reálná data. V rámci práce byla vytvořena aplikace, napsaná v Python 3.6, pro účely společnosti ZF Friedrichshafen AG a proto je připojena rovněž dokumentace k aplikaci.

Klíčová slova: modální analýza, dynamická měření, Modal Assurance Criterion a zobecnění, lineární regrese, statistické testy

Contents

1	Introduction	1
2	Modal analysis	3
2.1	Experimental modal analysis	4
2.2	Experimental mass and stiffness matrices - Pseudoinverse method	6
2.2.1	Tikhonov regularization	7
2.2.2	IC symmetric	9
2.2.3	Schulz method	10
2.2.4	PR2-Schulz method	10
2.2.5	Comparison of methods in terms of condition number . . .	11
3	Statistical modeling	13
3.1	Statistical models	13
3.1.1	Linear model of regression	13
3.1.2	Model of two-way ANOVA with interactions	16
3.1.3	Model of three-way ANOVA with interactions	17
3.2	Statistical testing	18
3.2.1	Statistical tests for regression	19
3.2.2	Statistical tests for ANOVA	20
4	Correlation criteria for two data sets comparison	23
4.1	Nonselective criteria	24
4.1.1	MAC – Modal Assurance Criterion	24
4.1.2	FDAC – Frequency Domain Assurance Criterion	26
4.1.3	MWMAC – Mass Weighted Modal Assurance Criterion . .	26
4.1.4	PLR – Pairing and Linear Regression	29
4.2	Selective criteria	31
4.2.1	COMAC – Coordinate MAC	32
4.2.2	SC – Selective Criterion	33
4.2.3	SC with COMAC	35
4.2.4	SDOFC – Selective Degrees of Freedom Criterion	36
4.2.5	SDOFC with COMAC	36
4.3	Criteria applied on real data	37

5	Statistical processing of multiple data sets	47
5.1	Basic criteria	48
5.1.1	MM – More Measurement	48
5.1.2	MMV – More Measurements and Variants	54
5.1.3	MMV 2	59
5.2	Scheffe tests	62
5.2.1	SMMV - Scheffe test for MMV	62
5.2.2	SMMV 2 - Scheffe test for MMV 2	65
5.3	Possible Extensions	66
5.4	Criteria applied on real data	67
5.4.1	Data 5×2	67
5.4.2	Data 10×4	81
6	MAC application	91
6.1	Requirements	91
6.1.1	Installation of Anaconda package	91
6.1.2	How to run the MAC app	92
6.2	MAC and MAC MV apps	94
6.2.1	File selections	94
6.2.2	Format of selected data	97
6.2.3	Starting the calculations	99
6.3	MAC app Input Data – Output Options	101
6.3.1	Input Data	101
6.3.2	Criteria	102
6.3.3	Criteria Parameters	103
6.3.4	Output Options	104
6.4	MAC MV app Input Data – Output Options	104
6.4.1	Input Data	104
6.4.2	Criteria	105
6.4.3	Criteria Parameters	106
6.4.4	Output Options	106
6.5	Useful remarks	107
6.5.1	One more option how to use the MAC app	107
6.5.2	How long the calculations should approximately take	109
6.5.3	Other comments	110
6.5.4	Program code	111
A	Matrices X and Y	115
	Nomenclature	117
	Bibliography	119

Chapter 1

Introduction

The text focuses on the topic of modal analysis and the statistical methods for data processing from acoustic and dynamic measurements and simulations. The standard criteria are usually used for comparison of measurement and simulation and for updating numerical models describing the real system or for a comparison of two measurement data with displaying differences in vibrational behaviour (e.g. visualizing which system is stiffer in terms of vibration) [1], [2], [3]. The main aim of this work was to generalize the basic criteria primarily for comparison of more data sets and apply them to real data and create an application with standard and improved criteria for ZF Friedrichshafen AG usage. The program was written in Python 3.6.

The text is composed of five main chapters.

The chapter Modal analysis serves as an introduction into the vibrational analysis and dynamic measurements and additionally contains a new approach for approximation of a mass (and stiffness) matrix for measurement data.

Whereas the chapter Statistical modeling is an introduction into the statistical analysis, the theory of linear regression models and statistical testing. The presented concepts and tools of both these chapters are used in the next two chapters.

In the chapter Correlation criteria for two data sets comparison, we present the known basic criteria, especially so called MAC and FDAC, as well as their new extensions. All of these criteria are suitable for comparison of two data and they are based on elementary correlation criteria and modal properties. The first main generalization is through the inclusion of the mass matrix obtained from measurement data into the criteria and the second new development are selective criteria which are intended to capture and determine the differences in the data and give a more detailed view of data. All criteria are applied on real data and the results are demonstrated.

The chapter Statistical processing of multiple data sets is the major chapter devoted to new generalizations of the basic criteria through application of the statistical theory of linear regression models. These criteria have the same purpose as the standard criteria, to compare the data involved and detect the differences.

The only, essential, difference is in comparing many data sets together. In conclusion of the chapter, there is again a demonstration of the criteria on real data.

The last large chapter MAC application is a user documentation for the programmed tool containing information about the installation, description of the application options and required format of input data and few more useful notes about the program itself.

Chapter 2

Modal analysis

Modal analysis is a way of studying vibrating systems by finding system resonant (natural) frequencies and modes of vibration. The concept is that, under certain conditions on the system, any structural vibration can be described as a weighted superposition of each natural mode [4], [5].

Let us focus on linear (linearised) time-invariant N degree of freedom systems that can be modelled by a set of coupled second order differential equations of motion as

$$M\ddot{x} + D\dot{x} + Kx = f, \quad (2.1)$$

where M , D , K are mass, damping and stiffness matrices, $x = x(t)$ is a time-varying vector of generalized coordinates (e.g. x , \dot{x} , \ddot{x} are displacement, velocity and acceleration vectors) and $f = f(t)$ is a vector of external forces. In classical structural models, the proportional viscous damping (Rayleigh damping) is assumed, i.e. the damping matrix can be written as a linear combination

$$D = c_1M + c_2K,$$

where $c_1 \in \mathbb{R}$, $c_2 \in \mathbb{R}$ and both matrices M and K are considered real symmetric, matrix M positive definite and matrix K positive semidefinite.

For the undamped, unforced system, the equations of motion (2.1) become

$$M\ddot{x} + Kx = 0. \quad (2.2)$$

Assuming the general solution of the form

$$x = ve^{i\omega t},$$

v designate a vector of amplitudes and ω an angular frequency of the motion, and introducing it into the equations (2.2), a generalized matrix eigenvalue problem

$$(K - \lambda M)v = 0, \quad (2.3)$$

where $\lambda = \omega^2$, is obtained.

We assume that a solution to the eigenvalue problem (2.3) are N distinct eigenvalues and corresponding N eigenvectors. Eigenvalues are numbers $\lambda_n = \omega_n^2$, $n = 1, 2, \dots, N$, for which the eigenvalue problem has a nontrivial solution v_n , the resonant frequencies are given by $\omega_n = \sqrt{\lambda_n}$, and the nontrivial solutions v_n are called eigenvectors or modal vectors or mode shapes. We will use the term modal vectors from now on.

Due to the M symmetric positive definite, K symmetric positive semidefinite and the assumption on distinct eigenvalues, we get N distinct nonnegative eigenvalues and corresponding N linearly independent real eigenvectors that comprise an orthogonal basis of mass and stiffness matrices [6], [7]. By arranging modal vectors as columns into a matrix V and eigenvalues on diagonal of a diagonal matrix Λ , a modal matrix V and a spectral matrix Λ arises.

Modal vectors comprise an orthogonal basis of mass and stiffness matrices, i.e. both V^*MV and V^*KV are diagonal matrices, and because of proportional damping, a matrix V^*DV is a diagonal matrix as well (matrices M , D , K are diagonalizable in the same basis). The modal expansion theorem holds, the vector of generalizes coordinates x can be expressed in modal (natural) coordinates $q = q(t)$ as

$$x = Vq.$$

By expressing the equations of motion (2.1) in modal coordinates and multiplying it by matrix V^* from the left, it is possible to have the system described by a set of N uncoupled equations of motion. Modal vectors may be M -orthonormalized (eventually also K -orthonormalized) and then for modal vectors holds

$$\begin{aligned} V^*MV &= I = VMV^*, \\ V^*KV &= \Lambda = VKV^*. \end{aligned}$$

When the system under study does not comply with the classical assumptions, the resonant frequencies of the system may not be identical to those of associated undamped system and both eigenvalues and modal vectors may be complex valued.

2.1 Experimental modal analysis

Experimental modal analysis is the study of modal properties of vibrating systems in a frequency domain, using an experimental frequency response function (a transfer function) information.

The transfer function $H(z)$ is, as a complex function of a complex variable z , an input-output representation of the system differential equation. For the linear time-invariant system with continuous time input $x(t)$ and output $y(t)$, the transfer function is defined as a linear mapping of the Laplace transform of the input $X(z) = \mathcal{L}(x(t))$ to the Laplace transform of the output $Y(z) = \mathcal{L}(y(t))$

with the following form,

$$H(z) = \frac{Y(z)}{X(z)}.$$

The roots of the equation $Y(z) = 0$ are the system zeros and the roots of the equation $X(z) = 0$ are the system poles or the system eigenvalues.

The frequency response function is the transfer function evaluated for $z = i\omega$, where i is an imaginary unit. Considering a set of N differential equations of motion with harmonic excitation as

$$M\ddot{x} + D\dot{x} + Kx = Fe^{i\omega t},$$

where F is a vector of amplitudes, then, assuming the solution of the form $x = Ae^{i\omega t}$ and introducing it into the equations,

$$[(i\omega)^2M + i\omega D + K]A = F,$$

the frequency response function can be expressed as

$$H(i\omega) = [(i\omega)^2M + i\omega D + K]^{-1},$$

where $H(i\omega)$ is a $N \times N$ matrix, whose element at position r, s (the r -th row, the s -th column) designate the response in the r -th position due to the unit excitation in the s -th position. Applying a harmonic force at resonant frequency results in a deflection shape that is proportional to the modal vector corresponding to this frequency.

In a dynamic measurements of the experimental frequency response function of mechanical or acoustical systems, usually the input is force (or acceleration) and the output is acceleration (or displacement or velocity). The structure is excited by external forces using an impact hammer (a hammer impact testing), an electrodynamic shaker or a measurement on a test bench. The output is measured in selected sensor positions per individual channels with translational degrees of freedom denoted as DOF 1, 2, 3 (or as ‘u’, ‘v’, ‘w’), eventually rotational degrees of freedom DOF 4, 5, 6 (or ‘phiu’, ‘phiv’, ‘phiw’).

Typically amplitudes and phases as a function of frequency for individual channels (sensor positions) are plotted as a graph of the frequency response function. Amplitudes have peaks (maxima) corresponding to resonant frequencies. If the frequency response function is expressed in real and imaginary parts, then the real parts equal zeros and imaginary parts have peaks (maxima or minima, either above or below zero) at resonant frequencies and the peak values of the imaginary parts may be taken as the modal vectors, for acceleration or displacement responses [8]. The selection of peaks itself is not entirely unambiguous due to the noise in the data (and all the measurement errors) and a requirement for global peak selection, it is necessary to select such frequency values (and corresponding modal vectors) that occur as peak at multiple channels.

For the modal vectors approximated by the deflection shape (experimental frequency response function) corresponding to resonant frequencies, the designation v (and V for the modal matrix) is also used in the following text. Therefore

also the resulting modal matrix do not have to be square, the length of modal vectors gained from the measurements depends on the number of channels and sensor positions.

2.2 Experimental mass and stiffness matrices - Pseudoinverse method

In case of Rayleigh damping, two dominant experimental damping coefficients may be measured and when mass and stiffness matrices are known, it is possible to determine the coefficients c_1 , c_2 and hence the damping matrix [5].

Mass and stiffness matrices are usually obtained from finite element models. However, when experimental modal analysis is used, the mass and stiffness matrices remain unknown. For the case we do not have the finite element model of the structure, we introduce a new approach, named pseudoinverse method, to approximate the mass and stiffness matrices. The matrices thus obtained may be also compared to matrices from the finite element simulation or used for additional eigenvalues and eigenvectors computation (and then serve as an accuracy indicator of the pseudoinverse method).

The approach of pseudoinverse method has been practically implemented only to obtain a mass matrix and theoretically can be used to the stiffness matrix as well.

Based on the knowledge of the (measured) modal and spectral matrices, the mass and stiffness matrices can be derived from equations

$$\begin{aligned} V^*MV &= I, \\ V^*KV &= \Lambda, \end{aligned}$$

so the normalization of modal vectors with respect to the mass matrix M is presumed (some methods to determine a scaling factor for modal vector normalization are shown in [9]). Since the modal matrix V may not be square, the Moore-Penrose pseudoinverse V^+ is used to get the matrices M and K as least squares solutions (minimum norm solutions) or as projections into a subspace of smaller size matrices. If the modal matrix V is of type $L \times N$, where $N < L$, (which is not at all a common case for measurement data) some rows of the modal matrix are deleted to ensure the system is not overdetermined. Let R denote the number of degrees of freedom (channels) per sensor, *AUTOMAC* an AUTOMAC matrix (see below, in the section 4.1.1) and *AUTOMAC_R* the AUTOMAC matrix for modal matrices with R rows corresponding to some sensor removed. While $N < L$, iteratively R rows per each sensor of both (the same duplicated) modal matrices are deleted so that the L^2 norm

$$\|AUTOMAC - AUTOMAC_R\|$$

is minimal. However, the square dimension of the modal matrix is the ideal shape for calculating M and K matrices.

Several numerical methods with regularisations have been extended and combined to obtain the mass matrix and named as

- Tikhonov and Schulz,
- IC symmetric and Schulz,
- Tikhonov and PR2-Schulz,
- IC symmetric and PR2-Schulz.

In these algorithms, Tikhonov regularization or an initial condition designated as IC symmetric serves to get the first approximation of matrices M and K (resp. matrix $(VV^*)^{-1}$) and Schulz or PR2-Schulz methods applied to improve their accuracy. The description of the individual (modified) procedures follows, methods together with implemented algorithms are presented, and a comment on a comparison of methods in terms of condition number is added at the end of this section.

2.2.1 Tikhonov regularization

Tikhonov regularization is a classical regularization method [1], [10], that improves conditioning of a problem (and allows numerical finding an approximate pseudoinversion) by seeking the solution, matrices M and K , in the least squares sense (as a projection) with a constraint eliminating solutions of large norm. We consider obtaining the mass and stiffness matrices separately, different regularization parameters and thus approximation of the pseudoinversion V^+ (resp. matrix $(VV^*)^{-1}$) are used.

For the mass matrix holds

$$M = V^{*+}V^+ = (VV^*)^{-1}V V^*(VV^*)^{-1} = (VV^*)^{-1}$$

(matrix V has full column rank thus the matrix (VV^*) is regular).

Applying a simple Tikhonov regularization with one regularization parameter $\beta_1 \in (0, 1)$ to the problem leads to two variants of mass matrix computation,

$$M = (VV^* + \beta_1 E)^{-1}V V^*(VV^* + \beta_1 E)^{-1}$$

or

$$M = (VV^* + \beta_1 E)^{-1}. \tag{2.4}$$

The second variant (2.4) is used, because less matrix multiplication implies fewer numerical errors and thus better results (experimentally verified).

The stiffness matrix can be calculated as

$$K = (VV^* + \beta_2 E)^{-1}V \Lambda V^*(VV^* + \beta_2 E)^{-1},$$

with the regularisation parameter $\beta_2 \in (0, 1)$.

The regularization parameters can be adjusted based on different tools, known are e.g. Morozov's discrepancy principle, L-curve method or a generalized cross-validation. However, parameters β_1 and β_2 may be also chosen for fixed V as minima

$$\begin{aligned}\beta_1 &= \min_{\beta} \varphi_1(V, \beta), & \varphi_1(V, \beta) &:= \text{cond}((VV^* + \beta)^{-1}), \\ \beta_2 &= \min_{\beta} \varphi_2(V, \beta), & \varphi_2(V, \beta) &:= \text{cond}((VV^* + \beta)^{-1}V),\end{aligned}$$

where $\text{cond}(\cdot)$ designate a condition number and $\beta \in (0, 1)$. It can be shown that both functions $-\varphi_1(V, \beta)$ and $-\varphi_2(V, \beta)$, $\beta \in (0, 1)$, are weakly unimodal (there exists a value c such that functions are weakly monotonically increasing for $0 < \beta \leq c$ and weakly monotonically decreasing for $c \leq \beta < 1$) and functions $\varphi_1(V, \beta)$, $\varphi_2(V, \beta)$, $\beta \in (0, 1)$, have the minimum, for fixed V . Which means, the minimum can be found e.g. by a golden section search.

We chose to select the parameter β_1 for fixed V numerically close to minimum by a golden section search with a termination condition

$$|a - b| \leq \varepsilon \quad \vee \quad \|V^* M_n V - I\| \geq 1.1 \cdot \|V^* M_{\beta=0} V - I\|,$$

where a and b are boundaries of the search interval, ε is a tolerance parameter, M_n is the current approximation of the mass matrix and $M_{\beta=0}$ is a mass matrix obtained without any regularization. That is, the parameter β_1 is chosen so that the conditional number of the resulting mass matrix M is the lowest possible while the L^2 norm error is less than 1.1 multiple of L^2 norm error for the mass matrix obtained without any regularization.

The value of β_2 may be found by a golden section search with analogous termination condition

$$|a - b| \leq \varepsilon \quad \vee \quad \|V^* K_n V - \Lambda\| \geq 1.1 \cdot \|V^* K_{\beta=0} V - \Lambda\|,$$

where a and b are boundaries of the search interval, ε is a tolerance parameter, K_n is the current approximation the stiffness matrix and $K_{\beta=0}$ is the stiffness matrix obtained without any regularization. That is, the parameter β_2 can be chosen so that the conditional number of the approximate pseudoinversion V_2^+ is the lowest possible while similar condition (as in case of mass matrix) for norms for the stiffness matrix holds.

Matrix $(VV^* + \beta_1 E)^{-1}$ (ev. $(VV^* + \beta_2 E)^{-1}$) is then the first approximation A_0^{-1} for Schulz and PR2-Schulz methods.

In addition to the mentioned simple Tikhonov regularization, other types of regularization, such as for example

- generalized (e.g. [11]) or higher order of Tikhonov regularization,
- TSVD – Truncated singular value decomposition, where singular values of inverted matrix smaller than a tolerance ε are truncated to zero (the disadvantage is that, unlike a smooth behaviour of Tikhonov filtering function, a TSVD filtering function has discontinuous cut off behaviour, which may result in oscillations),

- iterative regularization (steepest descent, conjugate gradients, Landweber-Fridman), where the regularization parameter equals to the number of iterations,

could be applied as well [10].

2.2.2 IC symmetric

IC symmetric is a denomination for an initial approximation of an inversion of symmetric positive definite matrix A of type $N \times N$,

$$A_0^{-1} = \frac{I}{\|A\|},$$

that, analytically, satisfies a necessary and sufficient condition of convergence for the Schulz and PR2-Schulz method,

$$\|I - AA_0^{-1}\|_S < 1,$$

with a known bound

$$\|I - AA_0^{-1}\|_S < 1 - \frac{1}{\sqrt{N} \text{cond}(A)},$$

where $\|\cdot\|_S$ designate a spectral norm [12], [13]. When searching for the mass matrix M , or stiffness matrix K , the IC symmetric condition is applied to matrix

$$A = VV^*.$$

Numerically, approximations of the matrix VV^* are forced to be real symmetric (trimming imaginary parts) with the assumption of positive definite (checked at the end of all the algorithms) in a Python implementation.

Besides the above mentioned initial condition, another initial approximation might be considered as well. A classic example is

$$\tilde{A}_0^{-1} = \frac{A^*}{\|A\|_1 \|A\|_\infty},$$

with a bound

$$\|I - A\tilde{A}_0^{-1}\|_S < 1 - \frac{1}{N(\text{cond}(A))^2},$$

for not only symmetric positive definite matrix A [13] (with trimming of imaginary parts and checking of positive definiteness of resulting mass and stiffness matrices only). Nevertheless, because of more matrix multiplication with such an initial approximation the sufficient condition of convergence, and convergence itself, is numerically less feasible (in addition, as expected, in numerical experiments applied to real data the trimming of imaginary parts led to better results when used on the matrix VV^* instead of on resulting mass matrix).

2.2.3 Schulz method

Schulz method is a Newton's method generalized for matrix (pseudo-) inversion problems, exhibits local quadratic convergence, but require a suitable initial approximation [12]. For a given initial matrix A_0^{-1} the iterations are defined by

$$A_{n+1}^{-1} = 2A_n^{-1} - A_n^{-1}AA_n^{-1}.$$

New approximations of A^{-1} are calculated while the error norm

$$\|I - V^*A^{-1}V\|$$

($\|I - V^*A^{-1}V\|$ in case of stiffness matrix) decreases and the maximum limit of iterations is not reached (if the decreasing stops, a few more iterations are additionally calculated in case of further error norm decrease).

Schulz method of the third order (Chebyshev-Sen-Prabhu) [14], where

$$A_{n+1}^{-1} = 3A_n^{-1} - 3A_n^{-1}AA_n^{-1} + A_n^{-1}AA_n^{-1}AA_n^{-1},$$

is also implemented. It is a cubically convergent method, less iterations, but more computations per iterations are needed.

To speed up the algorithms runs, secant methods, where iterations do not require a large amount of computational work, or just another fast iterative algorithm for matrix inversion approximation [13] could be used.

2.2.4 PR2-Schulz method

PR2-Schulz method is a combination of the Schulz method and an extension of Richardson's method using the matrices A_n^{-1} as preconditioners that results in an effect of accelerating the convergence of the iterative method. Nevertheless, some numerical instability can occur at the last iterations for very ill-conditioned problems [12]. PR2-Schulz method generalized to obtain the mass matrix led to an algorithm

$$\begin{aligned} r_0 &= I - V^*A^{-1}V, \\ d &= A^{-1}VrV^*A^{-1}, \\ \text{for } n &= 1, 2, \dots, N_{\max} \text{ do :} \\ u &= V^*dV, \\ \delta_n &= \frac{\langle r_n^*, u \rangle}{\|u\|^2}, \\ M_{n+1} &= M_n + \delta_n d, \\ r_{n+1} &= I - V^*M_{n+1}V, \\ A^{-1} &= 2A^{-1} - A^{-1}AA^{-1}, \\ d &= A^{-1}VrV^*A^{-1}, \end{aligned}$$

where at the beginning the matrix A^{-1} is a given initial approximation, $\langle r_n^*, u \rangle$ is a L^2 (Frobenius) inner product of r_n^* and u and r_n is an error matrix which is recomputed in every loop to avoid loss of precision (otherwise $r_{n+1} = r_n - \delta_n u$ could be used). The principle of the method is that the new approximation of mass matrix M_{n+1} is calculated as the addition of the old approximation M_n and the correction term $\delta_n d$, where the coefficient δ_n is chosen so that the new error norm squared $\|r_{n+1}\|^2$ is minimal — the formula is obtained from a necessary condition of extreme, using a Fermat's theorem.

In the Python implementation, the new iterations are calculated while the error norm decreases, while $r_{n+1} < r_n$, and the maximum limit of iterations N_{\max} is not reached (if the decreasing stops, a few more iterations are additionally calculated in case of the further error norm decrease).

2.2.5 Comparison of methods in terms of condition number

For ill-conditioned problems a small error in the initial input data can lead to much larger errors in the output results, regardless of the algorithm used for returning the output result. An ill-conditioned problem is indicated by a large condition number. For rectangular matrix with full column rank, a condition number of a matrix A can be defined as

$$\text{cond}(A) = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)},$$

where $\sigma_{\max}(A)$ is the largest and $\sigma_{\min}(A)$ the smallest singular value of the matrix A . The conditional number of the singular matrix is infinite. The inverted value of the conditional number indicates the distance of the given matrix from the set of singular matrices measured in the spectral norm [6].

For the idea, what does a ‘large’ condition number mean, the inverted value of float digits of precision in Python (`1/sys.float_info.epsilon`) gives a value of approximately $4.5 \cdot 10^{15}$.

In the implemented methods, the condition number of the input matrix V and of the resulting mass matrix M is computed (and reported) and at the beginning of Schulz and PR2-Schulz methods, the norm $\|I - AA_0^{-1}\|_S$ is computed as well. In the case of a higher condition number, the PR2-Schulz method should be avoided and the conditioning may be improved using a Tikhonov regularization. Therefore, the recommended methods are primarily ‘IC symmetric and Schulz’ and ‘Tikhonov and Schulz’.

Chapter 3

Statistical modeling

Statistical modeling stands for a set of statistical models that embodies also statistical assumptions, estimates and associated statistical tests derived from the models.

In this chapter, we introduce a model of linear regression and two ANOVA (Analysis of Variance) models primarily following on from [15], [16]. ANOVA models can also generally be understood as regression models, we present a model of two-way ANOVA with interactions and a model of three-way ANOVA with interactions. Finally we mention basic information about statistical testing and provide statistical tests that we use. We apply this theory in other chapters (partly in the chapter 4 and primarily in the chapter 5).

3.1 Statistical models

Introduction of individual models, the model of linear regression, the model of two-way ANOVA with interactions and the model of three-way ANOVA with interactions, follows.

3.1.1 Linear model of regression

Although a regression model with complex numbers exists as well (see [17], [18]), we assume all the following real.

Let us have a random vector $y = (y_1, \dots, y_R)^*$ and a regression matrix of given numbers $X = (x_{rs})$ of type $R \times S$, $S < R$ with full column rank. Suppose that the vector y is given by

$$y = X\beta + \varepsilon,$$

where $\beta = (\beta_1, \dots, \beta_S)^*$ is a vector of unknown regression parameters and $\varepsilon = (\varepsilon_1, \dots, \varepsilon_R)^*$ a vector of unknown random deviations (errors) satisfying

1. variables ε_r , $r = 1, \dots, R$, are independent,
2. $E\varepsilon = 0$, where 0 is a vector of zeros,
3. (homoscedasticity) $\text{var } \varepsilon = \sigma^2 I$, where $\sigma^2 > 0$ is an unknown parameter,
4. $\varepsilon \sim N(0, \sigma^2 I)$.

We call this model a regression model, see [15], [16].

With the regression model we try to explain the larger number of random variables y_1, \dots, y_R by known influences X and a relatively small number of parameters β_1, \dots, β_S .

In the regression model, the columns of matrix X are called regressors. The regression model tries to relate y to a function of $X\beta$, therefore y also stands for an explained (dependent) variable and the columns of X for an explanatory (independent) variables (y_r is the r -th observation of the s -th explained variable, x_{rs} is the r -th observation of the s -th explanatory variable). The regression model is called linear since the β parameters are represented in the model in linear form.

If any explanatory variable takes the value 1 for all r , if for some s and all r is $x_{rs} = 1$, then β_s is called the regression intercept. A regression model without the intercept term has no regression intercept and the regression mapping $\tilde{y} = Xb$, where b is an estimate of β (see below), is forced to pass through the origin in this case.

Parameter estimates

Parameters β_1, \dots, β_S are usually estimated by the least squares method. The estimates denoted as b_1, \dots, b_S , $b := (b_1, \dots, b_S)^*$ are given by

$$b = (X^*X)^{-1}X^*y$$

(matrix X has full column rank, thus the matrix X^*X is regular).

The parameter σ^2 is estimated by statistic s^2 , so called residual variance,

$$s^2 = \frac{S_e}{R - S},$$

where S_e is a residual sum of squares calculated as

$$S_e = y^*y - b^*X^*y. \tag{3.1}$$

When we define estimated values as

$$\tilde{y} = Xb,$$

that are the best approximations for y , and residuals

$$e_r = y_r - \tilde{y}_r, \quad r = 1, \dots, R,$$

the residual sum of squares can be defined as

$$S_e = \sum_{r=1}^R e_r^2$$

[15], [16].

From the assumption of normally distributed errors $\varepsilon \sim N(0, \sigma^2 I)$ it follows that variable y has a normal distribution $y \sim N(X\beta, \sigma^2 I)$ and because b is derived from y by linear transformation, it has a normal distribution as well and $b \sim N(\beta, \sigma^2 (X^* X)^{-1})$ [15, p. 100].

It is also known, that the vector b and the variable s^2 are independent and that the variable S_e/σ^2 has a χ^2 probability distribution with a parameter $R - S$, $S_e/\sigma^2 \sim \chi^2(R - S)$, [15, p. 100]. That is, $(R - S)s^2/\sigma^2 \sim \chi^2(R - S)$.

Coefficients of determination

The coefficient of determination is a statistical measure of how well the regression curve $\tilde{y} = Xb$ fits the observed data y .

Let define a total sum of squares as

$$S_T = \sum_{r=1}^R (y_r - \bar{y})^2,$$

where \bar{y} is an average value of y_r , $r = 1, \dots, R$,

$$\bar{y} = \frac{1}{R} \sum_{r=1}^R y_r.$$

For $S_T \neq 0$ and $R > 1$ we define a coefficient of determination

$$R^2 = 1 - \frac{S_e}{S_T} \tag{3.2}$$

and an adjusted coefficient of determination

$$R_a^2 = 1 - \frac{S_e/(R - S)}{S_T/(R - 1)}.$$

In regression with the intercept term, the definition (3.2) is equivalent to

$$R^2 = \frac{\sum_{r=1}^R (\tilde{y}_r - \bar{y})^2}{S_T}, \tag{3.3}$$

it ranges between zero and one and it is interpreted as a proportion of variance of the values y_r , $r = 1, \dots, R$, that was explained by the regression model ($1 - R^2$ is a proportion of variance of the values y_r , $r = 1, \dots, R$, that was not explained

by the regression model). However, for the model without the intercept term, the two relations (3.2) and (3.3) are not equivalent and R^2 is not bounded by the values of zero and one. Value of R^2 calculated from (3.2) lies in the interval $(-\infty, 1]$ and R^2 calculated from (3.3) lies in the interval $[0, +\infty)$ [16, p. 99].

We will also define a modified coefficient of determination in the form

$$R_0^2 = 1 - \frac{S_e}{S_{TT}},$$

with

$$S_{TT} = \sum_{r=1}^R (y_r)^2$$

(the reason will be obvious later, because it is connected with a special regression model introduced in the chapter 5).

3.1.2 Model of two-way ANOVA with interactions

In the following we again focus only on real numbers. We will be using the bold letters for ANOVA parameters to distinguish ANOVA and regression parameters later in the text.

Let us have a random variables \mathbf{y}_{sr}^n (take attention on top indexes here), $\mathbf{s} = 1, \dots, \mathbf{S}$, $\mathbf{n} = 1, \dots, \mathbf{N}$, $\mathbf{r} = 1, \dots, \mathbf{R}$ (we restrict to the same number of observations in subclasses). Suppose that the variables \mathbf{y}_{sr}^n are given by

$$\mathbf{y}_{sr}^n = \boldsymbol{\mu} + \boldsymbol{\alpha}_s + \boldsymbol{\beta}^n + \boldsymbol{\lambda}_s^n + \boldsymbol{\varepsilon}_{sr}^n,$$

where $\boldsymbol{\mu}$, $\boldsymbol{\alpha}_s$, $\boldsymbol{\beta}^n$, $\boldsymbol{\lambda}_s^n$ are unknown parameters and $\boldsymbol{\varepsilon}_{sr}^n$ are unknown random deviations (errors) that are assumed to be normally distributed with zero mean and constant variance $\boldsymbol{\sigma}^2 > 0$, $\boldsymbol{\sigma}^2$ is an unknown parameter, $\boldsymbol{\varepsilon}_{sr}^n \sim N(0, \boldsymbol{\sigma}^2)$, and $\boldsymbol{\varepsilon}_{sr}^n$ being mutually independent, for all $\mathbf{s} = 1, \dots, \mathbf{S}$, $\mathbf{n} = 1, \dots, \mathbf{N}$, $\mathbf{r} = 1, \dots, \mathbf{R}$. We assume reparameterization conditions

$$\begin{aligned} \sum_{s=1}^{\mathbf{S}} \boldsymbol{\alpha}_s &= 0, \\ \sum_{n=1}^{\mathbf{N}} \boldsymbol{\beta}^n &= 0, \\ \sum_{s=1}^{\mathbf{S}} \boldsymbol{\lambda}_s^n &= 0, \quad \forall n = 1, \dots, \mathbf{N}, \\ \sum_{n=1}^{\mathbf{N}} \boldsymbol{\lambda}_s^n &= 0, \quad \forall s = 1, \dots, \mathbf{S}, \end{aligned}$$

$\mathbf{SN} + 1$ linearly independent parameters $\boldsymbol{\mu}$, $\boldsymbol{\alpha}_s$, $\boldsymbol{\beta}^n$, $\boldsymbol{\lambda}_s^n$ and $\boldsymbol{\sigma}^2$, $\mathbf{R} \geq 2$ and for the number of observations \mathbf{SNR} , that $\mathbf{SNR} \geq \mathbf{SN} + 1$. We call this model a model of two-way ANOVA with interactions [15], [16].

The parameters α_s, β^n are called the main effects and λ_s^n the interactions. Because of two main effects in the ANOVA model, it is a two-way ANOVA model.

For all parameters $\mu, \alpha_s, \beta^n, \lambda_s^n$ and σ^2 we can get the estimates by the least square method as well [15, p. 165].

3.1.3 Model of three-way ANOVA with interactions

We will be again using the bold letters for ANOVA parameters and consider them real. Unlike the two-way ANOVA model we introduce three main effects and more interactions between them within the model.

Let us have a random variables y_{slr}^n (with top indexes here), $s = 1, \dots, S, l = 1, \dots, L, n = 1, \dots, N, r = 1, \dots, R$. Suppose that the variables y_{slr}^n are given by

$$y_{slr}^n = \mu + \alpha_s + \beta^n + \gamma_l + \lambda^{AB^n}_s + \lambda^{BC^n}_l + \lambda^{AC}_{sl} + \lambda^{ABC^n}_{sl} + \varepsilon_{slr}^n,$$

where $\mu, \alpha_s, \beta^n, \gamma_l, \lambda^{AB^n}_s, \lambda^{BC^n}_l, \lambda^{AC}_{sl}, \lambda^{ABC^n}_{sl}$ are unknown parameters and ε_{slr}^n are unknown random deviations (errors) that are assumed to be normally distributed with zero mean and constant variance $\sigma^2 > 0$, σ^2 is an unknown parameter, $\varepsilon_{slr}^n \sim N(0, \sigma^2)$, and ε_{slr}^n being mutually independent, for all $s = 1, \dots, S, l = 1, \dots, L, n = 1, \dots, N, r = 1, \dots, R$. We assume reparameterization conditions, in a simplified way written as

$$\begin{aligned} \sum_{s=1}^S \alpha_s &= 0, \\ \sum_{n=1}^N \beta^n &= 0, \\ \sum_{l=1}^L \gamma_l &= 0, \\ \sum_{s=1}^S \sum_{n=1}^N \lambda^{AB^n}_s &= 0, \\ \sum_{l=1}^L \sum_{n=1}^N \lambda^{BC^n}_l &= 0, \\ \sum_{s=1}^S \sum_{l=1}^L \lambda^{AC}_{sl} &= 0, \\ \sum_{s=1}^S \sum_{l=1}^L \sum_{n=1}^N \lambda^{ABC^n}_{sl} &= 0, \end{aligned}$$

$SNL + 1$ linearly independent parameters $\mu, \alpha_s, \beta^n, \gamma_l, \lambda^{AB^n}_s, \lambda^{BC^n}_l, \lambda^{AC}_{sl}, \lambda^{ABC^n}_{sl}$ and σ^2 , $R \geq 2$ and for the number of observations $SLNR$, that

$SLNR \geq SLN + 1$. We call this model a model of three-way ANOVA with interactions [15].

The main effects are $\alpha_s, \beta^n, \gamma_l$, the parameters $\lambda^{AB^n}_s, \lambda^{BC^n}_l, \lambda^{AC}_{sl}$ are called interactions of the first order and $\lambda^{ABC^n}_{sl}$ interactions of the second order.

The parameters can be estimated by the least square method as well [15, p. 169].

When the parameters $\gamma_l, \lambda^{BC^n}_l, \lambda^{AC}_{sl}$ and $\lambda^{ABC^n}_{sl}$ are all zero valued, we obtain the two-way ANOVA model with $\lambda_{sl} = \lambda^{AC}_{sl}$.

3.2 Statistical testing

Statistical testing provides a mechanism for making a decision about a chosen hypothesis. The main intent is to determine whether there is enough evidence to reject the hypothesis or not.

The statistical test requires a pair of hypotheses,

- a null hypothesis H_0 , which is a statement being tested,
- an alternative hypothesis H_1 , which is the negation of the null hypothesis.

The null hypothesis is generally assumed to be true (until evidence is indicative of the opposite) and it is usually our wish to reject the null hypothesis. The test procedure is constructed so that the probability of incorrectly rejecting the null hypothesis (rejecting the null hypothesis, when it is in fact true) is ‘small’.

The process of statistical testing is associated with two types of errors,

- a type I error, that denotes an incorrect rejection of a true null hypothesis,
- a type II error, denoting a failure to reject a false null hypothesis.

A significance level, usually denoted by letter α , is a maximum admissible probability of the type I error. The probability of the type II error is usually denoted by a letter β and is related to a power of a test. There is generally no relationship between the significance level and the power of a test, it can be only shown that, under certain conditions, the decrease of α leads to growth of β and vice versa.

In statistical testing, firstly a null and an alternative hypothesis are specified and then the significance level is set. Often a value of $\alpha = 5\%$ or $\alpha = 1\%$ is chosen (meaning that it is acceptable to have a 5% or 1% probability of incorrect rejection). Finally the testing of the null hypothesis against the alternative hypothesis is done according to specific random variable and critical value of a critical region. If the value of the random variable falls in the critical region, we reject the null hypothesis (for concrete examples see below).

An alternative approach of testing is using a p-value. The p-value is the smallest significance level α , on which the null hypothesis is rejected. A small p-value indicates a strong evidence against the null hypothesis. Thus, when the p-value

is less than the determined significance level α , the null hypothesis is rejected, otherwise the null hypothesis is not rejected. In special cases (as in our cases later), the p-value may be calculated as a complement of a cumulative distribution function of the random variable (when the null hypothesis is valid) evaluated in the critical value. We present the results of applied statistical tests primarily by the p-value.

3.2.1 Statistical tests for regression

Let us introduce two known theorems used for statistical testing within the regression models. We use the notations and estimates introduced in the previous sections here.

Theorem 3.1 ([15, p. 101]). *Let us denote*

$$\beta = \begin{bmatrix} \beta_1^\circ \\ \beta_2^\circ \end{bmatrix}, \quad b = \begin{bmatrix} b_1^\circ \\ b_2^\circ \end{bmatrix}, \quad (X^*X)^{-1} = \begin{bmatrix} T & U \\ U^* & W \end{bmatrix},$$

where vectors β_1° and b_1° have q_1 elements, vectors β_2° and b_2° have q_2 elements, $q_2 > 0$, $q_1 + q_2 = S$, T is of type $q_1 \times q_1$ and W of type $q_2 \times q_2$.

Then a random variable

$$Z = \frac{1}{q_2 s^2} (b_2^\circ - \beta_2^\circ)^* W^{-1} (b_2^\circ - \beta_2^\circ)$$

has the F distribution (Fisher probability distribution) with parameters q_2 and $R - S$, $Z \sim F(q_2, R - S)$.

We use the theorem 3.1 for testing the hypothesis

$$H_0 : \beta = 0,$$

against the alternative hypothesis

$$H_1 : \beta \neq 0,$$

that is, for testing whether the whole vector of regression parameters is zero valued. Therefore we set $\beta = \beta_2^\circ$, $b = b_2^\circ$, $(X^*X)^{-1} = W$, $q_2 = S$ and Z is calculated as

$$Z = \frac{1}{S s^2} b^* X^* X b.$$

If

$$Z > F_{1-\alpha}(S, R - S),$$

where $F_{1-\alpha}(S, R - S)$ is a $1 - \alpha$ quantile of the F distribution with parameters S and $R - S$, then hypothesis H_0 is rejected at the significance level α . Alternatively, we can compute a p-value, as

$$p = 1 - P(\mathbf{X} \leq Z),$$

where $P(\mathbf{X} \leq Z)$ is a cumulative distribution function of the F distribution evaluated at Z . If $p < \alpha$, the hypothesis H_0 is rejected at the significance level α .

In the case of $N := q_1 < S$ (if we want to test not all but fewer regression parameters), the theorem 3.1 is equivalent to the following theorem 3.2 with the same distributed variable Z [15].

Theorem 3.2 ([16, p. 100]). *Let M be a set of N different natural numbers, $M \subset \{1, \dots, S\}$, $N < S$, and define S_e the residual sum of squares of the regression model with all S regressors and S_{eM} the residual sum of squares of a regression model without all regressors with indexes $m \in M$.*

Then a random variable

$$Z = \frac{S_{eM} - S_e}{S_e} \frac{R - S}{N}$$

has the F distribution (Fisher probability distribution) with parameters N and $R - S$, $Z \sim F(N, R - S)$.

The theorem 3.2 is used for testing a null hypothesis

$$H_0 : \forall m \in M : \beta_m = 0,$$

i.e. that N selected regression parameters, $N < S$, are zero valued, against the alternative hypothesis

$$H_1 : \exists m \in M : \beta_m \neq 0.$$

The residual sum of squares S_{eM} , of a regression model without regressors with indexes $m \in M$, is a residual sum of squares calculated provided the hypothesis H_0 is valid. The null hypothesis H_0 is rejected at the significance level α , if

$$Z > F_{1-\alpha}(N, R - S),$$

where $F_{1-\alpha}(N, R - S)$ is a $1 - \alpha$ quantile of the F distribution with parameters N and $R - S$.

3.2.2 Statistical tests for ANOVA

Let us introduce the third theorem used for statistical testing, that we apply on ANOVA model. We use the notations and estimates introduced in the previous sections here as well.

Theorem 3.3 (Scheffe, [15, p. 147]). *Let $\varphi = (\varphi_1, \dots, \varphi_P)^*$ be a random vector, $\varphi \sim N(\psi, \sigma^2 W)$, where $W > 0$ is a known matrix and $\sigma^2 > 0$ an unknown parameter. Let A be a t dimensional subspace of \mathbb{R}^P and s^2 an independent estimate for σ^2 with ν degrees of freedom, i.e. $\nu s^2 / \sigma^2 \sim \chi_\nu^2$ with s^2 and φ independent.*

Then the probability that an inequality

$$|a^* \varphi - a^* \psi| \leq \sqrt{ts^2 F_\alpha(t, \nu)} a^* W a$$

holds true for all $a \in A$ simultaneously, is equal to $1 - \alpha$.

The Theorem 3.3 may be used to test the equality of multiple variables alongside. Let us have what is listed in assumptions of the Theorem 3.3 and let M be a set of Q natural numbers $M \subset \{1, \dots, P\}$, $Q \leq P$ and $W = (w_{pq})$ of type $P \times P$. With the Theorem 3.3 we can test a null hypothesis that Q selected parameters are identical, i.e.

$$H_0 : \forall p, q \in M : \psi_p = \psi_q,$$

against the alternative hypothesis

$$H_1 : \exists p, q \in M : \psi_p \neq \psi_q.$$

To verify whether the equality applies to all pairs (p, q) , $p \neq q$, $p, q \in M$, we introduce vectors a_{pq} , $p, q \in M$, of lengths of P that have the p -th element equal to one, the q -th element equal to -1 and other elements zero valued. Then $\psi_p = \psi_q$ for some $p, q \in M$ holds true if and only if $a_{pq}^* \psi = 0$. The set of these vectors a_{pq} generates a subspace A of dimension $t = Q - 1$ in \mathbb{R}^P , $a_{pq}^* \varphi = \varphi_p - \varphi_q$, $a_{pq}^* W a_{pq} = w_{pp} + w_{qq} - w_{pq} - w_{qp}$. According to Scheffe theorem, the inequality

$$\|\varphi_p - \varphi_q\| \leq \sqrt{(w_{pp} + w_{qq} - w_{pq} - w_{qp}) (Q - 1) s^2 F_\alpha(Q - 1, \nu)}$$

holds true for all pairs (p, q) , $p, q \in M$, simultaneously with a probability of at least $1 - \alpha$. If for some pair (p°, q°) the inverse inequality applies, i.e.

$$|\varphi_{p^\circ} - \varphi_{q^\circ}| > \sqrt{(w_{p^\circ p^\circ} + w_{q^\circ q^\circ} - w_{p^\circ q^\circ} - w_{q^\circ p^\circ}) (Q - 1) s^2 F_\alpha(Q - 1, \nu)},$$

the null hypothesis H_0 is rejected. In addition, we have directly designated the pair (or several such pairs) where the equality $\psi_{p^\circ} = \psi_{q^\circ}$ is rejected at the significance level α [15, p. 149].

In our case we take a vector of estimated ANOVA parameters as the vector ψ (the exact use of the Theorem 3.3 is then described in the section 5.2).

Chapter 4

Correlation criteria for two data sets comparison

This chapter is dedicated to an overview of classical methods (correlation criteria) and their extension that can be used for two data sets comparison. The chapter is divided into two parts, basic nonselective criteria and selective criteria developing them.

For most of the criteria, the data sets consist of a vector of resonant frequencies and a matrix of modal vectors together with channel specifications which designate modal vector components. The data can originate from simulation, the modal analysis using a finite element method, or from measurements, from the experimental modal analysis (see the section 2). In the criterion called FDAC, the input data sets consist of a vector of frequencies (not only resonant) and a matrix of deflection shapes again together with channel specifications which designate the deflection shapes components.

The following methods are used to compare two data, it can be two different measurements of the same system, two measurements or two simulations of slightly different systems or a comparison of measurement and simulation of the same system. We will shortly use “two variants” from now on. The result of the methods is a “degree of similarity” of the variants, an indication of the most similar positions and vice versa positions of differences, across frequencies and across channel specifications (modal vectors elements) or a global overall decision which input data describes a tougher, stiffer, system, relative to oscillations.

The description of individual methods follows.

4.1 Nonselective criteria

4.1.1 MAC – Modal Assurance Criterion

MAC is a standard criterion used to determine the degree of similarity of modal vectors. MAC criterion provides the correlation of each modal vector of the first system with each modal vector of the second system (i.e. the result is a matrix of these correlations) over all the channels (sensor positions) involved (see Fig. 4.1). Modal vectors used in the comparison may be obtained from simulation or measurements.

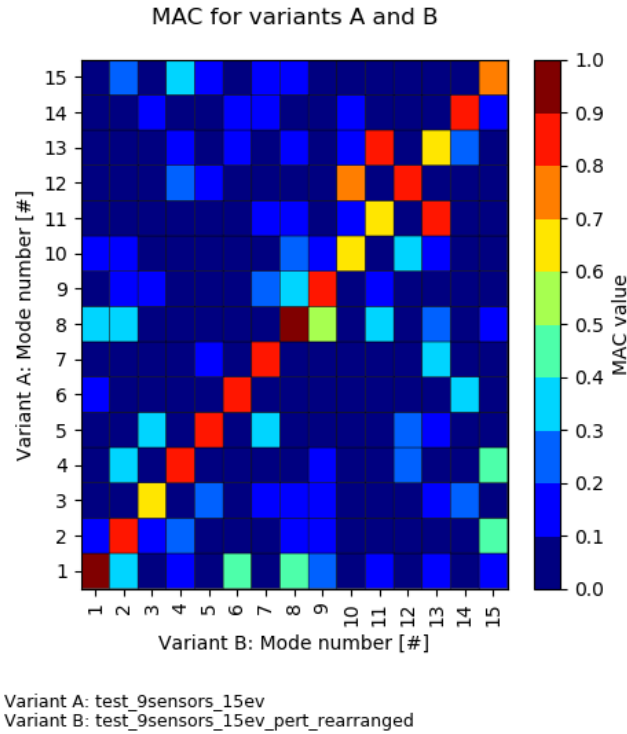


Figure 4.1: Example of MAC.

If we denote two modal vectors v_1 and v_2 , the MAC value is given by

$$MAC_{v_1, v_2} = \frac{|v_1^* v_2|^2}{v_1^* v_1 v_2^* v_2}.$$

MAC values range from zero to one. For two identical modal vectors (when the displacements in all sensor positions are the same) the MAC has the value of one and for completely different modal vectors is the MAC value close to zero (may not be zero because modal vectors are not necessarily L^2 -orthogonal).

Usually two vectors are considered correlated when the MAC value is greater than 0.9 and uncorrelated when the MAC value is less than 0.5.

MAC is insensitive to vector scaling, but sensitive to sensor location and their quantity and to the level of response at each sensor. In case that measured values

across individual sensors differ significantly, the normalization of data per sensors is for consideration [19]. The false correlation in the MAC matrix is also caused by not L^2 -orthogonality of modal vectors ([20] shows the impact of the diagonal non-identity mass matrix on the MAC nondiagonal values). Another issue is that for complex conjugate modal vectors, when the classical assumptions of structural dynamic are not complied with, the MAC criterion is not sufficient [19] (preferable methods are also introduced in [19]).

AUTOMAC

The MAC applied to one data set, one variant (if modal vectors of one system are compared only to themselves) is called AUTOMAC. The AUTOMAC is a standard indicator of how well the sensor values represent the system, each modal vector should be unambiguously distinguishable from others and the resulting matrix of AUTOMAC ideally diagonal.

MAC linearly scaled

A MAC linearly scaled is a newly displayed MAC. The MAC value for two modal vectors v_1 and v_2 can be interpreted as a square of cosine of an angle for these two vectors γ_{v_1, v_2} as well,

$$MAC_{v_1, v_2} = \cos^2(\gamma_{v_1, v_2})$$

(γ is a Hermitian angle for complex modal vectors [21]).

MAC linearly scaled is the MAC after a transformation

$$MAC \text{ linearly scaled}_{v_1, v_2} = 1 - \frac{\cos^{-1}(\sqrt{MAC_{v_1, v_2}})}{\frac{\pi}{2}}.$$

It shows linearly distributed value of MAC. (see Fig. 4.2).

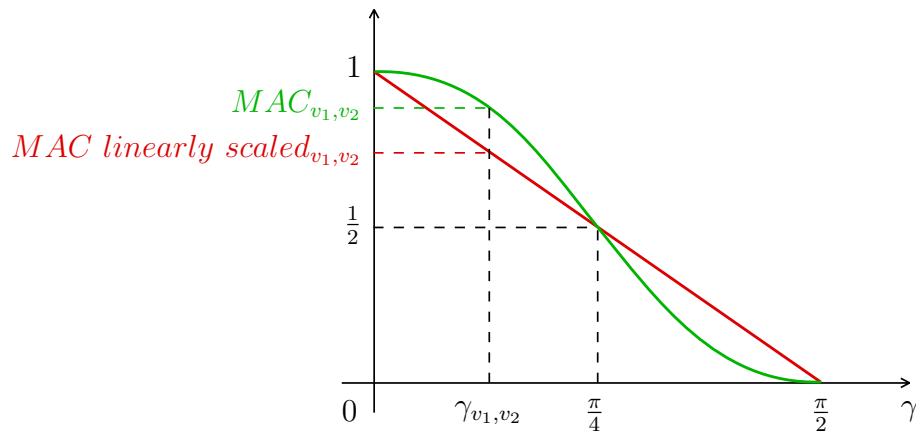


Figure 4.2: Dependence of MAC and MAC linearly scaled on modal vectors angle.

As we have already mentioned, two vectors are considered correlated when the standard MAC value is greater than 0.9, which with rounding corresponds to an angle less than 18 degrees, and uncorrelated when the MAC value is less than 0.5, which is equivalent to an angle greater than 45 degrees. Therefore the MAC value of 0.9 is equivalent to the rounded MAC linearly scaled value of 0.8 (which corresponds to the fact that 18 degrees is 20 percent of 90 degrees).

4.1.2 FDAC – Frequency Domain Assurance Criterion

FDAC is a MAC applied to frequency response functions - deflection shapes, it includes amplitude and phase information (it is complex valued). Given two deflection shapes h_1 and h_2 , the FDAC value is as follows,

$$FDAC_{h_1, h_2} = \frac{|h_1^* h_2|^2}{h_1^* h_1 h_2^* h_2}.$$

FDAC is, as well as MAC, insensitive to vector amplitudes scaling and to phase shift, but sensitive to sensor locations and their quantity and to the level of response at each sensor.

4.1.3 MWMAC – Mass Weighted Modal Assurance Criterion

MWMAC is a new tool, it is a MAC including mass matrix weighting of modal vectors (i. e. the modal properties in the criterion) wherein the mass matrix is not known at the input (when the mass matrix is known from simulations, a MOC criterion can be used, see below). Because modal vectors are M-orthogonal, MWMAC criterion improves the resulting matrices with smaller nondiagonal (or noncorresponding) values.

The mass matrix is obtained by the pseudoinverse method, by Tikhonov and Schulz, IC symmetric and Schulz, Tikhonov and PR2-Schulz or IC symmetric and PR2-Schulz algorithms (described in the section 2.2). That is, to obtain the mass matrices, the M-orthonormalization of input modal vectors is assumed.

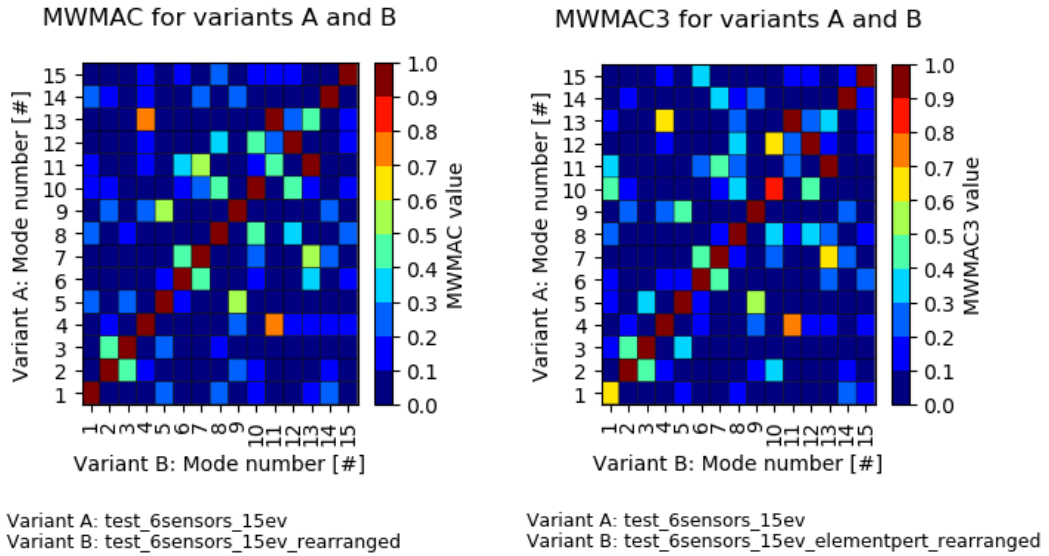
The MWMAC value for two modal vectors v_1 and v_2 is given by

$$MWMAC_{v_1, v_2} = \frac{|v_1^* \sqrt{M_1} \sqrt{M_2} v_2|^2}{v_1^* M_1 v_1 v_2^* M_2 v_2},$$

where $\sqrt{M_1}$ is a square root of a mass matrix gained for the first variant ($\sqrt{M_2}$ analogously), that is for a positive semidefinite matrix uniquely defined. A blocked Schur algorithm for computing the matrix square root was applied [22].

MWMAC is again insensitive to vector scaling (and to phase shift in case of complex modal vectors) and less sensitive to sensor selection and their quantity in comparison with MAC, but the accuracy of the calculated mass matrix M plays

an important role. Because of that the matrix M is generally just the projection (and because of inaccuracies), MWMAC values do not range strictly between zero and one. Due to the inclusion of mass matrix as a global information about the whole system, firstly, a local change at the input modal matrix generates a global change in the resulting MWMAC matrix (see Fig. 4.3), secondly, the MWMAC result is dependent on the number of sensors as well as on the number of modal vectors (see Fig. 4.4, 4.5).



(a)

(b)

Figure 4.3: (a) MWMAC, (b) MWMAC with the third element of the first modal vector perturbed.

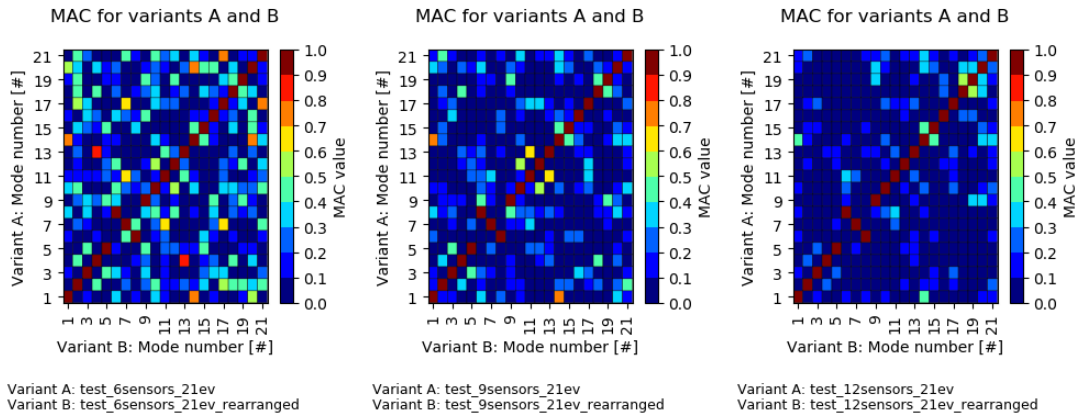
MOC – Mass Orthogonality Check

MOC is a known criterion that provides the similarity of two data sets through the check of M-orthogonality of all the modal vectors, when the mass matrix is known. Ideal case leads to pure diagonal (identity) matrix. In our case is this criterion primarily used as an indicator of how well the matrices M_1 , M_2 have been calculated. Therefore, these three types of MOC are calculated,

$$\begin{aligned}
 MOC_1 &= V_1^* M_1 V_1, \\
 MOC_2 &= V_2^* M_2 V_2, \\
 MOC_{12} &= V_1^* \sqrt{M_1} \sqrt{M_2} V_2,
 \end{aligned}$$

where V_1 is the modal matrix of the first variant, V_2 the modal matrix of the second variant.

The MOC_1 and MOC_2 have also similar meaning to the AUTOMAC, they indicate how well the sensor values represent the system with an embedded mass matrix obtained from these sensor values.



More sensors ➔

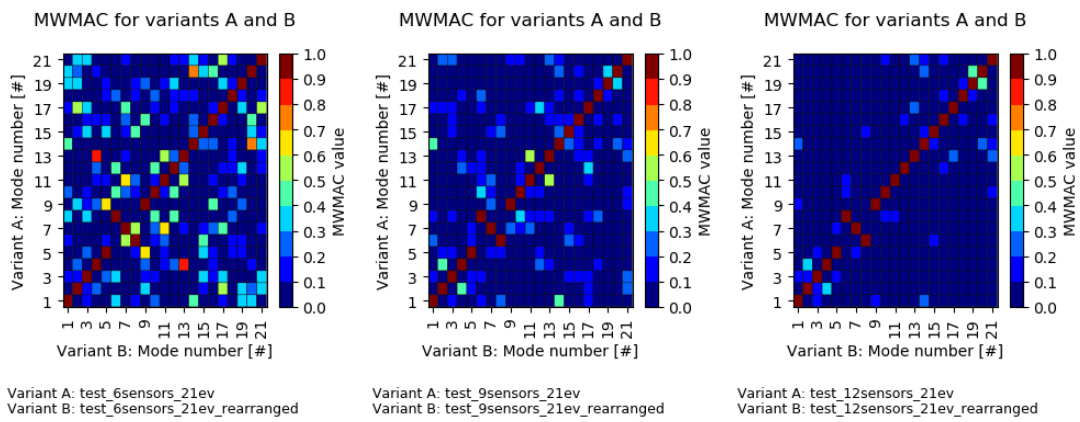


Figure 4.4: Dependence of MAC and MWMAC on number of sensors.

Less modal shapes ➔

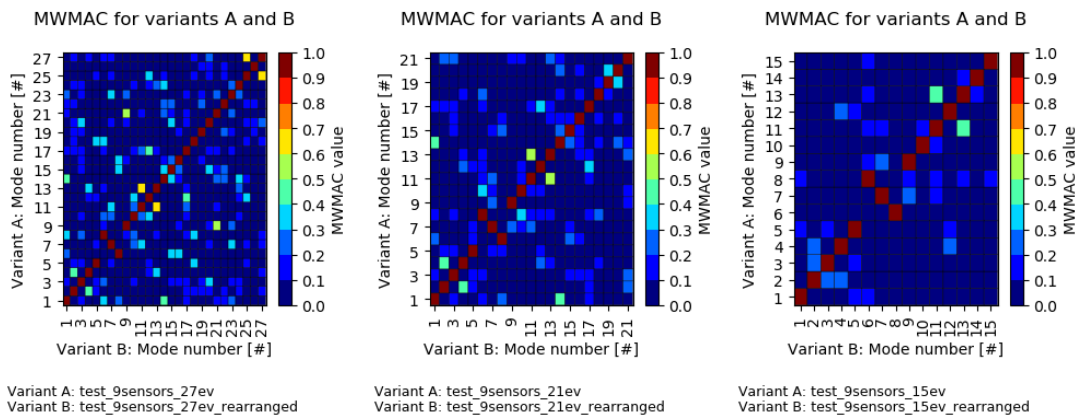


Figure 4.5: Dependence of MAC and MWMAC on number of modal vectors.

4.1.4 PLR – Pairing and Linear Regression

PLR is a global comparison of two variants based on MAC. Pairing is the designation for selecting pairs of the most similar modal vectors according to MAC, with assigning the corresponding frequencies, from the first and the second variant (see 4.6). Linear regression then denotes the weighted fitting of the assigned pairs of frequencies by a linear curve with only a slope coefficient (regression model without the intercept term, see the section 3.1.1), that indicates a stiffness ratio between the variants (see 4.7). The approach of the PLR global comparison is known and was only extended in the part of the pairing [23].

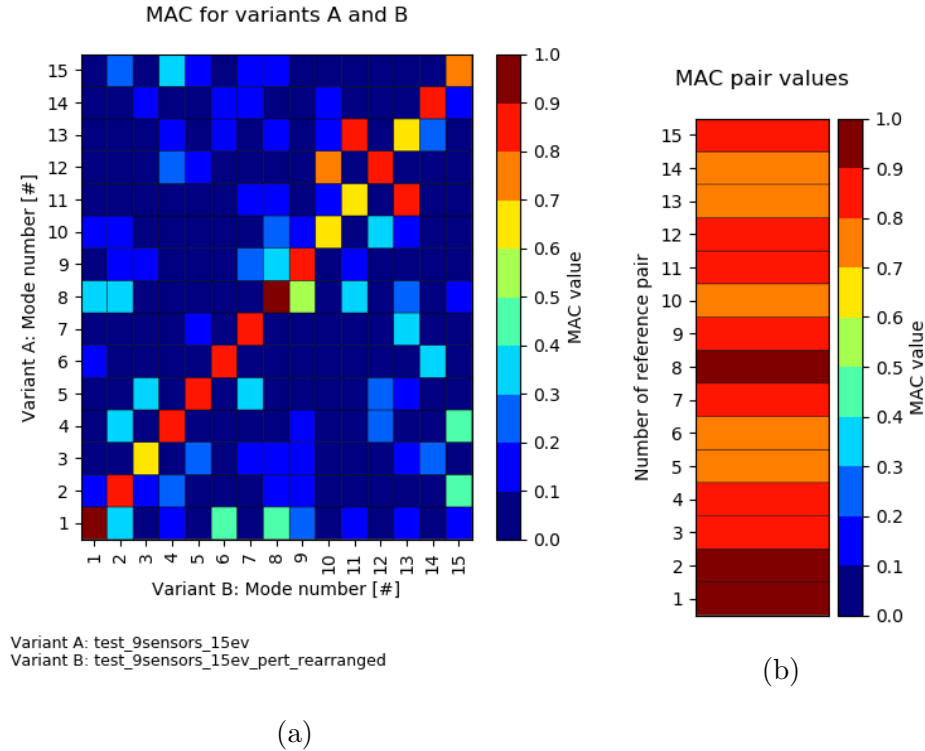


Figure 4.6: (a) MAC, (b) MAC paired values (per columns).

The PLR approach assumes sufficiently ‘sharp’ MAC results in the sense that in each column and each row of the MAC matrix there is at most one value close to one, indicating the correlation, and all other values indicating the uncorrelation. Because this assumption is not always met in reality, five possible types of pairing are considered,

- pairing per rows of MAC matrix (a reference variant is the first variant),
- pairing per columns of MAC matrix (a reference variant is the second variant),
- pairing per the shorter size of MAC matrix (a reference variant is the variant with less modal vectors contained),
- pairing per the longer size of MAC matrix (a reference variant is the variant with more modal vectors contained),

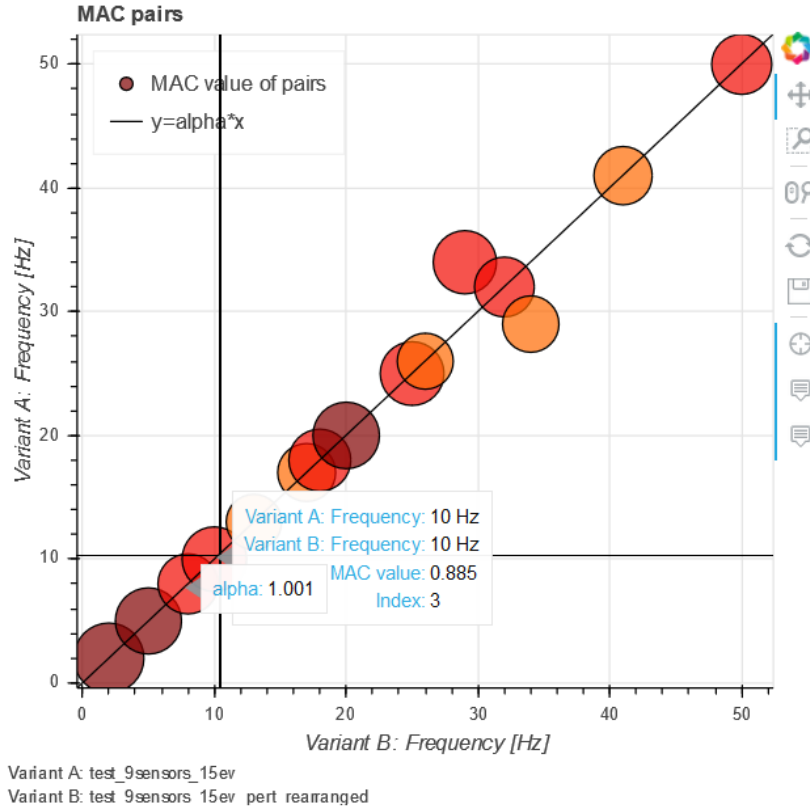


Figure 4.7: PLR interactive plot.

- pairing with unique maximization (no reference variant).

The first four approaches are based on ‘looking’ at the MAC matrix from one reference direction and searching for (a vector of) maximum values in the MAC matrix from this reference direction. e.g. with pairing per rows of MAC matrix the maximum value from each row is selected (in one column, multiple values can be selected). The last approach, pairing with unique maximization, is a new approach that searches for values so that each row and column contains at most one such value and the sum over all the selected values is the highest possible (the highest values may not be selected). Pairing with unique maximization is done by the Munkres algorithm (in a Python implementation, a module [28] was used).

In the linear regression the weighting is done by MAC values corresponding to pairs, the fitted linear function is of the form $\omega_1 = \alpha \omega_2$, where ω_1, ω_2 (take attention on top indexes here) are variables (frequencies) and $\alpha \in \mathbb{R}$ is a constant gained from a set of equations

$$\text{diag}(MAC) \omega^{1p} = \alpha \text{diag}(MAC) \omega^{2p},$$

ω^{1p}, ω^{2p} are vectors of the paired frequencies from the first and the second variant (in the s -th position of both vectors is the s -th pair of assigned frequencies, the letter p is used only to indicate the pairing), $\text{diag}(MAC)$ is a diagonal matrix with MAC values corresponding to assigned frequencies on a diagonal.

Denoting

$$\begin{aligned} y &= \text{diag}(MAC) \omega^{1p}, \\ X &= \text{diag}(MAC) \omega^{2p}, \end{aligned}$$

the value of α is given by

$$\alpha = (X^* X)^{-1} X^* y.$$

Together with the estimation of the regression coefficient α , a coefficient of determination R^2 is also calculated. If S denotes the number of elements of the vector y , then R^2 is calculated as

$$R^2 = 1 - \frac{S_e}{S_T} = 1 - \frac{\sum_{s=1}^S (y_s - (X\alpha)_s)^2}{\sum_{s=1}^S \left(y_s - \frac{\sum_{s=1}^S y_s}{S} \right)^2}.$$

The coefficient α represents the stiffness ratio between the variants, $\alpha > 1$ means that the first variant is stiffer than the second one (analogously for $\alpha < 1$).

The coefficient of determination expresses how well the regression curve passing through the origin approximates the linear relationship between the assigned pairs of frequencies. The value of R^2 equal to 1 indicates that the regression line perfectly fits the data and the lower the R^2 , the more local variations, with larger MAC values, between the assigned frequencies values. Note that R^2 in this case (of regression model without the intercept term) can be any number from the interval $(-\infty, 1]$ (see the section 3.1.1).

MWPLR – Mass Weighted Pairing and Linear Regression

Pairing and linear regression done with the MWMAC weighting instead of MAC.

4.2 Selective criteria

All selective criteria follow the pairing and serve as a root cause analysis attempting to capture which positions (at which frequencies and channel specifications) lead to poor correlations. Selective criteria can be applied to all paired modal vectors and frequencies or to paired modal vectors and frequencies that correspond to a higher MAC value only (the ‘high pairs only’ option in the Python implementation).

Selective criteria are divided into:

- with or without mass weighting (MW),

- over DOF 1 – 6 or over sensors,
- with component removal as ‘what if analysis’ or with COMAC (see below).

All following criteria except COMAC are new tools.

4.2.1 COMAC – Coordinate MAC

A known criterion COMAC compares two sets of paired modal vectors to identify which channels lead to poor correlation of these the most similar modal vectors. This criterion provides the correlation through all the pairs for every channel individually (i.e. the result is a vector of these correlations).

Let be S pairs of assigned modal vectors and let us designate matrices of paired modal vectors as V^{1p} , V^{2p} , where in the s -th column of both matrices is the s -th pair of assigned modal vectors (the letter p is used only to indicate the pairing; take attention on top indexes here). For the r -th channel, an element v_{rs}^{1p} of the matrix V^{1p} at position r , s (the r -th row, the s -th column) and an element v_{rs}^{2p} of the matrix V^{2p} , the COMAC value is defined by

$$COMAC_r = \frac{\sum_{s=1}^S |v_{rs}^{1p} (v_{rs}^{2p})^*|^2}{\sum_{s=1}^S (v_{rs}^{1p})^2 \sum_{s=1}^S (v_{rs}^{2p})^2}.$$

COMWMAC – Coordinate MWMAC

As COMWMAC was named a mass weighted (extended) version of COMAC criterion, that is applicable only to square modal matrices of the same sizes (and thereby mass matrices of the same sizes).

Let V^{1p} , V^{2p} be the matrices with paired modal vectors as columns (in the s -th column of both matrices is the s -th pair of assigned modal vectors). If the modal matrices V^{1p} , V^{2p} are not of the same square sizes, some rows and columns of the modal matrices are deleted in the Python implementation. Let R denote the number of degrees of freedom (channels) per sensor and $AUTOMAC_R$ the AUTOMAC for modal matrices with R rows removed. While the number of rows is greater than the number of columns, iteratively R rows per each sensor of both modal matrices are deleted so that the L^2 norm $\|AUTOMAC - AUTOMAC_R\|$ is minimal. Then the last S_1 and the last S_2 columns in the first and the second modal matrix are deleted so that the resulting matrices are square.

The mass matrices are obtained as inversions by the pseudoinverse methods (described in the section 2.2), where the M -orthonormalized modal vectors are assumed.

For two rows \tilde{v}_1 , \tilde{v}_2 of the modal matrices V^{1p} , V^{2p} as row vectors, the COMW-

MAC matrix value is calculated as

$$COMWMAC \text{ matrix}_{\tilde{v}_1, \tilde{v}_2} = \frac{|\tilde{v}_1 \sqrt{M_1} \sqrt{M_2} \tilde{v}_2^*|^2}{\tilde{v}_1 M_1 \tilde{v}_1^* \tilde{v}_2 M_2 \tilde{v}_2^*},$$

where $\sqrt{M_1}$ is the square root of the mass matrix obtained for a variant described by V^{1p} , $\sqrt{M_2}$ analogously. The COMWMAC matrix should be ideally an identity matrix and the diagonal of this matrix is taken as the COMWMAC, used for the further criteria.

4.2.2 SC – Selective Criterion

New criterion SC serves as a ‘what if’ analysis, where sensors are sequentially removed. If we denote a row vector of original MAC values for the paired modal vectors (without any sensor removal) as MAC and a row vector of MAC values for the paired modal vectors without the r -th sensor information (all of its channels-DOF) as MAC_r , then the r -th row of relative differences in the SC matrix is calculated as

$$SC_r = \frac{MAC - MAC_r}{MAC},$$

where the division is done element-wise (see Fig. 4.8).

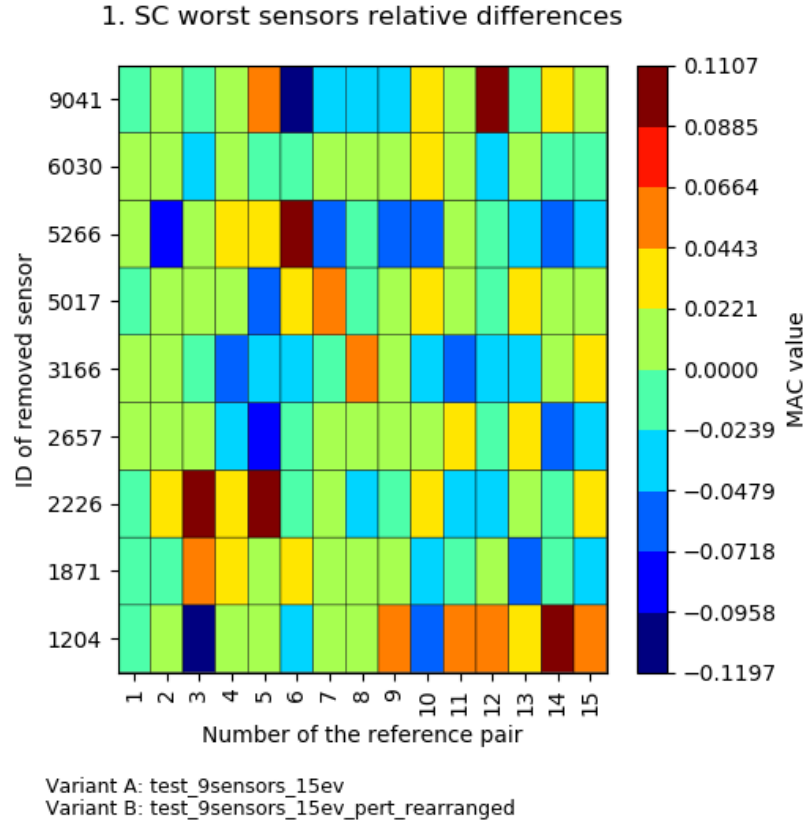


Figure 4.8: Example of SC.

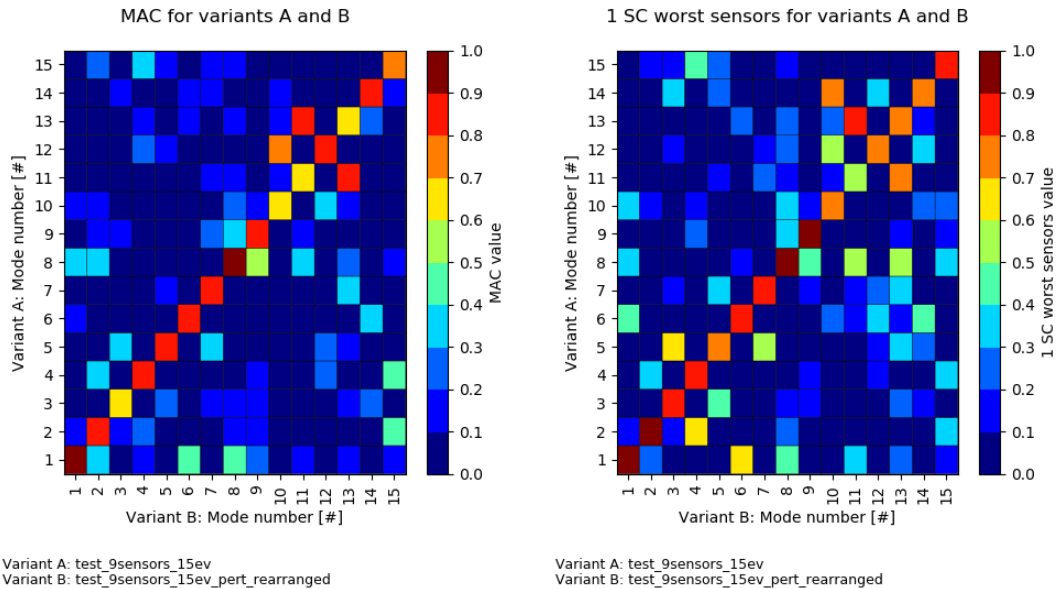
All the relative differences are less or equal to one. The value of SC_r element

- cca 1 (MAC_r close to 0) indicates that other sensors except the r -th capture the differences in the pair (the removed sensor had been capturing the similarity in the pair),
- cca 0 (MAC_r is similar to MAC) indicates no change,
- less than 0 (MAC_r is higher than MAC) indicates that the r -th removed sensor captures the differences in the pair.

Each MAC_r row of values is then averaged and the sensor with the minimal average value is selected as the worst sensor that is removed and the whole process is iteratively executed as many times as the chosen ‘number of selected sensors’.

During the process, the changes besides the pairs are not followed – after any sensor removal the nondiagonal (or noncorresponding) values can differ as well and new higher correlated pairs can occur (see Fig. 4.9). Thus, SC is meaningful only for data with unambiguous stable pairing (MAC matrix ‘sharp enough’), if the pairing is not changed after the sensor removal. Plus it is better if there is enough of sensors, so that after removing some of sensors, there is still enough of information for modal vectors distinguishing (the same reason).

In the implemented version, there is also a possibility to process the SC criterion only for pairs with higher MAC value greater than chosen limit (‘high pairs only’).



(a)

(b)

Figure 4.9: (a) MAC, (b) MAC for data after the ‘worst sensor’ removal.

SMWC – Selective Mass Weighted Criterion

As SMWC is named a mass weighted version of SC criterion using MWMAC instead of MAC.

4.2.3 SC with COMAC

Relating SC with COMAC, sensors are sequentially removed and the COMAC criterion is calculated. COMAC returns a vector of correlations of all paired modal vectors per channels, therefore the COMAC results are averaged within sensors. The resulting values range from zero to one and shows the correlation of the pairs per each sensor (see Fig. 4.10).

SC with COMAC is again meaningful only for data with unambiguous stable pairing (MAC matrix ‘sharp enough’).

Therefore, in the implemented version, there is also a possibility to process the SC with COMAC only for pairs with higher MAC value greater than chosen limit.

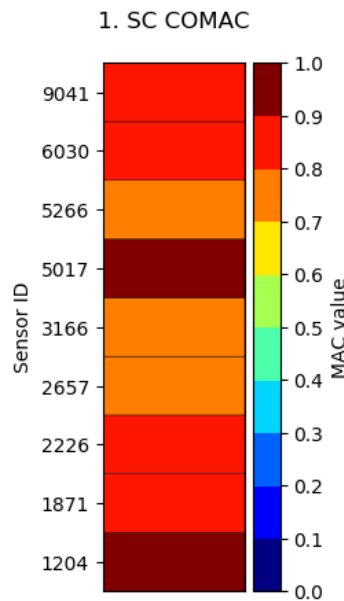


Figure 4.10: Example of SC with COMAC.

SMWC with COMWMAC

As SMWC with COMWMAC is named a mass weighted version of SC with COMAC (MWMAC instead of MAC and COMWMAC instead of COMAC), that is applicable only to M-orthonorma-lized modal vectors. Due to the possible removal of columns of rectangular modal matrices, some sensors may not be involved in the criterion comparison.

4.2.4 SDOFC – Selective Degrees of Freedom Criterion

New criterion SDOFC serves as a ‘what if’ analysis, where the data per degrees of freedom DOF 1 – 6 are considered individually. If we denote a row vector of original MAC values for the paired modal vectors as MAC and a row vector of MAC values calculated only for the r -th DOF for all sensors of paired modal vectors as MAC_r , then the r -th row of SDOFC matrix of relative differences is defined by

$$SDOFC_r = \frac{MAC - MAC_r}{MAC},$$

where the division is done element-wise.

All the relative differences are less or equal to one. The value of $SDOFC_r$ element

- cca 1 (MAC_r close to 0) indicates differences in the r -th DOF in the pair,
- cca 0 (MAC_r is similar to MAC) indicates that the r -th DOF in the pair is without the change,
- less than 0 (MAC_r is higher than MAC) indicates differences in the other than the r -th DOF in the pair.

SDOFMWC – Selective Degrees of Freedom Mass Weighted Criterion

As SDOFMWC is named a mass weighted version of SDOFC criterion using MWMAC instead of MAC.

4.2.5 SDOFC with COMAC

Relating SDOFC with COMAC, the mode shape data are considered separately for DOF 1 – 6 and the COMAC criterion is calculated. This criterion results are exactly the same as pure COMAC, the only difference is that it is rearranged and grouped by DOF. The resulting values range from zero to one and shows the correlation of the pairs per each channel. Pairs with low MAC values cause an overall poor COMAC correlation.

SDOFMWC with COMWMAC

As SDOFMWC with COMWMAC is named a mass weighted version of SDOFC with COMAC criterion (MWMAC instead of MAC and COMWMAC instead of COMAC), that is applicable only to M-orthonormalized modal vectors.

4.3 Criteria applied on real data

We show all the criteria used on data. The data comes from modal analysis of the construction-vehicle axle on the test bench. Three (slightly) different measurements were done. The first measurement is of the axle with attached planetary gear in its original configuration, this variant of data are denoted with 0.0D label. Then the position of the planetary gears were shifted, in the second measurement by only 4.5 degrees from the original settings, labelled with 4.5D, and in the third measurement by the largest extent by 22.5D degrees, marked with 22.5D label. The aim was to find out how the different positioning of planetary gears impacts on the resulting modal properties of the system, if and how these variants differ. We take always the variant 0.0D as the variant A and compare it to 4.5D or 22.5D as variant B and we take a look on comparison 0.0D – 22.5D in more detail.

One of each three data sets consists of a vector of approximately 100 natural frequencies from range of 0Hz to cca 1000Hz, a vector of 111 sensor ID specifications (37 sensors with 3 channels – 3 DOF per sensor) and a matrix of corresponding obtained modal vectors.

Firstly we show MAC for variants A:0.0D, B:4.5D only for frequencies less than 230Hz and selected modal vector of variant A for frequency 201Hz compared with three different modal vectors of variant B corresponding to different MAC values, so the MAC sensitivity can be seen (see Fig. 4.11, 4.12).

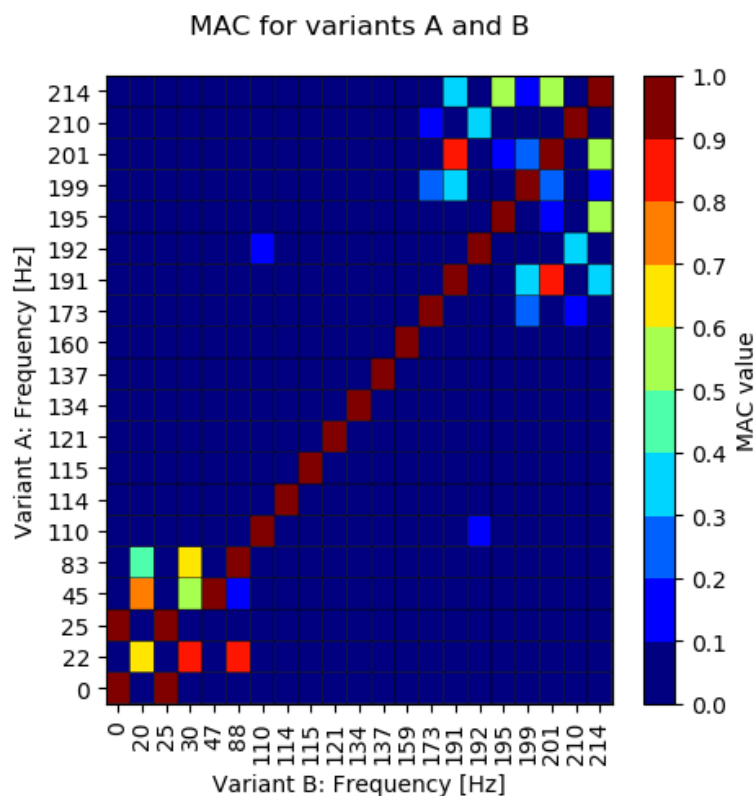
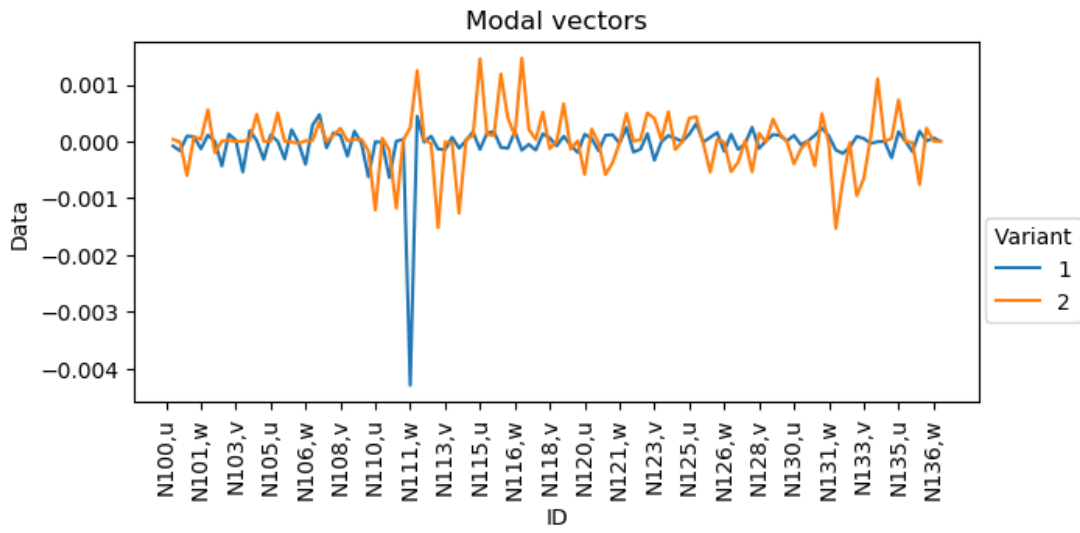
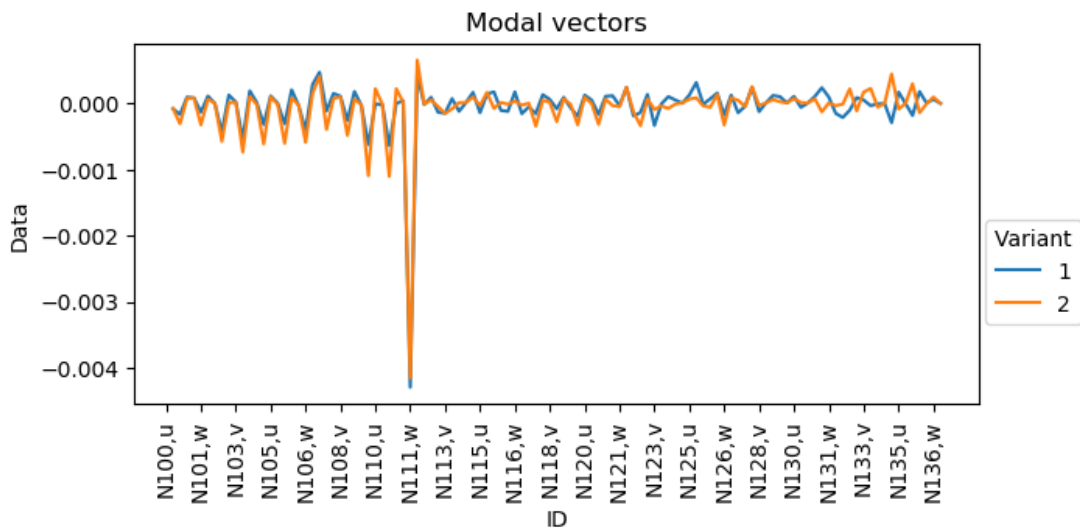


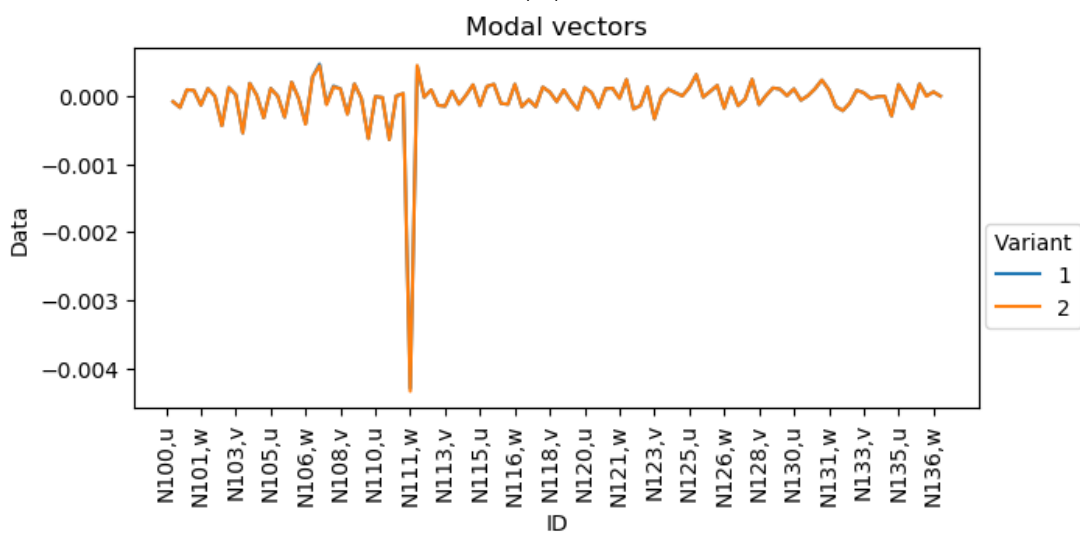
Figure 4.11: MAC for variants A:0.0D, B:4.5D.



(a)



(b)



(c)

Figure 4.12: Modal vectors for A:0.0D, B:4.5D and frequencies (a) A:201Hz, B:111Hz, (b) A:201Hz, B:191Hz, (c) A:201Hz, B:201Hz.

The whole MAC for A:0.0D – B:4.5D and also A:0.0D – B:22.5D comparison follows (see Fig. 4.13, 4.14). In the first case, the resulting MAC matrix is relatively close to a diagonal matrix, in the second case the corresponding pairs of modal vector are often shifted, because of some new natural frequencies and modal vectors without the pairs in the second variant of data.

We show AUTOMAC for variant A (see Fig. 4.15a), it is not a pure diagonal matrix. We display also the MAC, where we take into account the distances of the individual natural frequencies (ranges between them), for variants A:0.0D – B:22.5D (see Fig. 4.15b).

And there is compared MAC and MAC linearly scaled for variants A:0.0D – B:22.5D in Fig. 4.16. MAC linearly scaled is more sensitive (distinguishes between modal vectors more) around high and low MAC values.

Then we show results of MWMAC for A:0.0D – B:4.5D (see Fig. 4.17) and MWMAC and all three MOC criteria for A:0.0D – B:22.5D (see Fig. 4.18). MWMAC improved the results in both cases compared to MAC. For computing the mass matrices, the data corresponding to 4 sensors with all their 3 DOFs were deleted in the modal matrices (see 2.2).

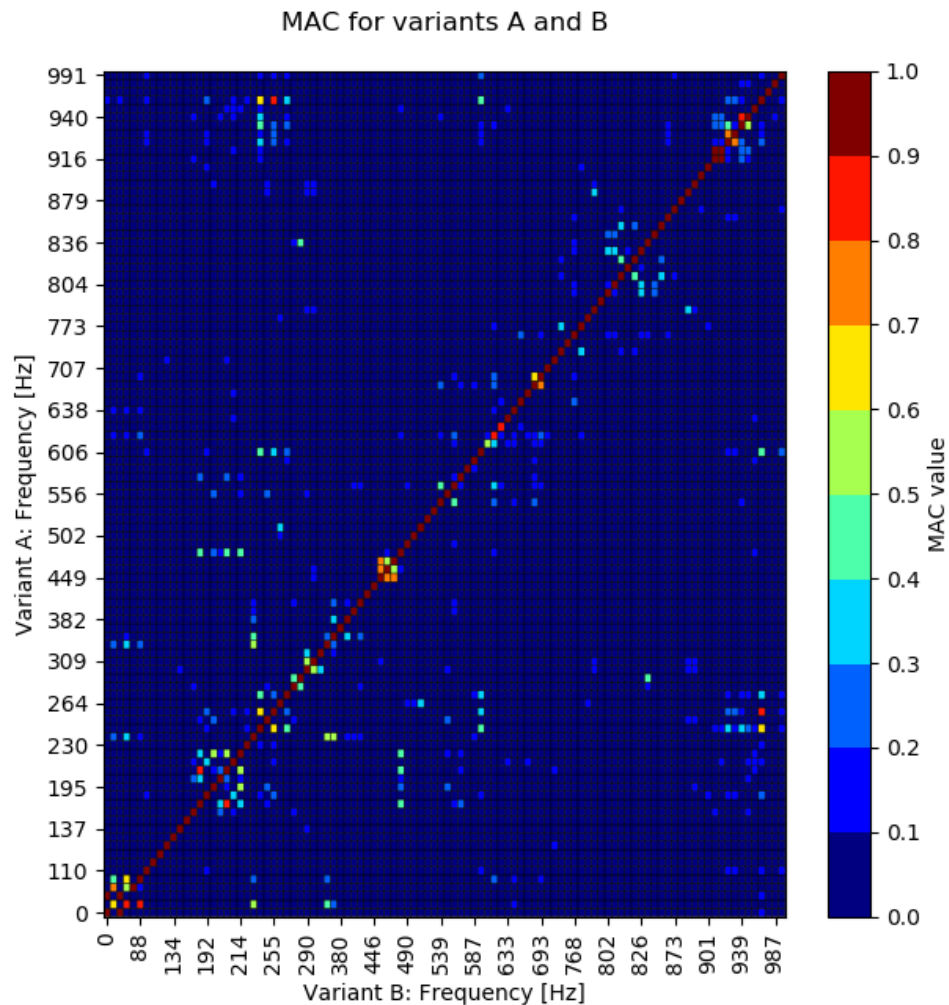


Figure 4.13: MAC for variants A:0.0D, B:4.5D.

MAC for variants A and B

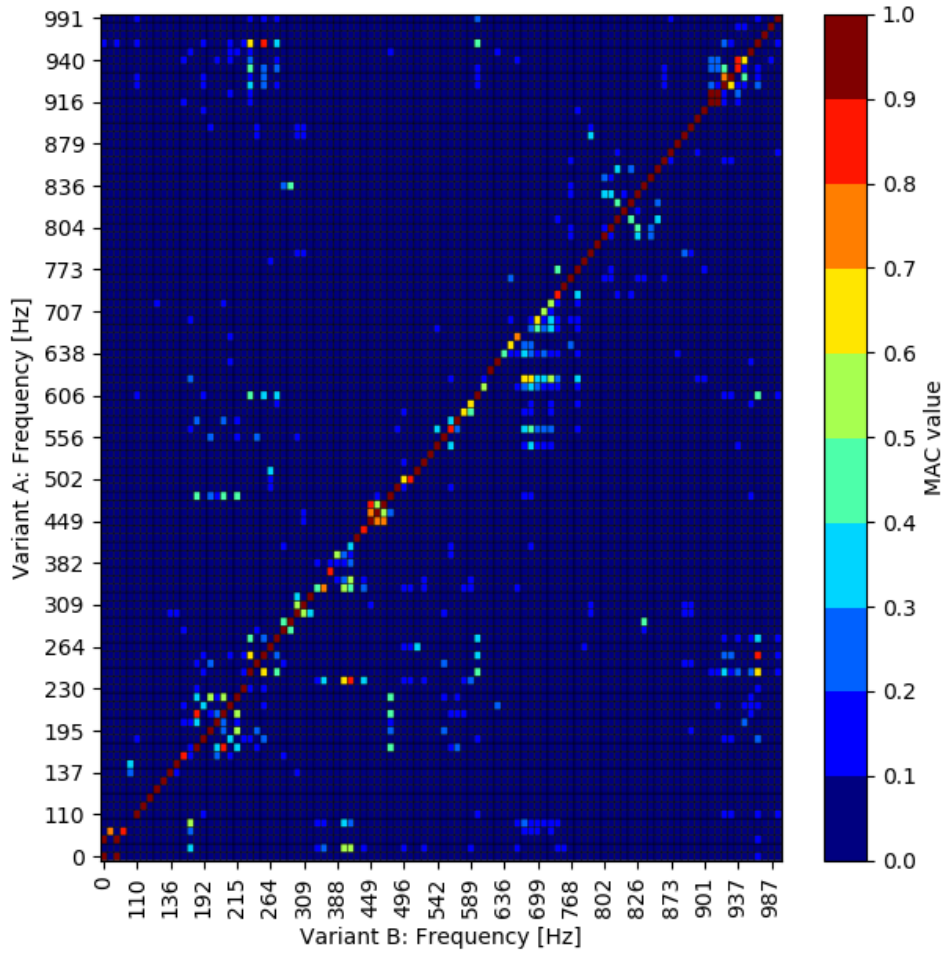


Figure 4.14: MAC for variants A:0.0D, B:22.5D.

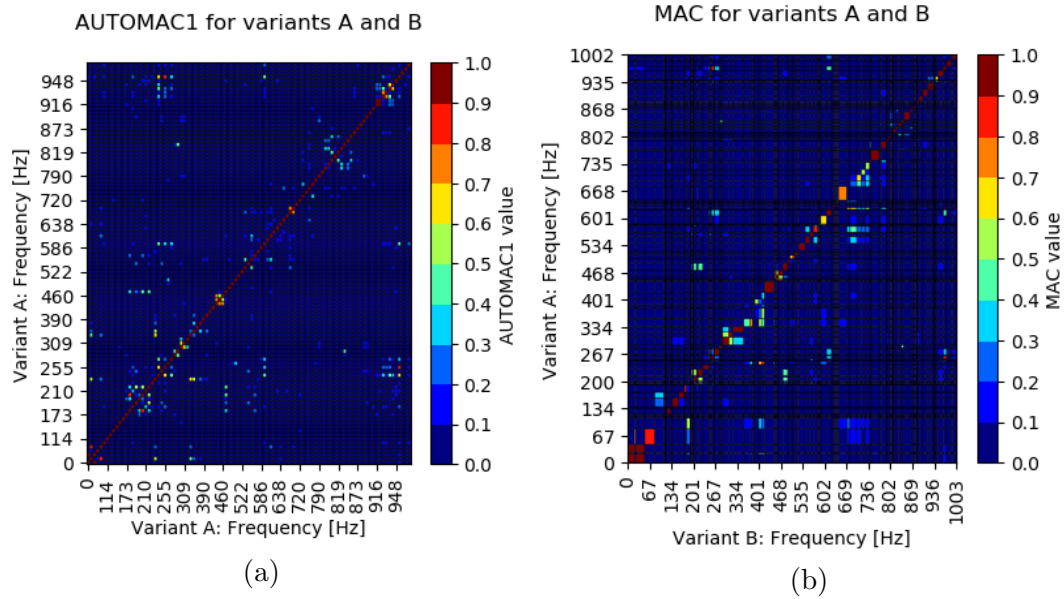


Figure 4.15: (a) AUTOMAC for variant A:0.0D, (b) MAC with frequency ranges for variants A:0.0D, B:22.5D.

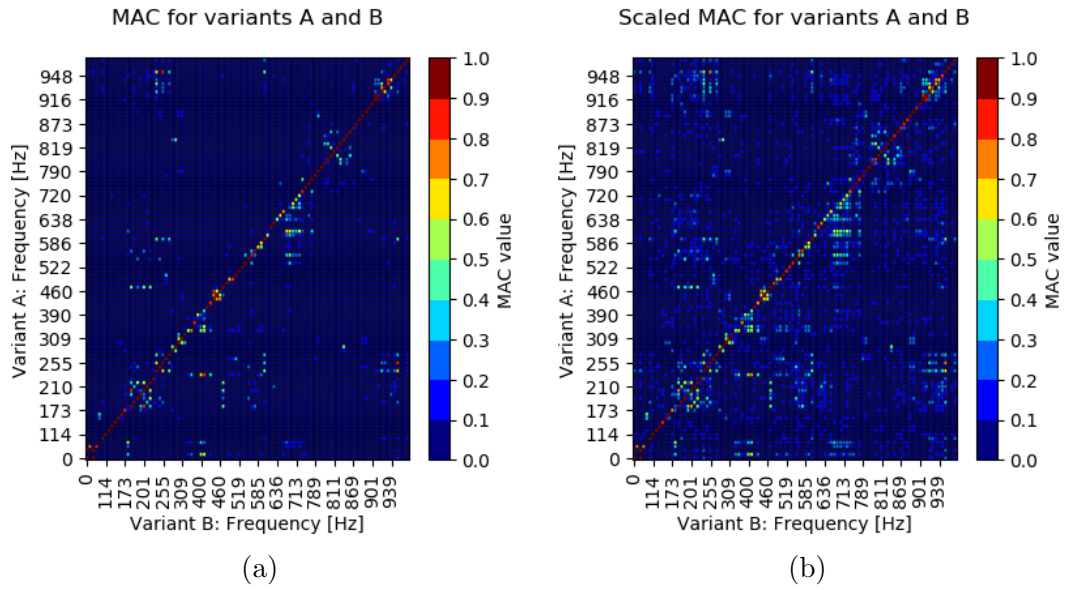


Figure 4.16: (a) MAC and (b) MAC linearly scaled for variants A:0.0D, B:22.5D.

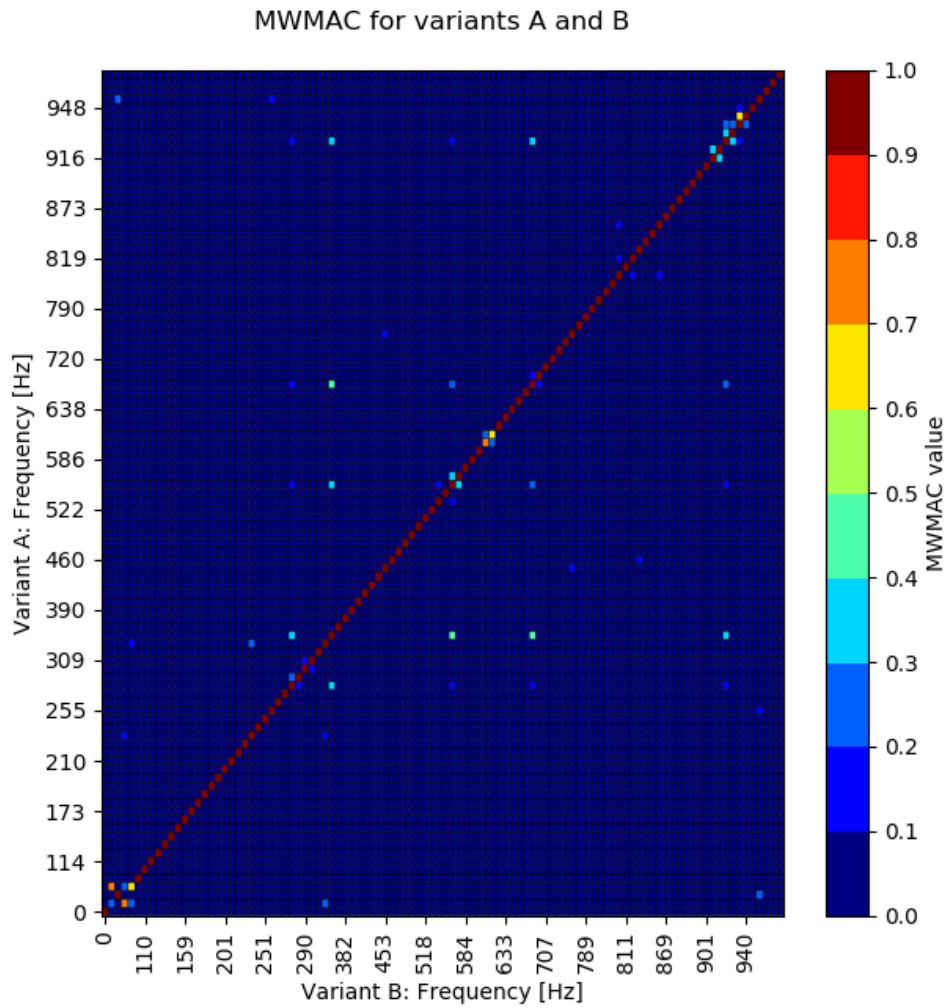


Figure 4.17: MWMAC for variants A:0.0D, B:4.5D.

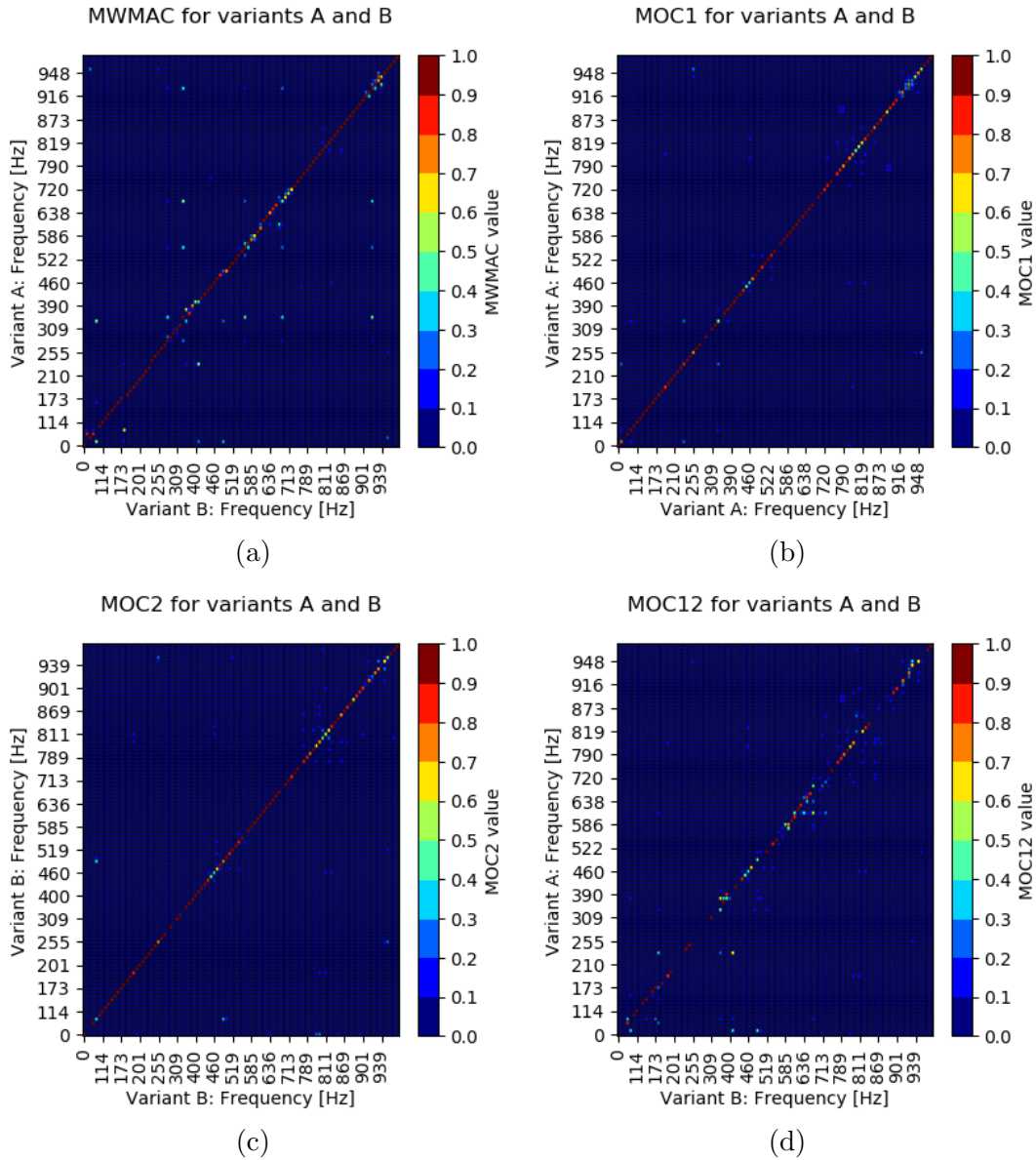


Figure 4.18: (a) MWMAC, (b) MOC1, (c) MOC2, (d) MOC12 for variants A:0.0D, B:22.5D.

From now on we continue only with the comparison of variants A:0.0D, B:22.5D and without the mass weighting. The reason is that the following criteria focus on identifying differences in the data and in this case we get the results displaying the largest differences in the data, well visible.

We performed the pairing per the shorter size of the MAC matrix, that is with the first reference variant A, and with PLR obtained the stiffness ratio 0.995 and the coefficient of determination 0.991, meaning that the second variant is stiffer than the first one in terms of vibrations. There is a screen shot of an interactive plot of MAC values for pairs together with the fitted linear curve in Fig. 4.19. Plot of pairs and so called absolute and relative frequency differences are displayed in Fig. 4.20. The absolute frequency differences are calculated as differences between assigned paired frequencies and they are in addition divided by the first reference frequencies and multiplied by 100% in case of relative differences.

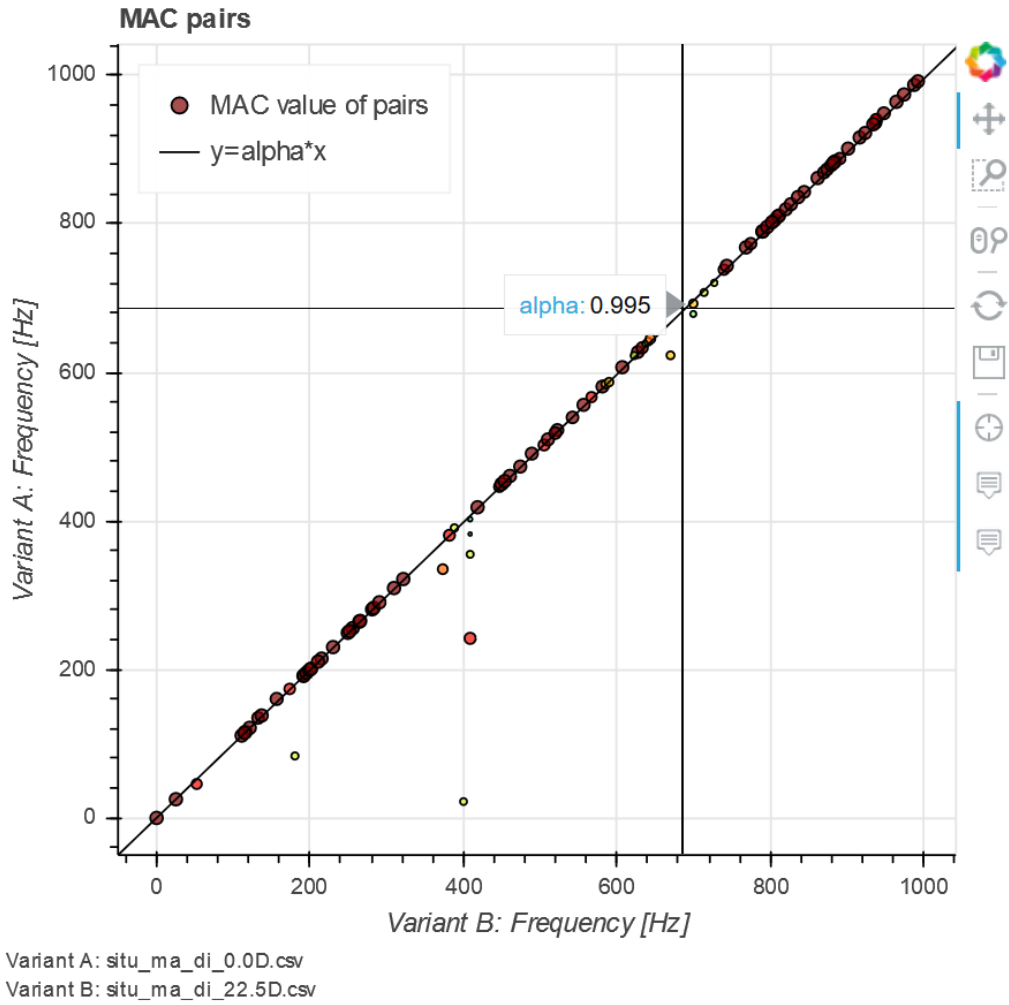


Figure 4.19: MWMAC for variants A:0.0D, B:22.5D.

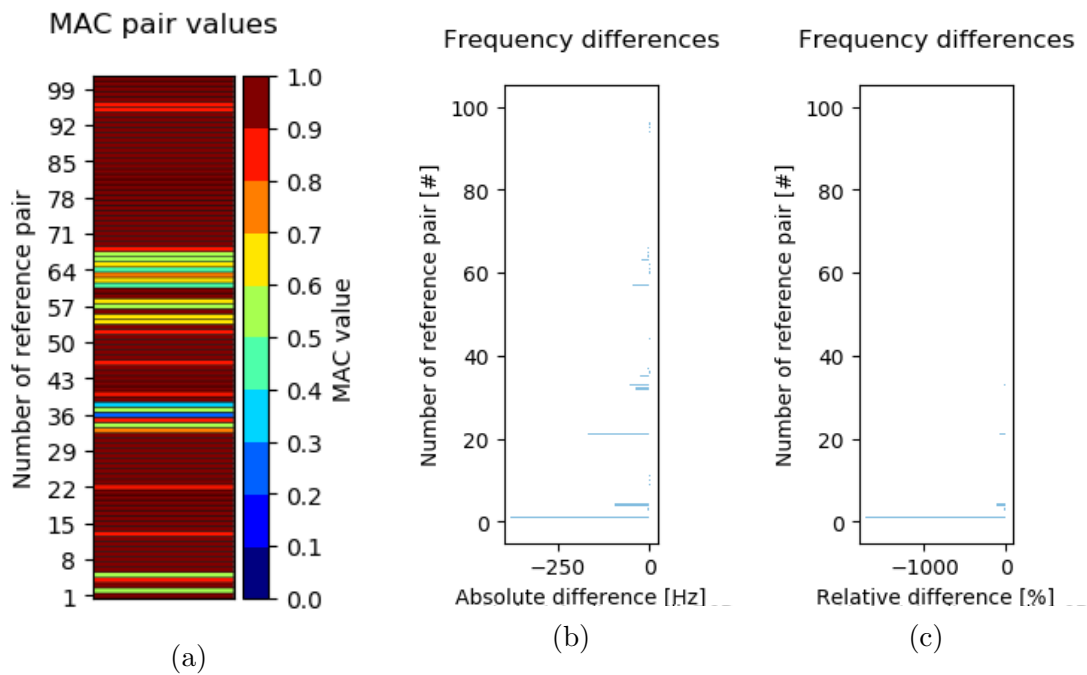


Figure 4.20: (a) Pairs, (b) absolute and (c) relative frequency differences.

The selective criterion SC follows (see Fig. 4.21). All the green and turquoise colors in the picture mean almost no differences after the sensor removal in the resulting MAC value for the pairs. The red color, e.g. for sensor ID 120 and the second pair of modal vectors, indicates that the other sensors except the one with ID 120 capture the differences in the second pair (and also in the 22-th and the 34-th pairs), that is, the sensor with ID 120 is capturing the similarities (except the scaling factor for modal vectors) in the second pair. Whereas the blue color, e.g. for the sensor ID 120 and the fifth pair of modal vectors, indicate that the sensor with ID 120 captures the differences in the fifth pair (and also in the 37-th and the 38-th pairs).

It is also possible to plot the whole MAC matrix after the sensor removal to check, what happens with non-paired MAC values, or to continue with the process after one ('the worst') sensor is removed as well. In all cases the results depend on the 'original' MAC values, it is always a comparison highlighting the differences in MAC (without some sensor information) relative to the original MAC matrix.

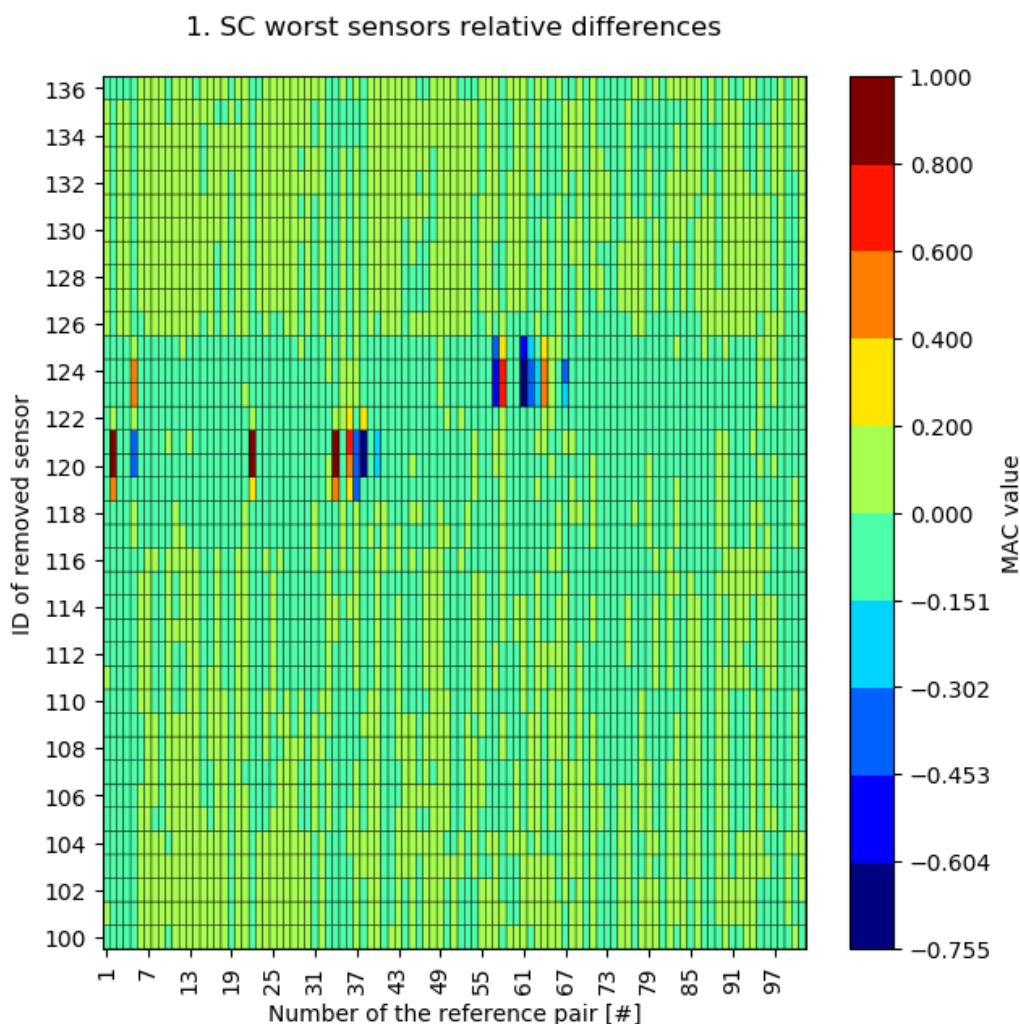


Figure 4.21: SC for variants A:0.0D, B:22.5D.

The COMAC criterion is sensitive to the values of all pairs, therefore, we firstly present the COMAC for all pairs (with the first reference variant A) and then the

COMAC for pairs with MAC value greater than 0.7 (see Fig. 4.22). Although all the MAC values in the last case are greater than 0.7 for the pairs, quite a lot of COMAC correlations across the pairs for sensors individually are still low. The best results are obtained for sensors with IDs (109) 116 and (130) 132, 133. In the previous SC criterion, after removal of these individual sensors, no large changes in new MAC values compared to the original MAC values for the pairs are displayed, i.e. no large similarities captured by e.g. the sensor 116 indicated using the MAC correlation across the sensors and the sensors except the sensor with ID 116.

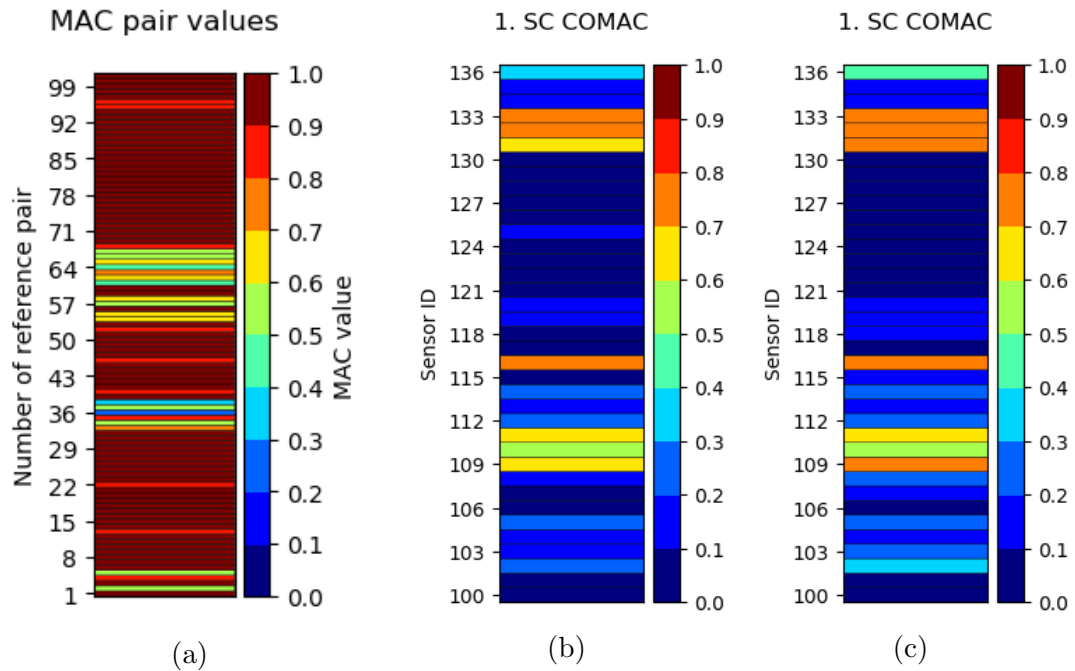


Figure 4.22: (a) Pairs, (b) SC with COMAC for all pairs, (c) SC with COMAC for the pairs with MAC value greater than 0.7.

Then we show the SDOFC criterion for all pairs of modal vectors (see Fig. 4.23). There are displayed the differences in MAC values when considering only the specific DOF separately compared to original MAC values for the pairs in the plot of SDOFC. All the green and turquoise colors again indicate minimal differences. The red color here, e.g. for the first DOF and the first pair of modal vectors, indicates the differences in the first DOF in the first pair. The (dark) blue color, e.g. in the second DOF and the 36-th pair of modal vectors, indicates similarities in the second DOF and in the 36-th pair.

Finally, we show the SDOFC with COMAC for all pairs of modal vectors (see Fig. 4.24). After comparing it to the COMAC result for sensors (see Fig. 4.22b), we get e.g. the information, that the high correlations for sensors with IDs (109) 116 and (131) 132, 133 are primarily in the second and the third translational DOF direction and all the correlations within the first DOF are low (less than 0.3). Then the high correlation just for the specific DOF can be seen.

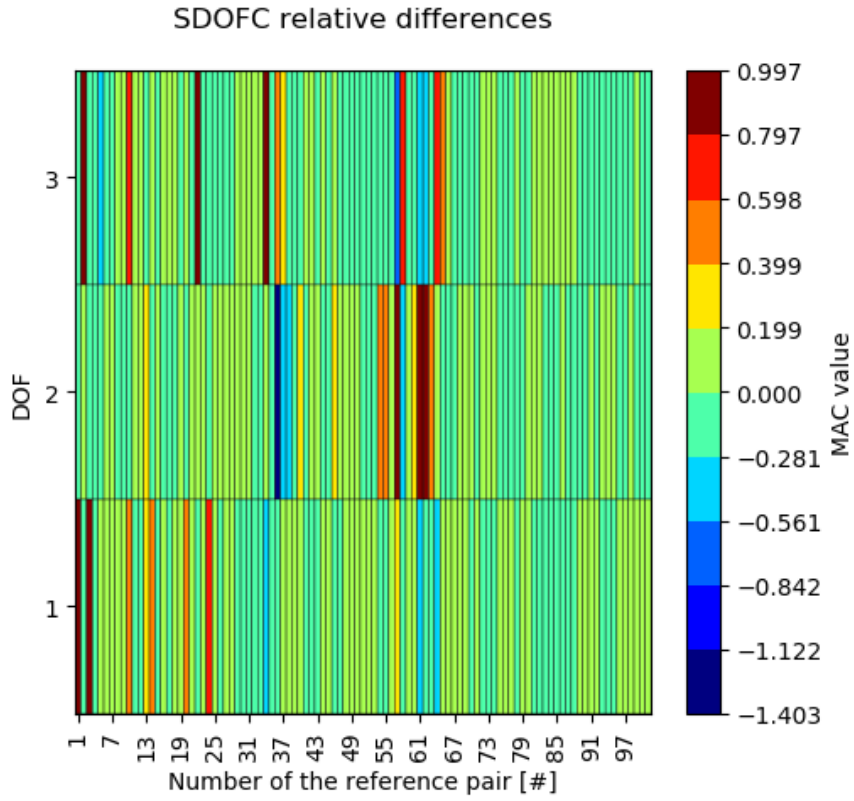


Figure 4.23: SDOFC for variants A:0.0D, B:22.5D.

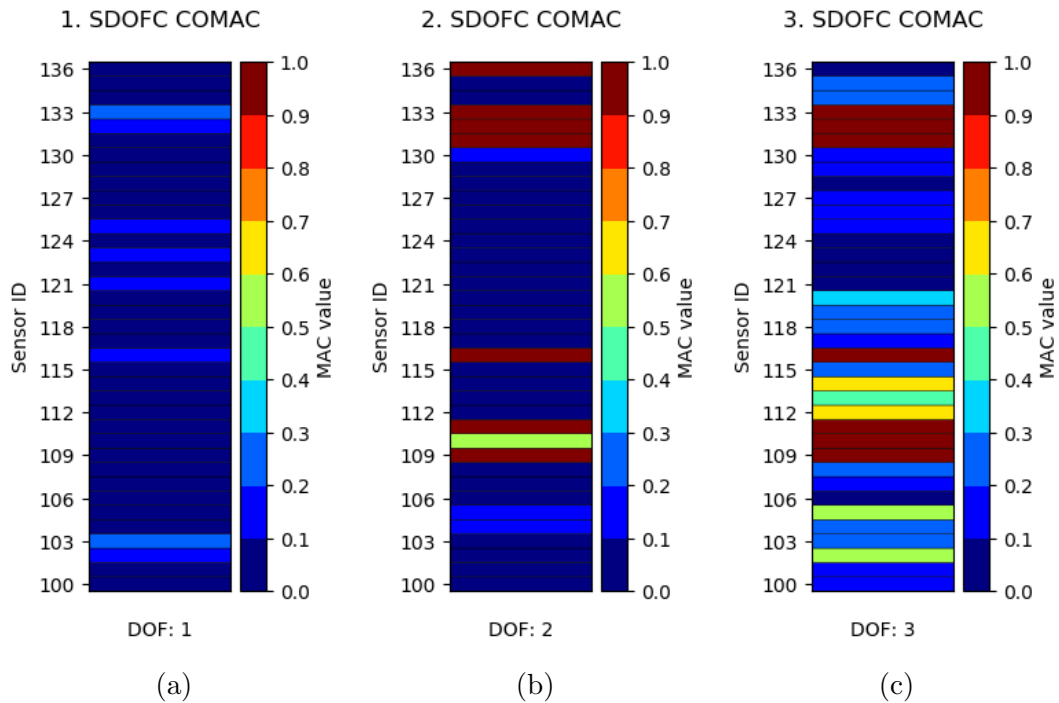


Figure 4.24: SDOFC with COMAC for all pairs (a) for the first DOF, (b) for the second DOF, (c) for the third DOF.

Chapter 5

Statistical processing of multiple data sets

This chapter is devoted to generalizations through regression and ANOVA approaches of MAC or FDAC criteria that can be used for comparison of multiple data sets. Of course, even in case of multiple input data, we could use MAC (FDAC) for each pair of input data and then process the results. However, our goal is to set up a criterion that provides a direct results similar to MAC (FDAC) for more input data. All the tools designed in this chapter are new procedures.

The data sets again consist of a vector of resonant frequencies and matrix of modal vectors together with channel specifications, which designate modal vector components, or all of them consist of a vector of frequencies, matrix of deflection shapes and channel specifications for deflection shape components. According to whether the methods are applied on modal vectors or deflection shapes, the methods are generalizations of MAC or FDAC. However, formulas are the same (just like formulas for MAC and FDAC are the same). Therefore, we will only use the terms ‘modal vectors’ and ‘resonant frequencies’ from now on, but the generalizations are applicable to deflection shapes as well.

Now, we want to compare more than two data sets. We distinguish between multiple variants of systems and more measurements of each variant of system. In the first criterion called MM, one reference variant is compared to the second variant, that may be measured more times. MM may be used for comparing e.g. one simulation and 10 measurements (of the same type) of the system or it may be used in the same way as MAC for just two variants. Other criteria serve for comparison of more variants, there is no reference variant (in the criteria themselves), but more measurements of each variant are required, making the criteria suitable e.g. for comparing 10 measured systems, each measured twenty times (in the same way). The result of the methods is a “degree of similarity” of all the variants (compared to the reference one in case of MM or all compared to all others) with an indication of differences across frequencies and eventually across channel specifications (modal vectors elements).

The detailed descriptions of individual methods follow.

5.1 Basic criteria

The basic criteria consist of a regression model and the associated statistical testing. The model and the testing is always introduced for one set of modal vectors and then applied as criterion independently for more combinations of modal vectors selection (as well as the MAC criterion is introduced for two modal vectors, but applied for all combinations of modal vectors). Three regression models were derived during the generalization that varies in use, however, the same type of statistical tests were then applied to each model.

5.1.1 MM – More Measurement

As a MM criterion, we named a basic criterion, on which we primarily try to show the consistency with MAC. MM criterion is a direct generalization of MAC.

MM criterion can be used to determine the degree of similarity of modal vectors from two variants, where the first variant is the reference variant once measured and the second variant may be more times measured. It is compared whether the modal vectors from the second variant are all (from all measurements) the same multiple of the modal vectors from the first reference variant. All modal vectors of the first variant with all the modal vectors of the second variant (n -th vector of all measurements) are independently compared in the same way.

In MM model the complex input data are decomposed into real and imaginary parts. In the case of two variants both once measured, it turns out that the modified coefficient of determination R_0^2 of the model is the same as the MAC criterion.

Let us firstly focus on a special case of MM criterion for two variants both once measured and then introduce the MM criterion in general.

MAC to MM transition

When comparing two modal vectors v^1 and v^2 , the MAC coefficient describes the degree of linearity of the vectors. When the MAC value for two vectors is equal to one, it means that (e.g.) the modal vector v^2 is just a scaled modal vector v^1 . When the MAC value is zero, the two modal vectors are completely different and we can not write the one vector using the other as its multiple. In fact, the MAC criterion is related to a complex regression model

$$v^2 = \beta v^1 + \varepsilon, \quad (5.1)$$

where β is a regression parameter and ε a vector of errors (take attention on top indexes here, we will be using them in the whole chapter), each component in (5.1) can be complex. In order to use classical regression analysis tools mentioned in the section 3.1.1, we write (5.1) with real and imaginary components as

$$v_{Re_s}^2 + i v_{Im_s}^2 = (\beta_{Re} + i \beta_{Im}) (v_{Re_s}^1 + i v_{Im_s}^1) + \varepsilon_{Re_s} + i \varepsilon_{Im_s}, \quad (5.2)$$

for every $s = 1, \dots, S$, where S is the number of modal vector elements. By writing the relationship (5.2) by real and imaginary components into two equations and rearranging it into a linear regression model with all components real, we obtain a regression model without the intercept term. This special case of MM model takes form

$$\underbrace{\begin{bmatrix} v_{Re1}^2 \\ \vdots \\ v_{ReS}^2 \\ v_{Im1}^2 \\ \vdots \\ v_{ImS}^2 \end{bmatrix}}_y = \underbrace{\begin{bmatrix} v_{Re1}^1 & -v_{Im1}^1 \\ \vdots & \vdots \\ v_{ReS}^1 & -v_{ImS}^1 \\ v_{Im1}^1 & v_{Re1}^1 \\ \vdots & \vdots \\ v_{ImS}^1 & v_{ReS}^1 \end{bmatrix}}_X \underbrace{\begin{bmatrix} \beta_{Re} \\ \beta_{Im} \end{bmatrix}}_{\beta} + \underbrace{\begin{bmatrix} \varepsilon_{Re1} \\ \vdots \\ \varepsilon_{ReS} \\ \varepsilon_{Im} \\ \vdots \\ \varepsilon_{ImS} \end{bmatrix}}_{\varepsilon}, \quad (5.3)$$

where y is a random vector, X a regression matrix of given numbers, β is an unknown vector of regression parameters and ε a vector of unknown random deviations, which we assume meet the regression assumptions (see the section 3.1.1).

Let us note that in the model (5.3), it can be shown that the estimation of the β coefficient by the least squares method

$$b = (X^*X)^{-1} X^*y$$

is equal to a value of MSF, so called Modal Scale Factor [3]. Renaming the elements of vector y as y_r , $r = 1, \dots, R$, $R = 2S$, the modified coefficient of determination takes the form

$$R_0^2 = 1 - \frac{S_e}{S_{TT}} = 1 - \frac{\sum_{r=1}^R (y_r - \tilde{y}_r)^2}{\sum_{r=1}^R (y_r)^2}, \quad (5.4)$$

where

$$\tilde{y} = Xb. \quad (5.5)$$

Theorem 5.1. *The modified coefficient of determination R_0^2 in the regression model (5.3) is equivalent to the MAC value,*

$$R_0^2 = MAC_{v^1, v^2}.$$

Proof. For purpose of this proof, let us denote

$$y_{Re} := \begin{bmatrix} v_{Re1}^2 \\ \vdots \\ v_{ReS}^2 \end{bmatrix}, \quad y_{Im} := \begin{bmatrix} v_{Im1}^2 \\ \vdots \\ v_{ImS}^2 \end{bmatrix}, \quad x_{Re} := \begin{bmatrix} v_{Re1}^1 \\ \vdots \\ v_{ReS}^1 \end{bmatrix}, \quad x_{Im} := \begin{bmatrix} v_{Im1}^1 \\ \vdots \\ v_{ImS}^1 \end{bmatrix}$$

and also

$$x := \begin{bmatrix} x_{Re} \\ x_{Im} \end{bmatrix}.$$

That is

$$y = \begin{bmatrix} y_{Re} \\ y_{Im} \end{bmatrix}, \quad X = \begin{bmatrix} x_{Re} & -x_{Im} \\ x_{Im} & x_{Re} \end{bmatrix},$$

where matrix X has full column rank, and

$$\begin{aligned} v_1 &= x_{Re} + i x_{Im}, \\ v_2 &= y_{Re} + i y_{Im}. \end{aligned}$$

Elements of real vectors x and y are denoted with y_r and x_r , $r = 1, \dots, R$, $R = 2S$ and with $\sum(\cdot)$ we understand $\sum_{r=1}^R(\cdot)$ within this proof.

Notice that

$$\begin{aligned} X^*X &= \begin{bmatrix} x_{Re}^* & x_{Im}^* \\ -x_{Im}^* & x_{Re}^* \end{bmatrix} \begin{bmatrix} x_{Re} & -x_{Im} \\ x_{Im} & x_{Re} \end{bmatrix} = \begin{bmatrix} \sum(x_r)^2 & 0 \\ 0 & \sum(x_r)^2 \end{bmatrix}, \\ (X^*X)^{-1} &= \begin{bmatrix} \frac{1}{\sum(x_r)^2} & 0 \\ 0 & \frac{1}{\sum(x_r)^2} \end{bmatrix} = ((X^*X)^{-1})^* \end{aligned}$$

and that

$$X^*y = \begin{bmatrix} x_{Re}^* & x_{Im}^* \\ -x_{Im}^* & x_{Re}^* \end{bmatrix} \begin{bmatrix} y_{Re} \\ y_{Im} \end{bmatrix} = \begin{bmatrix} x_{Re}^*y_{Re} + x_{Im}^*y_{Im} \\ x_{Re}^*y_{Im} - x_{Im}^*y_{Re} \end{bmatrix}.$$

By rewriting the relations we get

$$\begin{aligned} R_0^2 &= 1 - \frac{\sum(y_r - \tilde{y}_r)^2}{\sum(y_r)^2} = \frac{\sum(y_r)^2 - \sum(y_r - \tilde{y}_r)^2}{\sum(y_r)^2} = \frac{\sum 2y_r\tilde{y}_r - \sum(\tilde{y}_r)^2}{\sum(y_r)^2} = \\ &= \frac{\sum 2y_r [X(X^*X)^{-1}X^*y]_r - \left[\overbrace{(X(X^*X)^{-1}X^*y)^* (X(X^*X)^{-1}X^*y)}^{y^*X((X^*X)^{-1})^*X^*y} \right]_r}{\sum(y_r)^2} = \\ &= \frac{\sum 2[y^*X(X^*X)^{-1}X^*y]_r - [y^*X(X^*X)^{-1}X^*y]_r}{\sum(y_r)^2} = \\ &= \frac{\sum [y^*X(X^*X)^{-1}X^*y]_r}{\sum(y_r)^2} = \frac{\sum ([X^*y]_r)^2}{\sum(y_r)^2 \sum(x_r)^2} = \\ &= \frac{(x_{Re}^*y_{Re} + x_{Im}^*y_{Im})^2 (x_{Re}^*y_{Im} - x_{Im}^*y_{Re})^2}{\sum(y_r)^2 \sum(x_r)^2} = MAC_{v^1, v^2}. \end{aligned}$$

□

The modified coefficient of determination R_0^2 similarly to MAC_{v^1, v^2} ranges between zero and one and for two identical modal vectors have the value of one.

In the regression model with the intercept term, the standard coefficient of determination R^2 equals to the square of the correlation coefficient ρ^2 [16][p. 99]. For a random sample (v_r^1, v_r^2) , $r \in R$, $R \in \mathbb{N}$, the square of the correlation coefficient is given by

$$\rho^2 = \frac{\left| \sum_{r=1}^R (v_r^1 - \tilde{v}^1) (v_r^2 - \tilde{v}^2) \right|^2}{\sum_{r=1}^R |v_r^1 - \tilde{v}^1|^2 |v_r^2 - \tilde{v}^2|^2},$$

where

$$\tilde{v}^n = \frac{1}{R} \sum_{r=1}^R v_r^n, \quad \text{for } n = 1, 2.$$

The value of R_0^2 (MAC_{v^1, v^2}) may be interpreted as a square of a modified correlation coefficient in the MM regression model without the intercept term. By writing the MAC_{v^1, v^2} , the square of the modified correlation coefficient is given by

$$\rho_0^2 = \frac{\left| \sum_{r=1}^R v_r^1 v_r^2 \right|^2}{\sum_{r=1}^R |v_r^1|^2 |v_r^2|^2}.$$

The modified coefficient of determination R_0^2 have also again in the MM model the meaning of a measure of how well the regression line passing through the origin $\tilde{y} = Xb$ fits the observed data y .

Theorem 5.2. *In the regression model (5.3), the modified coefficient of determination R_0^2 is zero if and only if the vector b is zero.*

Proof. If $b = 0$, then $\tilde{y} = Xb = 0$ and $R_0^2 = 1 - 1 = 0$.

If $R_0^2 = 0$, we use the relation (3.1) for S_e and obtain

$$R_0^2 = 1 - \frac{S_e}{\sum_{r=1}^R (y_r)^2} = 1 - \frac{y^*y - b^*X^*y}{\sum_{r=1}^R (y_r)^2} = 1 - 1 + \frac{b^*X^*y}{\sum_{r=1}^R (y_r)^2} = 0.$$

The coefficient R_0^2 equals zero if and only if the term b^*X^*y equals zero. And because the first (resp. the second) zero valued element of X^*y implies the first (resp. the second) zero valued element of b , $b = (X^*X)^{-1}X^*y$, and the signs of the first (resp. the second) elements of vectors b and X^*y are the same, in all cases $b^*X^*y = 0$ implies $b = 0$.

Alternatively for the second part of the proof, we can write

$$b^*X^*y = ((X^*X)^{-1}X^*y)^*X^*y = (X^*X)^{-1}(X^*y)^*I(X^*y) = 0.$$

Because $(X^*X)^{-1} > 0$, it implies $X^*y = 0$ and thus $b = 0$. □

Therefore, we set the null and the alternative hypothesis as

$$\begin{aligned} H_0 &: \beta = 0, \\ H_1 &: \beta \neq 0 \end{aligned}$$

and test the null hypothesis against the alternative hypothesis using the theorem 3.1. That is, we are testing whether the theoretical MAC is zero, i.e. whether ‘the modal vectors are different’. We compute the value of statistic

$$Z = \frac{1}{2s^2} b^* X^* X b,$$

where

$$s^2 = \frac{y^* y - b^* X^* y}{2S - 2},$$

and the p value as a complement of a cumulative distribution function of the F distribution with parameters 2 and $2S - 2$ evaluated at Z . If we obtain $s^2 = 0$ in practical calculation, we set the p value to zero.

As the result of the MM criterion we provide the complement of p value $1 - p$ of the statistical test together with the R_0^2 .

Practically, we have implemented the described procedure for all combinations of modal vectors selection from both variants of input data and computed values of $1 - p$ and R_0^2 arranged into two matrices. The matrix with R_0^2 values equals to MAC matrix and matrix of $1 - p$ values provides analogous information. In both matrices the values close to zero indicate different modal vectors and values of one the same vectors except a multiple (linearly dependent vectors). The difference is that the R_0^2 (MAC) value is not affected by the number of elements of modal vectors S , whereas the size S is included in the statistical test and thus expressed in $1 - p$. However, because of the involvement of S in $1 - p$ value and not in R_0^2 , the results of both may differ.

By expressing the relation for Z statistics using R_0^2 we obtain (for $R_0^2 \neq 1$)

$$Z = (S - 1) \frac{R_0^2}{1 - R_0^2}. \quad (5.6)$$

From (5.6) evidently, for R_0^2 close to 1, the value of $1 - p$ may be close to 1 as well, but also smaller for small number of sensors S ($1 - p$ is larger for larger S). Contrariwise, for R_0^2 close to 0, the value of $1 - p$ may be close to 0 as well, but also larger for large number of sensors S .

MM in general

Let us denote a modal vector of the first reference variant as v^1 and the modal vectors of the second variant as v_1^2, \dots, v_L^2 , where L is the number of measurements of the second variant of input data, $L \geq 1$.

We want to set a model, where modal vectors from the second variant are all a multiple (the same multiple) of the modal vector from the first variant, that is the reference variant. We proceed from complex relations

$$\begin{aligned} v_1^2 &= \beta v^1 + \varepsilon_1, \\ &\vdots \\ v_L^2 &= \beta v^1 + \varepsilon_L \end{aligned}$$

and write them with real and imaginary components, for every $s = 1, \dots, S$, where S is the number of modal vectors elements, as

$$\begin{aligned} v_{Re_s1}^2 + i v_{Im_s1}^2 &= (\beta_{Re} + i \beta_{Im}) (v_{Re_s}^1 + i v_{Im_s}^1) + \varepsilon_{Re_s1} + i \varepsilon_{Im_s1}, \\ &\vdots \\ v_{Re_sL}^2 + i v_{Im_sL}^2 &= (\beta_{Re} + i \beta_{Im}) (v_{Re_s}^1 + i v_{Im_s}^1) + \varepsilon_{Re_sL} + i \varepsilon_{Im_sL}. \end{aligned} \quad (5.7)$$

By writing the relationships (5.7) by real and imaginary components into twice as much equations and rearranging into a linear regression model with all components real, we get the MM model

$$\underbrace{\begin{bmatrix} v_{Re11}^2 \\ \vdots \\ v_{ReS1}^2 \\ v_{Im11}^2 \\ \vdots \\ v_{ImS1}^2 \\ \hline v_{Re1L}^2 \\ \vdots \\ v_{ReSL}^2 \\ v_{Im1L}^2 \\ \vdots \\ v_{ImSL}^2 \end{bmatrix}}_y = \underbrace{\begin{bmatrix} v_{Re1}^1 & -v_{Im1}^1 \\ \vdots & \vdots \\ v_{ReS}^1 & -v_{ImS}^1 \\ v_{Im1}^1 & v_{Re1}^1 \\ \vdots & \vdots \\ v_{ImS}^1 & v_{ReS}^1 \\ \hline v_{Re1}^1 & -v_{Im1}^1 \\ \vdots & \vdots \\ v_{ReS}^1 & -v_{ImS}^1 \\ v_{Im1}^1 & v_{Re1}^1 \\ \vdots & \vdots \\ v_{ImS}^1 & v_{ReS}^1 \end{bmatrix}}_X \underbrace{\begin{bmatrix} \beta_{Re} \\ \beta_{Im} \end{bmatrix}}_{\beta} + \underbrace{\begin{bmatrix} \varepsilon_{Re11} \\ \vdots \\ \varepsilon_{ReS1} \\ \varepsilon_{Im11} \\ \vdots \\ \varepsilon_{ImS1} \\ \hline \varepsilon_{Re1L} \\ \vdots \\ \varepsilon_{ReSL} \\ \varepsilon_{Im1L} \\ \vdots \\ \varepsilon_{ImSL} \end{bmatrix}}_{\varepsilon},$$

where y is a random vector, X a regression matrix of given numbers, β is an unknown vector of regression parameters and ε a vector of unknown random deviations, which we assume meet the regression assumptions (see the section 3.1.1).

As the first part of the MM criterion result, we compute the modified coefficient of determination R_0^2 according to (5.4), (5.5), with $R = 2LS$. For two variants once measured, when $L = 1$, we have $R_0^2 = MAC_{v^1, v^2}$, as we have shown in Theorem 5.1.

It can be also shown that the modified coefficient of determination R_0^2 is zero if and only if the estimated vector b is zero in the general MM model (the proof is analogous to the proof of the Theorem 5.2). Thus we also set the null and the

alternative hypothesis as

$$\begin{aligned} H_0 : \beta &= 0, \\ H_1 : \beta &\neq 0 \end{aligned}$$

and using the theorem 3.1, as the second part of the MM criterion, compute the p value of the statistical test. The test statistic in this case is given by

$$Z = \frac{1}{2s^2} b^* X^* X b,$$

where

$$s^2 = \frac{y^* y - b^* X^* y}{2SL - 2},$$

and the p value is obtained as a complement of a cumulative distribution function of the F distribution with parameters 2 and $2SL - 2$ evaluated at Z . If we obtain $s^2 = 0$ in practical calculation, we set again the p value to zero.

The described procedure is implemented for a comparison of all combinations of the modal vectors from the first reference variant and the second variant. However, it is assumed that in the input data, the modal matrices and the vectors of natural frequencies of individual measurements of the second variant consist of sorted corresponding modal vectors and frequencies. Within the second variant, the first modal vectors from all measurements are taken to approximate the first modal vector, all second modal vectors approximate the second modal vector etc. (with deflection shapes that means that the same sampling rate is assumed).

The results of the MM criterion are two matrices, one for R_0^2 values, one for p values. A value of one shows that all the vectors are just a scaled version of the first reference modal vector and a value of zero indicates the total opposite, in both MM resulting matrices. The $1 - p$ values are affected by the size of modal vectors S and the number of measurements of the second variant L and R^2 are not.

5.1.2 MMV – More Measurements and Variants

As a MMV model, we denote a linear regression model of ANOVA type for more variants, each more times measured, with no reference variant. The MMV model is not a direct generalization of MAC, we can not get exactly MAC values as a special case of MMV and also we proceed a bit differently compared to the MV criterion. In the MMV model the complex input data are decomposed into amplitude and phase components and the same type of linear regression model is used for logarithm of amplitudes and phases (without logarithm) independently.

As in the case of previous correlation criteria, we select arbitrarily a single modal vector for each variant and show the statistic, that can be then used for all combinations of modal vectors selection in the same way.

Suppose we have N variants, each L times measured, $N > 1$, $L > 1$ and modal vectors are composed of S generally complex elements (S channels were used in

measurements). Let us have one modal vector per each variant and measurement $v_1^1, \dots, v_L^1, \dots, v_L^N$. Different number of measurements for each variant could be considered as well, but for simplicity we focus on the same number of measurement L for each variant $n, n = 1, \dots, N$.

This time, we introduce a new vector of unknown values (a new variable) μ , on which the first modal vector is also dependent and is not the reference to other modal vectors. Also, we now initially consider that each modal vector element is a different multiple of another modal vector element (we are assuming linearly independent modal vector). For all $l = 1, \dots, L$ and $s = 1, \dots, S$ we consider relations

$$\begin{aligned} v_{sl}^1 &= \mu_s + \varepsilon_{sl}^1, \\ v_{sl}^2 &= \mu_s \beta_s^2 + \varepsilon_{sl}^2, \\ &\vdots \\ v_{sl}^N &= \mu_s \beta_s^N + \varepsilon_{sl}^N \end{aligned}$$

and write all complex numbers except errors ε' in an exponential form, with amplitude and phase components, as

$$\begin{aligned} v'_{A_{sl}}{}^1 e^{iv_{P_{sl}}{}^1} &= \mu'_{A_s} e^{i\mu_{P_s}} + \varepsilon_{sl}^1, \\ v'_{A_{sl}}{}^2 e^{iv_{P_{sl}}{}^2} &= \mu'_{A_s} \beta_{A_s}'^2 e^{i(\mu_{P_s} + \beta_{P_s}^2)} + \varepsilon_{sl}^2, \\ &\vdots \\ v'_{A_{sl}}{}^N e^{iv_{P_{sl}}{}^N} &= \mu'_{A_s} \beta_{A_s}'^N e^{i(\mu_{P_s} + \beta_{P_s}^N)} + \varepsilon_{sl}^N. \end{aligned} \tag{5.8}$$

We could have also decomposed all complex numbers into real and imaginary parts, however we would failed with the following linearisation procedure.

The system (5.8) determines a nonlinear regression model with all components real. Because we want a linear model, we apply a complex logarithm along with the introduction of a new vector of errors. For complex logarithm of the product of $z_1, z_2 \in \mathbb{C} \setminus \{0\}$ the relation

$$\ln(z_1 z_2) = \ln |z_1| + \ln |z_2| + i(\arg(z_1) + \arg(z_2))$$

holds true.

Denoting

$$\begin{aligned} v_{A_{sl}}^n &:= \ln v'_{A_{sl}}{}^n, & n = 1, \dots, N, \\ \beta_{A_{sl}}^n &:= \ln \beta'_{A_{sl}}{}^n, & n = 2, \dots, N, \\ \mu_{A_s} &:= \ln \mu'_{A_s}, \end{aligned}$$

and introducing new errors, we get from (5.8)

$$\begin{aligned} v_{A_{sl}}^1 &= \mu_{A_s} + \varepsilon_{A_{sl}}^1, & v_{P_{sl}}^1 &= \mu_{P_s} + \varepsilon_{P_{sl}}^1, \\ v_{A_{sl}}^2 &= \mu_{A_s} + \beta_{A_s}^2 + \varepsilon_{A_{sl}}^2, & v_{P_{sl}}^2 &= \mu_{P_s} + \beta_{P_s}^2 + \varepsilon_{P_{sl}}^2, \\ &\vdots & &\vdots \\ v_{A_{sl}}^N &= \mu_{A_s} + \beta_{A_s}^N + \varepsilon_{A_{sl}}^N. & v_{P_{sl}}^N &= \mu_{P_s} + \beta_{P_s}^N + \varepsilon_{P_{sl}}^N, \end{aligned} \tag{5.9}$$

for all $l = 1, \dots, L$ and $s = 1, \dots, S$. Errors in (5.9) denoted with the letters ε_A , ε_P are different errors beside amplitudes and phases of errors ε' .

We look at relations (5.9) as two regression models separately. From now on, we continue with formulas with A indexes only, because the formulas are the same for both amplitudes and phases.

Finally, for all $n = 2, \dots, N$ and for $s = 2, \dots, S$, we replace the unknowns $\beta_{A_s}^n$ with new unknown additions $\Delta_{A_s}^n$ in the model by introducing

$$\beta_{A_s}^n = \beta_{A_1}^n + \Delta_{A_s}^n.$$

The obtained MMV model for amplitudes is given by

$$\underbrace{\begin{bmatrix} v_{A_{11}}^1 \\ \vdots \\ v_{A_{1L}}^1 \\ \vdots \\ v_{A_{S1}}^1 \\ \vdots \\ v_{A_{SL}}^1 \\ \hline v_{A_{11}}^2 \\ \vdots \\ v_{A_{1L}}^2 \\ \vdots \\ v_{A_{S1}}^2 \\ \vdots \\ v_{A_{SL}}^2 \\ \hline \vdots \\ \hline v_{A_{11}}^N \\ \vdots \\ v_{A_{1L}}^N \\ \vdots \\ v_{A_{S1}}^N \\ \vdots \\ v_{A_{SL}}^N \end{bmatrix}}_y = \underbrace{\begin{bmatrix} 1 & & & & & & \\ \vdots & & & & & & \\ & 0 & & & & & \\ 1 & & & & & & \\ & \ddots & & & & & \\ & & & & & & \\ & & & 1 & & & \\ & & & \vdots & & & \\ & & & 1 & & & \\ \hline 1 & & 1 & & & & \\ \vdots & & \vdots & & 0 & & \\ & 0 & & 0 & & & \\ 1 & & 1 & & & & \\ & \ddots & \vdots & & & & \\ & & \vdots & & & & \\ & & 1 & & 0 & & \\ & & \vdots & & \vdots & & \\ & & 1 & & 1 & & \\ \hline \vdots & & 0 & & \ddots & & 0 \\ \hline 1 & & & & & 1 & \\ \vdots & & & & & \vdots & \\ & 0 & & & & 0 & \\ 1 & & & & & 1 & \\ & \ddots & & & & \vdots & \\ & & & & & \vdots & \\ & & & & & \vdots & \\ & & & & & 1 & \\ & & & & & \vdots & \\ & & & & & 0 & \\ & & & & & \vdots & \\ & & & & & 1 & \\ & & & & & 1 & \end{bmatrix}}_X \underbrace{\begin{bmatrix} \mu_{A_1} \\ \vdots \\ \mu_{A_S} \\ \hline \beta_{A_1}^2 \\ \Delta_{A_2}^2 \\ \vdots \\ \hline \Delta_{A_S}^2 \\ \vdots \\ \hline \beta_{A_1}^N \\ \Delta_{A_2}^N \\ \vdots \\ \hline \Delta_{A_S}^N \end{bmatrix}}_\beta + \varepsilon. \quad (5.10)$$

where y is a random vector, X a regression matrix of given numbers, β is an unknown vector of regression parameters and ε a vector of unknown random deviations, which we assume meet the regression assumptions (see the section 3.1.1). Using the model on real data, we examine the fulfilment of assumptions in more detail (see the section 5.4).

The MMV model (5.10) is a regression model of type of two-way ANOVA with interactions not fulfilling the reparametrization conditions. The MMV model is

in the form

$$\mathbf{v}_{A_{sl}}^n = \boldsymbol{\mu} + \boldsymbol{\alpha}_s + \boldsymbol{\beta}^n + \boldsymbol{\lambda}_s^n + \boldsymbol{\varepsilon}_{sl}^n,$$

where the parameter $\boldsymbol{\mu}$ is set to zero, then $\boldsymbol{\alpha}_s = \mu_{A_s}$ for all $s = 1, \dots, S$, the first parameter $\boldsymbol{\beta}^1$ is set to zero, $\boldsymbol{\beta}^n = \beta_{A_1}^n$ for $n = 2, \dots, N$, as well as the interactions $\boldsymbol{\lambda}_s^n$ are set to zeros for $s = 1 \wedge n = 1, \dots, N$ and $n = 1 \wedge s = 1, \dots, S$ and $\boldsymbol{\lambda}_s^n = \Delta_{A_s}^n$ for $s = 2, \dots, S$, $n = 2, \dots, N$.

The first main effect $\boldsymbol{\alpha}_s$ are the sensor positions and the second main effect $\boldsymbol{\beta}^n$ are the variants. In this parametrization of MMV model e.g. coefficients $\boldsymbol{\beta}^n = \beta_{A_1}^n$ for $n = 2, \dots, N$ estimate the differences of additions of other variants compared to the first variant. The MMV model can be rewritten into the standard parametrization of two-way ANOVA with interactions meeting the reparametrization conditions (see 3.1.2), where then e.g. $\boldsymbol{\beta}^n$, $n = 1, \dots, N$, represent the additions of individual variants.

The reason for rewriting all $\beta_{A_s}^n$, $s = 2, \dots, S$ for all $n = 2, \dots, N$ as the sum of the first $\beta_{A_1}^n$ and the addition $\Delta_{A_s}^n$ was that, in the MMV model, we want to test, whether all the additions $\Delta_{A_s}^n$, $s = 2, \dots, S$, $n = 2, \dots, N$, are zero valued. If they are, that means that all modal vectors are scaled versions of others, all elements of one modal vector are the same multiple of all elements of another modal vector and this holds true for modal vectors of all variants. If all the additions are not zero valued, there is at least one nonzero addition and therefore at least one significantly different (and linearly independent) modal vector.

The situation of all additions $\Delta_{A_s}^n$, $s = 2, \dots, S$, $n = 2, \dots, N$, zero valued in MMV model corresponds to the equality of all parameters $\boldsymbol{\lambda}_s^n$ in the classical ANOVA parametrization meeting the reparametrization conditions.

We set the null and the alternative hypothesis as

$$\begin{aligned} H_0 : \quad & \forall n \in \{2, \dots, N\} \wedge \forall s \in \{2, \dots, S\} : \quad \Delta_{A_s}^n = 0, \\ H_1 : \quad & \exists n \in \{2, \dots, N\} \vee \exists s \in \{2, \dots, S\} : \quad \Delta_{A_s}^n \neq 0 \end{aligned}$$

and test the null hypothesis, using the theorem 3.2 and calculating the p value of the statistical test. The test statistic is given by

$$Z = \frac{S_{eM} - S_e}{S_e} \frac{S(L-1)N}{(S-1)(N-1)},$$

where S_e is the residual sum of squares of the MMV model and S_{eM} is the residual sum of squares of a regression model without all the regressors $\Delta_{A_s}^n$, and it has the F distribution with parameters $(S-1)(N-1)$ and $S(L-1)N$. Analogously we do the same for the phases, separately. With that we are testing, whether ‘the modal vectors are the same’ (linearly dependent).

Notice two things in the MMV model. In the case of phases, the input values are angle values from the basic interval of length 2π . Whenever the calculation involves a difference of two angles, an angle greater than π may be taken into account instead of the correct angle value (even in parameter estimates). Therefore, when differently large angles appear at the input, incorrect statistics F may

be calculated and the null hypothesis H_0 incorrectly rejected. The second note is about the assumption on the normal distribution of errors ε that are different unknown errors compared to the original unknown errors ε' , because the complex natural logarithm and the transformation into polar coordinates was used.

The result of the MMV criterion are p values for amplitudes and also p values for phases. Combining these two results, the maximum of p values for amplitudes and for phases for each set of modal vectors is also returned as a result of this criterion. Recall also that p values are affected by the length of modal vectors (number of channels) S , by number of measurements L and number of variants N .

The resulting p value equals to one indicates that all the amplitudes or phases of modal vectors are the same except amplitude scaling and phase shift (all the modal vectors are mutually linearly dependent) and p value close to zero indicates the differences (and in case of phases, the value close to zero can be falsely diminished by the inclusion of angles greater than π in estimates).

When implementing, in the case of two variants more times measured, all modal vectors of the first variant with all the modal vectors of the second variant are compared, resulting in matrices of p values as the MMV criterion result. For more variants, a comparisons for all combinations of all variants can be theoretically done as well, only, it would be more difficult to display the results. Therefore, for more variants, only the n -th modal vectors of the all variants, $n \in N$, are compared, resulting in vectors of p values. Another option is to use pairing that is done with the reference first variant per rows of matrices of p values (for the first reference and other variants) sequentially. In case of pairing, firstly matrices of p values for pairs of the first and another variants are calculated. Then the maximum value from each row in every matrix of p values is selected and the corresponding frequencies and modal vectors are assigned to each other, resp. to the frequencies of the first reference variant (the principle of pairing is the same as in case of PLR with pairing per rows in the section 4.1.4). Finally, only the sets of assigned modal vectors are compared, resulting again in vectors of p values only.

In the input data, the modal matrices (and the vectors of natural frequencies) from all measurements within one variant are assumed all with sorted corresponding modal vectors (and frequencies). The first modal vectors from all measurements of the variant are taken to approximate the first modal vector of the variant etc. (which means the same sampling rate for deflection shapes in all measurements of the variant).

With the resulting matrices or vectors of p values also a matrix or vector of classical coefficients of determination R^2 and adjusted coefficients of determination R_a^2 (see the section 3.1.1) are calculated to indicate how well the MMV model globally approximates the data. Value close to one indicates a good approximation (the coefficient of determination R^2 is not related to the linear dependence and independence of modal vectors). The adjusted coefficients of determination R_a^2 is a scaled version of R^2 taking into account also the number of variants, measurements and number of modal vector elements in the case of MMV. It could be

worth considering a derivation of some kind of (square of) partial correlation coefficient as in case of R_0^2 in MM criterion, however we are not providing it.

5.1.3 MMV 2

MMV 2 model is the second model of the same type as MMV. The only difference compared to MMV is that extra parameters expressing the linear dependencies between the measurements of the same variant are included in the model. Thus MMV 2 is also suitable for more variants each more times measured, but unlike MMV, in MMV 2 we expect that the second end every other measurement of some variant differs from the first measurement by a multiple (i.e. results of measurements within one variant are scaled by a multiple).

The derivation of the MMV 2 model is similar to the MMV model.

Suppose again N variants, each L times measured, $N > 1$, $L > 1$, S generally complex modal vectors elements and one modal vector per each variant and measurement $v_1^1, \dots, v_L^1, \dots, v_L^N$. We introduce a new vector of parameters μ , on which the first modal vector is also dependent, as in the MMV model, and moreover a matrix of parameters γ , which expresses the linear relations between the measurements of the same variant.

For all $s = 1, \dots, S$ we consider

$$\begin{aligned}
 v_{s1}^1 &= \mu_s + \varepsilon_{s1}^{\prime 1}, \\
 v_{s2}^1 &= \mu_s \gamma_2^1 + \varepsilon_{s2}^{\prime 1}, \\
 &\vdots \\
 v_{sL}^1 &= \mu_s \gamma_L^1 + \varepsilon_{sL}^{\prime 1}, \\
 \\
 v_{s1}^2 &= \mu_s \beta_s^2 + \varepsilon_{s1}^{\prime 2}, \\
 v_{s2}^2 &= \mu_s \beta_s^2 \gamma_2^2 + \varepsilon_{s2}^{\prime 2}, \\
 &\vdots \\
 v_{sL}^2 &= \mu_s \beta_s^2 \gamma_L^2 + \varepsilon_{sL}^{\prime 2}, \\
 \\
 &\vdots \\
 \\
 v_{s1}^N &= \mu_s \beta_s^N + \varepsilon_{s1}^{\prime N}, \\
 v_{s2}^N &= \mu_s \beta_s^N \gamma_2^N + \varepsilon_{s2}^{\prime N}, \\
 &\vdots \\
 v_{sL}^N &= \mu_s \beta_s^N \gamma_L^N + \varepsilon_{sL}^{\prime N}.
 \end{aligned}$$

Then we proceed in exactly same way as with MMV and obtain two identical systems for amplitudes and phases. Continuing only with the system for amplitudes and introducing $\Delta_{A_s}^n$ with $\beta_{A_s}^n = \beta_{A_1}^n + \Delta_{A_s}^n$ for all $n = 2, \dots, N$ and

$s = 2, \dots, S$, we get the MMV 2 model in the form

$$\begin{array}{c}
 \left[\begin{array}{c} v_{A11}^1 \\ v_{A12}^1 \\ \vdots \\ v_{A1L}^1 \\ \vdots \\ v_{AS1}^1 \\ v_{AS2}^1 \\ \vdots \\ v_{ASL}^1 \\ \hline v_{A11}^2 \\ v_{A12}^2 \\ \vdots \\ v_{A1L}^2 \\ \vdots \\ v_{AS1}^2 \\ v_{AS2}^2 \\ \vdots \\ v_{ASL}^2 \\ \vdots \\ v_{A11}^N \\ v_{A12}^N \\ \vdots \\ v_{A1L}^N \\ \vdots \\ v_{AS1}^N \\ v_{AS2}^N \\ \vdots \\ v_{ASL}^N \end{array} \right] = \underbrace{\left[\begin{array}{c|c|c|c} 1 & & & \\ 1 & 0 & & \\ \vdots & \vdots & & \\ 1 & & 0 & \\ \vdots & \ddots & & \\ & & & 0 \\ 0 & 1 & & \\ \vdots & \vdots & & \\ 1 & & & \\ \hline 1 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & & 1 & \\ \vdots & \ddots & \vdots & \ddots \\ & & 0 & \\ 0 & 1 & 1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & & \\ \hline \vdots & & 0 & \ddots & 0 \\ \hline 1 & & & 1 & & & \\ 1 & 0 & & 1 & 0 & & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \ddots & \vdots \\ 1 & & 0 & 1 & & 0 & 0 & 1 \\ \vdots & \ddots & & \vdots & \ddots & & \vdots & \\ & & & 1 & 1 & & & \\ 0 & 1 & & 1 & 0 & 1 & & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \ddots & \\ 1 & & & 1 & 1 & & & 0 & 1 \end{array} \right]}_X + \varepsilon,
 \end{array}
 \tag{5.11}$$

where y is a random vector, X a regression matrix of given numbers, β is an unknown vector of regression parameters and ε a vector of unknown random deviations (different compared to amplitudes and phases of errors ε'), which we assume meet the regression assumptions (see the section 3.1.1).

The MMV 2 model (5.11) may be considered as model of type of three-way ANOVA with interactions, without introducing parameters λ^{AC}_{sl} , $\lambda^{ABC^n}_{sl}$ and using the parametrization

$$y_{sl}^n = \mu + \alpha_s + \beta^n + \gamma_l + \lambda^{AB^n}_s + \lambda^{BC^n}_l + \varepsilon_{sl}^n,$$

where μ , β^1 , $\lambda^{AB^n}_s$ for $s = 1 \wedge n = 1, \dots, N$ or $n = 1 \wedge s = 1, \dots, S$, γ_l for all $l = 1, \dots, L$ and $\lambda^{BC^n}_l$ for $l = 1$ and all $l = 1, \dots, L$, are set to zeros

and $\alpha_s = \mu_{A_s}$ for $s = 1, \dots, S$, $\beta^n = \beta_{A_1}^n$ for $n = 2, \dots, N$, $\lambda^{AB^n}_s = \Delta_{A_s}^n$ for $s = 2, \dots, S$, $n = 2, \dots, N$ and $\lambda^{BC^n}_l = \gamma_l^n$ for $l = 2, \dots, L$, $n = 1, \dots, N$.

The first main effect α_s are the sensor positions, the second main effect β^n are variants and the third main effects γ_l are measurements. The MMV 2 model is also possible to rewrite into the standard parametrization of three-way ANOVA with interactions (without introducing parameters λ^{AC}_{sl} , $\lambda^{ABC^n}_{sl}$) that meets the reparametrization conditions.

The null and the alternative hypothesis are again set as

$$\begin{aligned} H_0 : \quad & \forall n \in \{2, \dots, N\} \wedge \forall s \in \{2, \dots, S\} : \Delta_{A_s}^n = 0, \\ H_1 : \quad & \exists n \in \{2, \dots, N\} \vee \exists s \in \{2, \dots, S\} : \Delta_{A_s}^n \neq 0 \end{aligned}$$

and the null hypothesis (for phases separately analogously) tested using the theorem 3.2 with p values calculation. The test statistic is now given by

$$Z = \frac{S_{eM} - S_e}{S_e} \frac{(S-1)(L-1)N}{(S-1)(N-1)},$$

where S_e is the residual sum of squares of the MMV 2 model and S_{eM} is the residual sum of squares of a regression model without all the regressors $\Delta_{A_s}^n$, and it has the F distribution with parameters $(S-1)(N-1)$ and $(S-1)(L-1)N$. With the hypothesis H_0 and H_1 we test whether there is one multiple for all elements of the modal vectors for all variants and thus whether all ‘the modal vectors are the same’ (linearly dependent).

The result of the MMV 2 criterion are p values for amplitudes, p values for phases and the maximum of both, p values for amplitudes and p values of phases. For phases, the same issue with incorrect angles greater than π can occur as in MMV criterion.

The resulting p value equal to one indicates that all the (amplitudes or phases of) modal vectors are the same (except multiples) and p value close to zero indicates the differences. In the case of MMV 2, p values are affected by the number of channels S , the number of measurements L and the number of variants N .

A matrices of values are computed for all combinations of modal vectors for two variants of input data, but a vectors of values are computed for more variants (just because all combinations would be difficult to visualize). In case of more variants, only the n -th modal vector of every variant are compared, $n \in N$, or the pairing with the reference first variant can be done and then the sets of assigned modal vectors are compared (the same approach as in MMV).

It is again assumed that the input modal matrices (and the vectors of natural frequencies) from all measurements within one variant are all with sorted corresponding modal vectors (and frequencies).

Together with resulting p values the coefficients of determination R^2 and adjusted coefficients of determination R_a^2 are calculated as global indicators of how well the MMV 2 model approximates the data, values close to one indicate good approximations.

5.2 Scheffe tests

The MMV and MMV 2 regression models may be rewritten to classical ANOVA models meeting the reparametrization conditions and the Theorem 3.3 applied. The statistical testing is denoted as the Scheffe test and it intends to serve as a root cause analysis, attempting to capture which positions (at which variants, which frequencies and sensor specifications) lead to poor results in MMV and MMV 2. The selective approach (following the pairing with the reference first variant) could be also used to the data, similarly to 4.2). The only issue is that, when multiple data are processed, this approach is just more complicated as there exist more combinations of removals.

The Scheffe tests has not been implemented and practically applied on data, only the theoretical introduction follows.

For Scheffe tests we assume firstly computing the basic criteria MMV and MMV 2 for the paired modal vectors with the first reference variant.

5.2.1 SMMV - Scheffe test for MMV

In the MMV model (5.10) for one set of assigned modal vectors (for some reference frequency) we test whether all the interactions λ_s^n , for that the reparametrization conditions applies, are of the same value. That is, whether the increments of individual interactions for sensor positions and variants are all identical, whereas the parameter μ , the additions for sensor positions α_s and additions for variants β^n may be arbitrary and we are not interested in them. This type of testing (but with another test and model parametrization) we have already done within the MMV criterion. However, if the interactions λ_s^n differ, it would be advantageous to know where, for which variant and sensor position, within the reference frequency. Thus, it is convenient to apply the Theorem 3.3 that results in the decision about the differences for all pairs of interactions (the rejecting or failure of rejecting of the null hypothesis together with indication of the pairs of interactions with differences).

Let us denote $\psi = (\mu, \alpha_1, \dots, \alpha_S, \beta^1, \dots, \beta^N, \lambda_1^1, \lambda_2^1, \dots, \lambda_S^N)^*$ a vector of parameters of MMV model meeting the reparametrization conditions, φ a vector of corresponding estimates by the least square method and rename the elements of vectors ψ as ψ_p and φ as φ_p , $p = 1, \dots, P$, $P = SN + S + N + 1$.

Using the Theorem 3.3 we want to test the equality of all interactions, i.e. to test a null hypothesis

$$H_0 : \forall p, q \in M_1 : \psi_p = \psi_q,$$

against the alternative hypothesis

$$H_1 : \exists p, q \in M_1 : \psi_p \neq \psi_q,$$

where M_1 is a set of SN numbers,

$$M_1 := \{S + N + 2, \dots, P\}.$$

The estimate φ is given by a linear transformation

$$\varphi = Cb,$$

where b is the estimate from the regression model (5.10) and C is a matrix

$$\left[\begin{array}{cccc|ccc|cccc} \frac{1}{S} & \frac{1}{S} & \cdots & \frac{1}{S} & \frac{1}{N} & 0 & \cdots & 0 & \frac{1}{N} & 0 & \cdots & 0 \\ 1-\frac{1}{S} & -\frac{1}{S} & & -\frac{1}{S} & 0 & 0 & & 0 & 0 & 0 & & 0 \\ -\frac{1}{S} & 1-\frac{1}{S} & & -\frac{1}{S} & 0 & \frac{1}{N} & & 0 & 0 & \frac{1}{N} & & 0 \\ & & \ddots & & & & \ddots & & & & \ddots & \\ -\frac{1}{S} & -\frac{1}{S} & & 1-\frac{1}{S} & 0 & 0 & & \frac{1}{N} & 0 & 0 & & \frac{1}{N} \\ \hline & & & & -\frac{1}{N} & -\frac{1}{SN} & & -\frac{1}{SN} & -\frac{1}{N} & -\frac{1}{SN} & & -\frac{1}{SN} \\ & & & & 1-\frac{1}{N} & \frac{1}{S}-\frac{1}{SN} & & \frac{1}{S}-\frac{1}{SN} & -\frac{1}{N} & -\frac{1}{SN} & & -\frac{1}{SN} \\ & & & & & & \ddots & & & & \ddots & \\ & & & & -\frac{1}{N} & -\frac{1}{SN} & & -\frac{1}{SN} & \cdots & 1-\frac{1}{N} & \frac{1}{S}-\frac{1}{SN} & \frac{1}{S}-\frac{1}{SN} \\ \hline & & & & 0 & \frac{1}{SN} & & \frac{1}{SN} & 0 & \frac{1}{SN} & & \frac{1}{SN} \\ & & & & 0 & \frac{1}{SN}-\frac{1}{N} & & \frac{1}{SN} & 0 & \frac{1}{SN}-\frac{1}{N} & & \frac{1}{SN} \\ & & & & & & \ddots & & & & \ddots & \\ & & & & 0 & \frac{1}{SN} & & \frac{1}{SN}-\frac{1}{N} & 0 & \frac{1}{SN} & & \frac{1}{SN}-\frac{1}{N} \\ & & & & 0 & \frac{1}{SN}-\frac{1}{S} & & \frac{1}{SN}-\frac{1}{S} & 0 & \frac{1}{SN} & & \frac{1}{SN} \\ & & & & 0 & \frac{1}{SN}+1-\frac{1}{N}-\frac{1}{S} & & \frac{1}{SN}-\frac{1}{S} & 0 & \frac{1}{SN}-\frac{1}{N} & & \frac{1}{SN} \\ & & & & & & \ddots & & & & \ddots & \\ & & & & 0 & \frac{1}{SN} & & \frac{1}{SN}-\frac{1}{N} & 0 & \frac{1}{SN}-\frac{1}{S} & & \frac{1}{SN}+1-\frac{1}{N}-\frac{1}{S} \end{array} \right]$$

with rows written for estimates

$$\varphi = (\mu^\circ, |\alpha^{\circ 1}, \alpha^{\circ 2}, \dots, \alpha^{\circ S}, |\beta^{\circ 1}, \beta^{\circ 2}, \dots, \beta^{\circ N}, |\lambda^{\circ 1}_1, \lambda^{\circ 1}_2, \dots, \lambda^{\circ 1}_S, \lambda^{\circ 1}_2, \lambda^{\circ 2}_2, \dots, \lambda^{\circ N}_S)^*$$

and columns corresponding to estimates of regression parameters

$$\beta = (\mu_{A1}, \mu_{A2}, \dots, \mu_{AS}, |\beta_{A1}^2, \Delta_{A2}^2, \dots, \Delta_{AS}^2, |\dots, |\beta_{A1}^N, \Delta_{A2}^N, \dots, \Delta_{AS}^N)^*.$$

We know from the regression model, that $b \sim N(Eb, \text{var } b)$ with $Eb = \beta$ and $\text{var } b = \sigma^2(X^*X)^{-1}$, $\sigma^2 > 0$ (see the section 3.1.1). Because φ is derived from b by the linear transformation, it has a normal distribution $\varphi \sim N(E\varphi, \text{var } \varphi)$ as well and for the parameters, we get the relations

$$\begin{aligned} E\varphi &= C Eb, \\ \text{var } \varphi &= C(\text{var } b)C^*, \end{aligned}$$

(see Theorem 2 in [15, p. 27]). That is, $\varphi \sim N(\psi, \sigma^2 W)$, where $\psi = C\beta$, $W := C(X^*X)^{-1}C^*$. We have also $W > 0$, because $(X^*X)^{-1} > 0$ and by checking the matrix C , we get $CC^* > 0$ as well.

Let us denote $\nu := SLN - SN$. We have the estimate s^2 of σ^2 for the MMV regression model in the form

$$s^2 = \frac{y^*y - b^*X^*y}{\nu},$$

$\nu s^2/\sigma^2 \sim \chi_\nu^2$ for the MMV model. And because vector b and variable s^2 are independent (see 3.1.1), we have also independent φ and s^2 .

Because of $S + N - 1$ reparametrization conditions for interactions, we practically test the equality only for all $p, q \in M_2$, $p \neq q$, $M_2 \subset M_1$, where

$$\begin{aligned} M_2 &:= M_1 \setminus (M_3 \cup M_4), \\ M_3 &:= \{S + N + 1 + nS \quad \forall n = 1, \dots, N - 1\}, \\ M_4 &:= \{S + N + 1 + S(N - 1), \dots, P\}. \end{aligned}$$

We introduce vectors a_{pq} , $p, q \in M_1$, of sizes P . We set the vectors a_{pq} zero valued and

- add 1 in the p -th element if $p \in M_2$,
- add numbers -1 at positions $p - (S - 1)$ to $p - 1$ if $p \in M_3$,
- add numbers -1 at positions $p - (nS)$, $\forall n = 1, \dots, N - 1$, if $p \in M_4$,
- add -1 in the q -th element if $q \in M_2$,
- add numbers 1 at positions $p - (S - 1)$ to $p - 1$ if $q \in M_3$,
- add numbers 1 at positions $p - (nS)$, $\forall n = 1, \dots, N - 1$, if $p \in M_4$.

Then $\psi_p = \psi_q$ for some $p, q \in M_1$ applies if and only if $a_{pq}^* \psi = 0$. The set of vectors a_{pq} generates a subspace A of dimension $t = SN - S - N + 1$ in \mathbb{R}^P .

According to the Scheffe Theorem 3.3, the inequality

$$\|\varphi_p - \varphi_q\| \leq \sqrt{ts^2 F_\alpha(t, \nu) a_{pq}^* W a_{pq}}, \quad (5.12)$$

where $t = SN - S - N + 1$, $\nu = SLN - SN$, holds true for all pairs (p, q) , $p, q \in M_1$, simultaneously with a probability of at least $1 - \alpha$. If for some pair (p°, q°) the inverse inequality applies, i.e.

$$\|\varphi_{p^\circ} - \varphi_{q^\circ}\| > \sqrt{ts^2 F_\alpha(t, \nu) a_{p^\circ q^\circ}^* W a_{p^\circ q^\circ}},$$

the null hypothesis H_0 is rejected at the significance level α .

Practically it is more demanding to display the results of the statistical testing. For vector of paired modal vectors with the first reference variant, vector of results of testing at chosen significance level α (the rejecting or failure of rejecting of the null hypothesis H_0) may be taken as a result of SMMV criterion. When rejecting the null hypothesis H_0 , the fulfilment of the equality (5.12) for all pairs (p, q) , $p, q \in M_1$, may be displayed in a matrix.

5.2.2 SMMV 2 - Scheffe test for MMV 2

The process of applying the Scheffe Theorem 3.3 for the MMV 2 model (5.11) is the same as with MMV. We denote ψ a vector of parameters of MMV 2 model meeting the reparametrization conditions and φ a vector of estimates. The estimate φ may be obtained by linear transformation $\varphi = Cb$, where b is the estimate from the MMV 2 model (5.11) and C is a matrix given by

$$\begin{array}{|c|c|c|}
 \hline
 \frac{1}{S} & \cdots & 0 \\
 \hline
 1-\frac{1}{S} & & 0 \\
 -\frac{1}{S} & \cdots & 0 \\
 \hline
 -\frac{1}{S} & & \frac{1}{N} \\
 \hline
 0 & \cdots & -\frac{1}{SN} \\
 & & -\frac{1}{SN} \\
 & & \frac{1}{S}-\frac{1}{SN} \\
 \hline
 0 & \cdots & 0 \\
 \hline
 0 & & \frac{1}{SN} \\
 & \ddots & \\
 0 & & \frac{1}{SN}+1-\frac{1}{N}-\frac{1}{S} \\
 \hline
 0 & & \\
 & & \frac{1}{LN}-\frac{1}{L} \\
 & & \frac{1}{LN}+1-\frac{1}{L}-\frac{1}{N} \\
 & & \frac{1}{LN} \\
 & & \frac{1}{LN}-\frac{1}{L} \\
 & & \frac{1}{LN} \\
 & & \frac{1}{LN} \\
 & & \frac{1}{LN} \\
 & & \frac{1}{LN}-\frac{1}{N} \\
 & & \frac{1}{LN} \\
 & & \frac{1}{LN} \\
 & & \frac{1}{LN}-\frac{1}{L} \\
 & & \frac{1}{LN}+1-\frac{1}{L}-\frac{1}{N} \\
 \hline
 \end{array}
 \begin{array}{|c|c|c|}
 \hline
 0 \\
 \hline
 0 \\
 \hline
 0 \\
 \hline
 \frac{1}{L} & \cdots & \frac{1}{L} \\
 \frac{1}{L} & \cdots & \frac{1}{L} \\
 \vdots & & \\
 \frac{1}{L} & \cdots & \frac{1}{L} \\
 \hline
 -\frac{1}{LN} & & -\frac{1}{LN} \\
 & \ddots & \\
 -\frac{1}{LN} & & -\frac{1}{LN} \\
 \hline
 0 \\
 \hline
 0 \\
 & & \frac{1}{LN}-\frac{1}{L} \\
 & & \frac{1}{LN}-\frac{1}{L} \\
 & & \frac{1}{LN} \\
 & & \frac{1}{LN} \\
 & & \frac{1}{LN}+1-\frac{1}{L}-\frac{1}{N} \\
 & & \frac{1}{LN} \\
 & \cdots & \frac{1}{LN} \\
 & & \frac{1}{LN}-\frac{1}{N} \\
 & & \frac{1}{LN} \\
 & & \frac{1}{LN} \\
 & & \frac{1}{LN}-\frac{1}{L} \\
 & & \frac{1}{LN}+1-\frac{1}{L}-\frac{1}{N} \\
 \hline
 \end{array}
 \begin{array}{|c|c|}
 \hline
 0 \\
 \hline
 0 \\
 \hline
 0 \\
 \hline
 0 \\
 \hline
 -\frac{1}{LN} & -\frac{1}{LN} \\
 \hline
 -\frac{1}{LN} & -\frac{1}{LN} \\
 \hline
 0 \\
 \hline
 \frac{1}{LN} & \frac{1}{LN} \\
 \frac{1}{LN}-\frac{1}{N} & \frac{1}{LN} \\
 \frac{1}{LN} & \frac{1}{LN} \\
 \frac{1}{LN}-\frac{1}{N} & \frac{1}{LN}-\frac{1}{N} \\
 \frac{1}{LN} & \frac{1}{LN} \\
 \frac{1}{LN}-\frac{1}{N} & \frac{1}{LN} \\
 \frac{1}{LN} & \frac{1}{LN} \\
 \frac{1}{LN}-\frac{1}{L} & \frac{1}{LN}+1-\frac{1}{L}-\frac{1}{N} \\
 \hline
 \end{array}$$

with rows written for estimates

$$\varphi = (\mu^{\circ}, |\alpha^{\circ}_1, \alpha^{\circ}_2, \dots, \alpha^{\circ}_S, |\beta^{\circ 1}, \beta^{\circ 2}, \dots, \beta^{\circ N}, |\gamma_1, \dots, \gamma_L, |\lambda^{AB_{\circ 1}}, \dots, \lambda^{AB_{\circ S}}, | \lambda^{BC_{\circ 1}}, \lambda^{BC_{\circ 2}}, \dots, \lambda^{BC_{\circ L}}, \lambda^{BC_{\circ 1}^2}, \lambda^{BC_{\circ 2}^2}, \dots, \lambda^{BC_{\circ L}^N})^*$$

and columns corresponding to estimates of regression parameters

$$\beta = (\mu_{A_1}, \dots, \Delta_{A_S}^N, |\gamma_{A_2}^1, \dots, \gamma_{A_L}^1, |\dots, |\gamma_{A_2}^N, \dots, \gamma_{A_L}^N)^*$$

From the assumptions on the regression model implies that $\varphi \sim N(\psi, \sigma^2 W)$, $\psi = C\beta$, $W := C(X^*X)^{-1}C^* > 0$, $\sigma^2 > 0$. Denoting $\nu := SLN - N(S + L - 1)$ we obtain also $s^2 = (y^*y - b^*X^*y)/\nu$ and $\nu s^2/\sigma^2 \sim \chi_\nu^2$ with φ and s^2 independent.

In MMV 2 model, we want to test, if interactions λ^{AB} are all the same, i.e.

$$\begin{aligned} H_0 : \forall p, q \in M_1 : \quad \psi_p &= \psi_q, \\ H_1 : \exists p, q \in M_1 : \quad \psi_p &\neq \psi_q, \end{aligned}$$

where M_1 is a set of indexes of interactions λ^{AB} within the vector ψ . That is, we can proceed with the same introduction of vectors a_{pq} for all $p, q \in M_1$, with the only difference that they contain more zeros on indexes corresponding to variables γ and λ^{BC} . The set of vectors a_{pq} then again generates a subspace A of the dimension $t = SN - S - N + 1$.

According to the Scheffe Theorem 3.3, we can check, whether the inequality

$$\|\varphi_p - \varphi_q\| \leq \sqrt{ts^2 F_\alpha(t, \nu) a_{pq}^* W a_{pq}},$$

where $t = SN - S - N + 1$, $\nu = SLN - N(S + L - 1)$, holds true for all pairs (p, q) , $p, q \in M_1$, if not, the null hypothesis H_0 is rejected at the significance level α . Results can be displayed in the same way as SMMV.

5.3 Possible Extensions

The whole approach can be extended also for mass weighted version of criteria using the weighting matrices in regression models. The weighting coefficients may be estimated also from the statistical point of view, e.g. by incorporating the estimated sample variance (each row in the regression model may be multiplied by the inverse of the square root of the estimated variance). Using an estimated variance stabilizing transformation (so that the new variables have approximately the same variance) could be an another, similar, option [16, p. 115]. Then the approaches of MM, MMV, MMV 2, can be also applied. And we can also proceed a bit differently, with statistical testing for rank deficient ANOVA models instead of MMV and MMV 2 regression models [15, p. 135].

In addition to the introduced statistics, other known statistical tools may be used. As for example, for estimates of regression parameters β , also an interval estimate a confidence interval may be calculated [16, p. 94] or for the measurement data, influential data points (diverging or outlying data points) can be detected [16, p. 102] or the influence rate, e.g. a Cook distance, determined [16, p. 104].

Regarding the MMV, MMV 2 criteria and regression or test assumptions, we only look at the data by plotting it (see later), however, many issues may be also formally statistically tested (to help in decisions about the fulfilment of the assumptions). E.g. it is possible to test, whether the residuals as estimated errors (or the data itself) have the normal distribution [16, p. 68] (and there exist also normality tests implemented in Python, see module ‘scipy.stats’, functions ‘.normaltest’, ‘.shapiro’, ‘.kstest’ or ‘.anderson’). In the case of measurement data, it

is known, that the measurements are more accurate for lower frequencies than for the larger. Thus, for high frequency ranges non-fulfillment of homoskedasticity can be expected. We focus only on smaller frequency ranges, but homoskedasticity could be tested [16, p. 109] as well. Also, it can be tested whether the three-way ANOVA (or more general model) extra parameters are zero valued and the two-way ANOVA applicable. Using the three-way ANOVA model and criterion MMV 2 generally is not a good idea, if the extra MMV 2 parameters are actually zero valued, we obtain misleading results, because in the statistical tests a higher than the actual number of parameters is used.

5.4 Criteria applied on real data

We show the criteria MM, MMV and MMV 2 applied on two different data sets.

5.4.1 Data 5×2

The first data consist of measurements of 10 similar parts where each one was once measured. 5 of these parts contain a crack and we already know which does and which does not. The hammer impact testing was used to obtain the data, which is quite accurate but exacting type of (‘laboratory’) measurement.

Each data set is composed of a vector of frequencies from 0Hz to 3200Hz with a sampling rate 0.5Hz, a vector of 50 sensor ID specifications and a matrix of corresponding deflection shapes (frequency response).

The data have been already analysed. From each measurement the natural frequencies were detected and because the natural frequencies of the parts with crack are lower than those without crack, the parts were distinguished. The natural frequencies of all measurements are covered by the intervals given by the values listed in the Table 5.1.

mean	±	spread
575.3	±	0.7
580.3	±	1.
1408.4	±	2.9
1626.8	±	1.5
2330.7	±	3.8
2430.65	±	6.05
2741.9	±	11.4
3142.4	±	9.2

Table 5.1: Intervals covering the natural frequencies in Hz.

Because we need every variant in the MMV and MMV 2 criteria twice or more times measured, we look at these data as two variants (just two parts, one with

crack and the second one without crack) where each one is 5 times measured.

Firstly, we introduce the results from the MM criterion, then MMV and MMV 2 criteria. In all cases we cut the frequencies with limits ‘500 < frequencies < 3200’.

MM

At first, we applied the MM criterion on just two chosen data sets, one data set for a part with the crack (Variant A) and one without the crack (Variant B). The R_0^2 (MAC) and $1 - p$ values give us similar results, however, the $1 - p$ values are higher for most pairs of compared modal vectors, see Fig. 5.1 (the color bar values are rounded to tens in Fig. 5.1).

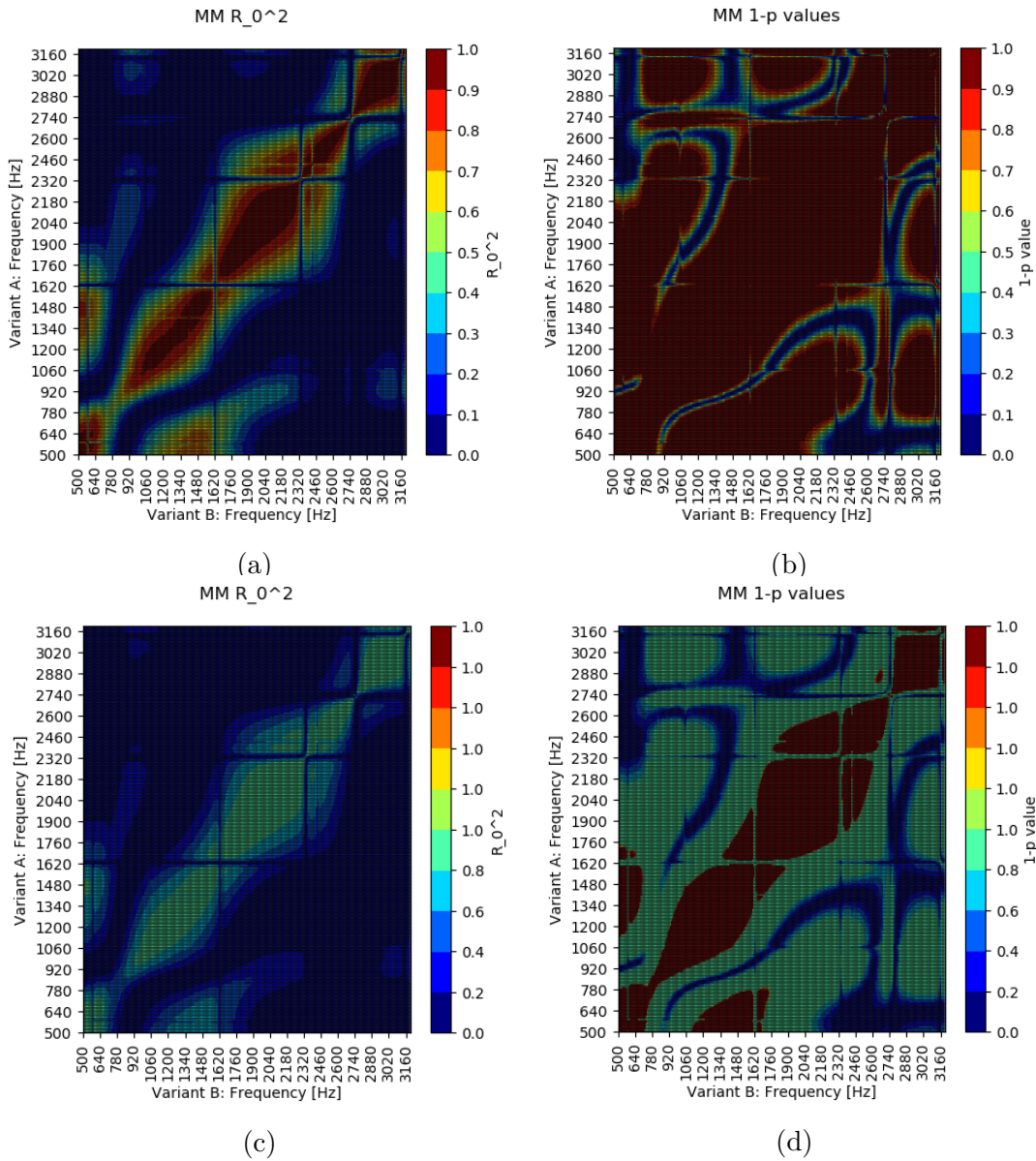


Figure 5.1: Result of MM criterion for 1&1 data, (a) R_0^2 , (b) $1 - p$ values, (c) R_0^2 with a high middle value close to 1, (d) $1 - p$ values with a high middle value close to 1.

We also obtained similar results when applied the MM criterion on 1 measurement of variant with the crack (Variant A) and the second variant 5 times measured without a crack (Variant B). The difference between obtained R_0^2 (MAC) and $1 - p$ values can be seen in the Fig. 5.2.

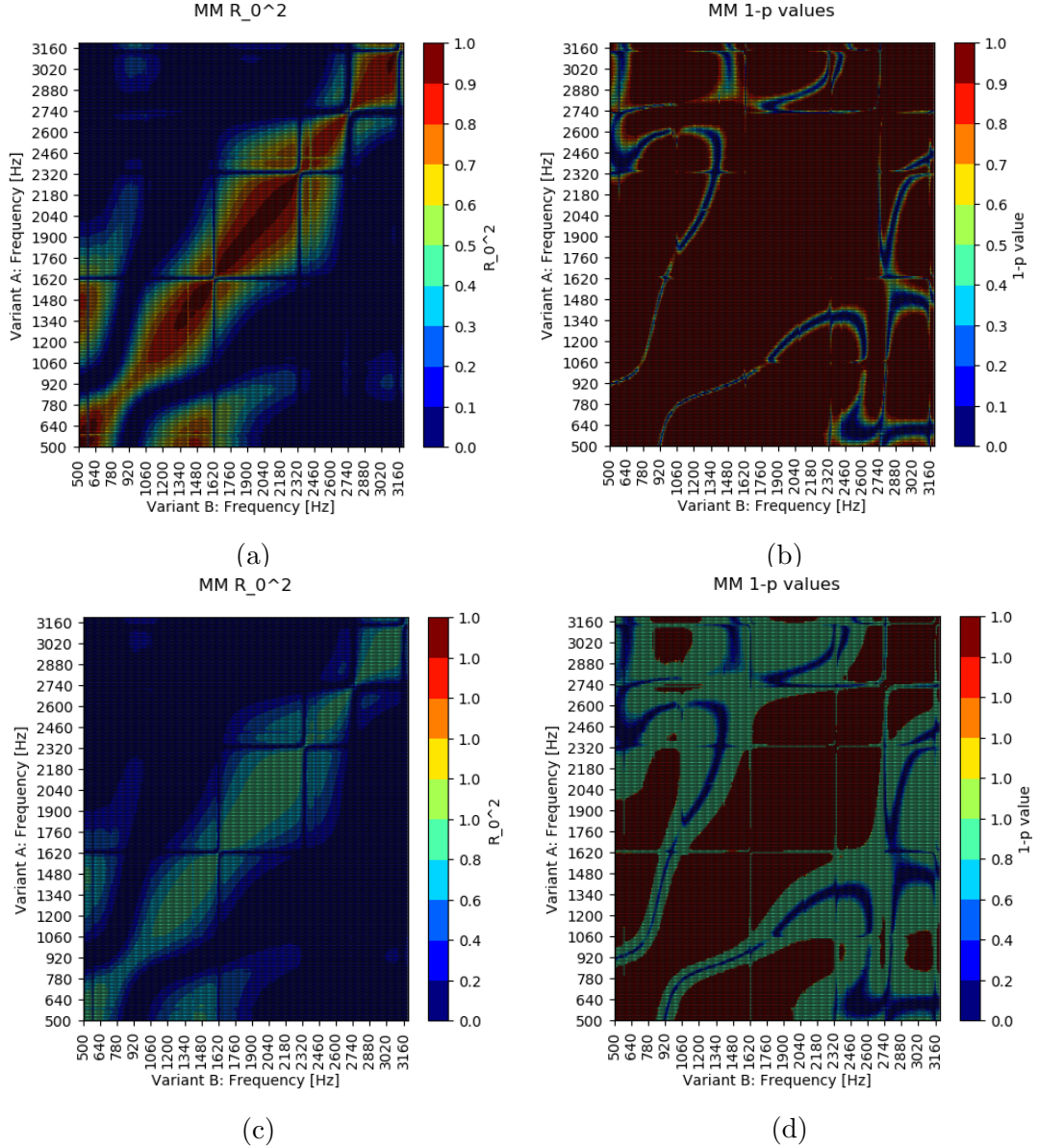


Figure 5.2: Result of MM criterion for 1&5 data (a) R_0^2 , (b) $1 - p$ values, (c) R_0^2 with and (d) $1 - p$ values with a high middle value close to 1.

MMV

As we have already said, we consider two variants of input data, one describing the part with crack and the second one without crack, each variant 5 times measured.

Let us firstly display the data itself. There is an example of frequency response from chosen sensors in Fig 5.3. In the whole data set, the curves for amplitudes

are quite similar, but differ for phases where the problem with the angle difference greater than π can be expected. We are also displaying all the data in a complex plane for one chosen frequency of 580.5Hz and the data distribution points out the same possible problem (see Fig 5.4).

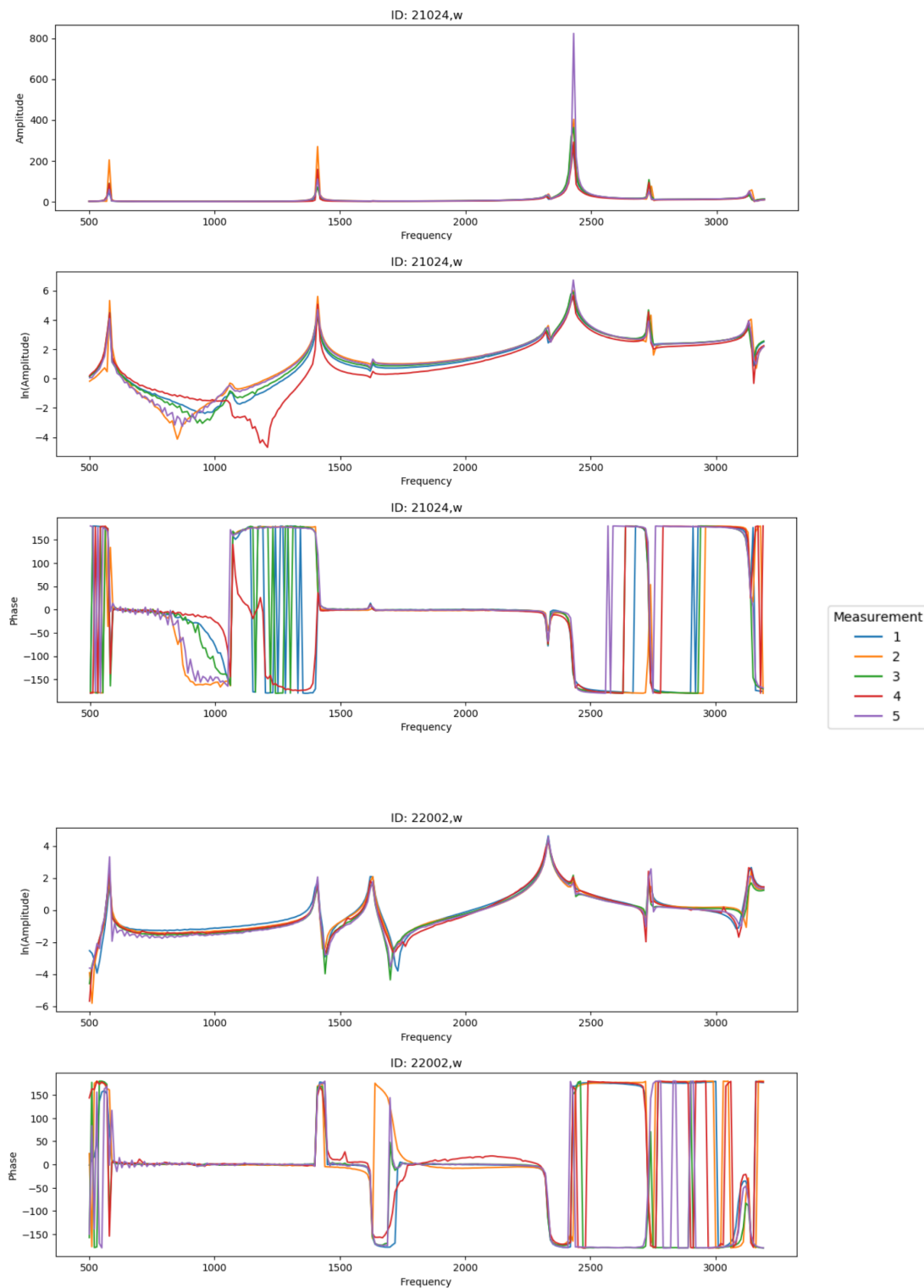


Figure 5.3: Frequency response from chosen sensors for all measurements of variant A.

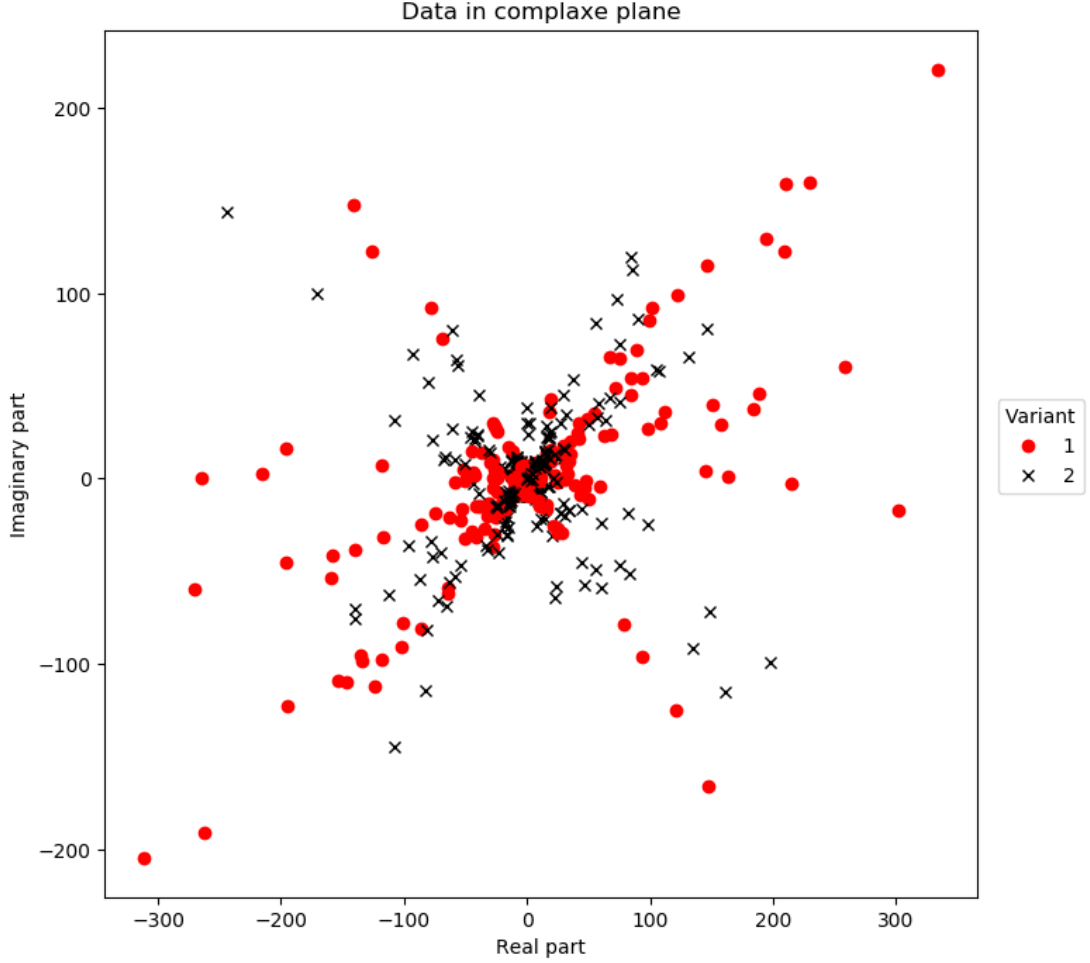


Figure 5.4: Data for fixed frequency 580.5Hz in a complex plane, all measurements and all variants.

For some frequencies, we plotted the natural logarithm of amplitudes and phases in degrees as a functions of number of measurements for all IDs and logarithmized amplitudes and phases as a functions of IDs for all measurements (see Fig. 5.5, 5.6), we call these graphs independency graphs. The situation was not entirely obvious for all frequencies, but in some cases, the observed data seemed to be dependent (the same may apply to the random deviations), which is inconsistent with the assumptions on the MMV regression model and can result in distorted conclusions. From the plotted (parts of) graphs we can e.g. observe, that some lines seems to (almost) be one equidistantly shifted from other.

As an example for the data dependence, we provide a standard correlation coefficient for pairs of vectors of logarithmized amplitudes, with vector elements corresponding to measurements, for the first ten sensors of the first variant and fixed frequency 580.5Hz in the Table 5.2. Some (nondiagonal) values are relatively high meaning relatively high correlation (and dependence) between the sensors information. The correlation coefficient r for vectors of logarithm of amplitudes m_1 and m_2 (values for all measurements from two sensors) is calculated as

$$r = \frac{m_1^* m_2 - L \overline{m_1^* m_2}}{\sqrt{(m_1^* m_1 - L \overline{m_1^* m_1}) (m_2^* m_2 - L \overline{m_2^* m_2})}} \quad (5.13)$$

with $L = 5$, the number of measurements, and with averages

$$\overline{m}_1 = \frac{1}{L} \sum_{l=1}^L m_{1l}, \quad \overline{m}_2 = \frac{1}{L} \sum_{l=1}^L m_{2l}.$$

Meas.	1	2	3	4	5	6	7	8	9	10
1	1.	0.81	0.78	0.82	0.74	0.75	-0.1	0.54	-0.23	-0.54
2	0.81	1.	0.99	1.	0.99	0.99	-0.5	0.81	0.09	-0.67
3	0.78	0.99	1.	0.99	1.	0.99	-0.51	0.83	0.07	-0.66
4	0.82	1.	0.99	1.	0.99	0.99	-0.5	0.81	0.06	-0.65
5	0.74	0.99	1.	0.99	1.	0.99	-0.52	0.81	0.14	-0.68
6	0.75	0.99	0.99	0.99	0.99	1.	-0.61	0.88	0.04	-0.58
7	-0.1	-0.5	-0.51	-0.5	-0.52	-0.61	1.	-0.84	0.15	-0.17
8	0.54	0.81	0.83	0.81	0.81	0.88	-0.84	1.	-0.3	-0.14
9	-0.23	0.09	0.07	0.06	0.14	0.04	0.15	-0.3	1.	-0.63
10	-0.54	-0.67	-0.66	-0.65	-0.68	-0.58	-0.17	-0.14	-0.63	1.

Table 5.2: Correlation coefficients for pairs of measurements within the first variant, corresponding to the Fig. 5.6, for frequency 580.5Hz.

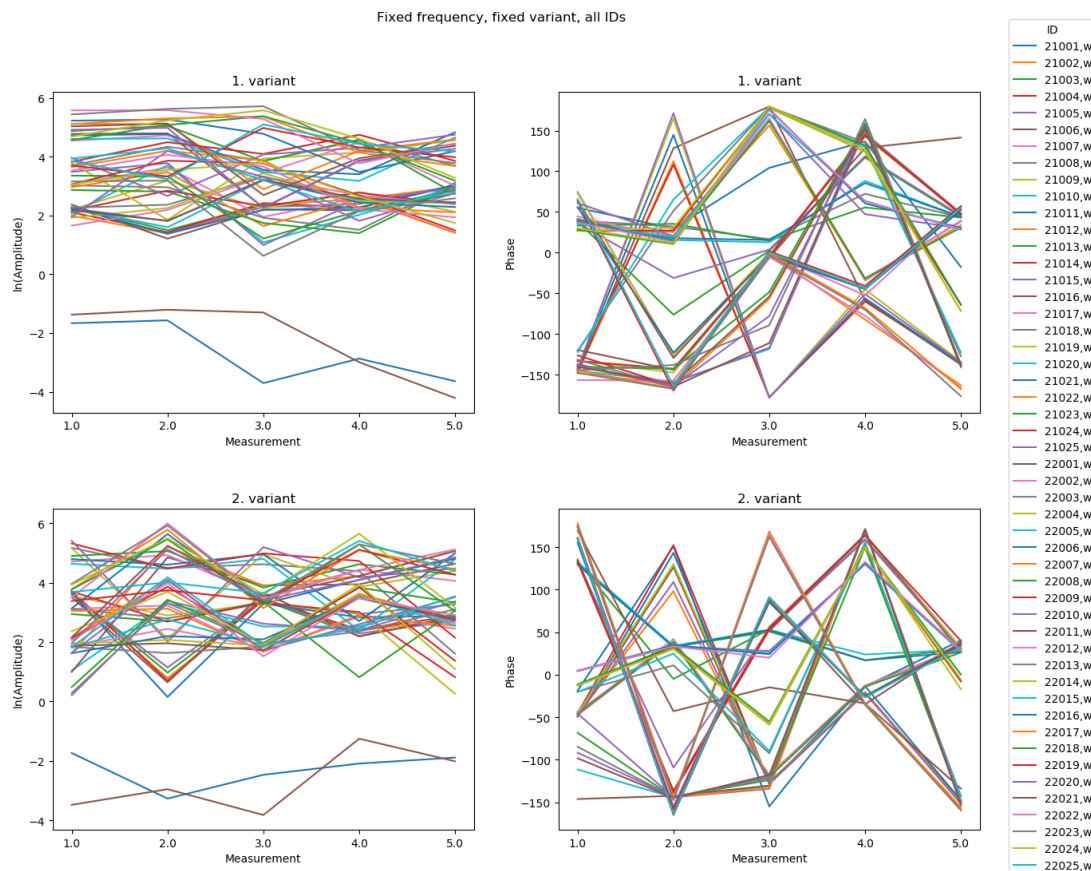


Figure 5.5: Independence graph for fixed frequency 580.5Hz.

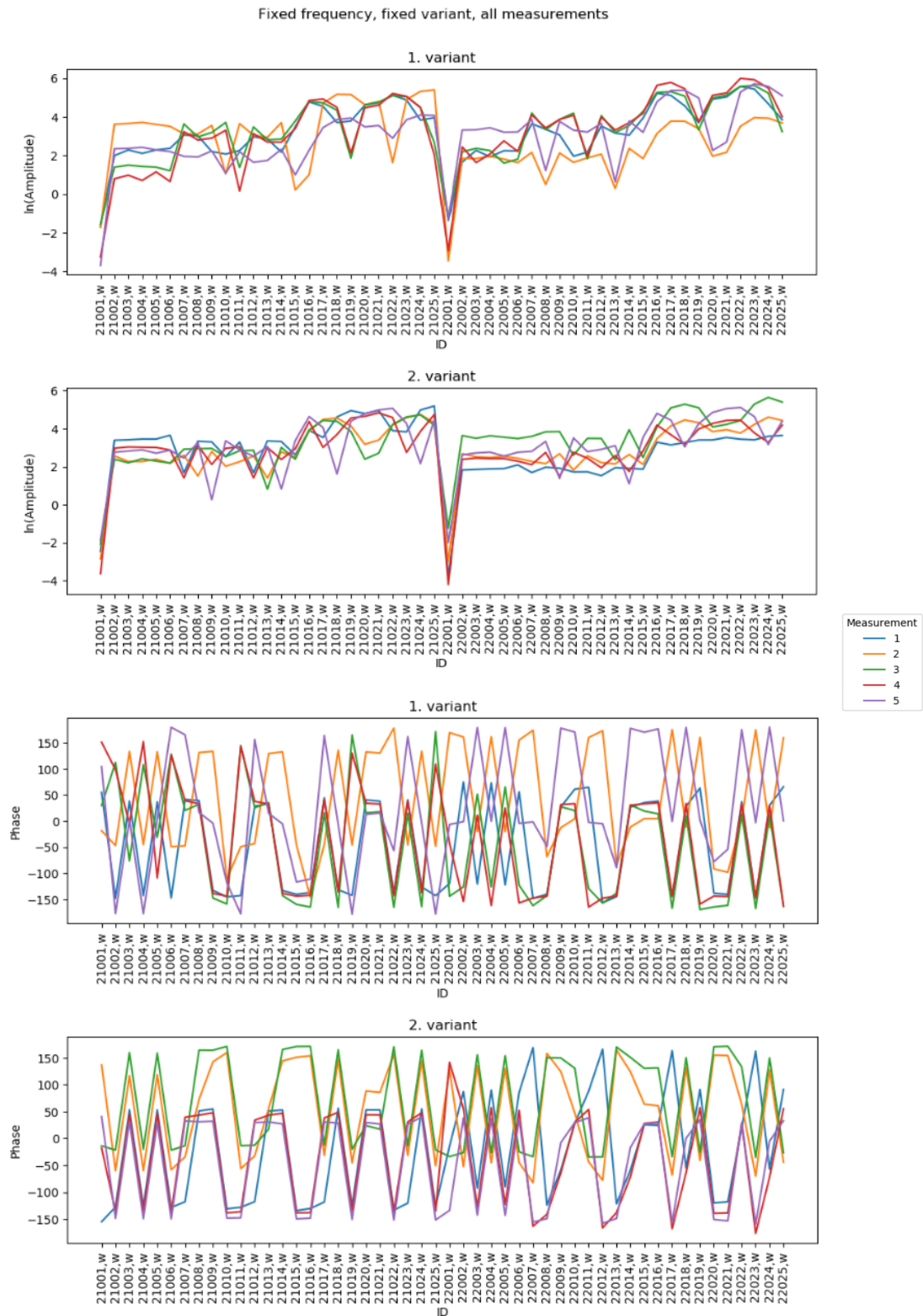


Figure 5.6: Independency graph for fixed frequency 580.5Hz.

After applying the MMV criterion we get the results displayed in Fig. 5.7, 5.9a with R^2 coefficient indicating the level of representativeness of the data by the MMV model in Fig. 5.8, 5.9b. From earlier analysis we got the result, that the parts with crack and without crack differ in areas of natural frequencies

2741.9 ± 11.4Hz and 3142.4 ± 9.2Hz and also slightly in region of 2330.7 ± 3.8Hz. We can see similar tendency in Fig. 5.7, in addition, there are low p values also in the area approximately 1900 – 2040Hz and the p values are close to one in the area around 2500Hz. The results for phases do not (entirely) match the previous results, also the maximum of estimated delta parameters was a value of cca 3.7π, which is irrelevant.

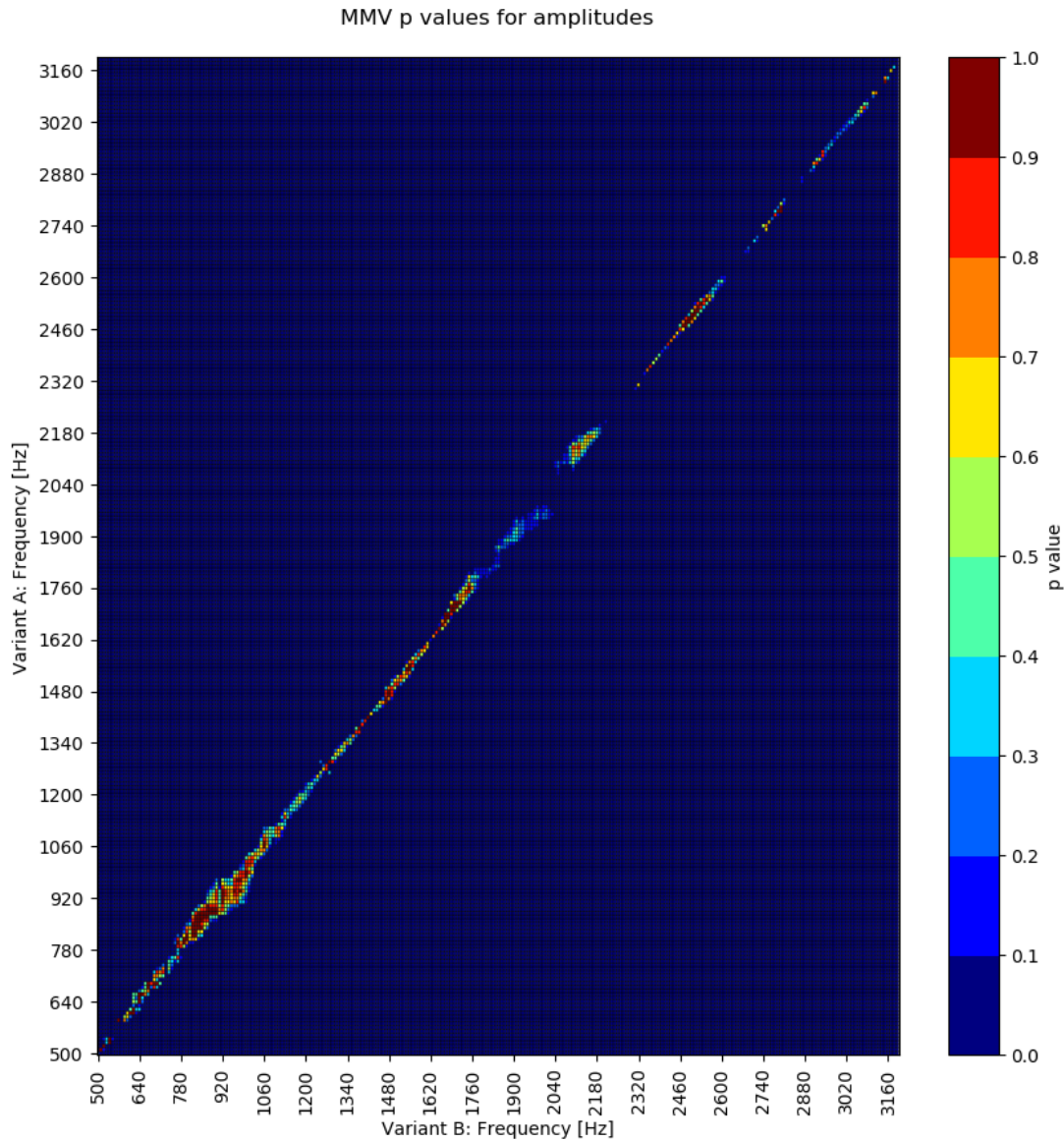


Figure 5.7: Result of MMV criterion for amplitudes.

When focusing only on the areas covering the natural frequencies and using the pairing of the most similar deflection shapes according to the reference first variant (Variant A with crack), we obtain the results in Fig. 5.10, where the differences of the parts with and without crack in areas of natural frequencies 2330.7 ± 3.8Hz, 2741.9 ± 11.4Hz and 3142.4 ± 9.2Hz are obvious. Also notice the lower value of R^2 for the frequency 580.5Hz (right below 581Hz) for which we plot the independency graphs (Fig. 5.5, 5.6) and listed the correlations (Table 5.2).

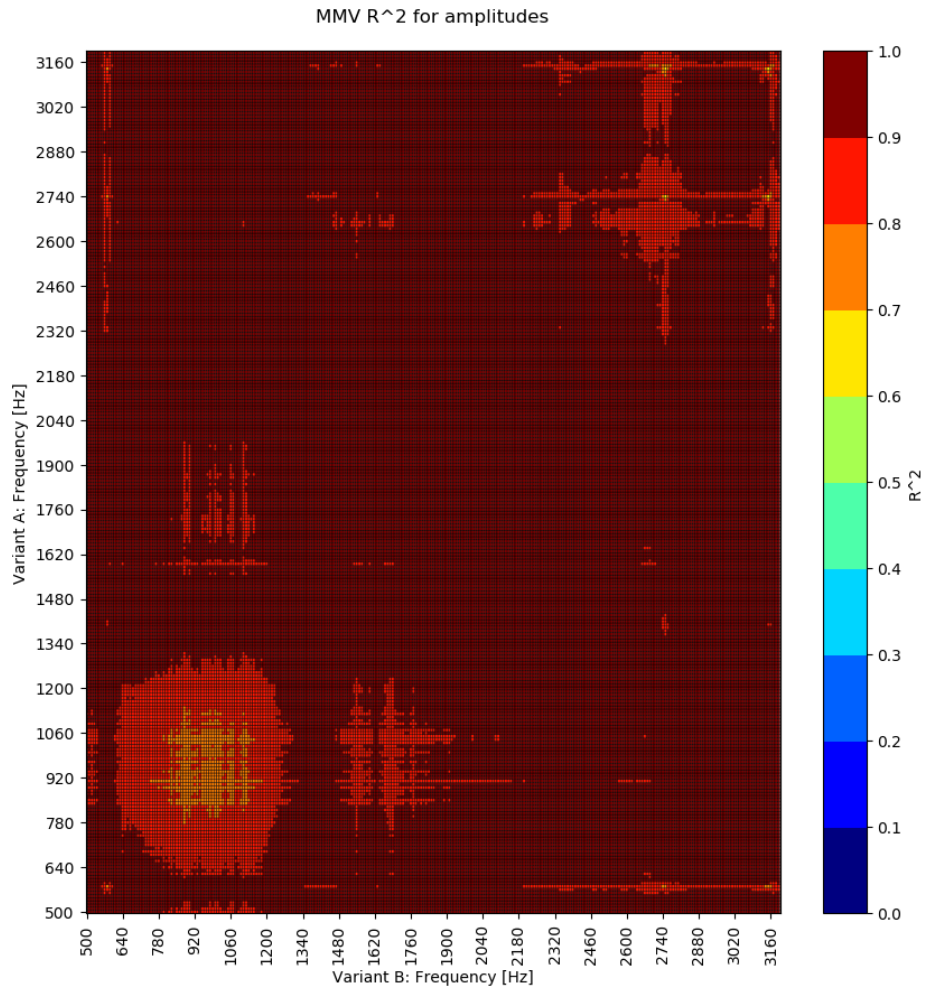


Figure 5.8: R^2 for amplitudes.

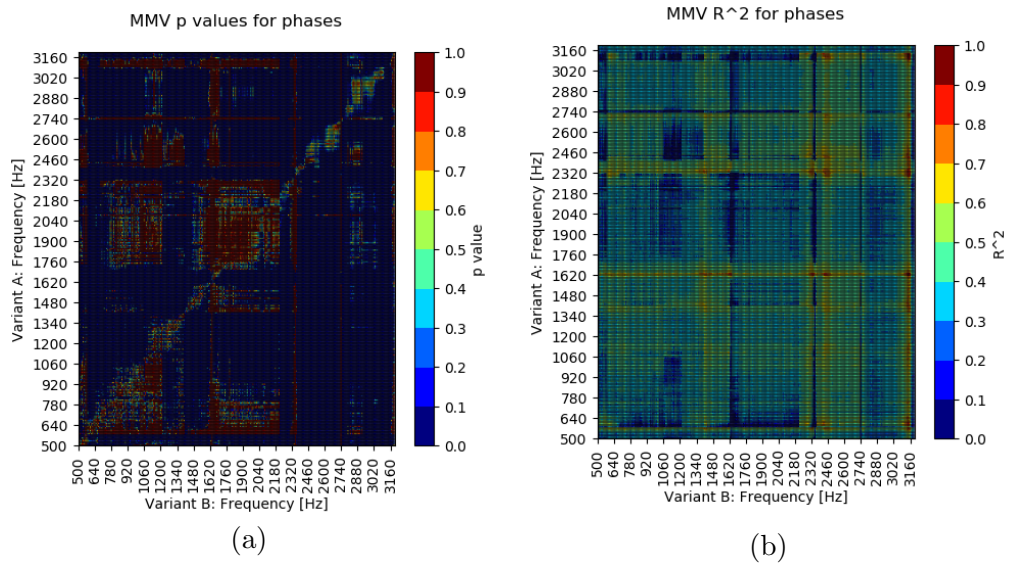


Figure 5.9: (a) Result of MMV criterion for phases, (b) R^2 for phases.

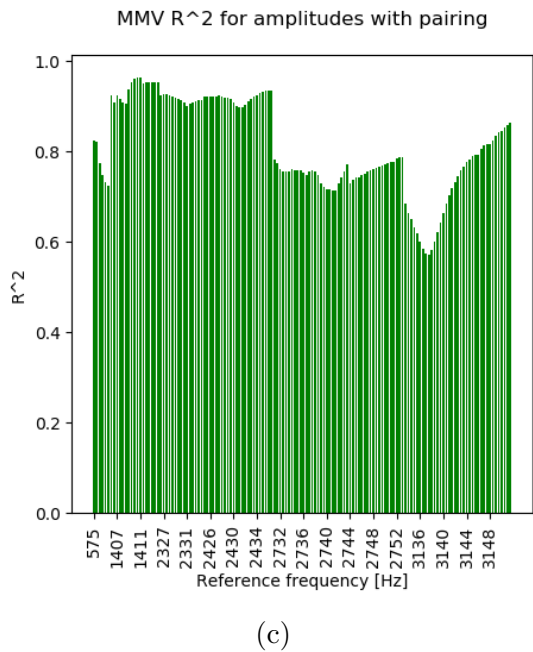
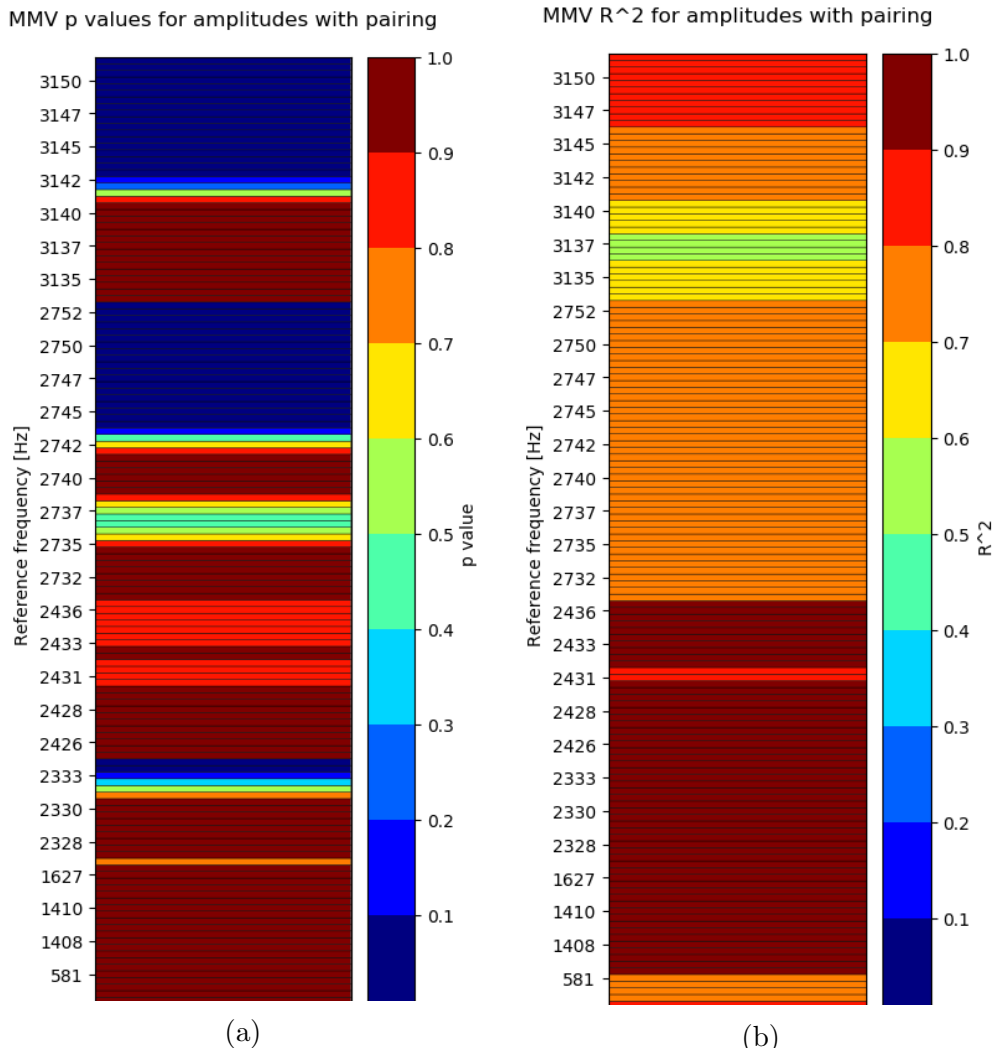


Figure 5.10: MMV criterion for amplitudes with pairing of the most similar deflection shapes, areas of the natural frequencies, (a) p values, (b) , (c) R^2 .

The results of the statistical testing at selected significance levels are shown in Fig. 5.11. For significance levels, let us say, 10–0.1%, the differences are detected and displayed in the expected areas of frequencies.

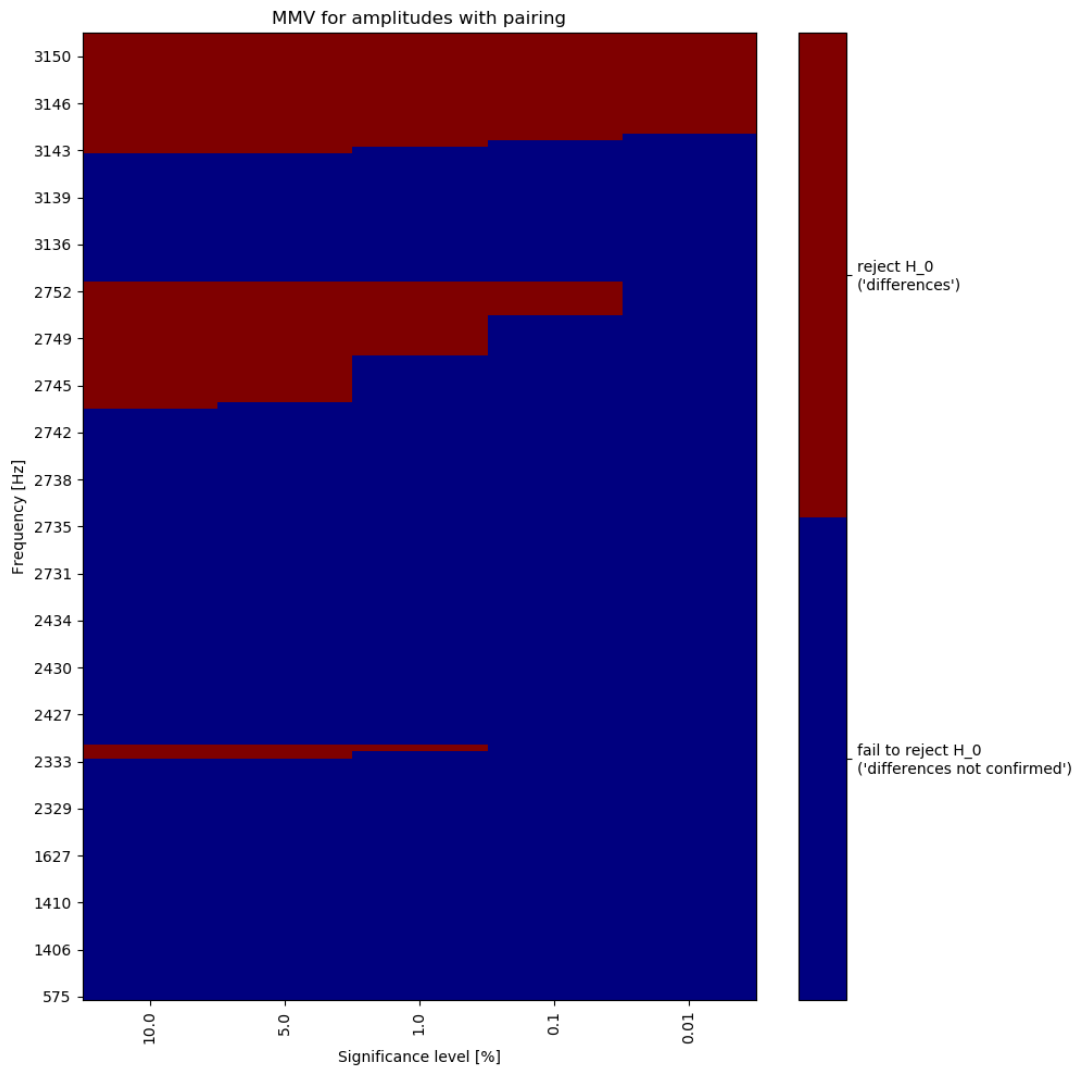


Figure 5.11: Results of statistical testing for amplitudes with pairing, areas of the natural frequencies.

To show how the criterion behaves for data we consider to be ‘the same’ (except the scaling coefficient for deflection shapes), we applied the MMV criterion also on data describing the parts with crack only. For this case we considered having two variants (both of them are with crack), where each variant is twice measured (we do not use the fifth measurement of the part with crack). Here follow the plots of p values and R^2 and results of statistical testing, all for amplitudes applied on paired deflection shapes only within the areas covering the natural frequencies (see Fig. 5.12).

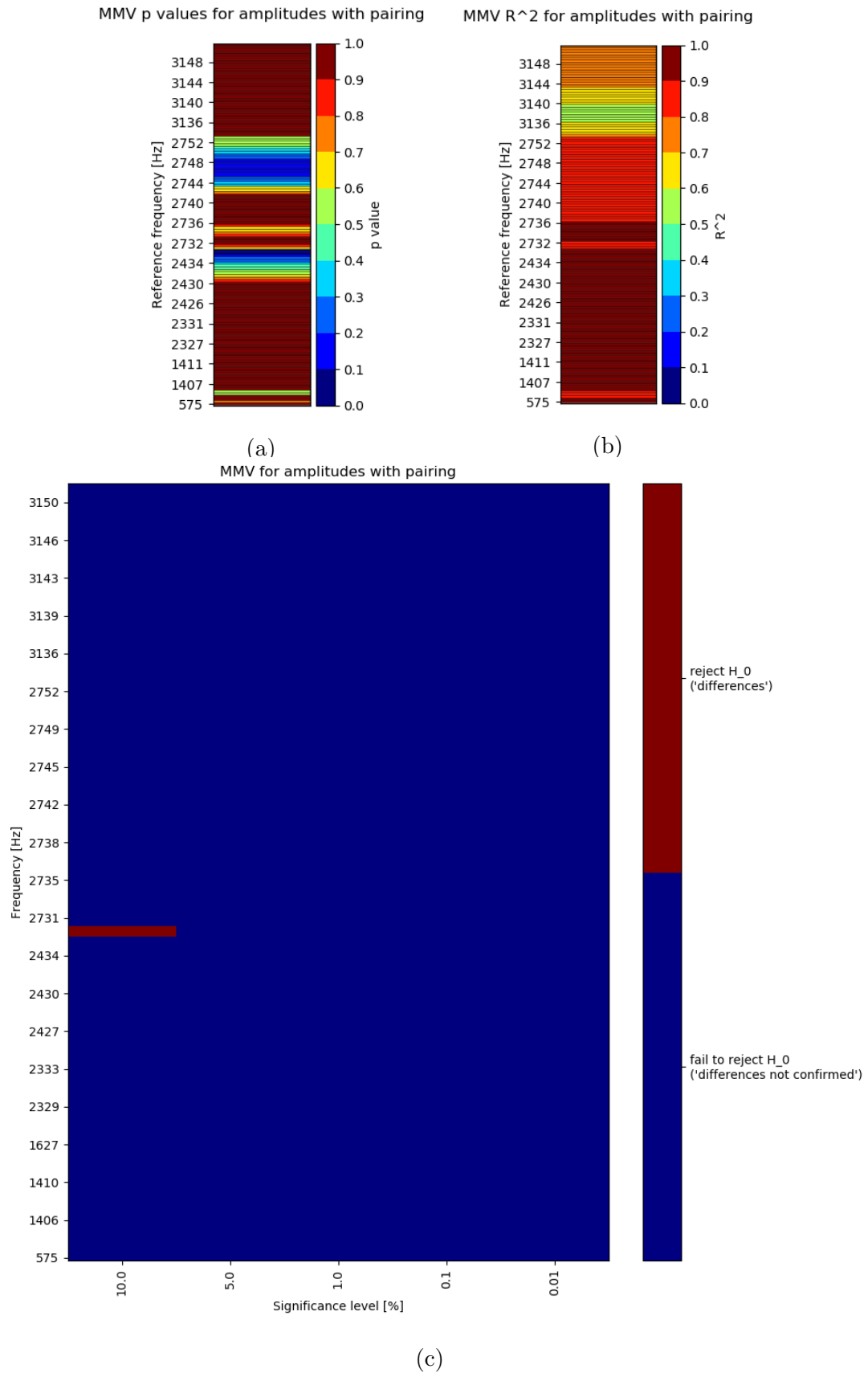


Figure 5.12: MMV criterion for amplitudes with pairing of the most similar deflection shapes, areas of the natural frequencies, data 2×2 of parts with crack only, (a) p values, (b) R^2 , (c) results of statistical testing.

MMV 2

Because the MMV model has actually sufficiently approximated the real data, applying the MMV 2 model on the data (with introduced parameters for scaled measurements within variants) led to similar results as MMV. Therefore, we just state here the MMV 2 p values with R^2 and results of statistical testing, for amplitudes with pairing for areas of the natural frequencies (see Fig. 5.13, 5.14 and compare to Fig. 5.10, 5.11).

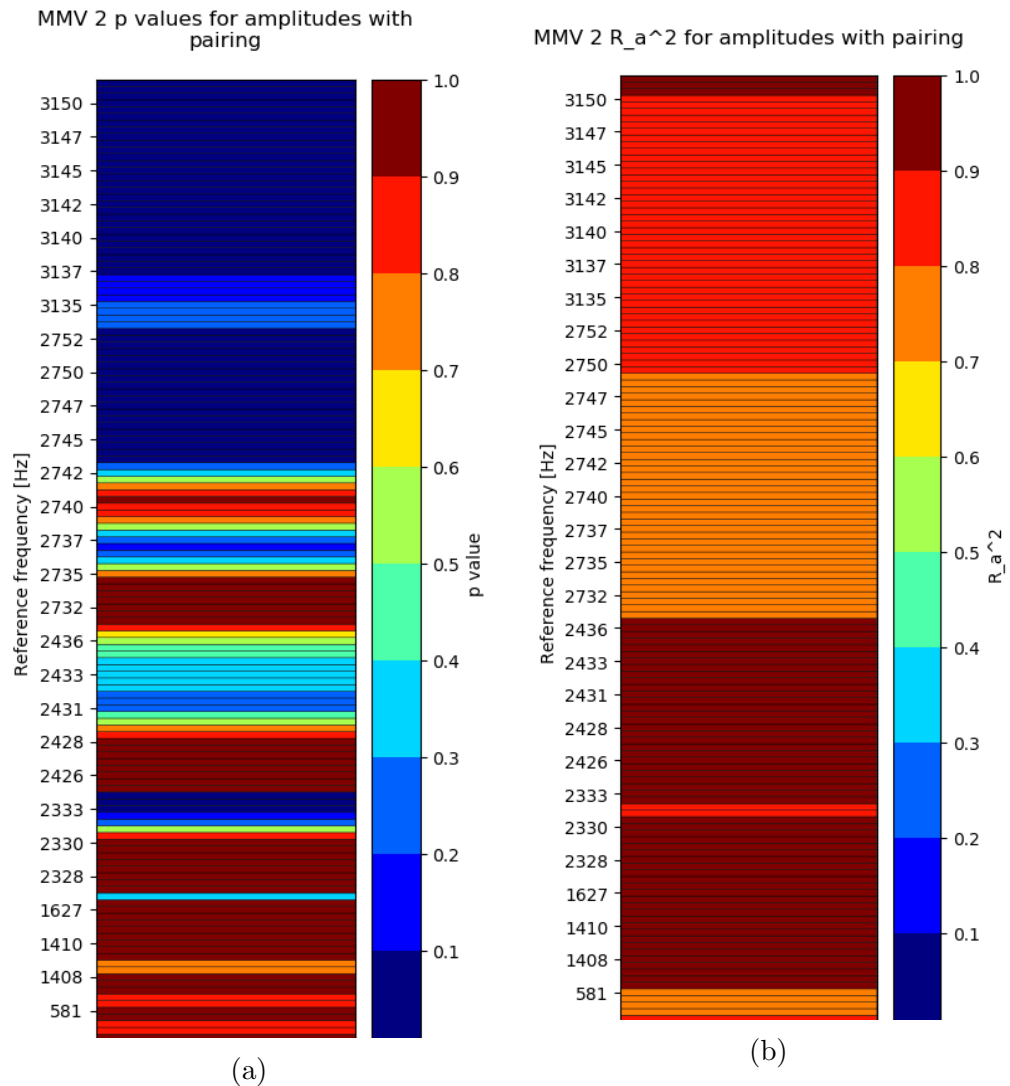


Figure 5.13: MMV 2 criterion for amplitudes with pairing of the most similar deflection shapes, areas of the natural frequencies, (a) p values, (b) , (c) R^2 .

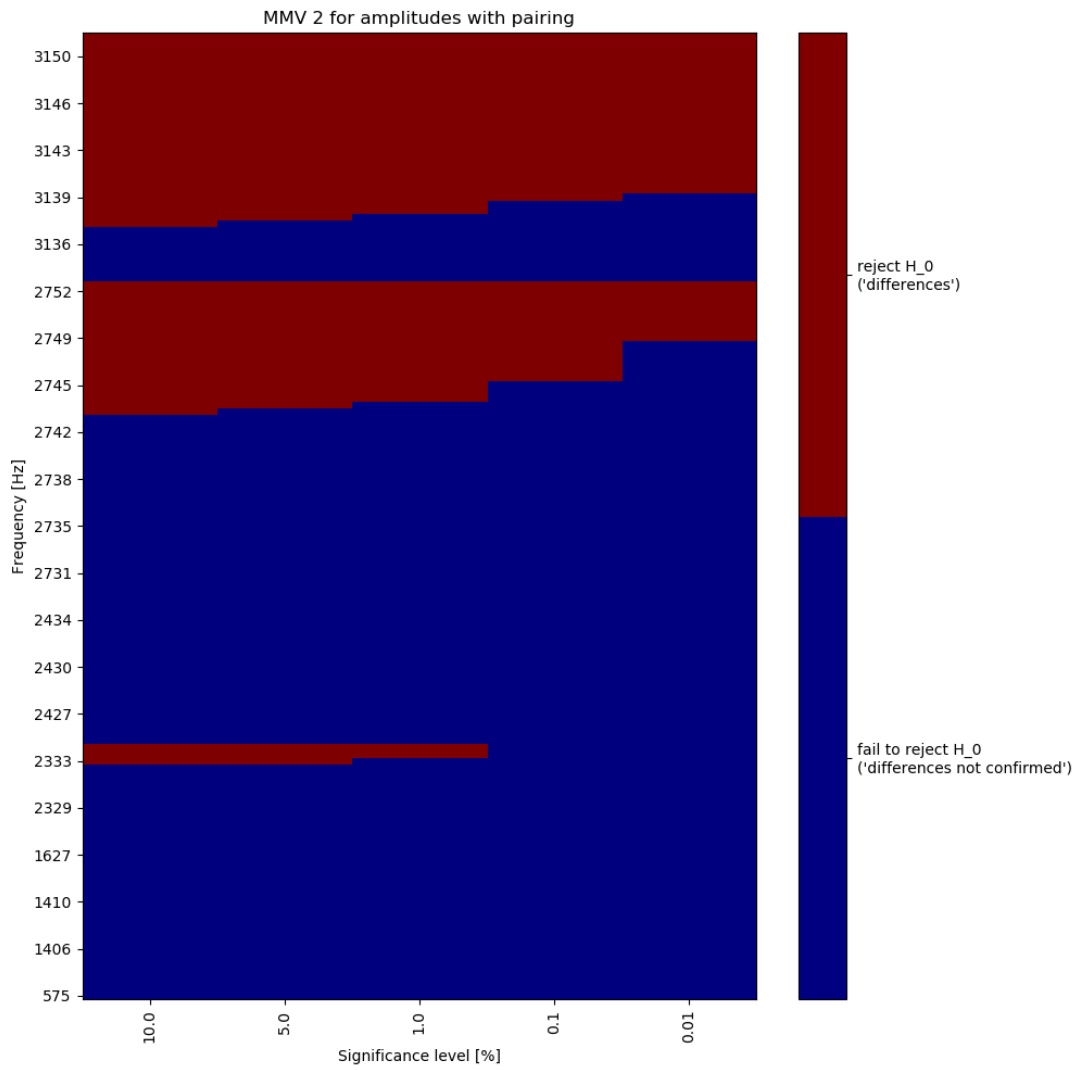


Figure 5.14: Results of statistical testing for amplitudes with pairing, areas of the natural frequencies.

5.4.2 Data 10×4

The second data, on which we show MMV and MMV 2 criteria, consist of 4 slightly different variants of 8HP transmission measurements, where each variant was 10 times measured. All of them were run up measurements on the test bench, that are frequent, but not very accurate measurements, with a relatively high error rate.

Each data set is composed of a vector of 64 frequencies from 500Hz to 2000Hz, a vector of 9 sensor ID specifications and a matrix of corresponding deflection shapes with size of 9×64 (the modal matrix ‘ V ’ has 9 rows and 64 columns).

MMV

At first, we display the data. There is an example of the frequency response from chosen sensors in Fig. 5.15. Although the phases appear to meet all assumptions of the MMV regression model, without expecting the problem with the angular difference, the phases of the data do not provide us with any information. They are averaged values where all of them are close to $\frac{\pi}{4}$, see also Fig. 5.16 of data for fixed frequency 703.125Hz in a complex plane, and we do not include them in further comparisons.

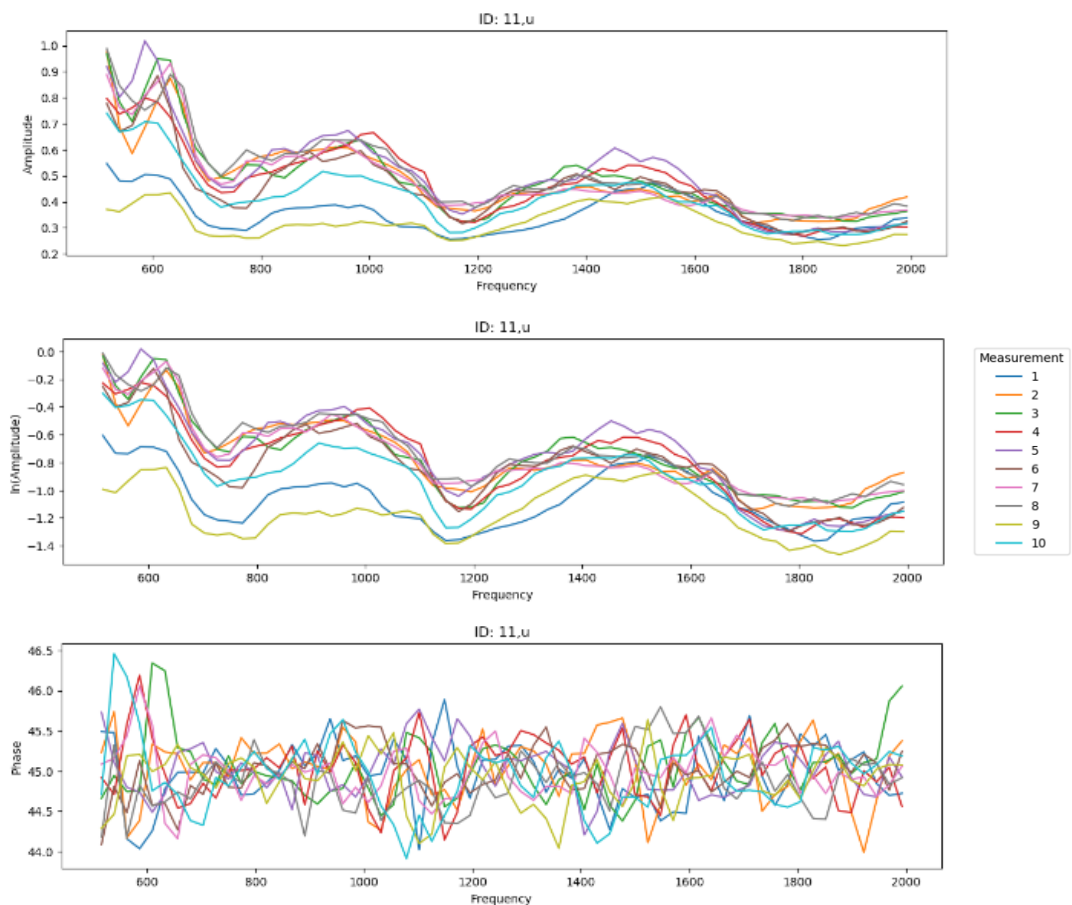


Figure 5.15: Frequency response, sensor 11, u , for all measurements of variant A.

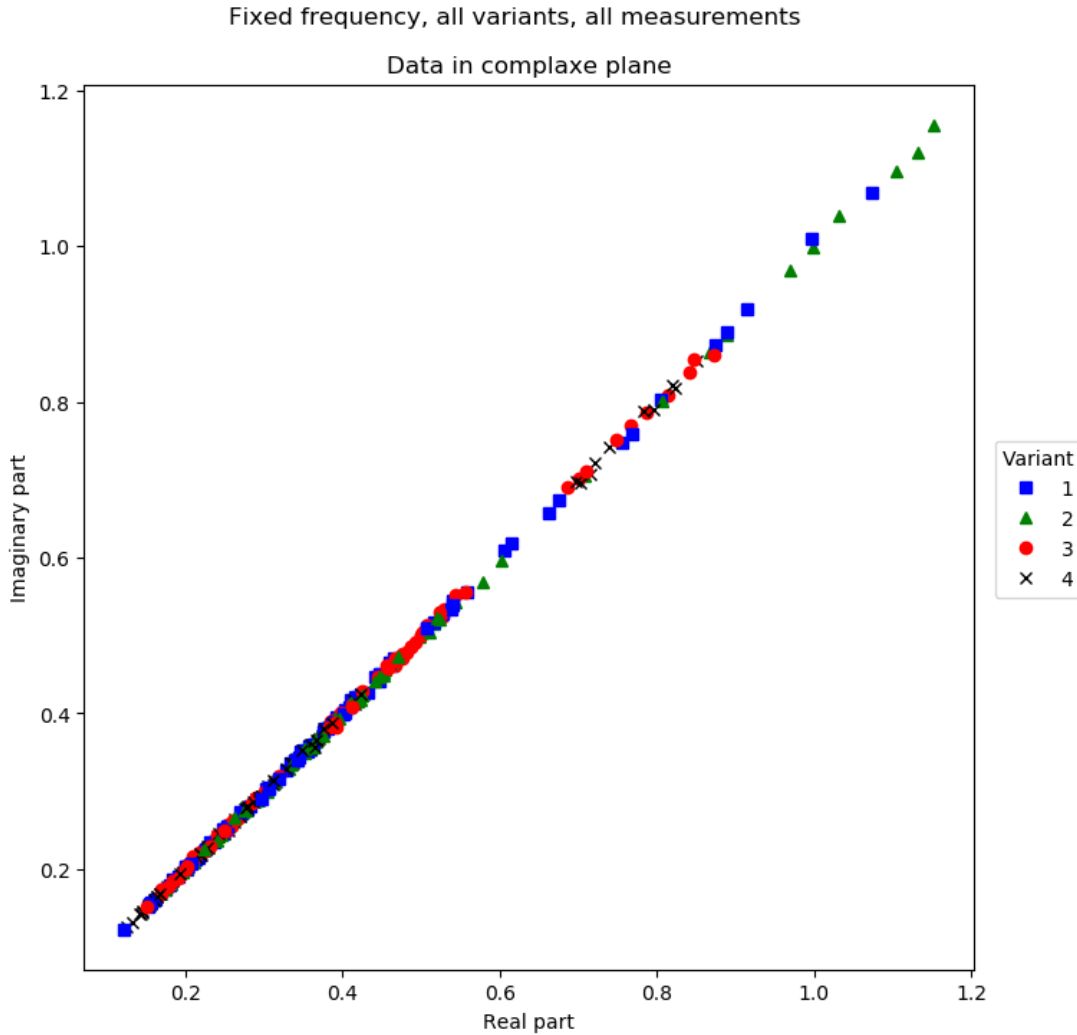


Figure 5.16: Data for fixed frequency 703.125Hz in a complex plane, all measurements and all variants.

We plotted the independency graphs as well and with these data, it was quite obvious that the MMV 2 model will be significantly more suitable to represent the data compared to MMV. As an example, we show the independency graphs only for the first and the second variant, for the fixed frequency 703.125Hz (see Fig. 5.17). Individual curves within both of these graphs appear to be all a scaled version of the others (and this is the reason for using the MMV 2 model including also the scaling parameter).

All values of the standard correlation coefficient for pairs of logarithmized amplitudes of individual measurements of the first variant, calculated again according to (5.13), ranges between 0.98 and 1.

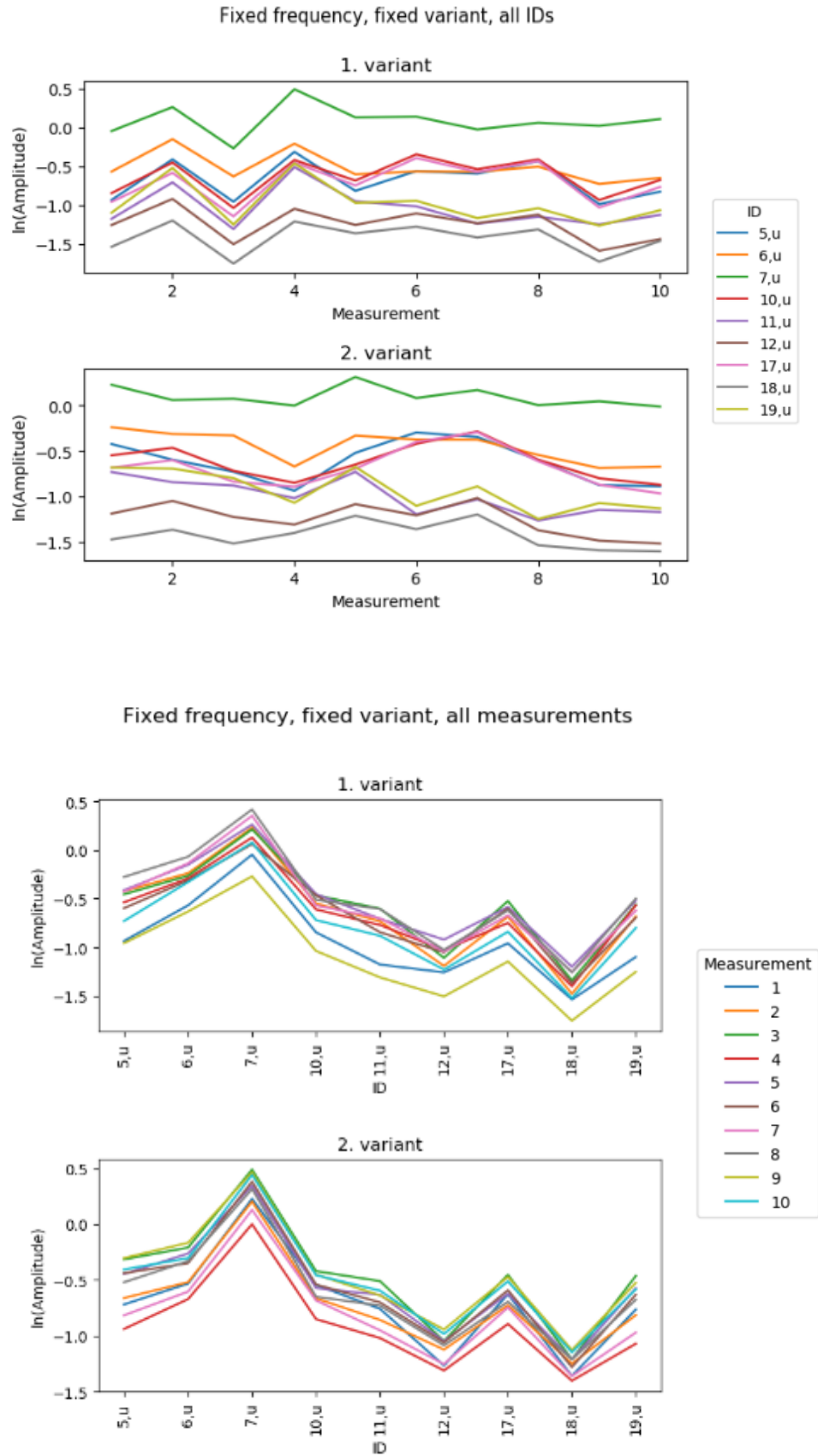


Figure 5.17: Independency graph for fixed frequency 703.125Hz.

We present a graph of data of all variants and measurements as well as the estimated data, that are obtained from the estimated regression parameters of the MMV model, for the fixed reference frequency 703.125Hz, see Fig. 5.18 (however, the corresponding paired deflection shapes are almost the same as deflection shapes all for the frequency 703.125Hz, which can be also seen comparing Fig. 5.18 and Fig. 5.17). Notice that the estimated data of variants 3 and 4 do not adequately capture the real data shapes. Introducing more parameters in MMV 2 model (additionally) improved the result, compare the p values for MMV and the MMV 2 criterion in Fig. 5.19.

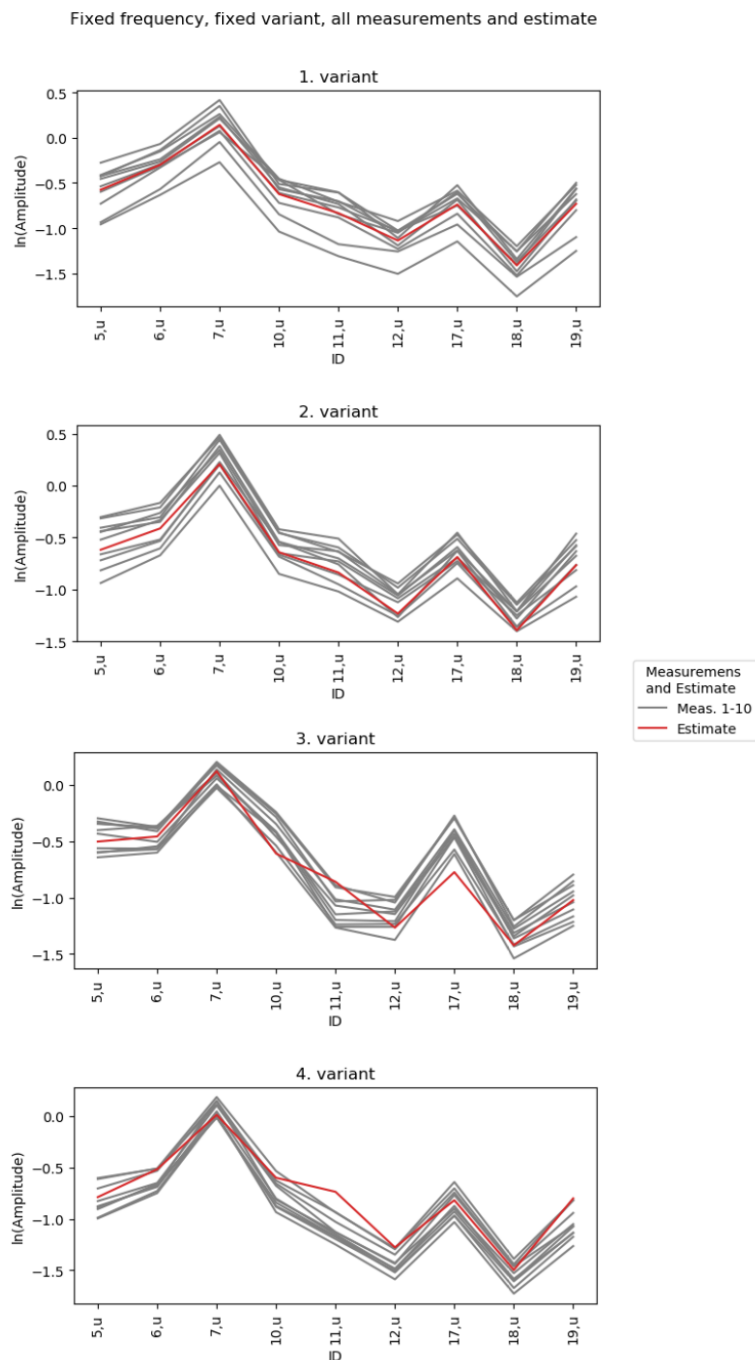


Figure 5.18: Graph of the data with MMV estimated data for fixed reference frequency 703.125Hz.

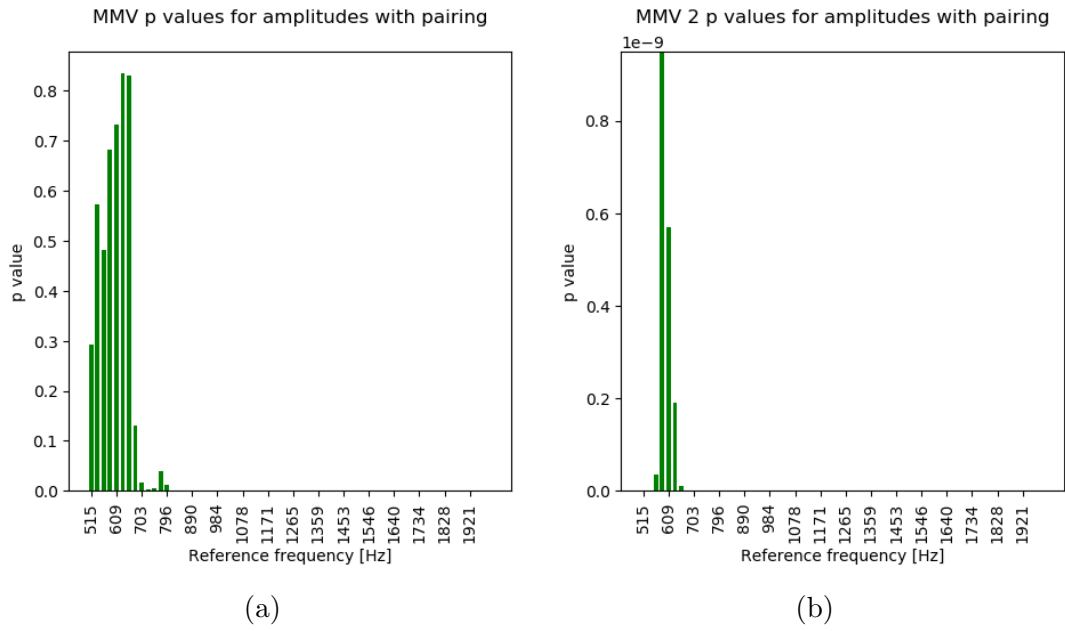


Figure 5.19: (a) MMV criterion and (b) MMV 2 criterion results, p values in bar plots for amplitudes with pairing of the most similar deflection shapes.

Let us now focus only on the results of MMV 2 criterion.

MMV 2

The resulting p values of MMV 2 together with R^2 are shown in Fig. 5.20 and they clearly indicate the different data across the entire frequency spectrum (the same is already visible in Fig. 5.19).

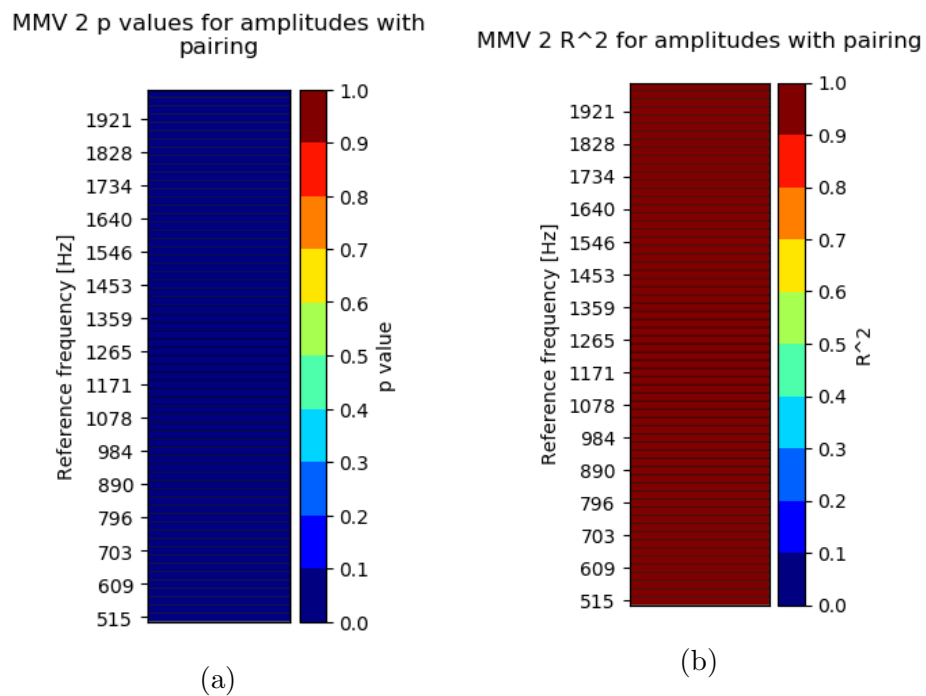


Figure 5.20: (a) MMV criterion and (b) MMV 2 criterion results, p values in bar plots for amplitudes with pairing of the most similar deflection shapes.

The results of statistical testing at chosen statistical significance levels are illustrated in Fig. 5.21, each time the zero hypothesis was rejected signifying all the paired deflection shapes different.

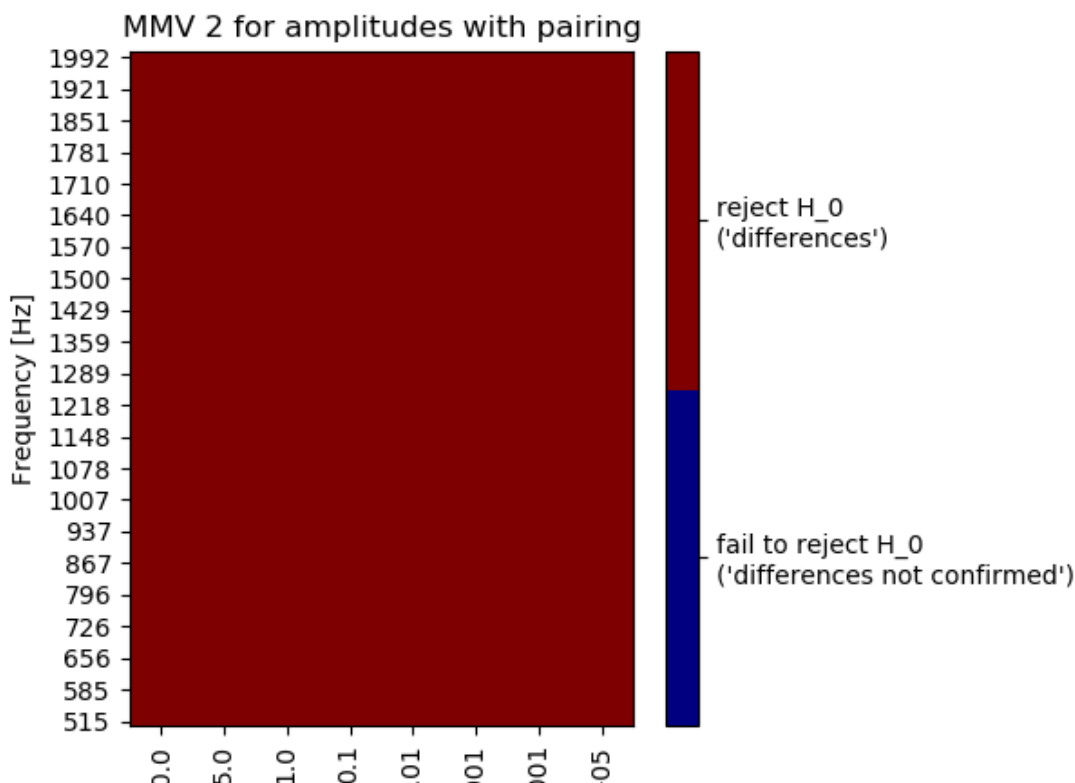


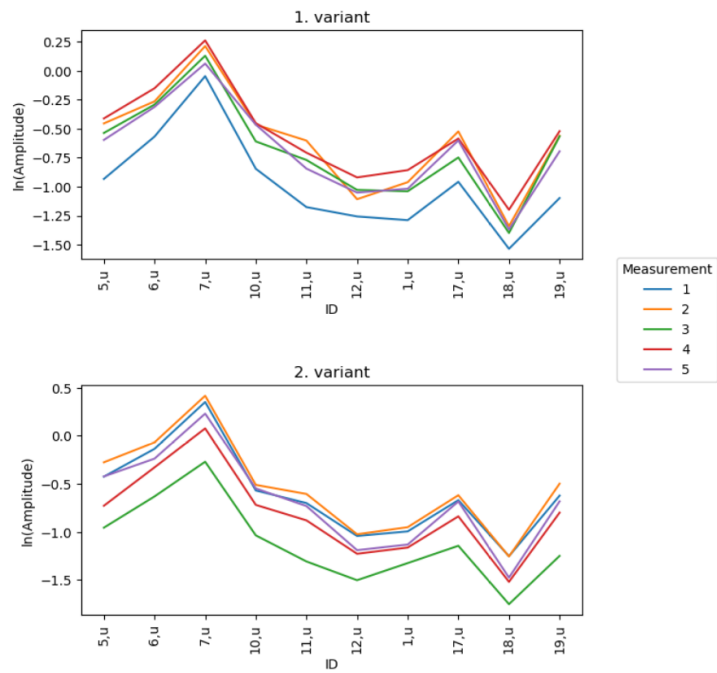
Figure 5.21: Results of statistical testing for amplitudes with pairing.

We do also the same trick as with the previous data of a part with and without crack, we take only the first variant of data and divide it into two fictive variants of the same type. Because the result of the criterion depends on the selection of measurements representing individual variants (and on the number of measurements as well), we perform this procedure twice. Firstly we take the first 5 original measurements as measurements of the first variant A and the last 5 measurements as measurements of the second fictive variant B. In the second case we take the 5 odd numbered original measurements as measurements of the first variant A and the even numbered to form the second fictive variant B.

Looking at the data, the second way of division seemed to split the measurement more ‘evenly’, as an example, we provide a graph of the data again for the fixed frequency 703.125Hz (see Fig. 5.22). The conjecture corresponds with obtained results in Fig. 5.23 – 5.25.

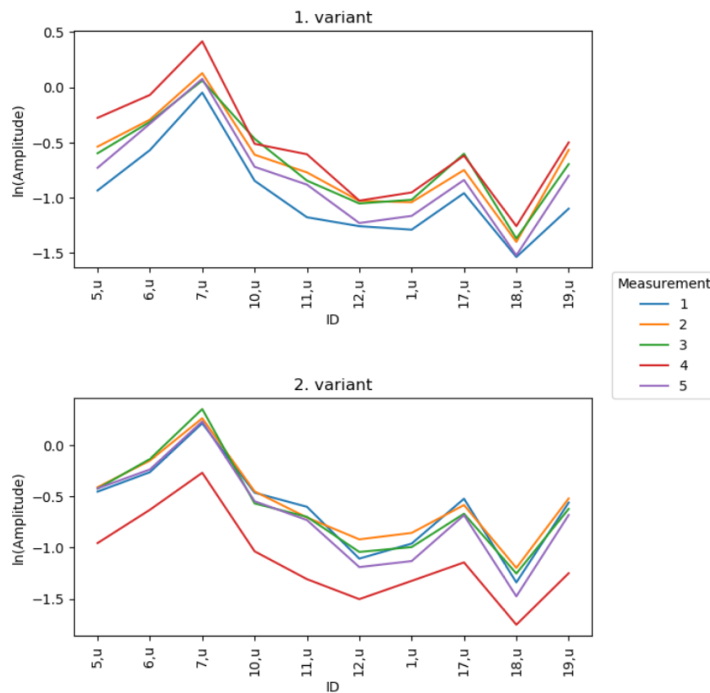
The plots of p values and R^2 follows as well as results of statistical testing for both types of measurement splitting, all results for amplitudes applied on paired deflection shapes (see Fig. 5.23 – 5.25), the influence of redistribution of measurements within variants on the results is worth noting.

Fixed frequency, fixed variant, all measurements, with MMV 2 pairing



(a)

Fixed frequency, fixed variant, all measurements, with MMV 2 pairing



(b)

Figure 5.22: Graph of the data 5×2 , fixed frequency 703.125Hz, (a) splitting of measurements to first 5 and last 5 (b) splitting to odd and even numbered original measurements.

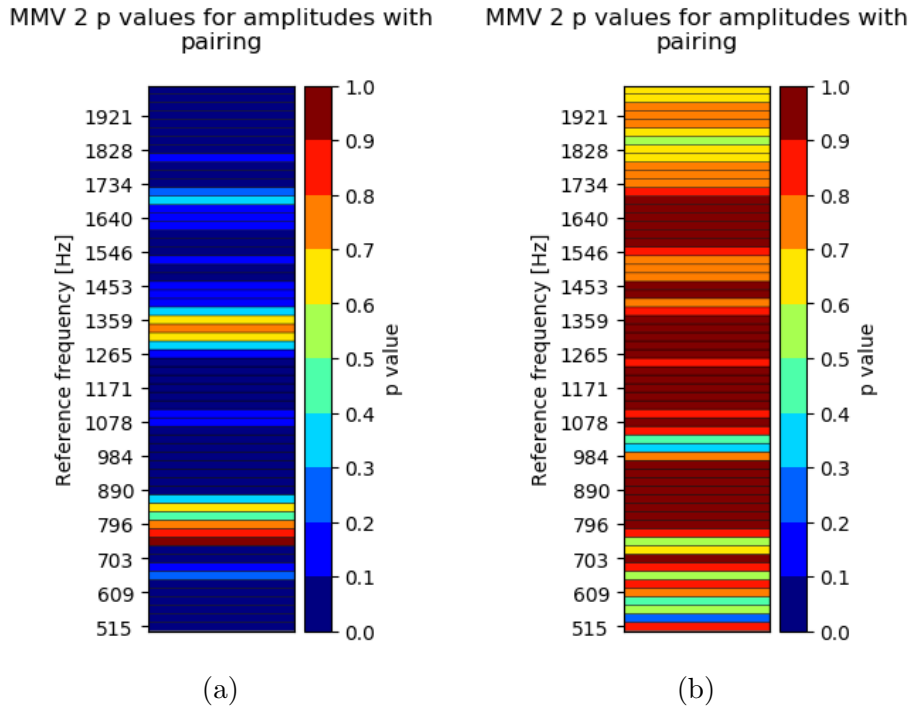


Figure 5.23: MMV 2 criterion p values, data 5×2 , (a) splitting of measurements to first 5 and last 5 (b) splitting to odd and even numbered measurements.

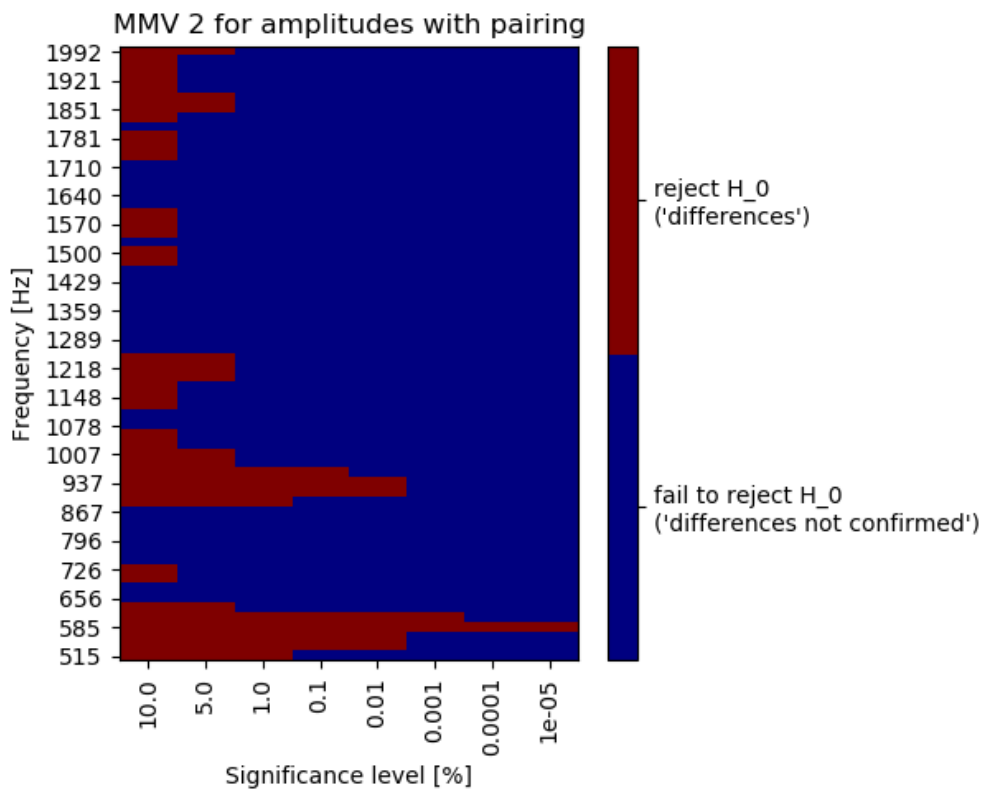


Figure 5.24: MMV 2 criterion statistical testing, data 5×2 , splitting of measurements to first 5 and last 5.

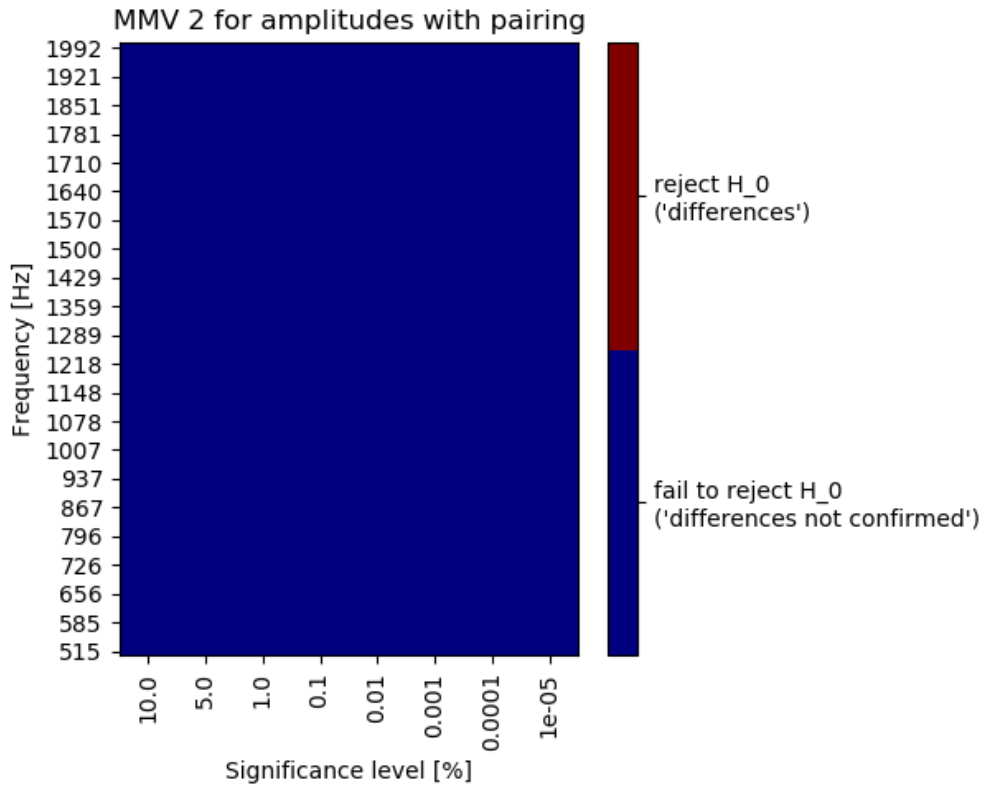


Figure 5.25: MMV 2 criterion statistical testing, data 5×2 , splitting to odd and even numbered measurements.

As the last example, we show the MMV 2 criterion applied only to the first two original variants of the data, each 10 times measured. In graphs in Fig. 5.18, these two variants for the frequency of 703.125Hz looks similar. The results are given in Fig. 5.26, 5.27 and show obvious differences in whole area of cca 930 – 1550Hz.

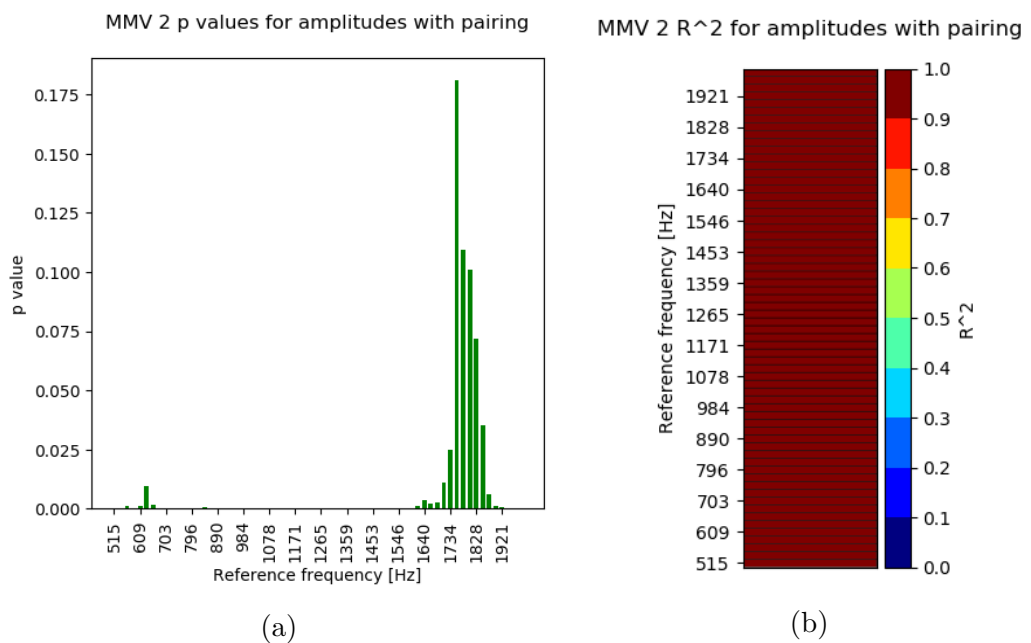


Figure 5.26: MMV 2 criterion, data 10×2 , (a) p values in bar plot, (b) R^2 .

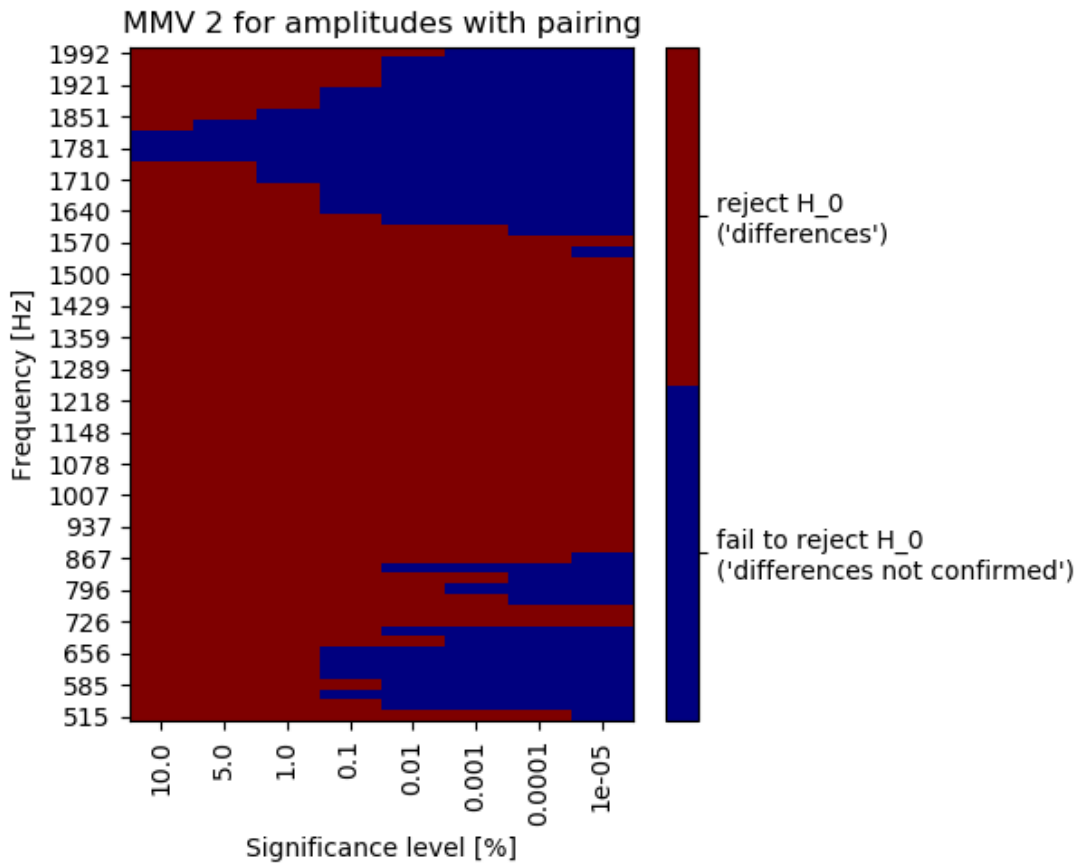


Figure 5.27: MMV 2 criterion, data 10×2 , statistical testing.

Chapter 6

MAC application

The basic information about the application itself are mentioned within this chapter. Two application windows were created and can be used, the first one was named MAC and contains correlation criteria for two data sets comparison (described in chapter 4), the second one was named MAC MV (MAC More Variants) and contains the statistical processing of multiple data sets (described in chapter 5).

6.1 Requirements

The requirements for running the MAC and MAC MV apps follows.

6.1.1 Installation of Anaconda package

Download Anaconda 5.0.0 or newer with Python 3.6 version from here (see Fig. 6.1): <https://www.anaconda.com/download/>.

Run the installer and go through with following settings.

- Select an installation for ‘Just Me’,
- as a destination folder, select drive D:\ anaconda\
(the path can not contain spaces or unicode characters, if you select a different folder, remember the directory path),
- do not add Anaconda to your PATH and not register Anaconda as your default Python 3.6 (see Fig. 6.2),
- after the installation, check or uncheck ‘Learn more about Anaconda Cloud’ and ‘Learn more about Anaconda Support’ as you wish and click the ‘Finish’ button.

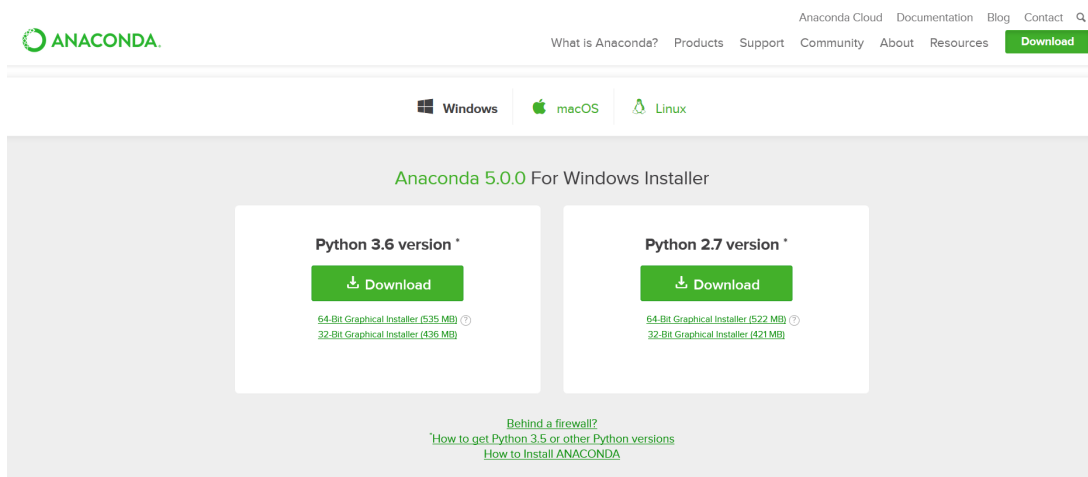


Figure 6.1: Anaconda download website.

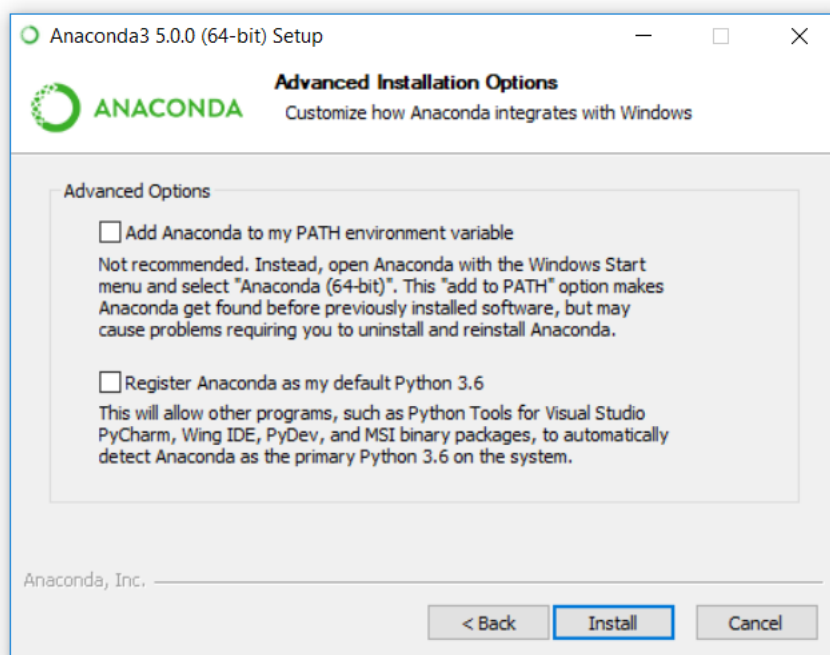


Figure 6.2: Anaconda installation settings.

6.1.2 How to run the MAC app

Copy the whole ‘MAC’ folder with everything inside into your drive D.

The MAC folder contains files ‘mac_gui.py’ and ‘mac_mv_gui.py’ and a folder ‘needed’. The tools, you can run and use, are ‘mac_gui.py’ (MAC for two input data) and ‘mac_mv_gui.py’ (MAC MV for more input data). The folder ‘needed’ contains the necessary modules and files to make the applications work properly – you can take a look inside, but **do not edit or change** anything in it (unless you know, what you are doing). If you ever want to move the ‘mac_gui.py’ or

‘mac_mv_gui.py’ file, move it together with the ‘needed’ folder.

Double click the ‘mac_gui.py’ and set the program from ‘Other programs’ to launch it (see Fig. 6.3).

- If you want to run only a graphical version of the application, choose D:\anaconda\python.pyw,
- if you want to run the graphical version and see the command line as well, choose D:\anaconda\python.py.

When choosing to launch the MAC app with displaying the command line, it is possible to see which calculation is in progress and how long the finished calculation has taken.

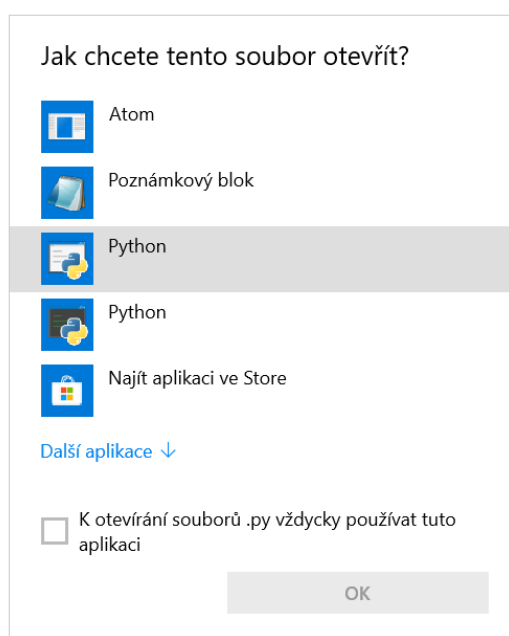


Figure 6.3: Set Python from Anaconda to run the application.

Now you should be able to run the MAC app with double click the ‘mac_gui.py’.

To run the MAC MV app, double click ‘mac_mv_gui.py’ and set the program to launch it in the same way as with MAC app.

Batch file

Using a batch file can be another way to run the MAC (or MAC MV) app anywhere from your PC. A simple batch file, a simple script, where only the Python and the MAC app are called, may be written. Open a Notepad and write firstly the path to the ‘python.exe’ program in quotes, then a space and the path to the ‘mac_gui.py’ app in quotes. If the standard paths were used, the Notepad should contain:

“D:\anaconda\python.exe” “D:\anaconda\mac_gui.py”

Save the file with a ‘.bat’ extension as ‘starter.bat’, without the ‘.txt’ extension, anywhere you want and run the MAC app with double click the ‘starter.bat’.

6.2 MAC and MAC MV apps

If you followed instructions in the section 6.1, you can run the MAC app (or MAC MV app) with double click the ‘mac_gui.py’ (‘mac_mv_gui.py’) or the ‘starter.bat’ file.

The main window of both applications is structured as follows. There are the ‘Browse...’ buttons for selection of your input data in the upper part of the window, the main tabs ‘Input Data’, ‘Criteria’, ‘Criteria Parameters’ and ‘Output Options’ in the middle of the window and the ‘Start’ button to start evaluations of selected calculations at the bottom of the window (see Fig. 6.4a). And there is a ‘Help’ with the documentation in the main ribbon.

With pointing to a button or a cell in the application window a tooltip appears in most cases, to specify what the button or cell is used for (see Fig. 6.4b).

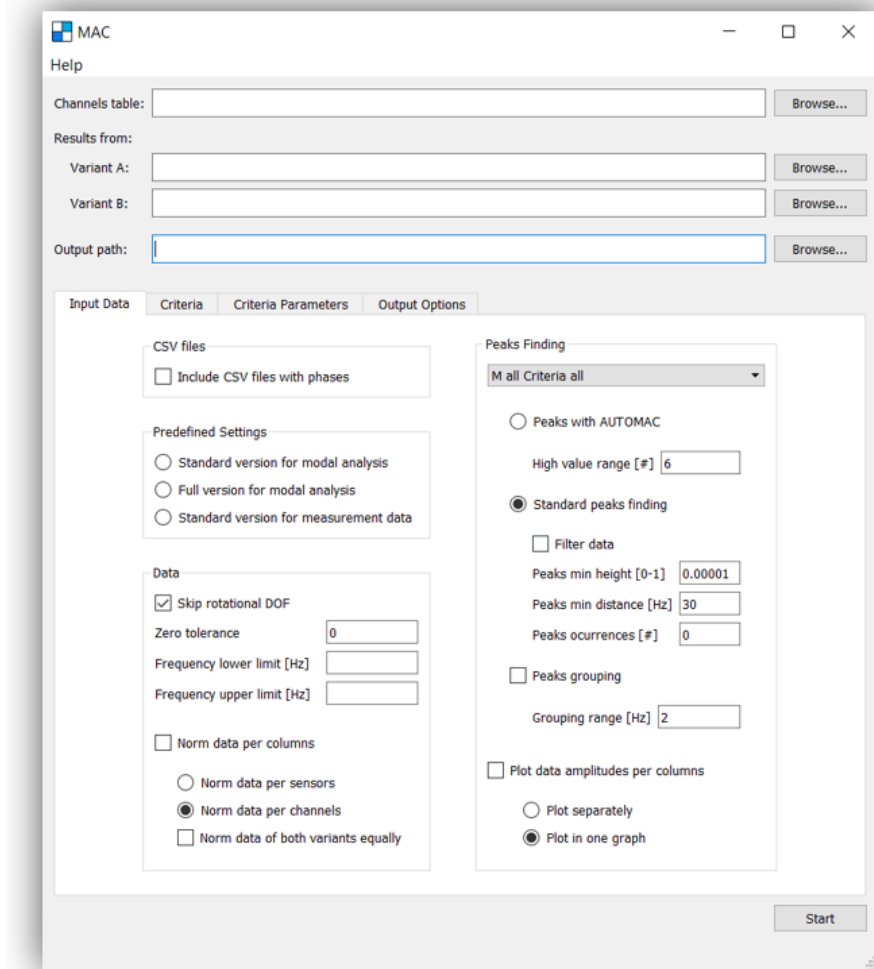
6.2.1 File selections

By clicking on the ‘Browse...’ button the standard dialog box for selecting a file or folder opens. When selecting a file, there is a drop-down list with supported file formats at the bottom right corner of the dialog box (see Fig. 6.5).

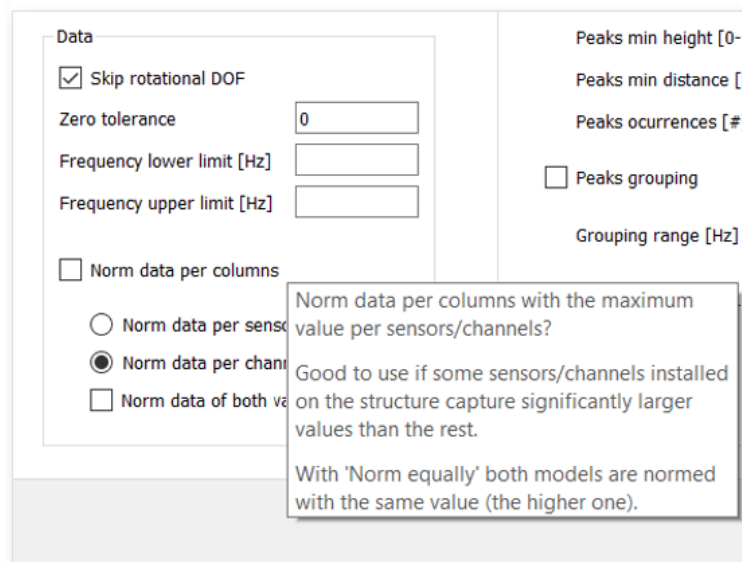
In MAC MV app for Variant A a dialog box for selecting a file opens, there should be selected a first measurement of the first reference variant of input data (reference variant for MM criterion or pairing, frequencies of this variant are then displayed in plots). For Data B in MAC MV app a dialog box for selecting a folder opens. In the selected folder should be places all other measurements of the first reference variant and all other variants and their measurements (together with the first measurement of the first reference variant, everything from this folder will be imported and used in calculations).

Supported file formats are CSV (*.csv), PUNCH (*.pch) and UFF (*.uf, *.uff and *.unv).

One CSV file with amplitudes or two CSV files, the first one with amplitudes and the second one with phases, can be considered as the input data of one measurement of one variant. CSV files with amplitudes are selected in the dialog box and for inclusion of the phases there is a check button in the tab ‘Input Data’ in the app – in this case the CSV files including phases must be located in the same folder, where CSV files with amplitudes are, and have exactly the same name as files with amplitudes and end with ‘p.csv’. e.g. if the file with amplitude is named ‘xyz_ac.csv’, then the file with phases must be named as ‘xyz_acp.csv’.



(a)



(b)

Figure 6.4: (a), (b) MAC application.

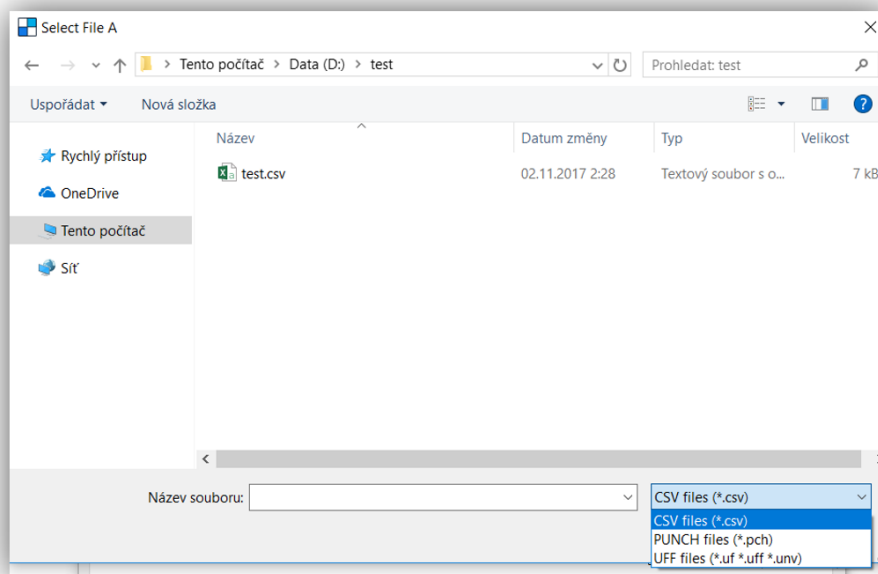


Figure 6.5: Dialog box for selecting a file.

In case of more variants of input data, in MAC MV app, individual measurements of the same variant must be named with the same name of variant and end with number: first measurement with 1, second with 2, third with 3, etc. right before the file extension and eventual ‘p’ letter. For MM criterion, there must be exactly two variants, one reference variant once measured and the second variant once or more times measured. For MMV and MMV 2 criteria, there must be at least 2 variants, each 2 times measured, and the number of measurement must be the same for each variant. According to selected reference variant, the pairing with the reference variant is done.

Here is an example for MMV criterion for 2 measurements of the first variant ‘xxx’ saved in CSV files with amplitudes and phases and 2 measurements of a second variant ‘yyy’ saved in PUNCH files. The data should be saved e.g. in folders ‘A’ and ‘B’ as

- folder ‘A’: ‘xxx_1.csv’, ‘xxx_1p.csv’,
- folder ‘B’: ‘xxx_2.csv’, ‘xxx_2p.csv’, ‘yyy_1.pch’, ‘yyy_2.pch’

and in the MAC MV application window, in Variant A should be selected the file ‘xxx_1.csv’ and in Data B should be selected the whole folder ‘B’.

After calculations are done, an ‘OUT’ folder is created in the selected output path. In the MAC app, the results are denoted as FDAC (MWFDAC etc.), if some complex input data were imported, otherwise the results are denoted as MAC (i.e. the notation FDAC or MAC is based only on the complexity of the data, it is not determined whether there are actually modal vectors or deflection shapes at the input). In the MAC MV app, for both deflection shapes and modal vectors the notations MM, MMV, MMV 2 are used. It is also practically possible to have the calculations done for one input data with modal vectors and others

with deflection shapes, in both MAC (detected as FDAC) and MAC MV apps.

6.2.2 Format of selected data

Files with input data must be arranged in a certain way. In all supported file formats, CSV (*.csv), PUNCH (*.pch) and UFF (*.uf, *.uff and *.unv), it is assumed that frequencies and sensor numbers are sorted from the smallest value to the largest and channel specifications are unambiguous. Only three possibilities of DOF specifications are considered, either DOF 1–6 ('u', 'v', 'w', 'phiu', 'phiv', 'phiw') or DOF 1–3 ('u', 'v', 'w') or just one arbitrary DOF.

CSV

CSV files contain only real numbers and the following is expected (see Fig. 6.6).

- There are frequencies in the first column and channel specifications in the first row and the corresponding data thereto.
- Channel specifications include sensor numbers and translational DOF specifications 'u', 'v', 'w' and/or rotational DOF specifications 'phiu', 'phiv', 'phiw', which are listed in mentioned order and separated by a comma (the DOF specifications follow the comma).
- The 'u', 'v', 'w', 'phiu', 'phiv', 'phiw' specifications are repeated for each sensor number in the same way, listed in the same order throughout the document. For example, there can not be columns with channel specifications '50,u', '50,v', '50,w' and '51,u', '51,w' only (the column '51,v' is missing).
- The data in the CSV file including phases are arranged identically as the data in CSV file with amplitudes and the values are given in degrees. Any row or column can not miss or be redundant and all frequencies and channel specifications must be the same in both files with amplitudes and phases.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Frequency	N50,u	N50,v	N50,w	N50,phiu	N50,phiv	N50,phiw	N62,u	N62,v	N62,w	N62,phiu	N62,phiv	N62,phiw
2	2.5E+01	-1.4E-05	-5.3E-06	-2.8E-06	-1.2E-05	-4.2E-04	-5.1E-05	-1.1E-05	5.8E-06	3.3E-06	-5.1E-05	-5.9E-04	-4.2E-04
3	8.4E+01	2.1E-06	8.4E-07	-8.0E-06	-1.1E-05	-2.0E-04	2.7E-06	-1.9E-06	-1.7E-06	-8.3E-06	6.4E-05	7.2E-05	-2.0E-04
4	1.2E+02	-1.8E-08	3.7E-05	1.9E-07	-1.9E-06	-1.4E-05	1.8E-06	2.4E-08	-4.2E-04	-1.3E-07	5.6E-04	-8.3E-06	-4.2E-04
5	1.7E+02	1.0E-05	-8.3E-06	-2.0E-04	2.4E-08	-8.3E-06	-4.2E-04	-8.8E-06	-6.5E-06	-5.1E-05	-5.9E-04	-4.2E-04	-1.9E-06
6	3.4E+02	-5.4E-06	1.8E-06	-1.4E-05	-8.8E-06	-1.1E-05	7.2E-05	7.2E-05	-2.5E-05	6.4E-05	7.2E-05	-2.0E-04	2.4E-08
7	3.8E+02	3.7E-05	-9.8E-06	-1.9E-06	-4.2E-04	2.7E-05	3.0E-05	3.0E-05	3.7E-05	2.7E-05	3.0E-05	-1.2E-05	-5.1E-05
8	5.8E+02	-6.3E-05	-1.4E-05	2.4E-08	-2.0E-04	-5.1E-05	-2.5E-05	-2.5E-05	-2.6E-05	-4.2E-04	-2.5E-05	-1.4E-05	6.4E-05
9	6.5E+02	-4.2E-04	3.9E-04	1.3E-05	-1.2E-05	6.4E-05	3.3E-06	3.3E-06	-3.8E-04	7.2E-05	-4.2E-04	-4.3E-04	5.6E-04
10	8.1E+02	1.7E-08	3.1E-04	4.4E-05	7.2E-05	5.6E-04	-1.2E+02	1.8E-04	2.8E-04	3.0E-05	-4.2E-04	-1.1E-05	2.7E-06
11	8.8E+02	-1.2E-05	2.7E-05	-2.4E-05	3.0E-05	-5.9E-04	3.0E-05	-1.2E-04	-8.3E-06	-2.5E-05	-2.0E-04	1.2E+02	1.8E-06
12	8.8E+02	-6.6E-04	-6.5E-05	-3.6E-06	-8.3E-06	7.2E-05	-1.4E-05	-6.5E-04	2.7E-05	3.3E-06	-1.2E-05	-2.5E-05	2.7E-05
13	8.8E+02	-4.3E-04	-2.5E-05	-1.4E-05	1.8E-06	3.0E-05	-2.6E-05	-4.1E-04	2.2E-05	-4.5E-05	-3.8E-04	3.3E-06	-2.5E-05
14	9.4E+02	6.1E-06	1.1E-04	4.2E-04	-9.8E-06	-2.5E-05	-2.5E-05	-8.9E-05	1.3E-04	4.1E-04	2.8E-04	1.4E+02	3.3E-06
15													
16													
17													
18													

Figure 6.6: Example of CSV input file.

PUNCH

In PUNCH files, the following is expected.

- At the beginning of the PUNCH document is line, that starts with either ‘\$REAL-IMAGINARY OUTPUT’ or ‘\$MAGNITUDE-PHASE OUTPUT’ or ‘\$REAL OUTPUT’, between the first lines starting with ‘\$TITLE’ and ‘\$SUBCASE ID’.
- The whole document then may be structured either in blocks by ‘POINT ID’ with frequencies stated before the letter ‘G’ or in block by ‘EIGENVALUE’ with sensor numbers stated before the letter ‘G’.
- The blocks are separated by ‘POINT ID’ or ‘MODE’, the file contains at least two blocks and the number of frequencies or the number of sensor numbers per each block is constant and the values are listed in the same order throughout the document.
- According to the minimal number of elements in all parts of the document between the letters ‘G’, the number of DOF is set. Only four possibilities are considered, either all parts have lengths one (the number of DOF is one), all parts have lengths of three (the number of DOF is three), all parts have lengths of six (the number of DOF is six) or some parts have lengths of three and some lengths of six (the number of DOF is three).
- In case of ‘REAL-IMAGINARY’, firstly real and then imaginary components of numbers are listed within the parts. In case of ‘MAGNITUDE-PHASE’, firstly amplitudes, then phases are listed and phases are given in radians. And in case of ‘REAL OUTPUT’, there are only real components of number.
- The numbers in parts are sorted in the order corresponding to specifications ‘u’, ‘v’, ‘w’, ‘phiu’, ‘phiv’, ‘phiw’ (firstly the real, then the imaginary components are sorted this way etc.). At the end of data importing, the DOF specifications ‘u’, ‘v’, ‘w’, ‘phiu’, ‘phiv’, ‘phiw’ are added to sensor numbers in this order (for DOF 1–6, for less DOF analogously).

UFF

All UFF files data may be read by setts with Python Module `pyuff.py` [29].

In UFF files with `.uf` or `.uff` extension, the following is assumed.

- The key names are: ‘x’ for frequencies, ‘ref_dir’ for DOF specifications, ‘id4’ for a type of data and ‘data’ for the data of the modal matrix. The sensor numbers are made of combination of integer numbers denoted with ‘rsp_node’ and ‘rsp_node’, placing digits next to each other and separating them by extra zero.
- The file contains at least 2 sensor numbers. The frequencies per each sensor are the same and the type of data as well.

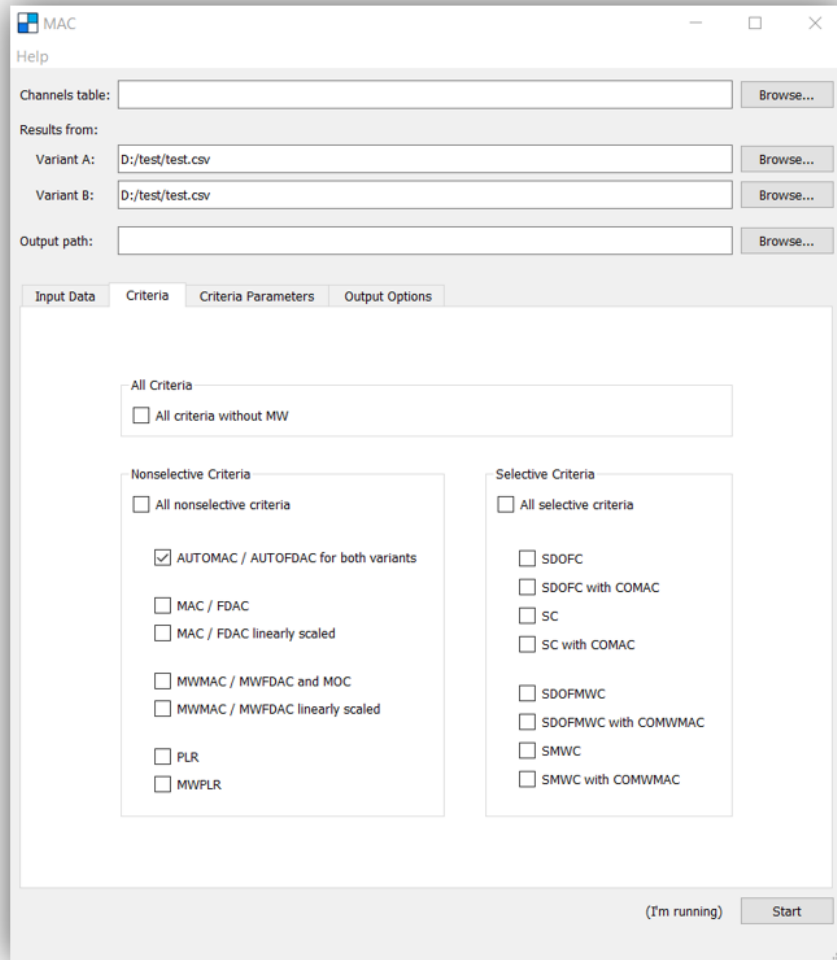
- The type of data may be complex only, the importing is not implemented for other data types. The specification for 'id4' must end with 'omplex' (it may be 'komplex' or 'complex') meaning that the complex data are saved with real and imaginary parts.
- The DOF specifications are not changing for individual sensor numbers and they are sorted in the same way within the whole document. The DOF specifications are integers between one and six. The numbers 1, 2, 3, 4, 5, 6 are renamed into specifications 'u', 'v', 'w', 'phiu', 'phiv', 'phiw' in this order. Any other eventual value or in case of not standard number of DOF, all DOF specifications are renamed into just one DOF 'u'.
- The sets of data consists of vectors of data per individual channels (sensors and DOFs) and they are sorted (the first sensor and all its DOFs, the second sensor and all its DOFs etc.).

In UFF files with .unv extension, the following is assumed.

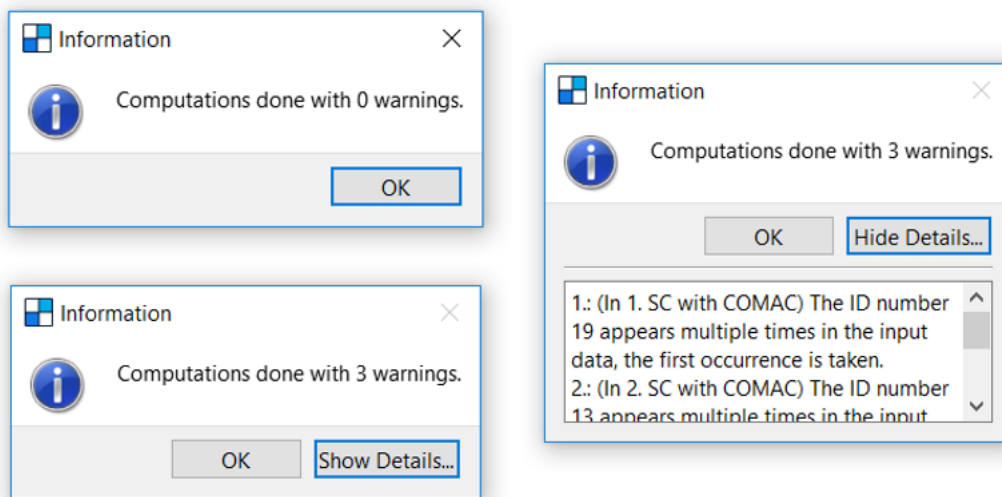
- In the first four sets are just basic information.
- The key names are: 'freq' for frequencies, 'node_nums' for sensor numbers, 'n_data_per_node' for the number of DOF per sensors, 'data_type' for a type of data and the data of the modal matrix are accessed per DOF labels 'r1', 'r2' up to 'r6'.
- The file contains at least 2 sensor numbers. The frequencies per each sensor are the same and the type of data as well.
- The type of data may be complex only, the importing is not implemented for other data types. The specification for 'data_type' must be value 5 meaning that the complex data are saved with real and imaginary parts.
- The DOF specifications are not changing for individual sensor numbers and they are sorted in the same way within the whole document. The DOF specifications are set according to the number of DOF, for 6 DOF as 'u', 'v', 'w', 'phiu', 'phiv', 'phiw' in this order (and analogously for 3 DOF or 1 DOF – only DOF 'u' is used in this case).
- The sets of data are arranged per sorted DOFs (accessible with keys 'r1' up to 'r6' in case of 6 number of DOF) and the sets themselves are arranged per sorted sensor numbers.

6.2.3 Starting the calculations

After setting all the values in tabs 'Input Data' – 'Output Options' (see specifications for MAC and MAC MV below separately) and pressing the 'Start' button, the application status changes in the lower-right corner of the app next to the 'Start' button into 'I'm running' (see Fig. 6.7a). When computations are done, the error or informative window, with or without warnings details, pops up (see Fig. 6.7b) as well as the application status changes again into 'Finished'.



(a)



(b)

Figure 6.7: (a) The change of application status and (b) informative pop-up windows.

If the button ‘Export XLSM for the overview of main results’ in the ‘Output Options’ tab have been checked, the XLSM file opens when the informative pop-up window is closed.

All outputs are saved in the ‘Output path’ or in the folder, where the application was started (in the current working directory) if the ‘Output path’ has not been selected.

6.3 MAC app Input Data – Output Options

Here is an overview of adjustable options in the ‘Input Data’ – ‘Output Options’ tabs in the MAC app. The following information is contained in the tooltips in the application as well.

6.3.1 Input Data

CSV files

‘Include CSV files with phases’ – In case CSV files were selected as the input, do you want to include CSV files with phase information in calculations too? CSV files including phases must have the same name as files including amplitudes and end with ‘p.csv’.

Predefined Settings

Three basic predefined settings.

Data

‘Skip rotational DOF’ – Let or skip rotational degrees of freedom (if there are rotational DOF included in the input data)?

‘Zero tolerance’ – Sensors with a maximum value less than or equal to ‘Zero tolerance’ and DOF (e.g. all translational DOF ‘u’) with a maximum value less than or equal to ‘Zero tolerance’ are skipped.

‘Frequency lower limit [Hz]’ – Data less than ‘Frequency lower limit’ are cut off. An empty cell means nothing cut off regardless of the frequency range of the input data.

‘Frequency upper limit [Hz]’ – Data greater than ‘Frequency upper limit’ are cut off. An empty cell means nothing cut off regardless of the frequency range of the input data.

‘Norm data per columns’ – Norm data per columns with the maximum value per sensors/channels? Good to use if some sensors/channels installed on the structure capture significantly larger values than the rest. With ‘Norm equally’ both variants are normed with the same value (the higher one).

Peaks Finding

A combo box – Choose whether all the data should be used or only peaks found, for computation of mass matrix ‘M’ and other ‘Criteria’ (e.g. ‘M all Criteria peaks’ means all applicable input data are used for mass matrix computation, but only the peaks found are applied in the criteria calculations).

‘Peaks with AUTOMAC’ – Selects a peak at AUTOMAC diagonal value if all values in the corresponding row (and column), except the diagonal value and values in \pm ‘High value range [#]’, are less than 0.5.

‘Standard peaks finding’ – Search for peaks through each column of the data.

‘Filter data’ – Remove noise from the data? Method doesn’t need any parameter and works well, but is slow for data with many columns (channels).

‘Peaks min height [0-1]’ – Relative minimum of peak height (value between zero and one) – lower values are not identified as peaks.

‘Peaks min distance [Hz]’ – Minimum frequency distance between peaks (value in Hz) – in case of more peaks found closer to each other only the maximum is preserved as peak.

‘Peaks occurrences [#]’ – How many times the peak must appear through the columns (channels)?

‘Peaks grouping’ – Group the found peaks at points that are nearby – every point is grouped with another one if it’s closer than \pm ‘Grouping range [Hz]’ to each other, finally only averages of groups retained.

‘Plot data amplitudes per columns’ – Plot amplitudes of data per columns? Plot them into one graph or into two graphs separately?

6.3.2 Criteria

In this tab it is possible to select which criteria to calculate.

All criteria from the section 4 may be selected.

6.3.3 Criteria Parameters

Mass Weighted Criteria

‘Regularisation’ – Which type of regularisation in mass weighted versions of criteria should be used? ‘Tikhonov and Schulz’ or ‘IC Symmetric and Schulz’ are mostly recommended.

‘Reg. loops limit’ – Limits for loops in regularization.

‘Reg. loops limit extension’ – Limits extension for loops in regularization – so many more loops are made and it is checked whether the value of the result is improved.

‘Schulz method of third order’ – Higher order of Schulz method? Implies faster convergence (less steps needed) with orderly similar results.

‘M with arbitrary precision’ – Compute the mass matrix M with an arbitrary precision floating point arithmetic? Computation of M with the arbitrary precision is much slower!

‘Decimal precision’ – Set the exact decimal precision (valid only for M with arbitrary precision = True).

Pairing and Linear Regressions

‘Type of pairing’ – Type of pairing used for linear regression and selective criteria. The ‘With unique maximization’ option uses a Munkres algorithm (also called the Hungarian algorithm or the Kuhn-Munkres algorithm) to find an unambiguous pairing such that the sum of values through this pairing is the maximum possible.

Selective DOF Criteria

‘High pairs only [0-1]’ – Let only pairs with (MW)MAC / (MW)FDAC value greater than ‘High pairs only’ (value from zero-one interval) for SDOFC computations?

‘All plots’ – Plot whole (MW)MAC / (MW)FDAC matrices for each degree of freedom?

Selective Criteria

‘High pairs only [0-1]’ – Let only pairs with (MW)MAC / (MW)FDAC value greater than ‘High pairs only’ (value from zero-one interval) for SC computations?

‘Number of selected sensors [#]’ – The number of sensors selected and iteratively removed as the worst in SC.

‘All plots’ – Plot whole (MW)MAC / (MW)FDAC matrices after the worst sensors removals?

‘Check of pairs’ – Check if pairing is the same after the worst sensors removal?

6.3.4 Output Options

Plot Parameters

‘Show frequencies’ and ‘Show mode / deflection numbers’ – Frequencies or mode / deflection numbers in the plot?

‘Faster plots’ – Generate only a quick less-pretty graphs?

‘Zero-one color set’ – Which color set to use in graphs with values ranging from zero to one?

‘Neg.-pos. color set’ – Which color set to use in graphs with positive and negative values?

‘Cbar midparameter [0-1]’ – Relative middle parameter of the color set (value between zero and one).

‘Figure height [pixel]’ – Figure height, width is given automatically (square image size for basic criteria). The minimum size to fit all the labels is cca 500 pixels.

MS Excel Output File

‘Export XLSM for an overview of main results’ – Create an Excel XLSM (Macro-Enabled Workbook) file for an overview of the main results. The XLSM file is opened automatically after the calculations have finished in the app.

6.4 MAC MV app Input Data – Output Options

Here is an overview of adjustable options in the ‘Input Data’ – ‘Output Options’ tabs in the MAC MV app. The following information is contained in the tooltips in the application as well.

6.4.1 Input Data

CSV files

‘Include CSV files with phases’ – In case CSV files were selected as the input, do you want to include CSV files with phase information in calculations too? CSV

files including phases must have the same name as files including amplitudes and end with 'p.csv'.

Data Plots

'Plot data per columns' – Plot data per columns? For all IDs, for every variant (and all variant together), graphs of amplitude, ln(amplitude) or phase as a function of frequency for all measurements are plotted.

'Independency plots' – Plot data for MMV or MMV 2 distinction? For fixed frequency, for every variant, graphs of ln(amplitude) or phase as a function of IDs for all measurements or as a function of measurements for all IDs are plotted.

Data

'Skip rotational DOF' – Let or skip rotational degrees of freedom (if there are rotational DOF included in the input data)?

'Zero tolerance' – Sensors with a maximum value less than or equal to 'Zero tolerance' and DOF (e.g. all translational DOF 'u') with a maximum value less than or equal to 'Zero tolerance' are skipped.

'Frequency lower limit [Hz]' – Data less than 'Frequency lower limit' are cut off. An empty cell means nothing cut off regardless of the frequency range of the input data.

'Frequency upper limit [Hz]' – Data greater than 'Frequency upper limit' are cut off. An empty cell means nothing cut off regardless of the frequency range of the input data.

'Extract every n -th frequency' – Extract every n -th frequency and n -th row of all input data?

6.4.2 Criteria

In this tab it is possible to select which criteria to calculate.

'MM' – More Measurements. Criterion for two variants, the first variant once measured, the second variant once ore more times measured.

'MMV' – More Measurements and Variants. Criterion for two or more variants, every variant is at least twice measured. The linear dependencies between measurements of the same variant are not expected.

'MMV 2' – More Measurements and Variants 2. Criterion for two or more variants, every variant is at least twice measured. The linear dependencies between measurements of the same variant are expected.

6.4.3 Criteria Parameters

Pairing

‘Pairing’ – With pairing? Pairing with the first reference variant A is performed.

‘Pairing with amplitudes’ – Include only amplitude information in the pairing?

‘Pairing with amplitudes and phases’ – Include both amplitude and phase information in the pairing?

‘Independency plots with pairing’ – Plot data for MMV or MMV 2 distinction for paired set of frequencies? For fixed reference frequency, for every variant, graphs of $\ln(\text{amplitude})$ or phase as a function of IDs for all measurements or as a function of measurements for all IDs are plotted. Valid only for `MMV = True` or `MMV 2 = True` and if `Pairing = True` as well.

MMV and MMV 2

‘Significance levels alpha’ – Which vector of significance levels should be used for statistical testing?

‘Alpha plots’ – Plots of results of statistical tests at the chosen significance levels? Valid only for `MMV = True` or `MMV 2 = True`.

MMV

‘MMV plots with estimates’ – Plot graphs with estimates for MMV? For fixed frequency, for every variant, graphs of $\ln(\text{amplitude})$ or phase as a function of IDs for all measurements are plotted. Valid only for the case of `MMV = True`.

6.4.4 Output Options

Plot Parameters

‘Show frequencies’ and ‘Show mode / deflection numbers’ – Frequencies or mode / deflection numbers in the plot?

‘Faster plots’ – Generate only a quick less-pretty graphs?

‘Zero-one color set’ – Which color set to use in graphs with values ranging from zero to one?

‘Binary color set’ – Which color set to use in graphs with binary values?

‘Cbar midparameter [0-1]’ – Relative middle parameter of the color set (value between zero and one).

‘Figure height [pixel]’ – Figure height, width is given automatically (square image size for basic criteria). The minimum size to fit all the labels is cca 500 pixels.

6.5 Useful remarks

6.5.1 One more option how to use the MAC app

There is one more way how to use the MAC and MAC MV tools, without the graphical window, only with double click on the icon but with own custom settings in one of the scripts. For this type of usage do the following.

Create a new folder ‘MAC_run_from_folder’ in your ‘MAC’ folder, here we will copy the necessary files and folders and use this ‘run from folder’ version of the MAC or MAC MV app (to ensure that the settings in the graphical version of MAC and MAC MV apps remain the original). Open your ‘needed’ folder (what is the ‘needed’ folder is described in the section 6.1.2) and select and copy these files and folders:

- mac_more_variants.py,
- mac.py,
- A,
- B,
- needed_for_mac

into the ‘MAC_run_from_folder’ folder.

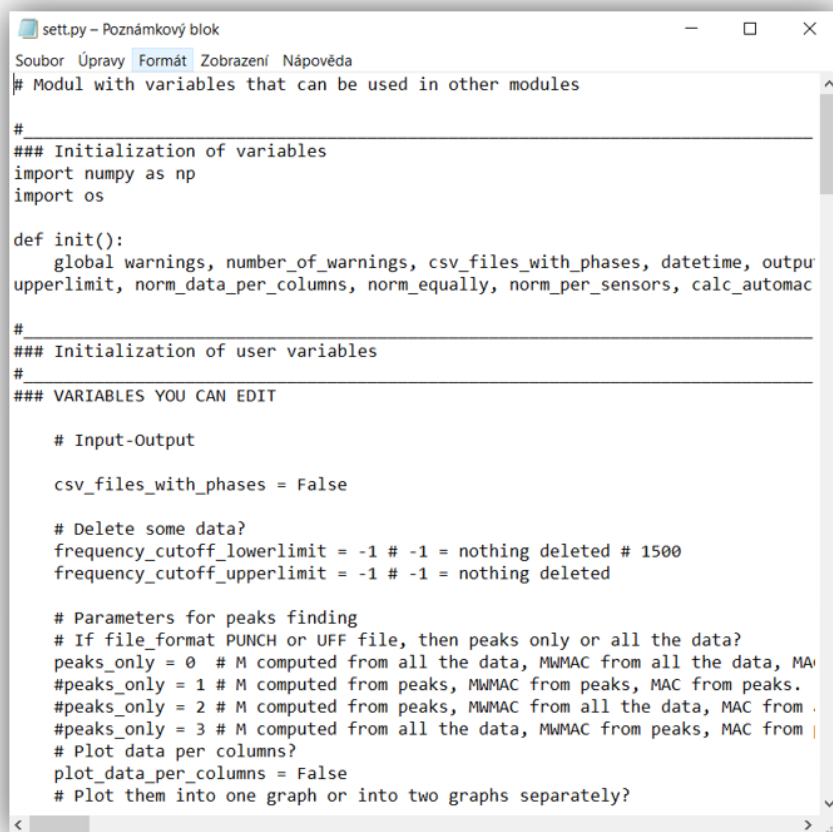
The file ‘mac_more_variants.py’ is the tool you will launch with a double click in case of two data sets comparison and the file ‘mac.py’ is the tool you will launch with a double click for multiple data sets comparison.

‘A’ and ‘B’ are empty folders where to put the input data, in A should be the data of the first measurement of variant A, in B the other data. Supported file formats are again PUNCH, UFF or CSV. If using MAC for two data sets comparison, in each folder A and B can be only one of PUNCH or UFF files or up to two CSV files – in this case, CSV files including phases must end again with ‘p.csv’ (there can not be more of your data in these folders, only the data you want to use in calculations when you run the ‘mac.py’). If using MAC MV for multiple data sets comparison, in A must be the first measurement of the first variant ending with the number 1 (denoting the first measurement) and other measurements and other variants must be in folder B (the same principle as with graphical user interface).

Folder ‘needed_for_mac’ contains again the necessary modules and files to make the tool work properly and, moreover, important files ‘sett_more_variants.py’ and ‘sett.py’ that needs to be opened and set to yours requirements. Before using ‘mac.py’, the ‘sett.py’ must be set and before launching ‘mac_more_variants.py’,

the 'sett_more_variants.py' must be set. We describe setting of 'sett.py' and all is analogous with 'mac_more_variants.py'.

Open the 'sett.py' in Notepad, PSPad, Spyder or any IDE (see Fig. 6.8).



```
sett.py - Poznámkový blok
Soubor Úpravy Formát Zobrazení Nápověda
# Modul with variables that can be used in other modules

#
### Initialization of variables
import numpy as np
import os

def init():
    global warnings, number_of_warnings, csv_files_with_phases, datetime, output_upperlimit, norm_data_per_columns, norm_equally, norm_per_sensors, calc_automac

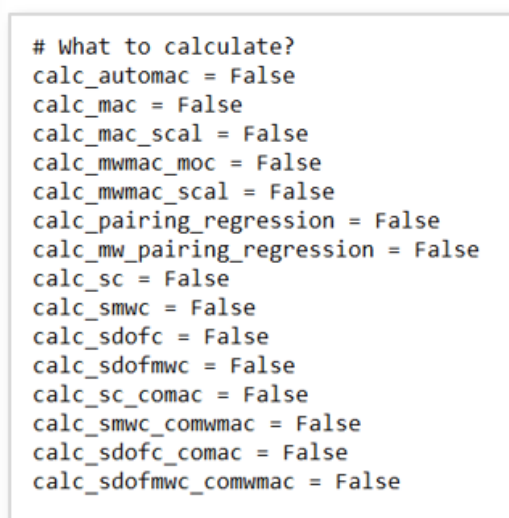
#
### Initialization of user variables
#
### VARIABLES YOU CAN EDIT

# Input-Output
csv_files_with_phases = False

# Delete some data?
frequency_cutoff_lowerlimit = -1 # -1 = nothing deleted # 1500
frequency_cutoff_upperlimit = -1 # -1 = nothing deleted

# Parameters for peaks finding
# If file_format PUNCH or UFF file, then peaks only or all the data?
peaks_only = 0 # M computed from all the data, MWMAC from all the data, MAC from all the data
#peaks_only = 1 # M computed from peaks, MWMAC from peaks, MAC from peaks.
#peaks_only = 2 # M computed from peaks, MWMAC from all the data, MAC from all the data
#peaks_only = 3 # M computed from all the data, MWMAC from peaks, MAC from all the data
# Plot data per columns?
plot_data_per_columns = False
# Plot them into one graph or into two graphs separately?
```

(a)



```
# What to calculate?
calc_automac = False
calc_mac = False
calc_mac_scal = False
calc_mwmac_moc = False
calc_mwmac_scal = False
calc_pairing_regression = False
calc_mw_pairing_regression = False
calc_sc = False
calc_smw = False
calc_sdfc = False
calc_sdfmwc = False
calc_sc_comac = False
calc_smw_comwmac = False
calc_sdfc_comac = False
calc_sdfmwc_comwmac = False
```

(b)

Figure 6.8: (a) The 'sett.py' file opened in Notepad and (b) variables for criteria calculations.

The character `#` is used for comments in the script, all after this character in the same line are just labels for better understanding. In a section ‘VARIABLES YOU CAN EDIT’ set values you want to use, especially select (write ‘True’ instead of ‘False’) which criteria you want to be calculated. Take attention on spacing, there can not be any redundant space character, otherwise the program will end with error (probably with an error message “Import of needed modules ... failed.”). Then save the file, close it and you can use the ‘mac.py’ by double click. If anything goes wrong, you overwrite something important etc., you can always delete the whole ‘sett.py’ file and copy it again from the ‘needed’ folder.

If you are using the Spyder or any IDE for launching the ‘mac_more_variants.py’ or ‘mac.py’ instead of double clicking on it, make sure to use the Python interpreter installed from the Anaconda. In Spyder the path to the Python interpreter can be set in ‘Tools > Preferences > Python interpreter’ by selecting ‘Use the following Python interpreter:’ with the standard path ‘D:\anaconda\python.exe’.

6.5.2 How long the calculations should approximately take

MAC app

Here is an example what can be seen from the command line, when all calculations were selected in the MAC app (run with graphical window) with the standard settings for loaded modal matrix size of cca 100 rows \times 100 columns containing 3 translational DOFs per sensors (on PC Fujitsu – Celsius W530 x64, processor Intel®Core™i5–4670 3.40GHz – 3.40GHz, RAM 8.00 GB, with Windows 7).

```
AUTOMAC
calculations time: 3.1029074305833513
MAC
calculations time: 3.276207595820119
MWMAC and MOC
calculations time: 8.226702792966238
PLR, MWPLR
calculations time: 2.396992050885924
SDOFC with COMAC
calculations time: 1.0331188938011167
SDOFC
calculations time: 0.22090833736314153
SC with COMAC
calculations time: 1.27199620142585
SC
calculations time: 0.8083703259637858
SDOFMWC with COMWMAC
calculations time: 8.1445031456564777
SDOFMWC
calculations time: 1.051228529634244
SMWC with COMWMAC
```

calculations time: 8.1948759094210345

SMWC

calculations time: 56.49866971421949

The AUTOMAC includes two calculations, one for both of the input data, MAC includes two calculations, the standard and the scaled version of MAC. The MW-MAC and MOC include calculations of both mass matrices M_1 , M_2 , three MOC criteria and the standard and the scaled version of MWMAC. Pairing per the shorter size was applied and the selective criteria ran with three sensors removal in this example.

All calculations times are in seconds. SC calculation is fast because the MAC values are computed for pairs only (the pairing is already done, in this case, and the MAC is not computed for all combinations of modal vectors, but only for modal vectors corresponding to pairs). On the contrary, SMWC is significantly the slowest because in every loop the mass matrices M_1 , M_2 are computed and the MWMAC is computed for all combinations of modal vectors.

MAC MV app

Here follows an example what is written in the command line, when MMV and MMV 2 calculations were selected in the MAC MV app (run with graphical window), with pairing (using only amplitudes) and plotting the ‘alpha plots’ for the vector of significance levels $[10, 1, 0.1, \dots, 0.00001]\%$ and other standard settings for one loaded modal matrix size of cca 9 rows \times 100 columns with 4 variants, each 10 times measured (on PC Acer Aspire E5-575G x64, processor Intel®Core™ i5-6200U 2.30GHz – 2.40GHz, RAM 8.00 GB, with Windows 10).

MMV

Calculations time: 9.29386730405005

MMV 2

Calculations time: 9.083900707797635

6.5.3 Other comments

Due to finite binary floating point representation of numbers in computers, floating point arithmetic in Python etc. [24], the MAC (FDAC) value can be slightly higher than the value of one (ev. less than zero). In scaling version of criteria the values above one are overwritten to the value of one before the transformation and values out of acceptable zero-one range are displayed as boundary values (zero or one) in graphs.

Regarding ‘big data’, in case of MAC app for two input data, the larger the data are, the more time calculations take. In case of MAC MV app for more input data, we encountered the problem of Memory Error in addition. The statistical models were implemented only in an elementary way in Python. The fact is, that we are

not dealing with sparse matrices and, moreover, the statistical testing is done for all combinations of modal vectors at the input for two data or when pairing is used. The Memory Error arising is dependent on more factors, on the PC system, where the memory is currently allocated, etc., thus it is not easy to specify some limit of input data sizes, for that the MAC MV app run without the Memory Error. In our numerical experiments, the error occurred when attempting to compute the MMV criterion for all combinations of modal vectors of two variant 10 times measured, with one modal matrix of size larger than 100 rows \times 6000 columns.

6.5.4 Program code

We are using several Python Modules (Python scripts), that are not our original work, these are

- mpmath [25],
- peakutils [26],
- gmpy2 [27],
- munkres [28],
- modified Module pyuff [29],
- tabulate [30]

and then other Anaconda and standard Modules.

In conclusion, we provide an example of the code for the MMV criterion, where the omitted parts are indicated by three dots. Only the code of MMV for all combinations of modal vectors for two variants was left. For the construction of matrices X and Y (Y is matrix of vectors y from the chapter 5 for all combinations of modal vectors) are attached schemes in appendix A.

```

"""
Calculate Z statistic for MMV.
"""
def z_calculate(Y, X, XXX, XM, XXXM, coef1, coef2, coef3,
               indexes, angle_recalculation=False):
    b = np.dot(XXX, Y)
    if angle_recalculation:
        # b[indexes,:][b[indexes,:] > 180] = 360 - b[indexes,:]
        for i in indexes:
            for j in range(len(b[i])):
                if b[i, j] > 180:
                    b[i, j] = 360 - b[i, j]
    ...
    YY = np.dot(X, b)
    S_e = np.sum((Y - YY) ** 2, axis=0) # vector

    b_M = np.dot(XXXM, Y)
    YY_M = np.dot(XM, b_M)

```

```

S_eM = np.sum((Y - YY_M) ** 2, axis=0)

Z = (S_eM - S_e) / S_e * (coef2 / coef1)

Y_aver = np.sum(Y, axis=0) / len(Y)
Y_aver = np.reshape(np.ones(len(Y)), (len(Y),1)) * Y_aver
S_T = np.sum((Y - Y_aver) ** 2, axis=0)
R2 = 1 - S_e/S_T
R2a = 1 - (S_e/coef2)/(S_T/coef3)

return Z, R2, R2a, b

"""
Method for the case of more measurements, more variants, concept
of no reference variant.
Rearrange data into regression model and calculate statistics Z
that have the F distribution.
Returns arrays of booleans, 1 means rejection of the hypothesis,
0 non-rejection, p values and R^2.

Parameters:
data: list of numpy arrays with transposed modal matrices,
sorted that the first is the first reference variant, all
measurements of it, then other variants and all their
measurements,
N: number of variants,
S: number of modal vector elements (number of sensor
specifications),
L: number of measurements,
frequencies: list of numpy arrays with vectors of frequencies,
pairing, pairing_with_phases, pairing_with_sums: specifications
for pairing
"""
def mmv_calculate(data, N, S, L, frequencies, pairing,
pairing_with_phases, pairing_with_sums):
...
# Rearrange data into regression model  $Y_a = X * b_a$ ,  $Y_p = X$ 
*  $b_p$ ,
data_a = np.log(np.array([np.abs(data[i]) for i in range(N *
L)]))
if sett_more_variants.mac_name == "FDAC":
data_p = np.array([np.angle(data[i], deg=True) for i in
range(N * L)])

if pairing and mmv_calculate_matrices:
...
else:
if mmv_calculate_matrices:
R1, = data[0].shape
R2, = data[sett_more_variants.number_of_data - 1].shape
Y_a = np.zeros((N * S * L, R1 * R2), dtype=np.float64)
Y_a_blocks1_r1 = np.zeros((S * L, R2), dtype=np.float64)
Y_a_blocks1_r2 = np.zeros((S * L, 1), dtype=np.float64)
for r1 in range(R1):
for s in range(S):
Y_a_blocks2_r1 = data_a[0:L, r1, s]
Y_a_blocks1_r1[s * L: s * L + L] =

```

```

        Y_a_blocks2_r1.reshape((len(Y_a_blocks2_r1), 1)) #
        size: L x 1
        Y_a[0:S * L, r1 * R2:(r1 + 1) * R2] = Y_a_blocks1_r1
indexes = np.array(range(R1)) * R2
for r2 in range(R2):
    for s in range(S):
        Y_a_blocks2_r2 = data_a[L:2 * L, r2, s]
        Y_a_blocks1_r2[s * L: s * L + L] =
            Y_a_blocks2_r2.reshape((len(Y_a_blocks2_r2), 1)) #
            size: L x 1
        Y_a[(S * L): 2 * (S * L), indexes + r2] = Y_a_blocks1_r2
    if sett_more_variants.mac_name == "FDAC":
        ...
else:
    ...

X = np.zeros((N * S * L, N * S), dtype=np.float64)
X_ones = np.zeros(L, dtype=np.float64) + 1
X_block1 = np.zeros((S * L, S), dtype=np.float64)
for s in range(S):
    X_block1[s * L:(s + 1) * L, s] = X_ones
X_block2 = X_block1.copy()
for s in range(1, S):
    X_block2[s * L:(s + 1) * L, 0] = X_ones
X[:S * L, :S] = X_block1
for n in range(1, N):
    X[n * S * L:(n + 1) * S * L, :S] = X_block1
    X[n * S * L:(n + 1) * S * L, n * S:(n + 1) * S] = X_block2
XT = np.transpose(np.conjugate(X))
XXX = np.dot(np.linalg.inv(np.dot(XT, X)), XT)

# Model without all regressors
XM = X.copy()
indexes2 = []
for n in range(1, N):
    for s in range(1, S):
        indexes2.append(n * S + s)
XM = np.delete(XM, indexes2, axis=1)
XTM = np.transpose(np.conjugate(XM))
XXXM = np.dot(np.linalg.inv(np.dot(XTM, XM)), XTM)

coef1 = (S - 1) * (N - 1)
coef2 = N * S * (L - 1)
coef3 = N * S * L - 1
# Calculate statistics Z (Z_a, Z_p)
np.seterr(invalid='ignore') # because of possible zero
division in Z (S_e = 0)
Z_a, R2_a, R2a_a, b_a = z_calculate(Y_a, X, XXX, XM, XXXM,
    coef1, coef2, coef3, indexes2)
if sett_more_variants.mac_name == "FDAC":
    Z_p, R2_p, R2a_p, b_p = z_calculate(Y_p, X, XXX, XM, XXXM,
        coef1, coef2, coef3, indexes2, angle_recalculation)
else:
    b_p = None
np.seterr(invalid='warn')

...

```

```

np.seterr(invalid='ignore')
p_value_a = f_distribution.sf(Z_a, coef1, coef2, loc=0,
    scale=1)
if sett_more_variants.mac_name == "FDAC":
    ...
np.seterr(invalid='warn')

if pairing and mmv_calculate_matrices:
    ...
else:
    if mmv_calculate_matrices:
        p_value_a = np.reshape(p_value_a, (R1, R2))
        R2_a = np.reshape(R2_a, (R1, R2))
        R2a_a = np.reshape(R2a_a, (R1, R2))
        if sett_more_variants.mac_name == "FDAC":
            ...
            freq = np.array([frequencies[0],
                frequencies[sett_more_variants.number_of_data - 1]])
        if alpha_plots:
            ...
    else:
        ...

return mask_a, mask_p, mask_ap, p_value_a, p_value_p,
    p_value_ap, R2_a, R2a_a, R2_p, R2a_p, freq

```

Appendix A

Matrices X and Y

Schemes of matrices X and Y for the Python program code follows (see Fig. A.1, A.2).

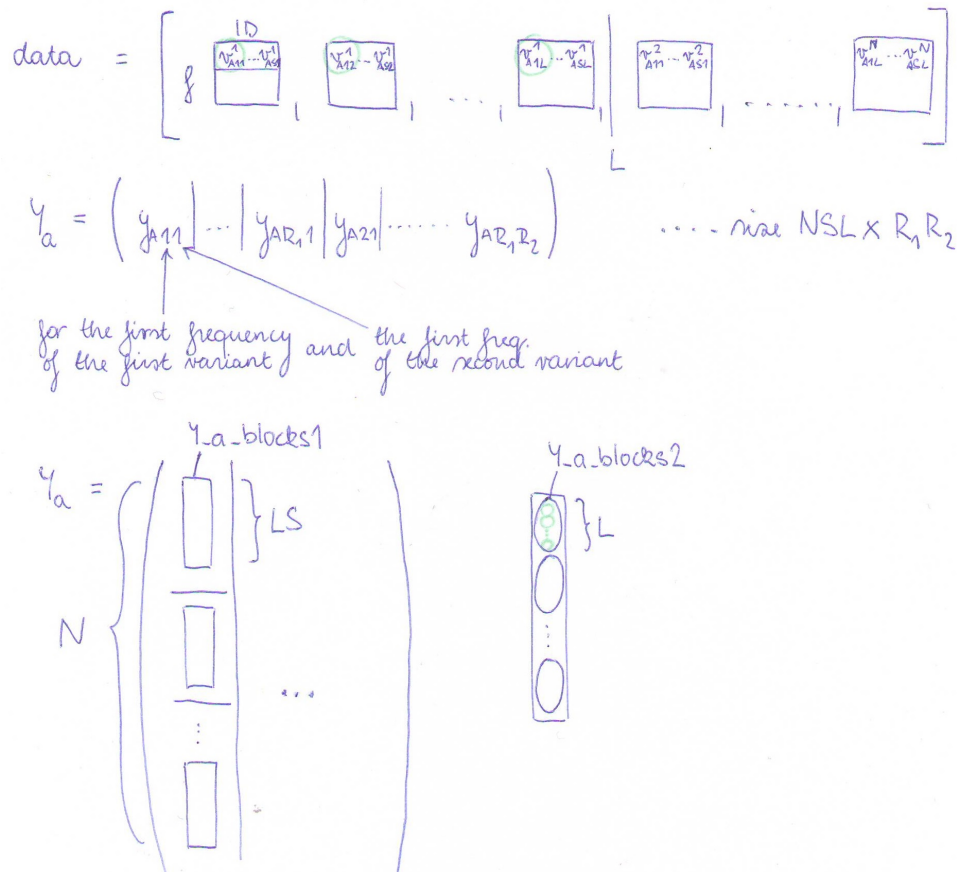


Figure A.1: A scan of a scheme for matrix Y construction.

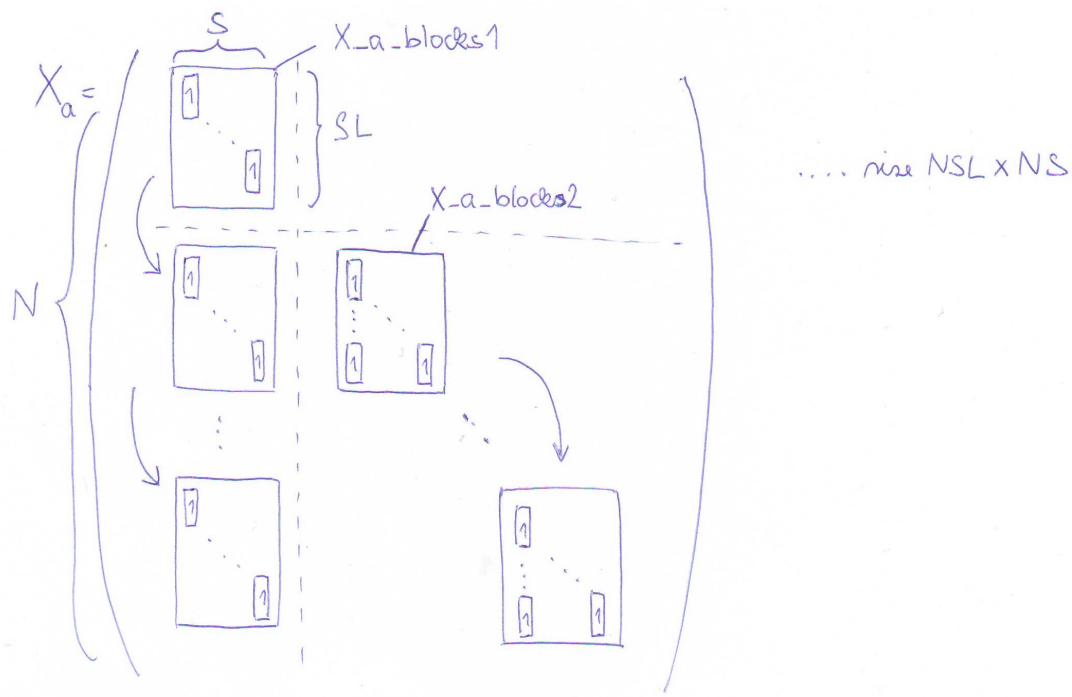


Figure A.2: A scan of a scheme for matrix X construction.

Nomenclature

$M, D, K,$	mass, damping and stiffness matrices,
$t,$	time,
$\omega,$	frequency,
$i,$	imaginary unit,
$v,$	modal vector (mode shape),
$V, \Lambda,$	modal and spectral matrices,
$I,$	identity matrix (unit matrix),
$R^2,$	coefficient of determination,
$\text{diag}(a),$	diagonal matrix with vector a on a diagonal,
$A^*, a^*,$	conjugate transpose (Hermitian conjugate) of matrix A or vector a (a transpose in case of A or a real),
$A^{-1},$	inversion of matrix $A,$
$A^+,$	Moore–Penrose pseudoinversion of matrix $A,$
$\text{cond}(\cdot),$	condition number,
$\sigma_{\max}(\cdot),$	the largest singular value,
$\sigma_{\min}(\cdot),$	the smallest singular value,
$\mathcal{L}(\cdot),$	Laplace transform,
$ \cdot ,$	absolute value,
$\arg(\cdot)$	the principal value of argument of complex number,
$\ \cdot\ ,$	L^2 (Frobenius) norm,
$\ \cdot\ _S,$	spectral norm (induced 2-norm),
$\ \cdot\ _1,$	induced (operator, Schatten) 1-norm,
$\ \cdot\ _\infty,$	induced (operator, Schatten) ∞ -norm,
$\mathbb{N}, \mathbb{R}, \mathbb{C},$	natural, real and complex numbers,
$E X,$	expected value of a random variable $X,$
$\text{var } X,$	variance of a random variable $X,$
$P X,$	probability of a random variable $X,$
$X \sim N(\mu, \sigma^2),$	a random variable X has a normal probability distribution with a mean μ and a variance $\sigma^2,$
$X \sim F(\nu_1, \nu_2),$	a random variable X has a F probability distribution (Fisher distribution) with parameters ν_1 and $\nu_2,$
$X \sim \chi_\nu^2,$	a random variable X has a χ^2 probability distribution (chi-squared distribution) with a parameter $\nu,$
$\alpha,$	significance level,
DOF,	degrees of freedom.
ANOVA,	analysis of variance.

Matrix norms

For matrix A of type $R \times S$ with an element $a_{r,s}$ at position r, s (the r -th row, the s -th column) the matrix norms are given by

$$\begin{aligned}\|A\| &= \left(\sum_{r=1}^R \sum_{s=1}^S |a_{r,s}|^2 \right)^{\frac{1}{2}}, \\ \|A\|_S &= \sigma_{\max}(A), \\ \|A\|_1 &= \max_{s \in \{1, \dots, S\}} \sum_{r=1}^R |a_{r,s}|, \\ \|A\|_\infty &= \max_{r \in \{1, \dots, R\}} \sum_{s=1}^S |a_{r,s}|.\end{aligned}$$

Bibliography

- [1] S. Ziaei Rad, “Methods for Updating Numerical Models in Structural Dynamics,” M.S. Thesis, Imperial College London, 1997.
- [2] J. F. Mercer, G. S. Aglietti, M. Remedios and A. M. Kiley, “Study of Correlation Criteria for Spacecraft-Launch Vehicle Coupled Loads Analysis,” in *34th Conference and Exposition on Structural Dynamics of Multiphysical Systems (IMAC XXXIV)*, Orlando, FL, USA, January 2016.
- [3] R. J. Allemang, “The Modal Assurance Criterion (MAC): Twenty Years of Use and Abuse,” in *Sound and Vibration*, vol. 37, pp. 14–21, August 2003.
- [4] W. T. Thomson and M. D. Dahleh, *Theory of vibrations with applications*, 5th ed., Harlow: Pearson Education, 2014, p. 524, ISBN 9781292042718.
- [5] V. Zeman and Z. Hlaváč, *Kmitání mechanických soustav*, 2nd ed., Plzeň: ZČU, 2004, p. 224, ISBN 80-7043-337-X.
- [6] J. D. Tebbens, I. Hnětynková, M. Plešinger, Z. Strakoš and P. Tichý, *Analýza metod pro maticové výpočty: Základní metody*, 1st ed., MatfyzPress, 2012, p. 328, ISBN 978-80-7378-201-6.
- [7] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe and H. van der Vorst, *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, Philadelphia: SIAM, 2000, p. 430, ISBN 978-0-89871-471-5.
- [8] M. H. Richardson, “Is It a Mode Shape, or an Operating Deflection Shape?,” *Sound and Vibration Magazine*, vol. 31, March 1997.
- [9] M. López Aenlle, R. Brincker and A. Fernández Canteli, “Some Methods to Determine Scaled Mode Shapes in Natural Input Modal Analysis,” in *Proc. IMAC XXIII*, January 2005.
- [10] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, 1st ed., New York: Springer-Verlag, 2005, p. 340, ISBN 978-0-387-22073-4.
- [11] M. Fuhry and L. Reichel, “A New Tikhonov Regularization Method,” *Numerical Algorithms*, vol. 59, pp. 433–445, March 2012.
- [12] O. Cahueñas, L. Hernández-Ramos and M. Raydan, “Pseudoinverse preconditioners and iterative methods for large dense linear least-squares problems,” *Bulletin of Computational Applied Mathematics*, vol. 1, pp. 25–47, Jun 2013.

- [13] F. Soleymani, “On a Fast Iterative Method for Approximate Inverse of Matrices,” *Communications of the Korean Mathematical Society*, vol. 28, no. 2, pp. 407–418, 2013.
- [14] V. Y. Pan, F. Soleymani and L. Zhao, “Highly Efficient Computation of Generalized Inverse of a Matrix,” *arXiv preprint arXiv:1604.07893*, April 2016.
- [15] J. Anděl, *Matematická statistika*, 2nd ed., Praha: SNTL, 1985, p. 352, ISBN 04-003-85.
- [16] J. Reif, *Metody matematické statistiky*, 2nd ed., Plzeň: ZČU, July 2004, p. 288, ISBN 80-7043-302-7.
- [17] S. Trampitsch, “Complex-Valued Data Estimation: Second-Order Statistics and Widely Linear Estimators,” Masterarbeit, Alpen-Adria-Universität Klagenfurt, 2013.
- [18] A. G. Hussin, N. A. Abdullah and I. Mohamed, “A Complex Linear Regression Model,” *Sains Malaysiana*, vol. 39, no. 3, pp. 491–494, 2010.
- [19] P. Vacher, B. Jacquier and A. Bucharles, “Extensions of the MAC Criterion to Complex Modes,” in *Proc. ISMA International Conference on Noise and Vibration Engineering*, Leuven, Belgium, 2010.
- [20] L. Aigner, “Modal Assurance Criteria Value for Two Orthogonal Modal Vectors,” in *Proc. IMAC XVI*, 1997.
- [21] K. Scharnhorst, “Angles in Complex Vector Spaces,” *Acta Applicandae Mathematicae*, vol. 69, pp. 95–103, October 2001.
- [22] E. Deadman, N. J. Higham and R. Ralha, “Blocked Schur Algorithms for Computing the Matrix Square Root,” in *Applied Parallel and Scientific Computing, PARA 2012, Lecture Notes in Computer Science*, vol. 7782, pp. 171–182, 2013.
- [23] M. Barth, J. Otta, “Simulation of Modal Analysis by FEM,” ZF Friedrichshafen AG, November 2016
- [24] “Python 3.6.3 documentation: Floating Point Arithmetic: Issues and Limitations,” 1 November 2017. [Online]. Available: <https://docs.python.org/3.6/tutorial/floatpoint.html>.
- [25] F. Johansson and others. “mpmath: a Python library for arbitrary-precision floating-point arithmetic,” November 2017. [Online]. Available: <http://mpmath.org/>.
- [26] L. H. Negri. “PeakUtils documentation,” November 2017. [Online]. Available: <http://peakutils.readthedocs.io/en/latest>.
- [27] C. Van Harsen. “gmpy2 documentation,” November 2017. [Online]. Available: <http://gmpy2.readthedocs.io/en/latest/>.

- [28] B. Clapper, “munkers – Munkres implementation for Python,” 1 November 2017. [Online]. Available: <http://software.clapper.org/munkres/>.
- [29] P. Čermelj, M. Mršnik, M. Pirnat, J. Slavič, B. Starc. “openmodal/pyuff,” November 2017. [Online]. Available: <https://github.com/openmodal/pyuff>.
- [30] S. Astanin. “gregbanks/python-tabulate,” November 2017. [Online]. Available: <https://github.com/gregbanks/python-tabulate>.

