Compensation of Thermal Effects on Tiltmeter Measurements With Moving Least Squares

Gianmarco Battista, Stefano Pavoni, and Marcello Vanali

***Abstract*— This article describes a method for compensating thermal effects on tiltmeter (TM) readings using operating temperature measurements. This practice is necessary when the contribution of thermal effects is comparable with the measurand inclination, which is typical of monitoring applications. In such contexts, several temperature-related phenomena take place concurrently and a customized compensation model is necessary. A tailored moving least square (MLS) approach is presented. This formulation is able to separate the contribution of the unknown measurand signal from the contribution of influence quantities, considered as exogenous inputs. Moreover, this method returns the model coefficients over time. However, the uncertainty in the temperature measurements, used as regressors, may induce bias in coefficient estimates. Therefore, a strategy for a posteriori evaluation and correction of the coefficients is proposed and validated with the Monte Carlo method (MCM). The effectiveness of this method is illustrated on a real case, consisting of the inclination monitoring of a wind-turbine tower.**

***Index Terms*— Measurement error models, Monte Carlo method (MCM), moving least squares (MLSs), structural health monitoring (SHM), thermal effects compensation, tiltmeter (TM).**

I. Introduction

THE structural health monitoring (SHM) is a branch of diagnostics that aims at long-term monitoring of the health status of structures [1], [2]. This is achieved by measuring physical quantities, related both to the structure and to the environment, that may prefigure a damage or reveal anomalies in operating conditions. Each SHM application requires a specific measurement system design and data analysis approach. In fact, how to convert data into information on the health of the structure is still an open question to be addressed by the research community [3], [4]. Another crucial aspect concerns the identification of the environmental physical quantities that influence the measurements of the structural parameters of interest [5]. Due to nonnegligible variations in environmental conditions, it is often necessary to "normalize" the structural response and compensate for interfering inputs. This enables comparison of monitored parameters between different conditions.

Inclination sensors are widely adopted in the field of SHM, using various technologies and configurations, such as wireless MEMS tilt sensors [6], [7] or wired fiber optic inclinometers [8]. O'Leary and Harker [9] developed a framework to measure the deflection of arbitrary structures, using the data from an array of tiltmeters (TMs) and an approach based on regularized least squares (LSs). They aim to retrieve the deflection curves from their local derivatives measured by the sensors. Additionally, the same authors addressed the issue of uncertainty on boundary conditions as well [10]. In many monitoring applications, the meaningful part of the inclination signal may be comparable or even smaller than the other interfering contributions. When this happens, a compensation is mandatory to reveal the information being sought and reach an adequate measurement accuracy. For example, Yang et al. [11] developed an approach to compensate the effects of vibrations in a drilling process on inclination measurements near the bit. The main interest of this work concerns the effects of the thermal effects on inclination sensor, as it is one of the influence quantities with the greatest impact on inclination measurements. Despite temperature-insensitive tilt sensors exist [12], [13], it is common to deal with sensors having a zero drift proportional to its operating temperature. In such a case, it is necessary to correct the angle measurement a posteriori. A possible approach is to develop a temperature compensation model in the TM calibration phase [14], using, for instance, a dedicated test rig [15] and a temperature reference. Although this option leads to great accuracy, it is not always feasible due to its greater complexity and cost. An example of this approach, adopted for SHM, is provided in [16]. A first stage of sensors characterization is performed in laboratory with a climatic chamber, and then, the temperature compensation is applied for on-site measurements on a real structure [17]. Nevertheless, in many monitoring applications, several temperature-related phenomena can take place concurrently. Some of these may depend on the experimental measurement context, rather than the sensor, and alter the measurand itself. Therefore, a laboratory characterization phase of these systematic effects is not always possible, meaning that the exact magnitude is unknown a priori. Consequently, to obtain a meaningful measure of inclination, these effects must also be compensated for. Despite its relevance, this topic is not addressed in the literature, at the best of authors' knowledge.

Manuscript received 7 July 2023; revised 20 December 2023; accepted 29 December 2023. Date of publication 23 February 2024; date of current version 1 March 2024. This work was realised in the project funded under National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4—Call for tender No. 3138 of 16/12/2021 of Italian Ministry of University and Research funded by the European Union—NextGenerationEU. Project code CN00000023, Concession Decree No. 1033 of 17/06/2022 adopted by the Italian Ministry of University and Research, CUP D93C22000400001, "Sustainable Mobility Center" (CNMS). The Associate Editor coordinating the review process was Dr. Boby George. *(Corresponding author: Gianmarco Battista.)*

The authors are with the Department of Engineering and Architecture, University of Parma, 43124 Parma, Italy (e-mail: gianmarco.battista@unipr.it).

Digital Object Identifier 10.1109/TIM.2024.3369139

This article presents a framework to compensate for systematic effects on inclination measurement, with particular focus on thermal effects. The proposed approach is compatible with typical SHM applications. The major contributions of this article are as follows.

1) The proposed technique performs a compensation using a measurement model customized to the specific application. It can be applied directly using the operational signals from the TM, together with the signals representing the phenomena to be compensated. The latter are considered the inputs to the measurement model.
2) A tailored version of moving LSs (MLSs) [18], [19] is formulated in order to estimate the unknown model coefficients over time. This method uses local regressions to split the contribution of influencing input quantities from the measurand signal.
3) A strategy for assessing and correcting the bias on the model coefficients caused by the uncertainty of the input quantities. The bias correction requires a realistic estimate of the uncertainty, in addition to the operational input signals. Its validity is demonstrated by means of the Monte Carlo method (MCM).

The theoretical basis of the technique is illustrated, and the effectiveness of the proposed approach is evidenced by its application to a real case, which consists of monitoring the inclination of a wind-turbine tower.

This article is structured as follows. Section II provides an overview of the monitoring test case, the measurement setup, and the measurement model. Section III illustrates the tailored MLS formulation for estimation of coefficients and reconstruction of inclination signal. In Section IV, compensation for thermal effects is applied to the experimental data. Two models are compared, highlighting how a proper compensation model is crucial. Section V describes the strategy for bias evaluation and correction. In Section VI, the MCM is exploited to characterize the effectiveness of bias correction and to propagate the instrumental uncertainty through the measurement algorithm. Section VII shows the comparison between the results with and without the bias correction. Finally, Section VIII draws the main conclusions about the approach proposed in this article.

II. Inclination Monitoring of Wind-Turbine Tower: Problem Definition and Measurement Setup

The case study presented in this article concerns the inclination monitoring of an onshore wind-turbine tower. The monitoring became necessary due to ground movements below the tower foundations, which induce a progressive inclination of the structure. The turbine is stopped for safety reasons and does not produce power during the whole monitoring period. The period under analysis is from February 2022 to April 2023. The objective is to monitor the subsidence state over time to identify dangerous trends. Therefore, the measurand is the change in the inclination of the tower along the direction of subsidence with respect to the initial state. The target accuracy is 0.01° to match the order of magnitude of expected variations. The most relevant interfering inputs identified are

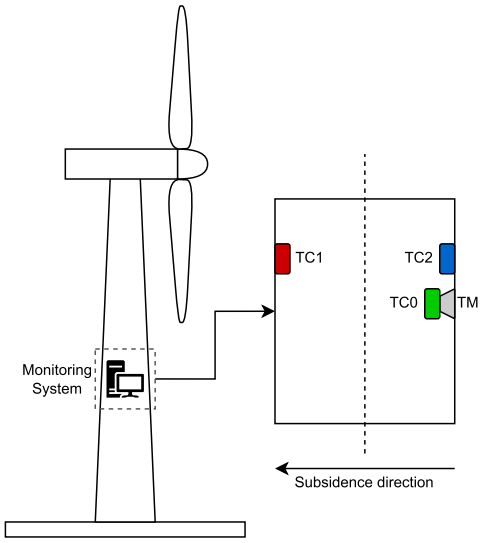

Fig. 1. Scheme of the sensors installed inside the wind-turbine tower: the TM, the thermocouple to measure the sensor operating temperature (TC0), and two thermocouples to measure the surface temperature on opposite sides of the tower in the subsidence direction (TC1 and TC2).

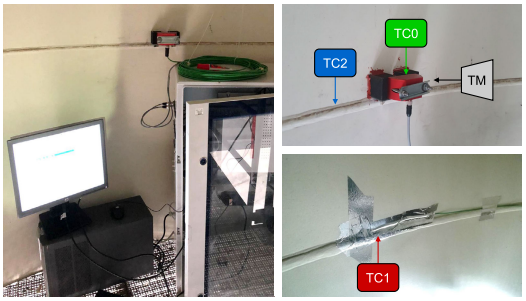

Fig. 2. Pictures of the location and installation of the sensors inside the wind-turbine tower. TM installed on the inner wall of the tower (left). Measurement locations of TC0 and TC2 thermocouples with respect to the TM (no picture available with the three sensors installed) (top right). Installation of thermocouple TC1. The same installation method is used for TC0 and TC2 (bottom right).

two thermal effects: the typical thermal zero drift of the sensor and the thermal bending of the tower due to temperature gradients in the structure. In fact, the incoming solar radiation on the cylindrical structure causes a temperature gradient which consequently induces a nonuniform thermal expansion of different sides of the tower. This causes a thermal bending, which depends on amount and direction of solar radiation. Both effects must be compensated, since they are the prevalent contributions on the sensor readings.

A. Measurement Setup

The tower has been instrumented with a TM on its inner surface, at a height of 6 m from the basement, together with three temperature sensors. The scheme of the measurement setup is depicted in Fig. 1. Some pictures of the measurement setup inside the tower are provided in Fig. 2. The inclination sensor is the Sisgeo 0S542MA1002 bi-axial TM, and one of the two sensing axes is aligned with the direction of subsidence, which is the only one considered in this work. The sensor has a measuring range of $\pm 10°$ and a resolution

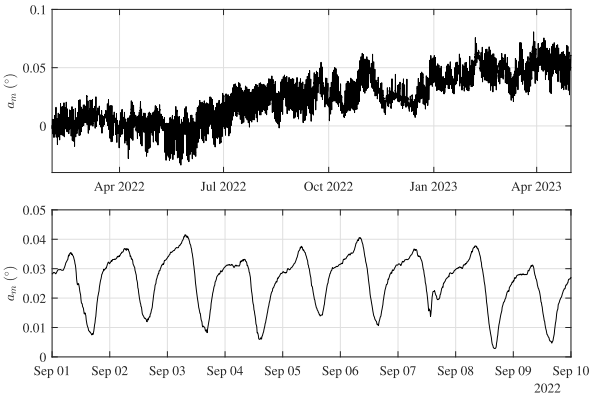

Fig. 3. Monitoring inclination signal from February 2022 to April 2023. The focus on a portion of the signal highlights the daily patterns induced by thermal effects on inclination measurements.

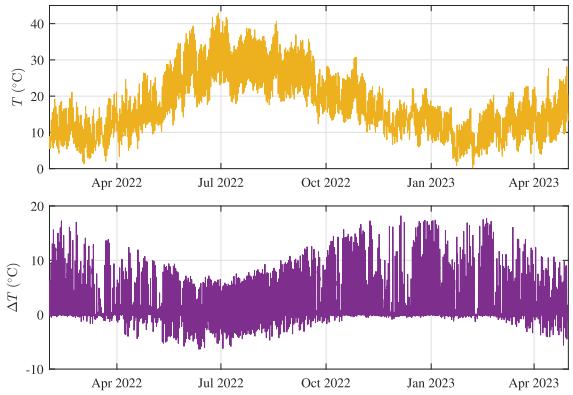

Fig. 4. Monitoring temperature signals, from February 2022 to April 2023, representing the thermal effects to be compensated.

of 0.01%FS (0.001°). The mechanical bandwidth is 18 Hz, which is more than adequate for quasi-static measurements required in this work. The technical specifications also report an accuracy of 0.020°, expressed as maximum permitted error (MPE), and an offset temperature dependence of ±0.003°/°C. This causes the zero temperature drift, but neither the sign nor the exact value of the "temperature coefficient" is known a priori for each sensor. The analog signal from the TM is a 4–20 mA current loop signal and is acquired by means of a National Instruments module NI 9203. The operating temperature of the TM is measured by a thermocouple (TC0). Two further thermocouples (TC1 and TC2) are installed to measure the temperature of the inner surface of the tower. The probes are positioned on opposite sides of the tower in the direction of subsidence. All thermocouples are type K and Class 1. According to the standard [20] *IEC 60584-1:2013*, the tolerance for them is ±1.5 °C, considering their operating range temperature (approximately from −10 °C to +50 °C). Thermocouple signals are directly acquired by National Instruments module NI 9211. No additional hardware is required for signal conditioning, since this module is designed specifically for thermocouples.

B. Experimental Data

The raw signal from each sensor is sampled at a frequency of 40 Hz. Then, the mean values over 10 min are computed and stored, so that 144 samples per day are available for each signal. This reduces the standard deviation given by the random components by a factor of $1/(10 \cdot 60 \cdot 40)^{1/2} \approx 6.5 \cdot 10^{-3}$. The time series so obtained are considered as the output of the sensors and are used for all analyses conducted in this work. In the following, the generic ith sampling time on the discrete time axis is denoted by t_i. The absolute angle measured by the TM is determined by the sensor fastening and its bias, in addition to the tower inclination. Since the interest is on the *variation of inclination*, the initial offset is subtracted from the entire TM signal. The offset is computed as the mean value of inclination in the first day of monitoring. The resulting signal for the monitoring task, indicated by $a_m(t_i)$, is depicted in Fig. 3. The accuracy of these values can be considered higher than the TM specifications. The TM measurements depend on a combination of systematic and random effects. The random component is reduced by the averaging process described above. The systematic component that is dependent on the input angle (e.g., nonlinearity) can be considered nearly constant during the entire monitoring period. This is because the maximum change of inclination is in the order of 0.1°, corresponding to 1%FS. Since $a_m(t_i)$ is the variation with respect to the initial value, the residual bias is considered negligible with respect to the target accuracy of this monitoring. For these reasons, in this application, the main contribution to the measurement accuracy is given by the systematic effects caused by thermal phenomena. By observing $a_m(t_i)$, an oscillatory pattern on daily basis is evident. This does not represent the actual movements of the tower, but it is the result of thermal effects that requires a proper compensation. The temperature signals used for this purpose are depicted in Fig. 4. The operating temperature of the TM $T(t_i)$ is measured by the thermocouple TC0. The temperature difference $\Delta T(t_i) = T_1(t_i) - T_2(t_i)$ represents the thermal gradient between opposite sides of the structure, along the subsidence direction. These are measured by TC1 and TC2, indicated, respectively, as $T_1(t)$ and $T_2(t)$.

C. Measurement Models

The actual tower inclination signal $a_s(t)$, i.e., the measurand, can be retrieved from measured signal $a_m(t_i)$ using a measurement model. A typical model for TM measurement to compensate the temperature drift is

$$a_m(t) = a_s(t) + T(t)a_T \tag{1}$$

where a_T represents the "temperature coefficient" of the sensor thermal drift. As mentioned earlier, the value of a_T is unknown and different from sensor to sensor. The measurement model (1) needs be customized to account for the thermal bending of the tower as well. The assumption is that this contribution on the inclinometer signal is proportional to ΔT. Consequently, the model becomes

$$a_m(t) = a_s(t) + T(t)a_T + \Delta T(t)a_{\Delta T} \tag{2}$$

where $a_{\Delta T}$ is the "thermal bending coefficient" pertaining to this particular installation. This model parameter is also

unknown and must be gathered from operational data. In both models, dynamic effects and lagged signals are intentionally neglected. The model (2) is adequate for monitoring purposes, since the main phenomena are included. The adequacy of the measurement model is assessed according to the guidelines provided by the GUM [21]. Ahead in this article, a detailed description is given of the measurement algorithm. The aim is to extract the inclination signal $a_s(t_i)$ and the model coefficients from the TM readings $a_m(t_i)$ and temperature measurements $T(t_i)$ and $\Delta T(t_i)$.

III. MLS Approach for Exogenous Inputs Compensation

The typical MLS approach [18], [19] is often used for curve and surface fitting and relies on the concept of "local regression" [22]. The same approach is also widely adopted for smoothing data as in Savitzky–Golay filter [23] or locally weighted scatterplot smoothing (LOWESS) [24], where data is locally fit with a set of polynomial functions. The MLS formulation proposed here relies on the same basis, but it is tailored to decompose the regressand signal $y(t)$, i.e., the explained variable, in two main contributions: the underlying signal $y_s(t)$ and the exogenous inputs $y_{\text{ex}}(t)$

$$y(t) = y_s(t) + y_{\text{ex}}(t)\,. \tag{3}$$

The term "exogenous" indicates that the value of a variable is determined outside the model and is imposed on it. This is what happens for influence quantities on TM readings. In this formulation, time t is considered as the independent variable to scroll the regression window of MLS, analogously to the moving average. By means of local regressions, the two macrocontributions can be estimated, thus recovering an approximation of the underlying signal and compensating for the exogenous inputs. The local regressions are performed with the weighted LSs (WLSs), which is an extension of the common OLS [25].

A. WLS Regression

A brief summary of WLS regression and its properties is reported here. Consider a linear model in the following form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \tag{4}$$

The column vector $\mathbf{y}$ contains n observations of the variable to explain, $\mathbf{X}$ is an $n \times m$ matrix of regressors (named also explanatory variables), and $\boldsymbol{\beta}$ is the vector of m model parameters. The vector $\boldsymbol{\varepsilon}$ contains the random error terms of each observation that are independent and identically distributed with zero mean and variance σ_ε^2. A diagonal $n \times n$ weighting matrix $\mathbf{W}$, having nonnegative entries on the main diagonal, is defined to assign different weights to each observation. The WLS estimate $\mathbf{b}$ of the unknown parameters $\boldsymbol{\beta}$ is

$$\mathbf{b} = \left(\mathbf{X}^\top \mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}^\top \mathbf{W}\mathbf{y}\,. \tag{5}$$

The WLS regression problem can always be rewritten as an analogous OLS problem. For this reason, in this article, they are considered analogous when their general properties are discussed. When the conditions of the Gauss–Markov theorem are fulfilled, the OLS estimates of regression coefficients are unbiased and have the minimum variance, among all the unbiased linear estimators [26]. For this reason, it is often referred to as the "best linear unbiased estimator" (BLUE).

B. Tailored Formulation of MLSs

Let $y(t_i)$ be the observed time series of the regressand signal and $S = \{t_i\}$ be the set of the whole time domain for which the data are available. The subsets of local data S_k are determined by choosing a time window of length L and a set of generic time instants t_k as window centers. The subscript k denotes the index of each regression window, for a total of K windows. The subsets S_k are defined as

$$S_k = \{t_i \quad : \quad |t_i - t_k| \le L/2\}\,. \tag{6}$$

The choice of the window centers is to some degree arbitrary. Clearly, the combination of L and the set of t_k must be such that the union of all S_k covers the entire time domain S. Therefore, the spacing between consecutive window centers must not exceed the window length L. The samples of $y(t_i)$, corresponding to the sampling instants included in S_k, are briefly denoted by y_{ik} and arranged in a column vector indicated with $\mathbf{y}_k$.

In each window, regressand $\mathbf{y}_k$ is fit with two different sets of regressors. The approximation of the underlying signal is assigned to a local polynomial of degree p. This choice makes it possible to approximate $y_s(t)$ and its derivatives within the regression window [27]. Instead, the contribution of all influencing input quantities $y_{\text{ex}}(t)$ is fit by a set of exogenous explanatory variables $q_1(t), \ldots, q_e(t)$, where e is the number of inputs considered. These signals are required to compensate for all influencing undesired effects acting on the observed signal $y(t)$. The set of polynomial basis functions is conveniently formulated on a normalized local abscissa defined as

$$x_k(t) = (t - t_k)/L\,. \tag{7}$$

Therefore, it follows that $|x_k(t)| \le 1/2$ within the window. The polynomial regression matrix $\mathbf{X}_{p,k}$ of degree p for the kth regression window can be written as

$$\mathbf{X}_{p,k} = \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ 1 & x_{\text{ik}}^1 & \cdots & x_{\text{ik}}^p \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} \tag{8}$$

where x_{ik} briefly denotes the local abscissa defined in (7) calculated for each time t_i in the subset S_k. The other set of regressors consists of the observed signals of exogenous explanatory variables $q_1(t_i), \ldots, q_e(t_i)$. These fill the columns of the matrix $\mathbf{Q}_k$ defined for the kth regression window

$$\mathbf{Q}_k = \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ q_{1,\text{ik}} & q_{2,\text{ik}} & \cdots & q_{e,\text{ik}} \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} \tag{9}$$

where the subscript ik denotes the signal sample at the time t_i in the subset S_k. Combining the two sets of regressors, the full matrix for each time window becomes

$$\mathbf{X}_k = \left[\mathbf{X}_{p,k}\ \mathbf{Q}_k\right]. \tag{10}$$

Consequently, the WLS estimate of local regression coefficients $\mathbf{b}_k$ is

$$\mathbf{b}_k = \left(\mathbf{X}_k^\top \mathbf{W}_k \mathbf{X}_k\right)^{-1} \mathbf{X}_k^\top \mathbf{W}_k \mathbf{y}_k\ . \tag{11}$$

The diagonal entries of the weighting matrix $\mathbf{W}_k$ are computed according to an arbitrary weighting function $w(x) \geq 0$ evaluated on the local abscissa; hence, $\text{diag}(\mathbf{W}_k) = w(x_{\text{ik}})$. This aspect is discussed in detail ahead in this article. The local approximation of the signal within the regression window is

$$\hat{\mathbf{y}}_k = \mathbf{X}_k \mathbf{b}_k \tag{12}$$

which is the representation of $\mathbf{y}_k$ in terms of all the chosen explanatory variables included in $\mathbf{X}_k$. However, the $m \times 1$ vector $\mathbf{b}_k$ can be divided into two subvectors.

1) $\mathbf{b}_{p,k}$ contains $p+1$ coefficients of local polynomial approximation of the underlying signal.
2) $\mathbf{b}_{ex,k}$ contains e coefficients accounting for the effect of the exogenous inputs considered in the model.

It holds that $m = e + p + 1$. This is a key feature of the MLS approach proposed in this article. Under the hypothesis that $y_s(t)$ can be approximated by a local polynomial, the contribution of exogenous inputs can be separated from the underlying signal. The two parts can be calculated using the following expressions:

$$\hat{\mathbf{y}}_{s,k} = \mathbf{X}_{p,k} \mathbf{b}_{p,k} \tag{13a}$$

$$\hat{\mathbf{y}}_{ex,k} = \mathbf{Q}_k \mathbf{b}_{ex,k}\ . \tag{13b}$$

Once the local regressions are available, they need to be merged on the entire time domain S. In such way, the MLS global approximations of the observed signal $\hat{y}(t_i)$ and of its macrocomponents $\hat{y}_s(t_i)$ and $\hat{y}_{\text{ex}}(t_i)$ are obtained. To get the global approximation of the underlying signal $\hat{y}_s(t_i)$ is the main objective of this article. At this stage, the shape of the weighting function $w(x)$ plays a crucial role. On the one hand, it has the usual role that weights play in WLS regressions through the matrix $\mathbf{W}_k$ in (11). On the other hand, it determines how local approximations are merged with adjacent ones. Consequently, its impact is twofold because it controls the MLS results both locally and globally. As usual in MLS approaches, the weights are symmetric with respect to the window center t_k and are a nonincreasing function of the distance from it. The global MLS reconstruction of a signal is obtained as weighted average of all the overlapping local WLS approximations that are available for the ith sample. From (12), the global MLS approximation of $y(t_i)$ is computed as

$$\hat{y}(t_i) = \frac{\sum_{k=1}^{K} \hat{y}_{\text{ik}}\ w_{\text{ik}}}{\sum_{k=1}^{K} w_{\text{ik}}} \tag{14}$$

where w_{ik} means $w(x_{\text{ik}})$. The same can be done for $y_s(t_i)$ and $y_{\text{ex}}(t_i)$ using, respectively, (13a) and (13b). This method of merging local data is analogous to the Shepard's interpolation [28]

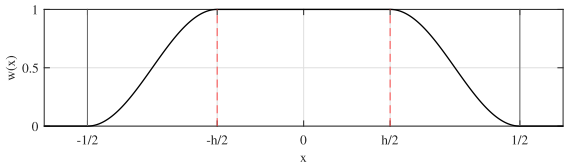

Fig. 5. Cosine tapered window function adopted as weighting for MLS fitting. The flatness ratio h controls the extent of the flat part versus extent of the cosine part.

A signal fit using this approach is continuous and differentiable only if the window function $w(x)$ has these properties [19]. An example is the "tricube" function proposed in [29]. For this application, the authors use the cosine tapered window function [30], also known as "Tukey window," that can be written as

$$w(x) = \begin{cases} 1, & \text{if } |x| < \dfrac{h}{2} \\ \dfrac{1}{2}\left(1 + \cos\left(\pi \dfrac{2|x| - h}{1 - h}\right)\right), & \text{if } \dfrac{h}{2} \leq |x| \leq \dfrac{1}{2} \\ 0, & \text{if } |x| > \dfrac{1}{2} \end{cases} \tag{15}$$

where h is the fraction of the window length L to be flat and hence is referred to as "flatness ratio". The function, depicted in Fig. 5, holds all the desired properties. An advantage is that the parameter h allows to easily control the shape of the function, giving more or less relevance to data close to the window edge. Moreover, the function thus defined has an identically null value outside the regression window.

In summary, this MLS formulation allows to estimate the underlying signal $\hat{y}_s(t_i)$, compensating the exogenous inputs considered in the model. As a further product, it also returns the coefficients $\mathbf{b}_{ex,k}$ representing the relationship between the exogenous inputs $q_1(t_i), \ldots, q_e(t_i)$ and the actual observed signal $y(t_i)$.

IV. Application of MLS to Experimental Data

The MLS for exogenous input compensation is suitable for estimating all unknowns in the models (1) and (2). The signal $a_m(t_i)$ measured by the TM is the observed variable to be explained, while the inclination signal of the tower $a_s(t_i)$ is the underlying signal to be retrieved with the MLS approach. The temperature $T(t_i)$ and the temperature difference $\Delta T(t_i)$ are the exogenous inputs to the model. It is important to note that it is not necessary to know the absolute value of the influence quantities, but it is important to know the "shape" of these signals, hence the variations with respect to an arbitrary offset. Here, the offset is 0 °C for both temperature signals. Each night/day thermal cycle is considered an "atom" for the following analyses; therefore, the sequence of window centers t_k is composed of time instants spaced one day apart that starts at 12:00 of the first day of monitoring. The window length is chosen to be $L = 15$ days, while the flatness ratio is $h = 1/L$. This means that each local regression is focused on the middle day of the window. The number of observations for the WLS

estimates is $n = 2160$, apart for the regression windows at the edges of the time domain. In such worst cases, the actual window length is at minimum $L/2$. The signal $a_s(t_i)$ is locally fit with a fourth-degree polynomial ($p = 4$). A general aspect must be pointed out regarding the choice of p. If the MLS polynomial approximation is adequate to fit the signal $a_s(t)$ in a given regression window, then the coefficients pertaining to exogenous inputs are not affected. Where the local polynomial regression is poor, it causes a bias on the other coefficients within the regression window.

In the following, the compensation of thermal effects is performed with both measurement model described in Section II. The aim is to highlight the impact of an incomplete compensation, with respect to a model customized to the application. The model (1) considers the TM operating temperature $T(t_i)$ as the only exogenous variable ($e = 1$). An estimate of the temperature coefficient a_T is returned for each regression window. Hence, the total number of parameters to be estimated for each local regression is $m = 6$. The model (2) is adopted to account for the thermal bending as well. Therefore, the temperature difference $\Delta T(t_i)$ between the opposite side of the tower in subsidence direction is included in the exogenous inputs considered ($e = 2$). The total number of parameters to be estimated for each local regression becomes $m = 7$. This customized model returns also an estimate for the thermal bending coefficient $a_{\Delta T}$. Both models assume a_T and $a_{\Delta T}$ to be constant within the regression window. Fig. 6 shows the results obtained from MLS fitting for both models. The results are compared in terms of the tower inclination signal $\hat{a}_s(t_i)$, obtained from (13a) and (14), and the model coefficients $\hat{a}_T(t_k)$ and $\hat{a}_{\Delta T}(t_k)$ estimated for each regression window. The difference between the inclination signal $\hat{a}_s(t_i)$ and the TM readings $a_m(t_i)$ is evident. This justifies the need to operate the compensation of thermal effects. Signals from both models have a similar trend and are close to each other for the most part of the monitoring period. However, model (1) provides significantly lower estimates for some periods with respect to model (2). Regarding the other model parameters, it is plausible to suppose the coefficient a_T constant over time, since it is a characteristic inherent to the sensor, while no assumptions can be made about $a_{\Delta T}$, since the daily solar radiation on structure changes with the seasons (intensity, direction, and so on). The estimates of $\hat{a}_T$ from model (1) appear to be inconsistent with the realistic hypothesis of a constant value over time. Conversely, model (2) returns more consistent estimates of the temperature coefficient of the sensor. Despite some fluctuations are still present, their magnitude is reduced. The estimates of the thermal bending coefficient $\hat{a}_{\Delta T}$, only available from model (2), suggest a seasonal trend, with the highest value in summer and the lowest in winter.

A. Analysis of Residuals

The analysis of residuals makes it possible to assess the adequacy of the models [26]. The residuals $r_m(t_i)$ are defined as

$$r_m(t_i) = a_m(t_i) - \hat{a}_m(t_i)\,. \tag{16}$$

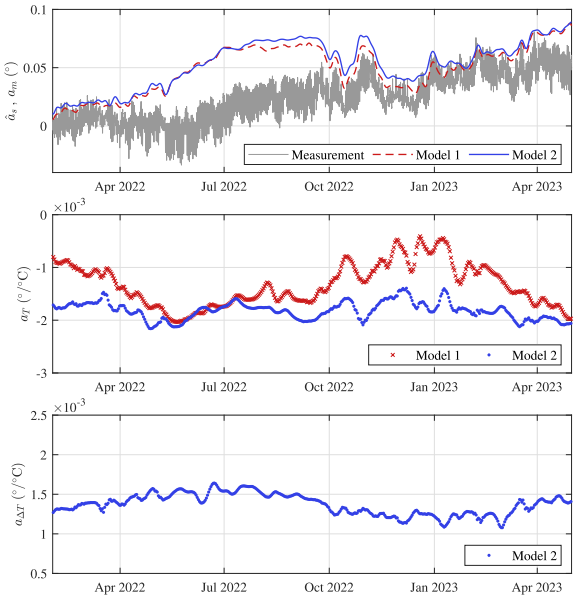

Fig. 6. Inclination signal $\hat{a}_s(t_i)$ and model coefficients $\hat{a}_T$ and $\hat{a}_{\Delta T}$ estimated with MLS adopting models (1) and (2).

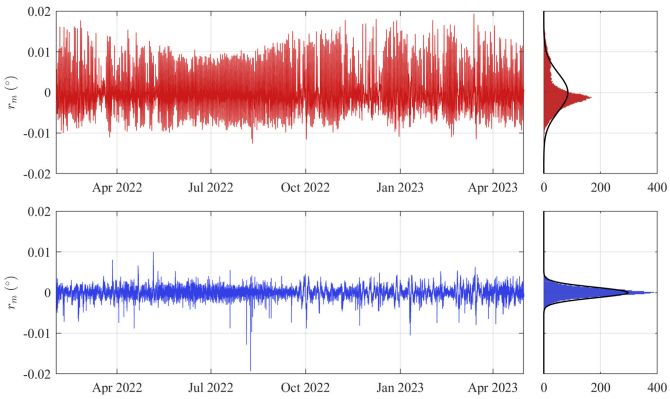

Fig. 7. Residuals $r_m(t_i)$ between measured signal $a_m(t_i)$ and predicted signal $\hat{a}_m(t_i)$ for model (1) on top and of model (2) on the bottom.

Gross information can be found out from verifications of their basic properties and some diagnostic plots. In this case, the residuals concern the global approximation with MLS.

Since the experimental data are in the form of time series, the residuals are first plotted versus time, together with their histograms, in Fig. 7. The reconstruction of measured signal with model (1) results heavily incomplete; in fact, the residuals have predominant oscillatory pattern versus time, as anticipated in Section II. In contrast, the residuals of model (2) no longer show such a clear pattern, although some small structured oscillations are still visible. The histograms of residuals supply information about the distribution of the observed error. For model (1), it is evident a serious departure from a Gaussian probability density function (pdf). Moreover, skewed residual distribution often suggests an incomplete modeling. Instead, the histogram of residuals for the model (2) is symmetrical and much closer to a Gaussian pdf.

A basic property of the residuals from LS regressions is that they must have a zero mean by definition. For both models,

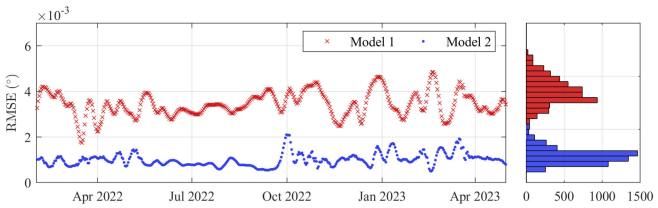

Fig. 8. On the left, the RMSE for model (1) and model (2) for each regression window. On the right, the histogram of RMSE for both models.

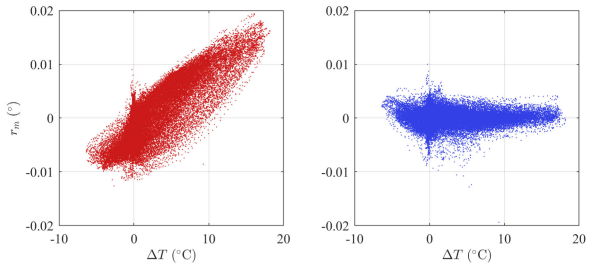

Fig. 9. Residuals $r_m(t_i)$ versus temperature difference $\Delta T(t_i)$ for model (1) on the left and of model (2) on the right.

this property is satisfied even after the operation of merging together the local WLS regression.

Another basic aspect to be evaluated is whether the error standard deviation σ_ε is constant across all observations. This can be estimated by the root-mean-square error (RMSE) of the residuals. Fig. 8 shows that the RMSE calculated for each regression window changes over time for both models. This is somehow expected, since the environmental conditions of the wind turbine can change considerably over time. Moreover, sporadic phenomena, not considered in the model, may occur.

The plot of residuals $r_m(t_i)$ versus $\Delta T(t_i)$ is depicted in Fig. 9, which indicates the clear presence of a relevant omitted variable in model (1). Conversely, the residuals of model (2) are independent of $\Delta T(t_i)$. This is also the cause of the difference in $\hat{a}_s(t_i)$ and $\hat{a}_T$ estimates between the two models. In Fig. 10, the sample autocorrelation function (ACF) is depicted to check the residuals correlation over time. The model (1) produces residuals with strong periodic correlation. The period is of one day. The residuals from model (2) do not show this behavior, having much weaker ACF, especially for lags greater than three days. Anyway, a relevant serial correlation occurs for lags of less than one day. This is somehow to be expected, since the data used for MLS fitting come from time series; thus, it is likely that the errors neighbored in time are similar. The analysis of residuals shows substantial differences in the results obtained with the two models. In particular, this analysis underlines the negative impact of a relevant omitted variable in the model (1). Therefore, it emphasizes the need of an adequate compensation of concurrent thermal effects on TM readings and the choice of a custom model as (2) is mandatory in such context. For these reasons, the following analyses are focused on model (2).

B. Potential Departure From LS Assumptions

When an adequate model of the measurand is adopted, the MLS fitting with exogenous inputs demonstrated to provide

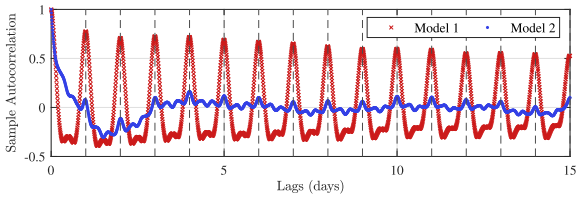

Fig. 10. Sample ACF of residuals $r_m(t_i)$ for model (1) and model (2).

plausible result. However, the trustworthiness of this result needs to be further investigated to identify possible flaws in the coefficient estimates. In fact, the "structure" of the data is determined by the operating conditions, so it may depart from the assumptions made in WLS regression. A detailed discussion about the problems occurring when certain assumptions are violated is available in [31]. Both OLS and WLS regressions produce unbiased coefficient estimates with minimum variance only if the hypotheses of Gauss–Markov theorem are met, thus holding the BLUE status. The undesirable effects of these violations can be divided into two categories: those that cause a bias in the estimate and those that alter the variance of the estimate.

The following violations of assumptions have an impact on the variance of local coefficient estimates. The RMSE depicted in Fig. 8 is heterogeneous over time. The minimum variance property of OLS/WLS derives directly from the homogeneity of the error variances, so heterogeneous variances degrade this property. A similar effect is produced by correlated errors. In this application, the usage of data in form of time series leads to serially correlated errors (Fig. 10). In such cases, the regression coefficients remain unbiased but the estimator may become inefficient. The collinearity between regressors may be an issue especially for the exogenous inputs. Since the operating conditions determine the temperature signals, different degrees of correlation occur during the monitoring. The Pearson correlation coefficient $\rho(T, \Delta T)$ for each regression window is shown in Fig. 11. In fact, this ranges from zero to about 0.5, thus meaning that there is no risk of collinearity, but only of partial correlation. The collinearity problem appears when the correlation between two exogenous inputs approaches the unity. Therefore, this is an aspect to be monitored to check the reliability of the compensation performed with MLS. The approximation of the dependent variable is not seriously affected by the use of correlated regressors. However, the greatest impact is on the coefficient estimates as it makes the LS solution unstable. A straightforward approach to handle this issue is to obtain additional data, in order to reduce the risk of singularity. In fact, this aspect is one of the main driver for the choice of the window length L.

The following violations of assumptions may lead to biased coefficient estimates. A possible cause is the presence of a relevant omitted variable, as demonstrated in the previous sections. The model (2) is considered adequate for the monitoring, since it a good compromise between its completeness and its complexity. Therefore, the effect of potential omitted variables is considered negligible. Some outliers may occur due to sporadic phenomena during the operating conditions,

Fig. 11. Pearson correlation coefficient between exogenous regressors T and ΔT for each regression window.

especially in long-term monitoring. These could cause bias, but the nature of MLS approach restricts their effect to the regression window where they occurs. The last aspect regards the errors in the explanatory variables. Indeed, one of the fundamental assumptions of OLS/WLS is that the explanatory variables are exactly known, while all the errors are on the explained variable. When this hypothesis is violated, the coefficient estimates are biased. In these cases, it is necessary to approach the field of so-called *errors in variable models* or *measurement error models* [32]. In the context of this monitoring, the exogenous input data come from experimental measurements and are thereby subject to measurement uncertainty. For this reason, the risk of biased coefficient estimate exists, and it is important to assess its impact on the final result. The next section is dedicated to a detailed discussion about this aspect.

V. Accounting for Uncertainty in Exogenous Explanatory Variables

The consequences of measurement errors in the explanatory variables on the regression outcomes depend on the error structure. The latter are comprehensively dealt with in [33], together with a detailed review of the analysis strategies to apply in such contexts. Despite an extensive literature on the subject, in [34], the authors remark the fact that regression techniques tailored to *errors in variables* cases are often neglected due to their increased complexity compared to OLS. For this reason, they analyze the consequences of applying OLS with errors in the independent variables, making assumptions and approximations that make the calculations tractable. This simple approach is used in this work to correct WLS estimates a posteriori and to assess the effect of measurement errors. This is achieved by introducing in the computation a plausible value of error variances on exogenous inputs, i.e., the measurement uncertainty.

In the following, typical additive error model is considered, i.e., the observed value of independent variables is considered as the sum of two contributions: the unobservable "exact" value and the measurement error. Therefore, the observed matrix of regressors is $\mathbf{Z}$, which is related to the exact matrix of regressors $\mathbf{X}$ as

$$\mathbf{Z} = \mathbf{X} + \mathbf{U} \tag{17}$$

where $\mathbf{U}$ is the $n \times m$ matrix of measurement errors for each observation of each regressor. The bias in the OLS estimate of coefficients occurs when $\mathbf{Z}$ is used in (4) instead of $\mathbf{X}$. The extent and the direction of the bias depend on several factors, such as the relative error magnitude versus the variance of data, the degree of correlation between the explanatory variables, and the structure of the error $\mathbf{U}$.

Suppose that all measurement errors in $\mathbf{U}$ have zero mean and constant variance for each regressor and are independent of each other and of the exact values (i.e., the covariance is null for each pair of entries). Under these hypotheses, the covariance matrix of the errors $\mathbf{U}$ has the following diagonal form:

$$\frac{1}{n}\mathbb{E}\left[\mathbf{U}^\top\mathbf{U}\right] = \mathbf{S}_\mathrm{U} = \begin{bmatrix} u_1^2 & \dots & 0 \\ & \ddots & \\ 0 & \dots & u_m^2 \end{bmatrix} \tag{18}$$

where u^2 denote the measurement error variances for each regressor. The approximate formula for the expected value of biased coefficients provided in [34] derives from the assumption of this error structure and "small" error magnitudes

$$\mathbb{E}[\mathbf{b}(\mathbf{U}, \boldsymbol{\varepsilon})] = \left(\mathbf{I} - \nu\, \mathbf{S}_\mathrm{X}^{-1}\mathbf{S}_\mathrm{U}\right)\boldsymbol{\beta} \tag{19}$$

where $\nu = (n - m - 1)$ and $\mathbf{X}^\top\mathbf{X}$ is expressed as $\mathbf{S}_\mathrm{X}$ for a more concise notation. Under the hypotheses just mentioned, the coefficients are always biased toward zero as the measurement errors increase. This phenomenon is often referred to as "regression dilution". Manipulating (19) and omitting the expectation operator, a convenient form of the same expression is obtained

$$\boldsymbol{\beta} = \left(\mathbf{S}_\mathrm{X} - \nu\mathbf{S}_\mathrm{U}\right)^{-1}\mathbf{S}_\mathrm{X}\mathbf{b}\,. \tag{20}$$

This expression means that, by introducing knowledge of the variances of the measurement error, it is possible to correct the biased OLS estimates of $\mathbf{b}$ and recover the unbiased values of the regression coefficients of $\boldsymbol{\beta}$, under the specified hypotheses. Since in practice, X is not available, Z can be used to get an estimate of unbiased regression coefficients as

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{S}_\mathrm{Z} - \nu\mathbf{S}_\mathrm{U}\right)^{-1}\mathbf{S}_\mathrm{Z}\mathbf{b} \tag{21}$$

where $\mathbf{S}_\mathrm{Z}$ denotes the matrix $\mathbf{Z}^\top\mathbf{Z}$. The expression (21) can be used to make an a posteriori correction of biased estimates of the coefficients derived from OLS or WLS, when the knowledge of the error variances on each regressor is introduced. This correction method is an original formulation presented in this article. However, it is derived after different assumptions and approximations, so it is important to characterize its behavior and trace the perimeter of its applicability.

VI. Uncertainty Analysis of the Measurement Process With MCM

In this section, the MCM [35] is used to conduct two different analyses to assess the severity of bias on coefficients, when the measurement errors are neglected. The aim is to identify the working range for the bias compensation formula (21). In the second analysis, the whole measurement process is simulated using experimental temperature signals and a test inclination signal. The objective of this experiment is to propagate the instrumental uncertainties through the measurement algorithm, i.e., MLS with model (2). On the one hand, it is possible to evaluate the uncertainty of the measurands (a_s, a_T, and $a_{\Delta T}$) over time. On the other hand, it is possible

to identify the occurrence of some critical periods where the compensation of thermal effects is inaccurate.

Before carrying out these analyses, it is necessary to assign a pdf to each input variable. The instrumental uncertainty for each sensor utilized is evaluated with a type B approach [36]. The accuracy of the measured signal a_m is discussed in Section II-B. Despite this, the upper limit corresponding to the MPE of the analog TM is used as instrumental uncertainty for the following analyses. In this case, where only the bounds of maximum errors are provided, the suggestion in clause *F.2.4.2* of the GUM [36] is pursued. A rectangular pdf with zero mean and the bounds equal to $\pm 0.020°$ is assumed, thus leading to $u(a_m) = 0.020/(3)^{1/2}\,°$ as standard uncertainty for sensor readings. All the three type K Class 1 thermocouples have the same uncertainty. Also in this case, a rectangular pdf with bounds ± 1.5 °C is assigned. It follows that the standard uncertainty for temperature measurements with TC0 is $u(T) = (3)^{1/2}/2$°C. The measurement of ΔT is obtained as difference between temperatures measured by TC1 and TC2. Hence, the resulting pdf is triangular, having bounds equal to ± 3.0 °C. The combined standard uncertainty is $u(\Delta T) = (6)^{1/2}/2$ °C. Since each sample of the signals is the result of ten minutes averaging, it has been checked that the uncertainty of the mean obtained from repeated measurements (type A) is negligible with respect to instrumental uncertainties. In fact, this contribution to the total uncertainties is at least two orders of magnitude lower than the instrumental ones.

A. Validation of Bias Correction Method

The presence of measurement errors in the regressors is simulated assuming the standard uncertainties $u(T)$ and $u(\Delta T)$ and the pdfs assigned in the previous section. The parameters that are varied in this analysis regard only statistics of exogenous inputs T and ΔT. These are the mean values, indicated as $\mu(T)$ and $\mu(\Delta T)$, the standard deviations, indicated as $\sigma(T)$ and $\sigma(\Delta T)$, and their Pearson correlation coefficient, indicated as $\rho(T, \Delta T)$.

Two different subsets of signals for $T(t_i)$ and $\Delta T(t_i)$ are pulled out from experimental data for which the Pearson correlation coefficients are, respectively, the minimum and the maximum recorded in the whole monitoring period, i.e., $\rho(T, \Delta T) \approx 0$ and $\rho(T, \Delta T) \approx 0.5$ (Fig. 11). Both subsets are limited to a length of three days ($n = 432$). This allows to reduce the computational burden but, at the same time, to perform an informative and meaningful analysis. These subsets are used to build the matrix of error-free regressors $\mathbf{X}$ and to simulate an ideal measured signal $\mathbf{y}$ for each case. The vector of exact regression coefficients is $\boldsymbol{\beta} = \left[a_0\ a_T\ a_{\Delta T}\right]^\top$. The nominal values are taken as the average of the MLS coefficients in Fig. 6, thus having $a_T = -1.8 \cdot 10^{-3}\,°/$°C and $a_{\Delta T} = 1.4 \cdot 10^{-3}\,°/$°C. The simulated inclination is constant and null for all the samples ($a_0 = 0$), so that it can be exactly reconstructed with a constant term, i.e., a polynomial of degree $p = 0$. In this way, there is no risk of incomplete reconstruction of the simulated inclination signal, and the analysis can be focused only on the effects of measurement errors of exogenous inputs. Moreover, any particular weighting is assigned to each observation in this analysis ($\mathbf{W} = \mathbf{I}$). For each combination of parameters, the procedure is the following.

1) Scale and offset T and ΔT to have the prescribed $\mu(\cdot)$ and $\sigma(\cdot)$ and build the exact regressors matrix $\mathbf{X}$.
2) Compute the exact observed signal as $\mathbf{X}\boldsymbol{\beta} = \mathbf{y}$.
3) Generate the measurement errors matrix $\mathbf{U}$ according to the assigned uncertainties and pdfs and compute the regressor matrix $\mathbf{Z}$ with measurement errors.
4) Regress $\mathbf{y}$ over $\mathbf{Z}$ with OLS to obtain the biased coefficient estimate $\mathbf{b}$.
5) Apply (21) to obtain the unbiased coefficient estimate $\hat{\boldsymbol{\beta}}$.

These steps are repeated over 10^4 trials to identify under which conditions the average estimate of $\mathbf{b}$ and $\hat{\boldsymbol{\beta}}$ differ from $\boldsymbol{\beta}$. Higher number of trials do not provide any further information about the trend of estimated coefficients. For this analysis, it is convenient to reason in terms of the ratio between the standard deviation $\sigma(\cdot)$ of a regressor and its standard uncertainty $u(\cdot)$, named here as "signal-to-noise ratio" (SNR). This quantity is expressed as $\sigma/u(\cdot)$ and controls the ability of the OLS to retrieve unbiased coefficients.

TABLE I
SUMMARY OF PARAMETERS FOR MCM TO VALIDATE THE BIAS CORRECTION STRATEGY

Parameter	$\mu(T)$	$\mu(\Delta T)$	$\sigma/u(T)$	$\sigma/u(\Delta T)$	$\rho(T, \Delta T)$
Minimum	0	0	0.1	0.1	≈0
Maximum	32	4	10	10	≈0.5

A summary of the values for the parameters tested with MCM is reported in Table I. A fine grid of values is tested for $\sigma/u(T)$ and $\sigma/u(\Delta T)$, which consists of 40 equally spaced values on a logarithmic scale covering the range specified in the table. A total of 1600 combinations are tested. These are repeated for each of the following combinations of the other parameters.

1) *Case 1:* Minimum $\mu(\cdot)$ and minimum $\rho(\cdot)$.
2) *Case 2:* Minimum $\mu(\cdot)$ and maximum $\rho(\cdot)$.
3) *Case 3:* Maximum $\mu(\cdot)$ and minimum $\rho(\cdot)$.
4) *Case 4:* Maximum $\mu(\cdot)$ and maximum $\rho(\cdot)$.

The first aspect to analyze is the amount of bias in the coefficients $\mathbf{b}$ returned by the OLS. The results state that both the "slope" coefficients a_T and $a_{\Delta T}$ are biased toward zero, thus confirming the occurrence of regression dilution. It is useful to focus the analysis on the relative error magnitude, which is mainly the function of $\sigma/u(T)$ and $\sigma/u(\Delta T)$. Although to a lesser extent, the estimation of coefficients is also influenced by $\rho(T, \Delta T)$. In fact, for uncorrelated regressors, the bias only depends on the SNR of corresponding regressor. Conversely, for partially correlated regressors, the estimation of one coefficient is sensitive to a poor estimation of the other. Generally, it results in a higher bias at the same SNR. In particular, the worst case occurs when for the other regressor holds $\sigma/u(\cdot) \approx 1$. Table II summarize the relative bias errors on slope coefficients and provides their order of magnitude in the worst cases. Instead, the mean values $\mu(T)$ and $\mu(\Delta T)$ do not affect the bias of a_T and $a_{\Delta T}$. Also for the

TABLE II

WORST CASE OF RELATIVE BIAS ERRORS (%) ON SLOPE COEFFICIENTS FROM OLS WITH MEASUREMENT ERRORS. THE WORST CASE CORRESPONDS TO $\sigma/u(\cdot) \approx 1$ FOR THE OTHER REGRESSOR

		$\sigma/u(\cdot)$		
		0.1	1	10
$\rho(\cdot)$	0	98%	50%	1%
	≈ 0.5	100%	67%	4%

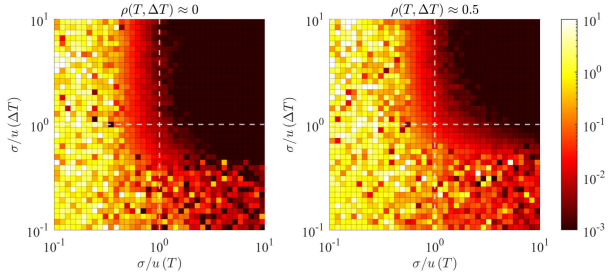

Fig. 12. Parametric analysis with MCM: absolute value of relative bias error on a_T after correction. Case 3 on the left and case 4 on the right.

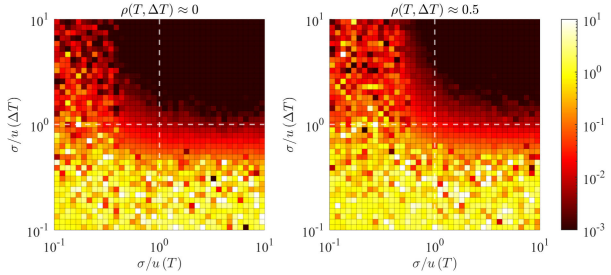

Fig. 13. Parametric analysis with MCM: absolute value of relative bias error on $a_{\Delta T}$ after correction. Case 3 on the left and case 4 on the right.

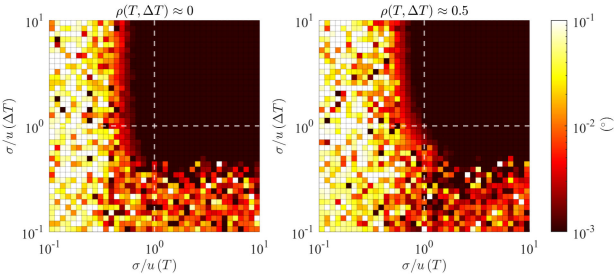

Fig. 14. Parametric analysis with MCM: absolute bias error on a_0 (°) after correction. Case 3 on the left and case 4 on the right.

constant term a_0, the SNR controls the bias. As the SNR of one regressor decreases, the severity of bias increases. Instead, the sign of bias on the constant term a_0 is not predefined; it can be either positive or negative. Indeed, magnitude and sign of the error depend on the combination of slope coefficients and regressor mean values. When the slope coefficients are poorly estimated, the bias error increases as the mean of the regressors increases. The bias error on a_0 is zero when $\mu(T)$ and $\mu(\Delta T)$ are zero (cases 1 and 2). For cases 3 and 4, the maximum absolute error is about 0.06°, occurring when $\sigma/u(T)$ is low. In these cases, the temperature T has the major effect due to greater value of the product $\mu(T)a_T$ with respect to $\mu(\Delta T)a_{\Delta T}$. The bias on a_0 tends to zero when the SNRs increase, regardless of the mean value of the regressors.

The second aspect to analyze is the effectiveness of bias correction strategy (21). This analysis assumes that the correction is performed with the exact value of error variances for each regressor; therefore, it is the ideal condition. For comprehensive visualization of the behavior of correction strategy, bias error maps are presented. All the color scales in the following maps are logarithmic, in order to enhance the order of magnitude of the values depicted. The maps in Figs. 12 and 13 pertain to the cases 3 and 4, but they are qualitatively similar to those of cases 1 and 2. Therefore, the latter are not explicitly shown here. This implies that the mean values of regressors have no relevant effect even on bias correction of the slope coefficients. The top-right quadrants of the error maps, i.e., where both the SNRs are greater than 1, demonstrate the effectiveness of bias reduction. In fact, in that area, the relative bias error is limited to 10^{-2}, i.e., to 1% of the coefficients. Similar behavior is noticeable in the error map of a_0 for cases 3 and 4 as depicted in Fig. 14. Here, the bias is reduced below 0.001° for the top-right quadrants. Even here, for cases 1 and 2, the expected bias is null when $\mu(T)$ and $\mu(\Delta T)$ are zero. The correlation between the regressors slightly affects the perimeter of applicability.

In the light of the results of this analysis, it is important to assess whether the available operational data for T and ΔT may lead to significant biases. To do this, realistic measurement uncertainties of the exogenous inputs need to be evaluated. Thereafter, it is possible to understand if the biases are relevant and if they can be mitigated with (21). This parametric analysis has been conducted by using the SNR $\sigma/u(\cdot)$ of each regressor as the measure of "quality" for the OLS estimates. A useful parameter, frequently used in this context, is the "reliability ratio" [32], which is defined as

$$\lambda(\cdot) = \frac{\sigma^2(\cdot)}{\sigma^2(\cdot) + u^2(\cdot)} = \frac{\frac{\sigma^2}{u^2}(\cdot)}{1 + \frac{\sigma^2}{u^2}(\cdot)} \tag{22}$$

where the relationship with the SNR $\sigma/u(\cdot)$ is explicitly indicated. In other words, the reliability ratio is defined as the variance of the error-free variable over the total variance. Depending on the context, it may be convenient to reason in terms of $\lambda(\cdot)$ or $\sigma/u(\cdot)$, even if they provide the same information. From the parametric analysis, it emerged that for each regressor, the SNR must be greater than 10 to avoid significant bias. This means that $\lambda(\cdot)$ must be close to 1; otherwise, the correction is needed. Recalling Figs. 12–14, for an effective correction, it is sufficient that SNR is greater than 1. This condition corresponds to $\lambda(\cdot) > 0.5$.

In practice, the sample variance $s^2(\cdot)$, which is computed for each regression window, is an approximation of the total variance, i.e., $s^2 \approx \sigma^2 + u^2$. It follows that a consistent estimator of the reliability ratio is [33]

$$\hat{\lambda}(\cdot) = \frac{s^2(\cdot) - u^2(\cdot)}{s^2(\cdot)} \tag{23}$$

where the numerator is bounded to zero if $s^2 < u^2$. The estimated reliability ratios $\hat{\lambda}$ are depicted in Fig. 15 for both regressors and confirm that the bias of OLS coefficients cannot

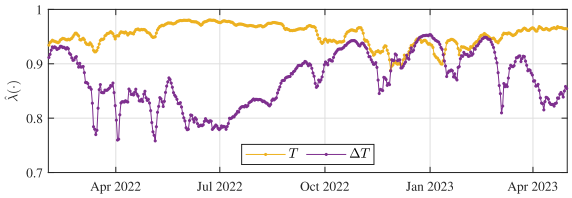

Fig. 15. Estimated reliability ratio $\hat{\lambda}(\cdot)$ for exogenous regressors T and ΔT for each regression window.

be neglected; anyway, the bias correction can be successfully applied. This parameter has been the main driver in the choice of the window length L. In fact, this parameter was set by looking for the minimum window width leading to an acceptable SNR for the entire monitoring period. In fact, for a short time window, there is an high risk of having some regression window that do not fulfill the condition of $\hat{\lambda} > 0.5$.

B. Simulation of Measurement Process for Instrumental Uncertainty Propagation

The model adopted here to simulate a measured signal is (2). The operational signals $T(t_i)$ and $\Delta T(t_i)$ are used as the ground truth to generate a simulated measured signal $a_m(t_i)$. The coefficients a_T and $a_{\Delta T}$ are the same of the previous parametric analysis. The test inclination signal is generated as it follows:

$$a_s(t_i) = 0.1 \sin\left(\frac{2\pi}{30}(t_i - t_1)\right). \tag{24}$$

The amplitude of the simulated oscillation is chosen to be of the same order of magnitude as the signal derived from the experimental data (Fig. 6). The period of the oscillation is one month. Preliminary tests, not reported here, indicate that the test inclination signal can be reconstructed correctly with the MLS polynomial approximation, applying the same parameters used for the experimental data ($p = 4$, $L = 15$, and $h = 1/L$). Therefore, its approximation does not affect the other coefficients. The parametric analysis described in the previous section suggests a useful expedient: the bias of the constant term is lower when the mean value of the regressors is nearly zero. For this reason, the mean value over the whole monitoring period is subtracted from each regressor. This limits from the root the maximum bias, independent of the application of bias correction. This practice is feasible in this application, since the interest is on the change of inclination rather than the absolute value. The mean values are $\mu(T) = 16.80$ °C and $\mu(\Delta T) = 1.91$ °C.

At this point, the MCM is applied to propagate the instrumental uncertainties and to characterize the outputs of the measurement algorithm over the monitoring period, accounting for the trend of exogenous input signals. Using this analysis, it is possible to identify the periods most prone to bias, since the exact values for $a_s(t_i)$, a_T, and $a_{\Delta T}$ are known and assess how uncertainty varies over time. The input quantities subject to uncertainty are $a_m(t_i)$, $T(t_i)$, and $\Delta T(t_i)$, whose standard uncertainties and pdf have been already assigned at the beginning of Section VI. For each MCM trial, the signals

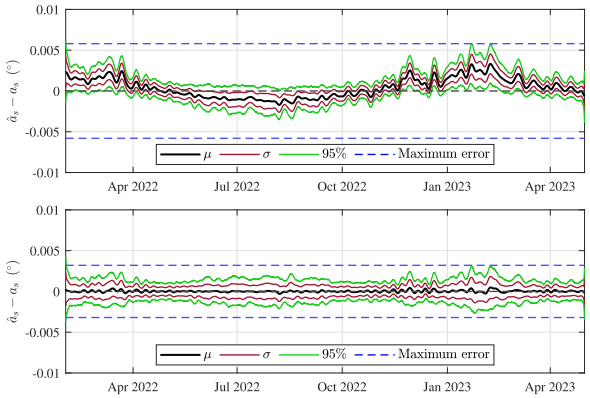

Fig. 16. Error on a_s from the analysis of the measurement process with MCM. The biased estimates from MLS on top. The estimates corrected for the measurement errors on bottom.

are generated by adding random errors to the ground truth signals. The random errors are drawn from the prescribed pdfs for each sample. Finally, the compensation with MLS is computed. The results presented hereafter are obtained from 10^3 trials with the same MLS parameters used on experimental data in Section IV. Both biased estimates from MLS and bias corrected with (21) are computed. Even in this analysis, the correction is made for the ideal condition of perfect knowledge of the measurement error variances on the exogenous inputs. To summarize the outcomes of the MCM, typical statistics are computed on the errors, i.e., mean, standard deviation, and bounds for covering 95% of the errors (2.5% and 97.5% percentiles).

The results are depicted in Fig. 16 for $\hat{a}_s(t_i)$, while in Figs. 17 and 18 for $\hat{a}_T(t_k)$ and $\hat{a}_{\Delta T}(t_k)$. The standard deviations over time depend on combinations of the degree of correlation between the exogenous inputs and their reliability ratios (Figs. 11 and 15). All the resulting distributions are almost symmetric and Gaussian (not reported in this article). Furthermore, the coverage factor results to be close to 1.96 for both inclination and coefficients over the monitoring period. The mean errors on biased estimates of $\hat{a}_s(t_i)$ are positive or negative depending on the period. This is the effect of subtracting the global mean values from exogenous regressors. Without this practice, the error would have been greater and always with the same sign. Fig. 16 shows a maximum error of $\pm 0.006°$, considering a coverage of 95%. By applying the bias correction, the maximum bias can be reduced to levels comparable with the TM resolution. The maximum error, with the level of confidence of 95%, decreases to $\pm 0.003°$. The edge of the monitoring period is excluded from the error evaluation, since only full size regression windows are considered. The slope coefficient a_T has different levels of bias in different periods with a trend that seems to match the correlation between the exogenous inputs (Fig. 11). In fact, the minimum bias occurs in the summer months, where the correlation is lower. Conversely, the bias on $a_{\Delta T}$ is mostly constant over the period, apart for some isolated peaks. Despite the bias is nonnegligible, it can be successfully corrected with (21), for almost the entire monitoring period. However, some values of $a_{\Delta T}$ still remain partially biased in the periods corresponding to low reliability ratios (Fig. 15). This is confirmed also by

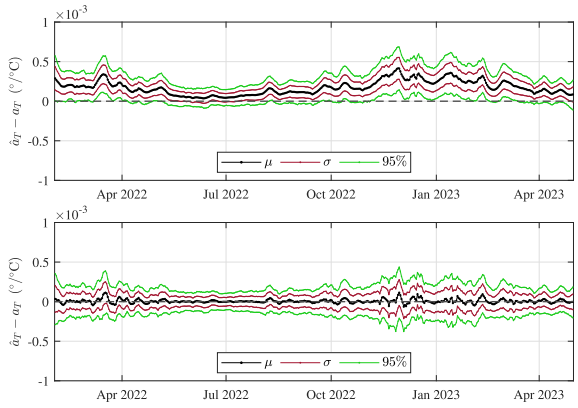

Fig. 17. Error on a_T from the analysis of the measurement process with MCM. The biased estimates from MLS on top. The estimates corrected for the measurement errors on bottom.

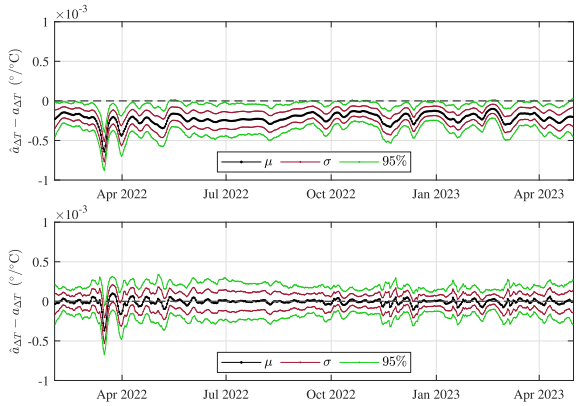

Fig. 18. Error on $a_{\Delta T}$ from the analysis of the measurement process with MCM. The biased estimates from MLS on top. The estimates corrected for the measurement errors on bottom.

Fig. 4, showing some days with very small variations of ΔT at end of March 2022.

VII. Final Results and Discussion

In this section, the thermal compensation of TM readings is carried out again from the experimental study case, in the light of what has been discussed so far in this article. As first point, the signals of exogenous inputs are offset to globally have zero mean in the monitoring period, differently from Section IV. This is done to limit the maximum bias on the constant term in the reconstruction of a_s, even if this introduce a global offset on the inclination value. Anyway, it is irrelevant since the focus is on variation of inclination. The second point is the need of a realistic evaluation of the measurement uncertainty for the exogenous inputs. It is complicated to get an accurate estimate of the "true" error variances present in the operational data. Therefore, the uncertainties of T and ΔT, discussed in Section VI, have been used for the correction of the MLS coefficients. This makes it possible to understand what impact the assumed level of uncertainty would have on the coefficients.

Wrapping all these considerations, the final results are computed and depicted in Fig. 19. The inclination signal $\hat{a}_s(t_i)$ is reconstructed by means of MLS, with and without bias correction. Their maximum absolute difference is 0.004°. This

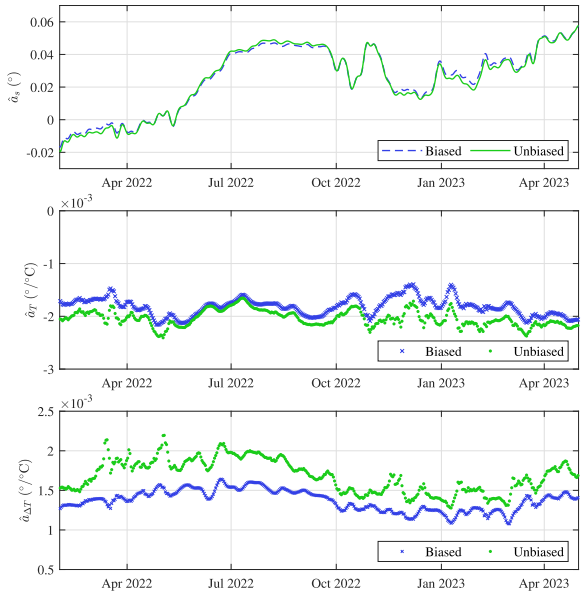

Fig. 19. Inclination signal $\hat{a}_s(t_i)$ and model coefficients $\hat{a}_T$ and $\hat{a}_{\Delta T}$ estimated with MLS adopting model (2) with and without measurement error compensation.

value can be linearly combined with the maximum expanded uncertainty from the MCM results depicted in Fig. 16 for the unbiased case ($\pm 0.003°$). Therefore, the total error after compensation of thermal effects and bias correction is lower than 0.01°. This means that the target on maximum error for this application is achieved. The values of $\hat{a}_T$ seem to converge to a constant value after the correction. This is in line with the hypothesis of constant value for temperature coefficient of the TM. Instead, the bias correction on the thermal bending coefficient estimates $\hat{a}_{\Delta T}$ enhances the seasonality of this coefficient. However, as emerged from the analysis with the MCM, the correction of this coefficient results more critical in some periods. This is due to insufficient variance of the operational data in those periods. A last consideration can be done about this aspect. Fig. 19 shows some values of $\hat{a}_{\Delta T}$ between March and May 2022 that appear "overcorrected", with respect to the neighborhood values, in correspondence of drops of the reliability ratio in Fig 15. This may suggest that the actual uncertainty in ΔT could be slightly lower than the assumed one.

VIII. Conclusion

In this article, a typical SHM problem is addressed, where compensation of thermal effects on TM readings is required. A compensation method based on operational data and MLS is presented. A typical monitoring case is shown, in which the thermal effects on the sensor and on the measurand are predominant over the measurand variations themselves. In fact, the temperature drift of the sensor and the thermal bending of the tower overlook variations in the inclination of the wind-turbine tower. The approach proposed in this article enables the users to successfully adopt a compensation model that accounts for more concurrent influence quantities on the measurement. The MLS approach has been tailored to the measurement compensation task. The proposed MLS formulation accounts for exogenous inputs on the measurement,

i.e., the influence quantities to compensate. The outcomes of this method are a local polynomial approximation of the underlying inclination signal and the estimates of the model parameters from operational data. Thereafter, the effect of measurement uncertainty on the LS coefficients has been studied in detail, since it may bias their estimates. With this purpose, a bias correction strategy is derived from other literature results. This method can be used to evaluate the severity of the bias that arises from the combination of operational data and measurement uncertainty of the exogenous inputs.

Two different analyses, conducted using the MCM, provide some important results about the proposed method and its application on the real monitoring case. The parametric analysis provides an overview of the conditions that make the bias on the model coefficients nonnegligible. Then, the analysis indicates the conditions under which the bias compensation strategy can be successfully applied. The evaluation of the quality of measurement compensation boils down to the monitoring of the reliability ratios. These can be estimated from the operational signals of exogenous inputs and from their uncertainties. The second analysis with MCM is the simulation of the measurement process to propagate the input uncertainties. The results indicate the different uncertainties over time, depending on the reliability ratios of the exogenous inputs and their correlation. Finally, the bias correction is applied to the MLS results from the experimental data. The difference between biased estimates and the unbiased ones can be considered an additional source of uncertainty the measurand. Therefore, considering the maximum bias from experimental data and the maximum error obtained from the MCM analysis, the target accuracy on the tower inclination (0.01°) is achieved.

In conclusion, the methodology proposed in this work is able to compensate for thermal effects on TM readings. Moreover, it provides a strategy to evaluate the quality of the compensation and the uncertainty during different monitoring periods. Furthermore, the MLS formulation and the bias compensation strategy are not limited to applications involving inclination measurements and thermal effects. In fact, the methods proposed in this article are general enough to perform the compensation of known systematic effects of influence quantities in several contexts. This is possible whenever the measurement model fulfills the hypotheses at the basis of the method.

REFERENCES

[1] C. R. Farrar and K. Worden, "An introduction to structural health monitoring," *Philos. Trans. Roy. Soc. A, Math. Phys. Eng. Sci.*, vol. 365, no. 1851, pp. 303–315, 2007, doi: 10.1098/rsta.2006.1928.

[2] M. Sun, W. J. Staszewski, and R. N. Swamy, "Smart sensing technologies for structural health monitoring of civil engineering structures," *Adv. Civil Eng.*, vol. 2010, pp. 1–13, Jun. 2010, doi: 10.1155/2010/724962.

[3] K. Worden, C. R. Farrar, G. Manson, and G. Park, "The fundamental axioms of structural health monitoring," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 463, no. 2082, pp. 1639–1664, Jun. 2007, doi: 10.1098/rspa.2007.1834.

[4] H.-N. Li, L. Ren, Z.-G. Jia, T.-H. Yi, and D.-S. Li, "State-of-the-art in structural health monitoring of large and complex civil infrastructures," *J. Civil Struct. Health Monitor.*, vol. 6, no. 1, pp. 3–16, Apr. 2015.

[5] S. Turrisi and F. Lucà, "On the use of multiple linear regression to compensate for the effect of environmental parameters in large structures tilt measurements," in *Proc. 4th Forum Nazionale delle Misure*, 2020, pp. 327–336.

[6] D. Zhang, C. Nie, M. Chen, H. Huang, and Y. Wu, "Wireless tilt sensor based monitoring for tunnel longitudinal settlement: Development and application," *Measurement*, vol. 217, Aug. 2023, Art. no. 113050.

[7] D. Ha, H. Park, S. Choi, and Y. Kim, "A wireless MEMS-based inclinometer sensor node for structural health monitoring," *Sensors*, vol. 13, no. 12, pp. 16090–16104, Nov. 2013.

[8] Y. Zhuang, Y. Chen, C. Zhu, R. E. Gerald, Y. Tang, and J. Huang, "A high-resolution 2-D fiber optic inclinometer for structural health monitoring applications," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6544–6555, Sep. 2020.

[9] P. O'Leary and M. Harker, "A framework for the evaluation of inclinometer data in the measurement of structures," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 5, pp. 1237–1251, May 2012.

[10] P. O'Leary and M. Harker, "Inverse boundary value problems with uncertain boundary values and their relevance to inclinometer measurements," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, May 2014, pp. 165–169.

[11] H. Yang, S. Gao, S. Liu, L. Zhang, and S. Luo, "Research on identification and suppression of vibration error for MEMS inertial sensor in near-bit inclinometer," *IEEE Sensors J.*, vol. 22, no. 20, pp. 19645–19655, Oct. 2022.

[12] P. F. C. Antunes, C. A. Marques, H. Varum, and P. S. André, "Biaxial optical accelerometer and high-angle inclinometer with temperature and cross-axis insensitivity," *IEEE Sensors J.*, vol. 12, no. 7, pp. 2399–2406, Jul. 2012.

[13] J. Guo, C. Zhu, Y. Tang, and J. Huang, "Temperature-insensitive inclinometer based on transmission line Fabry–Pérot resonators," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–10, 2022.

[14] X. Zhang and D. Li, "The error compensation of inclinometer based on polynomial fitting," *Adv. Aerosp. Sci. Technol.*, vol. 6, no. 1, pp. 1–8, 2021.

[15] S. Łuczak, "Dual-axis test rig for MEMS tilt sensors," *Metrol. Meas. Syst.*, vol. 21, no. 2, pp. 351–362, Jun. 2014. [Online]. Available: http://journals.pan.pl/Content/90250/PDF/Journal10178-VolumeXXI%20Issue2_15.pdf

[16] S. Wierzbicki, Z. Pióro, M. Osiniak, and E. Antoszkiewicz, "Inclinometer method of displacement measurements as an alternative to optical measurements in structural health monitoring—Laboratory tests," *Arch. Civil Eng.*, vol. 66, no. 2, pp. 147–164, Jan. 2020. [Online]. Available: http://journals.pan.pl/Content/115132/PDF/Paper_723_do%20druku.pdf

[17] S. Wierzbicki, Z. Pióro, M. Osiniak, and E. Antoszkiewicz, "Inclinometer method of displacement measurements as an alternative to optical measurements in structural health monitoring—On site tests," *Arch. Civil Eng.*, vol. 66, no. 3, pp. 109–124, Jul. 2020. [Online]. Available: http://journals.pan.pl/Content/117454/PDF/05.Paper_724%20do%20druku_B5.pdf

[18] P. Lancaster and K. Salkauskas, *Curve and Surface Fitting—An Introduction*. London, U.K.: Academic, 1986.

[19] D. Levin, "The approximation power of moving least-squares," *Math. Comput.*, vol. 67, no. 224, pp. 1517–1531, 1998.

[20] *Thermocouples—Part 1: EMF Specifications and Tolerances*, Standard IEC 60 584-1, Int. Electrotechnical Commission, 2013.

[21] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, *Guide to the Expression of Uncertainty in Measurement—Part 6: Developing and Using Measurement Models*, document JCGM GUM-6:2020, Joint Committee for Guides Metrology, 2020. [Online]. Available: https://www.bipm.org/documents/20126/2071204/JCGM_GUM_6_2020.pdf/d4e77d99-3870-0908-ff37-c1b6a230a337

[22] F. E. Harrell, *Regression Modeling Strategies*. Cham, Switzerland: Springer, 2015.

[23] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964, doi: 10.1021/ac60214a047.

[24] W. S. Cleveland and S. J. Devlin, "Locally weighted regression: An approach to regression analysis by local fitting," *J. Amer. Stat. Assoc.*, vol. 83, no. 403, pp. 596–610, Sep. 1988, doi: 10.1080/01621459.1988.10478639.

[25] N. R. Draper and H. Smith, *Applied Regression Analysis*. Hoboken, NJ, USA: Wiley, Apr. 1998.

[26] M. Kutner, *Applied Linear Statistical Models*. New York, NY, USA: McGraw-Hill, 2005.

[27] P. Breitkopf, A. Rassineux, and P. Villon, "An introduction to moving least squares meshfree methods," *Revue Européenne des Éléments Finis*, vol. 11, nos. 7–8, pp. 825–867, Jan. 2002, doi: 10.3166/reef.11.825-867.

[28] D. Shepard, "A two-dimensional interpolation function for irregularly-spaced data," in *Proc. 23rd ACM Nat. Conf.*, New York, NY, USA, 1968, pp. 517–524, doi: 10.1145/800186.810616.

[29] W. S. Cleveland, S. J. Devlin, and E. Grosse, "Regression by local fitting: Methods, properties, and computational algorithms," *J. Econometrics*, vol. 37, no. 1, pp. 87–114, 1988. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0304407688900772

[30] P. Bloomfield, *Fourier Analysis of Time Series: An Introduction* (Wiley Series in Probability and Statistics). Hoboken, NJ, USA: Wiley, 2004.

[31] J. O. Rawlings, S. G. Pantula, and D. A. Dickey, *Applied Regression Analysis: A Research Tool*. New York, NY, USA: Springer, 1998, pp. 325–340, doi: 10.1007/0-387-22753-9_10.

[32] W. A. Fuller, *Measurement Error Models*. Hoboken, NJ, USA: Wiley, Jun. 1987.

[33] R. J. Carroll, D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu, *Measurement Error in Nonlinear Models*. London, U.K.: Chapman & Hall, Jun. 2006.

[34] S. D. Hodges and P. G. Moore, "Data uncertainties and least squares regression," *J. Roy. Stat. Soc., C, Appl. Statist.*, vol. 21, no. 2, pp. 185–195, 1972, doi: 10.2307/2346491.

[35] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, *Evaluation of Measurement Data—Supplement 1 to the 'Guide to the Expression of Uncertainty in Measuremen'—Propagation of Distributions Using a Monte Carlo Method*, document JCGM 101:2008, Joint Committee for Guides Metrology, 2008.

[36] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement*, document JCGM 100:2008, Joint Committee for Guides Metrology, 2008. [Online]. Available: https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf/cb0ef43f-baa5-11cf-3f85-4dcd86f77bd6

Gianmarco Battista received the M.Sc. degree in mechanical engineering and the European Ph.D. degree in industrial engineering from the Polytechnic University of Marche, Ancona, Italy, in 2015 and 2019, respectively.

He continued working as a Post-Doctoral Researcher with the Department of Industrial Engineering and Mathematical Sciences, Polytechnic University of Marche from 2019 to 2022 apart for 2020. In that year, he worked as an NVH and a Sound Engineer at Ferrari S.p.A., Maranello, Italy. Since 2022, he has been an Assistant Professor with the University of Parma, Parma, Italy. His research activity focuses on noise and vibration measurements and structural health monitoring (SHM).

Stefano Pavoni received the M.S. degree in mechanical engineering from the University of Parma, Parma, Italy, in 2019, with a thesis on an automatic operational modal analysis (OMA) in the presence of harmonic excitations, where he is currently pursuing the Ph.D. degree in industrial engineering.

His research interests include structural health monitoring (SHM), damage identification techniques, and vibration analysis.

Marcello Vanali received the M.S. degree in mechanical engineering and the Ph.D. degree in applied mechanics from the Politecnico di Milano, Milan, Italy, in 1997 and 2003, respectively.

He is currently a Full Professor of mechanical measurements with the University of Parma, Parma, Italy. His research activity is mainly focused on structural health monitoring (SHM) applications, sensors development, and vibration control/analysis.

Open Access funding provided by 'Università degli Studi di Parma' within the CRUI CARE Agreement