Uncertainty propagation of missing data signals with the interval discrete Fourier transform

Marco Behrendt[1], Marco de Angelis[2], and Michael Beer[1,3,4]

[1]Institute for Risk and Reliability, Leibniz Universität Hannover, Callinstraße 34, 30167 Hannover, Germany

[2]Department of Civil and Environmental Engineering, University of Strathclyde, 16 Richmond Street, Glasgow G1 1XQ, United Kingdom

[3]Institute for Risk and Uncertainty, University of Liverpool, Peach Street, Liverpool L69 7ZF, United Kingdom

[4]International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, 1239 Siping Road, Shanghai 200092, China

ABSTRACT

The interval discrete Fourier transform (DFT) algorithm can propagate signals carrying interval uncertainty. By addressing the repeated variables problem, the interval DFT algorithm provides exact theoretical bounds on the Fourier amplitude and estimates of the power spectral density (PSD) function, whilst running in polynomial time. Thus, the algorithm can be used to assess the worst-case scenario in terms of maximum or minimum power, and to provide insights into the amplitude spectrum bands of the transformed signal. To propagate signals with missing data, an upper and lower value for the missing data present in the signal must be assumed, such that the uncertainty in the spectrum bands can also be interpreted as an indicator of the quality of the reconstructed signal. For missing data reconstruction, there are a number of techniques available that can be used to obtain reliable bounds in the time domain, such as Kriging regressors or interval predictor models. Alternative heuristic strategies based on variable – as opposed to fixed – bounds

can also be explored. This work aims to investigate the sensitivity of the algorithm against interval uncertainty in the time signal. The studies are conducted in different case studies using signals of different lengths generated from the Kanai-Tajimi PSD function, representing earthquakes, and the JONSWAP PSD function, representing sea waves as a narrowband PSD model.

INTRODUCTION

The consideration and quantification of uncertainties in real data records are of paramount importance for the design and simulation of buildings and structures and in engineering in general (Schuëller 2007; Kiureghian and Ditlevsen 2009; Nikolaidis et al. 2004). Even small measurement errors can lead to a wrong consideration of the input data and result in a disastrous interpretation of the simulation results, e.g. if an actually catastrophic result is shifted into an acceptable range by not taking uncertainties into account. Uncertainties should therefore be considered in any case and included in the simulation, also in order to determine possible safety margins. An overview of methods to model and quantify uncertainties is given, for instance in (Beer et al. 2013; Faes and Moens 2020).

In order to safely design or to assess the reliability and robustness of buildings and structures that are subject to environmental processes such as wind, earthquakes or waves and thus exhibit a dynamic behaviour, simulations are indispensable. Specifically in random vibrations (Soong and Grigoriu 1993; Roberts and Spanos 2003; Lutes and Sarkani 2004), spectral analysis (Priestley 1982; Newland 2012) and stochastic and structural dynamics (Lin and Cai 1995; Chopra 1995; Li and Chen 2009), the determination of the dynamic characteristics of such an environmental process is very important. In this regard, the power spectral density (PSD) function is an important tool as it is used to determine the governing frequencies of a signal and their amplitude. In the stationary case, the PSD function is based on the discrete Fourier transform (DFT), see for instance (Sneddon 1995). Therefore, the DFT is used ubiquitously when determining the spectral properties of a random signal and decomposing it into its harmonic components. The probably most famous implementation is the fast Fourier transform (FFT), first appeared in (Cooley and Tukey 1965; Cooley 1987).

However, if an uncertain signal is now to be transformed via the DFT, this cannot be accomplished with absolute certainty, since the DFT is not defined for non-discrete signals. Therefore, accounting for uncertainties in the data, such as missing data, should be combined with the DFT to obtain reliable results. Missing data in a signal, if not properly accounted for, can lead to severe erroneous results, as the spectral characteristics of the signal could thus be incorrectly determined. In particular, the estimation of the PSD function and the subsequent simulation of a structure can lead to incorrect results if the spectral characteristics, such as the peak frequency, are not determined correctly. Quantifying uncertainties of such a signal after transformation to the frequency domain is therefore of utmost importance.

Missing data in a signal can be reconstructed either by discrete points or in certain bounds represented by intervals. Such methods can be, for instance, least squares methods (Levenberg 1944), compressive sensing (Comerford et al. 2016; Comerford et al. 2017), autoregressive methods (Naghizadeh and Sacchi 2007; Naghizadeh and Sacchi 2010), interval predictor models (Campi et al. 2009; Sadeghi et al. 2019; Rocchetta et al. 2021) or Kriging (De Rubeis et al. 2005; Lin and Li 2020). However, since none of these advanced methods can guarantee, that the original data point is reconstructed with absolute certainty, a residual uncertainty remains. In fact, the DFT is very sensitive against small changes in the input signal, which will result in uncertain determination of the spectral characteristics of said signal. Therefore, only a simple reconstruction of the missing data may not be reliable enough and other methods must be sought which are capable of effectively transforming an uncertain signal to the frequency domain.

Some approaches for estimating PSD functions from signals with missing data have already been developed. In particular, approaches treating missing data as Gaussian distributed random variables and propagating them through the DFT (Comerford et al. 2015b; Zhang et al. 2017), while artificial neural networks are used in (Comerford et al. 2015a). Another approach was presented by (Liu and Kreinovich 2010), where the FFT and convolution were studied for signals with interval and fuzzy uncertainty. An algorithm to propagate interval signals through the DFT to obtain exact bounds on the Fourier amplitude, the so-called *interval DFT algorithm*, was derived by the authors of this work in (De Angelis et al. 2021), while the algorithm is described in details and applied to an example involving a dynamic structural analysis in (Behrendt et al. 2022b). Further insights can be found in (de Angelis 2022). The algorithm enables the quantification of uncertainties in time signals and to project them into the frequency domain by using interval arithmetic (Moore 1966; Moore 1979; Moore et al. 2009; Alefeld and Herzberger 2012). No assumptions are made about the dependence and distribution of the error over the time steps. The interval DFT algorithm fully addresses the repeated variables problem. Thus, the exact bounds on the Fourier amplitude and on an estimation of a PSD function can be computed, which can be used to analyse system responses in the frequency domain, taking into account these uncertainties.

The objective of this work is to investigate the capabilities of the interval DFT algorithm in missing data problems. It also aims to determine the severity of the missing data and the impact on the interval DFT algorithm and thus on the resulting bounds. The quantity used to measure uncertainty in this work is the area between the upper and lower bounds. An uncertain signal has an area between an estimation of the PSD function bounds greater than zero, whereas a PSD function without uncertainty has an area equal to zero, i.e. a discrete-valued PSD function. In addition, the findings of this work can be used to determine if a signal is considered insufficiently reliable to be used for frequency analysis. Preliminary studies on this work were conducted in (Behrendt et al. 2022a).

This work is organised as follows: Some theoretical background that is relevant for this work, such as the interval DFT algorithm, is provided in Section 2, while the problem of missing data is elaborated in Section 3. The capabilities of said algorithm in combination with missing data problems are explored in Section 4. The final conclusions are given in Section 5.

PRELIMINARIES

This section introduces some fundamental theoretical concepts that will be required in this work.

Interval analysis

An interval $\underline{\overline{x}} \in \mathbb{R}$ is defined as

$$\underline{\overline{x}} = [\underline{x}, \overline{x}] = \{\underline{x} \leq x \leq \overline{x}\}, \tag{1}$$

where $\underline{x}$ and $\overline{x}$ define the lower and upper bound, respectively. Every value between those bounds is a possible value. The interval is further defined by the interval midpoint

$$m_x = \frac{\underline{x} + \overline{x}}{2} \tag{2}$$

and the interval width

$$h_x = \overline{x} - \underline{x}. \tag{3}$$

Power spectral density estimation

Given a signal x_n, represented as a zero mean stochastic process. To examine the signal for its frequency components, it can be transformed into the frequency domain using the periodogram. The periodogram is the squared absolute value of the Fourier transform and reads as follows

$$\hat{S}_X(\omega_k) = \frac{\Delta t^2}{T} \left| \sum_{n=0}^{N-1} x_n \cdot e^{-\frac{2\pi i k n}{N}} \right|^2, \tag{4}$$

where Δt is the time step size, T is the total length of the record, n describes the data point index in the record, N is the total number of data points in the signal and k is the frequency number of $\omega_k = \frac{2\pi k}{T}$.

Generation of artificial time signals

To generate an artificial time signal for simulation purposes, the Spectral Representation Method (SRM) can be utilised (Shinozuka and Deodatis 1991). The SRM generates a time signal X_t based on an underlying PSD function S_X while carrying the spectral characteristics of this PSD function.

The SRM is

$$X_t = \sum_{m=0}^{M-1} \sqrt{4 S_X(\omega_m)\Delta\omega} \cos(\omega_m t + \varphi_m), \tag{5}$$

with $\omega_m = m\Delta\omega$, $m = 0, 1, 2, \ldots, M-1$, where M is the total number of frequency points, t as time coordinate and φ_m as uniformly distributed random phase angles in the range $[0, 2\pi]$.

As the underlying PSD function, a spectrum derived within the Joint North Sea Wave Observation Project (JONSWAP) (Hasselmann et al. 1973) will be used throughout this work. The JONSWAP PSD function is an extension of the Pierson-Moskowitz PSD function (Pierson Jr. and Moskowitz 1964) and is utilised to describe the dynamic behaviour of sea waves in the frequency domain. The PSD function reads as follows

$$S^J(\omega) = \frac{\alpha g^2}{\omega^5} \exp\left(-\frac{5}{4}\left(\frac{\omega_p}{\omega}\right)^2\right)\gamma^r \tag{6}$$

with

$$r = \exp\left(\frac{-(\omega - \omega_p)^2}{2\sigma^2\omega_p^2}\right).$$

In these equations α describes a spectral energy parameter, g is the gravity acceleration, ω_p describes the peak frequency, γ^r is the peak enhancement factor and σ the spectral width parameter. An example for the JONSWAP PSD function with $\alpha = 0.0081$, $w_p = 0.7$, $\gamma = 3.3$ and

$$\sigma = \begin{cases} 0.7 & \omega \leq \omega_p \\ 0.9 & \omega > \omega_p \end{cases}$$

is given in Fig. 1. The JONSWAP PSD function is characterised by its narrow band in the frequency domain and the very strong and sharp peak, thus many values distant from this peak are close to or equal to zero.

A second PSD function is used for verification, namely the Kanai-Tajimi PSD function, which

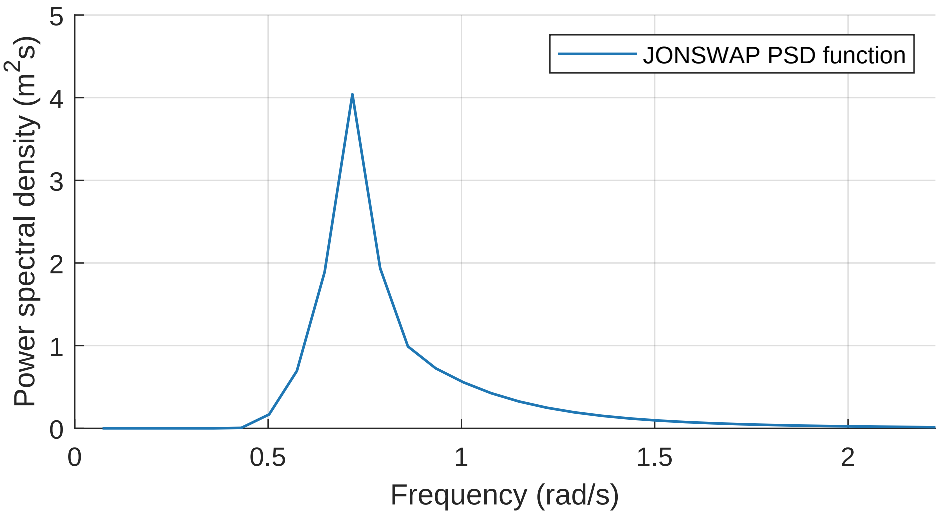

Fig. 1. Example for the JONSWAP PSD function.

is as follows

$$S^{KT}(\omega) = S_0 \frac{1 + 4\xi^2 \frac{\omega^2}{\omega_p^2}}{\left(1 - \frac{\omega^2}{\omega_p^2}\right)^2 + 4\xi^2 \frac{\omega^2}{\omega_p^2}}. \tag{7}$$

In this expression, $S_0 = 0.45$ is a constant, $\omega_p = 2\pi$ describes the peak frequency and $\xi = 0.25$ indicates the sharpness of the peak (Kanai 1957; Tajimi 1960). Furthermore, the upper cut-off frequency is defined to be $\omega_u = 4\pi$ rad/s. The Kanai-Tajimi PSD function with parameters $S_0 = 0.45$, $\omega_p = 2\pi$ and $\xi = 0.25$ is given in Fig. 2. In contrast to the JONSWAP spectrum, the Kanai-Tajimi spectrum has a broader range in the frequency domain and has only few values close to zero.

To investigate the sensitivity of the interval DFT algorithm, the two above PSD functions S^J and S^{KT} with the respective given parameters are used hereafter. For the investigations, two PSD functions with different shapes are used to find similarities or differences in the resulting bound PSD functions. With only one form of PSD function, drawing conclusions becomes more difficult. In particular, the two PSD functions mentioned are used because one of them has many values close to 0 and a sharp peak, and the other has many values distant from 0.

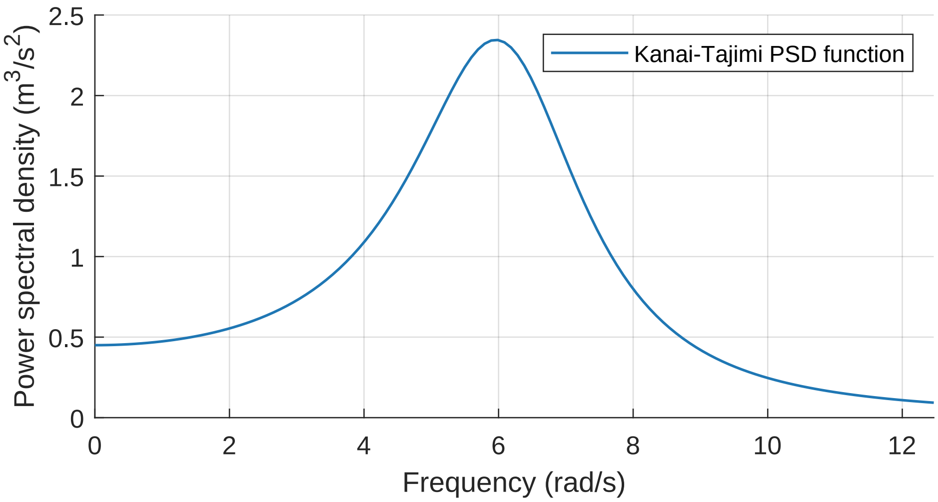

Fig. 2. Example for the Kanai-Tajimi PSD function.

The interval DFT algorithm

The DFT is applied to study the signal in the frequency domain. The DFT converts a signal $x = x_0, x_1, ..., x_{N-1}$ to a Fourier sequence $z = z_0, z_1, ..., z_{N-1}$ for $k = 0, ..., N-1$. Since many signals are subject to missing data, these must be taken into account during the transformation in order to obtain reliable results. One possibility is to reconstruct the data before the transformation. However, since the DFT is very sensitive to changes in the signal, as shown in Section 3, it is more reasonable to fill the missing data gaps with intervals and propagate them through the DFT. However, since the DFT is not able to transform such uncertainties, an algorithm was proposed that is capable to propagate interval uncertainties through the DFT and thus calculate exact bounds on the Fourier amplitude. This interval DFT algorithm is briefly described here, for a detailed explanation and examples the reader is referred to (Behrendt et al. 2022b).

Based on the *interval extension* of the DFT, obtained by replacing the real signal with their interval values for each frequency number k

$$\overline{\underline{z}}_k = \sum_{n=0}^{N-1} \overline{\underline{x}}_n e^{-\frac{i2\pi}{N}kn} = \sum_{n=0}^{N-1} \overline{\underline{x}}_n \cdot \left[\cos\left(\frac{2\pi}{N}kn\right) - i \cdot \sin\left(\frac{2\pi}{N}kn\right)\right], \qquad (8)$$

the algorithm computes two vertices for each iteration n of the sum in Eq. 8, resulting from the interval values of the n-th data point of the signal. In each iteration step, the vertices are added to the previous vertices. These vertices are represented in the 2-dimensional complex plane, where the real component is the x-coordinate and the imaginary component is the y-coordinate. From these vertices the convex hull is calculated, thus a polygon remains. The vertices of the convex hull are passed on to the next iteration step, while the remaining vertices have no influence on the calculation and are discarded. Once all data points of the signal have been iterated, the minimum and maximum distance of the convex hull to the origin of the coordinate system is determined, which defines the interval bounds of the absolute value of the transform

$$\overline{\underline{A}}_k = |\overline{\underline{z}}_k| = \sqrt{\left[\sum_{n=0}^{N-1} \overline{\underline{x}}_n \cos\left(\frac{2\pi}{N}kn\right)\right]^2 + \left[\sum_{n=0}^{N-1} \overline{\underline{x}}_n \sin\left(\frac{2\pi}{N}kn\right)\right]^2}. \tag{9}$$

The absolute values of the vertices in the convex hull are calculated for this purpose. If the origin of the coordinate system is within the convex hull, the lower bound is 0, otherwise it is defined by the minimum absolute value. The upper bound is always determined by the maximum absolute value. Thus, an upper and lower bound of the Fourier amplitude can be computed for each frequency number k.

MISSING DATA

A common problem when using real data records is that of missing data. The causes of missing data range from simple measurement errors to total sensor failure. It is possible that the sensor is damaged by the event it is supposed to record, e.g. an earthquake, and makes incorrect recordings or stops recording completely. In addition, the sensors may be temporarily unavailable due to maintenance. If the period of unavailability is sufficiently short, intervals can be used to bridge this gap. These causes introduce uncertainty into the time signal. Although there are various reconstruction methods available, as mentioned in Section 1, only simple reconstruction methods are used here. The main objective of this work is to investigate the performance and sensitivity of the interval DFT algorithm. Finding a suitable reconstruction method or assessing the quality of the reconstruction methods is not the aim of this work.

The reconstructed data are represented by intervals, accounting for uncertainties induced through the reconstruction. Thus, the reconstructed signal is passed to the interval DFT algorithm as an interval signal. Fig. 3 shows the signal under investigation, generated from the JONSWAP PSD function, with two examples each with missing data.

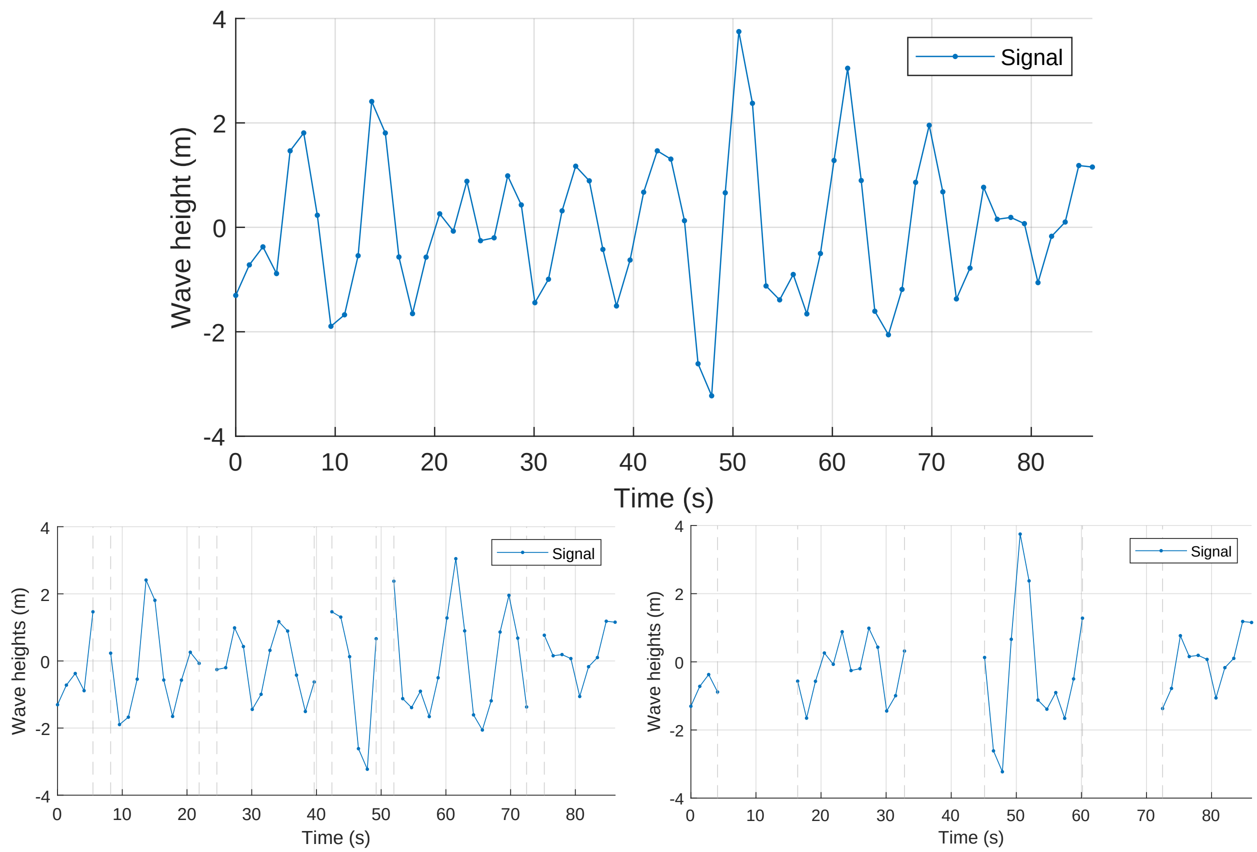

Fig. 3. A time signal generated from the JONSWAP PSD function (Eq. 6) with SRM (Eq. 5) consisting of 64 data points (top). The lower left plot has five missing data points, while the lower right plot has three missing data gaps with eight missing data points each.

If a signal in time domain is certainly known, it can be transformed to the frequency domain via the DFT without loss of information. In fact, the DFT is sensitive to small changes in the signal. To demonstrate the sensitivity of the spectrum to the missing data problem, the signal in Fig. 3, consisting of 64 data points, is investigated. The target PSD function, i.e. the PSD function of the signal without missing data computed with Eq. 4, is depicted with the PSD functions of the same signal with 5%, 10% and 25% missing data, which are reconstructed by linear interpolation between the two adjacent non-missing data points, see Fig. 4. The position of the missing data is randomly chosen. The interpolated values are treated as discrete values instead of intervals first. Although linear interpolation is not considered as a reconstruction method in this work, it can be used to illustrate the aforementioned sensitivity. It can be clearly seen that the transformations have the same shape and peak frequency, but are in part very different from the target spectrum and are not as smooth. Since reconstructed data accordingly do not allow a reliable transformation into the

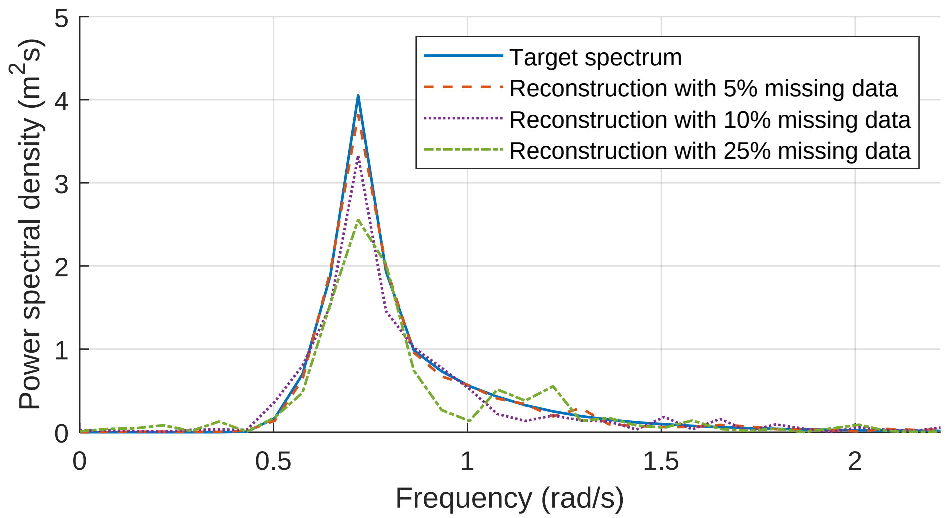

Fig. 4. Influence of the linear interpolation on the amplitude of the DFT.

frequency domain and do not take uncertainties into account, it is reasonable to derive bounds in which the actual spectrum may be located. The algorithm presented in Section 2 is applicable for this purpose.

In this work, two reconstruction methods are employed:

1. A method based on artificial inflation of the “true” data point using the sample standard deviation s (Eq. 10) of the entire signal before removing data. An interval whose width is $[-s, s]$ replaces the missing data.
2. A method that replaces the missing data by an interval determined by the minimum and maximum value of the entire signal.

The sample standard deviation s of the signal is defined as

$$s = \sqrt{\frac{\sum_{n=0}^{N-1} (x_n - \tilde{x})^2}{N-1}}, \tag{10}$$

where $\tilde{x}$ is the sample mean of the signal, n is the data point index and N the total number of data points.

Both methods serve only as very simple tools for determining the sensitivity of the interval DFT for signals with missing data. In practical applications, these methods should be replaced by advanced reconstruction methods, see examples in Section 1.

CASE STUDIES

In this section, the influence of missing data on the bounds of the estimated PSD function is investigated. Specifically, interval width, the number of missing data, the gap length, and the distribution of missing data within the signal are examined. The study is conducted as part of a Once-at-a-time (OAT) sensitivity analysis, such that only one of the above-mentioned influence factors is changed while the others are kept constant. The signal under investigation is generated by SRM (Eq. 5) with the underlying PSD function in Eq. 6 from (Hasselmann et al. 1973). The positions of the missing data in the signal are simulated in random order, uniformly within the length of the signal. A study is also conducted to investigate the influence of the position of the missing data, comparing the uniformly distributed missing data with binomially distributed missing data. In order to obtain the best possible comparison, the same signal is used in all studies of this work.

Sensitivity to interval width

Let h_x be the width of the interval gap, thus this determines the interval uncertainty. To investigate the sensitivity of the interval uncertainty h_x in time domain to the interval uncertainty in the frequency domain, missing data gaps of length $l_g \in \{1, 3, 5, 7, 9, 11\}$ are randomly generated, where the gap length is given as the number of missing time points.

The interval uncertainty h_x of these gaps is successively increased from 0.1 to 10. To determine the sensitivity, the area between the upper and lower bound of the resulting PSD is determined. The results are depicted in Fig. 5 for two signals generated from the JONSWAP PSD function with 64 and 128 data points.

The area between the bounds has a linear trend in the beginning which turns into a non-linear trend even with low interval uncertainty and small gaps. This non-linearity becomes stronger the larger the gap becomes. At many frequency points, the lower bound has already reached 0. For larger gaps, the lower bound is mostly zero, which explains why in Fig. 5 the start of the non-linear behaviour is appreciated for lower interval uncertainty.

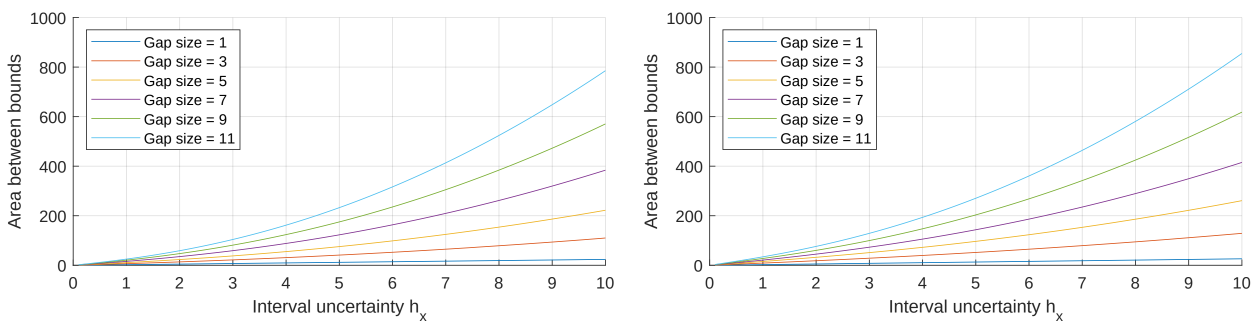

Fig. 5. Area between upper and lower bound for a signal with 64 data points (left) and a signal with 128 data points (right) for increasing interval uncertainty h_x and different lengths of the gap $l_g \in \{1, 3, 5, 7, 9, 11\}$ for the JONSWAP PSD function.

The same investigations are carried out for the Kanai-Tajimi PSD function, see Fig. 6. Two signals are considered, which were generated from the Kanai-Tajimi PSD function. One signal with 64 data points and the other one with 128 data points. As in the JONSWAP example before, first a linear trend can be observed, which turns into a non-linear trend as soon as the lower bound reaches 0. It should be noted, however, that in all cases the non-linear trend starts later and that it is not as strong as in the previous example with the JONSWAP PSD function (Fig. 5). This is due to the fact that fewer values of the lower bound of the PSD reach 0. In addition, these values need a relatively high interval uncertainty in the signal to reach zero in the frequency domain, since the target spectrum is significantly farther from 0 than, for example, the JONSWAP target spectrum. Therefore, the non-linear trend is not as significant.

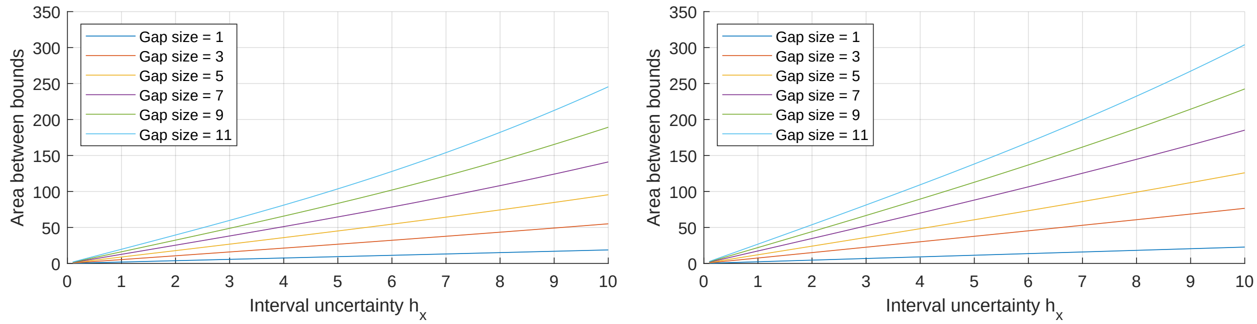

Fig. 6. Area between upper and lower bound for a signal with 64 data points (left) and a signal with 128 data points (right) for increasing interval uncertainty h_x and different lengths of the gap $l_g \in \{1, 3, 5, 7, 9, 11\}$ for the Kanai-Tajimi PSD function.

Number of missing data

In the following example, the interval uncertainty has been kept constant and corresponds to the sample standard deviation s of the signal (Eq. 10). The number of missing data points, on the other hand, has been gradually increased to investigate the influence of the number of missing data on the bounds of the PSD function. In Fig. 7, the reconstructed signals and the bounds of the estimated PSD functions are shown for 5%, 10%, 25% and 50% missing data in the signals, which consist of 64 and 128 data points, respectively. The results show that a small amount of missing data (e.g. 5% or 10%) can be captured well with the interval DFT algorithm. The bounds enclose the estimated PSD function of the discrete signal relatively tightly and are therefore very useful for quantifying the uncertainties. Also, the bounds of the PSD function for a higher amount of missing data in the signal (up to 50% in this example) can still be considered, despite the relatively wide bounds, e.g. for a worst-case consideration where only the upper bound is used.

In the following, the same example is shown, but the data was reconstructed using method (2), see Fig. 8 for the reconstructed signals and the bounds of the PSD functions in frequency domain.

The results also show here that small amounts of missing data can be mapped well in the frequency domain even with reconstruction method (2). With higher numbers of gaps, however, the determination of the bounds in the frequency domain reaches its limitation, as the computed bounds are very high and can no longer be used for practical purposes. For example, the bounds from the previous example with 50% missing data have a lower interval uncertainty than the signal

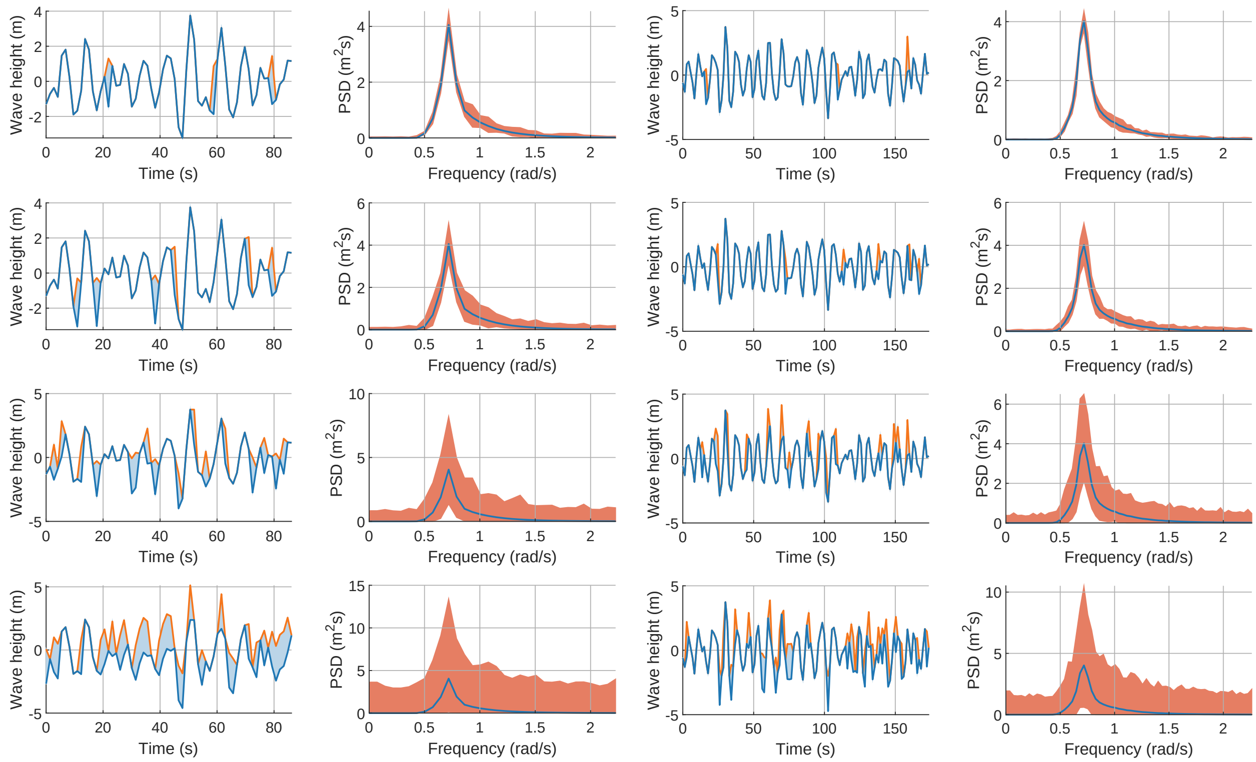

Fig. 7. Signal with 5%, 10%, 25% and 50% missing data (top to bottom) reconstructed using method (1) and corresponding bounded JONSWAP PSD function. On the left is the signal with 64 data points and the corresponding bounded PSDs, on the right is the signal with 128 data points and the corresponding bounded PSDs.

with 25% missing data in this example. This yields in particular that if there is little missing data, reconstruction can be carried out conservatively with wide intervals. Conversely, if the number of missing data is large, a method with a more accurate reconstruction is required.

The values for the given examples of the JONSWAP PSD function (Figs. 7 and 8) and longer signals are given in Table 1 for a comparison.

As a measure for uncertainty, the area between upper and lower bound is utilised. Fig. 9 shows this for an increasing number of missing data reconstructed with the two methods for the signal with 64 data points and 128 data points. Due to possible random fluctuations, as the position of missing data is randomly chosen, this simulation was carried out 100 times in order to average out these fluctuations. As expected, there is a significantly higher area between the bounds when using reconstruction method (2) compared to reconstruction method (1).

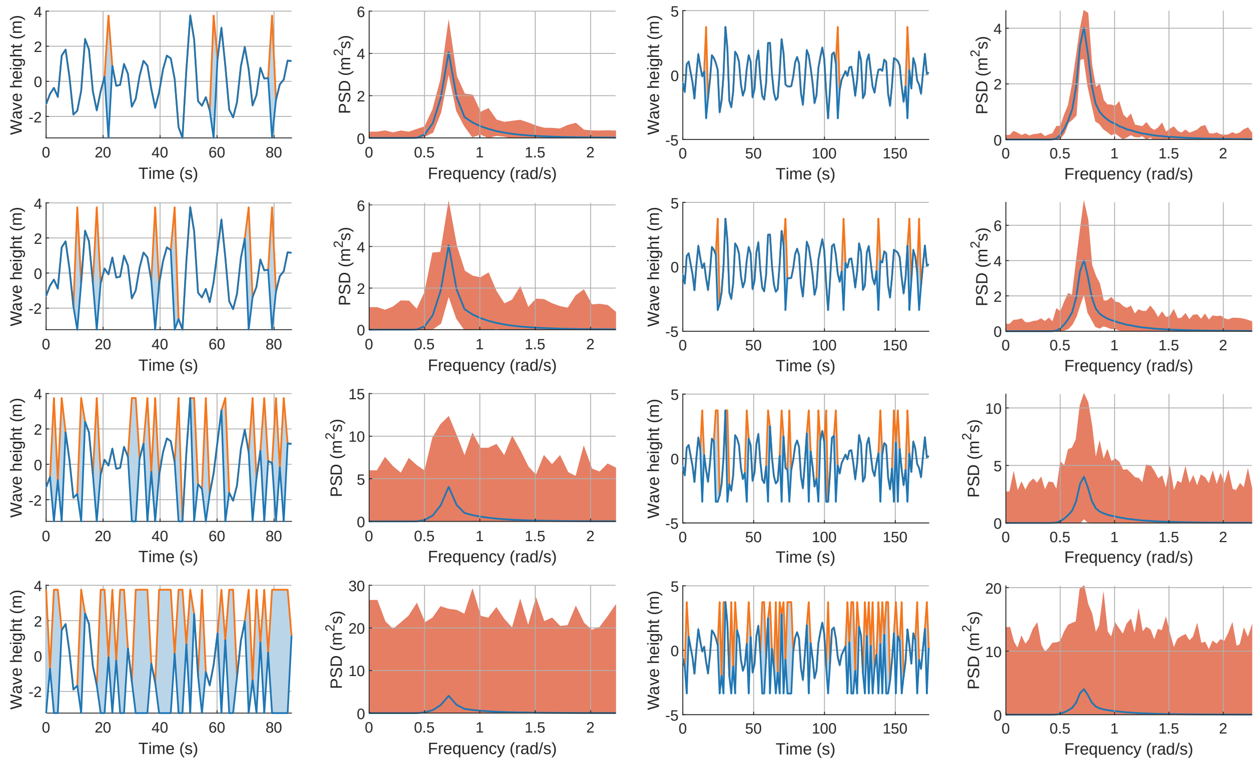

Fig. 8. Signal with 5%, 10%, 25% and 50% missing data (top to bottom) reconstructed using method (2) and corresponding bounded JONSWAP PSD function. On the left is the signal with 64 data points and the corresponding bounded PSDs, on the right is the signal with 128 data points and the corresponding bounded PSDs.

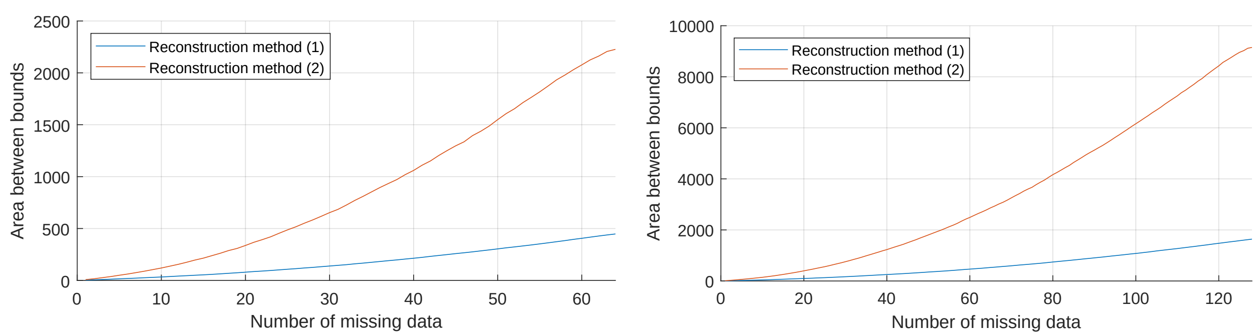

Fig. 9. Area between upper and lower bound investigated for the number of missing data for the JONSWAP PSD function. On the left for the signal with 64 data points, on the right for the signal with 128 data points.

The same investigations are carried out for a signal with 64 data points and a signal with 128 data points, generated from the Kanai-Tajimi PSD function. First, the sample standard deviation was utilised to reconstruct the missing data. The results are given in Fig. 10. Also, in these

TABLE 1. Area between upper and lower bound for the JONSWAP PSD function for the investigations on the number of missing data.

	Signal length	5%	10%	25%	50%
reconstruction method (1)	64	8.2261	18.3467	58.4300	150.0321
	128	12.3274	25.5648	76.5203	182.8820
	256	18.0512	37.1025	102.6569	241.2242
	512	228.5536	525.4112	2036.9685	6674.7539
	1024	674.2047	1678.9022	7233.0551	25114.9873
reconstruction method (2)	64	24.5173	59.7119	241.4550	719.4045
	128	34.5642	79.4739	288.2805	835.8317
	256	50.3423	109.8055	362.3446	1031.7899
	512	912.6849	2594.6585	13055.5258	47189.8391
	1024	4044.5176	13196.6865	71951.7682	270806.8143

examples it can be observed, that a lower number of missing data leads to practical usable results. For instance, a proportion of 5% or 10% missing data results in a bounded PSD with a moderate uncertainty. However, if the proportion is increased, the bounds can be very wide and are no longer useful for practical purposes, except for some worst-case scenario investigations. This becomes particularly clear in the examples with 50% missing data, since the bounds are very distant from the target spectrum and the shape of these bounds also barely shows similarities to the target spectrum.

In the second example, reconstruction method (2) was utilised for reconstructing the missing data in the signals. Again, the two identical signals with 64 data points and 128 data points as in the previous example are utilised. The results are depicted in Fig. 11. As expected, due to a higher uncertainty in the signal, the bounds of the PSD function also exhibit a higher uncertainty. Thus, the bounds are much wider. For low proportions of missing data, such as 5%, useful results can be obtained. However, with an increasing number of missing data the bounds become quickly very wide and are not useful anymore. The example with 10% missing data might be used for a worst-case scenario, more interval uncertainty in the signal will result in bounded PSD functions which are no longer practical.

An interesting observation is that already with a small amount of missing data (e.g. 5% or 10%) the resulting transformations become very spiky. This is specifically evident in the results of the signal with 128 data points. The reason for this could be that only a single missing data point

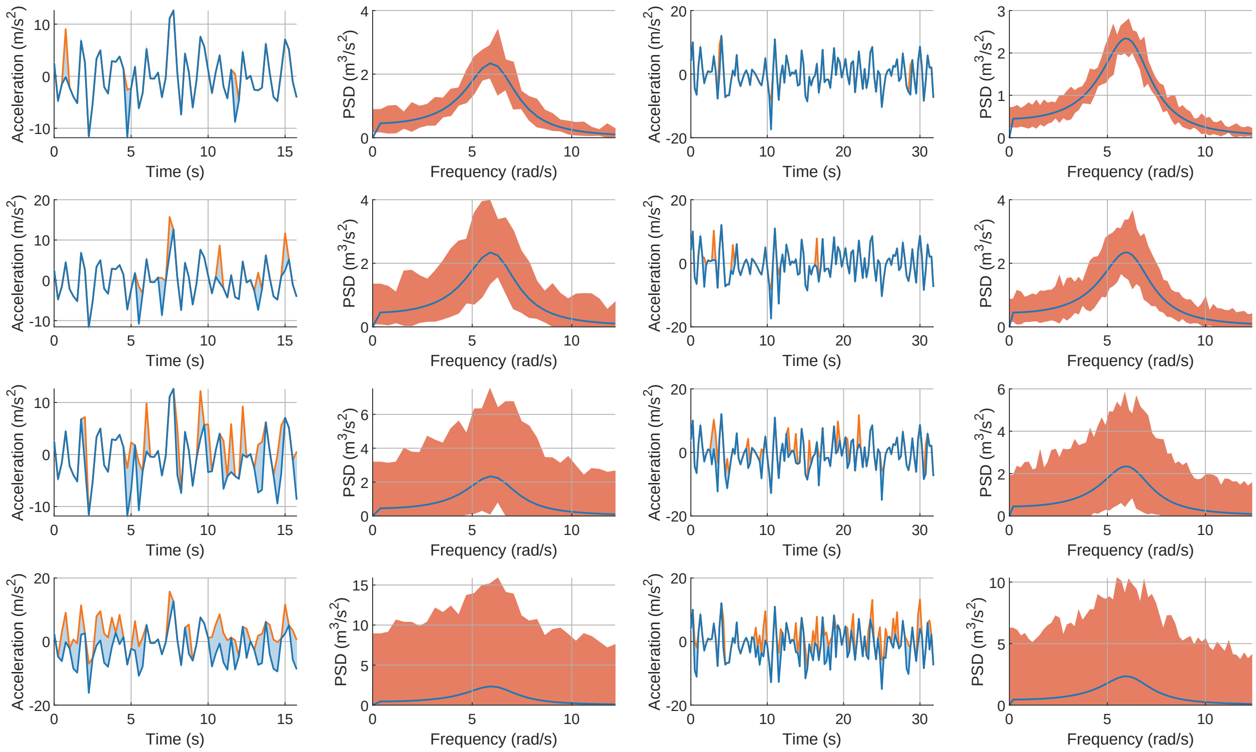

Fig. 10. Signal with 5%, 10%, 25% and 50% missing data (top to bottom) reconstructed using method (1) and corresponding bounded Kanai-Tajimi PSD function. On the left is the signal with 64 data points and the corresponding bounded PSDs, on the right is the signal with 128 data points and the corresponding bounded PSDs.

reconstructed with extreme values leads to a large distortion of the actual signal and its spectral characteristics. This result corresponds to the sensitivity analysis in Section 3, see Fig. 4, where the influence of linearly interpolated missing data points was investigated. In this example, the reconstruction with 25% missing data is also very spiky.

For better overview, the values for the examples of the Kanai-Tajimi PSD function (Figs. 10 and 11) and longer signals are given in Table 2.

Again, the area between the upper and lower bound is used as a measure of uncertainty. In Fig. 12 this is depicted for an increasing number of missing data reconstructed using the two methods for the signal with 64 data points and 128 data points. Also as expected, the range between the bounds is significantly larger for reconstruction method (2) than for reconstruction method (1).

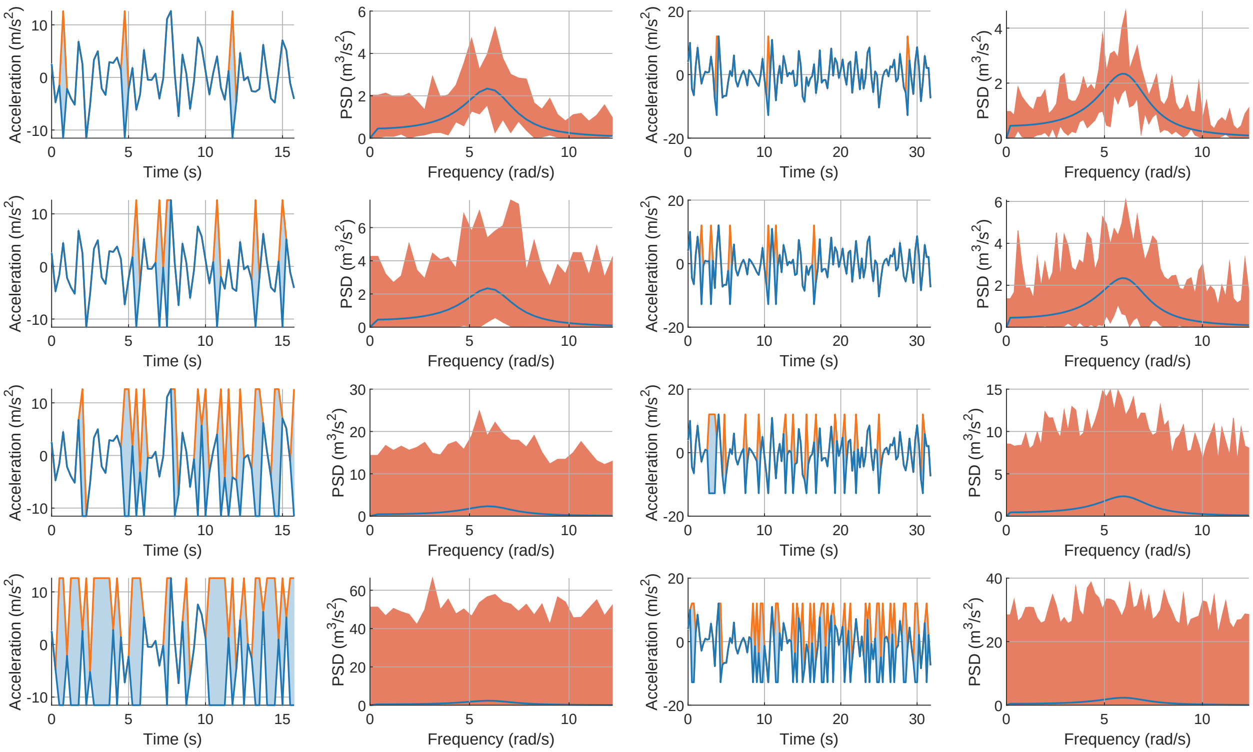

Fig. 11. Signal with 5%, 10%, 25% and 50% missing data (top to bottom) reconstructed using method (1) and corresponding bounded Kanai-Tajimi PSD function. On the left is the signal with 64 data points and the corresponding bounded PSDs, on the right is the signal with 128 data points and the corresponding bounded PSDs.

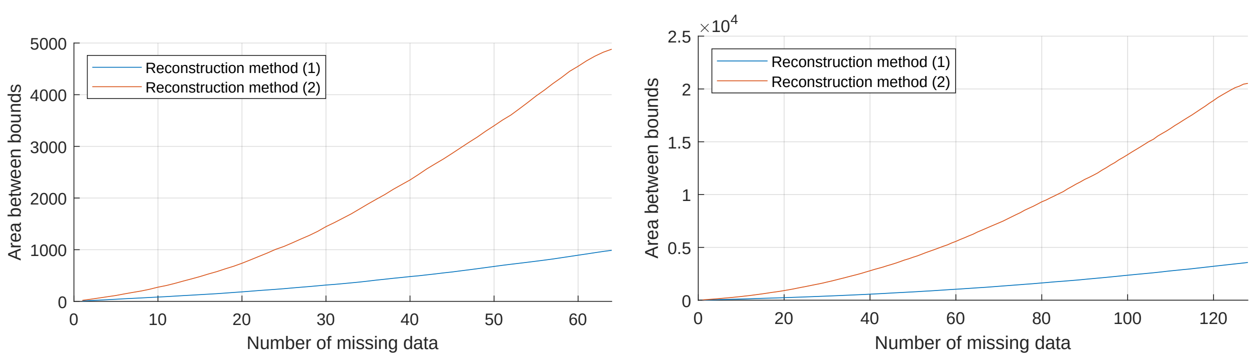

Fig. 12. Area between upper and lower bound investigated for the number of missing data for the Kanai-Tajimi PSD function. On the left for the signal with 64 data points, on the right for the signal with 128 data points.

TABLE 2. Area between upper and lower bounds for the Kanai-Tajimi PSD function for the investigations on the number of missing data.

	Signal length	5%	10%	25%	50%
reconstruction method (1)	64	23.3376	50.2520	136.9106	344.2631
	128	32.1355	67.9602	188.7277	424.1079
	256	46.3055	94.7233	256.1252	547.8648
	512	589.2857	1245.5312	4547.8639	14423.2745
	1024	1660.4582	3862.1247	15926.0403	53660.1195
reconstruction method (2)	64	61.6042	141.5591	522.5090	1628.4539
	128	81.9223	183.2325	663.2781	1927.7293
	256	120.1866	251.0239	755.2968	2022.6513
	512	2060.5408	5638.1834	27860.8249	100661.0096
	1024	7744.4521	24390.7841	129894.5636	481486.4526

Gap size of missing data

Recall that gap size is given as the number of missing time points, and it is also referred to as gap length. To determine the influence of the gap length, different scenarios were evaluated. The gap lengths $l_g \in \{1, 20, 40, 60\}$ were artificially inserted into the signals generated with the JONSWAP PSD function with a length of 64 data points and 128 data points. After the gaps were reconstructed using method (1), the corresponding transformations were computed. These are shown in Fig. 13. It can be seen that small gaps filled with the intervals provide a good transformation and the bounds

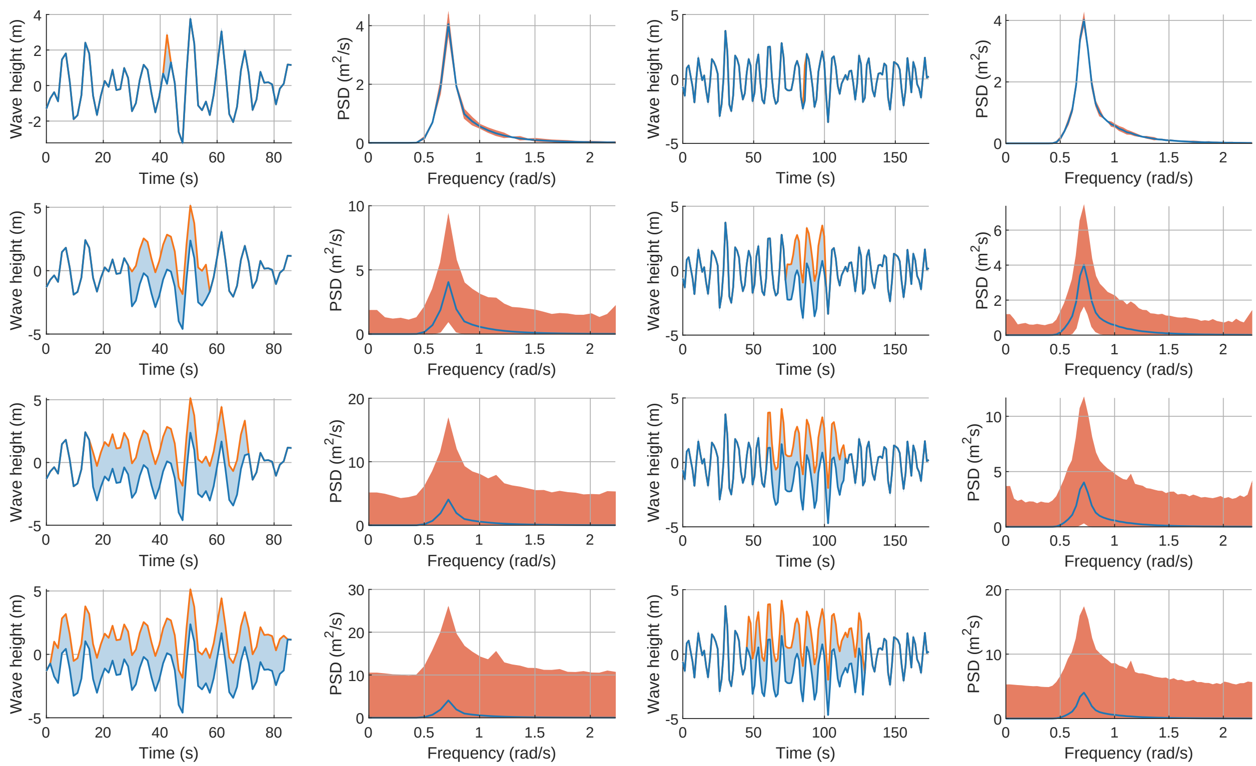

Fig. 13. Signals with a gap of length $l_g \in \{1, 20, 40, 60\}$ of missing data (top to bottom) reconstructed by method (1) and corresponding bounded JONSWAP PSD function. On the left is the signal with 64 data points and the corresponding bounded PSDs, on the right is the signal with 128 data points and the corresponding bounded PSDs.

are relatively tight around the target PSD function. The interval DFT algorithm can also handle larger gaps well, although the bounds of the transformation are comparatively large. Nevertheless, these can be used, for example, to design for a worst-case when only the upper bound with the largest power content is used for planning and simulation.

For completeness, the same investigations are carried out with reconstruction method (2). The same length of gaps l_g as in the previous example were inserted in the signal but reconstructed with minimum and maximum of the signal as intervals. The signals and transformations are given in Fig. 14. The reconstruction with the minimum and maximum of the signal already reaches its

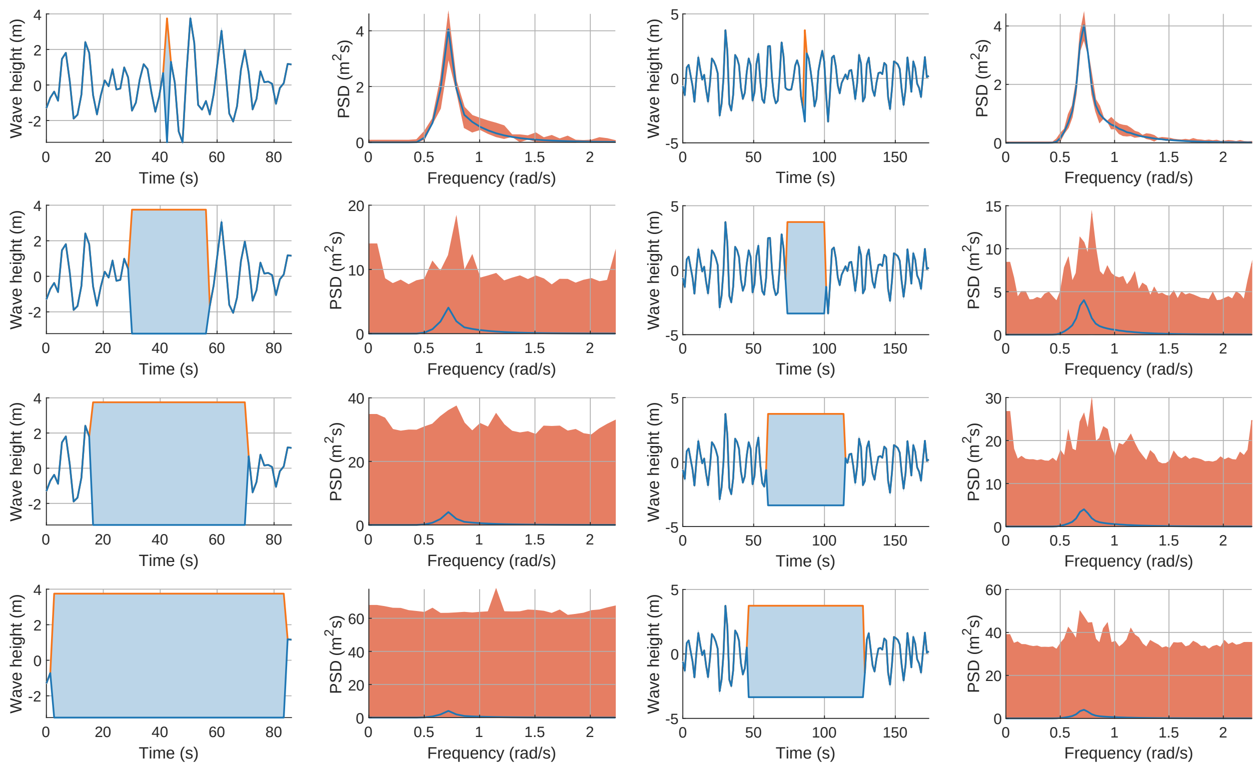

Fig. 14. Signals with a gap of length $l_g \in \{1, 20, 40, 60\}$ of missing data (top to bottom) reconstructed by method (2) and corresponding bounded JONSWAP PSD function. On the left is the signal with 64 data points and the corresponding bounded PSDs, on the right is the signal with 128 data points and the corresponding bounded PSDs.

limitations for smaller gaps. Although the interval DFT algorithm provides exact bounds, these are very large due to the highly conservative reconstruction method. Even a gap with 20 data points provides bounds that are very distant from the target PSD function. For even larger gaps, the shape of the exact transformation is no longer reflected. It is shown again in this example that too large intervals in the time domain lead to extremely large bounds in the frequency domain. To counteract this behaviour, the intervals in the time signal should be chosen reasonably.

The values for the examples of the JONSWAP PSD function (Figs. 13 and 14) and longer signals are given in Table 3 for a comparison.

TABLE 3. Area between upper and lower bounds for the JONSWAP PSD function for the investigations on the length of the gap size.

	Signal length	1	20	40	60
reconstruction method (1)	64	3.2585	77.0459	209.5725	407.5691
	128	3.5751	97.3246	245.7535	450.3334
	256	6.2338	134.1172	317.1225	558.9399
	512	7.9715	81.8397	381.5804	641.9508
	1024	11.5807	114.8309	501.2821	814.2205
reconstruction method (2)	64	8.8466	307.5685	1004.9875	2079.0797
	128	9.4480	376.9863	1144.8863	2307.9452
	256	16.7009	481.9337	1380.5991	2738.6998
	512	22.5879	258.8454	1646.1033	3193.1749
	1024	39.9040	439.4282	2612.9817	4969.5717

In the following, the area between upper and lower bound is determined for an increasing gap size. Since the length of the gap naturally corresponds to the number of missing data, no significant differences between Fig. 15 and Fig. 9 in the previous section can be detected. This indicates that the position of the missing data has a minor role in determining the uncertainty, but the number has a major role.

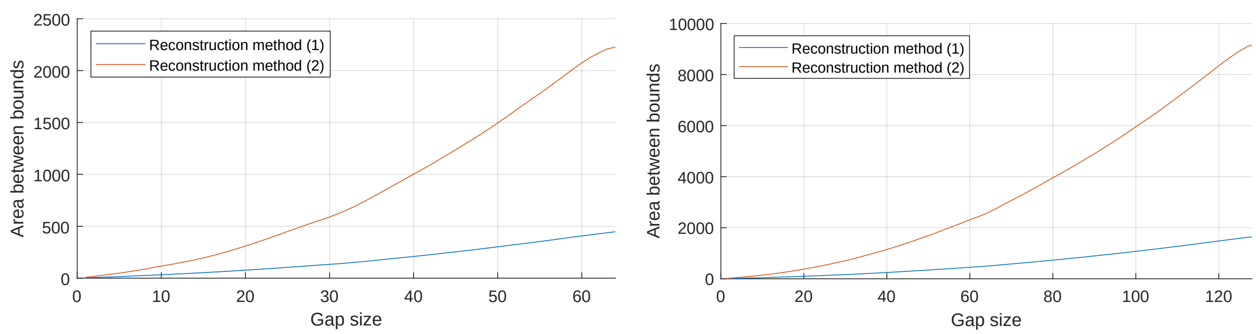

Fig. 15. Area between upper and lower bound investigated for the length of the gap for the JONSWAP PSD function. On the left for the signal with 64 data points, on the right for the signal with 128 data points.

The investigations on the influence of the gap size are carried out for the Kanai-Tajimi PSD function as well. First, missing data with a length of $l_g \in \{1, 20, 40, 60\}$ is simulated in the signals with length 64 and 128. Next, those missing data gaps are reconstructed using method (1). As it can be seen in Fig. 17, only small gaps, such as $l_g = 1$, will lead to useful results. Contrary to the example with the JONSWAP PSD function, a gap size of $l_g = 20$ can only be used for a worst-case scenario a the bounds are already very distant from the target PSD function. Even larger gaps only reflect roughly the shape of the target PSD function, but the bounds are far too wide to use them meaningfully in simulations.

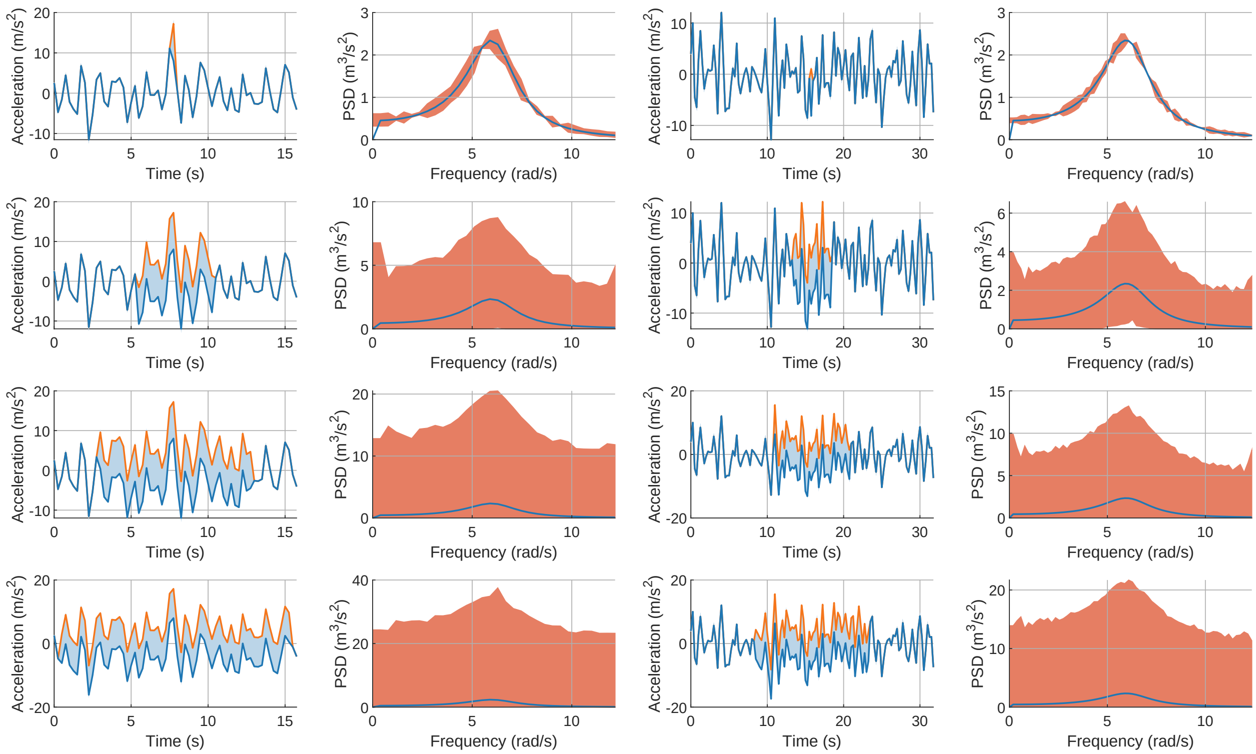

Fig. 16. Signals with a gap of length $l_g \in \{1, 20, 40, 60\}$ of missing data (top to bottom) reconstructed by method (1) and corresponding bounded PSD functions for the Kanai-Tajimi PSD function. On the left is the signal with 64 data points and the corresponding bounded PSDs, on the right is the signal with 128 data points and the corresponding bounded PSDs.

The same investigations are carried out with reconstruction method (2). As it can be seen in Fig. 17, only small gaps reconstructed with extreme values can be used in practical application. Even a reconstructed signal with a gap of $l_g = 20$ missing data may only be used for a worst-case scenario. In this example it can be observed again that the reconstruction with extreme values leads to a very spiky transformation, see for comparison Fig. 11. In addition, however, it can be observed that the signal with 60 missing data points in Fig. 11 is very smooth after reconstruction, as all intervals are identical. Although the reconstructed signal has nothing in common anymore with the original signal, it can be seen that the corresponding transformation looks much smoother than the transformations of signals with less missing data. These observations support the assumption that the spectral characteristics are distorted by the reconstruction with extreme values.

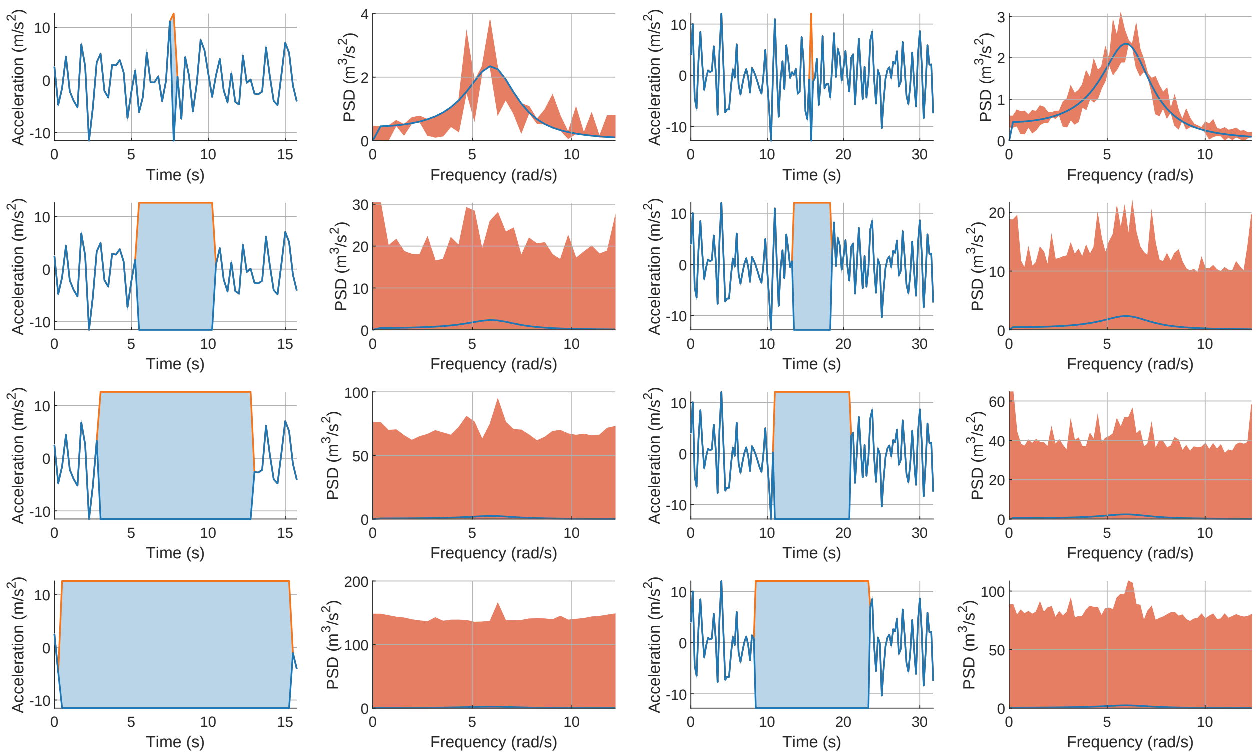

Fig. 17. Signals with a gap of length $l_g \in \{1, 20, 40, 60\}$ of missing data (top to bottom) reconstructed by method (2) and corresponding bounded PSD functions for the Kanai-Tajimi PSD function. On the left is the signal with 64 data points and the corresponding bounded PSDs, on the right is the signal with 128 data points and the corresponding bounded PSDs.

The values for the examples of the Kanai-Tajimi PSD function (Figs. 16 and 17) are given in Table 4 for a comparison.

The area between the bounds for an increasing gap size was investigated for the two signals with length of 64 and 129 data points generated for the Kanai-Tajimi PSD function. The results are given in Fig. 18. As for the example with the JONSWAP PSD function before, no significant differences between the length of the gap (Fig. 18) and the number of missing data (Fig. 12) can be observed. Thus, it can be concluded that the position of the missing data is of minor importance and rather

TABLE 4. Area between upper and lower bounds for the Kanai-Tajimi PSD function for the investigations on the length of the gap size.

	Signal length	1	20	40	60
reconstruction method (1)	64	8.6220	181.6365	467.8687	889.8082
	128	10.4224	237.4579	560.9135	1006.6424
	256	15.6008	322.9332	706.4987	1217.2839
	512	23.2987	223.0909	938.6201	1517.6868
	1024	32.1328	316.5083	1287.7693	1993.7773
reconstruction method (2)	64	19.3471	690.8520	2236.9974	4527.3797
	128	27.7411	868.6494	2639.2345	5291.6959
	256	40.1799	990.6577	2717.2864	5366.6247
	512	66.4938	649.1612	3744.2217	7133.9498
	1024	101.7266	1026.0292	5302.0325	9771.2658

the number of missing data points has the decisive influence. Nevertheless, for completeness the influence of the position and the distribution of the missing data will be investigated in the next section.

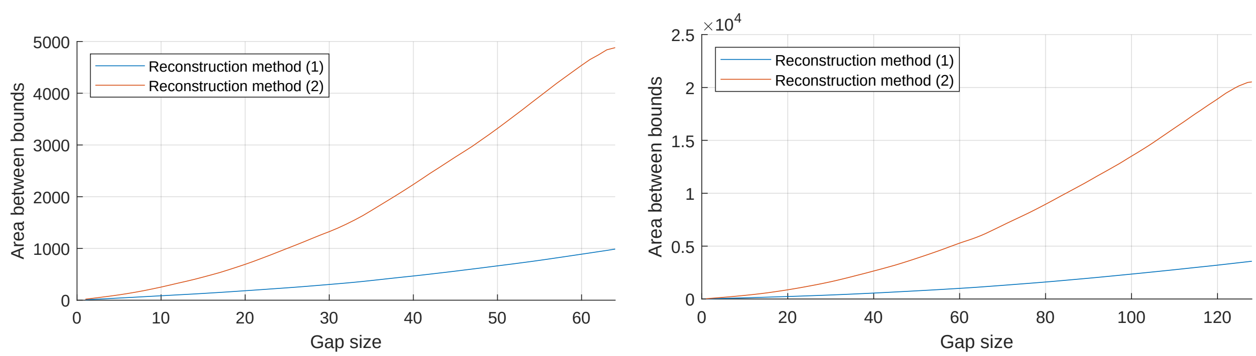

Fig. 18. Area between upper and lower bound investigated for the length of the gap for the Kanai-Tajimi PSD function. On the left for the signal with 64 data points, on the right for the signal with 128 data points.

Position and distribution of missing data

As expected, the position of the missing data can influence the shape and subsequently the area between the bounds. Although this will not affect a subsequent simulation and its results significantly, it is important to investigate this phenomenon. For the sake of brevity, this first investigation is carried out only for the JONSWAP PSD function and reconstruction method (1) as it has been shown in the previous sections that reconstruction method (2) cannot be used for real phenomena if the number of missing data is sufficiently high.

For this analysis, the signal generated from the JONSWAP PSD function with 64 data points was utilised in four different scenarios. In each of them, a single missing data point was randomly generated and reconstructed with method (1), i.e. the position of the the missing data is different in each of the four simulations, but due to the utilisation of reconstruction method (1) the interval uncertainty of this point is identical. As it can be seen in Fig. 19, each of the computation yields a slightly different shape of the bounds. The reason for this fluctuation in the bounds may be, that each point carries different information in terms of the spectral characteristics. Therefore, it is reasonable to expect a different shape of the bounds, if another point is missing. However, as the interval uncertainty is kept constant in each of the experiments, the bounds yield roughly the same area between the bounds, which corresponds to a similar total potential power of the PSD function.

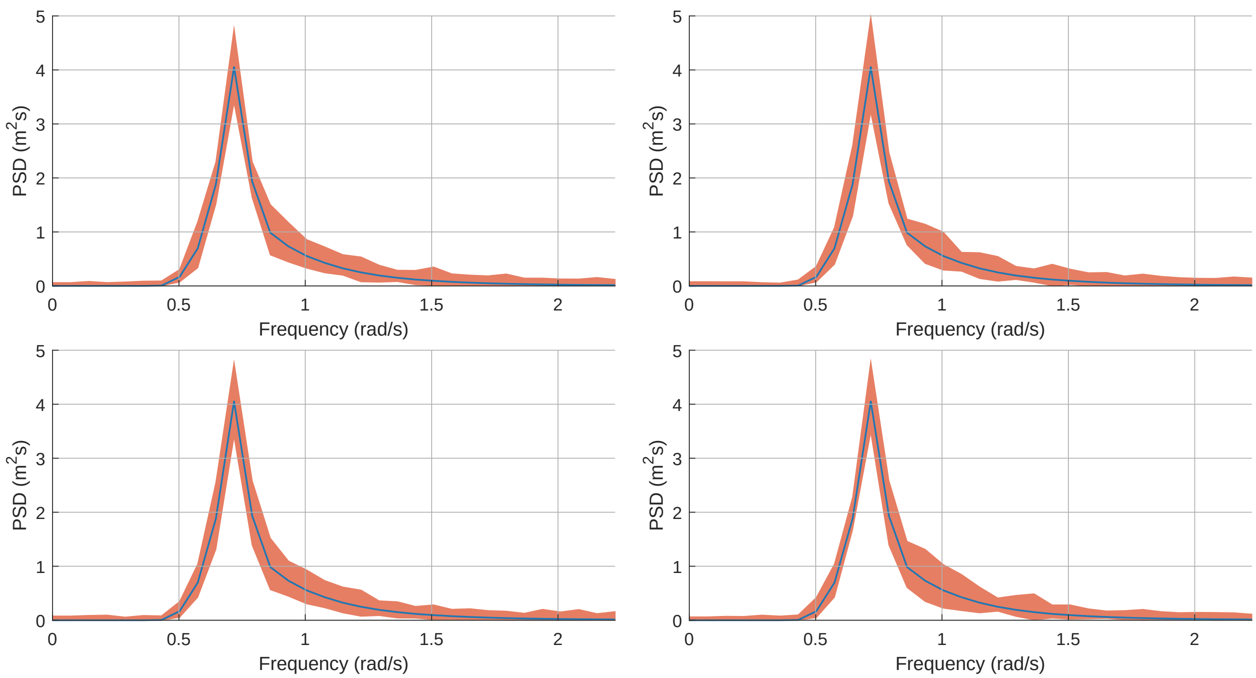

Fig. 19. Examples for different resulting bounded PSD functions depending on the position of the missing data point. Each transformed signal has exactly one reconstructed data point with identical interval uncertainty at different positions.

The impact of the position of the missing data serves as a motivation for the following investigation, namely the influence of the distribution of the missing data within a signal. For the investigations a uniform distribution and a binomial distribution were utilised to simulate the missing data and to investigate their influence on the transformation to the frequency domain. The bounded PSD functions of the reconstructed signal with 4, 8, 16 and 32 missing data are depicted in Fig. 20. It can be seen that the influence of the position of the missing data is of minor rel-

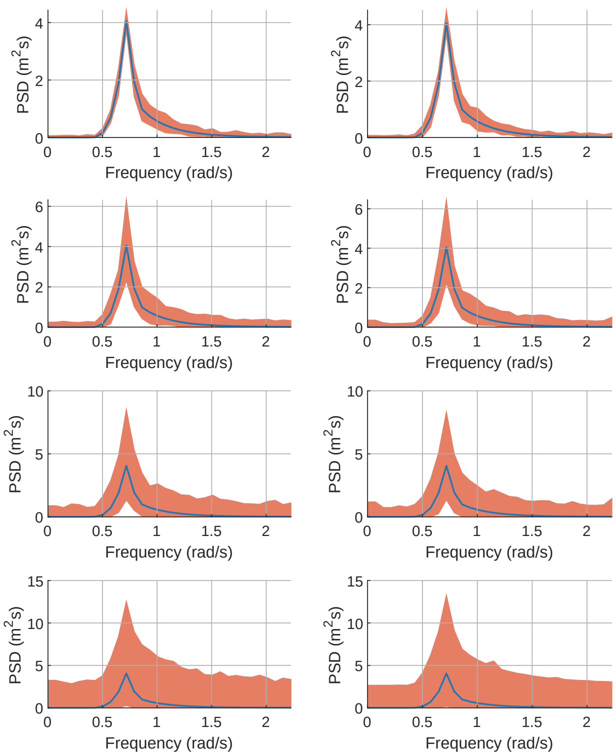

Fig. 20. Influence of the distribution of missing data within the signal for 4, 8, 16 and 32 missing data (top to bottom) reconstructed with method (1) for the JONSWAP PSD function. In the left column the corresponding bounded PSDs with uniformly distributed missing data, in the right column the corresponding bounded PSDs with binomially distributed missing data.

evance. Although the transformed signals shown are only specific cases, they are nevertheless representative for the general case. This statement can be supported by the fact that this simulation has been carried out several times, but the results are always identical. The interval transforms look almost identical in each case, regardless of the distribution of the missing data. In addition, the area between the bounds is also almost identical, see Table 5. Thus, these results support the assumptions on the position of the missing data in the beginning of this section.

The same investigations as above are carried out for the Kanai-Tajimi PSD function. Again, a

TABLE 5. Area between upper an lower bound for the JONSWAP PSD function for the distribution of the missing data within the signal.

	4	16	32	64
uniform distribution	12.382	24.784	57.840	153.712
binomial distribution	11.228	25.341	57.155	144.959

uniform distribution and a binomial distribution were utilised. The missing data were randomly generated within the signal for the scenarios of 4, 8, 16 and 32 points and reconstructed with method (1). The results of the bounded PSDs are given in Fig. 20. Similar to the previous example no major differences between the bounds can be observed. However, small fluctuations are evident. This is due to the position of the missing data within the signal. As the determined area between those bounds confirm, see Table 6, the differences are relatively small. Therefore, it can be concluded that the distribution of missing data has a similar effect as in the previous example. The position of the missing data has an influence in the sense that, depending on the spectral characteristics of the respective missing data points, they are passed on to the PSD bounds in a distorted way.

TABLE 6. Area between upper an lower bound for the Kanai-Tajimi PSD function for the distribution of the missing data within the signal.

	4	16	32	64
uniform distribution	29.015	64.227	137.103	345.842
binomial distribution	33.780	67.749	140.664	337.467

Although the values between the distributions in Tables 5 and 6 are not identical, a clear trend can be seen. This is because, as mentioned earlier, the exact position of the missing data point has an influence on the area between the bounds. The results of this investigations confirm that this influence is negligible, see Fig. 21 for a visual assessment.

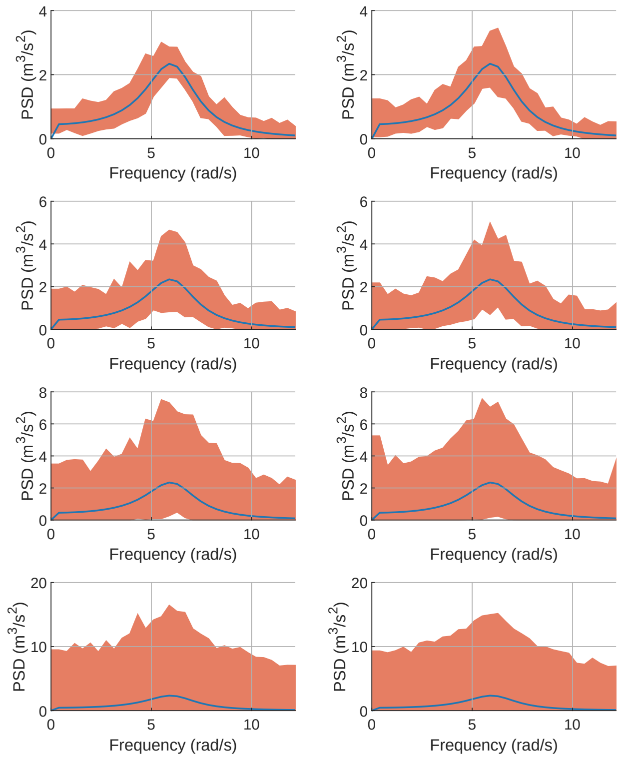

Fig. 21. Influence of the distribution of missing data within the signal for 4, 8, 16 and 32 missing data (top to bottom) reconstructed with method (1) for the Kanai-Tajimi PSD function. In the left column the corresponding bounded PSDs with uniformly distributed missing data, in the right column the corresponding bounded PSDs with binomially distributed missing data.

Interaction between number of missing data and interval width

As it was found that the number of missing data and the interval width of the reconstruction in the input signal have the highest influence on the area between the bounds of the PSD after propagation, the interaction between both is investigated in this section as a last case study. Again, for the sake of brevity, the investigations are carried out only for the signal with 64 data points for both, the JONSWAP PSD function and the Kanai-Tajimi PSD function. As per a space product, the missing data was increased and for each number of missing data, the interval uncertainty was successively increased from 0.1 to 10, similarly as in Section 4.

In Fig. 22 the results of the area between upper and lower bound for the JONSWAP PSD function are given. As it can be seen, a relatively low number of missing data combined with a low interval uncertainty in the signal will result in useful results. Also, a high number of missing data combined with a low interval uncertainty or a low number of missing data combined with a high interval uncertainty still provides useful results. However, when both quantities take on high values, a non-linear trend quickly produces results that are no longer useful for practical purposes. The corresponding area between the bounds is simply too large to obtain reasonable conclusions and meaningful results in subsequent simulations.

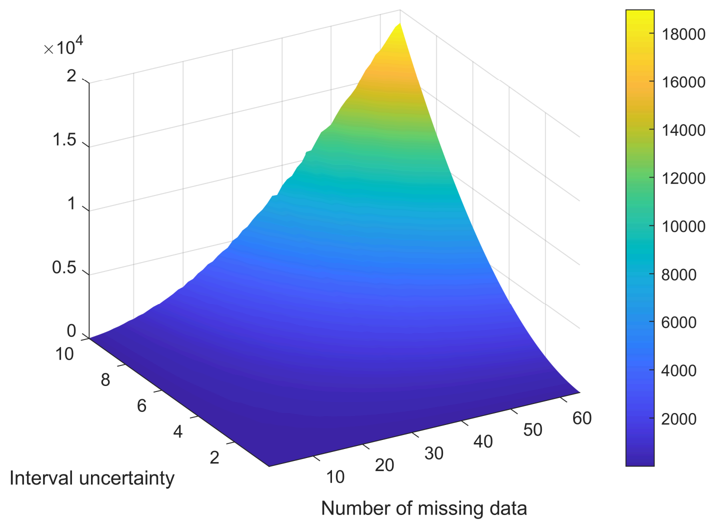

Fig. 22. Combination of number of missing data and interval uncertainty for the JONSWAP PSD function.

In Fig. 23 the results of the area between upper and lower bound for the Kanai-Tajimi PSD function are given. The rough shape of the surface is qualitatively identical to that of the JONSWAP PSD example (Fig. 22), so the same conclusions can be drawn. However, an interesting observation is the significant quantitative difference between the two surfaces. The reason for this might be that the Kanai-Tajimi PSD function has many values distant from zero in its analytical form. Thus, the nonlinear trend starts later, i.e. with a higher number of missing data and/or a higher interval uncertainty. While in the JONSWAP PSD function, many values are close to zero and the lower bound is thus very quickly zero, this non-linear trend starts much earlier. This results in a higher area between the bounds for the JONSWAP PSD function.

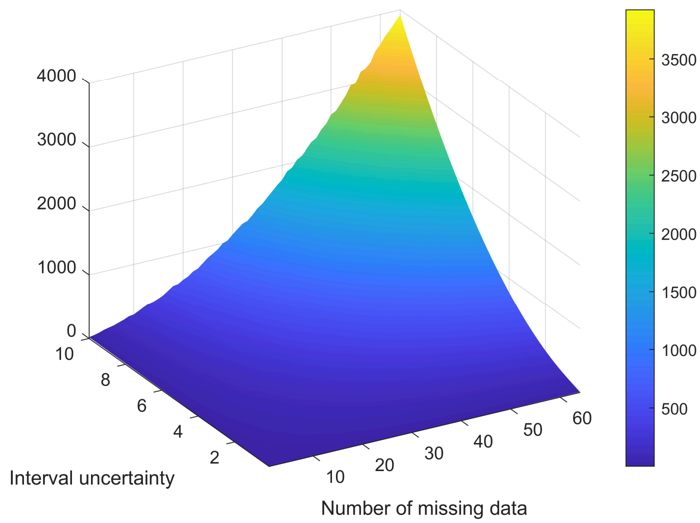

Fig. 23. Combination of number of missing data and interval uncertainty for the Kanai-Tajimi PSD function.

CONCLUSIONS

In this work, the interval DFT algorithm has been investigated for its ability to transform signals with missing data reconstructed by intervals. Different scenarios have been considered, such as the influence of the interval width, the number of missing data, the length of the gap of missing data and the distribution of the missing data in the signal. It was shown, that the largest influence was exerted by the interval uncertainty in the signal and the number of missing data, while the distribution of the missing data and their position is of minor importance. In addition, no indications could be found of an influence whether the data are missing at individual points or appear as a large gap. It was found that too large intervals often lead to extremely wide bounds, which are usually no longer usable for practical purposes. If the number of missing data is sufficiently small, however, a useful transformation can be computed even with a conservative estimation of the intervals, in which the bounds are close to the actual spectrum. With a larger number of missing data or larger gaps, it is also possible to plan for the worst-case by considering only the upper bound as this results in a total higher power of the spectrum, provided that the reconstruction method is not too conservative in determining wide interval widths. It has also been shown that the potential power content of the PSD function can change significantly depending on the choice of interval uncertainty. The results of this work can also be used to assess whether a signal and its reconstruction are considered overly uncertain to be used in practical applications. Further, it can be determined whether a sensor should be replaced to record a signal if its precision is too poor and the corresponding bound PSD yields too wide bounds. In summary, the interval DFT algorithm provides significant and conclusive results for signals with reconstructed data. It should be noted that the results are dependent on the quality of the reconstruction of the data. Thus, it is highly recommended that in the case of missing data, the interval DFT algorithm should be employed with an advanced reconstruction method in order to obtain practical in addition to reliable results.

DATA AVAILABILITY STATEMENT

Some or all data, models, or code generated or used during the study are available in a repository online in accordance with funder data retention policies. The software for computing the interval DFT can be accessed in a single instance via GitHub at: `https://github.com/interval-fourier-transform/application-to-missing-data`.

REFERENCES

Alefeld, G. and Herzberger, J. (2012). *Introduction to Interval Computation*. Computer Science and Applied Mathematics. Elsevier Science.

Beer, M., Ferson, S., and Kreinovich, V. (2013). "Imprecise probabilities in engineering analyses." *Mechanical Systems and Signal Processing*, 37(1), 4 – 29.

Behrendt, M., De Angelis, M., Comerford, L., and Beer, M. (2022a). "Assessing the severity of missing data problems with the interval discrete fourier transform algorithm." *Proceedings of the 32nd European Safety and Reliability Conference (ESREL 2022)*.

Behrendt, M., de Angelis, M., Comerford, L., Zhang, Y., and Beer, M. (2022b). "Projecting interval uncertainty through the discrete Fourier transform: An application to time signals with poor precision." *Mechanical Systems and Signal Processing*, 172, 108920.

Campi, M., Calafiore, G., and Garatti, S. (2009). "Interval predictor models: Identification and reliability." *Automatica*, 45(2), 382–392.

Chopra, A. K. (1995). *Dynamics of structures: Theory and applications to earthquake engineering*. Englewood Cliffs, N.J: Prentice Hall.

Comerford, L., Jensen, H., Mayorga, F., Beer, M., and Kougioumtzoglou, I. (2017). “Compressive sensing with an adaptive wavelet basis for structural system response and reliability analysis under missing data.” *Computers & Structures*, 182, 26–40.

Comerford, L., Kougioumtzoglou, I. A., and Beer, M. (2015a). “An artificial neural network approach for stochastic process power spectrum estimation subject to missing data.” *Structural Safety*, 52, 150–160 Engineering Analyses with Vague and Imprecise Information.

Comerford, L., Kougioumtzoglou, I. A., and Beer, M. (2015b). “On quantifying the uncertainty of stochastic process power spectrum estimates subject to missing data.” *International Journal of Sustainable Materials and Structural Systems*, 2(1-2), 185–206.

Comerford, L., Kougioumtzoglou, I. A., and Beer, M. (2016). “Compressive sensing based stochastic process power spectrum estimation subject to missing data.” *Probabilistic Engineering Mechanics*, 44, 66–76 Special Issue Based on Papers Presented at the 7th International Conference on Computational Stochastic Mechanics (CSM7).

Cooley, J. W. (1987). “The re-discovery of the fast Fourier transform algorithm.” *Microchimica Acta*, 93(1), 33–45.

Cooley, J. W. and Tukey, J. W. (1965). “An Algorithm for the Machine Calculation of Complex Fourier Series.” *Mathematics of Computation*, 19(90), 297–301.

de Angelis, M. (2022). “Exact bounds on the amplitude and phase of the interval discrete Fourier transform in polynomial time, <https://arxiv.org/abs/2205.13978>.

De Angelis, M., Behrendt, M., Comerford, L., Zhang, Y., and Beer, M. (2021). “Forward interval propagation through the discrete Fourier transform.” *The 9th international workshop on Reliable Engineering Computing*, 39–52.

De Rubeis, V., Tosi, P., Gasparini, C., and Solipaca, A. (2005). “Application of Kriging Technique to Seismic Intensity Data.” *Bulletin of the Seismological Society of America*, 95(2), 540–548.

Faes, M. G. and Moens, D. (2020). “Recent trends in the modeling and quantification of nonprobabilistic uncertainty." *Archives of Computational Methods in Engineering*, 27, 633–671.

Hasselmann, K. F., Barnett, T. P., Bouws, E., Carlson, H., Cartwright, D. E., Eake, K., Euring, J., Gicnapp, A., Hasselmann, D., Kruseman, P., et al. (1973). "Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP)." *Ergaenzungsheft zur Deutschen Hydrographischen Zeitschrift, Reihe A*.

Kanai, K. (1957). "Semi-empirical formula for the seismic characteristics of the ground." *Bulletin of the Earthquake Research Institute*, 35, 309–325.

Kiureghian, A. D. and Ditlevsen, O. (2009). "Aleatory or epistemic? does it matter?." *Structural Safety*, 31(2), 105–112 Risk Acceptance and Risk Communication.

Levenberg, K. (1944). "A method for the solution of certain non-linear problems in least squares." *Quarterly of applied mathematics*, 2(2), 164–168.

Li, J. and Chen, J. (2009). *Stochastic Dynamics of Structures*. John Wiley & Sons.

Lin, Q. and Li, C. (2020). "Kriging based sequence interpolation and probability distribution correction for gaussian wind field data reconstruction." *Journal of Wind Engineering and Industrial Aerodynamics*, 205, 104340.

Lin, Y.-K. and Cai, G.-Q. (1995). *Probabilistic structural dynamics: advanced theory and applications*. McGraw-Hill New York.

Liu, G. and Kreinovich, V. (2010). "Fast convolution and Fast Fourier Transform under interval and fuzzy uncertainty." *Journal of Computer and System Sciences*, 76(1), 63–76 Special Issue on Intelligent Data Analysis.

Lutes, L. D. and Sarkani, S. (2004). *Random Vibrations: Analysis of Structural and Mechanical Systems*. Butterworth-Heinemann.

Moore, R., Kearfott, R., and Cloud, M. (2009). *Introduction to Interval Analysis*. Cambridge University Press.

Moore, R. E. (1966). *Interval analysis*, Vol. 4. Prentice-Hall Englewood Cliffs.

Moore, R. E. (1979). *Methods and applications of interval analysis*. SIAM.

Naghizadeh, M. and Sacchi, M. (2010). "Seismic data reconstruction using multidimensional prediction filters." *Geophysical Prospecting*, 58(2), 157–173.

Naghizadeh, M. and Sacchi, M. D. (2007). "Multistep autoregressive reconstruction of seismic records." *GEOPHYSICS*, 72(6), V111–V118.

Newland, D. (2012). *An Introduction to Random Vibrations, Spectral & Wavelet Analysis: Third Edition*. Dover Civil and Mechanical Engineering. Dover Publications.

Nikolaidis, E., Ghiocel, D. M., and Singhal, S. (2004). *Engineering Design Reliability Handbook*. CRC press, 1 edition.

Pierson Jr., W. J. and Moskowitz, L. (1964). "A proposed spectral form for fully developed wind seas based on the similarity theory of S. A. Kitaigorodskii." *Journal of Geophysical Research (1896-1977)*, 69(24), 5181–5190.

Priestley, B. (1982). *Spectral Analysis and Time Series*. Academic Press.

Roberts, J. B. and Spanos, P. D. (2003). *Random Vibration and Statistical Linearization*. Courier Corporation.

Rocchetta, R., Gao, Q., and Petkovic, M. (2021). "Soft-constrained interval predictor models and epistemic reliability intervals: A new tool for uncertainty quantification with limited experimental data." *Mechanical Systems and Signal Processing*, 161, 107973.

Sadeghi, J., de Angelis, M., and Patelli, E. (2019). "Efficient training of interval neural networks for imprecise training data." *Neural Networks*, 118, 338–351.

Schuëller, G. I. (2007). "On the treatment of uncertainties in structural mechanics and analysis." *Computers & Structures*, 85, 235–243.

Shinozuka, M. and Deodatis, G. (1991). "Simulation of stochastic processes by spectral representation." *Applied Mechanics Reviews*, 44(4), 191–204.

Sneddon, I. (1995). *Fourier Transforms*. Dover Books on Mathematics. Dover Publications.

Soong, T. and Grigoriu, M. (1993). *Random Vibration of Mechanical and Structural Systems*. PTR Prentice Hall.

Tajimi, H. (1960). "A statistical method of determining the maximum response of a building structure during an earthquake." *Proceedings of the 2nd world conference of earthquake engineering*, Vol. 11, 781–797.

Zhang, Y., Comerford, L., Kougioumtzoglou, I. A., Patelli, E., and Beer, M. (2017). "Uncertainty quantification of power spectrum and spectral moments estimates subject to missing data." *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 3(4), 04017020.