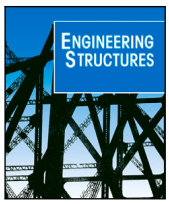
Engineering Structures 305 (2024) 117786

Contents lists available at ScienceDirect

Engineering Structures

journal homepage: www.elsevier.com/locate/engstruct



Heteroscedastic Gaussian processes for data normalisation in probabilistic novelty detection of a wind turbine

Clemens Jonscher *, Sören Möller, Leon Liesecke, Benedikt Hofmeister, Tanja Grießmann, Raimund Rolfes

Institute of Structural Analysis, Leibniz Universität Hannover, Appelstraße 9A, 30167 Hannover, Germany

ARTICLE INFO

Keywords:
BAYOMA
Heteroscedastic Gaussian process
Wind turbine tower
Structural health monitoring
Damage detectability

ABSTRACT

This study investigates the data normalisation of modal parameters of an operating concrete–steel hybrid onshore wind turbine tower considering also the identification uncertainty. In order to take into account the Environmental and Operational Condition (EOC)-dependent variance, sparse heteroscedastic Gaussian processes (GPs) are used for the data normalisation. Following a typical vibration-based *Structural Health Monitoring* (SHM) scheme, data normalisation of the natural frequencies and the mode shapes is performed first. Subsequently, a metric is defined which takes into account both the identification uncertainty and the operation-dependent uncertainty in order to enable novelty detection.

The data normalisation methods must be able to handle uncertainties of different magnitudes due to EOCs in the data. In this context, GPs can be a suitable tool. However, standard GPs assume homoscedasticity, which is an unrealistic assumption in the case of EOC-dependent variance. Using a heteroscedastic GP instead, the variance of the data is better mapped and allows comparison with the identification uncertainties of Bayesian operational modal analysis (BAYOMA), taking into account the specifics of closely spaced modes of the tower structure. This leads to a better interpretation of the data and enables the introduction of a probabilistic novelty metric.

This data normalisation approach, taking into account EOC-dependent uncertainties using heteroscedastic GPs, is being applied for the first time to a tower of a full scale 3.4 MW wind turbine in operation. Following this approach, it is possible to detect smaller changes in natural frequencies and second-order modal assurance criterion (S2MAC) compared to the assumption of homoscedasticity within the GP. In addition, a novelty was detected using the S2MAC during the period under study. Therefore, it can be illustrated that mode shape-based metrics tend to be more sensitive than purely frequency-based ones. However, it is difficult to assess the significance of such changes for structural integrity without further information.

1. Introduction

Wind energy plays a crucial role in the decarbonisation of the energy sector. One cost factor is operation and maintenance. Reliable monitoring systems are needed to reduce these costs. In vibration-based structural health monitoring (SHM), many methods are based on modal parameters estimated from measured vibration data [1]. The modal parameters in terms of natural frequencies, mode shapes and damping are often extracted using output-only or operational modal analysis (OMA) techniques, like *Bayesian Operational Modal Analysis* (BAYOMA) [2], that not only enables the identification of the modal parameters but also of the corresponding uncertainties. This is useful for tower structures because the mode shapes of closely spaced modes are challenging to identify. Here, two different parts of uncertainty need to be considered. The first relates to the uncertainty of the mode subspace (MSS), which is determined by the dominant vibration shapes. The second aspect relates to the alignment of the mode shape within the MSS [3]. As the closeness of the frequencies increases, so does the difficulty of identifying the alignment, as noted in [4]. It has been shown in [5] that the application of the Second Order Modal Assurance Criterion (S2MAC) [6] has the ability to eliminate this alignment uncertainty in the context of monitoring wind turbine towers. However, changing environmental and operational conditions (EOCs) also affect the structure's response [7]. This includes the variance of modal parameters, like natural frequencies and mode shapes, due to slowly changing EOCs, such as the nacelle position of wind turbine towers [5]. Therefore, the detection of structural damage under changing EOCs

* Corresponding author.
E-mail address: c.jonscher@isd.uni-hannover.de (C. Jonscher).

https://doi.org/10.1016/j.engstruct.2024.117786

Available online 1 March 2024

is an important issue in the SHM research community [8]. Typically, measurements are taken to gather information on the response of the structure and monitoring features are extracted. These features are then compared with those from the healthy state of the structure. To this end, Lucà et al. [9] use multiple vibration modes simultaneously to develop a damage metric based on the Mahalanobis squared distance. Similarly, Bull et al. [10] have applied this metric for damage detection in a properly operating Vestas V27 wind turbine. However, this method is limited by the variability of the EOC, which affects the specificity of the features in the training data. This could lead to an undue influence on the estimation of mean and covariance, resulting in so-called outlier masking, which leads to increased false negatives [11]. Thus, if the corresponding EOCs of the reference state differ from those of the current state, false damage detection may occur [12]. The confounding effects of EOC variations must therefore be removed [13,14] before a suitable metric can be derived to assess the condition of the structure. This process is referred to as data normalisation [15], which is the focus of this work.

Clearly, the applicability of data normalisation methods depends on the availability of reliable measurements of the EOCs [16]. When direct measurements of EOCs are available, it is possible to model the monitored parameters with respect to the EOC variations of the system and to use the error of the resulting model as a robust indicator for the structural condition [17–19]. Instead of treating the influences of the EOCs as a regression task, they can also be viewed in terms of classification. Here, the system behaviour is divided into discrete regimes based on the current EOC state. A clustering approach is conceivable in this context, which automatically recognises the different classes and then uses them for classification [20,21]. Projection methods are also an option. These methods attempt to capture the variations in monitoring parameters caused by changing EOCs without the need for direct measurements of the EOCs. In SHM, one of the methods that has been used to transform the measured data into a new EOC insensitive coordinate system is the principal component analysis (PCA) [22]. Cointegration represents another projection method [23]. Here, the fact that some monitored variables have common trends caused by the changing EOCs is taken into account. A disadvantage of these methods is that nonlinear effects can reduce their efficiency [24]. Therefore, nonlinear extensions have been proposed in recent years [25–27]. However, those methods are preferable when direct EOC measurements are impractical or difficult to perform [28]. In addition, these methods cannot provide information on the uncertainty of a prediction, which is particularly important for decision-making on potential damage in SHM. Therefore, several Bayesian approaches have been developed in recent years to quantify the uncertainties in the context of SHM [29]. For example, Bayesian models can be used for data-driven probabilistic damage detection, assuming that in the range of the damage-sensitive frequency band of the natural frequency, the relative proportions of the real and imaginary components in the frequency domain are not changed by EOCs, but only by the damage [30]. Another field of application in SHM is the forecast [31] or reconstruction [32] of measurement data with Bayesian models.

Since the Supervisory Control And Data Acquisition (SCADA) data for the investigated hybrid concrete–steel tower in a 3.4 MW onshore wind turbine are available, a regression-based approach to data normalisation is preferable. In this case, the EOCs are the input variables of the regression task. A simple approach would be to consider only the parameters at the standstill of the wind turbine and to compensate for the remaining variations, such as temperature, with a linear regression as Botz [33] proposed for the monitoring of a concrete steel tower of a 3 MW onshore wind turbine. An extension of this approach is to use static and dynamic multivariate linear regressions for different operating conditions, as presented for a steel tower of a 2 MW onshore wind turbine by Oliveira et al. [34]. Weijtjens et al. [35] additionally use a nonlinear regression model for data normalisation for monitoring a support structure of a 3 MW offshore wind turbine. Other nonlinear regression approaches make use of machine learning methods such as artificial neural networks (ANNs) [36].

Another option is Gaussian process (GP) regression [37], which is a nonparametric regression technique that is increasingly being used in several SHM applications [38–40]. It is well-suited for combination with BAYOMA because GPs include natural Bayesian foundations and inherently provide confidence intervals for model predictions. This makes it possible to develop a reliable damage detection method based on these confidence intervals. However, one well-known limitation of GP regression is the high computing and memory complexity for big data applications. To overcome this, sparse GPs can be applied to reduce the computation time. This is also of particular interest for online learning approaches [41]. Another limitation of the standard GP is the assumption of homoscedasticity, which means that the noise variance is assumed to be independent of the input variables. For systems with input-dependent (heteroscedastic) variances, this assumption may lead to erroneous estimates of the uncertainties due to noise variance. For example, in wind turbine support structures, the damping in the fore–aft (FA) direction is dependent on the rotor speed and wind speed due to aerodynamic damping. Higher damping leads to a higher identification uncertainty and thus higher variance. Thus, there is a functional dependence between the operational condition rotor speed and the identification uncertainties of the FA-modes. In this context the homoscedastic GP assumes that this uncertainty is constant for all inputs and therefore cannot account for noise changes due to, for example, aerodynamic damping. A more suitable regression approach is therefore the use of a GP with heteroscedastic noise [42–44]. These GPs can map epistemic (e.g.interpolation uncertainty due to missing data) as well as aleatoric (e.g. measurement uncertainty) uncertainties and are therefore able to map uncertainties of different amplitudes caused by EOCs and unrecorded influences.

Heteroscedastic GPs had already been used for some SHM applications, such as mapping temperature variations of a manifold space of natural frequencies of the Z24 bridge [45]. Additionally, Rogers et al. [46] used sparse heteroscedastic GPs to approximate the power curve of an offshore wind turbine, Wang et al. [47] forecasts strain measurements of a large-scale suspension bridge using variational heteroscedastic GP. Moreover, heteroscedastic GPs are applied in the field of guided waves [48] and acoustic emission [49]. For the appropriate treatment of non-stationary problems, Tolvanen et al. [50] presented an extended heteroscedastic GP.

The focus of this work is on the appropriate handling of the influences of natural frequencies and the S2MAC of the mode shapes identified with BAYOMA of an onshore wind turbine tower due to EOCs considering the uncertainties. For the data normalisation, sparse heteroscedastic GP will be used. Based on the normalised data, a probabilistic approach for novelty detection is then applied, using the uncertainty quantification provided by BAYOMA and the GP. The structure of this study is as follows: Section 2 introduces the theory of heteroscedastic sparse GP and in Section 3 the novelty metric is defined. Section 4 describes the wind turbine tower under investigation. In Section 5, the application of the heteroscedastic GP for predicting the identified natural frequencies and S2MACs using aggregated 10-min SCADA data as the input is examined. This is followed by an examination of the novelty detectability of theoretical system changes with the trained heteroscedastic GPs compared to the homoscedastic GPs as well as a more detailed examination of a real potential system change. Finally, in Section 6, the study is summarised and an outlook is given.

2. Gaussian process

In this paper, we only present a brief overview of Gaussian processes (GPs) and direct interested readers to Rasmussen et al. [37] for

more detailed information. GPs are used to model known input–output mappings of the form

$$y = f(\mathbf{x}) + \mathcal{N}(0, \sigma_n^2), \tag{1}$$

where y represents the observed target values and $f(\mathbf{x})$ is a function that is superimposed with additive Gaussian-distributed noise $\mathcal{N}(0, \sigma_n^2)$, assuming zero mean and variance σ_n^2. GPs are a powerful Bayesian machine learning tool with several desirable properties: They are able to make predictions and estimate probability distributions without specifying a particular parametric functional form, require few a priori inputs, and can model relationships between variables with high levels of noise. As shown in Eq. (1), GPs can be interpreted as prior probability distributions over functions $f(\mathbf{x})$. Formally, a GP can thus be defined as

$$f(\mathbf{x}) \sim \mathrm{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')), \tag{2}$$

where $\mathbf{x}$ and $\mathbf{x}'$ are a pair of inputs to the function of interest. According to Eq. (2), a GP is completely defined in terms of its mean $m(\mathbf{x})$ and its covariance function $k(\mathbf{x}, \mathbf{x}')$. In particular, the covariance function defines the smoothing characteristics of the underlying function. To make predictions, a joint Gaussian distribution between training and testing data is formed as follows: Given a set of training data in which the inputs are given in the matrix of multivariate training inputs $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ and target values $\mathbf{y}$, the joint Gaussian distribution can be formed for a testing point $\mathbf{x}_*$ and unknown output y_* using the training data. Assuming a zero-mean function, the joint distribution can be expressed as follows:

$$\begin{pmatrix} \mathbf{y} \\ y_* \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbb{I} & K(\mathbf{X}, \mathbf{x}_*) \\ K(\mathbf{x}_*, \mathbf{X}) & K(\mathbf{x}_*, \mathbf{x}_*) + \sigma_n^2 \end{bmatrix}\right), \tag{3}$$

where $K(\cdot)$ denotes the matrix of covariances defined by the covariance function, or kernel, $k(\cdot)$ between every data point. Given the joint distribution $p(\mathbf{y}, y_*)$, the conditional distribution $p(\mathbf{y}|y_*)$ can be recovered. This allows for the derivation of the expression for the posterior predictive mean and variance as follows:

$$y_* \sim \mathcal{N}(\mathbb{E}[y_*], \mathbb{V}[y_*]), \tag{4}$$

where the posterior predictive mean $\mathbb{E}[y_*]$ and variance $\mathbb{V}[y_*]$ are given by

$$\mathbb{E}[y_*] = K(\mathbf{x}_*, \mathbf{X})[K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbb{I}]^{-1} \mathbf{y}, \tag{5}$$

$$\mathbb{V}[y_*] = K(\mathbf{x}_*, \mathbf{x}_*) - K(\mathbf{x}_*, \mathbf{X})[K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbb{I}]^{-1} K(\mathbf{X}, \mathbf{x}^*) + \sigma_n^2, \tag{6}$$

respectively, expressed for the observations. To determine the values of the kernel hyperparameters $\boldsymbol{\Theta}$, the marginal likelihood of the predictions $p(\mathbf{y}|\mathbf{X}, \boldsymbol{\Theta})$ is maximised with respect to $\boldsymbol{\Theta}$. In log form, this is expressed as

$$\log p(\mathbf{y}|\mathbf{X}, \boldsymbol{\Theta}) = -\frac{1}{2}\mathbf{y}^{\mathrm{T}}(K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbb{I})^{-1}\mathbf{y} - \frac{1}{2}\log|K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbb{I}| - \frac{N}{2}\log 2\pi, \tag{7}$$

where N is the number of training samples. In this context, the marginal likelihood is used to quantify the joint probability of the data under a prior.

2.1. Sparse approximation of Gaussian process

In the previous section, it was stated that GPs provide powerful probabilistic models for regression. However, to adjust the hyperparameters and to use the GP for predictions, it is necessary to calculate the inverse of the covariance matrix including noise, denoted as $[K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbb{I}]^{-1}$. The computational complexity of this operation scales with $\mathcal{O}(N^3)$ and the memory scales with $\mathcal{O}(N^2)$, where N is the number of training samples. Therefore, GPs become impractical to train for large data sets with more than about ten thousand data points with current computer hardware [37]. To overcome this limitation, several approximation methods have been developed, which are referred to as *sparse approximations of Gaussian processes* [51]. These methods focus on a small set of samples and preserve the advantageous properties of GPs by simultaneously reducing the complexity to typically $\mathcal{O}(NM^2)$ for computational time and $\mathcal{O}(NM)$ for memory, where M is the number of inducing variables selected for the approximation, with $M < N$ [52]. We refer to [51] for a comprehensive review of fundamental sparse approximation methods. For this work, we adopt an approach based on Almosallam [53], who proposes a basis function model (BFM) approach to improve posterior variance accuracy. For describing sparse GPs using BFM, the underlying assumption is, given inputs $\mathbf{x}$ and target values $\mathbf{y}$, that the observed target y_i is generated by a linear combination of m non-linear functions $\boldsymbol{\phi}(\mathbf{x}_i) = [\phi_1(\mathbf{x}_i), \dots, \phi_m(\mathbf{x}_i)]^T \in \mathbb{R}^m$ of the inputs plus additive noise $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$:

$$y_i = \boldsymbol{\phi}(\mathbf{x}_i)^{\mathrm{T}}\mathbf{w} + \epsilon_i. \tag{8}$$

Here, $\mathbf{w}$ is a vector of length m representing the weights of the basis functions, which are represented by radial basis functions (RBFs) for this study. Compared to a standard GP, where the mean of the predictive distribution is a linear combination of n kernel functions, the BFM approach assumes that the form of the function is a linear combination of $m \ll n$ basis functions and integrates out its parameters [44]. To favour the simplest explanation that fits the data, a prior on the weights $\mathbf{w}$ can be formulated probabilistically by taking $p(\boldsymbol{w}|\alpha) = \mathcal{N}(\boldsymbol{w}|0, \mathbf{A}^{-1})$, where $\mathbf{A} = \alpha\mathbb{I}$ (with $\mathbb{I}$ as the identity matrix) is a prior precision on the weights $\mathbf{w}$, which is included in the parameter set $\boldsymbol{\Theta}$. It can be shown that the posterior probability of the weight vector $\mathbf{w}$ has the following distribution:

$$p(\mathbf{w}|\mathbf{y}, \mathbf{X}, \boldsymbol{\Theta}) = \mathcal{N}(\mathbf{w}|\overline{\mathbf{w}}, \boldsymbol{\Sigma}_w^{-1}), \tag{9}$$

where the mean and covariance matrix of the distribution are given by:

$$\overline{\mathbf{w}} = \beta \boldsymbol{\Sigma}_{\boldsymbol{w}}^{-1} \boldsymbol{\Phi}_x \mathbf{y}, \tag{10}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{w}} = \beta \boldsymbol{\Phi}_x \boldsymbol{\Phi}_x^{\mathrm{T}} + \mathbf{A}, \tag{11}$$

respectively, where $\beta = \sigma^{-2}$ is the precision parameter. By substituting $\boldsymbol{w} = \overline{\boldsymbol{w}}$ it is possible to express the marginal likelihood as:

$$p(\mathbf{y}|\mathbf{X}, \boldsymbol{\Theta}) = \mathcal{N}(\mathbf{y}|\boldsymbol{\phi}_{\mathbf{x}}{}^{\mathrm{T}}\overline{\mathbf{w}}, \beta^{-1}\mathbb{I})\mathcal{N}(\overline{\mathbf{w}}|0, \mathbf{A}^{-1})(2\pi)^{\frac{m}{2}}|\boldsymbol{\Sigma}_{\mathbf{w}}|^{-\frac{1}{2}}. \tag{12}$$

The log marginal likelihood can then efficiently be expressed in terms of the mean $\overline{\mathbf{w}}$ and the covariance $\boldsymbol{\Sigma}_w$ of the posterior distribution:

$$\begin{aligned} \log p(\mathbf{y}|\mathbf{X}, \boldsymbol{\Theta}) = &-\frac{\beta}{2}\|\boldsymbol{\Phi}_x\overline{\mathbf{w}} - \mathbf{y}\|^2 + \frac{n}{2}\log\beta - \frac{n}{2}\log 2\pi \\ &- \frac{\alpha}{2}\overline{\mathbf{w}}^T\overline{\mathbf{w}} + \frac{m}{2}\log\alpha - \frac{1}{2}\log|\boldsymbol{\Sigma}_w|, \end{aligned} \tag{13}$$

where $\boldsymbol{\Phi}_x$ is an $n \times m$ matrix whose (i, j) entry is given by $\phi_j(\mathbf{x}_i)$. Optimisation of the hyperparameters of the basis function, the precision β, and the weight precision α is achieved by maximising the log marginal likelihood, as defined in Eq. (13). The predictive distribution for a test target output y_* is given as follows:

$$p(y_*|\mathbf{x}_*, \mathbf{y}, \mathbf{X}, \boldsymbol{\Theta}) = \mathcal{N}(y_*|\mathbb{E}[y_*], \mathbb{V}[y_*]), \tag{14}$$

$$\mathbb{E}[y_*] = \boldsymbol{\Phi}(\mathbf{x}_*)^{\mathrm{T}}\overline{\mathbf{w}}, \tag{15}$$

$$\mathbb{V}[y_*] = \boldsymbol{\Phi}(\mathbf{x}_*)^{\mathrm{T}}\boldsymbol{\Sigma}_w^{-1}\boldsymbol{\Phi}(\mathbf{x}_*) + \beta^{-1}\mathbb{I}. \tag{16}$$

It is worth noting that BFMs are mathematically equivalent to the subset-of-regressors sparse GP method, as shown in [53].

2.2. Heteroscedastic Gaussian process

The prediction variance, as shown in Eq. (16), can be decomposed into two components: the model and the inherent noise variance. This decomposition reflects the typical classification of uncertainty sources into aleatoric and epistemic uncertainties [54]. The aleatoric uncertainty, which is irreducible, stems from inherent randomness, while the epistemic uncertainty arises from insufficient knowledge and can be

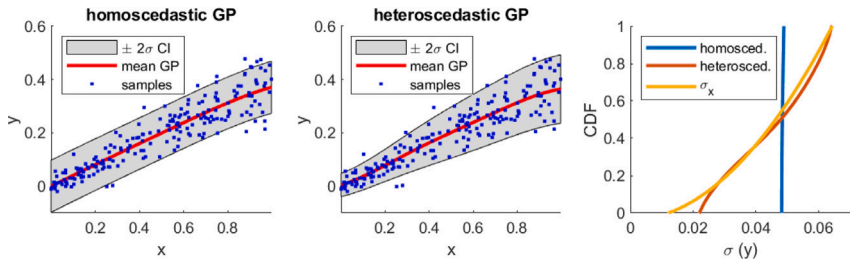


Fig. 1. Comparison of homoscedastic and heteroscedastic noise models of sparse GPs for input-dependent noise with two basis functions. The underlying data (200 samples) is obtained by sampling from $y \sim \mathcal{N}(0.4x, \sigma_x^2)$ with $\sigma_x^2 = 0.00015 + 0.004x$. For training, the samples are divided into 60% training, 20% validation and 20% testing.

reduced, if more knowledge is available. Therefore, model variance is an instance of epistemic uncertainty, and noise uncertainty is mainly a form of aleatoric uncertainty. The level of epistemic uncertainty of a test point $\mathbf{x}_*$ is caused by the density of the samples used during the training of the GP. This type of uncertainty can thus be reduced by incorporating more data in the training set. For the aleatoric uncertainty, homoscedasticity with constant precision β was assumed in the previous section. However, this assumption is inadequate for many applications wherein a more flexible heteroscedasticity model is required, in which noise variance is a function of the inputs. To account for this, one approach is to model the noise as a second GP [43]. Another approach is to use BFMs, which can be considered as a sparse GP method that defines a semiparametric form via a set of weights. To incorporate heteroscedasticity, Almosallam proposes a BFM approach [53], where a function for the variable and input-dependent noise is defined by a linear combination of basis functions of the form $\beta(\mathbf{x}) = \exp(\boldsymbol{\phi}(\mathbf{x})^{\mathrm{T}}\mathbf{v} + b)$. The exponential form is chosen to ensure the positivity of $\beta(\mathbf{x})$. Based on this, Almosallam expressed the posterior distribution as follows [53]:

$$p(\mathbf{w}|\mathbf{y}, \mathbf{X}, \boldsymbol{\Theta}) = \mathcal{N}(\mathbf{w}|\overline{\mathbf{w}}, \boldsymbol{\Sigma}_w^{-1}), \text{ where} \tag{17}$$

$$\overline{\mathbf{w}} = \boldsymbol{\Sigma}_{\boldsymbol{w}}^{-1}\boldsymbol{\Phi}_x\mathbf{B}\mathbf{y}, \tag{18}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{w}} = \boldsymbol{\Phi}_x\mathbf{B}\boldsymbol{\Phi}_x^{\mathrm{T}} + \mathbf{A}, \tag{19}$$

and $\mathbf{B} = \mathrm{diag}[\{\beta(\mathbf{x}_i)\}_{i=1}^n]$. For a testing point $\mathbf{x}_*$ and an unseen output y_*, the predictive distribution is Gaussian distributed as follows:

$$p(y_*|\mathbf{x}_*, \mathbf{y}, \mathbf{X}, \boldsymbol{\Theta}) = \mathcal{N}(y_*|\mathbb{E}[y_*], \mathbb{V}[y_*]), \text{ where} \tag{20}$$

$$\mathbb{E}[y_*] = \boldsymbol{\Phi}(\mathbf{x}_*)^{\mathrm{T}}\overline{\mathbf{w}}, \text{ and} \tag{21}$$

$$\mathbb{V}[y_*] = \boldsymbol{\Phi}(\mathbf{x}_*)^{\mathrm{T}}\Sigma_w^{-1}\boldsymbol{\Phi}(\mathbf{x}_*) + \beta(\mathbf{x}_*)^{-1}, \tag{22}$$

where $\beta(\mathbf{x}_*)^{-1}$ is the input-dependent noise uncertainty. A simple simulation model for illustration of the GP defined above is presented in Fig. 1. When homoscedasticity is assumed for the sparse GP, it cannot properly capture the varying level of noise variance across the input space, whereas a sparse heteroscedastic GP can. The cumulative distribution function (CDF) of the uncertainty of the GP illustrates this relationship. In the case of a purely aleatoric uncertainty, the uncertainty of the homoscedastic GP becomes a vertical line. Therefore, there are ranges where the homoscedastic GP overestimates the variance and ranges where it underestimates the variance. Comparing the underlying variance σ_x^2 of the sampled data and the heteroscedastic GP, it is noticeable that the GP overestimates the lowest variance. The more samples are included in the training, the better the variance of the GP fits the model. By better reflecting the variance, the probabilistic novelty metric described in the following section can make more reliable statements.

3. Novelty metric

There are several ways to detect novelty in data when describing data variation with a regression model, such as T^2 control charts [34]. Another probability-based distance metric uses the probability density function f of the identified monitoring parameters (MP) to calculate the probability of the MP lying in the confidence interval of the GP [55]

$$p_{\text{healthy}}(\text{CI}_{\text{lb}} \leq \text{MP}_i \leq \text{CI}_{\text{ub}}) = p_{MP,i} = \int_{\text{CI}_{\text{lb}}}^{\text{CI}_{\text{ub}}} f(\text{MP}_i)d\text{MP}_i, \tag{23}$$

where CI_{lb} and CI_{ub} are the lower and upper bound of the confidence interval of the GP. Assuming that the quantities used are statistically independent, one can multiply the probabilities of N different parameters $p_{\text{MP,i}}, i = 1, \ldots, N$, to obtain the joint probability

$$p_{\text{MP}} = \prod_{i=1}^{N} p_{\text{MP,i}}. \tag{24}$$

Although the independence of $p_{\text{MP,i}}, i = 1, \ldots, N$ is likely to be violated, the metric should perform well for real-world applications, as it known for example from the Naive Bayes classifier [56]. In order to reduce the influence of misidentifications and thus enable reliable statements, observations are often averaged [34]. In this study, a moving average over the $p_{\text{MP,i}}, i = 1, \ldots, N$ is used, whereby the window length must be specified. Furthermore, a threshold for detecting novelty must also be defined. Our approach to set the threshold is achieved by choosing the lower bound confidence interval of the averaged p_{MP} of the training period of the GP. In summary, the confidence interval of the GP, the window length of the moving average, and the lower bound of the confidence interval for determining the threshold must be defined. These parameters have a direct impact on the detectability of system changes. A too sensitive setting can lead to many false positives due to outliers. A too conservative setting leads to a significantly degraded damage sensitivity. The settings therefore have to be adjusted depending on the application.

4. Monitored tower of a wind turbine

The investigations are carried out on a steel hybrid tower of an onshore wind turbine, which is shown in Fig. 2. The accelerations of both horizontal measurement directions are recorded synchronously on five measurement planes. The modal parameters used in the following were identified with their identification uncertainties using BAYOMA [57]. In principle, all operational modal analysis methods that also determine the identification uncertainties, such as an extension of the stochastic subspace identification (SSI) [58], can be used for this approach. In the context of closely spaced modes identified using the SSI, the transformation of the complex mode shapes into the real space is not possible without errors due to the splitting of the phases in the complex plane according to spatial directions [4]. This can lead to an error in the calculation of the α_{S2MAC}. The applicability of BAYOMA, as well as the identification uncertainties, were investigated for the monitored tower in a previous study [5]. According to the study, the natural frequencies f_0 and the α_{S2MAC}, which is the angular representation of the second-order modal assurance criterion (S2MAC) [6], can be used as monitoring parameters of the first ($B1$) and fourth ($B4$) bending

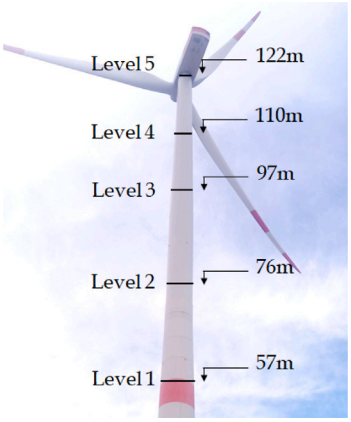


Fig. 2. Monitored steel–concrete hybrid tower of a wind turbine with the measurement setup.

mode pairs. The S2MAC compares an identified mode shape $\boldsymbol{\varphi}_i$ with a mode subspace spanned by the vectors $\boldsymbol{\psi}_j$ and $\boldsymbol{\psi}_k$. In this way, the dominant alignment uncertainty can be eliminated for closely spaced modes. For real mode shapes, the α_{S2MAC} is calculated as follows

$$\alpha_{\text{S2MAC,i,jk}} = \arccos\left(\sqrt{\frac{(\boldsymbol{\varphi}_i^T\boldsymbol{\psi}_j)^2 - 2(\boldsymbol{\varphi}_i^T\boldsymbol{\psi}_j)(\boldsymbol{\psi}_j^T\boldsymbol{\psi}_k)(\boldsymbol{\varphi}_i^T\boldsymbol{\psi}_k) + (\boldsymbol{\varphi}_i^T\boldsymbol{\psi}_k)^2}{1-(\boldsymbol{\psi}_j^T\boldsymbol{\psi}_k)^2}}\right). \tag{25}$$

The calculation of the α_{S2MAC} uncertainties for each data set is done using a Monte Carlo simulation with 3000 samples and the covariance matrix of the identified mode shape [5].

The study evaluates data from mid-October 2021 to January 2023. Only data indicating constant operation of the wind turbine, as described in more detail in [5], are considered. The time histories of the most probable values of the natural frequencies and the α_{S2MAC} are shown in Fig. 3. The modes are divided into fore–aft (FA) and side–side (SS) with respect to the main direction of vibration towards the nacelle.

5. Data normalisation with heteroscedastic Gaussian processes

The data normalisation of the natural frequencies and the α_{S2MAC} is done using the implementation of the sparse GP of Almosallam et al. [53]. The GPs use basis functions models (BFM), as described in more detail in Section 2. The widely used Radial Basis Functions (RBF) kernel is applied as basis function [37]. The choice is made due to the compatibility with large training data sets, as well as the existence homoscedastic and heteroscedastic implementation in MATLAB [59]. For each monitoring parameter, a heteroscedastic and a homoscedastic version of the GP is trained. The dimensionality should not be too high when selecting the input variables for the GP. Otherwise this can lead to large volumes of recorded data that may be imbalanced in their feature space, resulting in an inadequate representation of the data distribution leading to insufficient information to construct a reliable model. This outcome is known as the *curse of dimensionality*. In this study, the 10-min SCADA data sets used as inputs for the GP are listed in Table 1 with the statistical quantities of the training period. In addition to the parameters, the wind direction, rotor speed and pitch angle are also available. The wind direction has some data gaps and has little more information than the nacelle position during operation. The pitch angle can be determined indirectly from the power and the wind speed. Rotor speed is often used instead of the power for data normalisation, like [34]. However, for the wind turbine under study, the rotor speed data is not as reliable as the power data. SCADA data that are not physically explainable, such as 0 °C outside temperature in summer, or full power at zero wind speed, are further excluded. From a previous study [5] it was found that the natural frequencies during operation are dependent on the wind speed and power, and the α_{S2MAC} more on the nacelle position. The temperature has less impact on both. For learning, the data set from October 2021 to the end of September 2022 was randomly divided into three disjoint subsets to avoid overfitting: Training (60%), validation (20%) and testing (20%). The validation set is used to ensure that the GP is always evaluated on unknown data during training. The total training data contains data over one year to cover the usual EOC space. This long learning time is intended to compensate for the imbalance in the data caused by frequently occurring wind speeds and directions together with rare wind conditions. The same accounts for temperature imbalances within the temperature coverage. Based on the training data, the GPs learn the relationships between EOCs and MPs, such as the well-known relationship between natural frequencies and wind speed. Mode shapes are also influences by EOCs, meaning that shape-based metrics, like the α_{S2MAC} are also EOC-dependent. These are often influenced by the temperature and, in the case of wind turbines, even more by the nacelle position. These correlations are also learned by the GPs in the applied regression-based data normalisation approach.

For the underlying GP in this study, the decisive hyperparameter is the number of basis functions. For an initial guess of this hyperparameter, the Akaike Information Criterion (AIC) [60] based on the mean log likelihood probability is used. In order to obtain the most accurate prediction and sparse representation of the model, an Automatic Relevance Determination (ARD) is finally used, following the implementation of Almosallam [53]. Table 2 lists the resulting number of basis functions as well as the coefficients of determination (R^2) of the training, validation and test period. Here, if the predictions exactly match the considered measurements, R^2 is equal to one. A GP that always predicts the mean of the measurement data has $R^2 = 0$, and GP models that are worse than this have a negative R^2 score. Therefore, these values are used to evaluate the GPs and should always be greater than 0. For the validation of the proposed data normalisation scheme, two further periods from October 10th, 2022 to December 10th, 2022 (phase 1) and from December 19th, 2022 to January 18th, 2023 (phase 2) are considered, which are outside the original 12 month period. The R^2-values of the individual parameters, with the exception of the α_{S2MAC} of the first bending mode pair, change only slightly for all time

Table 1
Inputs of the GP with the statistical quantities from October 2021 to September 2022.

Input	Median	5% quantile	95% quantile
Wind speed in $\frac{\text{m}}{\text{s}}$	6.12	3.6	10.46
Power in kW	555	48	2725
Outside temperature nacelle in °C	9.68	2.17	23.52
Nacelle position in degrees	225	47.4	328.4

Table 2
Number of basis functions (#BF), as well as the coefficient of determination (R^2) of the heteroscedastic GPs. The R^2 are used to evaluate the models.

	$B1_{FA}$		$B1_{SS}$		$B4_{FA}$		$B4_{SS}$	
	f	α_{S2MAC}	f	α_{S2MAC}	f	α_{S2MAC}	f	α_{S2MAC}
#BF	9	24	13	23	5	46	23	18
R^2_{training}	0.55	0.544	0.34	0.602	0.561	0.691	0.633	0.795
$R^2_{\text{validation}}$	0.541	0.518	0.348	0.606	0.570	0.691	0.633	0.791
R^2_{testing}	0.531	0.51	0.354	0.594	0.561	0.68	0.614	0.795
$R^2_{\text{phase 1}}$	0.484	0.263	0.407	0.313	0.422	0.746	0.763	0.765
$R^2_{\text{phase 2}}$	0.645	−0.193	0.281	−1.878	0.615	0.691	0.588	0.722

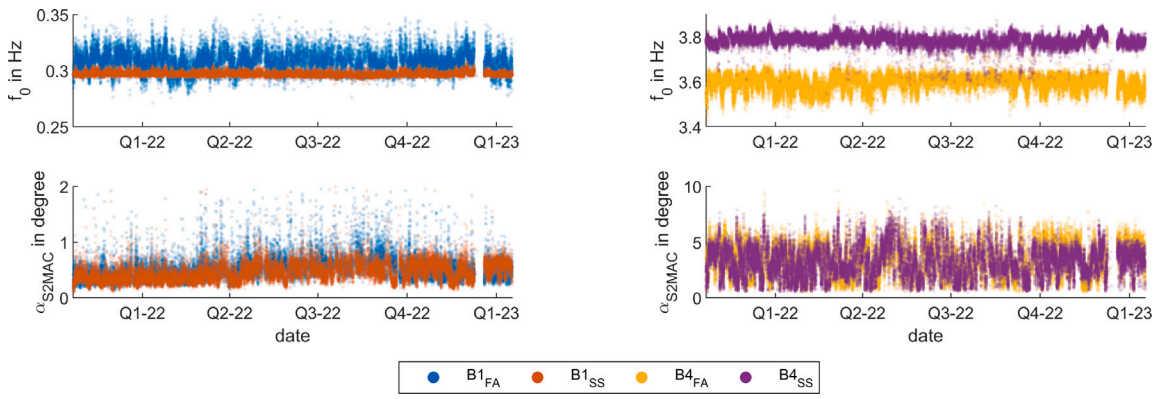


Fig. 3. Time course of the monitored natural frequencies for the bending mode pairs B1 and B4 and the corresponding α_{S2MAC}.

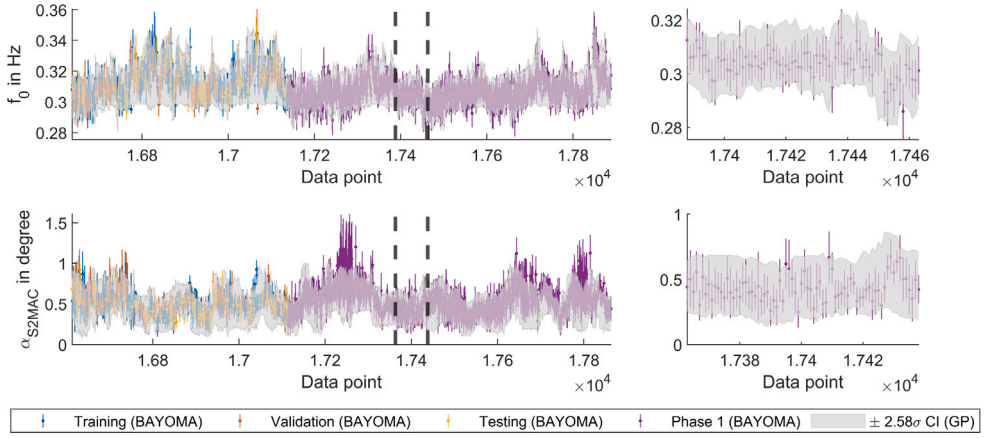


Fig. 4. Data section of the 99% CI identification results of BAYOMA for the natural frequency and the α_{S2MAC} for the first FA bending mode for 500 data points in the training period and 750 in the phase 1. The data sets of the training period are randomly divided into training, validation and testing data sets for training the GP. In addition, the predicted 99% CI of the heteroscedastic GP is shown. A zoom area in phase 1 is marked with black dashed lines and is shown on the right.

periods. This scatter occurs due to the varying EOC distributions of the different time periods. For the α_{S2MAC} of the first bending mode pair, the R^2 values of phase 1 are significantly lower and in phase 2 even less than 0. This mismatch is discussed in more detail in Section 5.2. Fig. 4 shows a part of the 99% confidence interval of the GP prediction of the natural frequency as well as the α_{S2MAC} for the first bending mode in the FA direction. The 99% confidence interval of the identification of BAYOMA is also shown. In the case of α_{S2MAC}, it can be seen around data set $1.72 \cdot 10^4$ that the identified values in phase 1 no longer correspond to the GP regression, which was already indicated by the R^2 score. This effect is explained in Section 5.2.

In the following it is investigated whether the use of a sparse heteroscedastic GP represents the data better than the sparse homoscedastic variant. Theoretically, for good training data density, the means of the heteroscedastic and homoscedastic GPs should be identical. Therefore the differences in the standard deviations of the two GPs are studied in more detail.

The homoscedastic GP has, as expected, a nearly constant standard deviation throughout the period. This suggests that there is a good data coverage of the EOCs throughout this time period. In contrast, the heteroscedastic GP has varying uncertainties in the observation period. The reason is the system variability due to EOCs of the wind turbine tower, which leads to higher identification uncertainties, as shown in [5]. When comparing the levels of the uncertainty, the identification uncertainties of BAYOMA are lower than the uncertainties determined from the heteroscedastic GP. In the ideal, probably unattainable case, the GP can explain the entire deterministic variation of the MP, so that the CDF of the standard deviation identification uncertainty corresponds to that of the heteroscedastic GP, as shown in Fig. 1. The deviation is plausible, as there are further uncertainties in addition to the identification uncertainties, such as the aggregated SCADA data not being fully synchronised and possibly not all influencing EOCs are included as input in the GP. Moreover, the assumptions of BAYOMA are violated, so that the actual uncertainties differ. The deviation can be regarded as epistemic uncertainty of the models. A similar trend in the uncertainties indicates that the GPs reflect the input-dependent uncertainties well. The largest deviations between the identification uncertainty and the GP uncertainty are observed in the cumulative distribution function of the α_{S2MAC} of the fourth bending mode pair. As already shown in [5], the α_{S2MAC} depends on the nacelle position, thus the high uncertainty is likely due to the fuzziness of the nacelle position. As known from other studies like [5], the identification of SS natural frequencies is more reliable than the FA natural frequencies and higher natural frequencies are less uncertain than lower ones. These trends are also represented in the uncertainties of the heteroscedastic GPs in Fig. 5. For the detection of state changes, it is important that the data normalisation represents the healthy state as accurately as possible. System states that are unreliable according to the uncertainty of the GP therefore give no added value, and lead to less reliable results. To increase accuracy, one possibility is to consider only certain operating states on the basis of prior knowledge. Another option is a purely data-based approach, which considers solely data sets where the standard deviation of the GP is below a certain value. This limit can be defined by a determined quantile using the training period data. The numerical value of this quantile must be chosen carefully, as it can greatly reduce the amount of data considered. For this application, a reasonable limit seems to be 99%.

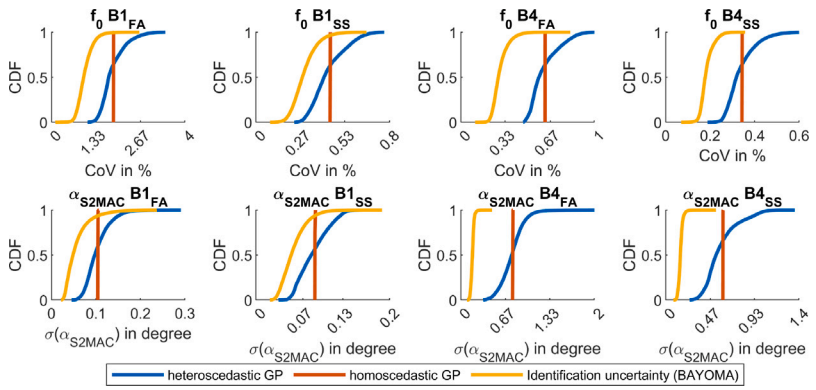


Fig. 5. Cumulative distribution function (CDF) of the standard deviation of the considered natural frequencies for bending mode pairs B1 and B4 and the corresponding α_{S2MAC} of the heteroscedastic and homoscedastic GP as well as of the identification results of BAYOMA during the training period from October 2021 to October 2022.

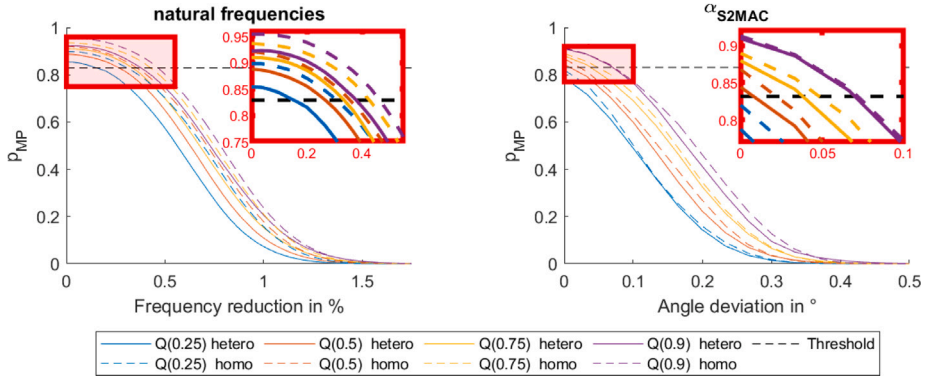


Fig. 6. Investigation of the detectability of the novelty using the natural frequencies and the α_{S2MAC} of the first and fourth bending mode pairs.

Table 3
Settings from the applied novelty metric.

Parameter	Setting
CI of the GP	99%
Window length moving average	144 (one day full data)
Lower bound of the CI of p_{MP}	0.5%

In conclusion, the heteroscedastic GP represents the uncertainties much more realistically and should be preferred to the homoscedastic formulation in the context of monitoring wind turbine supporting structures. In the following section, we will examine which changes in the investigated parameters can be detected.

5.1. Novelty sensitivity

This section examines whether the use of a heteroscedastic GP in the data normalisation of natural frequencies and α_{S2MAC} enables the detection of smaller system changes than a homoscedastic GP. For this purpose, the novelty metric p_{MP} from Section 3 with the settings listed in Table 3 is used. The natural frequencies and the α_{S2MAC} are considered separately. The investigation will be based on the phase 1 period from the beginning of October 2022 to the middle of December 2022. As suggested in the previous chapter, only data sets that are within the lower 99% quantile of the standard deviation of the GPs of the training period are included in the metric. The aim of the investigation is a sensitivity comparison between the homoscedastic and heteroscedastic GP to validate whether the consideration of the EOC-dependent variance enables the detection of smaller system changes.

As with Oliviera et al. [34], it is assumed that damage results in a reduction of the natural frequency. How the natural frequencies change in reality depends on the structure and the type and severity of the damage. For simplicity, in this study, it is assumed that the natural frequencies investigated decrease by the same percentage. In the case of the α_{S2MAC}, it is assumed that all angles increase by the same amount. It should be noted at this point that this is a simplified damage scenario. The results of the study are shown in Fig. 6. The Q(0.25) – where Q(0.25) is the 25%-quantile – Q(0.5), Q(0.75) and Q(0.9) of the averaged p_{MP} are shown for the homoscedastic and heteroscedastic GPs. Since damage cannot necessarily be detected equally well in every operating state due to higher variance of the GPs as well as the identification uncertainty, an average observation is performed using quantiles. The choice of these are based on representability and help to estimate how sensitive the novelty metric is. A novelty is detected when p_{MP} is less than the threshold. This means that if, for example, Q(0.25) is below the threshold, the lowest 25% of p_{MP} are below the threshold and these values detect the novelty. The data normalisation using a heteroscedastic GP can detect smaller system changes for all considered quantiles than the homoscedastic GP, as shown in Fig. 6. This is due to the fact that the uncertainties are better represented with a heteroscedastic noise model, as shown in Fig. 5. For example, when using the Q(0.9) a 0.45% change in natural frequencies can be detected when normalised with the homoscedastic GP and a 0.37% change in natural frequencies when normalised with the heteroscedastic GP.

Examining the novelty sensitivity of the α_{S2MAC} in Fig. 6, confirms the trend that the heteroscedastic GP can detect smaller changes than the homoscedastic one. However, the Q(0.25) of the α_{S2MAC} for both GP methods is already below the threshold in the initial state. In addition, the Q(0.9) of the two data normalisation methods are nearly identical in the threshold range. For further consideration, Fig. 7 shows the α_{S2MAC} separately for the modes. If only the second bending mode is considered, the effect already known from the natural frequencies becomes apparent, that in the initial state the quantiles are above the threshold and the data normalisation with the heteroscedastic GP can

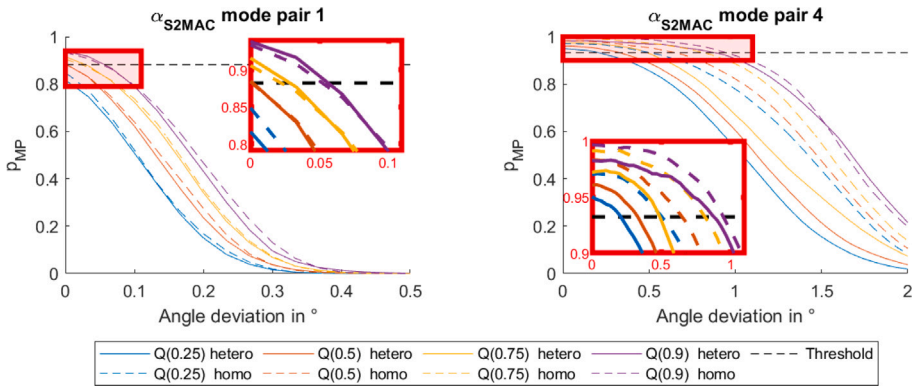


Fig. 7. Investigation of the detectability of novelty using the α_{S2MAC} of the first bending mode pair on the left and the fourth bending mode pair on the right.

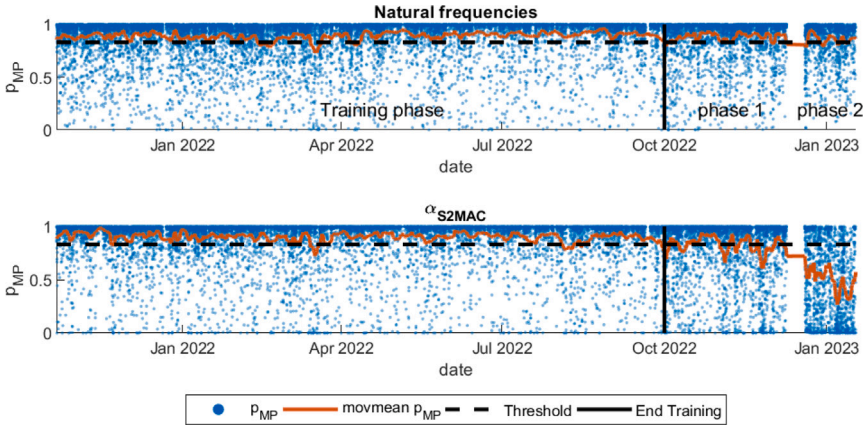


Fig. 8. Time course of the monitoring parameters for the entire observation period separated by natural frequencies and α_{S2MAC}.

detect smaller changes. However, the Q(0.9) has a significant distance from the Q(0.75), indicating a large variation of the variance of the GP, which is also evident in Fig. 5. The α_{S2MAC} of the first bending mode is more sensitive to smaller angle changes. It is noticeable that the Q(0.5) of the α_{S2MAC} of solely the first bending mode pair for both GP types is already below the threshold in the original state. This suggests a potential system change, which is discussed in the following section.

5.2. Potential system change

Considering the R^2 scores in Table 2, as well as the damage detectability presented in the previous chapter, the GP for the α_{S2MAC} of the first bending mode pair for phase 1 and especially for phase 2 no longer seems to correspond to the identified α_{S2MAC}. This section describes the deviation in more detail. Fig. 8 shows the evaluated novelty metric for the natural frequencies and α_{S2MAC} defined in Section 3 with the settings in Table 3 for the entire period. After the training period, it can be observed that the density of the p_{MP} is below the threshold. This is evident in the α_{S2MAC}, where the moving average drops significantly below the threshold. For phase 1, this behaviour can be assumed as a usual effect of the regression model, because it is well-known that the model is not as suitable for unknown data from outside the training period. However, in phase 2 after a two-week data gap, where the turbine was shut down due to icing, the moving average of the p_{MP} values of α_{S2MAC} drops significantly. In contrast, the probabilities of the natural frequencies, do not change significantly. The cause for these deviations is in particular the α_{S2MAC} of the first bending mode in SS direction. The most probable values of the α_{S2MAC} in the top panel and S2MAC in the bottom panel, as well as the 99% confidence interval of the GP, are shown in Fig. 9 for a half-year period. It can be seen, that the identified S2MAC and α_{S2MAC} values systematically exceed the 99% confidence intervals of the GP in phase 2. Hence, a novelty can be assumed.

There are various explanations for the detected novelty. The possibility that the GP did not properly approximate the system behaviour at rarely represented EOCs can be ruled out, as no unknown EOCs occurred and the novelty was detected over a longer period of time. This is supported by the fact that the training of the other parameters was successful. Another indication, that is contrary to a failure of the model itself, is that a change can be observed in the historical trend of the α_{S2MAC} of the first SS bending mode depending on the nacelle position, which is shown in the panels of Fig. 10. Compared to the training period, a higher α_{S2MAC} can be observed in phases 1 and 2 at a nacelle position of around 200°.

Another reasonable explanation for the inconsistent behaviour of the actual system and the GP is that the MSS of the first bending mode pair has changed slightly. Either the cause could lie in the measurement device, or there may actually be a small structural system change. In contrast to the natural frequencies, the relative amplitude ratio between the sensors also affects the mode shapes. This tends to make the metrics based on mode shapes not only more sensitive to system changes but also more prone to measurement errors than metrics based on frequency tracking. Omitting the top measurement level 5 for the calculation of the α_{S2MAC}, the increase of the α_{S2MAC} in phase 1 and 2 at around 200° is no longer visible, as can be seen in the lower panels of Fig. 10. In this case, the novelty can no longer be detected using the method presented, as shown in Fig. 11. It is therefore more likely that the reason for the detected novelty is a change in the measurement technology than a change in the structure. Basically, this illustrates a typical challenge in the context of SHM, i.e. deciding whether a novelty is linked to a relevant system change or to an effect unrelated to the observed structure itself. In order to make reliable decisions, rugged

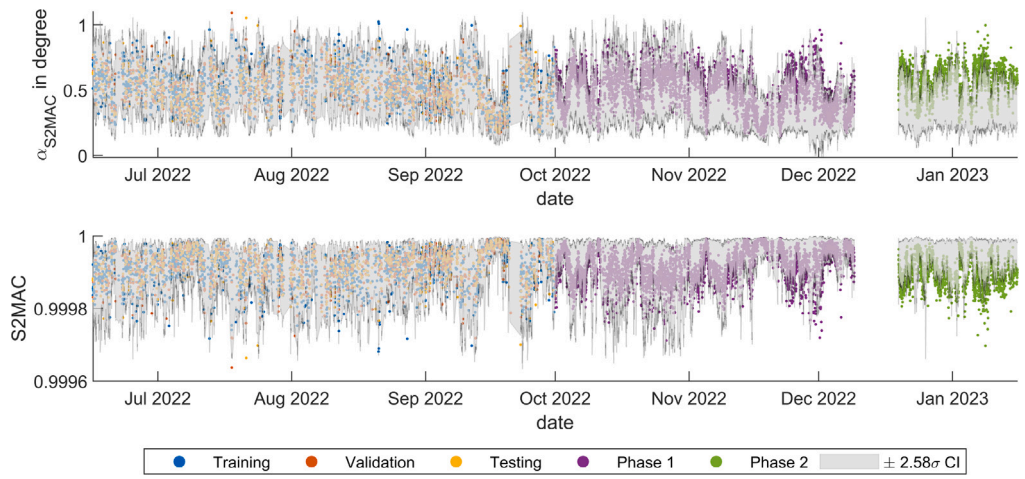


Fig. 9. Section of the identified and with the 99% CI of the GP predicted α_{S2MAC} and S2MAC of the first bending mode in SS direction from the middle of June 2022 to middle of January 2022.

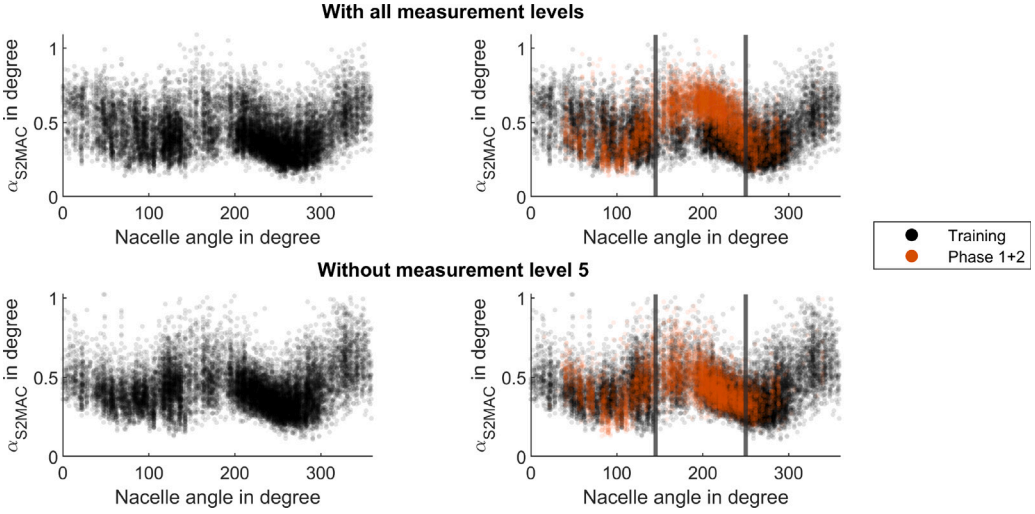


Fig. 10. The identified α_{S2MAC} of the first bending mode in SS direction depending on the nacelle angle. On the left only the training period, on the right the training and test periods. Upper panels α_{S2MAC} calculated with all measurement levels, lower panels without measurement level 5.

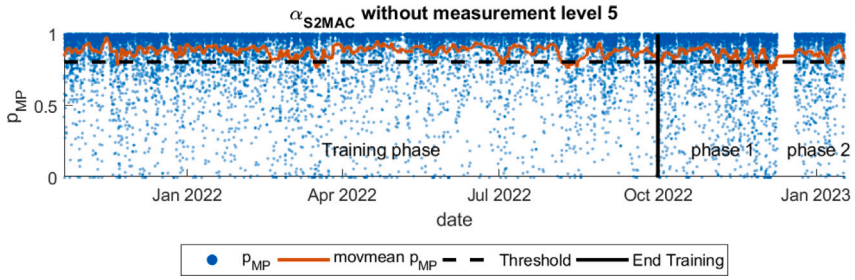


Fig. 11. Time course of the monitoring parameters α_{S2MAC} for the entire observation period without measurement level 5.

sensor technology as well as expert knowledge is essential for practical SHM applications.

6. Summary and outlook

In this study the data normalisation of natural frequencies and the α_{S2MAC} of bending modes of a tower of a full scale 3.4 MW wind turbine in operation was applied for the first time using sparse heteroscedastic GP regression considering the EOC dependent uncertainty. Unlike standard homoscedastic GP, the utilised GP incorporates heteroscedastic noise as a function of the input variables. In the context of wind turbine towers, this is a useful extension, since, for example, aerodynamic damping increases with wind speed, leading to larger identification uncertainties for the natural frequencies. When investigating the novelty detectability of a simplified damage scenario by decreasing natural frequencies, taking into account the identification and GP uncertainties, the data normalisation with a heteroscedastic GP can detect smaller changes than the homoscedastic variant. The results make it clear that the inclusion of input-dependent variance by means of heteroscedastic GPs are a useful addition to data normalisation

for this application. The α_{S2MAC} was used to investigate the damage detectability based on mode shapes. The results show that very small changes could be detected, which is also in line with experience from other studies [4,55]. In subsequent work, the detectability of various more realistic damage scenarios can be investigated with the aid of FE models, or small-scale experiments. In addition, an actual change in the α_{S2MAC} of the first bending mode was detected. A cause for the change could not be found so far. Plausible explanations are either a very small system change or minimal changes in the measurement device. The latter is more likely because, without the upper measuring level, the novelty is no longer detectable.

This study also revealed the strengths and weaknesses of mode shape-based condition metrics. These are often more damage-sensitive than natural frequencies but are also susceptible to changes in the measurement device. Therefore, it is essential for a monitoring system to be able to guarantee the reliability of the employed sensors over many years. Consequently, strategies on how to deal with possible ageing processes as well as ruggedisation and standardisation are required. In addition, the question of an optimal sensor setup is particularly relevant for such metrics. Here, it is crucial to clarify which system change would be detectable with a given distribution of sensors.

The data normalisation of natural frequencies and α_{S2MAC} with sparse heteroscedastic GPs has proven to be an adequate method for the presented application case. In the future, the identification uncertainties could be incorporated into the training of the GP as well. In addition, a further measurement campaign could investigate in detail which inputs should be included in the GP and whether, for example, the temperature gradient of the structure could have an influence, as in the case of bridges. As part of further research on wind turbine monitoring, different variants of sparse heteroscedastic GP need to be investigated and compared, but this is not the scope of this study. In addition, the incorporation of physical knowledge into the kernel or the basis function for the application domain could also be researched. To avoid a long learning period for each new wind turbine to be monitored, the transferability of the GP from one wind turbine to another with a transfer learning approach should also be investigated.

CRediT authorship contribution statement

Clemens Jonscher: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Sören Möller:** Writing – original draft, Formal analysis. **Leon Liesecke:** Writing – review & editing, Visualization. **Benedikt Hofmeister:** Writing – review & editing. **Tanja Grießmann:** Writing – review & editing, Supervision. **Raimund Rolfes:** Writing – review & editing, Supervision, Resources, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

We greatly acknowledge the financial support of the German Research Foundation (SFB-1463-434502799), the Federal Ministry for Economic Affairs and Climate Action of Germany (research projects *Deutsche Forschungsplattform für Windenergie*, FKZ 0325936E and *PreciWind-Präzises Messsystem zur berührungslosen Erfassung und Analyse des dynamischen Strömungsverhaltens von WEA-Rotorblättern*, FKZ 03EE3013B) that enabled this work. In addition, we are grateful to the *Deutsche WindGuard GmbH*, as well as the *Bremer Institut für Messtechnik, Automatisierung und Qualitätswissenschaft* (BIMAQ) for their support during the measurement campaign.

References

[1] Reynders E, Roeck G. Subspace identification for operational modal analysis. In: Maier G, Rammerstorfer FG, Salençon J, Schrefler B, Serafini P, Deraemaeker A, Worden K, editors. New trends in vibration based structural health monitoring. CISM international centre for mechanical sciences, vol. 520, Vienna: Springer Vienna; 2011, p. 55–106. http://dx.doi.org/10.1007/978-3-7091-0399-9{_}3.

[2] Au S-K, Zhang F-L, Ni Y-C. Bayesian operational modal analysis: Theory, computation, practice. Comput Struct 2013;126:3–14. http://dx.doi.org/10.1016/j.compstruc.2012.12.015.

[3] Au S-K, Brownjohn JM, Li B, Raby A. Understanding and managing identification uncertainty of close modes in operational modal analysis. Mech Syst Signal Process 2021;147:107018. http://dx.doi.org/10.1016/j.ymssp.2020.107018.

[4] Jonscher C, Liesecke L, Penner N, Hofmeister B, Grießmann T, Rolfes R. Influence of system changes on closely spaced modes of a large-scale concrete tower for the application to structural health monitoring. J Civ Struct Health Monit 2023;29(8):328. http://dx.doi.org/10.1007/s13349-023-00693-6.

[5] Jonscher C, Möller S, Liesecke L, Schuster D, Hofmeister B, Grießmann T, Rolfes R. Identification uncertainties of bending modes of an onshore wind turbine for vibration-based monitoring. Struct Control Health Monit 2024. publication in press.

[6] D'Ambrogio W, Fregolent A. Higher-order mac for the correlation of close and multiple modes. Mech Syst Signal Process 2003;17(3):599–610. http://dx.doi.org/10.1006/mssp.2002.1468.

[7] Farrar CR, Worden K. Structural health monitoring: a machine learning perspective. John Wiley & Sons; 2012.

[8] Farrar CR, Sohn H, Worden K. Data normalization : a key for structural health monitoring. 2001, URL https://www.osti.gov/biblio/975664.

[9] Lucà F, Manzoni S, Cigada A, Frate L. A vibration-based approach for health monitoring of tie-rods under uncertain environmental conditions. Mech Syst Signal Process 2022;167:108547. http://dx.doi.org/10.1016/j.ymssp.2021.108547.

[10] Bull T, Ulriksen MD, Tcherniak D. The effect of environmental and operational variabilities on damage detection in wind turbine blades. In: Proceedings of the 9th European workshop on structural health monitoring. 2018, URL http://www.bindt.org/events/ewshm-2018/. 9th European Workshop on Structural Health Monitoring.

[11] Filzmoser P, Garrett RG, Reimann C. Multivariate outlier detection in exploration geochemistry. Comput Geosci 2005;31(5):579–87. http://dx.doi.org/10.1016/j.cageo.2004.11.013.

[12] Limongelli MP, Manoach E, Quqa S, Giordano PF, Bhowmik B, Pakrashi V, Cigada A. Vibration response-based damage detection. In: Sause MGR, Jasiūnienė E, editors. Structural health monitoring damage detection systems for aerospace. Springer aerospace technology, Cham: Springer International Publishing; 2021, p. 133–73. http://dx.doi.org/10.1007/978-3-030-72192-3{_}6.

[13] Wah WSL, Chen Y-T, Roberts GW, Elamin A. Damage detection of structures subject to nonlinear effects of changing environmental conditions. Procedia Eng 2017;188:248–55. http://dx.doi.org/10.1016/j.proeng.2017.04.481.

[14] Lysgaard P, Amador SDR, Nielsen ST, Katsanos E, Brincker R. Vibration-based damage detection using input-output and output-only environmental models: A comparison. In: Pakzad S, editor. Dynamics of civil structures, volume 2. Conference proceedings of the society for experimental mechanics series, Cham: Springer International Publishing; 2021, p. 29–38. http://dx.doi.org/10.1007/978-3-030-47634-2{_}5.

[15] Sohn H. Effects of environmental and operational variability on structural health monitoring. Philos Trans R Soc Lond Ser A Math Phys Eng Sci 2007;365(1851):539–60. http://dx.doi.org/10.1098/rsta.2006.1935.

[16] Martinez-Luengo M, Kolios A, Wang L. Structural health monitoring of offshore wind turbines: A review through the statistical pattern recognition paradigm. Renew Sustain Energy Rev 2016;64:91–105. http://dx.doi.org/10.1016/j.rser.2016.05.085.

[17] Cross EJ, Koo KY, Brownjohn J, Worden K. Long-term monitoring and data analysis of the Tamar bridge. Mech Syst Signal Process 2013;35(1–2):16–34. http://dx.doi.org/10.1016/j.ymssp.2012.08.026.

[18] Spiridonakos MD, Chatzi EN, Sudret B. Polynomial chaos expansion models for the monitoring of structures under operational variability. ASCE-ASME J Risk Uncertain Eng Syst A 2016;2(3). http://dx.doi.org/10.1061/AJRUA6.0000872.

[19] Chandrasekhar K, Stevanovic N, Cross EJ, Dervilis N, Worden K. Damage detection in operational wind turbine blades using a new approach based on machine learning. Renew Energy 2021;168:1249–64. http://dx.doi.org/10.1016/j.renene.2020.12.119.

[20] Häckell MW, Rolfes R. Monitoring a 5MW offshore wind energy converter—Condition parameters and triangulation based extraction of modal parameters. Mech Syst Signal Process 2013;40(1):322–43. http://dx.doi.org/10.1016/j.ymssp.2013.04.004.

[21] Wickramarachchi CT, Rogers TJ, McLeay TE, Leahy W, Cross EJ. Online damage detection of cutting tools using Dirichlet process mixture models. Mech Syst Signal Process 2022;180:109434. http://dx.doi.org/10.1016/j.ymssp.2022.109434.

[22] García-Macías E, Ubertini F. Integrated SHM systems: Damage detection through unsupervised learning and data fusion. In: Cury A, Ribeiro D, Ubertini F, Todd MD, editors. Structural health monitoring based on data science techniques. Structural integrity, vol. 21, Cham: Springer International Publishing; 2022, p. 247–68. http://dx.doi.org/10.1007/978-3-030-81716-9{_}12.

[23] Cross EJ, Worden K, Chen Q. Cointegration: a novel approach for the removal of environmental trends in structural health monitoring data. Proc R Soc Lond Ser A Math Phys Eng Sci 2011;467(2133):2712–32. http://dx.doi.org/10.1098/rspa.2011.0023.

[24] Shi H. On nonlinear cointegration methods for structural health monitoring (Ph.D. thesis), University of Sheffield; 2018.

[25] Reynders E, Wursten G, de Roeck G. Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification. Struct Health Monit 2014;13(1):82–93. http://dx.doi.org/10.1177/1475921713502836.

[26] Sohn H, Worden K, Farrar CR. Statistical damage classification under changing environmental and operational conditions. J Intell Mater Syst Struct 2002;13(9):561–74.

[27] Dervilis N, Choi M, Taylor S, Barthorpe R, Park G, Farrar C, Worden K. On damage diagnosis for a wind turbine blade using pattern recognition. J Sound Vib 2014;333(6):1833–50.

[28] Fuentes R, Cross EJ, Gardner PA, Bull LA, Rogers TJ, Barthorpe RJ, Shi H, Dervilis N, FARRAR CR, Worden K. Structural health monitoring and damage identification. In: Allemang R, Avitabile P, editors. Handbook of experimental structural dynamics. New York, NY: Springer New York; 2020, p. 1–72. http://dx.doi.org/10.1007/978-1-4939-6503-8{_}23-2.

[29] Zhou K, Wang Z, Gao Q, Yuan S, Tang J. Recent advances in uncertainty quantification in structural response characterization and system identification. Probab Eng Mech 2023;103507.

[30] Wang Q-A, Dai Y, Ma Z-G, Ni Y-Q, Tang J-Q, Xu X-Q, Wu Z-Y. Towards probabilistic data-driven damage detection in SHM using sparse Bayesian learning scheme. Struct Control Health Monit 2022;29(11):e3070.

[31] Wang Q-A, Dai Y, Ma Z-G, Wang J-F, Lin J-F, Ni Y-Q, Ren W-X, Jiang J, Yang X, Yan J-R. Towards high-precision data modeling of SHM measurements using an improved sparse Bayesian learning scheme with strong generalization ability. Struct Health Monit 2024;23(1):588–604.

[32] Zhang Y-M, Wang H, Bai Y, Mao J-X, Xu Y-C. Bayesian dynamic regression for reconstructing missing data in structural health monitoring. Struct Health Monit 2022;21(5):2097–115.

[33] Botz M. Structural health monitoring der tragstruktur von windenergieanlagen (Ph.D. thesis), Technische Universität München; 2022.

[34] Oliveira G, Magalhães F, Cunha Á, Caetano E. Vibration-based damage detection in a wind turbine using 1 year of data. Struct Control Health Monit 2018;25(11):e2238. http://dx.doi.org/10.1002/stc.2238.

[35] Weijtjens W, Verbelen T, de Sitter G, Devriendt C. Foundation structural health monitoring of an offshore wind turbine—a full-scale case study. Struct Health Monit 2016;15(4):389–402. http://dx.doi.org/10.1177/1475921715586624.

[36] Ni YQ, Zhou HF, Ko JM. Generalization capability of neural network models for temperature-frequency correlation using monitoring data. J Struct Eng 2009;135(10):1290–300. http://dx.doi.org/10.1061/(ASCE)ST.1943-541X.0000050.

[37] Rasmussen CE, Williams CKI. Gaussian process for machine learning. London, England: The MIT Press; 2006.

[38] Avendaño-Valencia LD, Chatzi EN, Koo KY, Brownjohn JMW. Gaussian process time-series models for structures under operational variability. Front Built Environ 2017;3. http://dx.doi.org/10.3389/fbuil.2017.00069.

[39] Worden K, Cross EJ. On switching response surface models, with applications to the structural health monitoring of bridges. Mech Syst Signal Process 2018;98:139–56. http://dx.doi.org/10.1016/j.ymssp.2017.04.022.

[40] Möller S, Jonscher C, Grießmann T, Rolfes R. Investigations towards physics-informed Gaussian process regression for the estimation of modal parameters of a lattice tower under environmental conditions. In: Experimental vibration analysis for civil engineering structures. Springer International Publishing; 2023, http://dx.doi.org/10.1007/978-3-031-39117-0_41.

[41] Liu H, Ong Y-S, Shen X, Cai J. When Gaussian process meets big data: A review of scalable GPs. IEEE Trans Neural Netw Learn Syst 2020;31(11):4405–23. http://dx.doi.org/10.1109/TNNLS.2019.2957109.

[42] Le QV, Smola AJ, Canu S. Heteroscedastic Gaussian process regression. In: Dzeroski S, editor. Proceedings of the 22nd international conference on machine learning - ICML '05. New York, New York, USA: ACM Press; 2005, p. 489–96. http://dx.doi.org/10.1145/1102351.1102413.

[43] Kersting K, Plagemann C, Pfaff P, Burgard W. Most likely heteroscedastic Gaussian process regression. In: Proceedings of the 24th international conference on machine learning. 2007, p. 393–400.

[44] Almosallam IA, Jarvis MJ, Roberts SJ. GPz: non-stationary sparse Gaussian processes for heteroscedastic uncertainty estimation in photometric redshifts. Mon Not R Astron Soc 2016;462(1):726–39. http://dx.doi.org/10.1093/mnras/stw1618.

[45] Dervilis N, Shi H, Worden K, Cross EJ. Exploring environmental and operational variations in SHM data using heteroscedastic Gaussian processes. In: Pakzad S, Juan C, editors. Dynamics of civil structures, volume 2. Conference proceedings of the society for experimental mechanics series, vol. 51, Cham: Springer International Publishing; 2016, p. 145–53. http://dx.doi.org/10.1007/978-3-319-29751-4_15.

[46] Rogers TJ, Gardner P, Dervilis N, Worden K, Maguire AE, Papatheou E, Cross EJ. Probabilistic modelling of wind turbine power curves with application of heteroscedastic Gaussian process regression. Renew Energy 2020;148(10):1124–36. http://dx.doi.org/10.1016/j.renene.2019.09.145.

[47] Wang Q-A, Zhang C, Ma Z-G, Ni Y-Q. Modelling and forecasting of SHM strain measurement for a large-scale suspension bridge during typhoon events using variational heteroscedastic Gaussian process. Eng Struct 2022;251:113554.

[48] Wang H, Yuan S, Xu Q, Meng Y, Ren Y. A new GW-based heteroscedastic Gaussian process method for online crack evaluation. Struct Health Monit 2022;21(6):2874–89.

[49] Jones M, Rogers T, Worden K, Cross E. Heteroscedastic Gaussian processes for localising acoustic emission. In: Data science in engineering, volume 9: proceedings of the 39th IMAC, a conference and exposition on structural dynamics 2021. Springer; 2022, p. 185–97.

[50] Tolvanen V, Jylänki P, Vehtari A. Expectation propagation for nonstationary heteroscedastic Gaussian process regression. In: 2014 IEEE international workshop on machine learning for signal processing. MLSP, IEEE; 2014, p. 1–6.

[51] Quinonero-Candela J, Rasmussen CE. A unifying view of sparse approximate Gaussian process regression. J Mach Learn Res 2005;6:1939–59.

[52] Bauer M, van der Wilk M, Rasmussen CE. Understanding probabilistic sparse Gaussian process approximations. In: Proceedings of the 30th international conference on neural information processing systems. NIPS '16, Red Hook, NY, USA: Curran Associates Inc.; 2016, p. 1533–41.

[53] Almosallam IA. Heteroscedastic Gaussian processes for uncertain and incomplete data (Ph.D. thesis), University of Oxford; 2017.

[54] Hüllermeier E, Waegeman W. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. Mach Learn 2021;110(3):457–506. http://dx.doi.org/10.1007/s10994-021-05946-3.

[55] Jonscher C, Hofmeister B, Grießmann T, Rolfes R. Influence of environmental conditions and damage on closely spaced modes. In: Rizzo P, Milazzo A, editors. European workshop on structural health monitoring. Springer eBook collection, vol. 270, Cham: Springer International Publishing and Imprint Springer; 2023, p. 902–11. http://dx.doi.org/10.1007/978-3-031-07322-9_91.

[56] Lewis DD. Naive (Bayes) at forty: The independence assumption in information retrieval. In: Carbonell JG, Siekmann J, Goos G, Hartmanis J, van Leeuwen J, Nédellec C, Rouveirol C, editors. Machine learning: ECML-98. Lecture notes in computer science, vol. 1398, Berlin, Heidelberg: Springer Berlin Heidelberg; 1998, p. 4–15. http://dx.doi.org/10.1007/BFb0026666.

[57] Au S-K. Operational modal analysis: Modeling, Bayesian inference, uncertainty laws. 1st ed.. Singapore: Springer Singapore; 2017, http://dx.doi.org/10.1007/978-981-10-4118-1.

[58] Reynders EP. Uncertainty quantification in data-driven stochastic subspace identification. Mech Syst Signal Process 2021;151:107338.

[59] Almosallam IA. GPz 2.0 framework. 2021, URL https://github.com/OxfordML/GPz. Accessed on 2024-01-18.

[60] Akaike H. Information theory and an extension of the maximum likelihood principle. Sel Pap Hirotugu Akaike 1998;199–213.