Multivariate GP-VAR models for robust structural identification under operational variability

Avendaño-Valencia, Luis David; Chatzi, Eleni N.

Published in:
Probabilistic Engineering Mechanics

DOI:
10.1016/j.probengmech.2020.103035

Publication date:
2020

Document version:
Submitted manuscript

Citation for published version (APA):

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use
This work is brought to you by the University of Southern Denmark.
Unless otherwise specified it has been shared according to the terms for self-archiving.
If no other license is stated, these terms apply:

• You may download this work for personal use only.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.
Please direct all enquiries to puresupport@bib.sdu.dk

Download date: 15. Sep. 2023
Multivariate GP-VAR models for Robust Structural Identification under Operational Variability*

Luis David Avendaño-Valencia* a, Eleni N. Chatzi b

*Group of Applied AI and Data Science, Maersk-McKinney-Møller Institute, University of Southern Denmark. 
Campusvej 55, 5230 Odense M, Denmark.

bInstitute of Structural Engineering, Dept. of Civil, Environmental and Geomatic Engineering. ETH Zürich. 
Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland.

Abstract

While the concept of structural monitoring has been around for a number of decades, it remains under-exploited in practice. A main driver for this shortcoming lies in the difficulty to robustly and autonomously interpret the information that is extracted from dynamic data. This hindrance in properly deciphering the collected information may be attributed to the uncertainty that is inherent in i) the finite set of measured data, ii) the models employed for capturing the manifested dynamics, and more importantly, iii) the susceptibility of these systems to variations in Environmental and Operational Parameters (EOPs). In previous work of the authors, a Gaussian Process (GP) time-series approach has been introduced, which serves as a hierarchical input-output method to account for the influence of EOPs on structural response. This in turn enables a robust structural identification. In this scheme, the short-term dynamics are modeled by means of linear-in-the-parameters time-series models, while EOV dependence –acting on a long-term time scale– is achieved via GP regression of the model coefficients on measured EOPs. This work corresponds to a further advancement on this modeling approach, corresponding to its generalization to the vector response case. Particularly, the problem of global identification here is solved via an Expectation-Maximization algorithm tailored to the GP time-series model structure. Moreover, an EOP-dependent innovations covariance matrix is integrated in the model, which helps to capture variation in the vibration power. The resulting model does not only have the capability to represent the long-term response of a structure under variable EOPs, but also facilitates the enhanced tracking of modal quantities in contrast to traditional operational modal analysis techniques. The proposed approach is exemplified on the identification of the vibration response of a simulated wind turbine blade at different points along the blade axis in the flap-wise direction, under variability of both the acting wind speeds and ambient temperatures.

Keywords: Structural Health Monitoring (SHM), Environmental and Operational Variability (EOV), Gaussian Process (GPs) Vector AutoRegressive (GP-VAR) models, Data-Driven Condition Assessment, Wind Energy Infrastructure
1. Introduction

Despite its development since a number of decades and the maturity it has reached in terms of algorithmic implementation, Structural Health Monitoring (SHM) tools, relying on vibration-based monitoring approaches in particular, remain little exploited thus far into practice. A common shortcoming with respect to the reliability of such methods, when implemented onto field structures, pertains to the susceptibility of the SHM estimates to the variations of Environmental and Operational Parameters (EOPs). Structural properties, such as stiffness and even boundary conditions, are largely affected by EOPs [1, 2, 3]. This hardens the fulfillment of the SHM tasks, namely i) detection, ii) localization, iii) quantification and iv) prognosis [4], and further reduced our confidence in the Value of Information (VoI) that stems from monitoring data [5]. In order to prove the efficacy of SHM, it is essential to devise schemes rendering its implementation robust and almost autonomous. In doing so, exogenous effects such as varying EOPs ought to be accounted for. Such effects, which could include temperature, humidity, traffic, or even wind/wave loads, are of heterogeneous nature and might be acting across diverse temporal scales. For instance aeroleastic loads acting on a wind turbine present a variability within a temporal scale that is shorter than the variation of temperature or humidity. Here, we will refer to the first temporal scale, which is in the order of seconds or minutes, as short-term, whereas the second scale, which lies in the range of hours, days or months, is referred to as long-term.

Realizing the criticality of this issue, a number of works have appeared in recent literature for tackling the occurrence of Environmental and Operational Variability (EOV) [6, 7, 8]. These are largely classified into two main methodological fronts, namely output-only methods which operate purely on the basis of response measurements, and input-output or cause-effect methods, which establish a model of the response (output/effet) in relation to the measured EOPs (input/cause). In the first class, the key pursuit focuses on discovery of features that are latent within the output datasets (e.g. natural frequencies), and which may be projected on a space where the EOP influence is removed. The output-only class typically includes the so-called projection, or data normalization methods [6, 9]. As manifestation of such schemes, we here list the cointegration methods [10], as well as Principal Component Analysis (PCA) and factor analysis [11, 12]. These have in recent works been extended to account for nonlinear dependence on the affecting inputs (kernel-PCA [13]; regime-switching cointegration [14]). A major benefit of output-only schemes lies in the alleviation of the need to explicitly measure inputs, which might not always be quantifiable.

However, if access to some important inputs is ensured, then inclusion of this information into a stochastic scheme can enhance the accuracy of the resulting predictive models. This is a main pursuit of the second class of methods for tackling EOV, i.e., the input-output class, which commonly delivers a functional relationship between measured EOPs and monitored outputs. This relationship can be either deterministic or stochastic. In the former case, standard methods include regression schemes [15], while a more refined alternative is achieved via adoption of time series models, such as the Functionally Pooled (FP) time-series models [16, 17]. A deterministic

©2020. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/
*Corresponding author.
Email addresses: lday@mmmi.sdu.dk (Luis David Avendaño-Valencia), chatzi@ibk.baug.ethz.ch (Eleni N. Chatzi)
representation poses nonetheless severe limitations on our potential to explain further uncaptured or non-monitored effects. To this end, a stochastic formulation may instead be utilized, which effectively captures the uncertainty that is inherent in the measurements, as well as in the typically lacking capabilities of our models. This is achieved via stochastic functional dependence models, where response features are treated as stochastic variables whose dependence on EOPs is now formalized by a functional representation. As a next step, and once these models are established, the detection task can be rendered robust via use of suitable outlier detection schemes [18, 19].

The authors have in recent years worked on the derivation of various alternatives for stochastic functional dependence schemes, with the aim to develop a framework that is robust and as autonomous as possible for data-driven diagnostics. This includes Polynomial Chaos Expansion-based (PCE) models [20, 21, 22, 23], as well as Random Coefficient (RC) models [24, 25]. In more recent work [26] we introduce a hierarchical framework, able to distinguish among diverse temporal scales. In this context, the short-term dynamics and associated dependencies are accounted for by means of a linear-in-the-parameters time-series model, while the long-term stochastic dependency on measured EOPs is represented via Gaussian Process (GP) regression. This is termed a GP time-series model, which is particularly attractive in enabling the representation of complex dynamics (even non-stationary) over extended periods [27]. However, the explicit inclusion of inputs into the formulation often implies an increased dimensionality, and an uneconomical size for the resulting GP time-series models, whose training is hardened [28, 29].

This work correspond to a further advancement of our previous works on the topic, this time concerning to the application of GP time-series modeling method in vector response data. Particularly, two are the main contributions of this work: first, the problem of global identification is addressed here by means of an Expectation-Maximization algorithm specifically tailored to a GP Vector AutoRegressive (GP-VAR) model structure. In the core of the EM algorithm lies the estimation of the posterior model parameter PDF, which is done efficiently by means of a recursive Bayesian updating scheme. Second, an EOP-dependent innovations covariance matrix is now integrated in the model, which helps to capture variation in the vibration power, thus enhancing the modeling capabilities of the method. For that purpose, a functional-series expansion of the coefficients of the Choleski factorization of the innovations covariance is hereby considered. Both contributions come in addition to a dimensionality dimension scheme aimed at reducing the complexity of GP-VAR models, discussed in [29]. The resulting GP-VAR modeling framework does not only have the capability to represent the long-term variations in the response of a structure under variable EOPs, but also facilitates the enhanced tracking of modal quantities in contrast to traditional operational modal analysis techniques.

The capabilities of the proposed methodology are illustrated on the identification of the simulated vector acceleration response of a wind turbine blade, which undergoes yearly variations in both the acting wind speeds and ambient temperatures. The GP-VAR model identification process is thereby illustrated, focusing on the effect of different structural parameters on the final model performance. Furthermore, the modal parameter tracking capability of the GP-VAR models is contrasted with that achievable the by the Covariance-Driven Stochastic Subspace Identification (SSI-COV) operational modal analysis technique [30].

This work is organized as follows: Section 2 provides the main definitions related to the GP time-series modeling method, including the definition of the GP-VAR model 2.1 and the EOP-dependent innovations covariance matrix 2.3. Subsequently, the model identification issue is addressed in Section 3 and summarized in Section 4, while modal analysis based on the identified models is discussed in Section 5. An application example is provided in Section 6, involving the simulated vibration response of a wind turbine blade. Finally, the conclusions of this study are
2. Overview of Gaussian Process time-series models

2.1. Main definitions

A Gaussian Process (GP) time-series model of the response vector \( y[t] \in \mathbb{R}^n \) is defined as [26]:

\[
y[t] = \phi^T [t] \cdot \theta + \varepsilon[t], \quad \varepsilon[t] \sim \text{NID}(0_{n \times 1}, \Sigma_\varepsilon) \quad (1a)
\]

\[
\theta = W \cdot f(\xi) + u, \quad u \sim N(u|0_{n \times 1}, \Sigma_u) \quad (1b)
\]

where \( \phi[t] \in \mathbb{R}^{d \times n} \) is the regression matrix, \( \theta := \theta(\xi) \in \mathbb{R}^d \) is the parameter vector, and \( \varepsilon[t] \in \mathbb{R}^n \) is a zero-mean Normally and Independently Distributed (NID) vector innovations, with covariance matrix \( \Sigma_\varepsilon \in \mathbb{R}^{n \times d} \). The parameter vector follows a Gaussian Process (GP) regression on the EOP vector \( \xi \in \mathbb{R}^n \), and determined by the coefficient matrix \( W = \begin{bmatrix} w_1 & w_2 & \cdots & w_p \end{bmatrix}, \quad W \in \mathbb{R}^{d \times p} \), the GPR basis \( f(\xi) \in \mathbb{R}^p \), and the parameter covariance matrix \( \Sigma_u \in \mathbb{R}^{d \times d} \).

The GP time-series model is characterized by the conditional multivariate normal PDFs [26]:

\[
p(y[t]|\theta, \xi, \mathcal{P}) = N(y[t]|\phi^T [t] \cdot \theta, \Sigma_\varepsilon) \quad (2a)
\]

\[
p(\theta|\xi, \mathcal{P}) = N(\theta|W \cdot f(\xi), \Sigma_u) \quad (2b)
\]

where \( \mathcal{P} \triangleq \{ W, \Sigma_\varepsilon, \Sigma_u \} \) denotes the set of hyperparameters of the GP time-series model, while \( \mathcal{N}(x|\mu, \Sigma) \) indicates a (multivariate) normal distribution of the random variable \( x \) with mean \( \mu \) and covariance matrix \( \Sigma \).

2.2. GP Vector AutoRegressive models

The specific type of GP time-series model is determined by the structure of the regression matrix. In the present context, the selected time-series model corresponds to a Vector AutoRegressive (VAR) model, defined as follows [26]:

\[
y[t] = \sum_{i=1}^{n_a} A_i(\xi) \cdot y[t-i] + \varepsilon[t], \quad \varepsilon[t] \sim \text{NID}(0_{n \times 1}, \Sigma_\varepsilon) \quad (3)
\]

where \( A_i(\xi) \in \mathbb{R}^{n \times n} \) is the i-th AR coefficient matrix and \( n_a \) is the autoregressive order.

Eq. (3) can be cast into the regression form used in the definition of the GP time-series model (Eq.(1)) with the help of the following definitions [26]:

\[
y[t] = \phi^T [t] \cdot \theta + \varepsilon[t], \quad \phi[t] = I_{n_a} \otimes z[t], \quad \theta = \text{vec}(A(\xi)) \quad (4)
\]

where \( A(\xi) = [ A_1(\xi) \quad A_2(\xi) \quad \cdots \quad A_{n_a}(\xi) ] \cdot A(\xi) \in \mathbb{R}^{n \times n_a} \) is the coefficient matrix of the VAR model, \( z[t] = \begin{bmatrix} y^T [t-1] & y^T [t-2] & \cdots & y^T [t-n_a] \end{bmatrix}^T \), \( z[t] \in \mathbb{R}^{n \times 1} \) is the vector of regressors, \( I_{n_a} \) indicates the \( n \)-dimensional identity matrix, and \( \text{vec}() \) indicates the vectorization operation, which stacks the columns of the matrix in the argument into a column vector.
Often, the power of the vibration response of a structure depends on EOPs. For instance, the power of the vibration on aeroelastic structures, like wind turbines or long-span bridges, depends on the wind speed magnitude. For this case, it is useful to consider a time-series model with changing innovations covariance. Accordingly, a GP time-series model with EOP-dependent covariance matrix may be defined by means of the functional series expansion of the factorized innovations covariance matrix, as follows:

\[
\Sigma_x := \Sigma_\epsilon(\xi) = S(\xi) \cdot S^T(\xi) \\
S(\xi) = \text{vec}(S(\xi)) = \mathbf{W}_s \cdot f_s(\xi)
\]

(5a) \hspace{1cm} (5b)

where \( S(\xi) \in \mathbb{R}^{n \times n} \) is the square root of \( \Sigma_x \) (i.e., via the Cholesky factorization), \( s(\xi) \in \mathbb{R}^{n} \) is a column vector built by stacking the columns of \( \Sigma_x \), and \( \mathbf{W}_s \in \mathbb{R}^{np} \) and \( f_s(\xi) \in \mathbb{R}^{p} \) are the coefficient matrix and functional expansion basis of the factorized innovations covariance.

In contrast to the parameter vector \( \theta \), in the case of the innovations covariance a fully deterministic functional series expansion is appraised to avoid making the model too complex. The expansion is made on the square root matrix \( S(\xi) \) instead of directly on the innovations covariance matrix to ensure that the matrices resulting from the functional series expansion remain positive definite. In addition, if triangular square root matrices are used, as in the Cholesky factorization, then, the number of (non-zero) parameters in the functional series expansion is reduced from \( n^2 \) to \( n \cdot (n + 1)/2 \).

Under the EOP-dependent innovations covariance, the resulting GP time-series model is characterized by the PDFs:

\[
p(y[i] | \theta, \xi, \mathcal{P}) = N(y[i] | \phi^T[i] \cdot \theta, \Sigma_x), \quad \Sigma_x = S(\xi) \cdot S^T(\xi)
\]

(6a) \hspace{1cm} (6b)

while the hyperparameters in this case become \( \mathcal{P} = [\mathbf{W}, \mathbf{W}_s, \Sigma_\theta] \).

3. Identification of GP-VAR models

Model identification involves the estimation of the hyperparameters \( \mathcal{P} \) and the selection of the model structure, consisting of the VAR order \( n_a \) and the GPR basis dimensionality \( p \). For this purpose, a set of \( K \) experiments (trials) are available, comprised by the response vectors \( y_k[i], i = 1, \ldots, N \), and respective EOP vectors \( \xi_k \).

3.1. Maximum Likelihood Estimation of the hyperparameters of GP-VAR models

Maximum Likelihood (ML) estimation is based upon the optimization of the likelihood of the hyperparameters \( \mathcal{P} \) given the available data \( \mathcal{D} \), where the set of available data corresponds to \( \mathcal{D} = \{ Y_k, \xi_k \ | \ k = 1, \ldots, K \} \), with \( Y_k = [y_k^T[1] \ y_k^T[2] \ \cdots \ y_k^T[N]]^T \) representing the vector of stacked response time series at trial \( k \). Nonetheless, in the case of GP-VAR models, there is an additional set of unknown data \( \mathcal{U} = \{ \theta_k \ | \ k = 1, \ldots, K \} \) involved in the problem, comprised by the parameter vectors \( \theta_k \).
If both sets $\mathcal{D}$ and $\mathcal{U}$ were available, then the (complete data) log-likelihood of the GP-VAR model takes the form:\footnote{In the case of EOP-dependent innovations covariance, $\Sigma_x$ is substituted by $\Sigma_x = \mathbf{S}(\xi_i) \cdot \mathbf{S}(\xi_i)^T$, with $s(\xi_i) = \text{vec}(\mathbf{S}(\xi_i)) = W \cdot f_i(\xi_i)$. However, the complete notation for this case is avoided in the remainder of this section to facilitate the readability of the text, and will be recalled in the cases where it is strictly necessary.}

$$
\ln \mathcal{L}(\mathcal{P} | \mathcal{D}, \mathcal{U}) = \sum_{k=1}^{K} \sum_{t=1}^{N} \ln p(y_t, \theta_k | \xi_t, \mathcal{P}) \\
= \sum_{k=1}^{K} \left( \ln p(\theta_k | \xi_t, \mathcal{P}) + \sum_{t=1}^{N} \ln p(y_t | \theta_k, \xi_t, \mathcal{P}) \right) \\
= -\frac{1}{2} \sum_{k=1}^{K} \left( \ln 2\pi |\Sigma_\theta| + (\theta - W f(\xi_k))^T \cdot \Sigma_\theta^{-1} \cdot (\theta - W f(\xi_k)) \right) \\
+ \frac{N}{2} \ln 2\pi |\Sigma_y| + \sum_{t=1}^{N} (y_t - \phi(f(\xi_k))^T \cdot \Sigma_y^{-1} \cdot (y_t - \phi(f(\xi_k) \theta_k))) \\
= -\frac{1}{2} \sum_{k=1}^{K} \ln 2\pi |\Sigma_\theta| + \left( \sum_{t=1}^{N} (y_t - \phi(f(\xi_k))^T \cdot \Sigma_y^{-1} \cdot (y_t - \phi(f(\xi_k) \theta_k)) \right) \\
$$

(7)

However, since the parameter vectors are unavailable, optimization upon the complete data likelihood in Eq. (7) is unfeasible. Two alternatives to solve the optimization problem are discussed in the sequel.

3.2. The marginal likelihood method

A first alternative is to marginalize the likelihood with respect to the unknown parameters in order to obtain a new likelihood that is independent from the unknown data. This is achieved by means of the marginalization operation:

$$
p(y_t | \xi_t, \mathcal{P}) = \int_{\Theta} p(y_t, \theta | \xi_t, \mathcal{P}) d\theta = N(y_t | \hat{\theta}, S_x[t]) \\
\hat{\theta} = W f(\xi_k), \quad S_x[t] = \Sigma_x + \phi(f(\xi_k))^T \cdot \Sigma_y \cdot \phi(f(\xi_k)) \\
$$

(8)

where $\Theta \subseteq \mathbb{R}^d$ is the complete parameter space. Hence, optimization may be performed instead on the marginalized likelihood:

$$
\ln \mathcal{L}(\mathcal{P} | \mathcal{D}) = \sum_{k=1}^{K} \sum_{t=1}^{N} \ln p(y_t | \xi_t, \mathcal{P}) \\
= -\frac{1}{2} \sum_{k=1}^{K} \ln 2\pi |\Sigma_\theta| + \left( \sum_{t=1}^{N} (y_t - \phi(f(\xi_k))^T \cdot \Sigma_y^{-1} \cdot (y_t - \phi(f(\xi_k) \hat{\theta}))) \right) \\
$$

(9)

Thus, ML estimates of the GP coefficient matrix $W$ can be obtained by differentiating Eq. (9) with respect to $W$, equate to zero, and then solve for $W$, which leads to a sort of least squares type of solution. Nonetheless, this procedure is not possible when the covariance matrices are involved in the optimization, since the objective function is not quadratic anymore. In that case, it is necessary to utilize iterative non-linear optimization methods, which in the end turns out to be quite computationally inefficient. Despite of the problem of applicability for optimization, the
marginal likelihood is a key quantity for comparison among different candidate models. Indeed, in the context of Bayesian model selection, two models represented by hyperparameters $\mathcal{P}_a$ and $\mathcal{P}_b$, are compared in terms of the Bayes factor:

$$
B(\mathcal{P}_a, \mathcal{P}_b) = \frac{p(\mathcal{P}_a | \mathcal{D})}{p(\mathcal{P}_a | \mathcal{D})} = \frac{p(\mathcal{D} | \mathcal{P}_a) \cdot p(\mathcal{P}_a)}{p(\mathcal{D} | \mathcal{P}_b) \cdot p(\mathcal{P}_b)}
$$

(10)

where $p(\mathcal{P}_a)$ and $p(\mathcal{P}_b)$ represent the prior probabilities of both models. Typically, these priors are assumed to be equal, indicating that all the models are equally plausible, in which case, the Bayes factor reduces to the ratio between the marginal likelihoods of each model. A Bayes factor equal to one indicates that both models are equally suited for the data, while a value larger than one indicates an inclination towards model $a$, and a value lower than one indicates the opposite.

3.3. The Expectation-Maximization algorithm

A second alternative to cope with the problem of missing parameters is to use estimates of the missing data to produce updated estimates of the hyperparameters, which is the core idea behind the Expectation-Maximization (EM) algorithm. Accordingly, the EM algorithm is based upon optimization of the quantity:

$$
Q(\mathcal{P} | \mathcal{P}_i) = E_{\mathcal{U} | \mathcal{D}, \mathcal{P}_i} \left[ -\ln L(\mathcal{P} | \mathcal{D}, \mathcal{U}) \right] = E \left[ -\ln L(\mathcal{P} | \mathcal{D}, \mathcal{U}) \right] | \mathcal{D}, \mathcal{P}_i
$$

(11)

where $\mathcal{P}_i$ indicates the current value of the hyperparameters, and $E_{\mathcal{U} | \mathcal{D}, \mathcal{P}_i} \{ \cdot \}$ indicates the conditional expectation with respect to data $\mathcal{D}$ and current hyperparameter values $\mathcal{P}_i$. Applying the previous definition into the GP-VAR model complete data likelihood, yields:

$$
Q(\mathcal{P} | \mathcal{P}_i) = \frac{1}{2} \sum_{k=1}^{K} \left( \ln 2\pi |\Sigma_k| + \hat{\theta}_k^T \cdot \Sigma_k^{-1} \cdot \hat{\theta}_k + \text{tr}(\Sigma_k^{-1} \cdot P_k) \right)
$$

$$
+ N \ln 2\pi |\Sigma| + \sum_{t=1}^{N} \left( \hat{\epsilon}_k[t] \cdot \Sigma_k^{-1} \cdot \hat{\epsilon}_k[t] + \text{tr}(\Sigma_k^{-1} \cdot \phi_k[t] \cdot P_k \cdot \phi_k[t]) \right)
$$

(12)

where

$$
\hat{\theta}_k = \hat{\theta}_k - Wf(\xi_k),
$$

$$
\hat{\epsilon}_k[t] = y_k[t] - \phi_k[t] \cdot \hat{\theta}_k
$$

(13)

and where $\hat{\theta}_k$ and $P_k$ are the mean and covariance matrix of the conditional distribution of $\theta$ given data $Y_k, \xi_k$ and hyperparameters $\mathcal{P}_i$, otherwise known as the posterior parameter distribution $p(\theta | Y_k, \xi_k, \mathcal{P}_i)$, which takes the form:

$$
p(\theta | Y_k, \xi_k, \mathcal{P}_i) = p(Y_k | \theta, \xi_k, \mathcal{P}_i) \cdot p(\theta | \xi_k, \mathcal{P}_i) \cdot p^{-1}(Y_k | \xi_k, \mathcal{P}_i) = N(\theta | \hat{\theta}_k, P_k)
$$

(14)

The gaussianity of the posterior distribution follows from the gaussianity of all the distributions involved in the right hand side of the equation. Details on the calculation of the parameter posterior mean and covariance matrix will follow next. The EM algorithm then attempts at finding the maximum likelihood estimates of the hyperparameters by the iterative application of the following steps:
• **Expectation step**: Calculate:

\[ Q(P|P_t) = E_{U,D,F} (- \ln L(P|D,U)) \]

which presently limits to calculate the posterior parameter distribution \( p(\theta|Y_k, \xi_k, P_t) \) for \( k = 1, \ldots, K \).

• **Maximization step**: Update the hyperparameter values according to:

\[ P_{t+1} = \arg \max_P Q(P|P_t) \]

The expectation and maximization steps are iterated until either the relative change on the hyperparameters or the objective function reaches a desired limit, or a maximum number of iterations is accomplished. Details on the expectation and maximization steps for the specific case of GP-VAR models are provided next. Further topics including convergence and adjustment of the EM algorithm can be found in [31].

### 3.4. The expectation step

The expectation step consists on the evaluation of the conditional expectation \( Q(P|P_t) \) as follows, which, as observed in Eq. (12), requires the calculation of the posterior parameter mean and covariance. The approach utilized in [26] for calculation of these quantities is formulated on direct calculations over the vector of stacked responses at each trial \( Y_k \) within the context of scalar vibration response, which leads to large sized covariance matrices.

In this section, a computationally efficient algorithm based on recursive Bayesian updating of the parameter posterior is formulated.

To start with, consider the factorization of the parameter posterior defined in Eq. (14):

\[
p(\theta|Y_k, \xi_k, P_t) = c \cdot p(Y_k|\theta, \xi_k, P_t) \cdot p(\theta|\xi_k, P_t) = c \cdot p(y_k[N], y_k[N-1], \ldots, y_k[1]|\theta, \xi_k, P_t) \cdot p(\theta|\xi_k, P_t)
\]

\[
= c \cdot p(y_k[N]|\theta, \xi_k, P_t) \cdot p(y_k[N-1], \ldots, y_k[1]|\theta, \xi_k, P_t) \cdot p(\theta|\xi_k, P_t)
\]

where \( c = p^{-1}(Y_k|\xi_k, P_t) \) and \( c_N = p^{-1}(y_k[N]|\xi_k, P_t) \). This factorization suggests that one can compute the posterior with data up to time \( t-1 \) and update it when a new piece of data at time \( t \) arrives. Accordingly, the posterior PDF can be updated with observations up to time \( t \) as follows,

\[
p(\theta|y_k[1], \ldots, y_k[t], \xi_k, P_t) = p(y_k[t]|\theta, \xi_k, P_t) \cdot p(\theta|y_k[1], \ldots, y_k[t-1], \xi_k, P_t)
\]

while the initial value is computed at \( t = 1 \) as follows,

\[
p(\theta|y_k[1], \xi_k, P_t) = p(y_k[1]|\theta, \xi_k, P_t) \cdot p(\theta|\xi_k, P_t)
\]

Now, since both \( p(y_k[t]|\theta, \xi_k, P_t) \) and \( p(\theta|\xi_k, P_t) \) are Gaussian (see Eqn. (2)), then, the posterior is Gaussian too, of the form [26],

\[
p(\theta|y_k[t], \ldots, y_k[1]|\xi_k, P_t) := N(\hat{\theta}|\hat{\theta}_k[t], P_{\hat{\theta}}[t])
\]
where the posterior mean and covariance matrix constructed from data up to time $t$, $\hat{\theta}_i[t]$ and $P_{ii}[t]$ respectively, are recursively updated on times $t = 1, \ldots, N$ via the equation set:

$$
\hat{\theta}_i[t] = \hat{\theta}_i[t - 1] + K \cdot (y_i[t] - \phi_i^T[t] \cdot \hat{\theta}_i[t - 1])
$$

(19a)

$$
P_{ii}[t] = (I_d - K \cdot \phi_i^T[t]) \cdot P_{ii}[t - 1]
$$

(19b)

$$
K = P_{ii}[t - 1] \cdot \phi_i[t] \cdot \phi_i^T[t]
$$

(19c)

$$
S_i[t] = \Sigma_{ii} + \phi_i^T[t] \cdot P_{ii}[t] \cdot \phi_i[t]
$$

(19d)

with initial values:

$$
\hat{\theta}_i[0] = W_i \cdot f(\xi_i) \quad P_{ii}[0] = \Sigma_{ii}
$$

(20)

and where $P_i = [W_i, \Sigma_{ii}, \Sigma_{ii}]$ denote the hyperparameters of the current EM iteration. Moreover, in the case of EOP-dependent covariance matrix the covariance matrix $\Sigma_{ii} := \Sigma_{ii}(\xi_i)$ at each trial $k$ is obtained from the functional series expansion in Eq. (5) with coefficient matrix $W_i$.

There is an notorious similarity between the recursive updating equations shown in Eq. (19) and the Kalman filter working as a parameter estimator [32]. However, in the present case the recursive posterior updating algorithm is much simpler since there are no dynamic constraints on the parameter evolution, while the algorithm is set to converge to the complete data posterior PDF $p(\theta | Y_i, \xi_i, P_i)$. As a result, the values $\hat{\theta}_i[N]$ and $P_{ii}[N]$ correspond to the mean and covariance matrix of $p(\theta | Y_i, \xi_i, P_i)$, which are to be used in the calculations of the maximization step.

3.5. The maximization step

After having computed the mean and covariance matrix of the posterior PDFs $p(\theta | Y_i, \xi_i, P_i)$, for all trials $k = 1, \ldots, K$, then the maximization step involves the calculation of the hyperparameters that maximize the objective function $Q(P | P_i)$. Two cases are observed here, the first where the innovations covariance matrix is constant, and the second where the innovations covariance matrix depends on the EOPs, as defined in Sec. 2.3.

3.5.1. Constant innovations covariance

Expressions to calculate updated hyperparameter values are obtained after calculating the partial derivatives of $Q(P | P_i)$ with respect to $W_i, \Sigma_{ii}$ and $\Sigma_{ii}$ and equating to zero. After solving for each one of the quantities of interest, the following hyperparameter update equations are obtained [26]:

$$
W_{ii+1} = \left( \sum_{k=1}^{K} \hat{\theta}_k \cdot f^T(\xi_i) \right) \cdot \left( \sum_{k=1}^{K} f(\xi_i) \cdot f^T(\xi_i) \right)^{-1}
$$

(21a)

$$
\Sigma_{ii+1} = \frac{1}{KN} \cdot \sum_{k=1}^{K} \sum_{i=1}^{N} \hat{\xi}_i[t] \cdot \hat{\xi}_i^T[t] + \phi_i^T[t] \cdot P_{ii} \cdot \phi_i[t]
$$

(21b)

$$
\Sigma_{ii+1} = \frac{1}{K} \cdot \sum_{k=1}^{K} \sum_{i=1}^{N} \hat{\theta}_k \cdot \hat{\theta}_k^T + P_{ii}
$$

(21c)

3.5.2. EOP-dependent innovations covariance

In this case, the update equations for the GP coefficient matrix and the parameter covariance matrix remain the same. On the other hand, the updated innovations covariance coefficient of projection matrix $W_i$, as in Eq. (5), are obtained via a two-step procedure.
**Calculate individual covariances.** The innovations covariances at each trial \( k = 1, \ldots, K \) are obtained by calculating the derivative of \( Q(P|P_i) \) with respect to \( \Sigma := \Sigma_e(\xi_k) \) and equating to zero, which yields:

\[
\Sigma_e(\xi_k) = \frac{1}{N} \sum_{i=1}^{N} \tilde{e}_k[i] \cdot \tilde{e}_k[i] + \phi_k^T[i] \cdot P_{th} \cdot \phi_k[i] \tag{22}
\]

Subsequently, the factors \( \hat{S}(\xi_k) \) of the Choleski decomposition of \( \hat{\Sigma}_e(\xi_k) \) are calculated.

**Calculate the projection coefficients.** Updates of the projection coefficients of the functional series expansion are calculated from the estimates of the vectorized Choleski factors \( \hat{s}(\xi_k) = \text{vec}(\hat{S}(\xi_k)) \), as follows:

\[
W_{n,i1} = \left[ \sum_{k=1}^{K} \hat{s}(\xi_k) \cdot f'_i(\xi_k) \right] \cdot \left( \sum_{k=1}^{K} f_i(\xi_k) \cdot f'_i(\xi_k) \right)^{-1} \tag{23}
\]

**4. Summary of the EM algorithm-based identification procedure**

A summary of the EM algorithm-based identification procedure for the GP-VAR model with constant innovations covariance is provided in Table 1.

**5. GP-VAR model based analysis**

Once a GP-VAR model is available, it is possible to evaluate the dynamic characteristics of the vibration. For instance, modal analysis based on a GP-VAR model with parameter matrices \( A_i(\xi), i = 1, \ldots, n_{a} \) under conditions \( \xi \), is obtained via the eigenvalue decomposition of the matrix:

\[
A(\xi) = \begin{bmatrix}
A_1(\xi) & A_2(\xi) & \cdots & A_{n_{a}-1}(\xi) & A_{n_{a}}(\xi) \\
I_{n} & 0_{n} & \cdots & 0_{n} & 0_{n} \\
0_{n} & I_{n} & \cdots & 0_{n} & 0_{n} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0_{n} & 0_{n} & \cdots & I_{n} & 0_{n}
\end{bmatrix} \tag{24}
\]

which leads to a set of \( n_{a} \) eigenvalues \( \lambda_i(\xi) \) with respective eigenvectors \( \psi_i(\xi) \). In turn, the natural frequencies \( f_n(\xi) \), damping ratios \( \zeta_i(\xi) \), and mode shapes \( \psi_i(\xi) \in \mathbb{R}^{n} \) for conditions \( \xi \), are extracted from the corresponding eigenvalues and eigenvectors, as follows:

\[
\begin{align*}
f_n(\xi) & = \left| \ln \lambda_i(\xi) \right| \frac{f_s}{2\pi} \\
\zeta_i(\xi) & = \cos \arg \ln \lambda_i(\xi) \\
\psi_i(\xi) & = \left[ v_{i,1}(\xi) \ v_{i,2}(\xi) \ \cdots \ v_{i,n}(\xi) \right]^T
\end{align*} \tag{25}
\]

where \( f_s \) indicates the sampling frequency, and \( v_{i,j}(\xi) \) indicates the \( j \)-th entry of the vector \( \psi_i(\xi) \).
Table 1: Expectation-Maximization algorithm for the identification of the hyperparameters of GP-VAR models.

**Required data:**
- The set of $K$ trials, comprised by the response vectors $y_k[t], t = 1, \ldots, N$, and the respective EOP vector $\xi_k$.
- Structural parameters, including VAR model order $n_q$ and functional expansion order $p$.
- Maximum number of iterations of the EM algorithm $N_{iter}$ and thresholds for the minimum allowable variation in the objective function and the hyperparameters, $\rho_Q$ and $\rho_p$, respectively.

**Initialization:**
- Define initial hyperparameter values $\mathcal{P}_0 = \{W_0, \Sigma_{a_0}, \Sigma_{b_0}\}$ as $W_0 = 0$, $\Sigma_{a_0} = s_a^2 \cdot I_n$, and $\Sigma_{b_0} = s_b^2 \cdot I_n$, with $s_a^2$ and $s_b^2$ defined as large positive scalars (estimates in first EM iteration will approximate conventional ML estimates).

**EM iteration:**
- Define iteration index $i = 1$, and
- **Expectation step:** For each trial $k = 1, \ldots, K$:
  - Define initial values $\hat{\theta}_i[t] = \mathbf{W}_i \cdot f(\xi_k)$ and $\mathbf{P}_a[i] = \Sigma_{a_0}$.
  - For $t = 1, \ldots, N$ calculate:
    - $\hat{\theta}_i[t] = \hat{\theta}[t] - 1 + K \cdot (y_k[t] - \phi_k[t] \cdot \hat{\theta}_i[t - 1])$
    - $\mathbf{P}_a[i] = (I_q - K \cdot \phi_k[t]) \cdot \mathbf{P}_a[i - 1]$
    - $\mathbf{K} = \mathbf{P}_a[i - 1] \cdot \phi_k[t] \cdot S_k[t]$
    - $S_k[t] = \Sigma_{a_0} + \phi_k[t] \cdot \mathbf{P}_a[i] \cdot \phi_k[t]$
  - Keep final values $\hat{\theta}_i[N] := \hat{\theta}_i$, and $\mathbf{P}_a[N] := \mathbf{P}_a$ for the next step.
- **Maximization step:** Calculate hyperparameter update $\mathcal{P}_{i+1} = \{W_{i+1}, \Sigma_{a_{i+1}}, \Sigma_{b_{i+1}}\}$ as:
  - $W_{i+1} = \sum_{k=1}^{K} \hat{\theta}_i[1] \cdot f'(\xi_k)$
  - $\Sigma_{a_{i+1}} = \frac{1}{KN} \sum_{k=1}^{K} \sum_{t=1}^{N} (\mathbf{e}_k[t] \cdot \mathbf{e}_k'[t] + \phi_k[t] \cdot \mathbf{P}_a[i] \cdot \phi_k[t])$
  - $\Sigma_{b_{i+1}} = \frac{1}{K} \sum_{k=1}^{K} \sum_{t=1}^{N} (\hat{\theta}_i[t] \cdot \mathbf{P}_a[i] + \mathbf{P}_a[i])$

where $\hat{\theta}_i = \hat{\theta}_i - \mathbf{W} f(\xi_k)$, and $\mathbf{e}_k[t] = y_k[t] - \phi_k[t] \cdot \hat{\theta}_i$.

- **Convergence check:** Stop iteration if $i = N_{iter}$ or if:
  - $|Q(\mathcal{P}) - Q(\mathcal{P}_{i-1})| \leq \rho_Q$
  - $||\mathcal{P}_i - \mathcal{P}_{i-1}|| \leq \rho_p$

Otherwise, make $i = i + 1$ and return to **Expectation step**.

In addition, the Power Spectral Density (PSD) matrix of the vibration response under condi-
tions $\xi$ can be calculated as follows:

$$
P_{yy}(\omega, \xi) := 
\begin{bmatrix}
P_{y_1,y_1}(\omega, \xi) & P_{y_1,y_2}(\omega, \xi) & \cdots & P_{y_1,y_n}(\omega, \xi) \\
P_{y_2,y_1}(\omega, \xi) & P_{y_2,y_2}(\omega, \xi) & \cdots & P_{y_2,y_n}(\omega, \xi) \\
\vdots & \vdots & \ddots & \vdots \\
P_{y_n,y_1}(\omega, \xi) & P_{y_n,y_2}(\omega, \xi) & \cdots & P_{y_n,y_n}(\omega, \xi)
\end{bmatrix}
$$

$$
= D^{-T}(\omega, \xi) \cdot \Sigma(\xi) \cdot D^{-1}(\omega, \xi)
$$

(26)

where $j$ indicates the imaginary unit, $T_s = 1/f_s$ the sampling period, and $\omega$ frequency in radians per time unit.

6. Application example

6.1. Data description

The data used in this application example corresponds to the simulated vibration response of a 40 m aluminum wind turbine blade in the flap-wise direction, as thoroughly explained in [27]. The simulation is based on a numerical model of the blade composed by 4 Euler-Bernoulli beam elements with proportional damping in a cantilever configuration. Lift and drag forces on the blade are calculated from wind field simulations obtained from the TurbSim turbulent wind simulation package [33]. The Runge-Kutta method is used to integrate the aeroelastic model on a period of 600 seconds (10 minutes) with a sampling rate of 40 Hz, for a given 10 minute wind field time series and corresponding temperature. Uncertainty in the vibration response of the blade is introduced by means of variable temperature and wind speed. A whole year of temperature and wind speed variations are simulated according to average values measured in north-central Switzerland, provided by the Federal Office of Meteorology and Climatology and The Swiss Wind Power Data Website. The wind speed and temperatures used in the simulation are shown in Figure 2. Temperature linearly modifies the modulus of elasticity of the blade according to constants calculated for aluminum for the range of temperatures from -50 to 50 Celsius. From each wind speed and temperature pair, a single 10 minute simulation is created every hour, making a total of 8760 realizations. Four vertical accelerations measured at the DOFs $q_i$, $i = 1, 3, 5, 7$ (as in Fig. 1) re-sampled at 10 Hz are considered in the further analysis. The simulations performed in [27] include also extreme weather and damage scenarios. However, the present analysis is limited only to the data obtained from the healthy blade on the year-long normal weather simulation.

6.2. Identification of the vibration response in the healthy state with GP-VAR models

6.2.1. Determination of the model structure

Selection of the structural parameters of the GP-VAR model is carried out according to the guidelines provided in [26, 27]. This involves the determination of the VAR model order $n_a$ and the GP functional expansion order $p$. The selection of these quantities is described next.
Figure 1: Schematic representation of the blade: (a) Front/lateral view of the blade of length $L$, with distributed aerodynamic forces $f_a(t, x)$ and wind speed $v_w(t)$; (b) FEM of the blade with the deflection and rotation DOFs based on 4 Euler-Bernoulli beam elements, each of length $L/4$.

Figure 2: Yearly temperature and wind speed fluctuations used as environmental inputs in the blade vibration simulation.

VAR model order. In an initial stage, the selection of the VAR model order $n_a$ follows the traditional system identification procedure, where selection is based on the Residual Sum of Squares over the Series Sum of Squares (RSS/SSS), the Bayesian Information Criterion (BIC) and the frequency stabilization plot on a single 10 minute trial. For this purpose, VAR models in the range of orders $n_a = 1, \ldots, 30$ are inspected. The resulting RSS/SSS and BIC curves are displayed in the top row of Fig. 3. The RSS/SSS indicates increasing modeling accuracy for increasing model order, which tends to stabilize about order $n_a = 14$. On the other hand, the BIC reveals a global minimum at $n_a = 14$. The frequency stabilization plot, shown in the bottom part of Fig. 3 confirms the previous outcomes, demonstrating that the natural frequencies (those with Modal Assurance Criterion (MAC) over 0.9999) tend to stabilize from order $n_a = 10$ and beyond. Accordingly, hereby the autoregressive order of the GP-VAR model is selected as $n_a = 14$.

GP functional expansion order. Subsequently, the order of the functional expansion used for the GPR of the VAR model parameters is performed. To this end, ML estimates of the parameters and innovations covariance of VAR models of the selected order $n_a = 14$ are calculated on the entire set of 8 760 trials. For each trial, estimates of the parameter vector and covariance matrix
Figure 3: Selection of the VAR model order for a single trial of the blade vibration response. Top, RSS/SSS and BIC curves obtained for VAR models with orders in the range \( n_a = 5, \ldots, 30 \). Bottom, frequency stabilization plot. Colored areas indicate stable modes with \( MAC \geq 0.9999 \). Color shade indicates the damping ratio. Gray crosses indicate the remaining spurious modes.

At each trial \( k \) are obtained from the following optimization problem:

\[
\hat{\theta}_k, \hat{\Sigma}_\varepsilon_k = \arg \max_{\theta, \Sigma_{\varepsilon}} \ln L(\theta | Y_k)
\]

\[
\ln L(\theta | Y_k) = -\frac{1}{2} \sum_{i=1}^{N} \left( \ln 2\pi |\Sigma_{\varepsilon}| + \left( y_i[t] - \phi^T [t] \theta \right)^T \Sigma_{\varepsilon}^{-1} \left( y_i[t] - \phi^T [t] \theta \right) \right)
\]

Then, the complete set of trials is split into a training set, corresponding to the first six months of data, while the data from the remaining months is left for validation. In the training set, a GPR is calculated to approximate the functional relationship from temperature and mean wind speed to the VAR model parameters at each trial. For this purpose, a tensor product multivariate Hermite basis is constructed as follows:

\[
f(\xi) = h_1(\xi_1) \otimes h_2(\xi_2) \quad h_i(\xi_i) = \left[ H_{i,0}(\xi_i) \quad H_{i,1}(\xi_i) \quad \ldots \quad H_{i,p_i-1}(\xi_i) \right]^T, \quad \forall i = 1, \ldots, q
\]

where \( q = 2 \) is the number of EOPs (\( \xi_1 \): temperature; \( \xi_2 \): wind speed), \( p_i \) is the order of the univariate basis for the \( i \)-th EOP, \( \otimes \) denotes the Kronecker product, and \( H_{i,j}(\xi_i) \), \( j = 1, \ldots, p_i \)
stands for the Hermite polynomial of order \( j \), defined as follows:

\[
H_{i,j+1}(\xi_i) = \xi_i H_{i,j}(\xi_i) - jH_{i,j-1}(\xi_i) \quad H_{i,0}(\xi_i) = 1 \quad H_{i,1}(\xi_i) = 2\xi_i
\]

The RSS/SSS is calculated in both training and validation sets for a maximum univariate functional expansion order \( p_i = 5 \) for \( i = 1, 2 \). In addition, the Bayes factor defined in Eq. (10) based upon the likelihood of the estimated parameters given GPRs with different functional expansion basis, are calculated. The obtained Bayes factor and RSS/SSS calculated on the validation set are displayed in Fig. 4. The results demonstrate that orders in the range from \( p_i = 2, 3 \) are preferable. In the present case, knowing that the influence of the temperature on the blade dynamics is linear, a linear relation (\( p_1 = 2 \)) is subsequently selected for this variable. On the other hand, the influence of the wind speed is not well known, then three basis orders are to be considered for this variable, namely \( p_2 = (2, 3, 4) \). The selection of the definitive functional expansion order shall be demonstrated on the fully optimized GP-VAR model in the next section.

Figure 4: Selection of the GPR functional expansion order for the VAR model parameters. Left plot, Bayes ratio; right plot, validation RSS/SSS of the parameter vector.

6.2.2. Construction of the GP-VAR model for the complete blade vibration dataset

The EM algorithm is used to estimate the hyperparameters of GP-VAR models with structural parameters \( n_a = 14, p_1 = 2 \) and \( p_2 = [2, 3, 4] \), as suggested by the results presented on the previous section. Likewise, initial hyperparameter values are calculated from sample averages obtained from the ML estimates of the parameter vectors and innovations covariances. For each model structure, the EM algorithm is run for a total of 100 iterations or until the thresholds \( \rho_Q = 10^{-4} \) and \( \rho_P = 10^{-6} \) (defined in Table 1) are met. The prior and posterior RSS/SSS\(^2\), as well as the marginal likelihoods calculated on the training and validation sets of each one of the three optimized GP-VAR models are shown in Fig. 5. The results demonstrate that the prior RSS/SSS is approximately the same for all the models, while the posterior RSS/SSS tends to increase with increasing model structure. Contrariwise, the marginal likelihood demonstrates a preference towards the models with \( p_2 = (3, 4) \). The difference from the prior to the posterior RSS/SSS is relatively small (\( \leq 4\% \)), which demonstrates that the prior model provides an accurate estimation of the vibration response of the blade. In addition, the difference on the performance obtained

\^2\)The prior RSS/SSS is the RSS/SSS obtained from the prediction error \( e[t] = y[t] - \theta^T \Phi[t] \), with \( \hat{\theta}_k = \mathcal{W}(\xi_k) \). The posterior RSS/SSS is the RSS/SSS obtained from the posterior error \( \tilde{e}[t] = y[t] - \Phi[t] \tilde{\theta}_k \), where \( \tilde{\theta}_k \) is the maximum a posteriori (MAP) estimate of the parameter vector, corresponding to the mean of the distribution \( p(\theta | Y_k, \xi_k, \tilde{\theta}_k) \).
in training and validation sets is not dramatic, which suggests that the models do not overfit
the training data. Following the marginal likelihood criteria, the GP-VAR model with \( n_a = 14, \)
\( p_1 = 2 \) and \( p_2 = 3 \) is selected.

![Figure 5: Final performance of the optimized GP-VAR models with structural parameters \( n_a = 14 \) and \( p = [p_1, p_2] \), with \( p_1 = 2 \) and \( p_2 = [2, 3, 4] \), in terms of average of the prior RSS/SSS (left), posterior RSS/SSS (middle) and marginal likelihood (right) calculated on the training and validation sets (each with 4380 trials).](image)

6.3. Modal analysis based on the optimized GP-VAR models

The optimized GP-VAR models can now be used to calculate dynamic characteristics of the
blade vibration. To start with, Fig. 6 displays the first four modes of the blade as a function
of time obtained from the prior parameter estimates of the optimized GP-VAR models with
structural parameters \( n_a = 14, p_1 = 2, \) and \( p_2 = [2, 3, 4] \) evaluated on the temperature and
wind speeds used in the validation set. The obtained natural frequency trajectories are contrasted
with those extracted from the Covariance-driven Stochastic Subspace Identification (SSI-COV)
method. The procedure involves the extraction of stable modes in models with orders in the range
[10, 30] with the clustering approach presented in [30]. Although small differences are noticeable
on the natural frequencies extracted from the different GP-VAR models, the trajectories mostly
coincide on the prediction of the variation of the natural frequencies. In contrast, there is a
significant increment on the variability of the natural frequency estimates provided by the SSI-COV
method at each trial, while a notorious bias is found in the first natural frequency. This
indicates that, in addition to providing an accurate global model for the vibration response of the
blade, a carefully built GP-VAR model also helps to interpolate and remove unwanted variability
on the modal estimates.

Fig. 7 provides boxplots indicating the distribution of the MAC calculated between the refer-
ence mode shapes \( \psi_r(\xi) \), which are calculated based on the original FEM used to simulate the
blade, and the model-based mode shapes \( \psi_\xi(\xi) \) obtained with the optimized GP-VAR models and
with the SSI-COV method at each individual trial. As in the case of the natural frequencies, it
is evident that the mode shape estimates obtained from the GP-VAR models are more consistent
than those obtained with the SSI-COV method at individual realizations. In general, the mode
shape estimates tend to be more consistent on the first two modes, while the MAC of the third and
fourth modes is reduced in comparison. However, in all cases, the obtained MACs are always
over 0.99, which indicates at least a fair correspondence between the reference and estimated mode shapes.

Fig. 8 displays the expected value and 90% confidence intervals of the first four natural frequencies and their respective damping ratios as a function of temperature in the range \([-10, 30\) C with wind speed fixed at 12 m/s. Similarly, Fig. 9 shows the same quantities for the case of changing wind speed in the range \([2, 16]\) m/s with temperature fixed at 10 C. The expected values are calculated from the prior parameter matrices of the three obtained GP-VAR models with \(n_d = 14, p_1 = 2\) and \(p_2 = [2,3,4]\). The confidence intervals are calculated from 1000 random parameter vectors obtained after a Monte Carlo sampling of the prior parameter probability of the respective GP-VAR models. In the case of changing temperature, the three optimized GP-VAR models provide very consistent predictions of both the expected values of natural frequencies.
Figure 7: Distribution of the modal assurance criterion (MAC) computed between the reference (FEM-based) mode shapes $\psi_i^{(r)}$ and the model-based mode shapes $\psi_i^{(\xi)}$ obtained from the optimized GP-VAR models and from the SSI-COV method at each individual trial. MAC has been rescaled on a logarithmic scale to facilitate comparison. GP-VAR M1: GP-VAR model with $p_1 = 2, p_2 = 2$; GP-VAR M2: GP-VAR model with $p_1 = 2, p_2 = 3$; GP-VAR M3: GP-VAR model with $p_1 = 2, p_2 = 4$; VAR: VAR models from individual trials.

and damping ratios. On the other hand, in the case of changing wind speed, the mean natural frequencies and their corresponding confidence intervals are more consistent in the middle values (5 to 14 m/s), while the curves start to spread out around the edges. However, the actual change on the frequency values is negligible. Otherwise, the respective expected values and confidence intervals of the damping ratios are less consistent among the evaluated GP-VAR models, where inconsistencies are again more notorious around the edges. Although this effect may be due to the type of functional basis used in the GP-VAR models, significant variations on the damping ratio are expected with the wind speed due to the aeroelastic damping phenomenon. Following the fit criteria displayed in Fig. 5, the most plausible curves are those of the GP-VAR model with $p_1 = 2$ and $p_2 = 3$. A further study considering other types of functional basis could confirm these findings.

Lastly, Fig. 10 provides a similar analysis of the expected PSD of the vibration on the third sensor as a function of temperature and as a function of wind speed, with the other quantity fixed at a constant value, based on the GP-VAR model with $p_1 = 2$ and $p_2 = 3$. In this case, in addition to the changes on the natural frequencies and damping ratios observed in Figs. 8 and
9, it is clear that there are also changes in the vibration power, which are effectively represented via the EOP-dependent innovations covariance matrix. Particularly, in Fig. 10(b) is observed the substantial effect of the wind speed on the vibration power of the blade.

Figure 8: GP-V AR model-based prediction of the mean and 90% confidence intervals of the first four natural frequencies and damping ratios of the blade vibration as a function of temperature in the range [-10, 30] C with wind speed fixed at 12 m/s. Confidence intervals displayed as shaded areas are calculated from a 1000 sample Monte Carlo simulation of the parameter vectors of GP-V AR models at each temperature.

7. Conclusions

This work has been devoted to the generalization of Gaussian Process time-series models to the vector vibration response case, in the form of Vector AutoRegressive (VAR) models. In particular, the issue of global identification of the model, including an innovations covariance
Figure 9: GP-VAR model-based prediction of the mean and 90% confidence intervals of the first four natural frequencies and damping ratios of the blade vibration as a function of average wind speed (AWS) in the range [2, 16] m/s with temperature fixed at 10 °C. Confidence intervals displayed as shaded areas are calculated from a 1000 sample Monte Carlo simulation of the parameter vectors of GP-VAR models at each wind speed.

Different model structures where contrasted, demonstrating the consistency of the estimates, at least in the most part of the operational range. In addition, modal quantities derived from matrix dependent on the EOPs, has been tackled by means of an Expectation-Maximization (EM) algorithm. Moreover, the computationally expensive calculations involved in the expectation step have been relaxed by the application of a recursive Bayesian updating scheme on the mean and covariance of the posterior density of the GP-VAR model parameters. The GP-VAR modeling approach has been demonstrated in a simulation example consisting on the simulated vector vibration response of a wind turbine blade, under variability of both the acting wind speed and ambient temperature.
Figure 10: Expected value of the power spectral density of the blade vibration on sensor No. 3 based on the optimized GP-VAR model with \( n_a = 14, p_1 = 2, \) and \( p_2 = 3, \) as a function of: (a) Temperature in the range \([-10, 30]\) C with wind speed fixed at 12 m/s; (b) Wind speed in the range \([2, 20]\) m/s with temperature fixed at 10 C.

the obtained GP-VAR models were contrasted with those obtained from the SSI-COV method, which yields independent estimates at individual trials. As expected, the uncertainty on the modal parameter estimates was drastically reduced on the GP-VAR modeling approach.

This work is an additional contribution to the previous ones discussing the general setting of the GP time-series models [26], and the application of linear transformation to reduce model complexity in the case of GP-VAR models [29]. Moreover, GP time-series models can be used in the context of damage detection, as discussed in [27].

As a final thought, it is the authors feeling that the GP time-series modeling approach and the like (v.g.r. those based on Functionally Pooled and Polynomial Chaos Expansions), bear an immense potential for the effective response analysis and health monitoring of real-life structures operating in the field, required for a wide acceptance of these methods by industry.

Acknowledgements

Prof. E.N. Chatzi gratefully acknowledges the support of the ERC Starting Grant WINDMIL (ERC-2015-StG 679843) on the topic of “Smart Monitoring, Inspection and Life-Cycle Assessment of Wind Turbines”, as well as the ERC Proof of Concept (PoC) Grant, ERC-2018-PoC WINDMIL RT-DT on “An autonomous Real-Time Decision Tree framework for monitoring and diagnostics on wind turbines”.

References
