Synchronized Load Quantification from Multiple Data Records for Analysing High-rise Buildings

1st Marco Behrendt
Institute for Risk and Reliability
Leibniz Universität Hannover
Hannover, Germany
behrendt@irz.uni-hannover.de

2nd Wonsiri Punurai
Department of Civil and Environmental Engineering
Mahidol University
Bangkok, Thailand
wonsiri.pun@mahidol.ac.th

3rd Michael Beer
Institute for Risk and Reliability
Leibniz Universität Hannover
Hannover, Germany
beer@irz.uni-hannover.de

***Abstract*—To analyse the reliability and durability of large complex structures such as high-rise buildings, most realistically, it is advisable to utilize site-specific load characteristics. Such load characteristics can be made available as data records, e.g. representing measured wind or earthquake loads. Due to various circumstances such as measurement errors, equipment failures, or sensor limitations, the data records underlie uncertainties. Since these uncertainties affect the results of the simulation of complex structures, they must be mitigated as much as possible. In this work, the Procrustes analysis, finding similarity transformations between two sets of points in n-dimensional space is used and is extended to uncertainties so that data records can be analysed regarding the uncertainty. To find the best matching of two sets of points the Kabsch algorithm is used. In this manner, a basis is created to simulate and assess the reliability of high-rise buildings under load due to wind and earthquakes.**

***Keywords*—earthquake simulation, Procrustes analysis, uncertainty quantification, random vibrations, RESET**

I. INTRODUCTION

High-rise buildings are subject to environmental processes such as wind and earthquake loads. In order to make predictions as to whether high-rise buildings withstand certain loads, it is necessary to record and analyse data of these environmental processes. Often, these data are subject to uncertainties, as the measuring instruments operate not accurately enough or are subject to other influences. Common reasons for uncertain or limited data include measurement errors and sensor limitation or equipment failure. In the case of measurement errors, the data is recorded incorrectly and is different from the actual data. Likewise, the sensors may have certain limitations and data can only be detected up to a certain threshold. If the equipment fails, parts or even the entire earthquake will not be recorded. Device failures can be caused by the environmental process itself. This results in missing data within the measurement series, which must be reconstructed.

To quantify these uncertainties, more than just one sensor should be positioned in a high-rise building so that the data from different sensors can be compared and potential uncertainties can be detected and mitigated as much as possible. The comparison of two data sets results in an orthogonal Procrustes problem [1,2,3]. Two similar data sets are compared and adapted to each other as best as possible. Often, as in this work, the Kabsch algorithm [4,5] is used to solve the orthogonal Procrustes problem. The Kabsch algorithm finds a wide field of application. For example, it is used in depth restoration of images [6], shape analysis [7], noise compensation for speaker recognition [8], network intrusion detection [9] and calibration of laser sensors in mobile robotics [10].

In this work, the various influences that make the measured signals uncertain are examined in more detail. The influence of noise, missing data and rotated sensors are considered. First, the strength of the influence of these factors is determined and then a sensitivity analysis is performed. It determines which influences affect the results most and whether they distort the results too much.

This work is organised as follows. Section II provides a general overview and set up of the problem and describes the Kabsch algorithm. Section III first shows the results of the influence of the factors which may disturb the measured data. These are noise, missing data and rotated sensors. Then, a sensitivity analysis is performed considering these factors. It is followed by the conclusion in Section IV.

II. PROBLEM

The load recording is based on two sensors positioned in a high-rise building on the same floor as shown in Fig. 1. Both sensors measure movement and displacement of the high-rise building during the earthquake. It is necessary to place two or more sensors, as it is not always possible to place just one sensor in the centre of the building due to obstacles such as columns or walls. To measure the displacement in the middle of the building, the measured data of the sensors are transformed into the middle via (1). Therefore, the distances Δx and Δy respectively between both sensors and the distances x_l, x_r, y_t and y_b between the sensors and the middle of the building must be known.

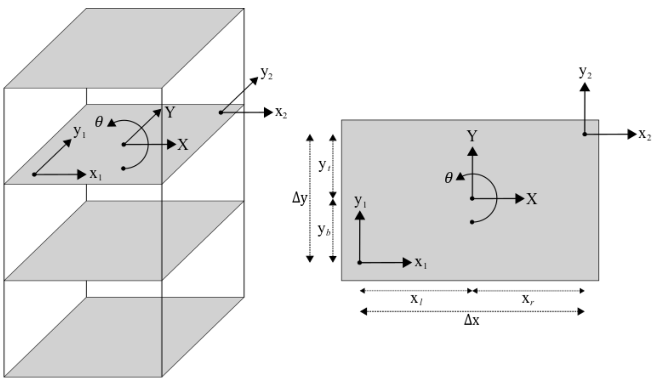

Fig. 1. *Setup of the Problem*

$$\begin{pmatrix} X(t) \\ Y(t) \\ \Theta(t) \end{pmatrix} = \begin{pmatrix} \frac{y_t}{\Delta y} & 0 & \frac{y_b}{\Delta y} & 0 \\ 0 & \frac{x_r}{\Delta x} & 0 & \frac{x_l}{\Delta x} \\ \frac{1}{\Delta y} & \frac{-1}{\Delta x} & \frac{-1}{\Delta y} & \frac{1}{\Delta x} \end{pmatrix} \begin{pmatrix} x_1(t) \\ y_1(t) \\ x_2(t) \\ y_2(t) \end{pmatrix} \quad (1)$$

Furthermore, to calculate the mode shape magnitude it is necessary to place sensors in more than just one storey. An example is depicted in Fig. 2.

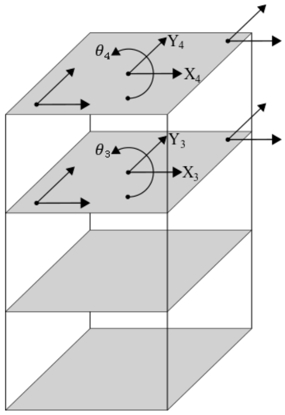

Fig. 2. Sensors on two storeys to calculate mode shape magnitude

To calculate the mode shape magnitude from 3rd to 4th floor (2) is considered. In this equation ω_x, ω_y and ω_θ are the natural frequencies.

$$\left(\left| \frac{X_3(\omega_i)}{X_4(\omega_i)} \right|, \left| \frac{Y_3(\omega_i)}{Y_4(\omega_i)} \right|, \left| \frac{\theta_3(\omega_i)}{\theta_4(\omega_i)} \right| \right) \quad (2)$$

with $i \in \{x, y, \theta\}$.

Because of the reasons given in section I, the data often cannot be measured accurately enough, it is necessary to compare the measured data of the two sensors with each other to estimate the uncertainty. The comparison of the two signals of the sensors results in an orthogonal Procrustes problem, since the data is mapped to each other. For this, the Kabsch algorithm is often used, which is considered in more detail in the following section.

A. *Kabsch algorithm*

The orthogonal Procrustes problem is a matrix approximation problem in which the best orthogonal matrix R is found to map two sets of points. The approximation is according to

$$R = \arg\min ||RY - X||_F \quad \text{subject to} \quad R^T R = I. \quad (3)$$

To solve this problem the Kabsch algorithm is used to find the best approximation for the rotation matrix R based on the least root mean square. The algorithm is explained in the following. Given two sets of points X and Y with points x_i and y_i ($i=1,\dots,N$). For every point x_i of X exists a corresponding point y_i in Y. The algorithm starts with calculating the centroid for each of the two sets of points. The centroids of the points sets are shifted to the origin of the coordinate system. In the next step, the covariance matrix $A=\tilde{X}^T\tilde{Y}$ of the centred matrices $\tilde{X}$ and $\tilde{Y}$ is calculated. A singular value decomposition is applied to the covariance matrix $A=VSW^T$ to estimate the rotation matrix R. To ensure that the estimated rotation matrix R is according to a right-hand-sided coordinate system the determinant $d=\det(W^T V)$ has to be calculated. Then the estimation of the rotation matrix is

$$R = W \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix} V^T \quad (4)$$

for the two-dimensional case, which is considered in this work. Now, R can be applied to the centred set of point $\tilde{Y}$ and the two sets of points are mapped as best as possible.

The Kabsch algorithm is subject to two limitations. First, if data is missing due to system errors, the data must be missing at the same point in time in both data sets. It is

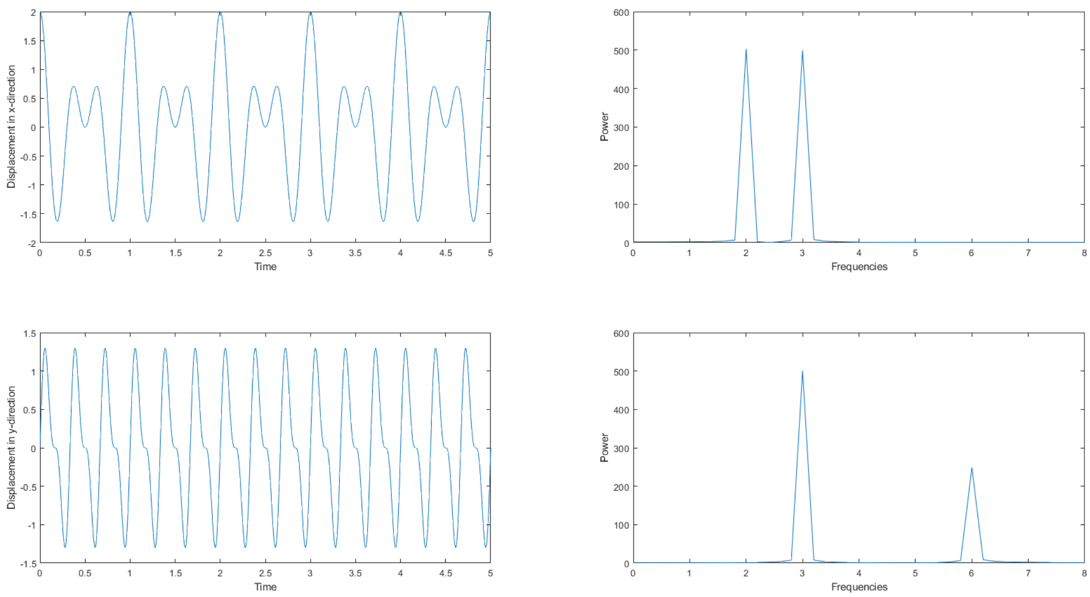

Fig. 3. *Sensor motion in x- and y-direction and corresponding frequencies*

necessary that the compared sets of points consist of the same number of data points, which obviously is not the case when data is missing. Second, the position of both sensors must be known for a correct reconstitution of the data in the centre.

Because the sensors are independent from each other, there is no guarantee that the data will be missing in the same point in time. It is assumed that at least the time steps at which the data is missing are known and can be filled up, so that both data sets have the same number of data points. Then the Kabsch algorithm can be applied. The second limitation is not a problem in this work either, as the position of the sensors must be known anyway for the meaningful validation of the data.

III. Results

In this chapter the influence on the reconstruction of the individual factors will be discussed and later a sensitivity analysis will be performed. The reconstruction is evaluated by the least root mean square error as shown in (5).

$$E = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|x_i - y_i|^2} \tag{5}$$

Since the simulations were influenced by random, each simulation was performed 10000 times. From this, the mean μ_e and standard deviation σ_e were calculated to obtain meaningful results. Outliers thus have less influence.

(6) are used as the measured signal in *x*- and *y*-direction for both sensors.

$$\begin{aligned} x &= \cos(2\pi 2t) + \cos(2\pi 3t) \\ y &= \sin(2\pi 3t) + 0.5\sin(2\pi 6t) \end{aligned} \tag{6}$$

with $t = [0; 5]$, $\Delta t = 0.005$.

The signal is also shown in Fig. 3 with the corresponding frequencies for both directions.

An example of the following calculations is depicted in Fig. 4, where the first sensor measured undisturbed data and the second sensor was rotated by angle φ.

A. Influence of the individual factors

First, the influence of noisy data is considered. After that, the missing data will be discussed. Finally, the influence of rotated sensors is determined.

The level of noise is characterized as signal-to-noise ratio (SNR). The SNR describes the ratio between the power of a signal to the power of the background noise, thus

$$\mathrm{SNR} = \frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}} \tag{7}$$

Since the noise was placed randomly on the data, each simulation was performed 10000 times and the mean and standard deviation were calculated. Initially, only one sensor was placed with noise and rotated by 20°. After applying the Kabsch algorithm, the error was determined. The results are shown in Fig. 5. It turns out that at a very low SNR the error is very high, and the results are poorly reconstructed. However, from an SNR of about 40, the influence of noise is limited. The error here is almost 0, while it converges to 0 with increasing SNR. The same simulation was performed when both sensors were noisy. However, there were no significant differences from the previous results.

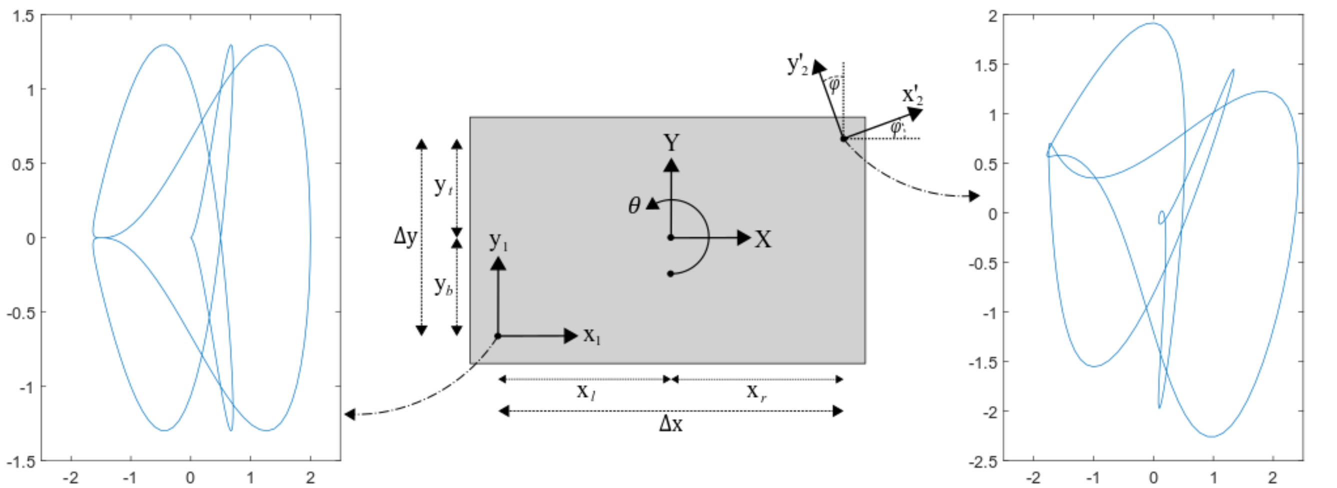

Fig. 4. *Two sensors and the corresponding measured motion*

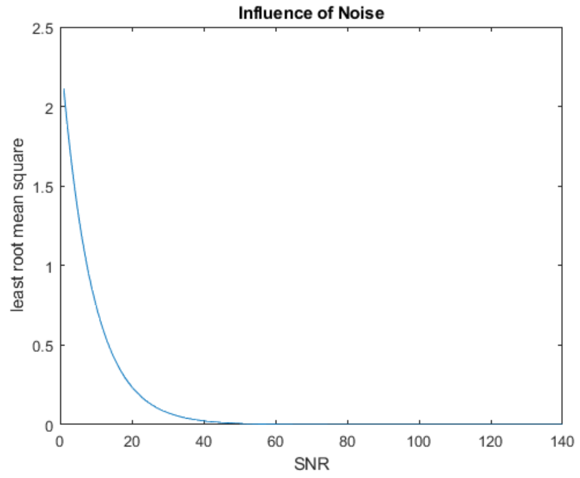

Fig. 5. *Influence of noise*

To determine the influence of the missing data on the reconstruction, they were randomly removed at a sensor and compared to the original data after reconstruction. The missing data was reconstructed by calculating the mean of the adjacent data points. Since the reconstruction of the data points is independent of a rotation of the sensors, this was initially not considered.

In the first step, only missing data was generated in one sensor. The number of missing data was increased, and the error measured. The results are shown in Fig. 6. The error increases almost linearly with the number of missing data points until the range of 700-750 missing data points is reached. From this point the error increases drastically. The data can thus be reconstructed up to of about 700 missing data points with a moderate error. Thus, it can also be said that the ratio of the number of missing to the number of all data points matters and the error is limited only by the calculation of the mean value up to a ratio of 70%.

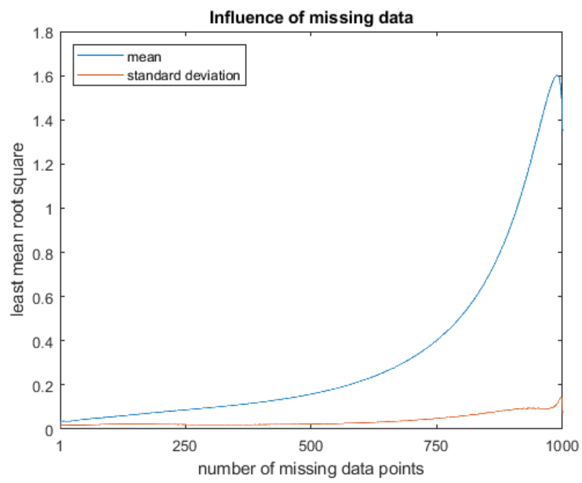

Fig. 6. *Influence of missing data for one sensor*

In the second case, the data at both sensors were randomly removed, reconstructed over the mean of the adjacent data points, and the error calculated. The data are not necessarily missing at the same points in time. The results are shown in Fig. 7. Again, a high number of missing data hinders a good reconstruction. The error increases moderately with the number of missing data points. The reconstruction becomes too inaccurate after a number of about 750 missing data points. However, it should be noted that the missing data points of two sensors hardly influence each other. Thus, the error remains approximately the same at any number of missing data points of the first sensor as the missing data points in the second sensor are increased. Even if data points are missing in both sensors, the ratio of 70% is easy to reconstruct.

Fig. 7. *Influence of missing data for two sensors*

Next, the influence of the rotation angle on the reconstruction of the data is determined. Since the Kabsch algorithm already provides good reconstructions with a small error for a fixed angle, it is considered here how a time-varying angle relates to the reconstruction using the Kabsch algorithm. In the first case, the data is rotated evenly over time. In a second case, it is considered how a time-varying angle relates to the reconstruction. This is particularly interesting considering the simulation of earthquakes.

For the first case, the sensor was rotated evenly over the entire measurement period, until an angle of 20° was reached at the last point in time. These data were reconstructed using the Kabsch algorithm. The error of the data is only 0.1297. The results are shown in Fig. 8.

For the second case, the change of the angle per time step was determined at random. In each time step, a random discrete uniformly distributed number between -3 and 3 was generated and added to the angle of the previous time step. Thus, the angle changes randomly over time, which is more realistic for the simulation of earthquakes. An example simulation is shown in Fig. 9. Since each of these simulations are influenced by random, it was performed 10000 times. The mean value of these simulations is 0.5198 with a standard deviation of 0.1983.

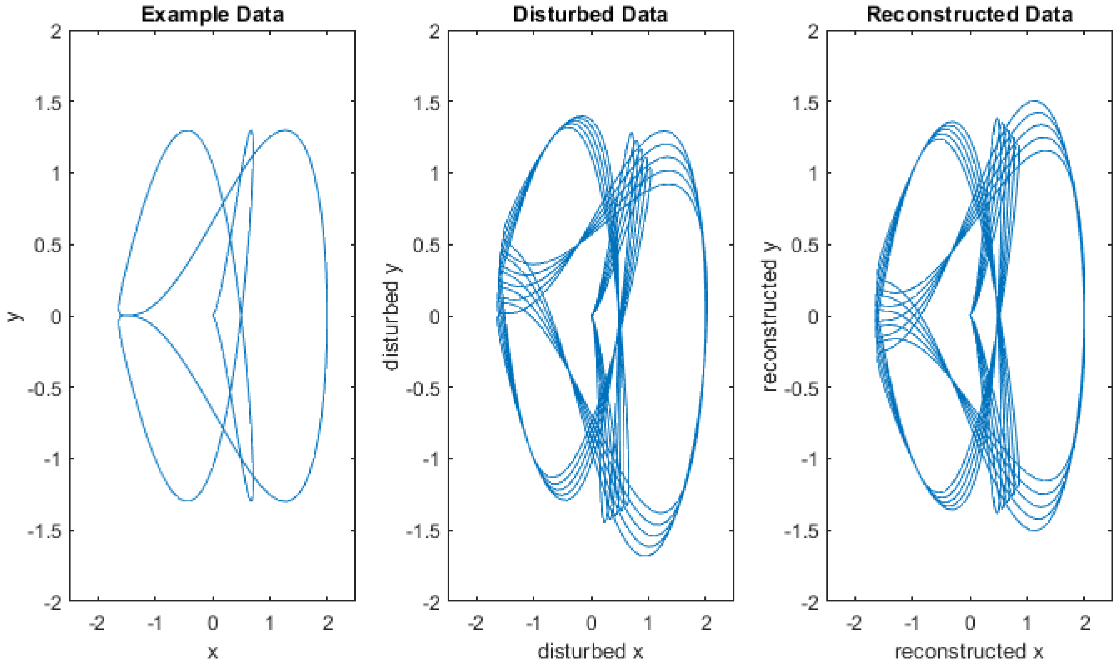

Fig. 8. *Influence of evenly rotated sensors*

B. *Sensitivity analysis*

In the following section, the individual influencing parameters described and analysed in the previous section are examined by Monte Carlo simulations and the error of reconstruction is calculated using the Kabsch algorithm in order to do a sensitivity analysis. First, the influences are examined individually, later all factors are applied together. For each Monte Carlo simulation 10000 calculations were performed and the mean μ_e and standard deviation σ_e of the least root mean squared error were calculated. In addition, first one sensor and then both sensors with the respective influences were disturbed for all factors. For this section the motion of the sensors is different from each other to make the problem more realistic. For the first sensor

$$x = \cos(2\pi 2t) + \cos(2\pi 3t)$$
$$y = \sin(2\pi 3t) + 0.5\sin(2\pi 6t) \tag{8}$$

are used and for the second sensor

$$x_2 = 1.2\cos(2\pi 2t) + 1.1\cos(2\pi 3t)$$
$$y_2 = 1.3\sin(2\pi 3t) + 0.8\sin(2\pi 6t) \tag{9}$$

with $t = [0; 5]$, $\Delta t = 0.005$ are used.

For this case it is obvious that a perfect reconstruction is impossible, and an error will remain even if there are no influences disturbing the sensors.

First, the influence of noise with Monte Carlo simulation was examined. The signal-to-noise ratio, which was applied to the sensors, was determined with an exponentially distributed random variable with Parameter λ=0.05 as an SNR of 140 is high enough to observe no influence. In table 1, the results are shown both in case of one sensor and in case both sensors were disturbed. It turns out that there are hardly any differences between the two cases. This is consistent with the results from section III A.

The number and position of missing data was only simulated. The number was determined by an exponentially distributed random variable with parameter

$$\lambda = \frac{1}{0.15N} \tag{10}$$

with N as the total number of data points. The position in the data record at which the data is missing was determined by discrete uniform distributed random variables. Before measuring the error, the missing data points are reconstructed via the mean value of the adjacent data points. The results are shown in table 1.

The first case of the rotated sensors, where a sensor is evenly rotated over time, is not considered here since it is deterministic. For the second case, where the sensors rotate randomly due to the earthquake ground motion, the results are shown in table 1. For this case the change of the angle to the previous time step is determined by a normally distributed random variable with the parameters μ_e=0 and σ_e=2.

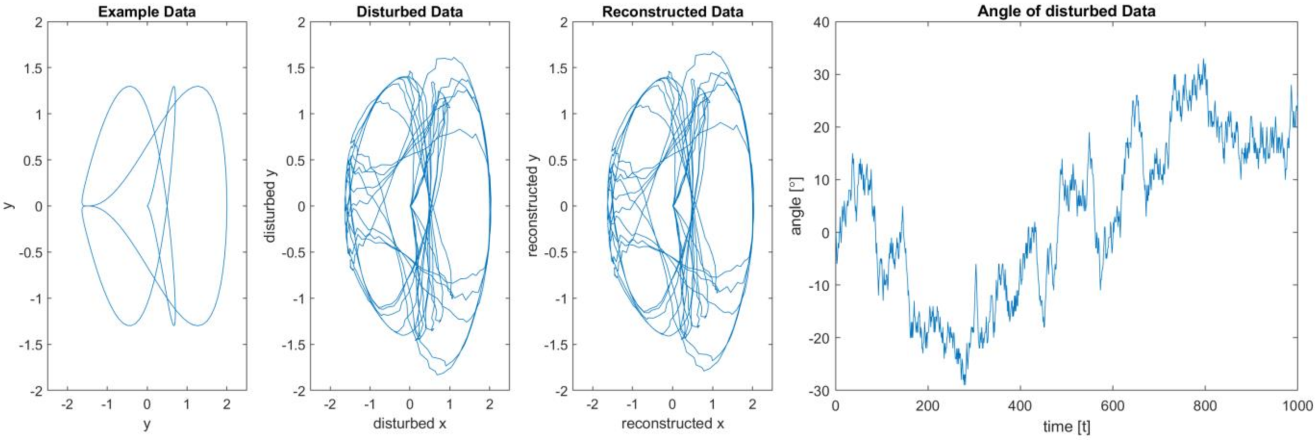

Fig. 9. *Influence of randomly rotated sensors*

As the last simulation, all influencing factors were applied in combination. For the generation of the individual influences, the same types of random variables with the same parameters were used as in the previous simulations. The results are shown in table 1.

TABLE 1: *RESULTS OF THE SENSITIVITY ANALYSIS*

	# disturbed sensors	μ_e	σ_e
Noise	1	0.3419	0.0729
	2	0.3453	0.1033
Missing Data	1	0.3501	0.062
	2	0.3673	0.0921
Rotation	1	0.6795	0.1971
	2	0.8763	0.2693
All factors	1	0.6866	0.2059
	2	0.8845	0.2789

IV. CONCLUSION

As the results in section III A show, the noise produces the smallest error. Only with an extremely poor SNR the error is too big, so that the data cannot be used. The error converges very fast to 0 with increasing SNR. In addition, the error behaves regardless of whether one or two sensors are noisy.

The reconstruction is relatively robust against a high number of missing data. Since these are reconstructed with the mean value of the adjacent data points, records can contain up to 70% missing data while the error is in a moderate range. Only when the number of missing data is higher, the error rises sharply, and the measurements cannot be used. The limit of 70% missing data was shown for one as well as two sensors.

With a sensor that rotates evenly over time, only a small error is calculated. The larger the angle in the last time step, the bigger the error becomes. Since this case is rather unrealistic, this was done here for the angle of 20° as an example. A more realistic example is the second case considered. The error is higher for this case, which is due to the fact that there can be a higher variation in the change of the angle. In addition, there can be significant outliers.

To make the problem more realistic, two slightly different signals were generated for a sensitivity analysis in section III B. The signals were disturbed by the respective factors and the error was calculated.

Here, too, the noise and the missing data have a much smaller influence on the reconstruction than the rotation of the sensors. While it hardly makes a difference in the noise or the missing data, if one or both sensors are disturbed, a much larger difference can be seen in the rotation of sensors.

If all factors are applied at the same time, the mean and standard deviation is similar, as with the rotated sensors. This means that the rotated sensors have the biggest impact on the error and the other two tend to be negligible. All in all, the error in rotated sensors or applying all factors at the same time is quite large, but a use of the measured data for the analysis of the lifetime of high-rise buildings is possible. Due to the calculated standard deviations, different variations of the mean value can be used for simulations.

In this work the problem was first described and the used Kabsch algorithm explained. Thereafter, the influence of each factor was carried out. A sensitivity analysis followed, in which the factors were presented realistically. Most recently, the results were interpreted.

ACKNOWLEDGMENT

Part of this work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 730888.

Part of this work was also funded by the Deutsche Forschungsgemeinschaft (German Research Foundation) grants BE 2570/4-1 and CO 1849/1-1 as part of the project 'Uncertainty modelling in power spectrum estimation of environmental processes with applications in high-rise building performance evaluation'.

REFERENCES

[1] Schönemann, P. H.: A generalized solution of the orthogonal procrustes problem. Psychometrika 31(1), 1-10 (1966).

[2] Gower, J. C.: Generalized procrustes analysis. Psychometrika 40(1), 33-51 (1975).

[3] Lourenço, P., Guerreiro, B.J., Batista, P., Oliveira, P. and Silvestre, C.: Uncertainty characterization of the orthogonal procrustes problem with arbitrary covariance matrices. Pattern Recognition. 61, 210–220 (2017).. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[4] Kabsch, W.: A solution for the best rotation to relate two sets of vectors. Acta Cryst. A32, 922-923 (1976).

[5] Kabsch, W.: A discussion of the solution for the best rotation to relate two sets of vectors. Acta Cryst. A34, 827-828 (1978).

[6] Rotman, D. and Gilboa, G.: A depth restoration occlusionless temporal dataset. 3D Vision (3DV). 176--184 (2016).

[7] Dogan, G., Bernal, J. and Hagwood, C. R.: A fast algorithm for elastic shape distances between closed planar curves. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 4222-4230 (2015).

[8] Kheder, W. B., Matrouf, D., Ajili, M. and Bonastre, J. F.: Iterative Bayesian and MMSE-based noise compensation techniques for speaker recognition in the i-vector space. In Proc. Speaker Lang. Recognit. Workshop. 60-67 (2016).

[9] Isbudeen, N., Ajaz, S. and Buhari, P. A.: Network Intrusion Detection in Wireless Network using Kabsch Algorithm, In Proceedings of the IEEE International Conference on Computational Intelligence and Computing Research. 1457-1461 (2011).

[10] Juric-Kavelj, S. and Petrovic, I.: Bias Estimation and Calibration of a Pair of Laser Range Sensors. In SyRoCo. 611-615 (2012).