Modelling Long-Term Vibration Monitoring Data with Gaussian Process Time-Series Models

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

Published in:
IFAC-PapersOnLine

DOI:
10.1016/j.ifacol.2019.12.343

Publication date:
2019

Document version:
Final published version

Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 52-28 (2019) 26–31

Modelling long-term vibration monitoring data with Gaussian Process time-series models ⋆

Luis David Avendaño-Valencia * **Eleni N. Chatzi** **

* *Machine Learning and AI Group, Maersk-McKinney-Moller Institute, University of Southern Denmark. Campusvej 55, 5230 Odense M, Denmark. (e-mail: ldav@mmmi.sdu.dk).*

** *Institute of Structural Engineering, Dept. of Civil, Environmental and Geomatic Engineering. ETH Zürich. Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland, (e-mail: chatzi@ibk.baug.ethz.ch)*

Abstract: Gaussian Process (GP) time-series models are a special type of models for Linear Parameter Varying (LPV) systems in which the parameters are represented as stochastic variables following a Gaussian Process regression of the scheduling variables. GP time-series models are ideal for the representation of LPV systems where some of the scheduling variables are uncertain or immeasurable, as is the case in most real-life Structural Health Monitoring (SHM) applications. In this work, a fully parametric version of GP is adopted, most suitable for identification based on large datasets typically originated in SHM campaigns. Here, the model identification problem is addressed via global and local approaches, while is demonstrated that the latter case corresponds to a sub-optimal version of the global optimization. Finally, the GP time-series modelling methodology is demonstrated on the identification of the simulated vibration response of a wind turbine blade, where temperature and wind speed act as scheduling parameters.

Keywords: Uncertain Systems; Parametric Methods; System Identification.

1. INTRODUCTION

Vibration-based Structural Health Monitoring (SHM) refers to a group of methods aiming at damage assessment of structural systems, on the basis of the characteristics of their vibration response (Sohn, 2007). Practical application of SHM –hereafter used in the specific context of vibration-based SHM– comes with several challenges (Alampalli, 2000; Sohn, 2007), from which the following two are of main interest in this work: (*i*) the variability on the structure's dynamic characteristics stemming from changing environmental and operational conditions, and (*ii*) the enormous volumes of monitoring data acquired, which includes vibration and Environmental and Operational Parameters –EOPs–. Within the SHM framework, uncertainty on the vibration response has been treated via statistical methods aiming at isolating the effects of EOPs from those of damage on characteristic quantities. Examples include linear and non-linear transformations (Yan et al., 2005; Deraemaeker et al., 2008), or cointegration methods (Cross and Worden, 2012). An alternative –and perhaps more robust– approach consists on accepting the variable nature of the structural dynamics, which implies adopting a Linear Parameter Varying (LPV) structure for the dynamics, where EOPs take the place of scheduling parameters of the dynamic system (Kopsaftopoulos et al., 2018). Hence, the aim is to construct a global model capable of representing the variability of the structural dynamics in terms of measurable EOPs.

Traditional LPV system identification techniques have been used for this purpose, hinged on functional basis expansions of the parameters of the LPV model, where a deterministic relation from the EOPs (scheduling variables) to the LPV parameters is enforced (Tóth, 2008; Sakellariou and Fassois, 2016; Kopsaftopoulos et al., 2018). These methods can capture the main trend on the variation of the dynamics due to EOPs, but fall short on quantifying the uncertainty arising from stochastic scheduling parameters (EOPs) and unmeasurable variables. When input parameters are stochastic, Polynomial Chaos Expansions (PCE) facilitate the quantification of the uncertainty propagating from stochastic inputs to the variables of interest (Spiridonakos et al., 2016). However, a complete treatment of the uncertainty coming from stochastic inputs and missing variables can only be achieved with a fully stochastic treatment of the LPV parameters, as provided for instance by Gaussian Processes (GP) (Golabi et al., 2017; Darwish et al., 2018). Nonetheless, GPs and related kernel-based system identification methods have been initially introduced for a different purpose: facilitate the identification process by casting the function to be approximated –*v.gr.* the parameter of an LPV system– into an infinite dimension functional space determined by a kernel function (Kocijan et al., 2005; Golabi et al., 2017; Pillonetto et al., 2014; Darwish et al., 2015; Svensson

⋆ Prof. E.N. Chatzi gratefully acknowledges the support of the ERC Starting Grant WINDMIL (ERC-2015-StG 679843) on *"Smart Monitoring, Inspection and Life-Cycle Assessment of Wind Turbines"*.

10.1016/j.ifacol.2019.12.343

and Schön, 2017). Then, the tedious process of functional basis selection is replaced by the selection of the best function subset on the basis of the available data. However, the main limitation of non-parametric GP-based system identification methods is their computational burden (Bijl et al., 2017). This characteristic cripples them for application on large data-sets, as in SHM campaigns, which can easily reach millions of data samples. The application of sparse non-parametric GP methods has been proposed as a solution to alleviate the computational complexity of GP-based system identification methods (Bijl et al., 2017), while parameterized versions of GP system identification approaches have been investigated in (Svensson and Schön, 2017) and in previous works of the authors (Avendaño-Valencia et al., 2017; Avendaño-Valencia and Chatzi, 2017) in the specific context of SHM. These *fully-parametric* GP time-series models have a sharply reduced computational cost in contrast to their non-parametric counterparts and are amenable to large datasets, while keeping much of the modelling flexibility and the benefits of the inference framework of GPs. Nonetheless, the reappearance of the selection of the functional basis into the identification loop may be considered as a minor downside.

Following previous works of the authors, this work further develops a fully parametric approach for GP-based LPV system identification. Two identification methods are discussed in this work. The first based on the estimation of locally "frozen" models which are subsequently interpolated via GP regression to obtain a global model of the dynamics under variable EOPs –the local approach–. The second aiming at obtaining a global model that accommodates the complete training data set, which is attempted by means of an Expectation-Maximization (EM) algorithm – the global approach–. Accordingly, the main contribution of this work consists on providing a general comparison of the identification process on GP time-series models, including a numerical evaluation of the global and local identification schemes on a simulation example.

This work is organized as follows: Section 2 provides an overview of the Gaussian Process time-series modelling methodology, including the main definitions, a description of the identification process, and a discussion on the selection of scheduling parameters. Section 3 provides an application example involving the identification of vibration response of a wind turbine blade. Finally, Section 4 provides the final conclusions and outlook of this work.

2. OVERVIEW OF GAUSSIAN PROCESS TIME-SERIES MODELS

2.1 Main definitions

A *fully parametric* GP time-series model of the signal $y[t] \in \mathbb{R}$ is defined as (Avendaño-Valencia et al., 2017):

$$y[t] = \boldsymbol{\phi}^T[t] \cdot \boldsymbol{\theta} + \boldsymbol{\varepsilon}[t], \quad \boldsymbol{\varepsilon}[t] \sim \text{NID}(\mathbf{0}_{n\times 1}, \sigma_\varepsilon^2) \tag{1a}$$

$$\boldsymbol{\theta} = \boldsymbol{W} \cdot \boldsymbol{f}(\boldsymbol{\xi}) + \boldsymbol{u}, \quad \boldsymbol{u} \sim \mathcal{N}(\boldsymbol{u} \,|\, \mathbf{0}_{d\times 1}, \boldsymbol{\Sigma}_\theta) \tag{1b}$$

where $\boldsymbol{\phi}[t] \in \mathbb{R}^d$ is the regression vector, $\boldsymbol{\theta} := \boldsymbol{\theta}(\boldsymbol{\xi}) \in \mathbb{R}^d$ is the parameter vector, and $\varepsilon[t] \in \mathbb{R}$ is a zero-mean *Normally and Identically Distributed* (NID) innovations, with variance σ_ε^2. The parameter vector follows a parameterized form of a GP regression on the vector of scheduling parameters $\boldsymbol{\xi} \in \mathbb{R}^q$, determined by the coefficient matrix $\boldsymbol{W} = [\boldsymbol{w}_1\ \boldsymbol{w}_2\ \cdots\ \boldsymbol{w}_p]$, $\boldsymbol{W} \in \mathbb{R}^{d\times p}$, the functional basis $\boldsymbol{f}(\boldsymbol{\xi}) \in \mathbb{R}^p$, and the parameter covariance $\boldsymbol{\Sigma}_\theta \in \mathbb{R}^{d\times d}$.

The GP time-series model may also be represented in the kernelized form, as (Darwish et al., 2018):

$$y[t] = \boldsymbol{\phi}^T[t] \cdot \boldsymbol{\theta} + \boldsymbol{\varepsilon}[t], \quad \boldsymbol{\varepsilon}[t] \sim \text{NID}(0, \sigma_\varepsilon^2) \tag{2a}$$

$$\boldsymbol{\theta} \sim \mathcal{GP}(\mathbf{0}, \boldsymbol{K}(\boldsymbol{\xi}, \boldsymbol{\xi}_*) + \boldsymbol{\Sigma}_\theta) \tag{2b}$$

where $\mathcal{GP}(\cdot,\cdot)$ denotes a GP with the prescribed mean and covariance, and where $\boldsymbol{K}(\boldsymbol{\xi}, \boldsymbol{\xi}_*) = \boldsymbol{W} \cdot \boldsymbol{f}(\boldsymbol{\xi}) \cdot \boldsymbol{f}^T(\boldsymbol{\xi}_*) \cdot \boldsymbol{W}^T$ denotes the covariance kernel. In the forthcoming analysis, a fully parametric form of the GP time-series model is preserved with the aim of improving the model compactness and, in turn, reducing the computational burden of the method.

In the context of structural health monitoring, the scheduling variable is often referred to as the vector of *Environmental and Operational Parameters* (EOPs). In the same context, it is generally assumed that EOPs remain constant over a period of analysis of length T, so that $\boldsymbol{\xi}[t] := \boldsymbol{\xi}$. However, for the sake of generality, such an assumption is broken on the following presentation of GP time-series models, while all the related methods remain unchanged, unless explicitly stated.

The GP time-series model is characterized by the conditional multivariate normal PDFs (Avendaño-Valencia et al., 2017):

$$p(y[t] \,|\, \boldsymbol{\theta}, \boldsymbol{\xi}, \mathcal{P}) = \mathcal{N}(y[t] \,|\, \boldsymbol{\phi}^T[t] \cdot \boldsymbol{\theta}, \sigma_\varepsilon^2) \tag{3a}$$

$$p(\boldsymbol{\theta} \,|\, \boldsymbol{\xi}, \mathcal{P}) = \mathcal{N}(\boldsymbol{\theta} \,|\, \boldsymbol{W} \cdot \boldsymbol{f}(\boldsymbol{\xi}), \boldsymbol{\Sigma}_\theta) \tag{3b}$$

where $\mathcal{P} \triangleq \{\boldsymbol{W}, \boldsymbol{\Sigma}_\theta, \sigma_\varepsilon^2\}$ denotes the set of *hyperparameters* of the GP time-series model, while $\mathcal{N}(\boldsymbol{x} \,|\, \boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$ indicates a (multivariate) normal distribution of the random variable $\boldsymbol{x}$ with mean $\boldsymbol{\mu}_x$ and covariance matrix $\boldsymbol{\Sigma}_x$.

The linear-in-the-parameters regressive form of Equation (1a) encompasses a large group of time-series models, which differ in the specific form of the regression and parameter vectors. For instance, an AR model is characterized by a regression and parameter vectors of the form (Ljung, 1999, Ch. 4):

$$\boldsymbol{\phi}[t] := [y[t-1]\ \ y[t-2]\ \cdots\ \ y[t-n_a]]^T \tag{4a}$$

$$\boldsymbol{\theta} := \boldsymbol{\theta}(\boldsymbol{\xi}) = [a_1(\boldsymbol{\xi})\ \ a_2(\boldsymbol{\xi})\ \cdots\ \ a_{n_a}(\boldsymbol{\xi})]^T \tag{4b}$$

where n_a is the order of the AR model. More complex model structures, including time-dependent, non-linear or vector ARMAX models, can also be embedded within a GP time-series model. For simplicity in the presentation, the remaining of this work is devoted to the AR model case, while extensions to more complex model types are discussed in (Avendaño-Valencia et al., 2017, 2018; Avendaño-Valencia and Chatzi, 2019).

2.2 The identification problem

On a monitoring campaign, a set of input-output data $\mathcal{D}$ is obtained following K individual experiments on the structure. At each experiment k, an excitation signal $X_k = \{x_k[1], x_k[2], \cdots, x_k[T]\}$ and scheduling trajectory $\boldsymbol{\xi}_k = \{\boldsymbol{\xi}_k[1], \boldsymbol{\xi}_k[2], \cdots, \boldsymbol{\xi}_k[T]\}$ give rise to a response signal $Y_k = \{y_k[1], y_k[2], \cdots, y_k[T]\}$, all measured on the period T with common sampling rate f_s. The data obtained

from experiment k is represented as $\mathcal{D}_k = \{X_k, \boldsymbol{\xi}_k, Y_k\}$, while the complete set of input-output data is noted as $\mathcal{D} = \{\mathcal{D}_1, \mathcal{D}_2, \cdots, \mathcal{D}_K\}$.

The identification problem consists on obtaining the model $\mathcal{M} = \{\mathcal{P}, \mathcal{S}\}$, determined by hyperparameters $\mathcal{P} := \{\boldsymbol{W}, \boldsymbol{\Sigma}_\theta, \sigma_\varepsilon^2\}$ and structure $\mathcal{S}$ that best fit the entire input-output data set $\mathcal{D}$. Fitness is defined in terms of the model's likelihood, so that the identification problem is posed in terms of the following optimization problem:

$$\{\hat{\mathcal{P}}, \hat{\mathcal{S}}\} = \arg\max_{\hat{\mathcal{P}}, \hat{\mathcal{S}}} \ln \mathcal{L}(\mathcal{P}|\mathcal{D}, \Theta) \tag{5a}$$

$$\ln \mathcal{L}(\mathcal{P}|\mathcal{D}, \Theta) = -\frac{K}{2} \ln |2\pi\, \boldsymbol{\Sigma}_\theta| - \frac{KT}{2} \ln 2\pi\, \sigma_\varepsilon^2 - \frac{1}{2} \sum_{k=1}^{K} \left(\boldsymbol{u}_k^T\, \boldsymbol{\Sigma}_\theta^{-1}\, \boldsymbol{u}_k + \frac{1}{2\sigma_\varepsilon^2} \sum_{t=1}^{T} \hat{\varepsilon}_k^2[t|t-1] \right) \tag{5b}$$

where $\Theta := \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \cdots, \boldsymbol{\theta}_K\}$ denotes the set of parameter realizations of each experiment, $\boldsymbol{u}_k := \boldsymbol{\theta}_k - \boldsymbol{W} \cdot \boldsymbol{f}(\boldsymbol{\xi}_k)$ is the vector of parameter innovations, and $\hat{\varepsilon}_k[t|t-1] := y_k[t] - \boldsymbol{\phi}_k^T[t] \cdot \boldsymbol{\theta}_k$ is the model's one-step-ahead prediction error evaluated at time t on experiment k. Given that the parameter set Θ is not available, direct optimization of Eq. (5b) is infeasible. Different alternatives exist to solve this problem, including the marginalized likelihood –or empirical Bayes– method, and the Expectation-Maximization –EM– algorithm (Gupta and Chen, 2011). The former is based upon the optimization of a marginalized likelihood with respect to the nuisance parameters Θ, while the second is based on the iterative updating of the hyperparameters based on estimates of the unknown parameters. In both cases, a single global model is obtained, which optimally represents the response dynamics of the structure on the operational range observed at the monitoring campaign. Application of these *global* optimization methods often requires considerable computational resources, so a reduced-complexity method may come handy. For that case, it is possible to calculate locally "frozen" models and interpolate the obtained parameters via GP regression, leading to a sub-optimal model in the sense of the likelihood in Eq. (5b). Such an approach is regarded as a *local* one (Tóth, 2008), in the sense that is hinged on the approximation of locally-stationary models. In the sequel, global and local methods for the optimization of the hyperparameters of parametric GP time-series models are further elaborated.

Global identification approach Global identification can be either performed by the marginal likelihood or EM methods. Although the marginal likelihood method is regularly used on the optimization of GP-based methods, its application on large datasets is unappealing, due to the computational resources (particularly memory) required on the calculation of the likelihood. On the other hand, the EM algorithm operates on data from single experiments, while information condensation is carried out on the maximization step, while updating the hyperparameter estimates. More explicitly, in the EM algorithm the unknown parameter set Θ is replaced by estimates obtained with an initial guess of the hyperparameter values –*Expectation step*–. Then, hyperparameters are updated on the basis of the newly obtained parameter estimates –*Maximization step*–. Both steps are iterated until convergence. The EM algorithm is fully described in (Avendaño-Valencia et al., 2017) for scalar-response GP time-series models, and extended to the vector-response case in (Avendaño-Valencia and Chatzi, 2019). Here, a compendium of the expectation and maximization steps is presented.

Expectation Step: The mean $\tilde{\boldsymbol{\theta}}_k$ and covariance matrix $\boldsymbol{P}_k$ of the posterior parameter probability $p(\boldsymbol{\theta}|D_k, \mathcal{P}_i) = \mathcal{N}(\boldsymbol{\theta}\,|\,\tilde{\boldsymbol{\theta}}_k, \boldsymbol{P}_{\theta_k})$ are computed. Such quantities are efficiently computed by means of the algorithm (Avendaño-Valencia and Chatzi, 2019):
Initialization:

$$\tilde{\boldsymbol{\theta}}_k[0] = \boldsymbol{W}_i\, \boldsymbol{f}(\boldsymbol{\xi}_k) \qquad \boldsymbol{P}_{\theta_k}[0] = \boldsymbol{\Sigma}_{\theta_i} \tag{6}$$

For $t = 1, \ldots, T$*:*

$$\tilde{\boldsymbol{\theta}}_k[t] = \tilde{\boldsymbol{\theta}}_k[t-1] + \boldsymbol{K} \cdot (\boldsymbol{y}_k[t] - \boldsymbol{\phi}_k^T[t] \cdot \hat{\boldsymbol{\theta}}_k[t-1]) \tag{7a}$$

$$\boldsymbol{P}_{\theta_k}[t] = (\boldsymbol{I}_d - \boldsymbol{K} \cdot \boldsymbol{\phi}_k^T[t]) \cdot \boldsymbol{P}_{\theta_k}[t-1] \tag{7b}$$

$$\boldsymbol{K} = \boldsymbol{P}_{\theta_k}[t-1] \cdot \boldsymbol{\phi}_k[t] \cdot s_e^{-2}[t] \tag{7c}$$

$$s_e^2[t] = \sigma_{\varepsilon_i}^2 + \boldsymbol{\phi}_k^T[t] \cdot \boldsymbol{P}_{\theta_k}[t] \cdot \boldsymbol{\phi}_k[t] \tag{7d}$$

where $\mathcal{P}_i = \{\boldsymbol{W}_i, \boldsymbol{\Sigma}_{\theta_i}, \boldsymbol{\Sigma}_{\varepsilon_i}\}$ denotes the hyperparameters of the current EM iteration.

The resulting values $\tilde{\boldsymbol{\theta}}_k[T] := \tilde{\boldsymbol{\theta}}_k$ and $\boldsymbol{P}_{\theta_k}[T] := \boldsymbol{P}_{\theta_k}$ correspond to the mean and covariance matrix of $p(\boldsymbol{\theta}\,|\,D_k, \mathcal{P}_i)$ to be used in the calculations of the maximization step.

Maximization Step: Updated hyper-parameter values $\mathcal{P}_{i+1}$ are obtained with the update equations (Avendaño-Valencia et al., 2017):

$$\boldsymbol{W}_{i+1} = \left(\sum_{k=1}^{K} \tilde{\boldsymbol{\theta}}_k\, \boldsymbol{f}^T(\boldsymbol{\xi}_k) \right) \cdot \left(\sum_{k=1}^{K} \boldsymbol{f}(\boldsymbol{\xi}_k)\, \boldsymbol{f}^T(\boldsymbol{\xi}_k) \right)^{-1} \tag{8a}$$

$$\boldsymbol{\Sigma}_{\theta_{i+1}} = \frac{1}{K} \sum_{k=1}^{K} \left(\boldsymbol{\delta}_k\, \boldsymbol{\delta}_k^T + \boldsymbol{P}_{\theta_k} \right) \tag{8b}$$

$$\sigma_{\varepsilon_{i+1}}^2 = \frac{1}{K\,T} \sum_{k=1}^{K} \sum_{t=1}^{T} \left(\tilde{e}^2[t|t-1] + \boldsymbol{\phi}_k^T[t]\, \boldsymbol{P}_{\theta_k}\, \boldsymbol{\phi}_k[t] \right) \tag{8c}$$

where $\boldsymbol{\delta}_k = (\tilde{\boldsymbol{\theta}}_k - \boldsymbol{W}_i\, \boldsymbol{f}(\boldsymbol{\xi}_k))$ and $\tilde{e}[t|t-1] = y_k[t] - \boldsymbol{\phi}^T\, \tilde{\boldsymbol{\theta}}_k$.

The expectation and maximization steps are iterated until a specific number of iterations, say N_{iter}, is reached, or until convergence, which may be assessed by evaluating if the norm of the difference between the current and previous values of the marginal likelihood and hyper-parameter estimates is lower than a pre-specified threshold.

Local identification approach Local identification consists on estimating locally–stationary models at different operating points dictated by the scheduling variable(s). The local models are then interpolated to obtain a model that approximates the global behaviour of the system on the observed range of operating conditions. The identification process is greatly simplified with the local approach, although the obtained model may be sub-optimal in the sense of the likelihood in Eq. (5b). The local identification follows two steps: (*i*) identification of locally-stationary models for individual trials with data D_k; (*ii*) estimation of a GP regression representing the parameter vectors of individual trials as a function of the scheduling variable. This

may only be achieved on the condition that the scheduling variable $\boldsymbol{\xi}_k$ remains constant at each experiment k.

The estimates of the parameter vectors of locally-stationary models $\hat{\boldsymbol{\theta}}_k$ are obtained by the maximum likelihood method based on the observation density in (3a), leading to the optimization problem (Avendaño-Valencia et al., 2017):

$$\{\hat{\boldsymbol{\theta}}_k, \hat{\sigma}^2_{\varepsilon_k}\} = \underset{\boldsymbol{\theta}\in\mathbb{R}^d, \sigma^2_{\varepsilon_k}\in\mathbb{R}^+}{\arg\max} \ln \mathcal{L}(\boldsymbol{\theta}, \sigma^2_{\varepsilon}|D_k) \tag{9a}$$

$$\ln \mathcal{L}(\boldsymbol{\theta}, \sigma^2_{\varepsilon}|D_k) = -\frac{T}{2}\ln 2\pi\sigma^2_{\varepsilon_k} - \frac{1}{2\sigma^2_{\varepsilon_k}}\sum_{t=1}^{T}\hat{\varepsilon}^2_k[t|t-1] \tag{9b}$$

$$\hat{\varepsilon}_k[t|t-1] = y_k[t] - \boldsymbol{\phi}^T_k[t]\cdot\boldsymbol{\theta} \tag{9c}$$

where $\ln \mathcal{L}(\boldsymbol{\theta}, \sigma^2_{\varepsilon}|D_k)$ is the log-likelihood of the locally-stationary model based on data D_k, $\hat{\varepsilon}_k[t|t-1]$ is the model's (one-step-ahead) prediction error with respective variance $\sigma^2_{\varepsilon_k}$. Given the linear-in-the-parameters nature of the model, the solution to the optimization problem in Equation (9a) translates into the calculation of a least-squares type of solution for the parameter vector and prediction error variance (see *v.gr.* (Ljung, 1999, Sec. 7.4)).

Subsequently, the obtained parameter vector estimates, collected in the set $\hat{\Theta} = \{\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2, \cdots, \hat{\boldsymbol{\theta}}_K\}$, are regressed on the respective set of (static) scheduling parameters Ξ. Following from (3b), maximum likelihood estimates of the GP coefficient matrix and parameter covariance are obtained as the solution to the optimization problem:

$$\{\hat{\boldsymbol{W}}, \hat{\boldsymbol{\Sigma}}_\theta\} = \underset{\boldsymbol{W}\in\mathbb{R}^{d\times p}, \boldsymbol{\Sigma}_\theta\in\mathbb{R}^{d\times d}}{\arg\max} \ln \mathcal{L}(\boldsymbol{W}, \boldsymbol{\Sigma}_\theta|\hat{\Theta}) \tag{10a}$$

$$\ln \mathcal{L}(\boldsymbol{W}, \boldsymbol{\Sigma}_\theta|\hat{\Theta}) = -\frac{1}{2}\sum_{k=1}^{K}\left(\ln|2\pi\,\boldsymbol{\Sigma}_\theta| + \hat{\boldsymbol{u}}^T_k\,\boldsymbol{\Sigma}^{-1}_\theta\,\hat{\boldsymbol{u}}_k\right) \tag{10b}$$

$$\hat{\boldsymbol{u}}_k = \hat{\boldsymbol{\theta}}_k - \boldsymbol{W}\cdot\boldsymbol{f}(\boldsymbol{\xi}_k) \tag{10c}$$

where $\hat{\boldsymbol{u}}_k$ is the GPR estimation error on the parameter vector (or *prior* GPR error estimate).

3. ILLUSTRATIVE EXAMPLE

3.1 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. A full description of the simulation procedure is provided in (Avendaño-Valencia and Chatzi, 2017). In short, the blade is simulated via 4 Euler-Bernoulli beam elements in a cantilever configuration, while aerodynamic loads are calculated considering a specific blade profile under turbulent wind excitation. Numerical integration is used to calculate the vibration response of the aeroelastic model on a period of 600 seconds, sampled at 40 Hz. At each simulation, the blade temperature, 10 minute average wind speed and turbulence intensity are kept constant. Thus, a realization of the wind speed time-series is generated based on the prescribed average wind speed and turbulence intensity with the software TurbSim (Jonkman, 2009), which is subsequently used as excitation to the blade. A whole year of temperature and average wind speed variations are simulated according to normal values measured in north-central Switzerland, provided by the *Federal Office of Meteorology and Climatology* and *The Swiss Wind Power Data Website*. A single 10 minute simulation is created for every hour, making a total of 8 760 realizations. The analysis presented here is based on the blade tip vibration response sampled at 10 Hz.

In the present configuration, both the wind speed and blade temperature modify the system dynamics. Temperature linearly modifies the modulus of elasticity of the blade according to constants calculated for aluminium on the temperatures range $[-50, 50]$ C, while, the wind speed determines the power of the vibration and modifies the damping properties of the blade. Note however, that only average temperature and wind speeds are available, while the instantaneous values of the wind speed are unavailable, thus increasing the uncertainty of the system's dynamics.

3.2 Identification of the blade response

A GP AutoRegressive (GP-AR) model is selected to represent the tip vibration of the blade, using the 10 minute average wind speed and temperature as scheduling variables. In this case, the GP-AR model is ideal for capturing the stochastic nature of the parameter-varying system structure stemming from the stochastic turbulent wind characteristics. Identification of the GP-AR model is hinged on the entire simulation set. Data from the first half of the year (4 380 records) is used for identification (training set), while the remaining half (4 380 records) is used for model validation (validation set). The order of the AR component of the model is selected by evaluation of the Bayesian information criterion and the frequency stabilization plot on a single record, which leads to an order $n_a = 41$.

Local identification Local identification is accomplished as described in Sec. 2.2. ML estimates of the parameter vectors of locally-stationary AR models are calculated on the complete data set, while the parameter vectors corresponding to the records in the training set are used to calculate the hyperparameters of the GP-AR model. Multivariate Hermite polynomials are selected as the GPR functional basis. The multivariate basis is constructed by Kronecker products of univariate basis on the average wind speed and temperature, according to the relation (Avendaño-Valencia et al., 2019)

$$\boldsymbol{f}(\boldsymbol{\xi}) = \boldsymbol{f}_1(\xi_1)\otimes\boldsymbol{f}_2(\xi_2) \tag{11}$$

where $\boldsymbol{\xi} = [\xi_1\ \xi_2]^T$ is the vector of scheduling parameters, and $\boldsymbol{f}_i(\cdot), i = 1, 2$ are univariate Hermite polynomials, each one of order p_i (Spiridonakos et al., 2016). The functional basis orders are selected as $p_1 = 2$ for the temperature (considering its linear influence on the dynamics), while for wind speed the values $p_2 = \{1, 2, 3, 4\}$ are evaluated.

Global identification Global identification is carried out by means of the EM algorithm, as explained in Sec. 2.2. Initial values of the hyperparameters are selected as those obtained with the local identification procedure. The EM algorithm is run for a maximum of 100 iterations on the training set, while is set to stop if the normalized change on hyperparameter updates is less than 10^{-3} or the normalized change on the marginal likelihood is less than 10^{-3}. This procedure is repeated for the four model structures considered on the local identification method, namely $n_a = 41$, $p_1 = 2$ and $p_2 = \{1, 2, 3, 4\}$.

Performance comparison Comparison of the GP-AR models obtained with the local and global approach is made in terms of the model's predictive ability in the validation set, and in terms of the accuracy of the prior parameter estimates provided by the model. The model's predictive ability is measured by the *Residual Sum of Squares over the Series Sum of Squares* (RSS/SSS), which is defined as follows:

$$RSS/SSS[\%] = \frac{\sum_{t=1}^{T} \hat{\varepsilon}^2[t|t-1]}{\sum_{t=1}^{T} y^2[t]} \times 100\%$$

where $\hat{\varepsilon}^2[t|t-1]$ is the prediction error based on the a-priori parameter estimates $\hat{\boldsymbol{\theta}}_o = \hat{\boldsymbol{W}}\,\boldsymbol{f}(\boldsymbol{\xi})$. Additionally, the accuracy of the prior parameter estimates is measured by calculating the difference with the posterior parameter estimates, in the *Mean Prior-to-Posterior Squared Difference* (MPPSD) defined as:

$$MPPSD[\%] = \frac{(\tilde{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_o)^T \cdot (\tilde{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_o)}{\hat{\boldsymbol{\theta}}_o^T \cdot \hat{\boldsymbol{\theta}}_o} \times 100\%$$

The first measure quantifies the prediction error of the model before calculating parameter estimates with data, while the second measure is a distance between the prior and posterior parameter estimates, thus evaluating the quality of the functional dependency approximation of the AR model parameters on the scheduling parameters.

Fig. 1 displays the RSS/SSS and MPPSD calculated on the validation set of GP-AR models obtained with the local and global identification approaches. The predictive ability of the different models appears to have some differences with different functional basis orders and on the different identification approaches, however, the difference is actually insignificant. On the other hand, the MPPSD demonstrates that after applying the global identification method, there is a significant increment on the representation accuracy of the GP regression on the model parameter estimates. Fig. 2 displays the first and second AR parameters of the GP-AR model obtained with the global and local approaches, as well as the distribution of the average wind speeds corresponding to the training set. It is observed that while the parameter estimates correspond well in largely populated areas, the outcomes start to disagree as the training data becomes scarce.

In summary, although there is no significant change in the predictive ability of the model when using the global approach, there is a substantial gain on the capability of the model to produce accurate estimates of the parameter vector at individual realizations of the process. This means that subsequent model-based analysis of the dynamics (*v.gr.* modal analysis under changing EOPs), is more reliable when using the global identification results. Nonetheless, the local identification approach provides a fairly good approximation of the AR model parameter functional dependency, which is useful for simple analyses and predictive models.

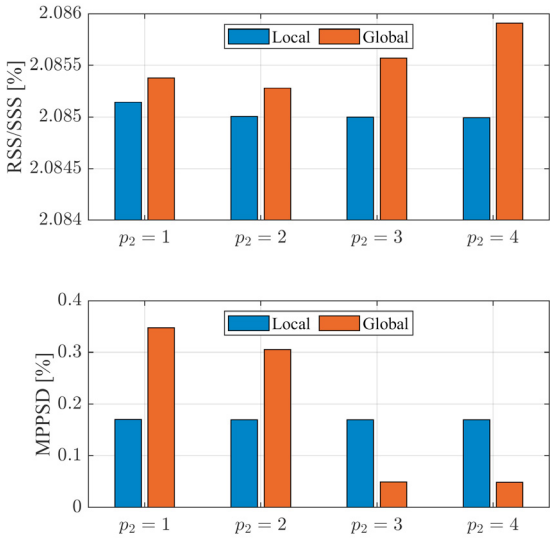

Fig. 1. Average RSS/SSS and MPPSD calculated on the validation set of GP-AR models with increasing functional basis order, obtained with the local and global identification methods.

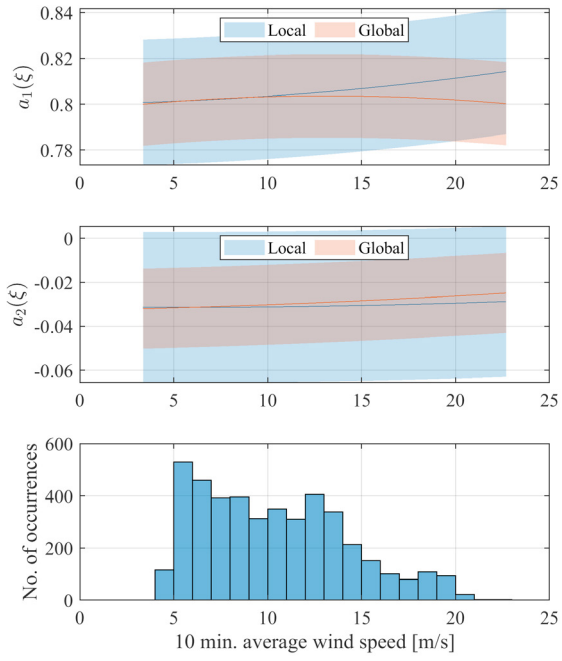

Fig. 2. Global and local estimates of the first and second parameters of the GP-AR model as a function of the wind speed, and histogram of the average wind speeds of the training set.

4. CONCLUSION

This work has been devoted to the modelling of uncertain LPV systems with GP time-series models. An overview of local and global methods for the identification of these models has been provided. The methods discussed in this work were illustrated on the identification of the vibration response of a simulated wind turbine blade, where the variation on the system dynamics stems from the ambient temperature, the average wind speed and turbulence. The application example demonstrated no major difference on the predictive ability of the GP-AR models obtained

with the local and global approach. However, the main differences were observed on the model's capability to provide accurate prior parameter vector estimates and in turn enhanced estimates of the dynamic characteristics of the system under changing scheduling parameters, which demonstrate the benefits of the global identification approach. On the other hand, damage detection methods based on the GP time-series model predictive accuracy are feasible based on locally identified models, which can be useful for SHM applications.

REFERENCES

Alampalli, S. (2000). Effects of testing, analysis, damage, and environment on modal parameters. *Mechanical Systems and Signal Processing*, 14(1), 63 – 74.

Avendaño-Valencia, L.D. and Chatzi, E.N. (2017). Sensitivity driven robust vibration-based damage diagnosis under uncertainty through hierarchical Bayes time-series representations. *Procedia Engineering*, 199, 1852 – 1857. X International Conference on Structural Dynamics, EURODYN 2017.

Avendaño-Valencia, L.D. and Chatzi, E.N. (2019). Multivariate GP-VAR models for vibration-based structural identification under varying operational conditions. *Under Review - Probabilistic Engineering Mechanics*, Special Issue of the 8th Computational Stochastic Mechanics Conference, 1–22.

Avendaño-Valencia, L.D., Chatzi, E.N., Koo, K.Y., and Brownjohn, J.M. (2017). Gaussian Process Time-Series Models for Structures under Operational Variability. *Frontiers in Built Environment*, 3, 69.

Avendaño-Valencia, L.D., Chatzi, E.N., and Tcherniak, D. (2019). Gaussian Process models for mitigation of operational variability in the structural health monitoring of wind turbines. *Under Review - Mechanical Systems and Signal Processing*, 1–34.

Avendaño-Valencia, L.D., Tatsis, K., and Chatzi, E.N. (2018). Gaussian process vector AR surrogates for identification of structures under varying operational conditions. In P. Spanos and G. Deodatis (eds.), *Proceedings of the 8th Conference on Computational Stochastic Mechanics*, 1–12.

Bijl, H., Schön, T.B., van Wingerden, J.W., and Verhaegen, M. (2017). System identification through online sparse Gaussian process regressin with input noise. *IFAC Journal of Systems and Control*, 2, 1–11.

Bishop, C.M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, Berlin, Heidelberg.

Cross, E.J. and Worden, K. (2012). Cointegration and why it works for shm. *Journal of Physics: Conference Series*, 382(1), 012046.

Darwish, M., Cox, P., Pillonetto, G., and Tóth, R. (2015). Bayesian identification of LPV Box-Jenkins models. In *2015 54th IEEE Conference on Decision and Control (CDC)*, 66–71.

Darwish, M.A.H., Cox, P.B., Proimadis, I., Pillonetto, G., and Tóth, R. (2018). Prediction-error identification of LPV systems: A nonparametric gaussian regression approach. *Automatica*, 97, 92 – 103.

Deraemaeker, A., Reynders, E., Roeck, G.D., and Kullaa, J. (2008). Vibration-based structural health monitoring using output-only measurements under changing environment. *Mechanical Systems and Signal Processing*, 22(1), 34 – 56.

Golabi, A., Meskin, N., Tóth, R., and Mohammadpour, J. (2017). A bayesian approach for LPV model identification and its application to complex processes. *IEEE Transactions on Control Systems Technology*, 25(6), 2160–2167.

Gupta, M.R. and Chen, Y. (2011). *Theory and Use of the EM Algorithm*. Now Foundations and Trends.

Jonkman, B.J. (2009). TurbSim User's Guide: Version 1.50. Technical report, National Renewable Energy Laboratory, U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Battelle, CO, U.S.A.

Kocijan, J., Girard, A., Banko, B., and Murray-Smith, R. (2005). Dynamic systems identification with gaussian processes. *Mathematical and Computer Modelling of Dynamical Systems*, 11(4), 411–424.

Kopsaftopoulos, F., Nardari, R., Li, Y.H., and Chang, F.K. (2018). A stochastic global identification framework for aerospace structures operating under varying flight states. *Mechanical Systems and Signal Processing*, 98, 425 – 447.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall PTR, 2nd edition.

Pillonetto, G. (2018). System identification using kernel-based regularization: New insights on stability and consistency issues. *Automatica*, 93, 321–332.

Pillonetto, G., Chiuso, A., and De Nicolao, G. (2011). Prediction error identification of linear systems: A nonparametric Gaussian regression approach. *Automatica*, 47, 291–305.

Pillonetto, G., Dinuzzo, F., Chen, T., De Nicolao, G., and Ljung, L. (2014). Kernel methods in system identification, machine learning and function estimation: A survey. *Automatica*, 50, 657–682.

Rizvi, S.Z., Velni, J.M., Abbasi, F., Tóth, R., and Meskin, N. (2018). State-space LPV model identification using kernelized machine learning. *Automatica*, 88, 38–47.

Sakellariou, J. and Fassois, S. (2016). Functionally pooled models for global identification of stochastic systems under different pseudostatic operating conds. *Mechanical Systems and Signal Processing*, 72-73, 785 – 807.

Sohn, H. (2007). Effects of environmental and operational variability on structural health monitoring. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 365(1851), 539–560.

Spiridonakos, M.D., Chatzi, E.N., and Sudret, B. (2016). Polynomial chaos expansion models for the monitoring of structures under operational variability. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 2(3), B4016003.

Svensson, A. and Schön, T.B. (2017). A flexible state-space model for learning nonlinear dynamical systems. *Automatica*, 80, 189–199.

Tóth, R. (2008). *Modeling and Identification of Linear Parameter Varying Systems an Orthonormal Basis Function Approach*. Ph.D. thesis, TU Delft.

Yan, A.M., Kerschen, G., Boe, P.D., and Golinval, J.C. (2005). Structural damage diagnosis under varying environmental conditions - Part I: A linear analysis. *Mechanical Systems and Signal Processing*, 19(4), 847 – 864.