

Article

A Data-Driven Diagnostic Framework for Wind Turbine Structures: A Holistic Approach

Simona Bogoevska ¹, Minas Spiridonakos ², Eleni Chatzi ^{2,*}, Elena Dumova-Jovanoska ¹ and Rudiger Höffer ³

¹ Faculty of Civil Engineering, University Ss. Cyril and Methodius, Skopje 1000, Macedonia; simona.bogoevska@gf.ukim.edu.mk (S.B.); dumova@gf.ukim.edu.mk (E.D.-J.)

² Department of Civil, Environmental and Geomatic Engineering, ETH, Zürich CH-8093, Switzerland; mspyridonakos@gmail.com

³ Department of Civil and Environmental Engineering, Ruhr-University Bochum, Bochum 44801, Germany; ruediger.hoeffler@ruhr-uni-bochum.de

* Correspondence: chatzi@ibk.baug.ethz.ch; Tel.: +41-44-633-6755

Academic Editor: Vittorio M. N. Passaro

Received: 31 January 2017; Accepted: 21 March 2017; Published: 30 March 2017

Abstract: The complex dynamics of operational wind turbine (WT) structures challenges the applicability of existing structural health monitoring (SHM) strategies for condition assessment. At the center of Europe's renewable energy strategic planning, WT systems call for implementation of strategies that may describe the WT behavior in its complete operational spectrum. The framework proposed in this paper relies on the symbiotic treatment of acting environmental/operational variables and the monitored vibration response of the structure. The approach aims at accurate simulation of the temporal variability characterizing the WT dynamics, and subsequently at the tracking of the evolution of this variability in a longer-term horizon. The bi-component analysis tool is applied on long-term data, collected as part of continuous monitoring campaigns on two actual operating WT structures located in different sites in Germany. The obtained data-driven structural models verify the potential of the proposed strategy for development of an automated SHM diagnostic tool.

Keywords: wind turbines; data-driven framework; uncertainty propagation; operational spectrum; time varying autoregressive moving average (TV-ARMA) models; polynomial chaos expansion (PCE)

1. Introduction

The complexity related to the interacting subsystems of WT structures (namely the rotating blades, moving yaw mechanism, and pitch angle changes) and the alternating aerodynamic loads redefining the operational regime, result in a complex vibrating system necessitating adoption of efficient monitoring and diagnostic methods. The understanding of the intricate behavior of operating WT systems has preoccupied the research community since as early as the 1990s, resulting in development of operational modal analysis (OMA) methods, when in 1988 output-only testing was utilized for the first time on the 110 m high Éole turbine [1]. Although directly related with the birth of the nowadays widely implemented OMA techniques, the assessment of WT infrastructure has still remained a multidisciplinary, rather challenging task. Moreover, with Europe's strategic planning currently focusing on renewable energy management, WTs are resurging into the spotlight for both the industrial and research community [2]. In this context, modern WT structures have to fulfil the growing demands for higher productivity and reduced downtime, which in turn calls for improved and automated SHM strategies, ensuring early-stage damage detection and structural diagnostics, reliability in power supply, as well as optimal operation and maintenance [3–6].

The difficulties related to assessment of operating WTs may be attributed among others sources to limited knowledge of the loading conditions (aeroelastic effects and rotor harmonics), the complexity of the multi-component WT system, varying operational regimes and environmental factors, as well as the typical uncertainties relating to incomplete and imperfect sensor data, modeling errors, complex and unique to the location soil-structure interaction effects.

Considering OMA methods applied on onshore operating WT structures, several studies based on short-term measurement campaigns successfully presented different concepts and novel measurement techniques [7–9]. In more recent works implementation of long-term strategies for tracking variations in identified modal parameters has been reported, thus shifting the focus towards automated SHM schemes and interlinked fatigue assessment strategies [10–15]. In [10] Devriendt et al. present an automated identification algorithm that combines an OMA tool (the poly-reference least squares complex frequency-domain estimator or covariance-driven stochastic subspace identification method) with a double clustering approach (hierarchical and fuzzy clustering algorithms). The proposed method is validated via successful tracking of minor changes in the dynamic behavior of an offshore wind turbine on data continuously collected during a two-week interval, for idling, i.e., parked, condition of the structure. Iliopoulos et al. [11] propose and validate a response estimation methodology for acceleration predictions at arbitrary points of an offshore WT structure via fusion of a finite element model (FEM) of the structure, estimated OMA-based modal parameters and a limited number of measured acceleration time histories. In the work of Maes et al. [12], a full-scale verification of three response estimation algorithms for the prediction of dynamic strains along the tower of an offshore monopile wind turbine is presented. The proposed hybrid filtering approaches combine a system model, derived from an updated FEM of the structure (based on OMA identification of the parked system), and a limited set of acceleration, and combined acceleration and strain measurements. The fatigue assessment strategy is verified for parked and rotating conditions of the structure. Male and Lourens in [13] illustrated a strategy for real-time monitoring of WT fatigue accumulation by means of a joint input-state estimator. The method comprises of measuring operational vibrations at selected locations along a WT tower and strains assessment at unmeasured locations on the WT lattice support structure. The method is verified on artificial signals generated via a full-order FEM, whereas input forces and states are estimated on erroneous design model.

A significant point to keep in mind however is that the OMA class of methods is developed for implementation with time invariant systems [16], while more refined schemes pertaining to non-stationary systems do exist but are less frequently applied onto civil systems [17,18]. However, the major challenge for delivering an efficient performance-based long-term framework lies precisely in the time varying nature of WT structures, linked to the changing operational regimes and varying environmental agents, as well as in the misinterpretation of this variability [19]. The latter may result in false alarms hindering effective operation of associated damage detection and control intervention. For addressing the aforementioned challenge research studies in this field are mainly based on two general approaches. The first approach relies on filtering out the influence of environmental factors from estimated performance indicators, and comprises models which employ measured environmental/operational variables [15,20] or models which are solely based on structural response estimates [21]. The alternative approach lies in incorporation of the measured environmental variables into models of measured vibration response in the form of extracted performance indicators, e.g., modal parameters [22–26].

Related to the first concept of eliminating benign variations from estimated structural features, several works dedicated to operating wind turbine (OWT) structures have recently emerged. In order to compensate for the environmental variability, Weijtjens et al. in [15] subtracted a (non)-linear regression model from estimated resonant frequencies of an operating offshore WT. However, the operational variability of the WT dynamics is tackled via adoption of a case-by-case modal tracking SHM algorithm. A similar strategy of fitting multivariate linear and dynamic regression models is applied in Oliveira et al. [20] on a 2MW onshore WT. In the recent work of Hu et al. [21], the authors

extracted a structural health index for a prototype of an operating 5MW WT, by removing temperature effects from selected natural frequency estimates based on a principal component analysis method.

Novel strategies integrating both structural response data and influencing agents within mathematical models have proven successful in several recent studies. Addressing the non-stationarity in measured response, Spiridonakos et al. in [22] employed time varying autoregressive models to capture the short-term dynamics, and a polynomial chaos expansion (PCE) tool for capturing the influence of operational and environmental agents, in tracking the performance of an actual operating WT tower located in Lübbenau, Germany. The use of PCE for constructing metamodels of structural response that are adept in incorporating external influences in a long-term scale was initially introduced and successfully applied for the purpose of damage detection of the benchmark SHM project of the Z24-bridge in Switzerland by Spiridonakos and Chatzi in [23]. In a preliminary investigation, Bogoevska et al. in [24,25] examined the capacity of this framework for effective reduction of the input set dimension facilitating an implementation of the proposed approach in an automated fashion. It is also worth mentioning that a similar approach, this time relying on random coefficient (RC) linear parameter varying (LPV) AR models has appeared in the work of Avendano et al. in [18,26]. In [26] the RC-LPV-AR methodology is assessed on simulations of an operating wind turbine blade under different damage scenarios. The study confirms the potential of data-driven methodologies in the accurate modeling of the non-stationary and uncertain vibration response of operating WT blades, albeit exclusively demonstrated on simulation data. Indeed, automated strategies have so far mostly been demonstrated in either a laboratory setting or on numerical simulations [27–29].

The research study presented herein employs actual field data, from operating turbines, in order to extend the PCE-AR tool, developed by the authoring team [22–25], highlighting two main aspects: the robustness of the strategy, as well as the high potential of the proposed method for automated structural health assessment. The first aspect is demonstrated via implementation on two separate operating WT structures located in Germany for an extended period of monitored structural responses and larger set of influencing agents. This is accomplished by a bi-component tool, which combines the parametric smoothness priors time varying autoregressive moving average (SP-TARMA) method, for identifying structural performance indicators (short-term framework), with a PCE probabilistic model, for quantifying the uncertainty in the identified structural performance indicators (long-term framework). Towards the development of an autonomous diagnostic tool, capable of tracking and diagnosing structural condition during the complete WT life-cycle, a twelve-month tracking of an estimated PCE-SPTARMA diagnostic index is further demonstrated. In this context, as a preliminary step, the correlation between alerting values of the obtained diagnostic index and novel fluctuations in measured operating conditions is verified by means of outlier analysis.

The proposed combined PCE-SPTARMA approach, in contrast to strategies incorporating commonly applied OMA methods, intrinsically captures the time variability of the system dynamics and delivers a holistic model, alleviating mode-by-mode or case-by-case modeling, for long-term tracking of the structural behavior. Furthermore, the results of the presented study demonstrate the effectiveness and high potential of the proposed method for automated condition assessment of large real world structures, operating in a wide range of conditions.

2. The Bi-Component Framework

The novel approach proposed herein relies on observation of the time-variability of WT structures in a dual-reference time-frame, namely: (a) the short-term scale, relating to periodic fluctuations induced by the rotational nature of components of the WT system itself as well as the nature of the input loads, and (b) the long-term scale, in which phenomena associated to constantly changing loading and environmental conditions perturb structural behavior.

In satisfying the first goal, time-varying autoregressive moving average (TARMA) models deliver a suitable tool, capable of tracking the short-term variability of the structure. This parametric modeling technique proves adept in capturing the periodic variability and non-stationary dynamics

of the operating structure, primarily attributed to: (i) forces linked to the environmental conditions, namely aerodynamic loads which originate from the linear increase in the vertical wind profile in the rotor-swept area, translated into cyclic load changes on the rotating rotor, and (ii) forces related to the operating conditions, i.e., gravitational forces whose direction relative to the blade is shifted during rotation, causing alternating periodic loading conditions.

The long-term front is on the other hand focused on delivering the big picture, and therefore aims at accurate description of the WT dynamics over the complete operating spectrum of the structure. The typical stimuli linked to the long-term structural variability are changing within the range of hours, or even minutes and are usually related to changes in the wind flow profile, stochastic loads caused by wind turbulence, as well as variation of environmental conditions. For the purpose of modeling the evolution of the system response as a function of varying acting loads and operational conditions, the TARMA model is combined with the PCE method. The PCE model builds an expansion of a random output variable, e.g., an extracted TARMA model feature, on polynomial basis functions, which are orthogonal with respect to the probability density of the system's random inputs, e.g., monitored environmental/loading parameters. The synergy between the two models enables a holistic assessment of a structure interacting with its environment. A schematic overview of the proposed strategy is presented in Figure 1.

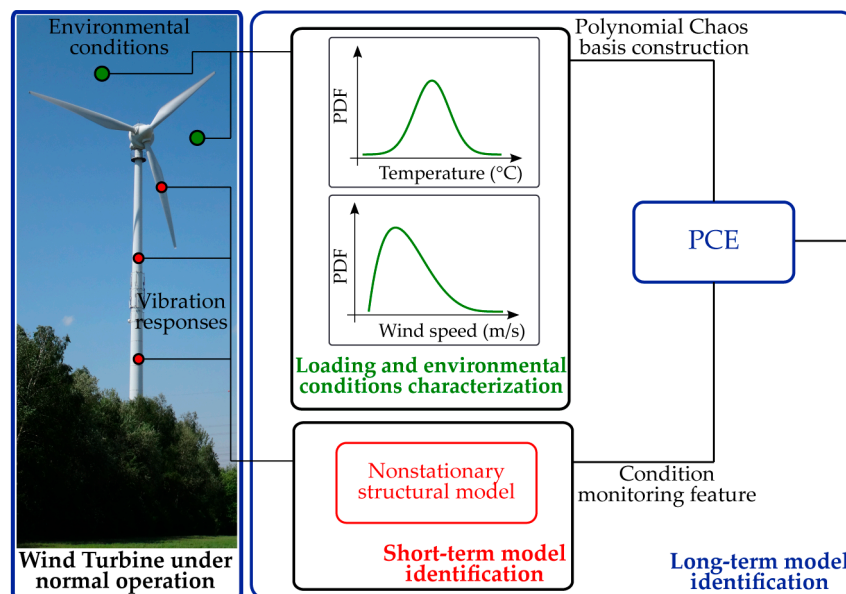


Figure 1. Schematic overview of the proposed bi-component monitoring strategy.

Considering for example a WT structure in regular operation, under the influence of ambient loading and changing environmental factors, a selected output feature (statistical moments of model residual) from the TARMA fitted model (step I) could be expanded on an appropriately constructed PC basis. By proper utilization of the TARMA model residuals (output parameter) and of the experimentally estimated operational and environmental variables (input parameter), the PCE model is finally constructed (step II). This coupled strategy provides the means for quantifying the uncertainty of WT dynamics due to randomness of the operational and environmental influencing factors. Furthermore, it delivers performance indicators in a standardized manner, which in turn contribute to development of a sophisticated damage detection strategy, where false fault alarms are timely recognized. An overview of the involved steps is presented in Figure 2. The following sections summarize the theoretical background underlying these steps. For more detailed information on the applied models, the interested reader is referred to [30–40].

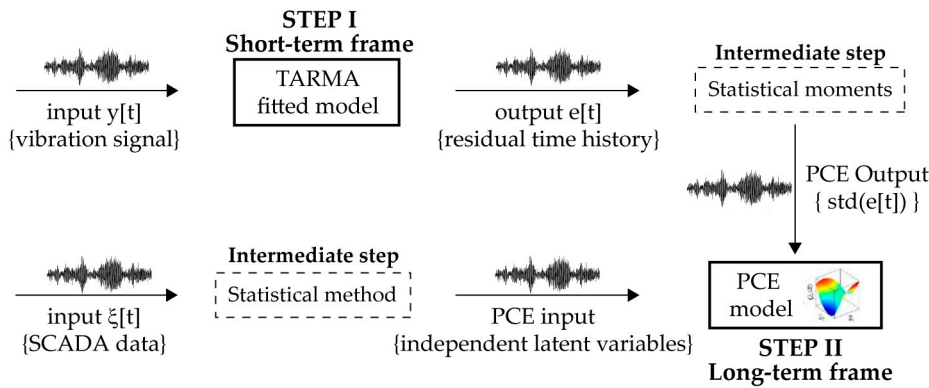


Figure 2. Steps of the proposed bi-component monitoring framework.

2.1. Step I: Short-Term Framework

2.1.1. The Model Structure

For a non-stationary signal $y[t]$ the TARMA model, provides a compact parameterized formulation given in Equation (1), [30]. Similar to the descriptive nature of its stationary counterpart (i.e., the ARMA model), the TARMA model describes a regressive form of the present output $y[t]$ on its past realizations $y[t-1], \dots, y[t-n]$, taking into account modeling error and time-variability of the AR/MA parameters of the model:

$$y[t] + \sum_{i=1}^{n_a} a_i[t] \cdot y[t-i] = e[t] + \sum_{i=1}^{n_c} c_i[t] \cdot e[t-i] \quad e[t] \sim NID(0, \sigma_e^2[t]) \quad (1)$$

where t designates discrete time (with $i = 1, 2, \dots, N$) of the observed nonstationary signal $y[t]$, $e[t]$ is the residual sequence, i.e., the unmodeled part of the signal, assumed to be normally identically distributed with zero mean and time-varying variance $e[t] \sim NID(0, \sigma_e^2[t])$, and $a_i[t]$, $c_i[t]$ the time-varying AR and MA parameters respectively, of AR, MA orders n_a, n_c .

In general, the identification of a specific TARMA model comprises two steps: (i) proper selection of the model structure (the n_a/n_c order for AR and MA parameters); (ii) estimation of the AR/MA parameters and the innovations variance $\sigma_e^2[t]$. However, for the specific subclass of smoothness-priors time-varying autoregressive moving average (SP-TARMA) models, the structure selection problem is further expanded to the selection of the smoothness constraints' order κ . The order κ is directly related to an additionally imposed stochastic structure upon the AR and MA parameters. In order to constrain their time evolution, the SP-TARMA model enhances Equation (1) with the following set of supplementary stochastic difference equations, [30]:

$$(1 - B)^\kappa a_i[t] = w_{a_i}[t] \quad w_{a_i}[t] \sim NID(0, \sigma_{w_{a_i}}^2[t]) \quad (2)$$

$$(1 - B)^\kappa c_i[t] = w_{c_i}[t] \quad w_{c_i}[t] \sim NID(0, \sigma_{w_{c_i}}^2[t]) \quad (3)$$

where B is the backshift operator ($B^k x[t] = x[t-k]$), κ designates the difference equation order, and $w_i[t]$ is a zero-mean, Gaussian sequences with time-dependent variance, uncorrelated, mutually uncorrelated and also uncorrelated with $e[t]$.

Thus, the full SP-TARMA model may be completely described by: (i) a model for the system response $y[t]$, Equation (1), and (ii) a model which controls the time evolution of the AR and MA parameters of the first model, Equations (2) and (3). The selected AR parameter order n_a represents the memory of the model and defines the number of past values of the signal, which are to be included into Equation (1). On the other hand, the stochastic smoothness constraints reflect our prior knowledge

regarding the evolution of the underlying dynamics. The chosen order of the stochastic difference equations κ introduces the number of past values of AR/MA parameters to be incorporated in the stochastic model of the constraints, and thus determines the behavior of the time varying AR/MA coefficients. For example for order $\kappa = 1$, the stochastic difference equations represent the random walk model, while for $\kappa = 2$ each approximated AR/MA parameter locally follows a straight line and is distributed as a Gaussian white noise sequence [31]. The time evolution smoothness of the AR/MA parameters is furthermore influenced, in a reciprocal sense, by the variance $\sigma_{w_i}^2[t]$. Therefore, a third user selected parameter, the residual variance ratio $\nu = \sigma_w^2[t]/\sigma_e^2[t]$, controls the equivalent memory of the estimation algorithm between the two limit values: $\nu \rightarrow 0$, which implies a locally deterministic (polynomial) parameter evolution, and $\nu \rightarrow \infty$ which implies no structure on parameter evolution [30]. An optimal value for ν may be achieved by means of statistical approaches, such as minimization of the AIC (Akaike information criterion) or the BIC (Bayesian information criterion). Both criteria employ a penalty term that controls the increase of the model order, thus ensuring adequate modeling precision without overfitting the N-sample-long modeled signal y^N , Equations (4) and (5):

$$AIC = -2 \cdot \ln \mathcal{L}(y^N) + 2 \cdot d \tag{4}$$

$$BIC = -\ln \mathcal{L}(y^N) + \frac{\ln N}{2} \cdot d \tag{5}$$

where $\mathcal{L}(\cdot)$ is the model likelihood, N the number of signal samples, and d is the number of independently estimated model parameters [30].

2.1.2. Model Parameter Estimation

We reformulate the general (κ -th order) smoothness constraints equations, Equations (2) and (3), in the following form:

$$z[t] = F \cdot z[t - 1] + G \cdot w[t] \tag{6}$$

where:

$$z[t] = [a_1[t] \dots a_{n_a}[t] c_1[t] \dots c_{n_c}[t] \dots a_1[t - k + 1] \dots a_{n_a}[t - k + 1] c_1[t - k + 1] \dots c_{n_c}[t - k + 1]]_{k \cdot (n_a + n_c) \times 1}^T \tag{7}$$

$$w[t] = [w_{a1}[t] w_{a2}[t] \dots w_{a n_a}[t] w_{c1}[t] w_{c2}[t] \dots w_{c n_c}[t]]_{(n_a + n_c) \times 1}^T \tag{8}$$

and F, G are matrices of the following forms (depending upon the value of κ):

$$\begin{aligned} \kappa = 1 : & \quad F \triangleq I_{n_a + n_c}, & \quad G \triangleq I_{n_a + n_c}, \\ \kappa = 2 : & \quad F \triangleq \begin{bmatrix} 2 \cdot I_{n_a + n_c} & -I_{n_a + n_c} \\ I_{n_a + n_c} & 0_{n_a + n_c} \end{bmatrix}, & \quad G \triangleq \begin{bmatrix} I_{n_a + n_c} \\ 0_{n_a + n_c} \end{bmatrix}, \\ \kappa = 3 : & \quad F \triangleq \begin{bmatrix} 3 \cdot I_{n_a + n_c} & -3 \cdot I_{n_a + n_c} & I_{n_a + n_c} \\ I_{n_a + n_c} & 0_{n_a + n_c} & 0_{n_a + n_c} \\ 0_{n_a + n_c} & I_{n_a + n_c} & 0_{n_a + n_c} \end{bmatrix}, & \quad G \triangleq \begin{bmatrix} I_{n_a + n_c} \\ 0_{n_a + n_c} \\ 0_{n_a + n_c} \end{bmatrix}, \end{aligned} \tag{9}$$

and so on,

(I_n and O_n designate the $n \times n$ dimensional identity and zero matrices, respectively.)

The TARMA (n_a, n_c) representation of Equation (1) may then be expressed as:

$$y[t] = h^T [t, z^{t-1}] \cdot z[t] + e[t] \tag{10}$$

with:

$$h [t, z^{t-1}] = [-y[t-1] \dots - y[t-n_a]; e [t-1, z^{t-1}] \dots e [t-n_c, z^{t-n_c}]; 0 \dots 0]_{k \cdot (n_a+n_c) \times 1}^T \quad (11)$$

z^t designates a vector containing all state vectors $z[t]$ up to time t .

Hence, the SP-TARMA (n_a, n_c) model may be completely expressed in state-space form, Equations (6) and (10). However, due to the unknown residual sequence $e [t, z^t]$ in $h [t, z^{t-1}]$, the full SP-TARMA case leads to a nonlinear state estimation problem. For this reason, i.e., for the recursive estimation of the state vector $z[t]$, an Extended Least Squares (ELS)-like algorithm is employed. This leads to replacement of the theoretical prediction errors $e [t, z^t]$ in Equation (11) by their posterior estimates $\hat{e}[t, t]$, [30]:

$$h [t] = [-y[t-1] \dots - y[t-n_a]; \hat{e}[t-1, t-1] \dots \hat{e}[t-n_c, t-n_c]; 0 \dots 0]_{k \cdot (n_a+n_c) \times 1}^T \quad (12)$$

$$\hat{e}[t|t] = y[t] - h^T [t] \cdot \hat{z}[t|t] \quad (13)$$

Then, for a given and time-invariant order κ and residual variance ratio ν , the estimation of the state-space model may be achieved via the ordinary Kalman Filter (KF) scheme (Figure 3). The innovations (one-step-ahead prediction error) variance, $\sigma_e^2 [t]$, may be estimated via a window of certain length, centered at time instant t , which slides over the prediction error (residual) sequence (more details and a summary of the normalized version of the KF method is provided in [30]).

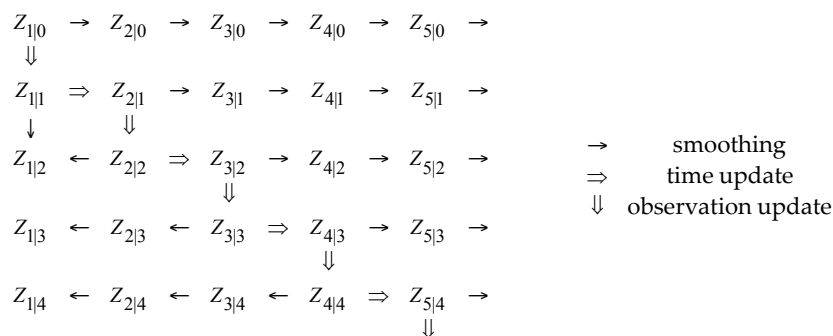


Figure 3. Schematic overview of the KF scheme (adapted from [32]).

2.1.3. SP-TARMA Based Dynamic Analysis

When the best-fit SP-TARMA model is identified, the time-varying dynamics of the structure may be recovered by the frozen-time spectrum [30]:

$$S(\omega, t) = \left| \frac{1 + \sum_{i=1}^{n_c} c_i [t] e^{-j\omega T_s i}}{1 + \sum_{i=1}^{n_a} a_i [t] e^{-j\omega T_s i}} \right|^2 \sigma_e^2 [t] \quad (14)$$

where ω is the frequency in rad/sec, T_s the sampling period in s, and j the imaginary unit. The frozen-time natural frequency $\omega_{ni} [t]$ and damping ratio $\zeta_i [t]$ of the system are related to the $\lambda_i [t]$ discrete-time ‘frozen’ pole of the TARMA model’s frozen frequency response function:

$$\omega_{ni} [t] = \frac{|\ln \lambda_i [t]|}{T_s} (rad/sec) \quad \text{and} \quad \zeta_i [t] = -\cos(\arg(\ln \lambda_i [t])) \quad (15)$$

2.2. Step II: Long-Term Framework (Modeling Uncertainty)

The time domain identification method described in Section 2.1, accounting for the non-stationarity of the WT system, is combined with a PCE tool in order to deliver a SHM framework capable of describing the comprehensive structural dynamics, in a long-term scale.

In contrast to the local representative nature and high computational cost of the well-known Monte Carlo collocation method and other similar sampling techniques, the PCE non-sampling approach is focused on constructing a functional dependence of the model output on the set of independent random variables (RVs) (following a certain probability law) that parametrizes the input data, usually expressed in terms of a series [33]. The polynomial chaos (PC) approach draws its origins in the modeling of stochastic processes via Gaussian RVs by employment of Hermite polynomials [34]. Later on, Cameron and Martin [35] demonstrated that for any arbitrary stochastic process with finite variance, the PC expansion converges in the \mathcal{L}_2 sense. The latter was generalized by Xiu et al. [36] to different continuous and discrete distributions using orthogonal polynomials of the so called Askey-scheme [37].

The main concept underlying the PCE framework is to represent a random quantity by means of expansion, comprising functions of random variables multiplied with deterministic coefficients. More precisely, PCE modeling generates an expansion of a random output variable on polynomial chaos basis functions, which are orthonormal with respect to the probability space of the system's random inputs. For a system S featuring M random input parameters represented by independent random variables $\{\xi_1, \dots, \xi_M\}$, gathered in a random vector Ξ of prescribed joint Probability Density Function (PDF) $p_{\Xi}(\xi)$, the system output, denoted by $Y = S(\Xi)$, will also be random [38]. Provided that Y has finite variance, it can be expressed as follows:

$$Y = S(\Xi) = \sum_{d \in N^M} \theta_d \varphi_d(\Xi) \quad (16)$$

where θ_d are unknown deterministic coefficients of projection, d is the vector of multi-indices of the multivariate polynomial basis (PC functions) $\varphi_d(\Xi)$, orthonormal to $p_{\Xi}(\xi)$. The PC polynomials are chosen so that the weights $w(x)$, which ensure the orthogonal property.

$\langle \varphi_i, \varphi_j \rangle = \int_{x_1}^{x_2} \varphi_i(x) \varphi_j(x) w(x) dx = \delta_{ij}$ (with $\delta_{ij} = 1$ if $i = j$ and 0 otherwise), resemble the PDF of the random variables. Several well-known families of orthogonal polynomials can be associated to a specific PDF, a more detailed overview may be found in [36].

The multivariate polynomials $\varphi_d(\Xi)$ used for constructing the PC model (for each term j of the PC series expansion) are obtained by tensor products of the corresponding univariate functions:

$$\varphi_{d_j}(\Xi) = \prod_{m=1}^M \varphi_{d_{j,m}}^1(\xi_m) \quad (17)$$

where M is the number of random input variables.

For ensuring timely computation, the functional series terms Equation (16) must be truncated to a finite number. The usual approach is the selection of the multivariate polynomial basis with total maximum degree $|d_j| = \sum_{m=1}^M d_{j,m} \leq P \forall j$. In this case the dimensionality of the functional subspace, i.e., the total number of series terms, equals:

$$T = \frac{(M+P)!}{M!P!} \quad (18)$$

where P is the maximum total degree of the multivariate polynomials. The truncated PCE model of Equation (16) to the first T terms yields a finite number of deterministic coefficients and the parameter vector θ_d may be estimated by solving Equation (16) in a least squares sense [38].

For illustrating the construction of the PCE model, let us consider the simple case of a 2D-PCE model for two independent input random variables ξ_1, ξ_2 . By employing the first four univariate Legendre polynomials as PC basis functions for a selected total degree $P = 3$, we may develop the

truncated PC expansion of Equation (19) (total number of terms T = 10, single index notation used). The construction of the 2D polynomials is presented in more detail in Table 1, and a schematic overview of the complete method is presented in Figure 4.

$$\begin{aligned} \tilde{Y} \equiv \sum_{j=0}^9 \theta_j \varphi_j = & \theta_0 + \theta_1 \zeta_1 + \theta_2 \zeta_2 + \theta_3 \frac{1}{2} (3\zeta_1^2 - 1) + \theta_4 \zeta_1 \zeta_2 + \theta_5 \frac{1}{2} (3\zeta_2^2 - 1) \\ & + \theta_6 \frac{1}{2} (5\zeta_1^3 - 3\zeta_1) + \theta_7 \frac{1}{2} (3\zeta_1^2 - 1) \zeta_2 + \theta_8 \frac{1}{2} (3\zeta_2^2 - 1) \zeta_1 + \theta_9 \frac{1}{2} (5\zeta_2^3 - 3\zeta_2) \end{aligned} \quad (19)$$

Table 1. Multivariate 2D Legendre polynomials for total polynomial order P = 3.

First Four 1D Legendre Polynomials $L_j = \varphi_{d,j}^1$	Single Index j	Multi-Index d	$\prod_{m=1}^M \varphi_{d_j,m}^1(\Xi)$	
$L_0 = 1$	0	(0,0)	$L_0 L_0$	$\varphi_0 = 1$
$L_1 = x$	1	(1,0)	$L_1 L_0$	$\varphi_1 = \zeta_1$
$L_2 = 0.5(3x^2 - 1)$	2	(0,1)	$L_0 L_1$	$\varphi_2 = \zeta_2$
$L_3 = 0.5(5x^3 - 3x)$	3	(2,0)	$L_2 L_0$	$\varphi_3 = 0.5(3\zeta_1^2 - 1)$
	4	(1,1)	$L_1 L_1$	$\varphi_4 = \zeta_1 \zeta_2$
	5	(0,2)	$L_0 L_2$	$\varphi_5 = 0.5(3\zeta_2^2 - 1)$
	6	(3,0)	$L_3 L_0$	$\varphi_6 = 0.5(5\zeta_1^3 - 3\zeta_1)$
	7	(2,1)	$L_2 L_1$	$\varphi_7 = 0.5(3\zeta_1^2 - 1)\zeta_2$
	8	(1,2)	$L_1 L_2$	$\varphi_8 = 0.5(3\zeta_2^2 - 1)\zeta_1$
	9	(0,3)	$L_0 L_3$	$\varphi_9 = 0.5(5\zeta_2^3 - 3\zeta_2)$

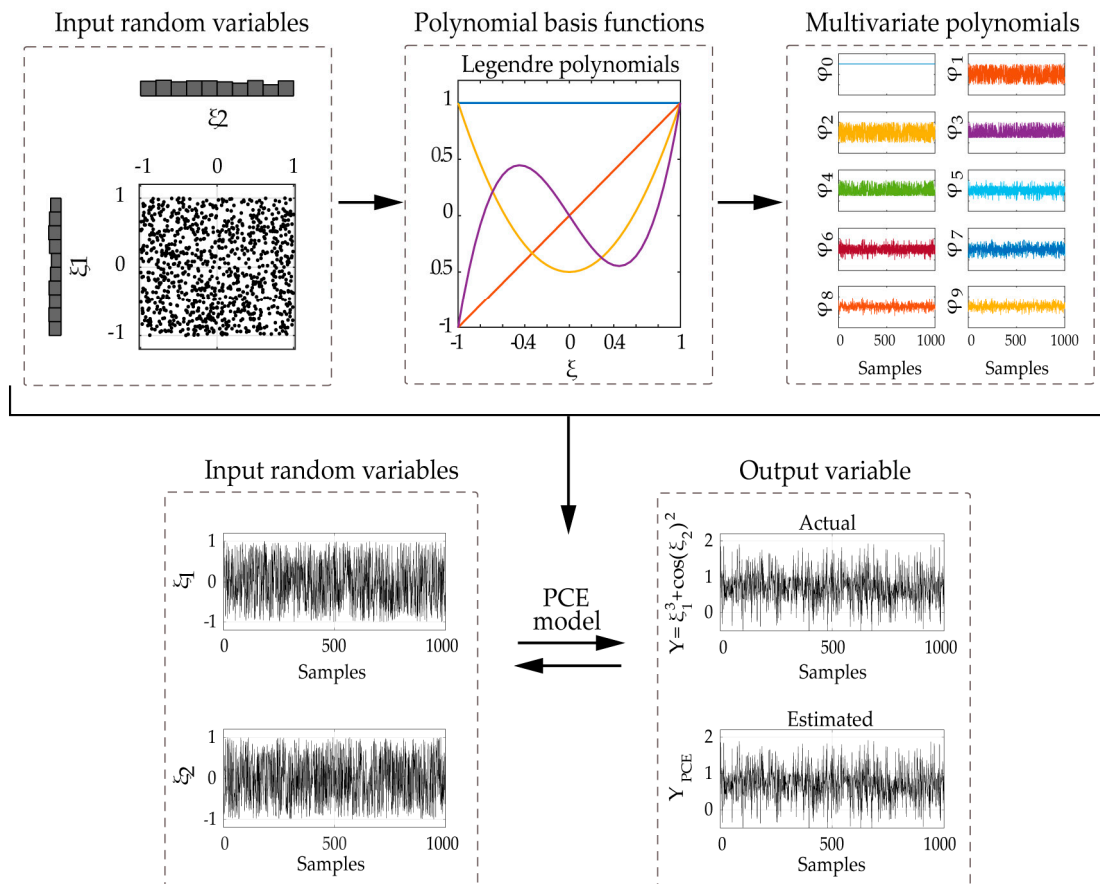


Figure 4. Schematic overview of the PCE method: example of two independent input random variables ζ_1, ζ_2 with uniform pdf and a system output variable $Y = \zeta_1^3 + \cos(\zeta_2)^2$.

It is worth mentioning that in case of subpar performance of the standard PCE due to the rather limiting assumption of prescribed families of input distributions, an arbitrary or adaptive PCE scheme could serve as viable alternatives [33,38,39]. Unlike the classical approach, which utilizes a fixed form of PC basis functions, these variants employ PC functions adapted to the specific input variable characteristics.

Probabilistic Approximation of Random Variables

The prerequisite of independent input parameters, necessary for efficient PCE modeling, is here ensured by utilization of a statistical technique capable of inferring independent (latent) variables that are intermixed within observed data. Independent Component Analysis (ICA) is a computational approach that in contrast to Principal Component Analysis is concerned with the higher order statistics of the observable data, i.e., minimization of mutual information. ICA eventually aims at identifying non-Gaussian and mutually independent components [40], which may then be fed into the PCE component of the proposed SHM framework.

The basic concept of ICA is in fact to describe how an observed random vector x is generated by a process of mixing unobservable (latent) variables. More specifically, if we consider n observed linear mixtures of n independent components s_i ($i = 1, \dots, n$) the following relation is assumed [40]:

$$x_j = a_{j,1}s_1 + a_{j,2}s_2 + \dots + a_{j,n}s_n \quad \text{for } j = 1, \dots, n \quad (20)$$

or by using vector-matrix notation:

$$x = As \quad (21)$$

with x designating the random observed variables vector, s the vector of unobserved independent variables and A the unknown mixing matrix:

$$A = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix}$$

Via inversion of the mixing matrix A , the vector of the independent latent variables s becomes:

$$s = A^{-1}x = Wx \quad (22)$$

ICA employs nonlinear optimization for estimating the mixing matrix A , which maximizes the non-Gaussianity of each one of the latent variables $w^T x$ (with w designating a column vector of matrix W). The adopted measure of non-Gaussianity, may be based on kurtosis, negentropy, and others [40]. A concise overview of the method, as well as a comprehensive flowchart of the ICA algorithm is presented in [23].

3. Steps of Application—Case Studies

3.1. Description of the Monitored Structures

The previously described bi-component SHM framework was initially implemented for tracking of the performance of an actual operating WT tower by utilizing 10 min long acceleration datasets, recorded every half hour for 29 days (from 18 December 2013 till 15 January 2014) [22]. Within this study the robustness of the strategy is further tested over a more extended time frame and a broader set of input parameters for the case of two operating WT structures, situated in different sites.

The first case (in the further text: case A) is a 2MW WT which is part of a wind farm located in Lübbenau, Germany, [22], where access is obtained through collaboration with Repower Wind Lübbenau GmbH. The vibration response of the WT is measured via triaxial accelerometers (STM

LIS344ALH MEMS sensors) at four distinct locations of the WT tower (Figure 5). The output-only vibration data is available in 10-min-long dataset records for each hour from 3 March 2015 till 6 May 2015, a total of 1388 datasets, measured at a 1600 Hz sampling rate. The SCADA data for the same time period are available in a 10-min average format.

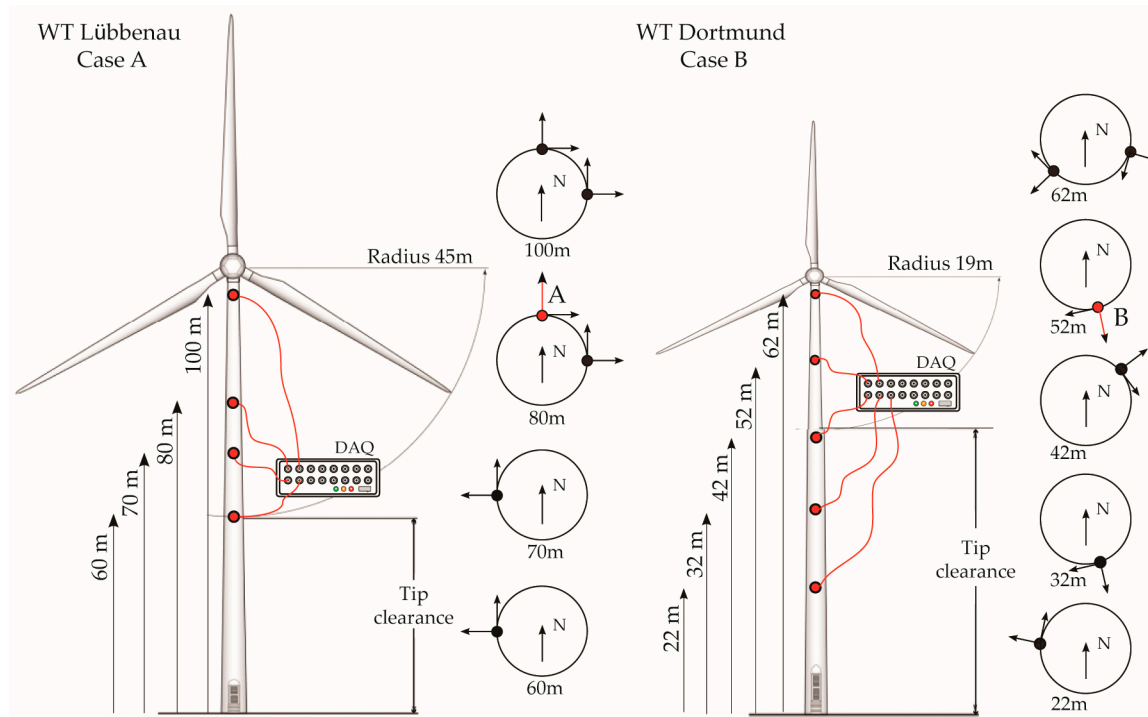


Figure 5. Schematic overview of the experimental setup (case A left and case B right).

The second structure under study (case B) is a 0.5MW WT erected in 1997, located in the vicinity of Dortmund, Germany [41], where access is obtained through collaboration with Dortmunder Energie- und Wasserversorgung GmbH (DEW21). The acceleration response of the WT is measured by triaxial accelerometers (PCB-3713D1FD3G MEMS sensors) at five positions along the shaft (Figure 5). The vibration response as well as the SCADA data are sampled with the same frequency of 100 Hz. The implementation of the proposed SHM strategy for this case is presented for acceleration records of four complete months of continuously monitored data (May to August 2010).

3.2. Operational Modal Analysis for Parked Conditions

The initial identification of modal properties for both structures was carried out on datasets corresponding to parked conditions. For this purpose, the stationary ARMA method (prediction-error based) was applied for acceleration response signals measured at 80 m (Case A) and 52 m (Case B) height (directions marked as A and B in Figure 5). In Figure 6 the estimated stabilization diagrams for ARMA model orders between 2 and 50 are presented, together with results of the stochastic subspace identification method, based on the canonical variate algorithm for comparing evaluation. The ARMA estimated natural frequencies, for damping ratios less than 5%, for selected model orders equal to 32 and 18, for case A and B respectively, are summarized in this figure as well. The spectrograms (short time Fourier transform; Hamming data window; NFFT = 512; overlap 98%), appended in the background of the same plot, clearly demonstrate the stationarity of identified frequencies within the explored time frames.

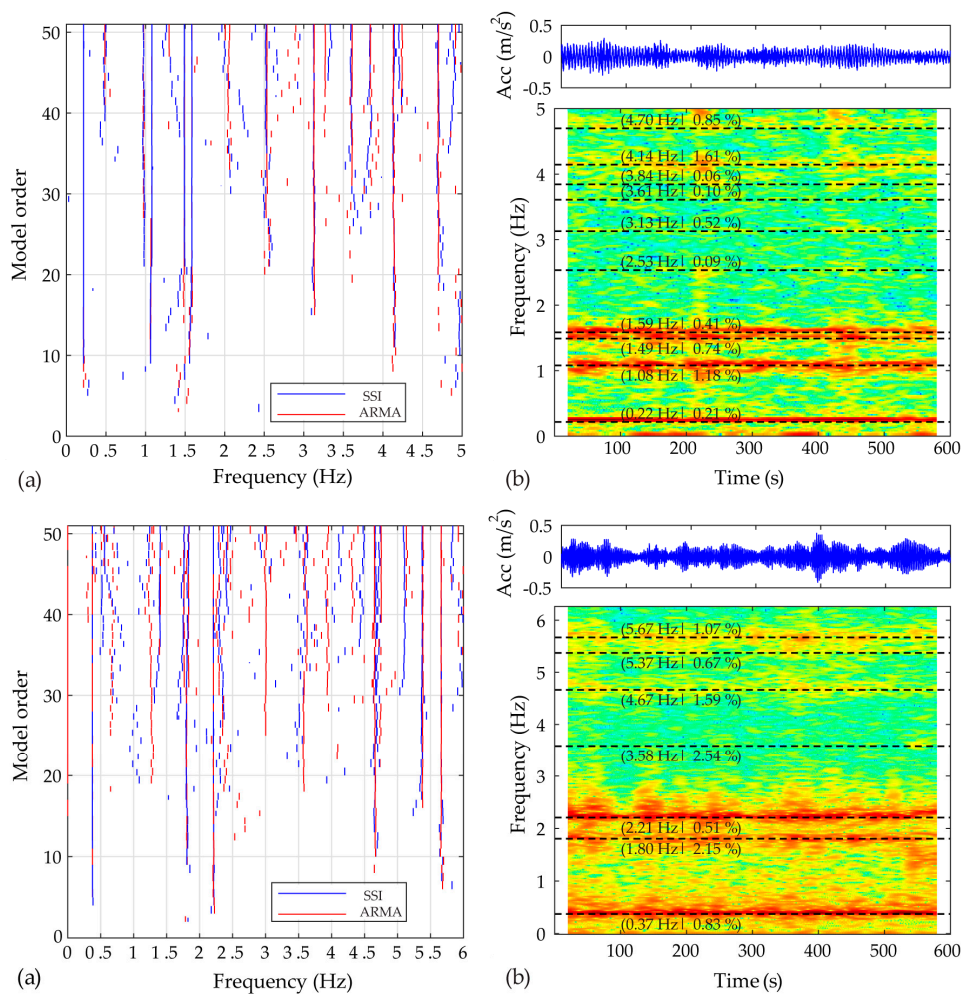


Figure 6. Dynamics of the parked WTs (Case A up and Case B bottom): (a) Stabilization plot for the stationary ARMA and SSI methods (model orders from 2 to 50); (b) Spectrogram and estimated natural frequencies and damping ratios based on an ARMA(32,32) model for Case A and ARMA(18,18) model for Case B (For interpretation of the references to color, the reader is referred to the web version of this article).

3.3. Short-Term Framework Implementation

In contrast to response that remains close to stationary, where the standard OMA techniques are commonly employed, the complex dynamic system of operating WT structures underlines the necessity for utilizing more efficient time-sensitive tools. The non-stationary dynamics of an operational WT structure can be clearly observed in Figure 7 by the spectrograms of a signal recorded at 2:00 p.m. 29 March 2015 (case A) and at 1:00 p.m. 31 June 2010 (case B). The SP-TARMA model (solid lines) is contrasted to the corresponding ARMA model (gray lines) and compared to the spectrograms in the background. As observed the SP-TARMA model is able to track the evolution of the time varying frequencies. The accuracy is different depending on the observed frequency. This precision is regulated by the residual variance ratio, which is directly linked to the evolution of the AR/MA coefficients, but therefore only indirectly linked to the evolution of the individual frequencies (Equation (15)). In separate work the optimal tuning of this parameter will be sought. However, we here demonstrate that this does not pose an issue for construction of the desired performance indicator. An alternate approach would be to filter the signal and separately treat the different frequency components, however there is a point to avoiding such practices as we here try to develop a framework that is minimally invasive with respect to the monitored signal.

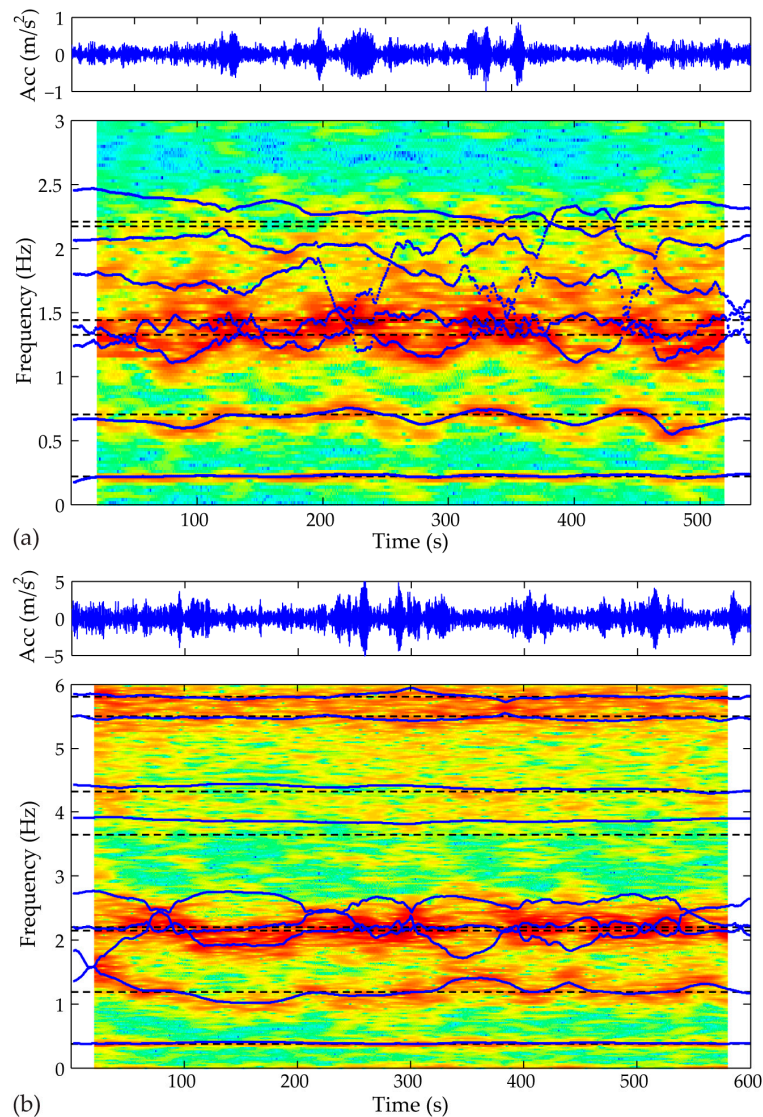


Figure 7. Dynamics of the WTs under normal operation: Stationary ARMA (n_a, n_c) versus nonstationary SP-TARMA (n_a, n_c) model estimates with spectrogram in the background (Short Time Fourier Transform; Hamming data window; NFFT = 512; overlap 98%). (a) Case A: (n_a, n_c) = (32, 32); (b) Case B: (n_a, n_c) = (18, 18).

The same figure reveals additional frequencies fluctuating around 0.7 Hz (Case A) and 1.2 Hz (Case B), which match to the 3P harmonics of the rotating WT structures (for the corresponding time frames mean RPM equal 15 and 24, for Case A and B respectively). It is worth noting that in contrast to strategies where existent harmonics interfere the process of structural identification, the short-term tool framework delivers a robust performance indicator, without the necessity of eliminating alternating dynamic effects introduced by the operating WT components. Moreover, the ability of the proposed tool for incorporated tracking of existing harmonics may further facilitate the detection of potential rotor related faults (i.e., unbalanced masses).

For the short-term framework implementation, the acceleration time histories for both cases were low-pass filtered and down-sampled to 12.5 Hz, cutoff frequency at 5 Hz (Case A) and 6 Hz (Case B), observed as 10-min long data sets. This results in total of 1388 data sets for Case A, and 17 706 data sets for case B. The acceleration records are utilized with their originally measured orientation. However, the underlying directional dependency is incorporated via the continuously measured WT nacelle

orientation (yaw angle), which is used as an input parameter in the long-term modeling framework of the next step.

After preprocessing the acceleration records, the first step of the proposed framework lies in fitting an appropriate SP-TARMA model to the actual 10-min structural response signals. Toward this end, the user-defined parameters of the SP-TARMA model (i.e., the model order n , the smoothness constraint order κ and the residual variance ratio ν) have to be properly tuned. The plots of the Bayesian and Akaike statistical criterion for model order selection are presented in Figure 8. The zoomed views (Figure 8b,c) indicate the range of values that may lead to adequate modelling of the nonstationary signal, without overfitting it. The BIC and AIC deliver similar results for both case studies. For Case A, however, the values minimizing the criteria functions do not adequately track the variation of the frequency at the vicinity of 0.7 Hz (Figure 9). Thus, in ensuring an automated operation of the short-term tool the values $n = 32$, $\kappa = 1$, $\nu = 0.01$, are selected. On the other hand, the inspection of the SP-TARMA estimates for Case B (Figure 10) reveals a good agreement with the parameter values that minimize the penalty functions ($n = 18$, $\kappa = 1$, $\nu = 0.0001$).

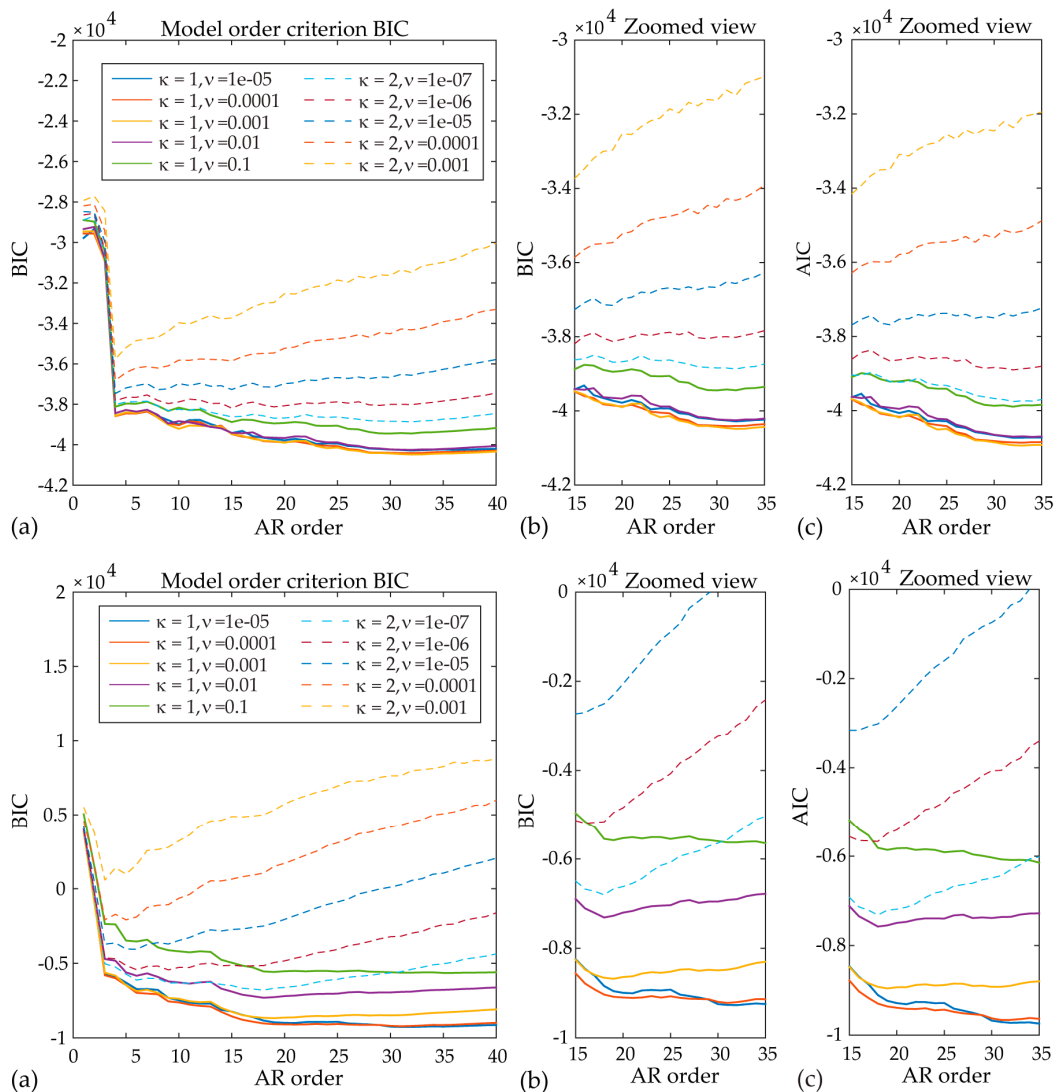


Figure 8. Statistical criterion for model order selection (Case A top and Case B bottom): (a) Bayesian statistical criterion; (b) Bayesian statistical criterion zoomed-in view; (c) Akaike statistical criterion zoomed-in view (For interpretation of the references to color, the reader is referred to the web version of this article).

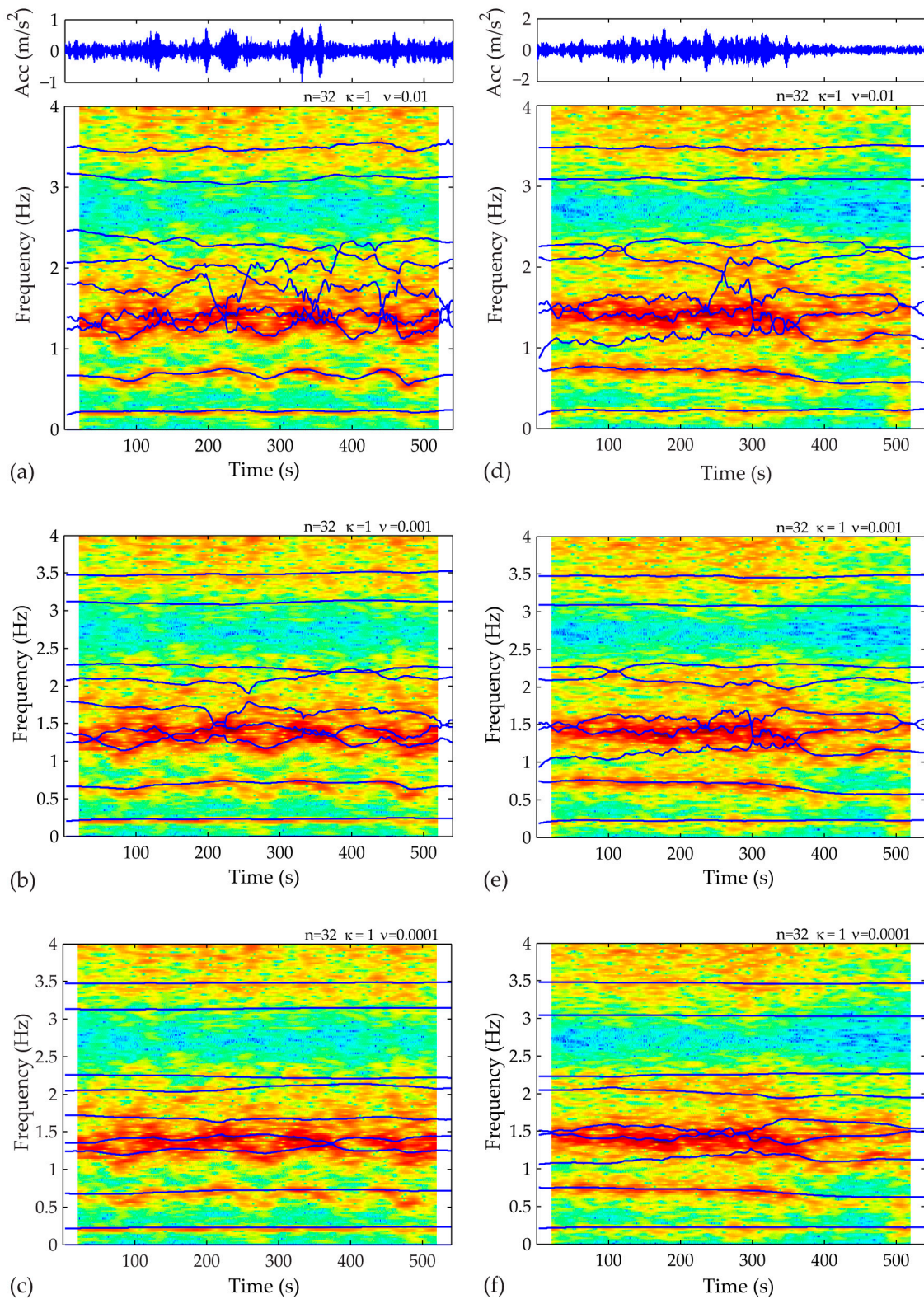


Figure 9. SP-TARMA ($n = 32$, $k = 1$) model tuning for Case A by varying the value of the ratio of model residuals. First sample of a 10-min acceleration signal fitted with: (a) $\nu = 0.01$; (b) $\nu = 0.001$; (c) $\nu = 0.0001$; Second sample of a 10-min acceleration signal fitted with: (d) $\nu = 0.01$; (e) $\nu = 0.001$; (f) $\nu = 0.0001$.

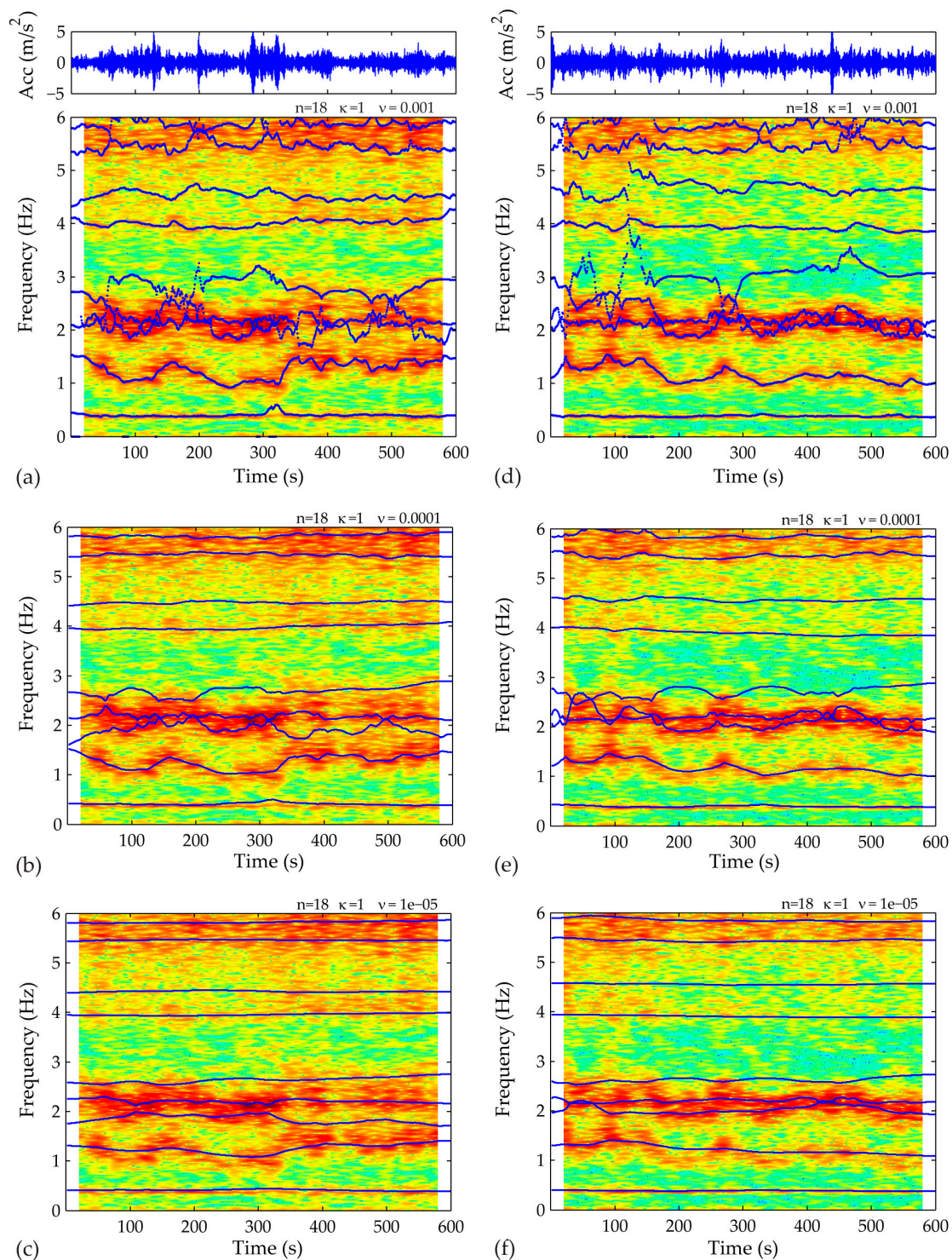


Figure 10. SP-TARMA ($n = 18$, $k = 1$) model tuning for Case B by varying the value of the ratio of model residuals. First sample of a 10-min acceleration signal fitted with: (a) $\nu = 0.001$; (b) $\nu = 0.0001$; (c) $\nu = 0.00001$; Second sample of a 10-min acceleration signal fitted with: (d) $\nu = 0.001$; (e) $\nu = 0.0001$; (f) $\nu = 0.00001$.

3.4. Long-Term Framework Application

For utilizing the long-term framework, input variables corresponding to operational and environmental measured data need to be selected. In Figure 11 the 10 min averages of the six

selected SCADA parameters for Case A are plotted. The correlation plots for each pair of chosen output variables is presented in Figure 12. It is apparent that several pairs may be assumed as correlated with correlation larger than 0.6 (marked with asterisk). In order to transform the input data to independent variables, thus satisfying the PCE method requirement, the ICA method is herein applied. The minimum required number of ICA variables fully describing the input data variance is revealed via inspection of the eigenvalues of the covariance matrices of the selected SCADA variables. For the first application case study, the corresponding three ICA-derived latent variables are presented in Figure 13a.

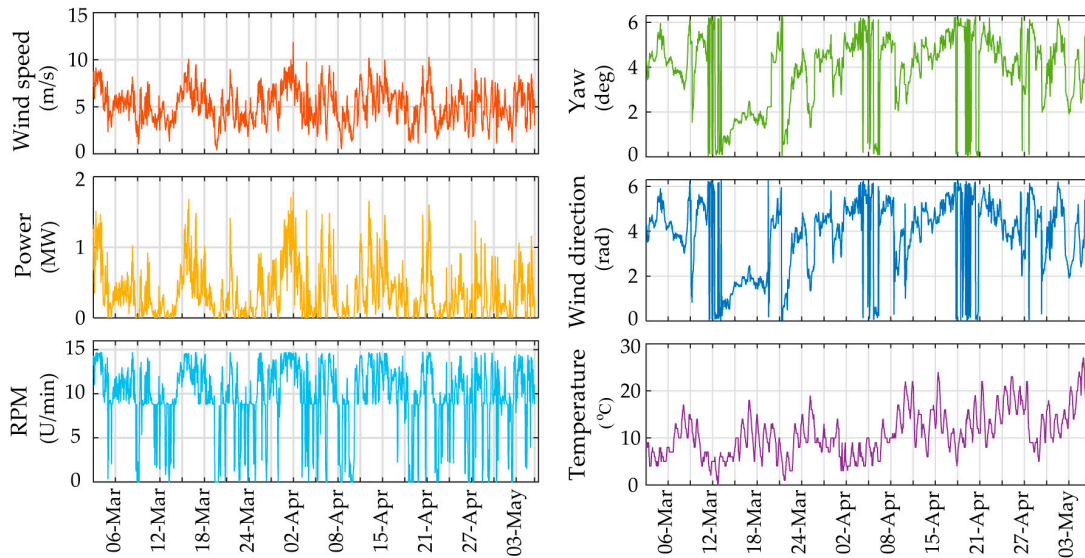


Figure 11. Time history plots of selected PCE input variables for 1388 datasets (Case A).

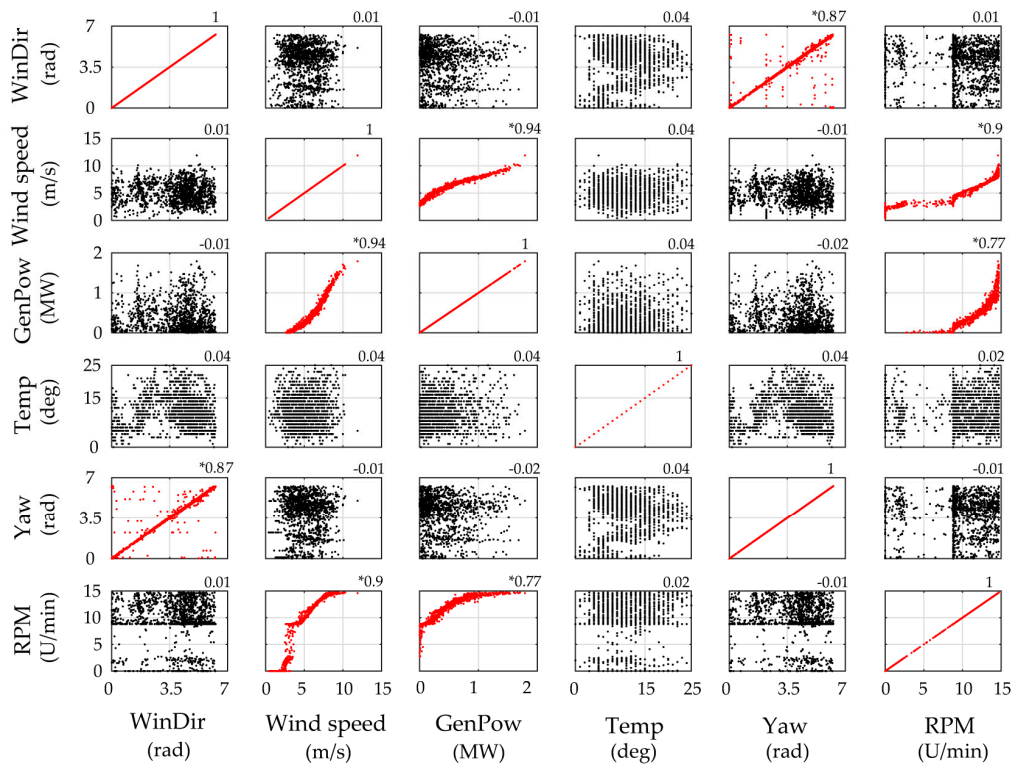


Figure 12. Scatter plots of selected PCE input variables and estimated correlation values (Case A).

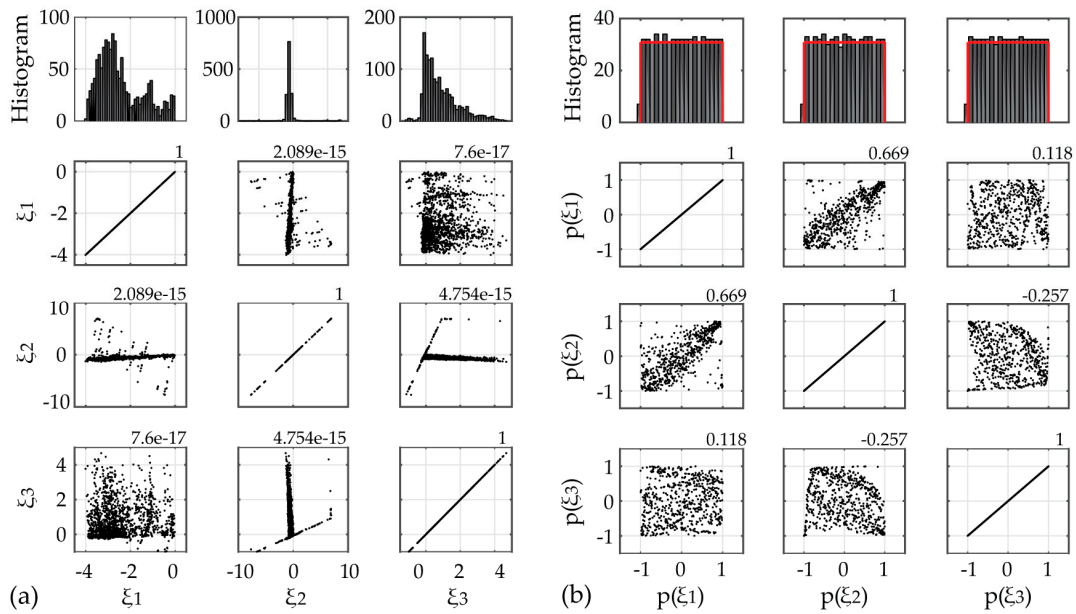


Figure 13. Input data for Case A: (a) Scatter plots of ICA-based input variables and estimated correlation values, histograms of the latent variables (upper row); (b) Scatter plots of the random vector Ξ of prescribed uniform PDF $p_{\Xi}(\xi)$, histograms of the random vector Ξ values (upper row).

For the purpose of constructing the random vector Ξ of prescribed joint PDFs $p_{\Xi}(\xi)$, the ICA estimates are further transformed into uniformly distributed variables via use of the non-parametrically estimated cumulative distribution functions, Figure 13b.

For the second case study the long-term framework is utilized for an expanded time frame of four full months. The time histories of the four selected input variables as well as their correlation values are plotted in Figure 14. The plots of the ICA-derived latent variables and scatter plots of the random vector Ξ of prescribed uniform PDF, after the twofold transformation of the input parameters to independent and uniform variables, are presented in Figure 15.

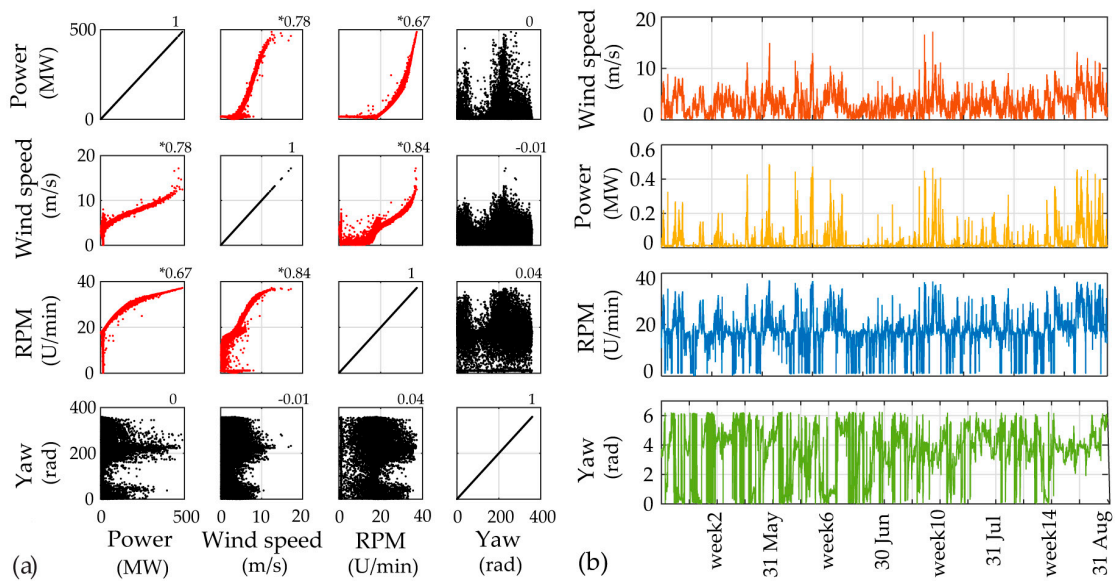


Figure 14. (a) Scatter plots of selected PCE input variables and estimated correlation values; (b) Time history plots of the selected PCE input variables for 17 706 datasets (Case B).

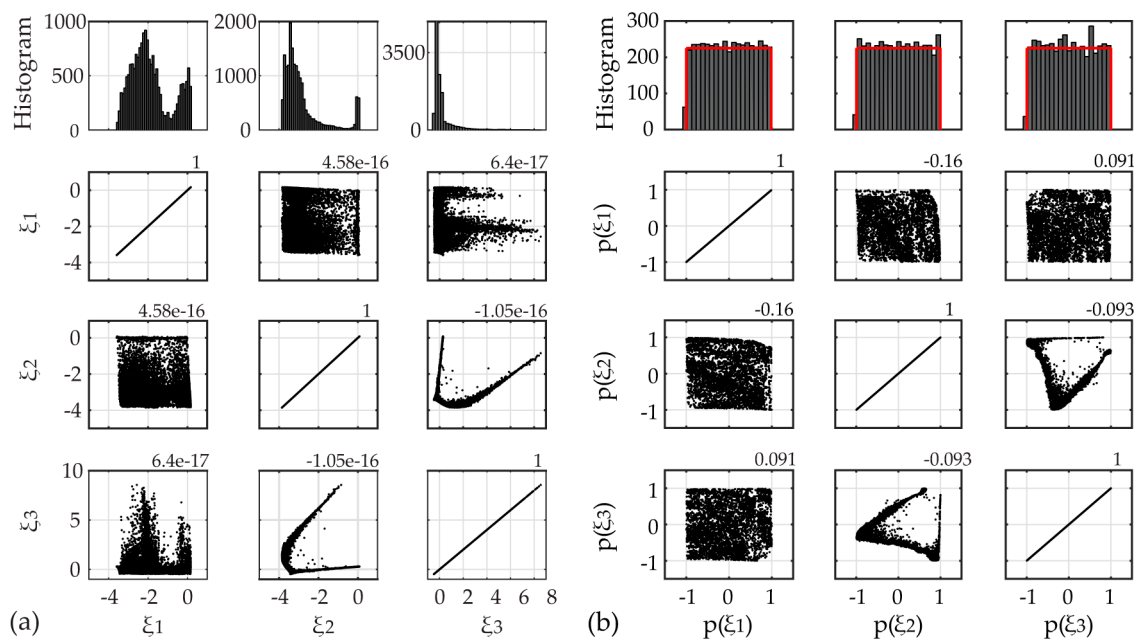


Figure 15. Input data for Case B: (a) Scatter plots of ICA-based input variables and estimated correlation values, histograms of the latent variables (upper row); (b) Scatter plots of the random vector Ξ of prescribed uniform PDF $p_{\Xi}(\xi)$, histograms of the random vector Ξ values (upper row).

As a last step, the SP-TARMA output variables and the PDFs of the measured operational input data are fed into the PCE (long-term) framework. In accordance with the uniform PDFs of the input data, the Legendre polynomials are selected as the PC functional basis. The maximum polynomial order is selected equal to three (case A) and five (case B). Further increasing the maximum order does not significantly improve the accuracy of the expansion. The standard deviation (std) of the SP-TARMA residuals for the 10 min intervals are selected as the PCE output parameter.

The standard deviation (std) of the residuals for each dataset and the PCE model estimates are plotted in Figures 16 and 17 for both cases. For case A, the total of 1388 data sets for the two month period, are divided into an estimation (45%) and validation (55%) period. The PCE errors are plotted in the same figure, along with the corresponding 99.7% confidence intervals calculated for the fitted Gaussian distribution of the estimation set errors. For case B, the total monitoring period of 4 months is divided into a one-month estimation period and a three-month validation period. As observed, the PCE model is capable of simulating the std(e) output variable with very good accuracy for both cases, and the model residual falls within the 99.7% confidence intervals for both sets. For the actual structures under study no damages were observed, with results verifying the applicability of the proposed framework for the continuous monitoring period of the two studied systems (summarized in Table 2).

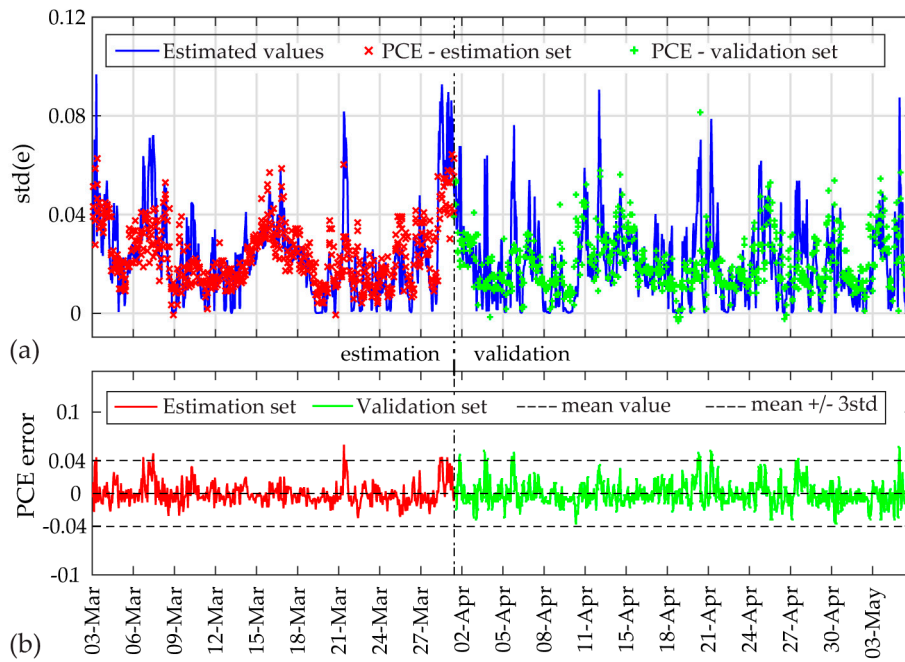


Figure 16. PCE estimates for Case A: (a) std of SP-TARMA(32,32,1,0.01) model residual and PCE model expansion values; (b) The PCE errors along with 99.7% confidence intervals (calculated for estimation set).

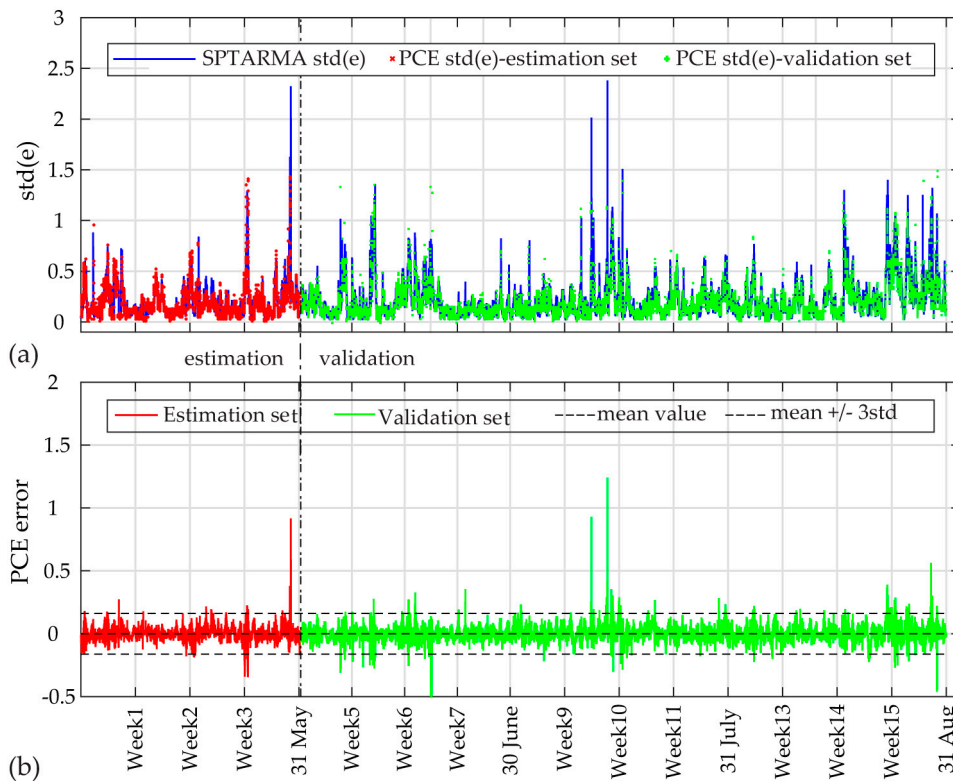


Figure 17. PCE estimates for Case B: (a) std of SP-TARMA(18,18,1,0.0001) model residual and PCE model expansion values; (b) The PCE errors along with 99.7% confidence intervals (calculated for estimation set).

Table 2. Summary of the applied bi-component tool for the two case studies.

PCE-TARMA Tool	SP-TARMA (Short-Term Modeling)			PCE (Long-Term Modeling)		
	Input	Model	Output	Input	Model	Output
Case A: 2MW WT Vibration response 10-min. datasets, 1/h, Mar–May 2015 SCADA data 10-min averages	- 1388 total datasets - filtered and down-sampled to 12.5 Hz	$n = 32, \kappa = 1,$ $\nu = 0.01$	residual time histories 10-min long	3 ICA-based variables from 6 SCADA variables	- order $P = 3$ - Legendre polynomials - estimation 45% of data - validation 55% of data	standard deviation of 10-min. long residual time histories (1388 values)
Case B: 0.5MW WT Vibration response 1 h datasets, cont., May–Aug 2010 sampling at 100 Hz SCADA data 1 h datasets sampling at 100 Hz	-17 706 total datasets -filtered and down-sampled to 12.5 Hz	$n = 18, \kappa = 1,$ $\nu = 0.0001$	residual time histories 10-min long	3 ICA-based variables from 4 SCADA variables	- order $P = 5$ - Legendre polynomials - estimation 25% of data - validation 75% of data	standard deviation of 10-min. long residual time histories (17,706 values)

4. PCE-TARMA Diagnostic Index

The SHM-based system diagnostics is built on training data from the baseline, or “healthy” structure under regular environmental and operational conditions. The obtained models are then implemented with newly acquired data, and deviations from the established normal conditions are detected as a novelty [28]. A robust diagnostic tool should be able to distinguish between true system changes and benign alarms, originating from new ranges of measured input data. In this case, instead of alarming for a potential maintenance intervention, model retraining should be carried out on the basis of the extended data set.

In this context, the developed PCE-TARMA model responsiveness to varying environmental and operational parameters (EOP) is tested for two, three and four-month long training periods, with an input of five distinct SCADA variables (Figure 18). Based on the previously described framework in Section 3, the long-term tracking is extended to one complete year time-frame of monitored data for the operating Case B WT structure. The PCE-TARMA model is refitted with the model parameters summarized in Table 2, Section 3.4. During the twelve-month tracking period influencing agents characterized by seasonal variations would bear a significant impact. Therefore, in this setting the temperature measured at the outer side of the WT tower shaft (at 20 m height) is incorporated as a PCE model input as well.

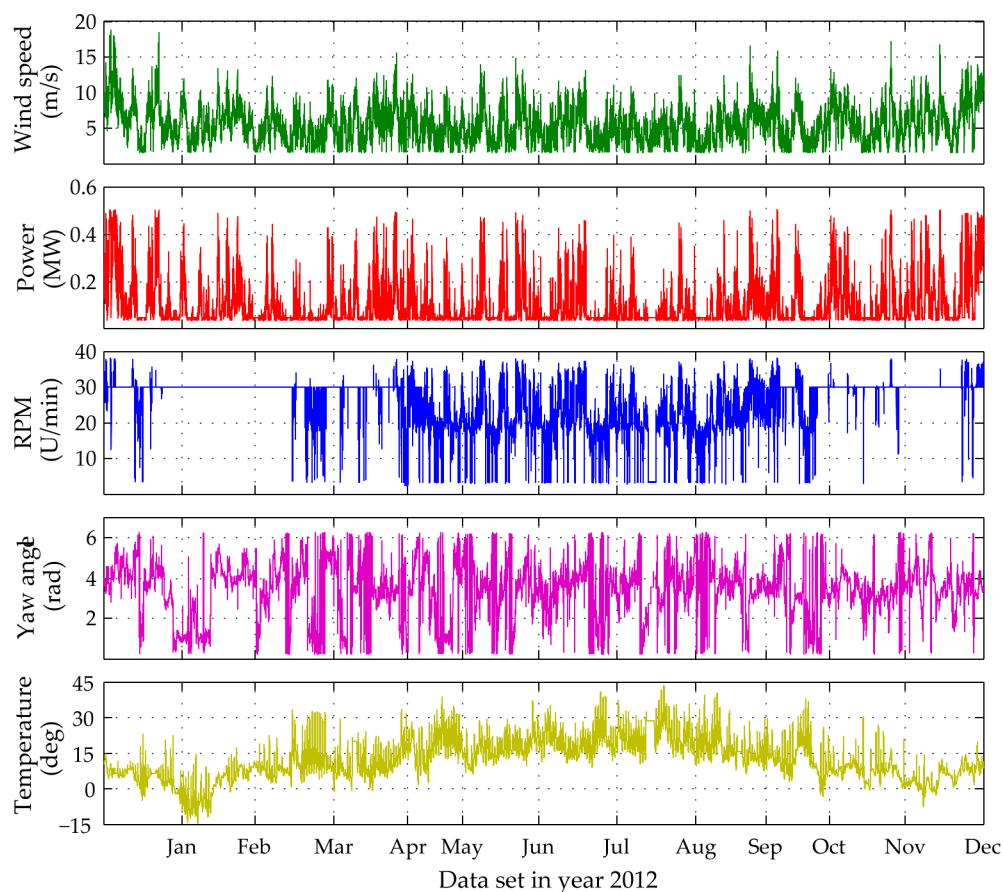


Figure 18. Time history plots of the selected PCE input variables for 49,956 datasets corresponding to complete year 2012.

In Figures 19–21 the validation sets of the estimated PCE residual demonstrate the workings of a consistent Diagnostic Index (DI) able to directly illustrate unfamiliar fluctuations in the EOP. Following

statistical outlier analysis the index values which exceed the thresholds of $\pm 3\text{std}$ can be linked to new data ranges of measured influencing agents.

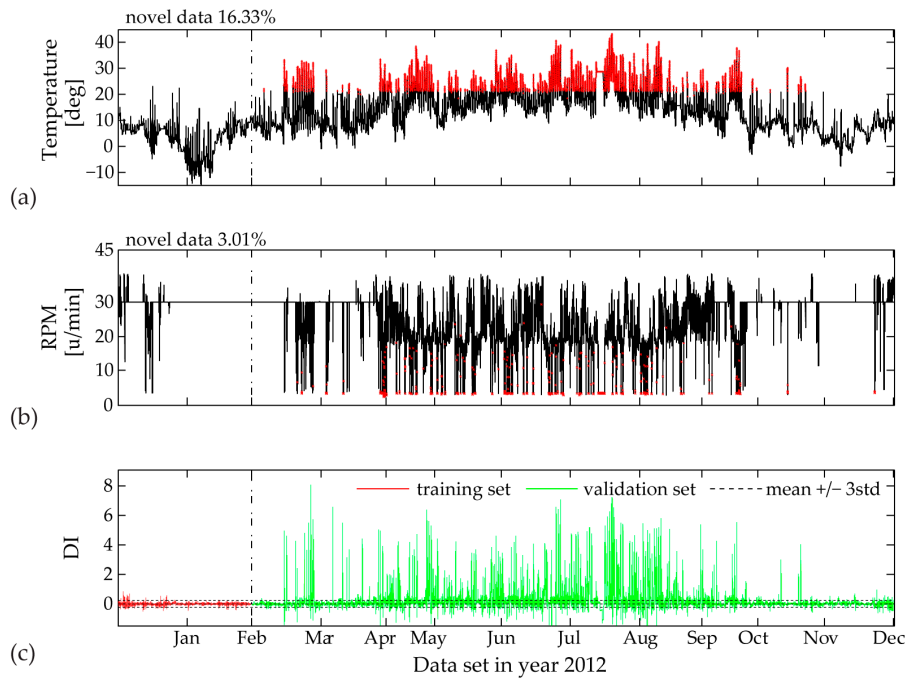


Figure 19. Two-month training set: (a) Identified novel data (red points) within time history of 10-min mean values of measured temperature; (b) Identified novel data (red points) within time history of 10-min mean values of measured RPM; (c) Diagnostic Index for the complete data sets of year 2012.

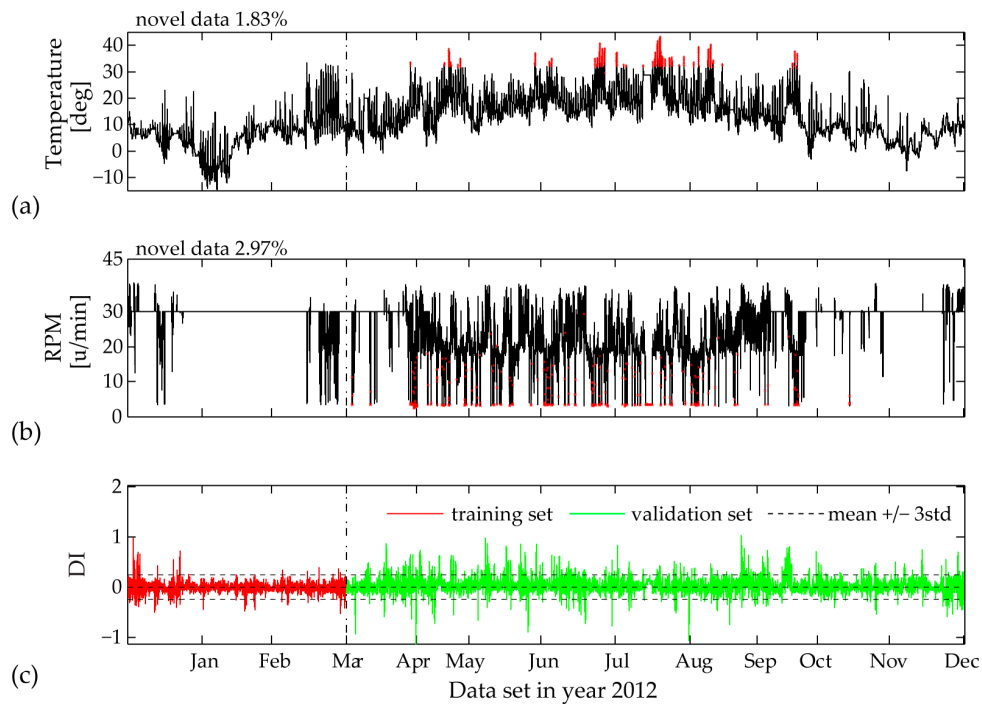


Figure 20. Three-month training set: (a) Identified novel data (red points) within time history of 10-min mean values of measured temperature; (b) Identified novel data (red points) within time history of 10-min mean values of measured RPM; (c) Diagnostic Index for the complete data sets of year 2012.

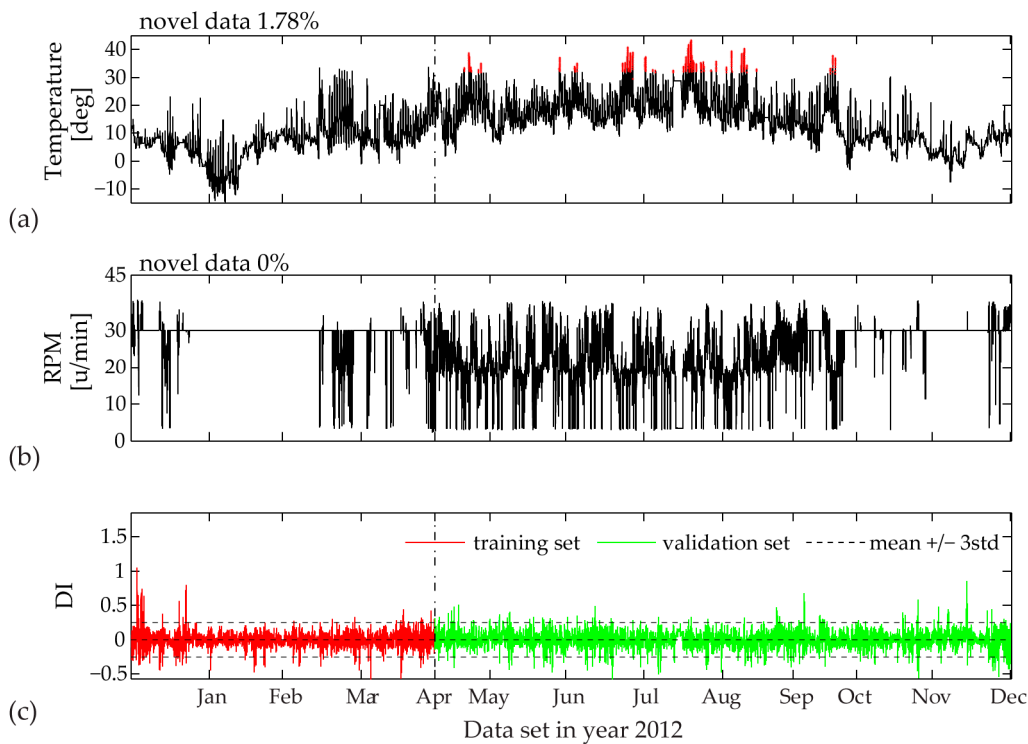


Figure 21. Four-month training set: (a) Identified novel data (red points) within time history of 10-min mean values of measured temperature; (b) Identified novel data (red points) within time history of 10-min mean values of measured RPM; (c) Diagnostic Index for the complete data sets of year 2012.

For this purpose, the well-known Mahalanobis Distance (MD) method is herein applied on the input data time histories. For a p -dimensional multivariate sample x_1, \dots, x_n , the MD is defined as, [42]:

$$MD_i = \sqrt{(x_i - t_{tr})^T C_{tr}^{-1} (x_i - t_{tr})} \quad \text{for } i = 1, \dots, n \quad (23)$$

where t_{tr} is the arithmetic mean and C_{tr} is the sample covariance matrix, both estimated for an input data set corresponding to a certain training period length. For normally distributed multivariate data, the MD values result approximately chi-square distributed with p degrees of freedom (χ_p^2), while the MD threshold is usually set as a certain quantile of χ_p^2 . However, an adaptive method that takes into account the actual empirical chi-square distribution function of the estimated MD (instead of a fixed quantile) is herein applied [42]. In this way thresholds (adjusted quantiles) are estimated for corresponding training sets for each of the five SCADA parameters. The x_i samples from the validation sets which have MD beyond the predefined threshold are defined as novelties.

False triggering alarms due to incomplete training set are mostly related to input parameters characterized by unpredictable range of values, i.e., temperature measurements and wind velocity. In this case for a two-month training period, the univariate MD plot of the temperature time history shows the highest percentage of 16.33% novel data (Figure 19). The DI distribution pattern, plotted in the same figure, is clearly linked to the new upcoming values between months March and November. If we further increase the training set to a three-month period (Figure 20), the MD outlier percentage for measured temperature decreases. Correspondingly, the DI becomes significantly reduced, although not yet within the ± 3 std limit values. In Figure 20, for the same three-month training period, an interesting observation is the contribution of the rotor rotation to novel data (2.97%), which is completely reduced to 0% after four months of model training (Figure 21). The reason for this occurrence is the relatively flat

range of RPM values between mid-January and mid-March, most probably due to a sensor malfunction, which later on is repaired.

The MD plots (Figures 20 and 21) demonstrate that for a full incorporation of various fluctuations in the measured temperature, the training period should be even extended to eight months. Additional univariate MD analysis for the measured wind velocity, power production and yaw angle for the complete year 2012 reveals no new data points within the validation sets of these inputs, even for a training period lengths of one to two months (January and February). The estimated minimum PCE model training set lengths for capturing the complete scope of the separate SCADA variables is summarized in Table 3.

Table 3. Minimum necessary training lengths for capturing the complete SCADA data value ranges.

Power	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Wind	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Yaw angle	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Rotor RPM	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Temperature	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug

5. Future Work

Statistical pattern recognition concepts [28] represent a favorable detection approach for SHM data-driven diagnostic strategies based on measured responses of structures in a healthy, or unknown baseline condition. Recent studies [15,43] reported successful application of control chart based strategies for monitoring structural changes in WT systems. Dervilis et al. in [27] investigated the applicability of various neural networks techniques towards online accurate monitoring of WT blades under continuous fatigue loading. In the field of bridge monitoring, Dervilis et al. in [44] developed residual outlier maps which reveal different spatial manifestation of detected outliers related to changing environmental and operational conditions, versus structural damages.

In this context, for completing the scope of the proposed tool, future work is targeted at testing the assessment capabilities of the PCE-SPTARMA method by expanding the validation data range to subsequent recorded years. Through long-term tracking of the estimated DI, typical operating regimes to specific structural behavioral patterns of the WT system will be interrelated. The final goal lies in development of an algorithm that separates benign pattern distortions, linked to EOP fluctuations, from malignant pattern alterations due to actual structural damage or system malfunction. Accordingly, the diagnostic tool triggers a model retraining or appropriate notification procedure.

6. Conclusions

Europe's growing needs for energy savings have placed WT structures in focus, albeit the tackling of complexities of these structures has long occupied both the academic and industrial communities. At the same time, the recent technological developments in sensors and monitoring solutions motivate the enhancement of the currently adopted SCADA stream for WTs with information of structural nature, able to guide operators in the management of these assets.

In materializing such a goal, it is important to account for the non-stationary response of operating WTs and the temporal variability of their identified modal parameters. In order to develop comprehensive dynamic models both issues need to be addressed. By merging environmental and operational variables with a time-varying model of vibrational response, the proposed bi-component tool serves as the first step towards automated condition assessment.

Successful implementation of the devised strategy on two distinct operating WT structures in Germany for different settings, namely, different sites, turbine types, SCADA duration, and different influencing agents, verifies the robustness of the approach. Furthermore, the outcomes of the complete one-year tracking of the obtained diagnostic index demonstrate the potential of the proposed tool for

incorporation within a holistic SHM damage detection framework, further extended via statistical pattern recognition methods, to be explored in a next step. Another interesting perspective lies in fusing the surrogate models developed as part of this diagnostic framework with a simulation-based design procedure, where simulations of the SP-TARMA model for diverse loadings may be used for predictions on processes such as fatigue, equivalent damage loads, etc.

Acknowledgments: The authors would like to acknowledge the support of COST Action TU1402 on “Quantifying the Value of Structural Health Monitoring” and the support received from the Ruhr-University Bochum Research School. E. Chatzi would further like to acknowledge the support of ERC Starting Grant #67984 WINDMIL on the topic of “Smart Monitoring, Inspection and Life-Cycle Assessment of Wind Turbines”.

Author Contributions: Simona Bogoevska has performed the main body of work in this paper by analyzing the data and carrying out the major part of scientific writing. Minas Spiridonakos and Eleni Chatzi provided invaluable guidance and support in the treatment of the data. Rudiger Höffer and Eleni Chatzi ensured access to and critical assessment of the employed data sets. Eleni Chatzi and Elena Dumova- Jovanoska provided further support in the scientific writing.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Carne, T.G.; James, G.H. The inception of OMA in the development of modal testing technology for wind turbines. *Mech. Syst. Signal Process.* **2010**, *24*, 1213–1226. [[CrossRef](#)]
2. WindEurope. *Wind in Power: 2015 European Statistics*; WindEurope: Brussels, Belgium, 2016.
3. Ciang, C.C.; Lee, J.-R.; Bang, H.-J. Structural health monitoring for a wind turbine system: A review of damage detection methods. *Meas. Sci. Technol.* **2008**, *19*, 1–20. [[CrossRef](#)]
4. Hameed, Z.; Hong, Y.S.; Cho, Y.M.; Ahn, S.H.; Song, C.K. Condition monitoring and fault detection of wind turbines and related algorithms: A review. *Renew. Sustain. Energy Rev.* **2009**, *13*, 1–39. [[CrossRef](#)]
5. Yang, B.; Sun, D. Testing, inspecting and monitoring technologies for wind turbine blades: A survey. *Renew. Sustain. Energy Rev.* **2013**, *22*, 515–526. [[CrossRef](#)]
6. Thöns, S. Monitoring Based Condition Assessment of Offshore Wind Turbine Support Structures. Ph.D. Thesis, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland, 2012.
7. Chauhan, S.; Tcherniak, D.; Basurko, J.; Salgado, O.; Urresti, I.; Carcangiu, C.E.; Rossetti, M. Operational modal analysis of operating wind turbines: application to measured data. In Proceedings of the 29th International Modal Analysis Conference (IMAC 2011), Jacksonville, FL, USA, 31 January–3 February 2011; Volume 5, pp. 65–81.
8. Ozbek, M.; Rixen, D.J. Optical measurements and operational modal analysis on a large wind turbine: Lessons learned. In Proceedings of the 29th International Modal Analysis Conference (IMAC 2011), Jacksonville, FL, USA, 31 January–3 February 2011; Volume 5, pp. 257–276.
9. Hansen, M.H.; Thomsen, K.; Fuglsang, P.; Knudsen, T. Two methods for estimating aeroelastic damping of operational wind turbine modes from experiments. *Wind Energy* **2006**, *9*, 179–191. [[CrossRef](#)]
10. Devriendt, C.; Magalhães, F.; Weijtjens, W.; De Sitter, G.; Cunha, L.; Guillaume, P. Structural health monitoring of offshore wind turbines using automated operational modal analysis. *Struct. Heal. Monit.* **2014**, *13*, 644–659. [[CrossRef](#)]
11. Iliopoulos, A.; Devriendt, C.; Guillaume, P.; Van Hemelrijck, D. Continuous fatigue assessment of an offshore wind turbine using a limited number of vibration sensors. In Proceedings of the 7th European Workshop on Structural Health Monitoring (EWSHM 2014), Nantes, France, 8–11 July 2014.
12. Maes, K.; Iliopoulos, A.; Weijtjens, W.; Devriendt, C.; Lombaert, G. Dynamic strain estimation for fatigue assessment of an offshore monopile wind turbine using filtering and modal expansion algorithms. *Mech. Syst. Signal Process.* **2016**, *76–77*, 592–611. [[CrossRef](#)]
13. van der Male, P.; Lourens, E. Operational vibration-based response estimation for offshore wind lattice structures. In Proceedings of the 33rd International Modal Analysis Conference (IMAC 2015), Orlando, FL, USA, 2–5 February 2015; Volume 7, pp. 83–96.

14. Oliveira, G.; Magalhães, F.; Cunha, Á.; Caetano, E. Automated modal tracking and fatigue assessment of a wind turbine based on continuous dynamic monitoring. In Proceedings of the 6th International Conference on Experimental Vibration Analysis for Civil Engineering Structures (EVACES 2015), Dubendorf, Switzerland, 19–21 October 2015.
15. Weijtjens, W.; Verbelen, T.; De Sitter, G.; Devriendt, C. Foundation structural health monitoring of an offshore wind turbine—A full-scale case study. *Struct. Heal. Monit.* **2015**, *15*, 389–402. [[CrossRef](#)]
16. Tcherniak, D.; Chauhan, S.; Hansen, M.H. Applicability limits of operational modal analysis to operational wind turbines. In Proceedings of the 28th International Modal Analysis Conference (IMAC 2010), Jacksonville, FL, USA, 1–4 February 2010.
17. Avendaño-Valencia, L.D.; Fassois, S.D. Stationary and non-stationary random vibration modelling and analysis for an operating wind turbine. *Mech. Syst. Signal Process.* **2013**, *47*, 263–285. [[CrossRef](#)]
18. Avendaño-Valencia, L.D.; Chatzi, E.N.; Spiridonakos, M.D. Surrogate modeling of nonstationary systems with uncertain properties. In Proceedings of the 25th European Safety and Reliability Conference (ESREL 2015), Zürich, Switzerland, 7–10 September 2015.
19. Ozbek, M.; Meng, F.; Rixen, D.J. Challenges in testing and monitoring the in-operation vibration characteristics of wind turbines. *Mech. Syst. Signal Process.* **2013**, *41*, 649–666. [[CrossRef](#)]
20. Oliveira, G. Vibration-Based Structural Health Monitoring of Wind Turbines. Ph.D. Thesis, University of Porto, Porto, Portugal, 2016.
21. Hu, W.-H.; Thöns, S.; Rohrmann, R.G.; Said, S.; Rücker, W. Vibration-based structural health monitoring of a wind turbine system Part II: Environmental/operational effects on dynamic properties. *Eng. Struct.* **2015**, *89*, 273–290. [[CrossRef](#)]
22. Spiridonakos, M.; Ou, Y.; Chatzi, E.; Romberg, M. Wind turbines structural identification framework for the representation of both short-and long-term variability. In Proceedings of the 7th International Conference on Structural Health Monitoring of Intelligent Infrastructure (SHMII 2015), Torino, Italy, 1–3 July 2015.
23. Spiridonakos, M.; Chatzi, E. Polynomial chaos expansion models for SHM under environmental variability. In Proceedings of the 9th International Conference on Structural Dynamics (EURODYN 2014), Porto, Portugal, 30 June–2 July 2014.
24. Bogoevska, S.; Spiridonakos, M.; Chatzi, E.; Dumova-Jovanoska, E.; Höffer, R. A novel bi-component structural health monitoring strategy for deriving global models of operational wind turbines. In Proceedings of the 8th European Workshop on Structural Health Monitoring (EWSHM 2016), Bilbao, Spain, 5–8 July 2016.
25. Bogoevska, S.; Spiridonakos, M.; Chatzi, E.; Jovanoska, E.D.; Höffer, R. A data-driven framework for comprehensive identification of operational wind turbines under uncertainty. In Proceedings of the 5th International Conference on Uncertainty in Structural Dynamics (USD 2016), Leuven, Belgium, 19–21 September 2016.
26. Avendaño-Valencia, L.D.; Chatzi, E.N. Non-stationary random coefficient models for vibration-based SHM in structures influenced by strong operational and environmental variability. In Proceedings of the 10th International Workshop on Structural Health Monitoring (IWSHM 2015), Stanford, CA, USA, 1–3 September 2015.
27. Dervilis, N.; Choi, M.; Taylor, S.G.; Barthorpe, R.J.; Park, G.; Farrar, C.R.; Worden, K. On damage diagnosis for a wind turbine blade using pattern recognition. *J. Sound Vib.* **2014**, *333*, 1833–1850. [[CrossRef](#)]
28. Martinez-Luengo, M.; Kolios, A.; Wang, L. Structural health monitoring of offshore wind turbines: A review through the statistical pattern recognition paradigm. *Renew. Sustain. Energy Rev.* **2016**, *64*, 91–105. [[CrossRef](#)]
29. Ou, Y.; Dertimanis, V.; Chatzi, E. Experimental Damage Detection of a wind turbine blade under varying operational conditions. In Proceedings of the 5th International Conference on Uncertainty in Structural Dynamics (USD 2016), Leuven, Belgium, 19–21 September 2016.
30. Spiridonakos, M.D.; Poulimenos, A.G.; Fassois, S.D. Output-only identification and dynamic analysis of time-varying mechanical structures under random excitation: A comparative assessment of parametric methods. *J. Sound Vib.* **2009**, *329*, 768–785. [[CrossRef](#)]
31. Akaike, H.; Kitagawa, G. *The Practice of Time Series Analysis (Ch. 11)*; Springer: New York, NY, USA, 1999.
32. Kitagawa, G.; Gersch, W. *Smoothness Priors Analysis of Time Series (Ch. 5)*; Springer: New York, NY, USA, 1996.
33. Le Maître, O.P.; Knio, O.M. *Spectral Methods for Uncertainty Quantification*; Springer: Dordrecht, The Netherlands, 2010.
34. Wiener, N. The homogeneous chaos. *Am. J. Math.* **1938**, *60*, 897–936. [[CrossRef](#)]

35. Cameron, R.H.; Martin, W.T. The orthogonal development of non-linear functionals in series of fourier-hermite. *Ann. Math.* **1947**, *48*, 385–392. [[CrossRef](#)]
36. Xiu, D.; Karniadakis, G.E. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.* **2002**, *24*, 619–664. [[CrossRef](#)]
37. Askey, R.; Wilson, J. *Some Basic Hypergeometric Polynomials That Generalize Jacobi Polynomials*; American Mathematical Soc.: Providence, RI, USA, 1985.
38. Blatman, G.; Sudret, B. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Probabilistic Eng. Mech.* **2010**, *25*, 183–197. [[CrossRef](#)]
39. Oladyshkin, S.; Nowak, W. Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. *Reliab. Eng. Syst. Saf.* **2012**, *106*, 179–190. [[CrossRef](#)]
40. Hyvärinen, A.; Oja, E. Independent component analysis: Algorithms and applications. *Neural Netw.* **2000**, *13*, 411–430. [[CrossRef](#)]
41. Lachmann, S.; Baitsch, M.; Hartmann, D.; Höffer, R. Structural lifetime prediction for wind energy converters based on health monitoring and system identification. In Proceedings of the 5th European & African Conference on Wind Engineering (EACWE 2009), Florence, Italy, 19–23 July 2009.
42. Filzmoser, P.; Garrett, R.G.; Reimann, C. *Multivariate Outlier Detection in Exploration Geochemistry*; Technical Report TS 03-5; Department of Statistics, Vienna University of Technology: Vienna, Austria, 2003.
43. Oliveira, G.; Magalhães, F.; Cunha, Á.; Caetano, E. Optimization of a wind turbine vibration-based SHM system. In Proceedings of the 8th European Workshop on Structural Health Monitoring (EWSHM 2016), Bilbao, Spain, 5–8 July 2016.
44. Dervilis, N.; Worden, K.; Cross, E.J. On robust regression analysis as a means of exploring environmental and operational conditions for SHM data. *J. Sound Vib.* **2015**, *347*, 279–296. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).