Bootstrap tests for model selection in robust vibration analysis of oscillating structures

Boris Kargoll, Mohammad Omidalizarandi, Jens-André Paffenholz, Ingo Neumann, Gaël Kermarrec, Hamza Alkhatib

Geodetic Institute, Leibniz University Hannover, Nienburger Str. 1, 30167 Hannover, Germany,
{kargoll,zarandi,paffenholz,neumann,kermarrec,alkhatib}@gih.uni-hannover.de

Key words: structural health monitoring; low-cost accelerometer; vibration analysis; damped harmonic oscillation; robust parameter estimation; model selection; bootstrap test

ABSTRACT

In this contribution, a procedure for deciding, whether the oscillation of a surveyed structure is damped or not, is proposed. For this purpose, two bootstrap tests under fairly general assumptions regarding auto-correlation and outlier-affliction of the random deviations ("measurement errors") are suggested. These tests are derived from an observation model consisting of (1) a parametric oscillation model based on trigonometric functions, (2) a parametric auto-correlation model in the form of an autoregressive process, and (3) a parametric stochastic model in terms of the heavy-tailed family of scaled t-distributions. These three levels, which generalize current observation models for oscillating structures, are jointly expressed as a likelihood function and jointly adjusted by means of a generalized expectation maximization algorithm. Closed-loop Monte Carlo simulations are performed to validate the bootstrap tests. Visual inspection of models fitted by standard least-squares techniques are shown to be insufficient to detect a small significant damped oscillation. Furthermore, the tests are applied to a controlled experiment in a laboratory environment. The oscillation was generated by means of a portable shaker vibration calibrator and measured by a reference accelerometer and a low-cost accelerometer.

I. INTRODUCTION

Models for periodic phenomena based on trigonometric functions have played an important role in geodesy for decades (Vanicek, 1969; Wells *et al*., 1985; Craymer, 1998; Pagiatakis, 1999; Mautz and Petrovic, 2005; Kaschenz and Petrovic, 2005; Psimoulis *et al*., 2008; Neitzel *et al*., 2011; Neitzel *et al*., 2012; Lehmann, 2014; Bogusz and Klos, 2016). Oscillations may be accurately measured in various ways, for instance, by means of a global navigation satellite system (GNSS) receiver, a ground based synthetic aperture radar (GBSAR), a terrestrial laser scanner (TLS), a laser tracker, an accelerometer and an image-assisted total station (cf. Neitzel and Schwarz, 2017; Omidalizarandi *et al*., 2018). The stochastic model employed for the inference of oscillation models greatly depends on the kind of observable.

Besides the variances of the random deviations, auto-correlations play a large role for such electronic instruments measuring at a high sampling rate and should therefore be taken into account in deformation analysis in general (cf. Kuhlmann, 2003). While covariance functions and covariance matrices have been traditionally employed for this purpose, autoregressive (AR) processes have become increasingly popular since the early 2000s in diverse fields of geodetic science (e.g., Schuh, 2003; Nassar *et al*., 2004; Park and Gao, 2008; Li, 2011; Luo *et al*., 2012). These studies demonstrated that AR models can be easily fitted to geodetic data in connection with least-squares estimation (LSE) or Kalman filtering techniques. While constituting a parsimonious time-domain model that can easily be combined with a functional observation model, AR processes can be transformed into and interpreted through a covariance function/matrix and a spectral density function. Therefore, the specification and fitting of an AR correlation model are adopted also in the current contribution.

Outliers constitute another phenomenon often found in geodetic data. A common approach to dealing with outliers consists of a robust (i.e., outlier-resistant) method of iteratively reweighted least squares (IRLS) based on an error law whose defining probability density function (pdf) has "longer tails" than a Gaussian bell curve (e.g., the L1-norm estimator and Huber estimator). Thus, when outliers are defined to be errors larger than three times the standard deviation of a random deviation (see Lehmann, 2013), they obtain substantially more probability mass via IRLS (reflected by lower weights) than with (uniformly weighted) LSE. To obtain estimation results that are as realistic as possible, it is desirable to match the shape of the pdf or the error law with the actual distribution of the residuals/outliers (cf. Wisniewski, 2014). To enable this, it makes sense to use a flexible family of error laws having suitable mathematical properties and having - besides a scale parameter that accounts for the variance of a random deviation - at least one shape parameter that controls the thickness of the tails. To foster automatization of the adjustment procedure, the

heavy-tailed family of scaled t-distributions is employed in this contribution. Its shape parameter/tuning constant is the degree of freedom, which can be estimated as part of an IRLS procedure (cf. Koch and Kargoll, 2013), so that this robust estimator has been called *self-tuning* (Parzen, 1979). Standard LSE corresponds to the special case, where the degree of freedom takes a large value.

Parameter estimation is an intermediate step that reduces the observations to sufficient statistics, which are then used to express the test statistic (cf. Kargoll, 2012). Under the standard assumptions of linear observation equations involving normally distributed random deviations and a given covariance or weight matrix, optimal (*uniformly most powerful invariant*) parameter tests are readily available (cf. Teunissen, 2003). As these assumptions do not hold under the given model, we make use of bootstrap tests (cf. McKinnon, 2007), which can be tailored to such nonstandard model assumptions. In combination with Monte Carlo (MC) simulation (cf. Koch, 2018), such tests do not require knowledge of the distribution of the test statistic employed. In Section II, we develop various bootstrap tests for deciding, whether an observed oscillation is damped or not. This methodology extends recent developments of bootstrapping techniques in geodesy for confidence intervals (Neuner et al., 2014), parameter estimation (e.g., Angrisano *et al.*, 2018) and uncertainty quantification (Lösler *et al.*, 2018), to the domain of hypothesis testing. In Section III, it is shown how MC simulations are carried out to estimate the type-I error probabilities and power functions of the bootstrap tests. The performance of these tests is assessed and contrasted with the standard F-Test. The limits of standard least-squares fitting and visual inspection of resulting model plots for the purpose of detecting a damped oscillation are also explored. Subsequently, the proposed tests are applied to measurements of an oscillation generated by a portable shaker vibration calibrator and recorded both by the associated reference accelerometer and a low-cost accelerometer. Section IV draws some conclusions and gives an outlook to potential applications of the tests.

II. METHODOLOGY

A. *Observation Model*

When observations describe a time-dependent undamped oscillation, one may use the model

$$h_t(\boldsymbol{\beta}) = \frac{a_0}{2} + \sum_{j=1}^{M} a_j \cos\left(2\pi f_j x_t\right) + b_j \sin\left(2\pi f_j x_t\right) \quad (1)$$

consisting of an unknown offset $\frac{a_0}{2}$ and a sum of sinusoids, which involve unknown coefficients a_j and b_j, unknown frequencies f_j, and specified equidistant time instances x_t for $t = 1, \dots, n$. The unknowns a_j, b_j, f_j $(j = 1, \dots, M)$ form the vector $\boldsymbol{\beta}$ of functional model parameters. In the case of a damped oscillation, the deterministic model (1) is extended to

$$h_t(\boldsymbol{\beta}, \boldsymbol{\xi}) = \frac{a_0}{2} + \sum_{j=1}^{M} \left[a_j \cos\left(2\pi f_j \sqrt{1-\xi_j^2}\, x_t\right) + b_j \sin\left(2\pi f_j \sqrt{1-\xi_j^2}\, x_t\right)\right] \times \exp\left(-2\pi \xi_j f_j x_t\right) \quad (2)$$

where ξ_j is the so-called damping ratio coefficient and $f_{jd} = f_j\sqrt{1-\xi_j^2}$ the corresponding damped frequency (cf. Amezquita-Sanchez and Adeli, 2015). Clearly, the undamped oscillation model (1) is nested inside the damped oscillation model (2) since the former results from the latter by setting all damping ratio coefficients equal to zero.

For both kinds of deterministic model, the observations l_t are considered to be subject to auto-correlated random deviations

$$e_t = \textstyle\sum_{j=1}^{p} \alpha_j e_{t-j} + u_t \quad (3)$$

The coefficients $\boldsymbol{\alpha} = [\alpha_1 \cdots \alpha_p]^T$ of this AR(p) model are treated as unknown parameters to take unknown forms of auto-correlation into account. The boundary conditions are simply fixed by setting $e_0 = e_{-1} = \cdots = e_{1-p} = 0$. The AR model order p has to be specified based on prior information or through a statistical model selection procedure. To set the level of precision and to model outliers of unknown absolute frequency and magnitudes, the white noise components $u_1, \dots, u_n$ are assumed to be stochastically independent and to follow a scaled, centred t-distribution with unknown scale factor σ^2 and degree of freedom ν, symbolically

$$u_t \sim t_\nu(0, \sigma^2). \quad (4)$$

This combined parametric auto-correlation and error model allows for a self-tuning, robust, maximum likelihood estimation of the parameters. This estimation is based on the log-likelihood function

$$\log L(\boldsymbol{\beta}, \boldsymbol{\xi}, \boldsymbol{\alpha}, \sigma^2, \nu; \boldsymbol{l}) = n \log\left[\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi\sigma^2}\,\Gamma\left(\frac{\nu}{2}\right)}\right] - \frac{\nu+1}{2} \times \textstyle\sum_{t=1}^{n} \log\left[1 + \frac{1}{\nu}\left(\frac{u_t}{\sigma}\right)^2\right] \quad (5)$$

where u_t is expressed through (3) and the observations equations as functions of $\boldsymbol{\beta}$, possibly $\boldsymbol{\xi}$, and $\boldsymbol{\alpha}$. The maximization of this log-likelihood function constitutes an alternative to a least-squares approach based on outlier-elimination and noise reduction through filtering. One advantage of the former approach is that a single algorithm and adjustment routine can be devised while avoiding some pre-processing steps. Furthermore, (5) allows for likelihood ratio (LR) tests about the model parameters or

constraints thereof, as the deterministic and stochastic model assumptions are included.

B. *Hypotheses*

Besides estimating model parameters, we wish to identify an adequate oscillation model from (1) and (2). Since parsimonious observation models are preferred over models that include unnecessary parameters, it is desirable to test the null hypothesis H_0 that the damping ratio coefficients $\boldsymbol{\xi}$ are equal to zero. The alternative hypothesis can simply be specified to be the negation of H_0, so that the problem is to test

$$H_0: \boldsymbol{\xi} = \mathbf{0} \qquad \text{versus} \qquad H_1: \boldsymbol{\xi} \neq \mathbf{0}. \tag{6}$$

A simple case is given by testing the damping ratio coefficient ξ associated with a single natural frequency of the oscillating structure (the vector $\boldsymbol{\xi}$ thus reduces to a scalar quantity).

One possibility of measuring the deviation from H_0 is to compute the weighted square sum $T = \hat{\boldsymbol{\xi}}^T \widehat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\xi}}\hat{\boldsymbol{\xi}}}^{-1} \hat{\boldsymbol{\xi}}/M$, based on estimates $\hat{\boldsymbol{\xi}}$ and their joint *a posteriori* covariance matrix $\widehat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\xi}}\hat{\boldsymbol{\xi}}}$. In the special case of uncorrelated and normally distributed random deviations (corresponding to $p = 0$ and $\nu \to \infty$), the kind of Wald (W) test statistic T would approximately follow Fisher's $F_{M,n-u}$-distribution, where $u = 4M + 1$ is the total number of functional parameters $\boldsymbol{\beta}, \boldsymbol{\xi}$ in model (2). However, due to the stochastic model based on the family of scaled t-distributions, the test statistic T might not follow any standard distribution to a sufficient level of approximation. This hindrance, however, does not prevent the solution of the testing problem (6) since simulation-based bootstrap tests do not require the specification of a test distribution. A second natural test statistic is based on the difference

$$LR = \log L(\tilde{\boldsymbol{\beta}}, \mathbf{0}, \tilde{\boldsymbol{\alpha}}, \tilde{\sigma}^2, \tilde{\nu}; \boldsymbol{l}) - \log L(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\xi}}, \hat{\boldsymbol{\alpha}}, \hat{\sigma}^2, \hat{\nu}; \boldsymbol{l})$$

of the log-likelihoods at the constraint and unconstraint maximum likelihood estimates. Like the test statistic T, the LR test statistic $T_{LR} = -2\ LR$ (cf. Section 2.5.6 in Kargoll, 2012) generally does not have a standard distribution, so that it will also be carried out by means of bootstrapping.

C. *The Bootstrap Tests*

The idea of a bootstrap test is to generate a large number B of observation vectors under the true H_0, to compute the B values that the test statistic takes for these generated measurement series, and to check whether the value of the test statistic obtained for the actual measurement results is extremely large in comparison to the test values obtained under H_0. These comparisons replace the comparison with a critical value derived from a fully specified test distribution (which is unknown in the present situation). As with a classical hypothesis test, the significance level α may be fixed in advance. Then, given a vector $\boldsymbol{l}$ of measurement results, the order M of the oscillation model (1)/(2) and the order p of the AR model (3), the following steps can be carried out in order to arrive at the test decision (see also Fig. 1.

1. Estimation step: The generalized expectation maximization (GEM) algorithm described in Alkhatib *et al*. (2018) is used to compute:
   a) Parameter estimates $\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\xi}}, \hat{\boldsymbol{\alpha}}, \hat{\sigma}^2, \hat{\nu}$, covariance matrix $\widehat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\xi}}\hat{\boldsymbol{\xi}}}$ and white noise components $\hat{u}_1, \dots, \hat{u}_n$ in the damped harmonic oscillation model (2). The nonlinear model is linearized within each GEM iteration (see the Appendix for the derivation of the Jacobi matrix);
   b) Parameter estimates $\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\alpha}}, \tilde{\sigma}^2, \tilde{\nu}$ and white noise components $\tilde{u}_1, \dots, \tilde{u}_n$ in the undamped harmonic oscillation model (1). The Jacobi matrix of this nonlinear model is obtained from the Jacobi matrix used in a) by setting $\boldsymbol{\xi} = \mathbf{0}$.
2. Testing step: The test statistic $T = \hat{\boldsymbol{\xi}}^T \widehat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\xi}}\hat{\boldsymbol{\xi}}}^{-1} \hat{\boldsymbol{\xi}}/M$ or $T_{LR} = -2\ LR$ is computed.
3. Simulation step: Firstly, B white noise vectors

$$\begin{aligned} \boldsymbol{u}^{(1)} &= \left[u_1^{(1)} \cdots\ u_n^{(1)}\right]^T \\ &\vdots \\ \boldsymbol{u}^{(B)} &= \left[u_1^{(B)} \cdots\ u_n^{(B)}\right]^T \end{aligned}$$

   are generated in one of the following two variants:
   a) *Parametric bootstrapping*: each white noise component $u_t^{(k)}$ $(t = 1, \dots, n; k = 1, \dots, B)$ is generated randomly from the fitted t-distribution $t_{\hat{\nu}}(0, \hat{\sigma}^2)$.
   b) *Nonparametric bootstrapping*: each $u_t^{(k)}$ is "randomly drawn" from the fitted white noise series $\hat{u}_1, \dots, \hat{u}_n$ "with replacement" by generating random numbers $\tau_t^{(k)}$ from the discrete uniform distribution $U(1, n)$. Each number $\tau_t^{(k)}$ defines an index, and the associated value $\hat{u}_{\tau_t^{(k)}}$ of the fitted white noise series defines the newly generated white noise component $u_t^{(k)} = \hat{u}_{\tau_t^{(k)}}$.

   Secondly, the white noise series are inserted into the fitted AR models to generate the coloured noise components $e_t^{(k)} = \sum_{j=1}^{p} \hat{\alpha}_j e_{t-j}^{(k)} + u_t^{(k)}$ based on the initial values $e_0^{(k)} = \cdots = e_{1-p}^{(k)} = 0$, resulting in the vectors

$$\begin{aligned} \boldsymbol{e}^{(1)} &= \left[e_1^{(1)} \cdots\ e_n^{(1)}\right]^T \\ &\vdots \\ \boldsymbol{e}^{(B)} &= \left[e_1^{(B)} \cdots\ e_n^{(B)}\right]^T \end{aligned}$$

   Thirdly, these realizations of the random deviations are added to the deterministic model

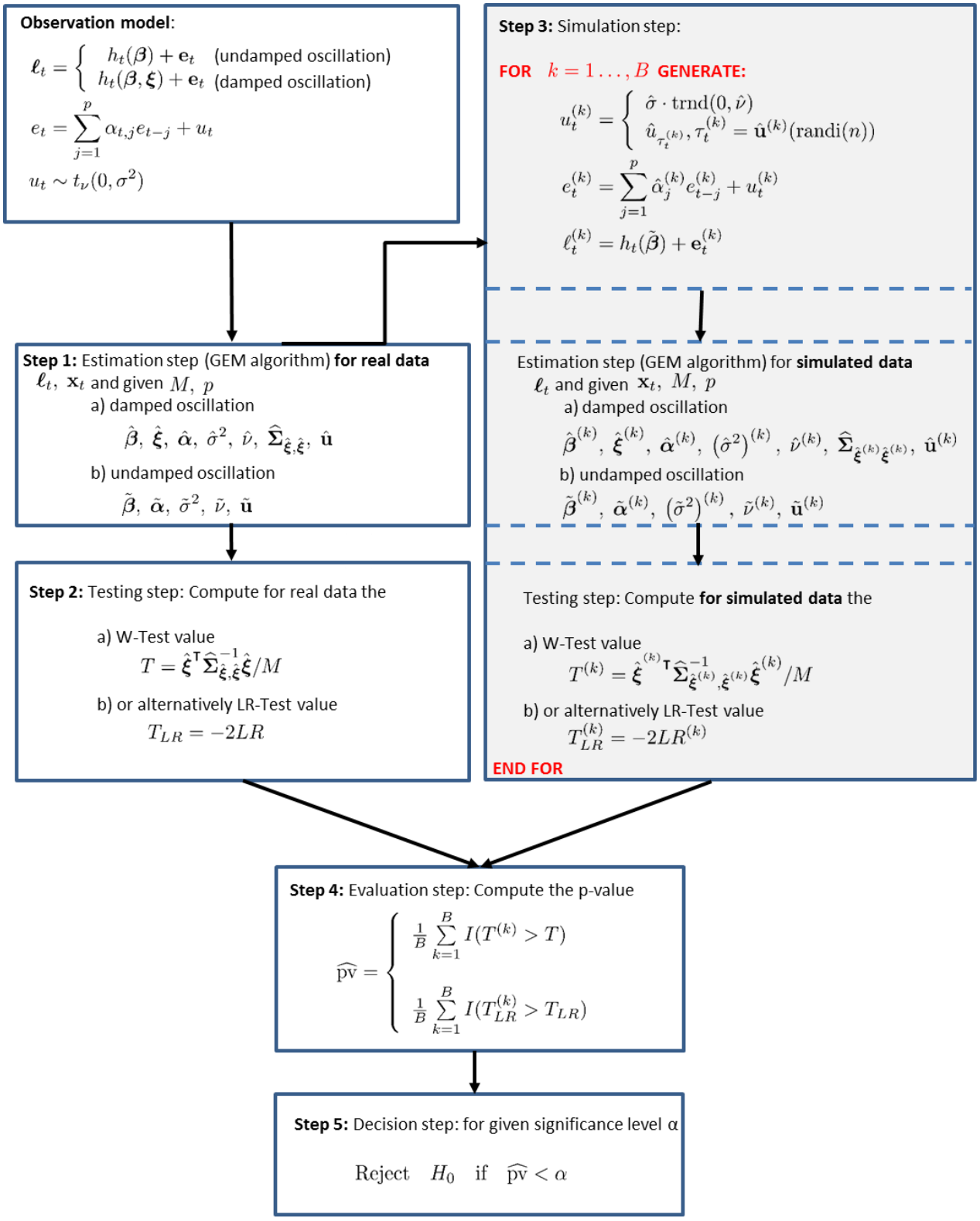

Figure 1. Flowchart of the computational steps of the bootstrap tests described in Section II.C. The random number generation within the simulation step is shown here by employing the MATLAB routines *trnd* (t-distribution) and *randi* (discrete uniform distribution).

fitted under H_0, that is, $l_t^{(k)} = h_t(\tilde{\boldsymbol{\beta}}) + e_t^{(k)}$, thereby producing the observation vectors

$$\boldsymbol{l}^{(1)} = \left[l_1^{(1)} \cdots\ l_n^{(1)}\right]^T$$
$$\vdots$$
$$\boldsymbol{l}^{(B)} = \left[l_1^{(B)} \cdots\ l_n^{(B)}\right]^T$$

The aforementioned GEM algorithm is then used to compute, for every $\boldsymbol{l}^{(k)}$, the parameter estimates $\hat{\tilde{\boldsymbol{\xi}}}^{(k)}$ and covariance matrix $\hat{\boldsymbol{\Sigma}}_{\hat{\tilde{\xi}}^{(k)}\hat{\tilde{\xi}}^{(k)}}$ required for evaluating the test statistic via $T^{(k)} = \hat{\tilde{\boldsymbol{\xi}}}^{(k)T}\hat{\boldsymbol{\Sigma}}^{-1}_{\hat{\tilde{\xi}}^{(k)}\hat{\tilde{\xi}}^{(k)}}\hat{\tilde{\boldsymbol{\xi}}}^{(k)}$. The algorithm also outputs $\log L(\hat{\tilde{\boldsymbol{\beta}}}^{(k)}, \mathbf{0}, \tilde{\boldsymbol{\alpha}}^{(k)}, \tilde{\sigma}^{2(k)}, \tilde{v}^{(k)}; \boldsymbol{l})$ and $\log L(\hat{\tilde{\boldsymbol{\beta}}}^{(k)}, \hat{\tilde{\boldsymbol{\xi}}}^{(k)}, \hat{\boldsymbol{\alpha}}^{(k)}, \hat{\sigma}^{2(k)}, \hat{v}^{(k)}; \boldsymbol{l})$ with respect to the adjusted constraint and unconstraint model, necessary for the determination of $T_{LR}^{(k)}$.

4. Evaluation step: The p-value is estimated by computing (cf. McKinnon, 2007, Section 2)

$$\widehat{\mathrm{pv}} = \frac{1}{B}\sum_{k=1}^{B} I\left(T^{(k)} > T\right),$$

or

$$\widehat{\mathrm{pv}} = \frac{1}{B}\sum_{k=1}^{B} I\left(T_{LR}^{(k)} > T_{LR}\right),$$

where the indicator function $I(.)$ takes the value 1 if the argument is true and the value 0 if the argument is false. Note that in case H_0 is false, the test value T tends to be large. As the test values generated under the true H_0 tend to be small, the indicator function tends to take the value 0, so that the p-value tends to be small in this case.

5. Decision step: H_0 is rejected if $\widehat{\mathrm{pv}} < \alpha$.

III. Results

A. *Simulation of the Bootstrap Tests*

Every statistical test is characterized by the complementary type-I and type-II error rates, which constitute its primary quality measures. On the one hand, it is important to know if the specified significance level α is truly reflected by the actual type-I error rate for the described bootstrap test. On the other hand, it is useful to study the power function of that test in order to obtain an impression of its sensitivity. Both measures can be estimated via Monte Carlo simulation.

To analyse the empirical type-I and type-II error rates for the bootstrap tests developed in Section II, Monte Carlo simulations based on $R = 500$ and $R = 1{,}000$ samples of $n = 1{,}000$ observations were carried out. For this purpose, firstly white noise samples

$$\boldsymbol{u}^{(1)} = \left[u_1^{(1)} \cdots\ u_{10{,}000}^{(1)}\right]^T$$
$$\vdots$$
$$\boldsymbol{u}^{(R)} = \left[u_1^{(R)} \cdots\ u_{10{,}000}^{(R)}\right]^T$$

were generated, on the one hand, using the t-distribution (4) with parameter values $\sigma = 0.001$ and $\nu = 3$, on the other hand using the centred normal distribution $N(0, \sigma^2)$ with the same scaling. Note that the vectors are 10 times longer than the actual number of observations. The reason for this is to eliminate the so-called warm-up effect created by the initial conditions (i.e., the zero values for time index values $t = 0, -1, \ldots$) in the recursive computation of the coloured noise samples by means of (3). The last 1,000 values of the generated coloured noise vectors thus truly reflect the characteristics of the AR process. The indexes 9,001…10,000 of these 1,000 values are shifted by -9,000 in order to obtain the indexing $t = 1, \ldots, n$ as defined for the models (1) – (4). Thus, the generated coloured noise vectors are denoted by

$$\boldsymbol{e}^{(1)} = \left[e_1^{(1)} \cdots\ e_{1{,}000}^{(1)}\right]^T$$
$$\vdots$$
$$\boldsymbol{e}^{(R)} = \left[e_1^{(R)} \cdots\ e_{1{,}000}^{(R)}\right]^T$$

In this simulation study, an AR(1) process with parameter value $\alpha_1 = 0.6828$ was applied. A set of true deterministic models was defined to consist of the oscillation model (2) with the parameter values $a_0 = 0.0016$, $a_1 = 0.0572$, $b_1 = -0.0950$, $f_1 = 16$ [Hz] and $\xi_1 = i \cdot 10^{-9}$ $(i \in \{0, \ldots, 51\})$ for the time instances $x_t = 67.6813\ [\mathrm{s}] + (t-1) \cdot 0.00512[\mathrm{s}]$. Adding the resulting true observations to the previously generated coloured noise vectors gave the observation samples

$$\boldsymbol{l}^{(1)} = \left[l_1^{(1)} \cdots\ l_{1{,}000}^{(1)}\right]^T$$
$$\vdots$$
$$\boldsymbol{l}^{(R)} = \left[l_1^{(R)} \cdots\ l_{1{,}000}^{(R)}\right]^T$$

under

- the two MC sample sizes $R = 500$ and $R = 1{,}000$,
- the two sample distributions $t_3(0, 0.001^2)$ and $N(0, 0.001^2)$, and
- the various damping ratio coefficient values ξ_1 listed before.

Note that the case $\xi_1 = 0$ $(i = 0)$ corresponds to the simulation of a true H_0, whereas the other non-zero values for ξ_1 (when $i \in \{1, \ldots, 51\}$) are associated with a true H_1.

Three tests were applied to all these observation samples at a significance level of $\alpha = 0.05$ under various settings:

1. both the parametric and the nonparametric bootstrap test based on the LR statistic $T_{LR}^{(k)}$ ("BS LR-Test") sampled $B = 99$ and $B = 999$ times,

2. both the parametric and the nonparametric bootstrap test based on the W statistic $T^{(k)}$ ("BS W-Test") for $B = 99$ and $B = 999$ times, and
3. the standard "F-Test" based on the previous W statistic with the distributional assumption $T \sim F_{1,995}$. With this test, the presence of the AR process and the non-normality of the white noise are ignored, so that the assumed redundancy is $n - (4M + 1) = 995$ with $M = 1$.

The odd choice of 99 or 999 bootstrap samples is motivated by the requirement of a Monte Carlo test that $\alpha(B + 1)$ is an integer (cf. McKinnon. 2007). To keep the computational burden of the Monte Carlo simulations manageable, the parameter ν was fixed at the true value 3 within the GEM algorithm. This yields a robust (though not self-tuning) estimator. The choice of the low degree of freedom $\nu = 3$ is also in line with certain applications of the Guide to the Expression of Uncertainty in Measurement (ISO/IEC, 2008), when the observables are explained by input quantities having type-A (i.e., statistically determined) standard uncertainties (cf. Sommer and Siebert, 2004). For the samples generated by means of the normal distribution, the GEM algorithm was run with the fixed degree of freedom $\nu = 120$, as this value leads to a close approximation of the sampling normal distribution $N(0, 0.001^2)$ by the t-distribution $t_{120}(0, 0.001^2)$ (cf. Koch, 2017).

Now, counting for each application of a test throughout the R MC runs the number of times that H_0 is rejected and dividing that number by R yields an estimate of the test's

1. type-I error rate if H_0 is true (i.e., if the MC simulation was carried out with $\xi_1 = 0$).
2. power if H_1 is true (i.e., if the MC simulation was carried out with $\xi_1 > 0$).

Table 1 shows that the empirical type-I error rate of the F-Test greatly differs from the specified significance level $\alpha = 0.05$ in case of Student white noise correlated by the AR(1) model. In contrast, the significance level is reproduced by the two bootstrap tests rather well already for $B = 99$ bootstrap samples. Increasing to $B = 999$ samples results in a correct second digit of type-I error rate. This confirms the finding of McKinnon (2007, p.3) that "it might be dangerous to use a value of B less than 999." Evidently, increasing in addition the number of MC runs does not further improve the approximation. Furthermore, the results for parametric and nonparametric bootstrapping are surprisingly similar, in view of their fundamentally different ways of generating the white noise series. The former leads to slightly closer approximations of α than the latter.

Table 1. Bold numbers: Average rejections of H_0 as estimates of the respective test's type-I error rate. Regarding the bootstrap (BS) Wald (W-) and likelihood ratio (LR-) Test, the numbers of the first row correspond to nonparametric (np), of the second row to parametric (p) bootstrapping.

R	500		1000	
F-Test	**0.006**		**0.007**	
B	99	999	99	999
BS W-Test: np	**0.044**	**0.052**	**0.061**	**0.053**
p	**0.048**	**0.052**	**0.042**	**0.051**
BS LR-Test: np	**0.044**	**0.052**	**0.061**	**0.054**
p	**0.046**	**0.052**	**0.043**	**0.051**

Generating Gaussian white noise without auto-correlations leads to the power functions shown in Fig. 1. As the graphs for the parametric and the nonparametric bootstrap tests coincide, only the latter are displayed. Figure 2 (top) demonstrates that all three tests correctly reproduce $\alpha = 0.05$ when H_0 is true (i.e., for the vanishing damping ratio coefficient $\xi_1 = 0$). For most $\xi_1 > 0$ the F-Test has slightly larger power than the bootstrap tests. As the F-Test is known to be *uniformly most powerful invariant* for linear models, the bootstrap tests could be expected to outperform the F-Test as the distribution of the F-Test statistic is not exact for a nonlinear model, as given by the damped harmonic oscillation model (2). Increasing the number of Monte Carlo runs or bootstrap samples might improve the performance of the bootstrap tests in this regard. More importantly, the bootstrap tests clearly outperform the F-Test in power when the Student distribution and the AR(1) model are used to generate the random deviations (see Fig. 2, bottom). The bootstrap W- and LR-Test again produce very similar results, which is not obvious since the standard Wald and likelihood ratio test are known to be equivalent mainly for linear models with Gaussian errors.

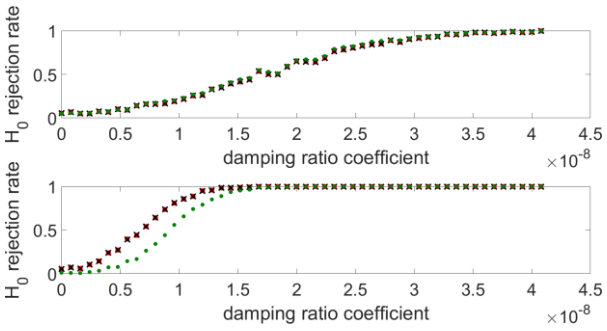

Figure 2. Average rejections of H_0 as estimates of the type-I error rate (for damping ratio coefficient $\xi_1 = 0$) and power (for $\xi_1 > 0$) for the F-Test (green dots), the bootstrap W-Test (dark crosses) and the bootstrap LR-Test (red dots). Top subplot: The bootstrap samples are generated with Gaussian white noise and without AR model (top subplot), alternatively with Student white noise and an AR(1) model (bottom subplot).

B. *Failure of Least Squares Curve Fitting and Visual Inspection of Model Plots*

Taking the bootstrap tests as the reference methods for detecting a damped oscillation, the performance of a simpler testing procedure based on LSE and visual inspection of the fitted oscillation model is now explored. For this purpose, two different damped oscillation models (2) were defined for 100 s of data with a sampling rate of 100 Hz using the parameter values $a_0 = 0, a_1 = 4.0, b_1 = -3.0, f_1 = 5$ [Hz] as well as the two alternative damping ratio coefficients (i) $\xi_1 = 1 \cdot 10^{-6}$ and (ii) $\xi_1 = 1 \cdot 10^{-5}$. Such small values were selected because the focus of the current study is on the reliable detection of a significant damping. In real applications, these levels could be larger, and the purpose of the test procedure would then be to test whether the damping ratio coefficient differs significantly from the permissible level available, e.g., from existing ISO standards. Note that the choice of the coefficient values a_1 and b_1 corresponds to the amplitude $A = \sqrt{a_1^2 + b_1^2} = 5.0$ and the phase angle $\varphi = \text{atan2}\left(\frac{-b_1}{a_1}\right) \cdot \frac{180}{\pi} = 36.87$ [degree]. The true observations resulting from this model were added to a white noise path generated from a scaled t-distribution with parameter values $\sigma = 0.2$ and $\nu = 4$.

In order for the LSE based on the linearization of the highly non-linear functional model (2) to converge, a precise initial value in particular of the frequency should be given. It is well known that the determination of oscillations with unknown frequencies is a challenging task requiring generally global optimization (cf. Mautz, 2001; Mautz and Petrovic, 2005). In Tables 2 and 3 it can be seen that LSE converges for the initial frequency value 5.005 [Hz] but diverges for the initial values 5.2 [Hz] and 5.3 [Hz]. MATLAB's curve fitting routine *fit*, which was applied using the robust fitting option since the simulated t-distribution gives ride to outliers, yields estimates similar to LSE for the initial frequency value 5.005 [Hz]; the estimates of the amplitude, phase angle and damping ration coefficients differ greatly from the true values for the two larger initial frequency values. The unreliable LSE can be improved to some extent by decreasing the step size in the computation of the parameter update. This Gauss-Newton method is also employed within the aforementioned GEM algorithm (applied in the estimation step of the bootstrap tests); see Alkhatib *et al*. (2018). The results in Tables 2 – 3 show that LSE with decreased step size converges for the initial frequency values 5.005 [Hz] and 5.2 [Hz], but produces strongly distorted estimates for the initial value 5.3 [Hz]. In contrast, the GEM algorithm, which includes the fitting of an AR error process and of a scaled t-distribution, approximates the true solution precisely for all three initial values. Although the generated observations contain only white noise, AR model orders of 1 or 2 were identified through the application of the white noise test described in Kargoll et al. (2018a). Thus, the parameter estimates by the GEM algorithm are not distorted by the additional low-order AR model estimation. Applying the bootstrap tests to the reference solution produced by the GEM algorithm results yields the results that the damping ratio coefficient $\xi_1 = 1 \cdot 10^{-6}$ is not significant (i.e., H_0 is accepted) whereas $\xi_1 = 1 \cdot 10^{-5}$ is significant (i.e., H_0 is rejected). In the cases where the LSE and the Gauss-Newton method both converge they produce identical oscillation models (shown in Fig. 3). Clearly, the significant damping by $\xi_1 = 1 \cdot 10^{-5}$ cannot be detected by visual inspection of the corresponding model plot (Fig. 3, bottom). It would not help either to zoom into a smaller time window since it needs to be analysed whether the entire segment analysed is damped or not. It may therefore be concluded that the employed curve fitting tool and the least squares methods, in connection with model plots, are inferior to a rigorous hypothesis test for damping such as the proposed bootstrap tests.

Table 2. Estimated parameters of the damped oscillation model with $\xi_1 = 1 \cdot 10^{-6}$ by three methods for three different initial frequency values. `NaN' indicates divergence of the iterations.

Parameters	f_1 [Hz]	A [-]	φ [degree]	ξ_1 [%]
True values	**5.0**	**5.0**	**36.87**	**0.0001**
Initial values	5.005	-	-	0.0000
MATLAB-fit	5.0029	5.0029	36.92	0.00012
LSE	4.9999	5.0008	36.95	0.00008
Gauss-Newton	4.9999	5.0008	36.95	0.00008
GEM alg.	4.9999	5.0043	36.92	0.00013

Parameters	f_1 [Hz]	A [-]	φ [degree]	ξ_1 [%]
True values	**5.0**	**5.0**	**36.87**	**0.0001**
Initial values	5.2	-	-	0.0000
MATLAB-fit	5.009	3.9039	-30.80	0.1357
LSE	NaN	NaN	NaN	NaN
Gauss-Newton	4.9999	5.0008	36.95	0.00008
GEM alg.	4.9999	5.0043	36.92	0.00013

Parameters	f_1 [Hz]	A [-]	φ [degree]	ξ_1 [%]
True values	**5.0**	**5.0**	**36.87**	**0.0001**
Initial values	5.3	-	-	0.0000
MATLAB-fit	5.028	2.8375	-37.95	0.2623
LSE	NaN	NaN	NaN	NaN
Gauss-Newton	5.1074	0.0000	-137.27	-1.5645
GEM alg.	4.9999	5.0043	36.92	0.00013

Table 3. Estimated parameters of the damped oscillation model with $\xi_1 = 1 \cdot 10^{-5}$ by three methods for three different initial frequency values. `NaN' indicates divergence of the iterations.

Parameters	f_1 [Hz]	A [-]	φ [degree]	ξ_1 [%]
True values	**5.0**	**5.0**	**36.87**	**0.001**
Initial values	5.005	-	-	0.0000
MATLAB-fit	5.0000	5.0022	36.90	0.00102
LSE	4.9999	4.9986	36.91	0.00099
Gauss-Newton	4.9999	4.9986	36.91	0.00099
GEM alg.	4.9999	5.0022	36.90	0.0010

Parameters	f_1 [Hz]	A [-]	φ [degree]	ξ_1 [%]
True values	**5.0**	**5.0**	**36.87**	**0.001**
Initial values	5.2	-	-	0.0000
MATLAB-fit	5.013	3.5886	-32.52	0.1412
LSE	NaN	NaN	NaN	NaN
Gauss-Newton	4.9999	4.9986	36.91	0.00099
GEM alg.	4.9999	5.0022	36.90	0.0010

Parameters	f_1 [Hz]	A [-]	φ [degree]	ξ_1 [%]
True values	**5.0**	**5.0**	**36.87**	**0.001**
Initial values	5.3	-	-	0.0000
MATLAB-fit	5.028	2.8496	-37.58	0.2608
LSE	NaN	NaN	NaN	NaN
Gauss-Newton	5.1255	0.0000	-61.15	-1.6032
GEM alg.	4.9999	5.0022	36.90	0.0010

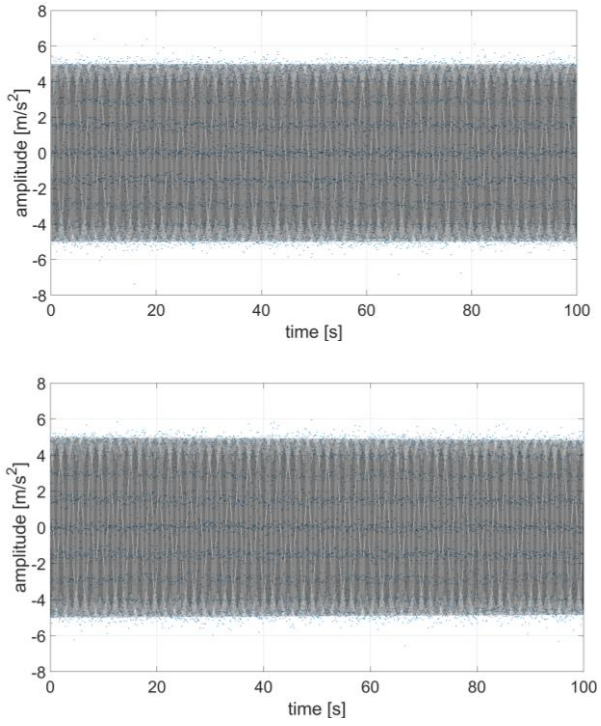

Figure 3. Oscillation time series of simulated acceleration data (blue dots) with damping ratio coefficient $\xi = 1 \cdot 10^{-6}$ (top) and $\xi = 1 \cdot 10^{-5}$ (bottom) and the model computed by converged LSE/Gauss-Newton method using the initial frequency value 5.005 [Hz].

C. *Controlled Excitation Experiment*

In the context of short-term deformation analysis of oscillating structures such as bridges, a geo-sensor network of low-cost accelerometers can be utilised for an accurate and robust vibration analysis of the structure (cf. Neitzel *et al.*, 2012). For this purpose, a proper deterministic model needs to be identified to truly characterise the global behaviour of a bridge structure such as natural frequencies, mode shapes and modal damping. Therefore, the time-dependent undamped oscillation model (1) considered by Kargoll *et al*. (2018a,b) and Omidalizarandi *et al*. (2018) is extended to the time-dependent damped harmonic oscillation model (2). In these previous studies, it was demonstrated in particular that the frequencies and amplitudes can be estimated robustly and accurately. Concerning the amplitudes, which are of great importance for a subsequent mode shape analysis, it should be noted that their estimation is directly influenced by the choice of the deterministic model. In the context of structural health monitoring, it is desirable to test whether the damping of a structure has a certain level or not, since deviations from that level would indicate a deterioration of the structure's intactness. As this level is usually non-zero, the hypotheses (6) and the test statistics should be adapted accordingly.

As a preparation for such more general testing problems, we consider in this contribution the simpler case of testing whether an oscillation with a single natural frequency is damped or not. For this purpose, a controlled excitation experiment was performed at the Institute of Concrete Construction of Leibniz University Hannover, using a portable shaker vibration calibrator (PSVC) 9210D and a low-cost accelerometer. The PSVC comes along with a highly accurate reference accelerometer of type PCB ICP quartz, which is used for the validation. The acceleration data are acquired by the low-cost accelerometer of type BNO055 (Bosch Sensortec) and reference accelerometer with an oscillation frequency of 4 Hz and sampling rates of 100 Hz and 200 Hz, respectively. For further information concerning the experimental setup and the used sensors, the reader is referred to Omidalizarandi *et al*. (2018).

In this section, the hypotheses are one-dimensional cases of (6). To analyse damping behaviour throughout the measured time series, the low-cost and reference acceleration datasets were firstly divided into 45 segments of consecutive 1000 and 2000 observations, respectively (each spanning 10 s). The auto-correlations of the accelerometer measurements are modelled by means of AR processes, which have previously been found to be an adequate class of models for this purpose (see, Nassar *et al*., 2014; Park and Gao, 2008). The order of the AR process was assumed to be $p = 15$, based on the experience with previous analyses of the

datasets (see Kargoll *et al.*, 2018a). Two tests were applied to all these observation samples at a significance level of $\alpha = 0.05$, based on the LR statistic $T_{LR}^{(k)}$ ("BS LR-Test") and the W statistic $T^{(k)}$ ("BS W-Test"). For both tests, nonparametric bootstrapping based on $B = 99$ and $B = 999$ times was performed. The parameter ν was fixed at the value 4 within the GEM algorithm, reflecting the expectation of a moderate number of outliers in the data.

In view of the setup of the controlled experiment, we expected to find no significant damping of the oscillation. To verify this, the p-values of both tests applied to all segments of the low-cost and reference accelerometer sensor data were calculated. The results for 99 bootstrap samples were found to be similar to those for 999 samples, so that only the former are shown in Figs. 4 and 5. It can be seen that all p-values are greater than the significance level, so that the rejection rate is 0, i.e., H_0 is always accepted, so that there is no evidence for significant damping ratio coefficients. This finding demonstrates that the PSVC device including the reference accelerometer produces an oscillation at the specified frequency with no significant damping.

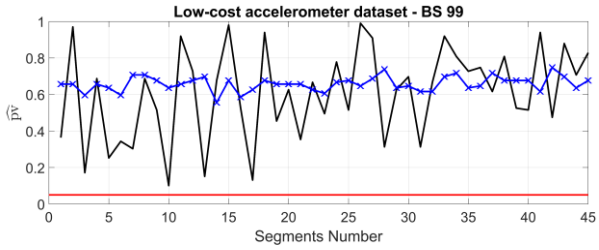

Figure 4. Low-cost accelerometer: The estimated p-value based on the bootstrap W-Test (black solid line) and the bootstrap LR-Test (blue cross line) corresponding to nonparametric bootstrapping with $B = 99$ samples. The red horizontal line shows the significance level of $\alpha = 0.05$.

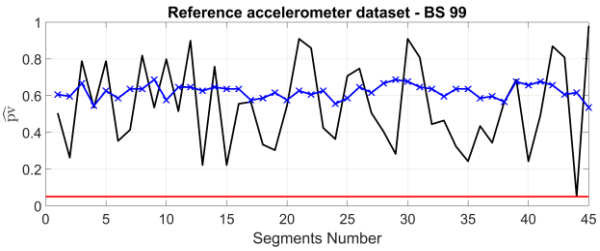

Figure 5. Reference accelerometer: The estimated p-value based on the bootstrap W-Test (black solid line) and the bootstrap LR-Test (blue cross line) corresponding to nonparametric bootstrapping with $B = 99$ samples. The red horizontal line shows the significance level of $\alpha = 0.05$.

IV. Conclusions and Outlook

A comprehensive observation model for a damped harmonic oscillation involving multiple frequencies, autoregressive random deviations and t-distributed white noise components can be adjusted by means of a GEM algorithm, which acts as a self-tuning robust estimator of all model parameters. Due to the intricacy of the model, test statistics in general do not have an exact standard distribution. Reasons for this are the non-linearity of the functional model (cf. Lehmann and Lösler, 2018), the non-normality of the estimator, and the non-normality of the random deviations. To test in particular whether the damping of an observed oscillation is significant or not, two bootstrap tests based on the well-known W and LR statistics were proposed since small significant damping is generally not visible to the eye in a plot of the oscillation model fitted by means of standard least-squares estimation or MATLAB's curve fitting tool. The bootstrap tests are carried out by means of randomly generated bootstrap samples, without resorting to critical values from a test distribution. The number of bootstrap samples is crucial. To reproduce the significance level of $\alpha = 0.05$ precisely by the empirical type-I error rate, it is recommended to generate at least 999 bootstrap samples. Both bootstrap tests have almost identical power functions, and it is also irrelevant whether parametric or nonparametric bootstrapping is carried out. The standard F-Test is slightly more powerful than the bootstrap tests when the random deviations are normally distributed and uncorrelated. In cases of Student and AR errors, however, the F-Test has an erratic type-I error rate and visibly reduced power, so that the bootstrap tests are clearly preferable in such situations. In the future, it is intended to extend the model selection procedure to determine whether certain sinusoids of within the damped harmonic oscillation model are significant or not. Furthermore, the test decisions could be contrasted with standard information criteria such as the AIC and BIC. The bootstrap tests confirmed that oscillations induced by a portable shaker vibration calibrator within a controlled experiment and observed by means of the reference and a low-cost accelerometer are practically undamped. The next step will be to adapt the procedure such that the damping ratio coefficient(s) can be tested against specified values in order to verify the structural health of an oscillating structure such as a bridge. Such a bootstrap testing procedure can be extended and used for a factory calibration of a portable shaker vibration calibrator measuring the accelerations by means of a reference accelerometer. In that case it is desirable to detect changes of all the parameters of the damped harmonic oscillation model, that is, possibly multiple frequencies, amplitudes and damping ratio coefficients simultaneously. In addition, the measurements of a low-cost accelerometer oscillating in combination with such a shaker can be investigated over a long period of time regarding aforementioned parameter changes.

V. Acknowledgements

The research was funded partly by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 386369985. In addition, the research was

partly funded and carried out within the scope of the collaborative project "Spatio-Temporal Monitoring of Bridge Structures Using Low Cost Sensors" with ALLSAT GmbH, which was supported by the German Federal Ministry for Economic Affairs and Energy (BMWi) and the Central Innovation Programme for SMEs (Grant ZIM Kooperationsprojekt, ZF4081803DB6). The authors acknowledge the Institute of Concrete Construction (Leibniz University Hannover) for providing the shaker table and the reference accelerometer used within the experiment.

References

Alkhatib, H., Kargoll, B., and Paffenholz, J.-A. (2018). Further Results on a Robust Multivariate Time Series Analysis in Nonlinear Models with Autoregressive and t-Distributed Errors. In: *Time Series Analysis and Forecasting*. *ITISE 2017*. Rojas, I., Pomares, H., and Valenzuela, O. (eds.) Contributions to Statistics, Springer, Cham, pp. 25-38.

Amezquita-Sanchez, J.P. and Adeli, H. (2015). A New Music-Empirical Wavelet Transform Methodology for Time-Frequency Analysis of Noisy Nonlinear and Non-Stationary Signals. Digital Signal Processing, Vol. 45, pp. 55-68.

Angrisano, A., Maratea, A., and Gaglione, S. (2018). A Resampling Strategy Based on Bootstrap to Reduce the Effect of Large Blunders in GPS Absolute Positioning. Journal of Geodesy, Vol. 92, pp. 81-92.

Bogusz, J. and Klos, A. (2016). On the Significance of Periodic Signals in Noise Analysis of GPS Station Coordinates Time Series. GPS Solutions, Vol. 20, pp. 655-664.

Craymer, M.R. (1998). The Least Squares Spectrum, Its Inverse Transform and Autocorrelation Function: Theory and Some Applications in Geodesy. PhD thesis, Graduate Department of Civil Engineering, University of Toronto, Canada.

ISO/IEC (2008). Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement (GUM:1995 with minor corrections). Geneva.

Kargoll, B. (2012). On the Theory and Application of Model Misspecification Tests in Geodesy. Deutsche Geodätische Kommission, Series C (Dissertations), No. 674, Munich.

Kargoll, B., Omidalizarandi, M., Loth, I., Paffenholz, J.-A., and Alkhatib, H. (2018a). An Iteratively Reweighted Least Squares Approach to Adaptive Robust Adjustment of Parameters in Linear Regression Models with Autoregressive and t-Distributed Deviations. Journal of Geodesy, Vol. 92, No. 3, pp. 271-297.

Kargoll B., Omidalizarandi, M., Alkhatib, H., and Schuh, W.-D. (2018b). Further Results on a Modified EM Algorithm for Parameter Estimation in Linear Models with Time-Dependent Autoregressive and t-Distributed Errors. In: *Time Series Analysis and Forecasting*. *ITISE 2017*. Rojas, I., Pomares, H., and Valenzuela, O. (eds.) Contributions to Statistics, Springer, Cham, pp. 323-337.

Kaschenz, J. and Petrovic, S. (2005). A Methodology for the Identification of Periodicities in Two-Dimensional Time Series. Zeitschrift für Geodäsie, Geoinformation und Landmanagement, Vol. 134, pp. 105-112.

Koch, K.R. (2017). Expectation Maximization Algorithm and its Minimal Detectable Outliers. Studia Geophysica et Geodaetica, Vol. 61, pp. 1–18.

Koch, K.R. (2018). Bayesian Statistics and Monte Carlo Methods. Journal of Geodetic Science, Vol. 8, pp. 18-29.

Koch, K.R. and Kargoll, B. (2013). Expectation-maximization algorithm for the variance-inflation model by applying the t distribution. Journal of Applied Geodesy, Vol. 7, pp. 217–225.

Kuhlmann, H. (2003). Kalman-Filtering with Coloured Measurement Noise for Deformation Analysis. In: *Proceedings of the 11th FIG International Symposium on Deformation Measurements*, FIG.

Lehmann, R. (2013). 3σ-Rule for Outlier Detection from the Viewpoint of Geodetic Adjustment. Journal of Surveying Engineering, Vol. 139, pp. 157-165.

Lehmann, R. (2014). Detection of a Sinusoidal Oscillation of Unknown Frequency in a Time Series - a Geodetic Approach. Journal of Geodetic Science, Vol. 4, pp. 136-149.

Lehmann, R. and Lösler, M. (2018). Hypothesis Testing in Non-Linear Models Exemplified by the Planar Coordinate Transformations. Journal of Geodetic Science, Vol. 8, pp. 98-114.

Li, L. (2011). Separability of Deformations and Measurement Noises of GPS Time Series with Modified Kalman Filter for Landslide Monitoring in Real-Time. Ph.D. thesis, Institute of Geodesy and Geoinformation, University of Bonn.

Lösler, M., Eschelbach, C., and Haas, R. (2018). Bestimmung von Messunsicherheiten mittels Bootstrapping in der Formanalyse. Zeitschrift für Geodäsie, Geoinformatik und Landmanagement zfv, Vol. 143, pp. 224-232.

Luo, X., Mayer, M., and Heck, B. (2012). Analysing Time Series of Residuals by Means of AR(I)MA Processes. In: *Proceedings of the VII Hotine-Marussi Symposium of Mathematical Geodesy*. Sneeuw, N., Novàk, P., Crespi, M., and Sansó, F. (eds.) International Association of Geodesy Symposia, Vol. 137, pp. 129-134, Springer, Berlin.

Mautz R. (2001). Zur Lösung nichtlinearer Ausgleichungs-probleme bei der Bestimmung von Frequenzen in Zeitreihen. Deutsche Geodätische Kommission C 532, München.

Mautz, R. and Petrovic, S. (2005). Erkennung von physikalisch vorhandenen Periodizitäten in Zeitreihen. Zeitschrift für Geodäsie, Geoinformation und Landmanagement, Vol. 130, pp. 156-165.

McKinnon, J. (2007). Bootstrap Hypothesis Testing. Queen's Economics Department Working Paper, No. 1127, Queen's University, Kingston, Ontario, Canada.

Nassar, S., Schwarz, K.-P., El-Sheimy, N., and Noureldin, A. (2004). Modeling Inertial Sensor Errors Using Autoregressive (AR) Models. Navigation, Vol. 51, No. 4, pp. 259-268.

Neitzel, F., Resnik, B., Weisbrich, S., and Friedrich, A. (2011). Vibration Monitoring of Bridges. Reports on Geodesy Vol. 1/90, pp. 331-340.

Neitzel, F., Niemeier, W., Weisbrich, S., and Lehmann, M. (2012). Investigation of Low-Cost Accelerometer, Terrestrial Laser Scanner and Ground-Based Radar Interferometer for Vibration Monitoring of Bridges. 6th European Workshop on Structural Health Monitoring, http://www.ndt.net/search/docs.php3?showForm=off&id=14063.

Neitzel, F. and Schwarz, W. (2017). Schwingungsunter-suchungen – Ein Beitrag zum Monitoring im Bauwesen. In:

Ingenieurgeodäsie. Freeden, W. and Rummel, R. (eds.) Handbuch der Geodäsie, Springer Spektrum.

Neuner, H., Wieser, A., and Krähenbühl, N. (2014). Bootstrapping: Moderne Werkzeuge für die Erstellung von Konfidenzintervallen. In: *Zeitabhängige Messgrößen - Ihre Daten haben (Mehr-)Wert*. Neuner, H. (ed.) Schriftenreihe des DVW, Vol. 74, pp. 151-170.

Omidalizarandi, M., Kargoll, B., Paffenholz, J.-A., and Neumann, I. (2018). Accurate Vision-Based Displacement and Vibration Analysis of Bridge Structures by Means of an Image-Assisted Total Station. Advances in Mechanical Engineering, Vol. 10, pp. 1-19.

Pagiatakis, S.D. (1999). Stochastic Significance of Peaks in the Least-Squares Spectrum. Journal of Geodesy, Vol. 73, pp. 67-78.

Park, M. and Gao, Y. (2008). Error and Performance Analysis of MEMS-based Inertial Sensors with a Low-Cost GPS Receiver. Sensors, Vol. 8, pp. 2240-2261.

Parzen, E. (1979). A Density-Quantile Function Perspective on Robust Estimation. In: *Robustness in Statistics*. Launer, L. and Wilkinson, G.N. (eds.) pp. 237-258, Academic Press, New York.

Psimoulis, P., Pytharouli, S., Karambalis, D., and Stiros, S. (2008). Potential of Global Positioning System (GPS) to Measure Frequencies of Oscillations of Engineering Structures. Journal of Sound and Vibration, Vol. 318, pp. 606–623.

Schuh, W.-D. (2003). The Processing of Band-Limited Measurements; Filtering Techniques in the Least Squares Context and in the Presence of Data Gaps. Space Science Reviews, Vol. 108, No. 1, pp. 67-78.

Sommer, K.D. and Siebert, B.R.L. (2004). Praxisgerechtes Bestimmen der Messunsicherheit nach GUM. Technisches Messen, Vol. 71, pp. 52–66.

Teunissen, P.J.G. (2003). Testing Theory; an Introduction. Delft University Press, Delft.

Vaníček, P. (1969). Approximate Spectral Analysis by Least-Squares Fit. Astrophysics and Space Science, Vol. 4, pp. 387-391.

Wells, D., Vaníček, P, and Pagiatakis, S.D. (1985). Least-Squares Spectral Analysis Revisited. Technical Report No. 84, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada.

Wiśniewski, Z. (2014). *M*-estimation with probabilistic models of geodetic observations. Journal of Geodesy, Vol. 88, pp. 941–957.

APPENDIX

The Jacobi matrix $\boldsymbol{A}$ with respect to the damped harmonic oscillation model (2) is based on the partial derivatives $\frac{\partial h_t(\boldsymbol{\beta},\boldsymbol{\xi})}{\partial a_0} = 0.5$ as well as

$$\frac{\partial h_t(\boldsymbol{\beta},\boldsymbol{\xi})}{\partial a_j} = \cos\left(2\pi f_j\sqrt{1-\xi_j^2}\,x_t\right),$$

$$\frac{\partial h_t(\boldsymbol{\beta},\boldsymbol{\xi})}{\partial b_j} = \sin\left(2\pi f_j\sqrt{1-\xi_j^2}\,x_t\right),$$

$$\begin{aligned}\frac{\partial h_t(\boldsymbol{\beta},\boldsymbol{\xi})}{\partial \xi_j} = \Bigg[&-a_j\sin\left(2\pi f_j\left[1-\xi_j^2\right]^{-\frac{1}{2}}x_t\right)\\ &+b_j\cos\left(2\pi f_j\left[1-\xi_j^2\right]^{-\frac{1}{2}}x_t\right)\Bigg]\\ &\times\left(-2\pi f_j\xi_j\left[1-\xi_j^2\right]^{-\frac{1}{2}}x_t\right)\exp\left(-2\pi f_j\xi_j x_t\right)\\ &+\Bigg[a_j\cos\left(2\pi f_j\left[1-\xi_j^2\right]^{-\frac{1}{2}}x_t\right)\\ &+b_j\sin\left(2\pi f_j\left[1-\xi_j^2\right]^{-\frac{1}{2}}x_t\right)\Bigg]\\ &\times\exp\left(-2\pi f_j\xi_j x_t\right)\left(-2\pi f_j x_t\right)\end{aligned}$$

$$\begin{aligned}\frac{\partial h_t(\boldsymbol{\beta},\boldsymbol{\xi})}{\partial f_j} = \Bigg[&-a_j\sin\left(2\pi f_j\left[1-\xi_j^2\right]^{-\frac{1}{2}}x_t\right)\\ &+b_j\cos\left(2\pi f_j\left[1-\xi_j^2\right]^{-\frac{1}{2}}x_t\right)\Bigg]\\ &\times\left(2\pi\left[1-\xi_j^2\right]^{-\frac{1}{2}}x_t\right)\exp\left(-2\pi f_j\xi_j x_t\right)\\ &+\Bigg[a_j\cos\left(2\pi f_j\left[1-\xi_j^2\right]^{-\frac{1}{2}}x_t\right)\\ &+b_j\sin\left(2\pi f_j\left[1-\xi_j^2\right]^{-\frac{1}{2}}x_t\right)\Bigg]\\ &\times\exp\left(-2\pi f_j\xi_j x_t\right)\left(-2\pi\xi_j x_t\right)\end{aligned}$$