Statistical correlation between environmental time series and data from long-term monitoring of buildings

Statistical correlation between environmental time series and data from long-term monitoring of buildings

R. Ceravolo[1], G. Coletta[1], G. Miraglia[1] and F. Palma[1].

[1] Department of Structural, Geotechnical and Building Engineering,
Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy.

Abstract: Within the context of civil structures, a monitoring system supported by an intelligent diagnostic features extraction allows to keep under observation the overall health state of a building. In most cases, the diagnostic features are influenced by Environmental and Operational Variations (EOVs) which cause fluctuations that can be confused with the appearance of damage, or worse, hide it. A useful strategy to get rid of those confounding effects consists in modelling the structural behaviour of the system, considering and predicting these harmless and reversible fluctuations. However, a model approximates a much more complex reality and therefore it is based on a reasonable number of components whose selection might turn out complicated. In this research, a large amount of heterogeneous experimental data is systematically analysed to investigate which have the greatest influence on structural behavior and therefore, could contribute for modelling the behaviour of a historic building for Structural Health Monitoring (SHM) purpose. Environmental data, measurements of static sensors and modal natural frequencies collected in more than 10 years are scanned and crossed in order to discover any correlations. The analysis of these time series, treated with mathematical and statistical tools, has led to some mechanical interpretations of the observed behaviour of the system, i.e. the Sanctuary of Vicoforte, a monumental Italian church which houses the largest masonry oval dome in the world. The results obtained, especially in terms of correlations between different factors affecting measurements, are deemed relevant in the practice of long-term monitoring of cultural heritage and existing buildings in general.

KEYWORDS:

Correlation analysis; Structural heath monitoring; Cultural heritage; Long-term monitoring; Sanctuary of Vicoforte; Environmental and operational variations.

1. INTRODUCTION

The growing interest in preserving and protecting the architectural heritage has encouraged the development and the application of modern structural monitoring techniques based on the analysis and the interpretation of data acquired from sensors [1-4]. Monitoring the health state of this type of buildings is an ingenious approach as it allows to evaluate the condition of conservation with minimal invasiveness and enable to readily establish the safety condition after sudden events such as earthquakes [5-8]. Furthermore, in presence of data from continuous time monitoring systems, symptom-based approaches [9-11] may prove suitable for perceiving anomalies in the structural behaviour, especially when models are strongly affected from a high level of uncertainty regarding the effective properties of materials, their current state, the construction techniques, the possible interventions stratified over the years, the structure-soil interaction and so on [12-14].

Innovative techniques in SHM often involve the study and analysis of dynamic response signals using statistical tools and machine learning algorithms [15]. In their strongest sense, they aim to predict the future response of the system and to identify occurrence, type, entity and/or location of possible damage in real time or near real time. When working on architectural assets, these techniques are particularly appreciated as they prove to be practically non-invasive and reversible and in addiction, helping to increase the level of knowledge [16-19], they lead to compliance with the minimum intervention principle, as that introduced in the International Council on Monuments and Sites (ICOMOS) guidelines [20].

In some of these procedures [21-26], damage detection is based on the analysis of variations in the trend of structural frequencies over time. In fact, since a significant damage affects the global stiffness of a system and the modal parameters depend on it, checking that the parameters are stable over time and do not present unexpected variations is a good way to investigate the health of the system. Complications arise because, in addition to damage, also several EOVs influence the data recorded on the structure [27-31], although they are actually harmless and reversible. Therefore, discriminating them from real damage is important to avoid wrong diagnoses. Previous studies [32-38], some concerning the behaviour of bridges, have pointed out the importance of the effects of EOVs on real structures in SHM approach.

In this paper, the quantities that most affect the structural behaviour of architectural heritage buildings are assessed by systematically investigating dependencies and correlations with environmental phenomena;

knowing how monitoring data depend on environmental phenomena allows to shed light on their annual fluctuations, giving us the extent of the values attributable to seasonal variations and actual anomalies which could be associated with changes in the structural system, i.e. damage. The information thus obtained are given as input of regression models, to assess whether the involvement of those correlated environmental measures could bring an advantage in the prediction of structural behavior, compared to models (which will be referred to as *reference* models), based only on homogeneous measurements, i.e. concerning same type of diagnostic feature.

The benchmark of this research is represented by the Sanctuary of Vicoforte, Piedmont, Italy, a masterpiece of Baroque architecture. This monumental 17th century church, covered by the world's largest masonry oval dome, 37.23 by 24.89 m, is equipped with a permanent static [39] and dynamic [40] monitoring system, which over the years have recorded long series of data, useful for the purposes of this analysis.

The layout of this paper is as follows. Section 2 discusses the relevant theoretical background behind correlation and the regression analysis exploiting one of the most common machine learning algorithms; in Section 3 the case study is described with particular focus on its monitoring systems; Section 4 refers to the processing of experimental data. The correlation analysis is exposed in Section 5 and its results are used in Section 6 to improve the prediction of structural behavior. Finally, conclusions are drawn in Section 7.

2. CORRELATION AND REGRESSION ANALYSIS

In the first part of this chapter, reference is made to theoretical notions useful for explaining subsequent elaborations. They are simple and consolidated concepts of statistics, effective in analysing experimental data. The last sub-paragraph briefly reports the theory underlying Support Vector Machine (SVM), a machine learning tool already used in the literature for SHM purpose [41-48], which will be used here to model two diagnostic variables of the structure on the basis of different predictor variables.

2.1 Correlation analysis

Correlation analysis is a method of statistical evaluation used to study the strength of a relationship between continuous variables. This particular type of analysis is useful when establishing possible connections between variables [49-51].

The existence of a correlation between series of values implies that if there is a systematic variation in one, it is also found on the other. The search for a possible correlation between the variables passes through the supposition, in the specific case, of the existence of a linear relationship between them. In general, variables may also exhibit a more complicated than linear relationship. As a matter of fact, two series that have a small or no linear correlation might have a strong nonlinear relationship. However, checking for linear correlation before fitting any model is a useful way to identify variables that have a simple relationship. In addition, a small deviation from linearity, to be detected from the scatter plot of the data, will not generally affect the magnitude of the *linear correlation coefficient*.

In cases when an estimate of uncertainties is available, it could be assessed whether the measurements approach the linear relationship, and in the positive case, the hypothesis of the linear relationship between the variables would be confirmed. Unfortunately, in many experiments it is impossible to have an a priori estimate of the uncertainty and one must use the data themselves to determine if they are linearly related. The linear correlation coefficient (or *Pearson correlation coefficient*) measures the extent to which a set of points $(x_1, y_1), \ldots (x_n, y_n)$ supports the hypothesized relationship. It is expressed by:

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \tag{1}$$

where σ_{xy} is the covariance and σ_x and σ_y are the variances of variables. By considering the definitions of variance and covariance the formula can be written in this form:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{2}$$

r is a measure of how well the supposed y function approximates the data, being defined in the range -1 and $+1$. The closer r is to the limits of the range, the closer the measured points are to the supposed line. On the contrary, if $|r|$ is very small, close to 0, it means that the points are linearly uncorrelated. The sign of r indicates the slope of the line: a positive and closer to the unit coefficient implies a *positive correlation*, i.e.

one variable increases simultaneously with the other; in the opposite case, a *negative correlation* is observed when a variable decrease while the other increases. Even if a strong linear relationship exists between x and y variables of the system, there is no expectation that the experimental measurements will be placed exactly on a line. Therefore, unit values will never be achieved. Similarly, as a discrete number N of points is contemplated, it is not expected to obtain perfectly 0 as a correlation coefficient between uncorrelated variables: if the variables are really uncorrelated, this coefficient will decrease as the number of points increases.

While a significant relationship may be identified by correlation analysis techniques, this cannot reveal causation. The cause cannot be determined by the analysis, nor should this conclusion be attempted. A significant relationship implies that there is more to understand and that there are extraneous or underlying factors that should be explored further in order to look for a cause. In this paper, the correlation analysis allows to identify which aspects and variables depend on each other, information that can generate usable insights as they are or starting points for further investigations.

If a large number of variables are associated with the system, a pre-processing via *Principal Component Analysis* could streamline the computational burden needed to perform correlation analysis.

2.2 Principal Component Analysis

Principal Component Analysis (PCA), also known as *Karhunen-Loève* transform, *Hotelling* transform or orthogonal decomposition, is a technique mathematically defined as an orthogonal linear transformation that turn the data to a new coordinate system such that the greatest variance, by some scalar projection of the data, comes to lie on the first coordinate (called the *first principal component*), the second greatest variance on the second coordinate, and so on [52]. This technique was first proposed in 1901 by Karl Pearson and then developed by Harold Hotelling in 1933, and it is part of the factor analysis. The aim is to reduce the number of variables describing a data set to a smaller number of latent variables, limiting the loss of information as much as possible. This technique can identify the most influencing set of data through the *Singular Value Decomposition* (SVD) algorithm, a factorization that generalizes the eigen-decomposition for any $m \times n$ matrix through an extension of the polar decomposition.

In particular, given a matrix of n variables and m observations $[O]_{m\times n}$, the SVD decomposes it in three matrices, as:

$$[O] = [U][S][V]^T \quad (3)$$

Where $[U]_{m\times m}$ and $[V]_{n\times n}$ are unitary matrices which stores the *left-singular vectors* and *right-singular vectors* of $[O]$ respectively, and $[S]_{m\times n}$ is a rectangular matrix with real and non-negative terms on the diagonal known as the *singular values*, sorted in descending order.

For instance, when a very large number of sensors are installed on a structure, the SVD algorithm can be used to choose the most representative time series of the system for each type of sensor. In particular, as many $[O]$ as the types of devices in the monitoring system are assembled, which collects the time series in columns. This avoids processing a large number of series containing the same information, thus reducing data elaboration burden. For the actual problem, for example, SVD was used to extract the singular values of the temperature sensors setup. Then, the most representative temperature sensors have been chosen starting from the Proper Orthogonal Mode (POM) with highest singular value. A representative sensor was thus selected by the channel with highest absolute amplitude of the first POM.

2.3 Regression Models for SHM

Various damage detection procedures provide for the creation of a model of the system that reproduces its structural behaviour [53-57]. This turns out to be an acute strategy because the comparison of monitoring data with those produced by the model, which simulates the *healthy* condition, would allow to understand if any anomaly is taking place in the real system. For the definition of the model, some studies exploit the dependence of the diagnostic feature on environmental factors, especially temperature, sometimes measured in different positions of the system. Other researches instead, provide to model one of the diagnostic features through machine learning algorithms, using analogous features as predictors. For example, in [41,54] a model is built on the one frequency of the structure, based on the trend of the others. The idea is that the relationship between these features remains unchanged under the effect of EOVs, and that it changes when damage occurs. The warning is manifested by a departure of the monitoring data from the model data. It seems clear that the more precise and realistic the model, the easier it is to notice a difference between the two trends.

In this research, a model of a diagnostic variable is built by exploiting both dependencies on environmental factors and on a similar diagnostic feature, seeking an improvement over the model that considers only one of them. The correlation analysis could be useful to understand which of all the available environmental variables can improve the prediction and therefore, bring a potential advantage for damage detection: in this regard, in section 6, different regression models of two diagnostic feature of the sanctuary are built by changing the input variables which have been selected based on the results of the correlation analysis. To create the aforementioned models, it was decided to use SVM, harnessing all the benefits of a machine learning based approach.

2.3.1 Support Vector regression

SVM analysis is a popular machine learning tool for classification and regression, introduced by Vapnik *et al.* [58,59]. Both learning problems are based on the same idea with the difference that in a classification problem, much more commonly used in SHM, the output is represented by class labels (i.e. it has a discrete domain), while in a regression problem, such as the one addressed here, the output is a continuous variable.

Given a training dataset composed of N P-dimensional observations $\boldsymbol{x}_i \in R^P$ and their scalar outputs $y_i \in R$, the goal of a regression problem is to find the function $y = f(\boldsymbol{x})$ that better generalizes a new data set, different from the training one. To do this, SVM considers a linear function, as:

$$y = \langle \boldsymbol{w}, \boldsymbol{x} \rangle + b \tag{4}$$

where $\boldsymbol{w}$ and b are the parameters to be adjusted to reach the best fit. The regression problem is based on the Structural Risk Minimization (SRM) principle, i.e. the use of an appropriate risk functional, which not only involve the minimization of the error on the training data set -*Empirical Risk Minimization* (ERM) principle- but also the minimization of an upper bound on the generalization error. This allows to avoid over-fitting and thereby improves the generalization performance on an unknown dataset. The structural risk also depends on a loss function. In the case of SVM a *ε-Insensitive Loss-Function* is used, which derives from the concept of *robust regression* introduced by Huber, to which an insensitive zone is added, defined by ε, named *insensitive parameter*. This implies that the loss function will be set equal to ε if the discrepancy between the expected and actual value is less than ε. Introducing some *slack* variables and Lagrange multipliers and considering the above-mentioned characteristics of the algorithm, the optimization problem is solved by maximizing an

objective convex functional. It depends on a constant *C*, called *box-constraint*, which is a positive value that controls the penalty imposed on observations that lie outside the margin defined by ε. These two parameters, which should be defined a priori, are usually selected by implementing a cross validation procedure.

The linear formulation can be easily generalized to the case of non-linear regression using the so-called *kernel methods*: through a mapping function, the data of the original space (in which they are difficult to separate) are mapped into a high-dimensional space, where they can be separated more simply. The linear SVM algorithm is then conducted in the feature space, which represents nonlinear SVM operation in the original space and thanks to the popular *kernel trick*, the computational burden is reduced as the calculation of the high dimensional coordinates of the points can be avoided.

2.3.2 Measures of goodness of fit

Regardless of the regression technique used, the regression model that best fits data can be defined using some measures of *goodness of fit* [60-64], which describe how well the model fits a set of experimental data. One of the most used measures is the *Root Mean Square Error* (RMSE) that is the standard deviation of the residuals, i.e. prediction errors. Residuals are a measure of how far from the regression line data points are. In other words, RMSE indicates how concentrated the data is around the line of best fit. This measure is commonly used in climatology, forecasting, and regression analysis to verify experimental results. RMSE is defined by the following equation:

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2} \tag{5}$$

Where *n* is the number of y_j and $\hat{y}_j$ is the predicted data. Then, in general, the lower the RMSE, the better the model fits the data. The higher the RMSE, the more the data moves away from the model, which in this specific case means that something unexpected is happening in the system such as damage or an anomaly in general.

3. CASE STUDY: CONTINUOUS DATA MONITORING ACQUIRED ON THE SANCTUARY OF VICOFORTE

The Sanctuary of Vicoforte, a monumental church of the 17th century, is a unique case in terms of importance in the Italian Cultural Heritage framework. The majestic dome that covers the structure, is the largest masonry oval dome in the world with the axes of 37.15 and 24.80 m (Figure 1) [65].

Figure 1: The Sanctuary of Vicoforte: external and internal views of the dome.

The construction of the Sanctuary began in 1596 and ended in 1735. Since the first years of its construction, this stunning masonry building has suffered from significant structural problems, partly due to the settlement of the foundations. In this regard, in 1983, concerns over the structural health state of the building prompted the decision to undertake analysis, monitoring activities and strengthening interventions [66].

The extended network of cracks that affects a large portion of the buildings becomes particularly severe in the drum-dome system, where the cracks are oriented along typical meridian directions with maximum openings at window locations and an increased density next to the buttress. In 1987, because of a severe cracking configuration and the possible yielding or rupture of the original three levels of circumferential iron rings located between dome and drum, a new strengthening system was put in place. The intervention consisted of 56 tie-bars tensioned, slightly stressed by jacks, placed within holes drilled in the masonry along 14 tangents of the perimeter. The interface between two adjacent bars consists of a steel frame placed in the masonry, to ensure continuity. The bars were equipped with tensioning devices and load cells for an instantaneous reading of the load values, placed at both ends. Bars were re-tensioned 10 years after installation, in 1997.

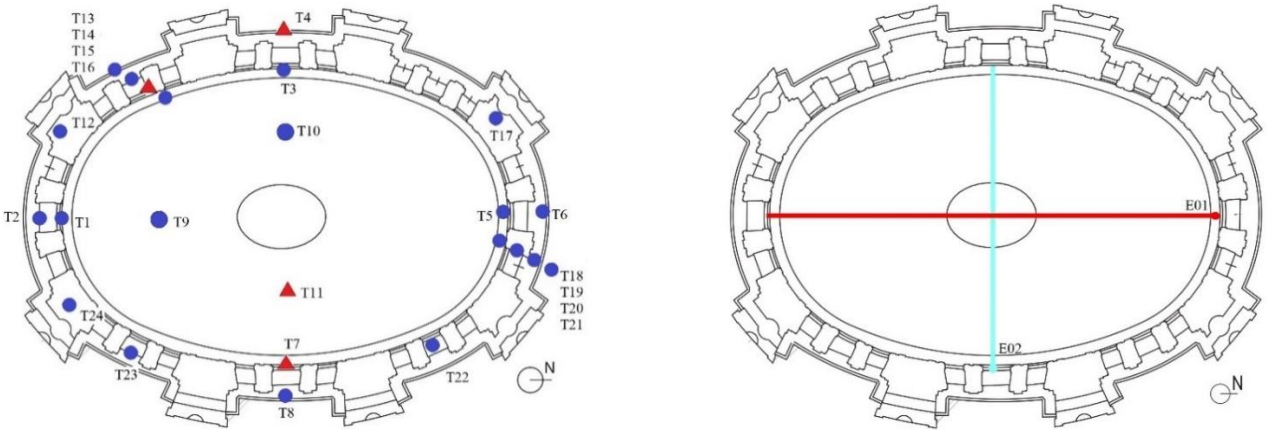
During the last decades, several non-destructive investigations were conducted, including some geophysical tests [67] and a dynamic identification campaign which allowed the updating of a first FE model of the structure [68,69].

In 2004 a static monitoring system was installed to check the effectiveness of the strengthening system and the propagation of the crack [39]. However, the static monitoring system provides only local information about the health state of the structure. In order to solve this shortcoming and to investigate the seismic behaviour of the Sanctuary, a permanent dynamic monitoring system was installed 10 years later [70].

3.1 Static monitoring system

In 1983 a first instrumentation to check the evolution of the significant crack patterns was installed. In the following decades, several upgrades of the static monitoring system followed. In 2004 the latest monitoring system was installed and the acquisition procedure was automated.

The devices composing the static monitoring system can be grouped into two main categories depending on the measurements recorded: sensors of strains, stresses and crack width and instruments that acquire measurements of environmental conditions. The static monitoring system includes 12 crack meters (of which 2 are damaged), 20 pressure cells to determine the stress in the masonry, 56 load cells to control the load condition of the bars, 2 orthogonal wire gauges to measure the axis of the dome, 24 temperature sensors and 3 piezometric electric cells. The layout of the thermometers, wire gauges, load cells and crack meters are reported in Figure 2.



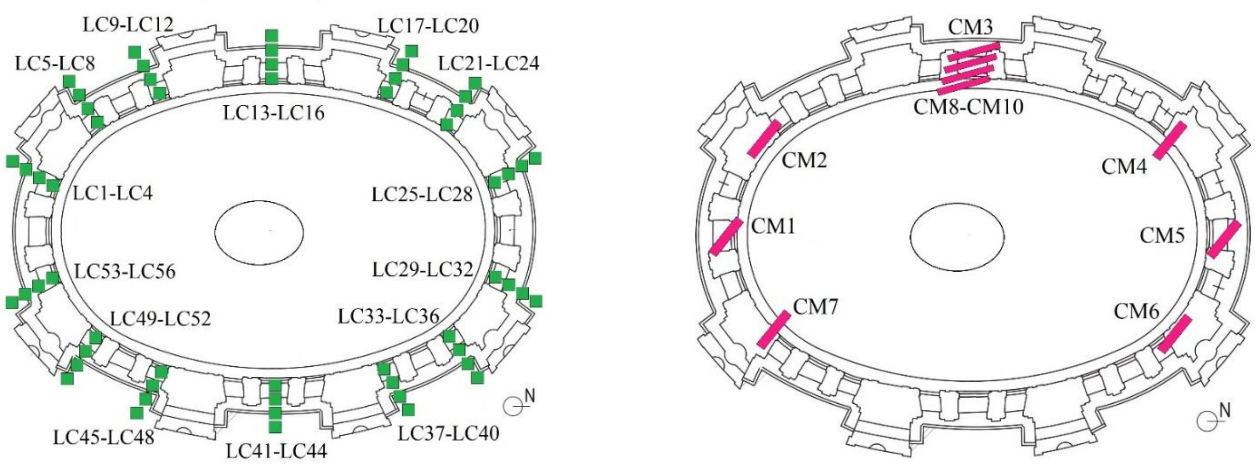


Figure 2: Layout in plan of thermometers, wire gauges, load cells and crack meters

The crack meters (LVDTs) check the evolution of the main cracks and are placed at the bottom of the dome. The temperature sensors are installed inside and outside the Sanctuary, in the hole of the staircases, on the tie bars and in the extrados of the dome. The load cells monitor the structural efficacy of the strengthening interventions. The two wire gauges measure the deformation of the minor and major axes of the dome.

The static monitoring system began collecting data in 2004 and all data acquired until 2015 were regularly processed and analysed, showing the substantial efficacy of the tie-bar system in containing displacements but also revealing measurement anomalies in some instruments [39]. The renewal of the static monitoring data acquisition and management system is now scheduled, also for integration with the dynamic one.

3.2 Dynamic monitoring system

In order to have a global perspective on the structural behavior of the Sanctuary, a permanent dynamic monitoring system was installed in December 2015. Nevertheless, the data acquisition started only in December 2016, after a calibration process of the procedure. The position of the 12 mono-axial piezoelectric accelerometers (PCB Piezotronic, model 393B12, seismic, high sensitivity, ceramic shear ICP® accel., 10 V/g, 0.15 to 1k Hz, Resonant Frequency ≥10,000 Hz, Overload Limit ±5000 g pk, Temperature Range −50 to +180 °F) were defined through sensor placement techniques [40]. As shown in Figure 3, three orthogonal accelerometers are located at the base of the crypt to record ground accelerations, and a set of nine accelerometers are located at different levels of the lantern-dome-drum system.

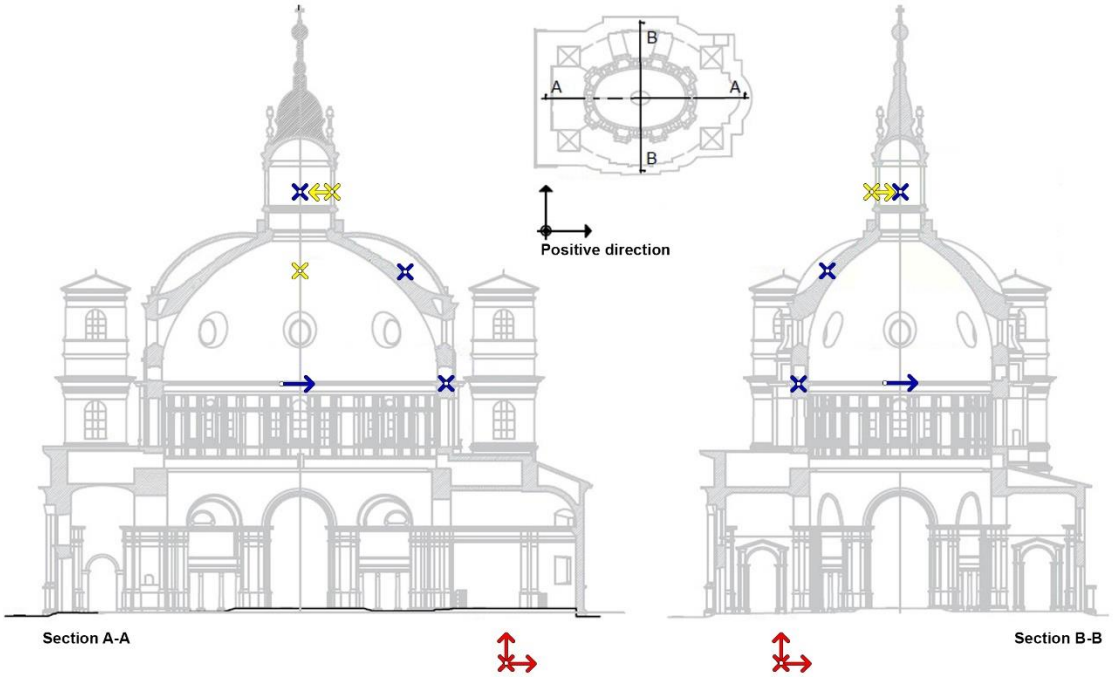


Figure 3: Location of accelerometers in plane and in sections A-A and B-B

The recorded accelerations are continuously processed to extract the modal parameters of the Sanctuary, using an automatic modal identification procedure [70] which include an algorithm belonging to the family of Stochastic Subspace Identification (SSI) techniques.

4. PROCESSING OF THE EXPERIMENTAL DATA

The first step in analysing the factors influencing static and dynamic structural behaviour of the Sanctuary is the collection, the organization and the processing of the experimental data: environmental measurements, data acquired from the static monitoring system and frequencies obtained by the dynamic monitoring system. These operations are covered in the next 3 paragraphs.

4.1 Environmental data

The environmental phenomena that are considered most significant for this case study, also based on the results of previous research [27-38], are involved in the correlation analysis. In particular, the time series of temperature and its daily excursion, humidity, wind, rain and snow have been selected. These have been requested and obtained from the ARPA Piemonte website [71] and refer to measurements recorded in

Mondovì (CN), the closest station to Vicoforte (about 8 km), in the period from 01/01/2004 to 22/06/2019. Their trend over time with a daily sampling is reported in Figure 4.

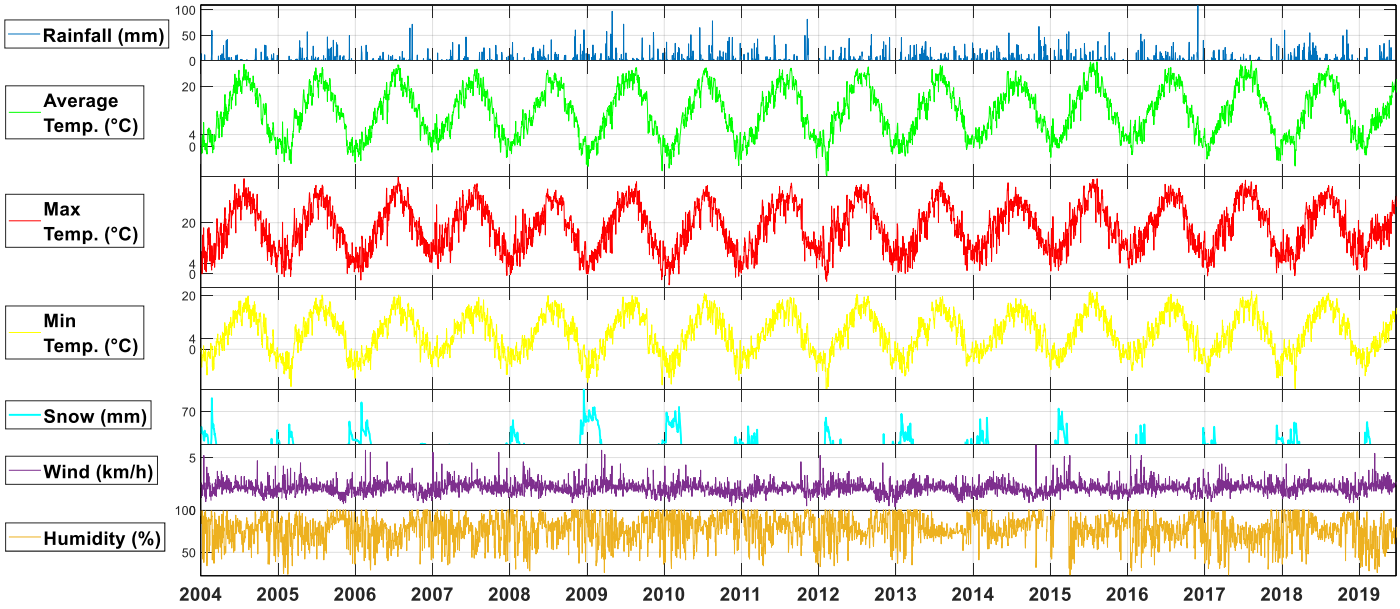


Figure 4: Environmental parameters. From the top to the bottom: rainfall, average temperature, maximum temperature, minimum temperature, snow, wind, humidity.

The graphs clearly show the seasonal trend of temperatures, which seems to follow an almost sinusoidal function. As it is reasonable to expect, snowfall occurs purely during the winter months, unlike rain, which also peaks during the warmer seasons. Wind and humidity also seem to imply a seasonal trend, even if much less than that shown by the temperature. Snowfall appears to intensify at minimum humidity values. To analyse the influence of water on the system (in the ground or in the pores of the masonry) two lines were drawn at 0 and 4°C as these represent two important points for the properties of the water: in the case of atmospheric pressure, at 0 ° C water shows the liquid/solid phase transition , whereas at 4 ° C it reaches its maximum density and minimum volume (these are properties of pure water but they can be considered a fairly reliable reference also for the water contained in the system) [72].

4.2 Dynamic data

The accelerometric signals acquired by the permanent dynamic monitoring system installed on the Sanctuary were processed through the automatic identification procedure described in [70]. In particular, each accelerometer records signals about 21 minutes long starting from the sixth minute of each hour. The signals are stored and subsequently loaded into the identification code. After a pre-processing of the acquired data,

which includes signal decimation, average and trend removal, band pass filter, low frequency component removal, signal normalization, the code selects the 5-minute signal range with higher RMS and supplies it as input for the dynamic identification process. Working with signals of 5 instead of 20 minutes allows to reduce processing times without having any influence on the quantity of identified modes: for the dynamics of the Sanctuary, 5 minutes is a time long enough to contain a suitable number of oscillations of the low structural frequency and not too long, in order to keep the calculation time at a reasonable level. It uses a time domain technique, the Stochastic Subspace Identification (SSI) algorithm. At the end, a statistical analysis of the results obtained in several identification session is carried out, evaluating the stability of the modes by varying the order of the system, as illustrated in Figure 5.

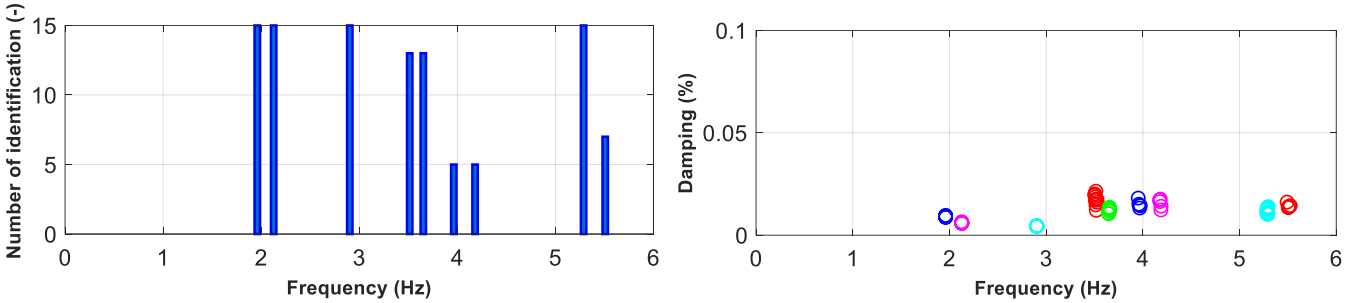


Figure 5: Example of stabilisation and cluster diagrams, output of the identification procedure

As it has been observed that the identifications from 18:00 to 6:00 return a much lower number of modes, probably due to the low level of excitement at night, the procedure processes only the daytime hours. Exclusively for the correlation analysis, only one observation per day was considered as the aim was to evaluate the relationship of the frequencies with the environmental quantities and the static parameters, both having daily sampling. It was established to use the data acquired at 12:10, because that is the time when the first and second frequencies were identified more in 2018. Figure 6 shows the trend of the first two frequencies of the Sanctuary during the year 2018, from 01/02/2018 until 31/01/2019. The first one (range 1.892-1.989 Hz), corresponds to the first bending mode in the Y direction (the direction of the minor axis of the dome oval), the second one (range 2.025-2.143 Hz) to the first bending mode in the X direction (the direction of the major axis). The gaps in the trends correspond to missing identifications, as it is quite common that some modes are not identified under general operational conditions. As the following analyses require almost continuous data, the higher modes have not been considered as they are more difficult to identify and consequently have sparsest trends.

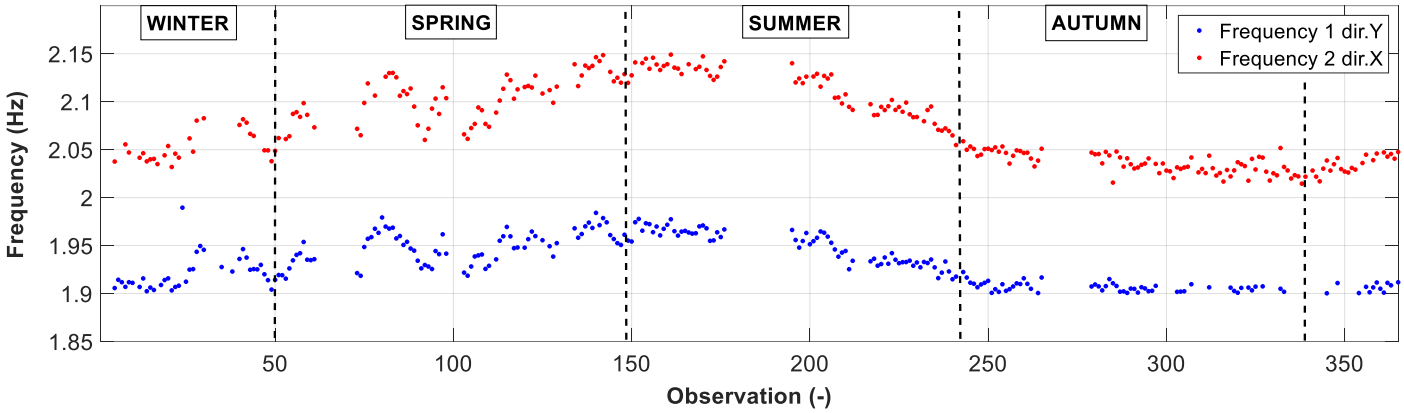


Figure 6: Annual fluctuation of the first two frequencies of the Sanctuary

Like the environmental parameters, the frequencies also seem to show a seasonal trend. At first glance it seems that the value of f_1 and f_2 increases in view of the summer and decreases in the colder months. The clearly non-stationary trend of the first two frequencies makes it difficult to recognize a possible appearance of damage and this is what arouses the need to create a predictive model. In fact, the reduction of stiffness (and consequently of frequency) associated with damage could be hidden by seasonal fluctuations and therefore could be catastrophically ignored.

4.3 Static data

As said, the static monitoring system began collecting data in 2004 and the analysis of data acquired in the following decade has been already performed and discussed in another paper [39]. Since the system has experienced periodic malfunctions over the years, mainly due to storms and lightning, the analysis has required some further elaboration to be used. In order to ground this study on verified data, only static monitoring information acquired until May 2015 was used in the present study, with the consequence that there is no contemporaneity between static and dynamic data. The plots in Figure 7 refer, respectively from top to bottom, to the data of the thermometers applied inside the structure, piezometers installed in the ground, LVDTs applied on the most significant cracks (above: all the time series; below: a single time series to better show the seasonal trend), convergence of the axes of the dome, pressure cells inserted into the masonry (as before, above: all the time series; below: a single time series to better show the seasonal trend) and load cells on the hooping system. Most of the static parameters also show seasonal fluctuations. This is certainly true for the temperature measurement of the internal walls of the Sanctuary, which is influenced by the external one. The

various thermometers seem to have a very similar temperature trend but shifted of some degrees as they are installed at different heights both internally and externally. The piezometer data are sparse. In most of the observations, no piezometric height data were recorded, probably due to sensor malfunction: for this reason, it was decided not to consider the piezometric data in subsequent evaluations. The crack meter data are very interesting: all LVDTs show seasonal data in which the cracks tend to open with the arrival of summer and to close again approaching the cold months. A very clear seasonal trend is described by the deformation records of the axes of the dome. The expansion of the axes is in line with that observed in the first two frequencies: it has peaks in correspondence with the summer months and minima in the cold seasons. In particular, although the static data acquisition is performed on a daily basis, it can be observed that the major axis deforms more than the minor one: it has been noticed that in the summer months, when the elongation of the axes is maximum, the ratio between the elongation of the minor and the major axis oscillates in the range 0.6÷0.7, in accordance with the ratio between the length of the axes, which is about 0.67. The pressure cells seem to have recorded coherent data up to about 2009. After this date, some sensors have trends completely discordant from the others, which could be due to local phenomena or sensor malfunction. For this reason, it was decided not to involve them in the correlation analysis. The recordings of the load cells in the bars appear to have credible trends with the exception of an anomalous time series (LC44) and some unusual data of another sensor in spring of 2009 (LC4). These series also seem to repeat the same cycle every year.

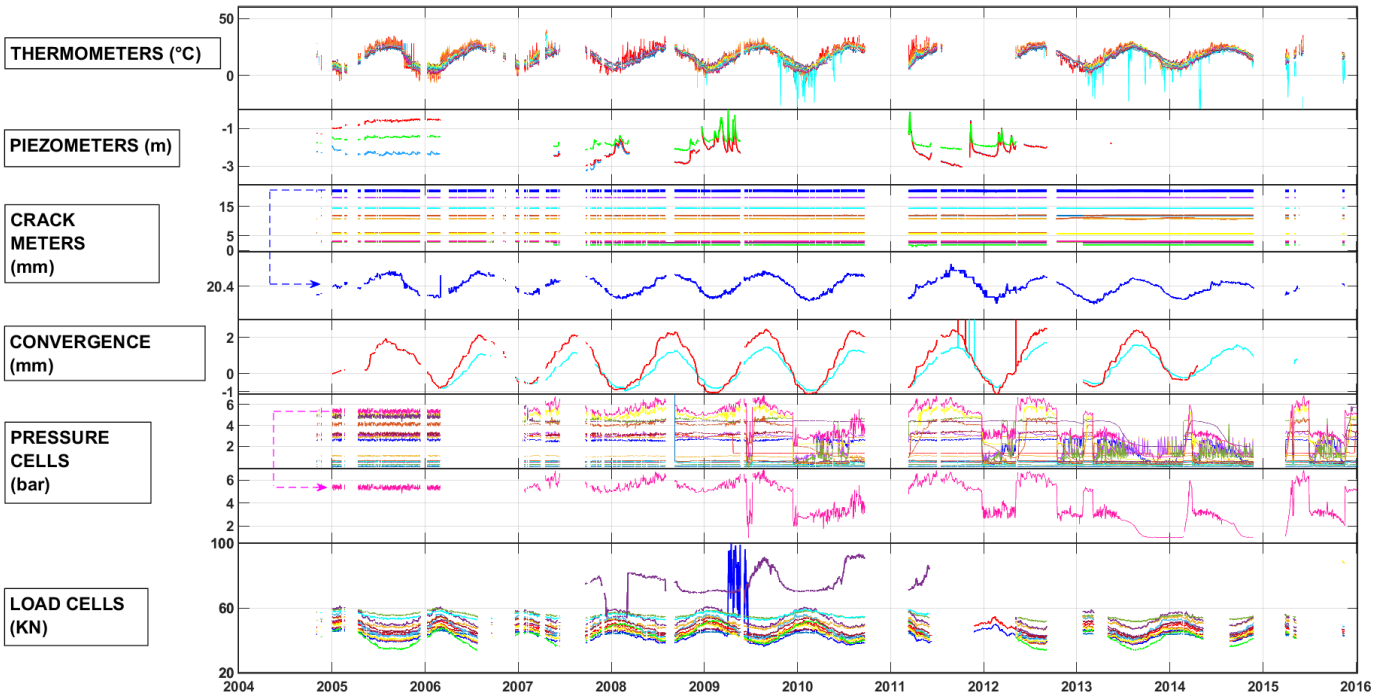


Figure 7: Static monitoring system. From the top to the bottom: thermometers, piezometers, crack meters (all and a selected time series), convergence, pressure cells (all and a selected time series), load cells.

5. CORRELATION ANALYSIS

In this section, the possible existence of correlation, i.e. any statistical relationship, is verified between heterogeneous data described in the previous chapter.

A first step before calculating the correlation coefficient and also a useful system to detect the possible dependence between variables is a visual examination of the data within a scatter plot. This operation gives the possibility of making a check on the value of r obtained, both in terms of sign (positive or negative dependence) and value (measure of the strength of relationship) and allows you to immediately notice any *strongly* non-linear dependencies that would be difficult to identify from the correlation coefficient, implicitly suggesting to use another type of approach. For *slightly* nonlinear relationships, the linear correlation coefficient is still an effective index since it generally does not change in magnitude. In addition, a visual examination enables you to identify any outliers and evaluate whether to eliminate them (if they are physically unachievable) as they could significantly influence the correlation coefficient.

First, environmental data were combined with static monitoring data. Not all sensors have been used in this analysis: for each type of data, one or a subset of representative sensors of the system was selected, through engineering and mathematical (PCA) assessments.

For example, as regards the 28 thermometers installed on the structure, they have been classified by level and exposure (inside or outside the building) and only one per category has been selected via PCA. In particular, thermometers 4, 7, 11 and 15 have been chosen, which are positioned respectively on the external and internal ellipse above the drum, inside the dome and on the hooping system. The hooping system is made up of 56 post-tensioned bars distributed over 4 levels in 14 positions of the oval. Of these, only 5 time series have been examined: with reference to the highest ring, the LC48 was selected by applying the PCA among the 14 cells and in addition, the cells adjacent to the attachment of strain gauges (i.e. LC04, LC16, LC32, LC44) have been involved. The PCA has also been applied to the data concerning the opening of the cracks, and the series of one of the most important cracks, placed on the West side has been selected. All the possible combinations have been scatter-plotted and the correlation coefficient has been calculated for each (Table 1). For reasons of space, only the most significant dependencies are reported in Figure 8.

Table 1: Correlation between environmental parameters and each type of static sensor

Combination		r				
Thermometer 4, 7, 11, 15	$T_{ext,med}$	0,94	0,94	0,93	0,89	
	$T_{ext,max}$	0,92	0,93	0,91	0,86	
	$T_{ext,min}$	0,89	0,9	0,87	0,87	
	Snow	-0,53	-0,53	-0,56	-0,57	
	Rain	-0,10	-0,13	-0,08	-0,06	
	Humidity	-0,16	-0,17	-0,17	-0,12	
	Wind	0,07	0,07	0,10	0,08	
Convergence minor and major axis	$T_{ext,min}$	0,80		0,43		
	$T_{ext,med}$	0,80		0,40		
	$T_{ext,max}$	0,76		0,37		
	Snow	-0,46		-0,27		
	Humidity	-0,03		0,06		
	Rain	-0,01		0,15		
	Wind	-0,007		-0,08		
Crack meter	$T_{ext,med}$	0,72				
	$T_{ext,min}$	0,71				
	$T_{ext,max}$	0,69				
	Snow	-0,42				
	Humidity	-0,06				
	Rain	0,04				
	Wind	-0,001				
Load cell	$T_{ext,med}$	-0,34	-0,87	-0,75	-0,70	-0,73

LC04, LC16, LC32, LC44, LC48	$T_{ext,min}$	-0,34	-0,86	-0,75	-0,67	-0,73
	$T_{ext,max}$	-0,32	-0,83	-0,70	-0,66	-0,69
	Snow	0,13	0,53	0,45	0,11	0,49
	Rain	0,09	0,08	0,04	0,04	0,05
	Wind	0,04	0,05	0,03	0,05	0,06
	Humidity	0,02	0,01	-0,02	0,01	0,01

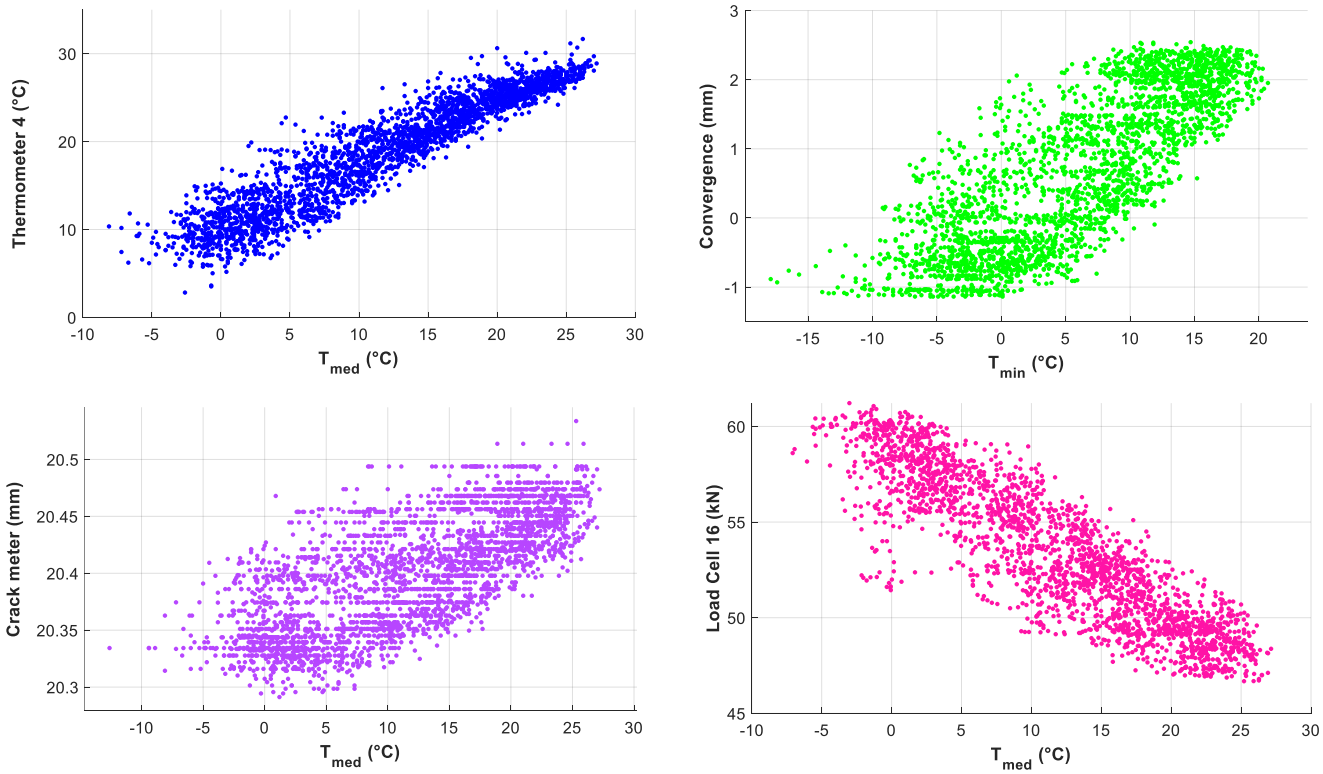


Figure 8: Correlation between static sensors and $T_{ext,med}$

The same procedure was repeated on the dynamic data, and the respective results are shown in Table 2. As above, in Figure 9 the scatter plots with the most significant dependencies are reported.

Table 2: Correlation between environmental parameters and frequencies

Combination		r
frequency 1	$T_{ext,max}$	0,72321
	$T_{ext,med}$	0,72063
	$T_{ext,min}$	0,64754
	Wind	0,2715
	Humidity	-0,2642
	Snow	-0,0877
	Rain	0,08523
frequency 2	$T_{ext,med}$	0,85214
	$T_{ext,max}$	0,84094
	$T_{ext,min}$	0,79134
	Wind	0,27
	Rain	0,13577
	Snow	-0,1181
	Humidity	-0,0285

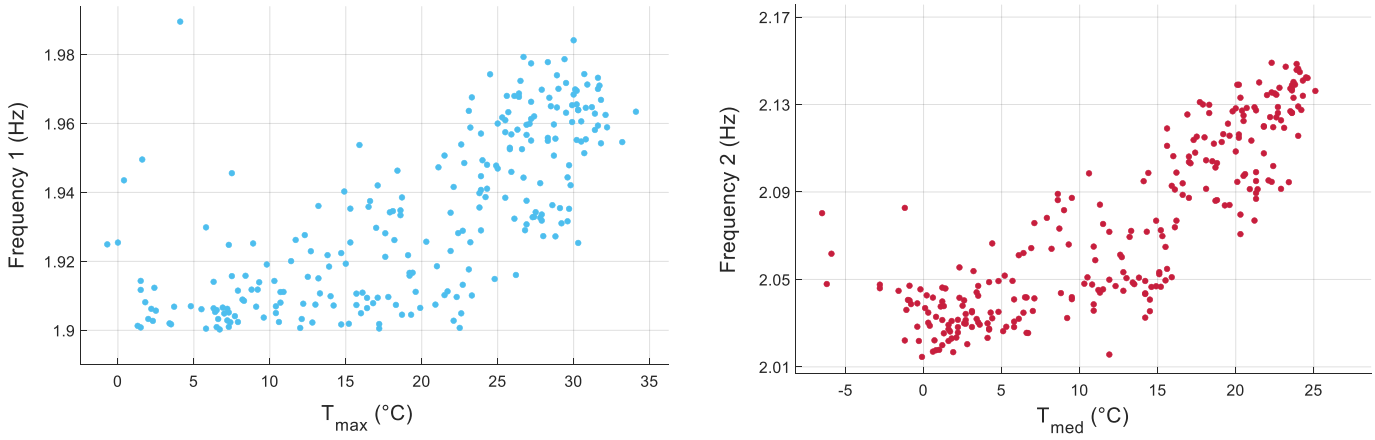


Figure 9: Stronger correlation between environmental parameters and frequencies.

The results of correlation analysis suggest that both static and dynamic behaviour of the Sanctuary are mostly affected by environmental temperature fluctuations. In fact, static and dynamic data series do not show a significant correlation with other environmental parameters taken into consideration, such as humidity, wind and rain; a weak correlation between snow and recorded data series is attributed to a secondary effect of temperature fluctuations. The most significant results among the static data are related to the analysis of masonry temperature, load in the bars and crack meters. The coefficients show that the increase in the external temperature corresponds to an increase of internal temperature in the masonry with a time lag that range from 10 up to 30 days depending on the position of the sensor, due to the thermal inertia of the material. The increase in temperature results in crack opening at the level of the drum-dome system, which is accompanied by a decrease of the load in the tie-bars. This is predictably due to the fact that the steel of the bars expands more than the masonry does, due to the difference in their thermal expansion coefficients. In this situation, the bars tend to compress and therefore the tension decreases [39]. The recordings of the wire gauges on the dome suggest that the masonry of the structure expands in the summer months and shrinks in cold periods, i.e. the elongation measurement shows a directly proportional trend to the external temperature.

The first natural frequency tends to increase as the external temperature rises, except for low temperature: a bilinear behaviour with slope change for low temperatures is suspected, as observed in other case studies too [73] (this aspect is the subject of current research of the authors). A simple normalization (each distribution

has been centred and scaled to have mean equal to 0 and standard deviation equal to 1, to get its *z-score*) and superimposition of the frequency and temperature data (Figure 10) allows us to make some observations.

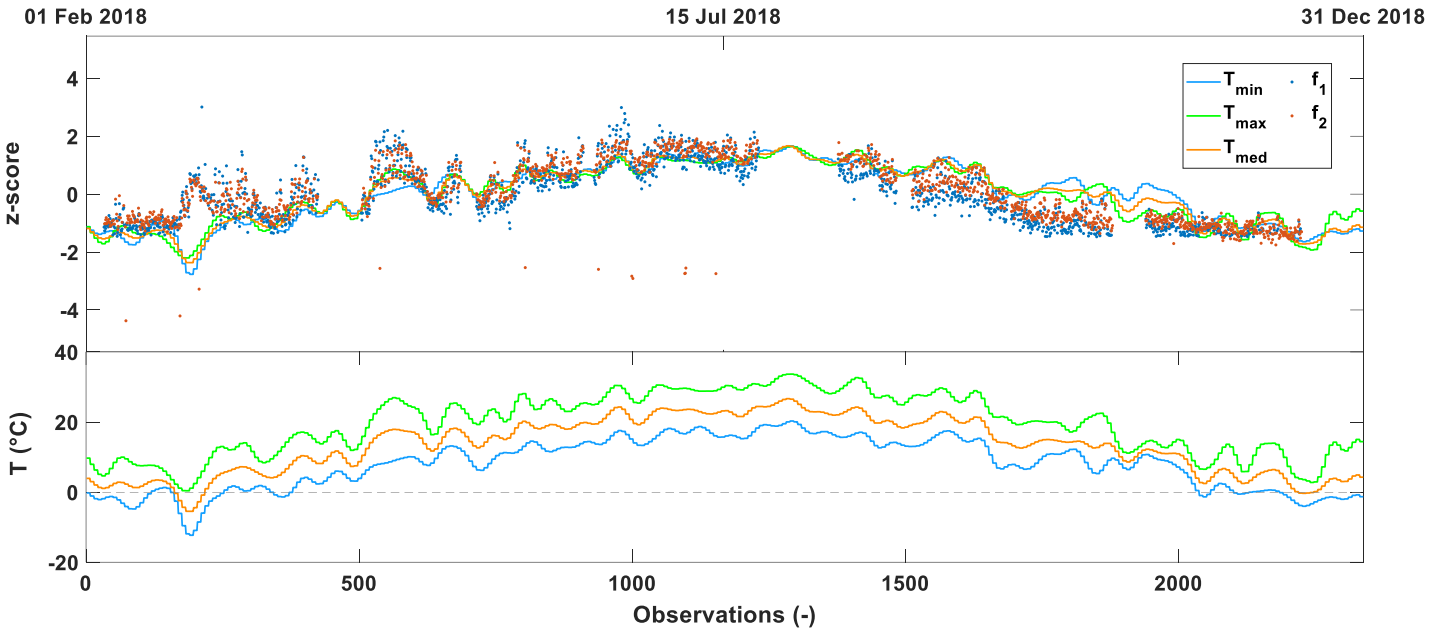


Figure 10: Above: normalized time series of frequencies with superimposed external temperatures. Below: Measured external temperature.

These are also data belonging to 2018, but considering seven frequency data per day, as they were obtained with the automatic procedure. The trend of frequencies seems to follow quite precisely that of temperatures in almost the whole set except in the first part. Around observation 200 the two frequencies show a peak while temperatures seem to display the same shape but reversed: note that in all the other peaks the two quantities seem to move parallel, but in these observations, they have the opposite behaviour. By placing a plot of non-normalized temperatures below the one just described, it was seen that this is the only range in which temperatures drastically drop below 0°C. This is a further indication of the bilinearity of the frequency-temperature behaviour when the latter approaches negative values. A possible interpretation, already suggested in a previous study on dynamic monitoring of the same building [41], is that related to the effect of ice, which is known to increase appreciably the structural stiffness [73]. An explanation of this effect could be found in the different values taken by the coefficient of thermal expansion of water/ice and solid material as the temperature changes. The differential values would be able to create a kind of "stress stiffening" effect at the micro scale, which would determine an increase of the elastic modulus at bigger scales. Then, because the coefficient of thermal expansion of liquid is much higher than that one of solid materials, all the changes in stiffness (and thus in natural frequencies), could be explained by the variation of the coefficient of thermal

expansion of water with temperature. However, this behaviour is still argument of study by the authors and the hypothesis will be verified once greater set of dynamic data will be available.

Figure 10 also exhibits a greater dispersion of the frequencies when the three temperatures diverge, that is, there is a greater thermal excursion during the day. Even more importantly, the increase in frequency seems to be in some contradiction with the observed increase in crack opening: generally, if the cracks open, a system should become weaker and therefore register a decrease in frequencies, yet the opposite would seem to happen. Probably other factors such as the stiffness of the ground, the depth of groundwater and the effect of the temperature on material parameters as mentioned above, etc. influence the dynamics of the Sanctuary, especially lower vibration modes.

6. APPLICATION OF THE RESULTS OF THE CORRELATION ANALYSIS TO STRUCTURAL BEHAVIOR MODELS

A model is always a simplification of reality as it is impossible (for both a perceptual and computational limit) to take into account all the phenomena that influence a system; on the other hand, some phenomena which seem to influence the system but which are actually unrelated, could deviate the model if considered. The careful selection of predictors also serves to avoid making a model unnecessarily too complex, by incorporating variables that contain the same information. For these reasons, the study of the most significant phenomena and the accurate selection of the most influencing predictors could lead to a model that more faithfully reproduces the real system behaviour without being too complicated (both from a computational and data acquisition point of view). In this chapter, the relationships discovered among the variables are used to strengthen the model of two diagnostic features: the first frequency of the sanctuary and the measurement of the load inside the circling system of the dome.

6.1 f_1 regression model

A regression model of the first frequency, created on a subset of data collected on the undamaged structure, is intended to reproduce the "healthy" behaviour of the sanctuary as the environmental / operational conditions vary. Theoretically, a model that reproduces in a very precise way the healthy behaviour of the structure,

should have less difficulty in recognizing an anomaly, even the most modest, and for this reason, building a performing model represents a challenge of great interest in the field of SHM and damage detection.

In this section the performances of different SVM regression models of the first frequency of the sanctuary will be compared. In particular, the model based on a single predictive variable, i.e. f_2 is compared with models that consider additional predictors, i.e. some environmental variables examined in the previous chapters. The data collected in 2018 were considered and observations in which the first or second modes were not identified were neglected. In Table 3, the different models with their combinations of predictor variables were reported together with their regression quality indicators. The training set of all the models consists of the range of observations between 280 and 700 (period from 28 March to 01 July), as the most significant oscillations of the whole data set occur in it. For each model, the SVM parameters (insensitive parameter, box constraint, Kernel function and its parameters in case of nonlinear regression) were selected through a five-fold cross validation. For reasons of space only the regression model with the best performance has been reported in Figure 11, and in order to facilitate comparison, it is overlapped on the reference model, which has f_2 as single predictor.

Table 3: goodness of fit indices of the regression models of f_1

Predictors of $f_{1,SVM}$	$RMSE$
f_2	0,007
f_2 , $T_{ext,med}$	0,0061
f_2 , $T_{ext,max}$	0,0064
f_2 , $T_{ext,min}$	0,0061

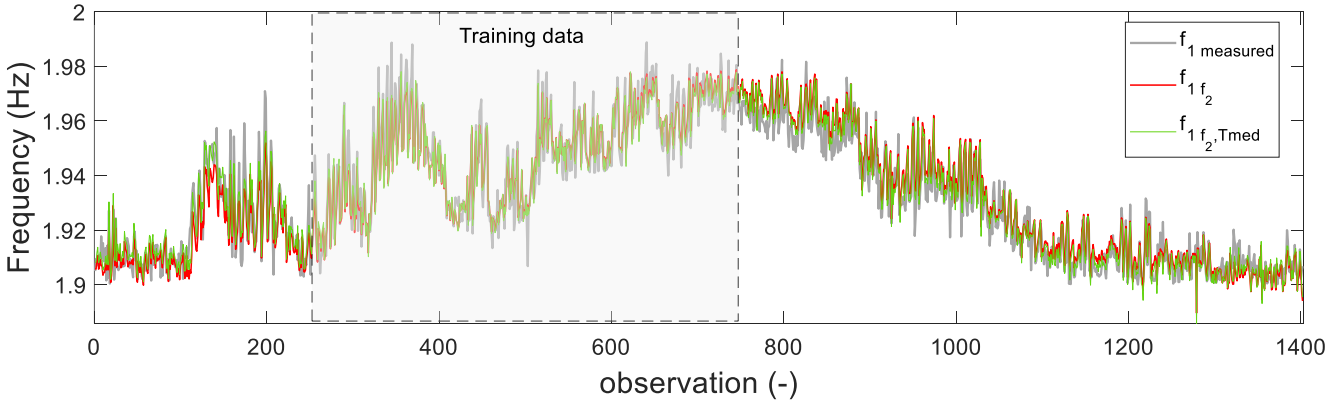


Figure 11: Comparison between the measured value of f_1 (gray), the reference model (red) and the model with the lower RMSE (green).

The residue of the models obtained with the addition of predictors from the correlation analysis is generally smaller than that obtained with the only predictor f_2: the RMSE has decreased. This means that the information (or simply, the data) selected through the correlation analysis, has led to a more precise prediction of the dynamics of the system under normal conditions.

6.2 Regression model of the load in the cerclage system

The same procedure done for the first structural frequency was repeated for the load measured in one of the bars forming the dome strengthening system. This variable can also be connected to an indicator of the structural health since the appearance of a damage most likely causes a redistribution of the loads and therefore a variation of the load within the cerclage system (see in [39] the discussion on the LC05 and CM02).

Unlike the previous case, the variable to be modelled -LC44- shows a trend conditioned not only by EOVs, but it exhibits an anomaly already declared in [39]. This represents a good opportunity to test the model also on data that deviate from the normal conditions of the structure. The load in the aforementioned cell is characterized by a non-monotonous trend, with an increase in tension in the years between 2008 and 2012. The same trend, albeit in a much lighter way, was recorded in some of the cells to the east, including the LC48, the bar adjacent to 44 (see Figure 2). It should be noted that, although this anomaly is appreciable even just by looking at the LC44 trend (see the last graph in Figure 7), here we want to evaluate a possible improvement in its prediction, which could be decisive in view of an automatic damage detection procedure (i.e. without an operator who visually examines the data but through the definition of appropriate damage thresholds) even in cases of more modest anomalies.

The reference model of the LC44 was created using the LC48 cell as a predictor and training it on the data corresponding to the normal condition (i.e. data from 2004 to 2008). Also in this case, the SVM algorithm has been applied and the same procedure for the selection of hyperparameters has been adopted. The reference model was then compared with models that exploit environmental variables as predictors. The results in terms of RMSE, both in the healthy and in the anomalous sections are summarized in the Table 4, while the model with lower $RMSE_{healty}$ is shown in the Figure 12. Unlike the "healthy" section, in the anomalous tract, the model that will have the highest $RMSE_{anomaly}$ is considered the one with the best performance, because it detects the irregularity of the data more clearly.

Table 4: Goodness of fit indices of the regression models of LC44

Predictors of $LC44_{SVM}$	$RMSE_{healty}$	$RMSE_{anomaly}$
$LC48$	0,2661	26,15
$LC48$, $T_{ext,med}$	0,2453	26,22
$LC48$, $T_{ext,max}$	0,2417	26,92
$LC48$, $T_{ext,min}$	0,234	27,01
$LC48$, T7	0,2601	27,58
$LC48$, T8	0,2752	27,34

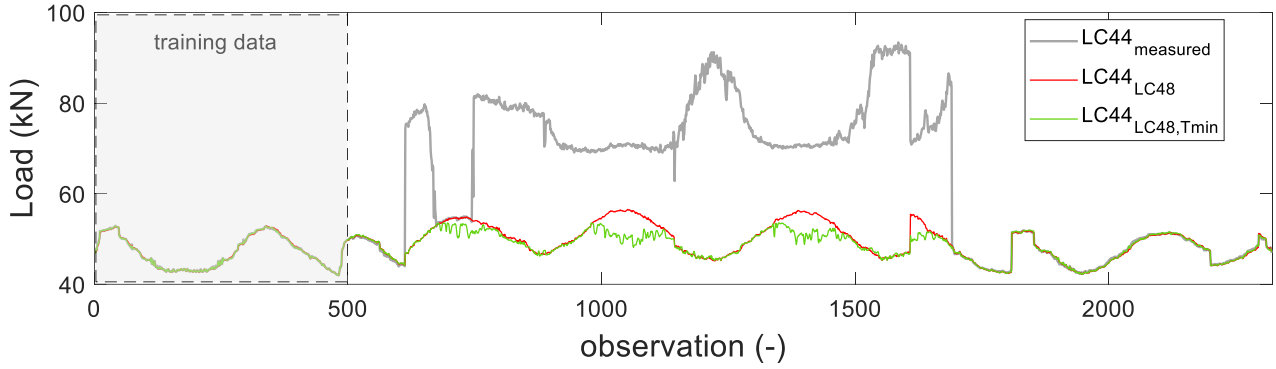


Figure 12: Comparison between the measured value of LC44 (gray), the reference model (red) and the model with the lower RMSE (green).

Compared to the case of the first frequency, two further models were created by exploiting the measurement of the thermometers applied on the structure, as they refer to the same period. In this case, the data collected by T7 and T8 thermometers were selected, which are closest to the LC44. The quality measures relating to these two models are shown in the last rows of the table. Both are midway between the reference model and the models that depend on the external temperature. It would be reasonable to think that the measurement of the thermometers installed directly on the masonry could give a more precise indication of the system temperature than the external temperature, because in some way it considers the thermal inertia of the material. On the other hand, unlike the load measurement in the cerclage system which reflects a global condition, the recordings of the thermometers have a local connotation and are strongly influenced by their location and exposure and this probably justifies the results obtained. An analysis of the spatial distribution of temperature in the system and the consequent positioning of the sensors could improve the reliability of the data in relation to this type of diagnostic analysis.

However, it must be emphasized that the annexation of the data concerning the temperature, both external and of the masonry, brings an improvement compared to the reference model.

7. CONCLUSIONS

This paper presents the systematic study of heterogeneous monitoring data and their potential application in the context of SHM. A theoretical indication of which phenomena can condition the structural behaviour, how and in which extent is given through an orderly and systematic interweaving of all the environmental and monitoring measures available for a system. Having heterogeneous measures available, coming from static, dynamic and environmental sensors, the data analysis was methodically organized setting up a first study by category, in which the variables vs. time are examined. In this phase it is also possible to select the data to be analysed in the second stage, through mathematical tools and engineering evaluations, in cases where there are many sensors of the same type available that return the same information. The intersection of the different categories of data constituted the second phase of this study. In particular, the static and dynamic monitoring vs. environmental data were studied, bringing to light both correlations that are already encountered in literature, such as that between frequencies and temperature, and unexpected ones, such as that between temperature and width of the cracks. The annexation of the data selected through the correlation analysis as predictors of regression models led to a promising improvement in making the prediction more accurate and therefore presumably more performing from the perspective of damage detection. Possible future works will include: the crossing of static and dynamic data, as soon as contemporary series are available; the study of the spatial distribution of temperature on a structure in relation to the external one, considering its geometry, material and phenomena such as thermal inertia and exposure and the consequent optimal positioning of the thermometers; in the specific case of the Vicoforte sanctuary, the bilinear behavior between temperature and frequencies and the unexpected trend of the opening of the cracks will be studied in depth.

This research and the correlations found could represent a reference for historical structures with similar dimensions. The undertaken procedure has the advantage to be easily generalizable and applicable to data other than that treated in the case study of the Sanctuary, depending on the type of monitoring system installed on the building. The results obtained could also represent a tool to guide the decision on which instruments to install / restore / enhance to optimize the health state monitoring of the specific building, of course by integration of economic assessments, and for structures belonging to the same category.

Acknowledgments

This research was partially supported by the Amministrazione del Santuario di Vicoforte and the Fondazione Cassa di Risparmio di Cuneo.

REFERENCES

[1] Sohn, H., Farrar, C. R., Hemez, F. M., Shunk, D. D., Stinemates, D. W., Nadler, B. R., & Czarnecki, J. J. (2003). A review of structural health monitoring literature: 1996–2001. *Los Alamos National Laboratory, USA*, 1-7.

[2] Brownjohn, J. M. (2007). Structural health monitoring of civil infrastructure. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *365*(1851), 589-622.

[3] Rytter, A. (1993). *Vibrational based inspection of civil engineering structures* (Doctoral dissertation, Dept. of Building Technology and Structural Engineering, Aalborg University).

[4] C. Boller, F.-K. Chang, Y. Fujino (2009). *Encyclopedia of structural health monitoring. Part 9: Civil Engineering Applications.* John Wiley & Sons Ltd.

[5] Ubertini, F., Cavalagli, N., Kita, A., & Comanducci, G. (2018). Assessment of a monumental masonry bell-tower after 2016 Central Italy seismic sequence by long-term SHM. *Bulletin of Earthquake Engineering*, *16*(2), 775-801.

[6] Russo, S. (2013). On the monitoring of historic Anime Sante church damaged by earthquake in L'Aquila. *Structural control and health monitoring*, *20*(9), 1226-1239.

[7] Boscato, G., Dal Cin, A., Russo, S., & Sciarretta, F. (2014). SHM of historic damaged churches. In *Advanced Materials Research* (Vol. 838, pp. 2071-2078). Trans Tech Publications Ltd.

[8] Ramos, L. F., Aguilar, R., Lourenço, P. B., & Moreira, S. (2013). Dynamic structural health monitoring of Saint Torcato church. *Mechanical Systems and Signal Processing*, *35*(1-2), 1-15.

[9] Gentile, C., Guidobaldi, M., & Saisi, A. (2016). One-year dynamic monitoring of a historic tower: damage detection under changing environment. *Meccanica*, *51*(11), 2873-2889.

[10] Bartoli, G., Chiarugi, A., & Gusella, V. (1996). Monitoring systems on historic buildings: the Brunelleschi Dome. *Journal of structural engineering*, *122*(6), 663-673.

[11] Lorenzoni, F., Casarin, F., Modena, C., Caldon, M., Islami, K., & da Porto, F. (2013). Structural health monitoring of the Roman Arena of Verona, Italy. *Journal of Civil Structural Health Monitoring*, *3*(4), 227-246.

[12] Forgács, T., Sarhosis, V., & Bagi, K. (2018). Influence of construction method on the load bearing capacity of skew masonry arches. *Engineering Structures*, *168*, 612-627.

[13] D'Altri, A. M., Sarhosis, V., Milani, G., Rots, J., Cattari, S., Lagomarsino, S., Sacco, E., Tralli, A., Castellazzi, G. & De Miranda., S. (2019). Modeling strategies for the computational analysis of unreinforced Masonry structures: Review and classification. *Archives of Computational Methods in Engineering*, 1-33.

[14] Bal, İ. E., Dais, D., Smyrou, E., & Sarhosis, V. (2020). Monitoring of a Historical Masonry Structure in Case of Induced Seismicity. *International journal of architectural heritage*, 1-18.

[15] Farrar, C. R., & Worden, K. (2012). *Structural health monitoring: a machine learning perspective*. John Wiley & Sons.

[16] Zonta, D., Glisic, B., & Adriaenssens, S. (2014). Value of information: impact of monitoring on decision-making. *Structural Control and Health Monitoring*, *21*(7), 1043-1056.

[17] Straub, D., Chatzi, E., Bismut, E., Courage, W., Döhler, M., Faber, M. H., Köhler, J., Lombaert, J., Omenzetter, P., Pozzi, M., Thöns, S., Val D. V., Wenzel, H., Zonta, D. (Aug 2017). Value of information: A roadmap to quantifying the benefit of structural health monitoring. *ICOSSAR - 12th International Conference on Structural Safety & Reliability*, Vienna, Austria.

[18] Cappello, C., Zonta, D., & Glišić, B. (2016). Expected utility theory for monitoring-based decision-making. *Proceedings of the IEEE*, *104*(8), 1647-1661.

[19] Masciotta, M. G., Ramos, L. F., & Lourenço, P. B. (2017). The importance of structural monitoring as a diagnosis and control tool in the restoration process of heritage structures: a case study in Portugal. *Journal of Cultural Heritage*, *27*, 36-47.

[20] ICOMOS charter- Principles for the analysis, conservation and structural restoration of architectural heritage (*2003*). Ratified by the ICOMOS 14th General Assembly in Victoria Falls, Zimbabwe, in 2003

[21] Farrar, C. R., Doebling, S. W., & Nix, D. A. (2001). Vibration–based structural damage identification. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *359*(1778), 131-149.

[22] Doebling, S. W., Farrar, C. R., Prime, M. B., & Shevitz, D. W. (1996). *Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review* (No. LA-13070-MS). Los Alamos National Lab., NM (United States).

[23] Bicanic, N., & Chen, H. P. (1997). Damage identification in framed structures using natural frequencies. *International Journal for Numerical Methods in Engineering*, *40*(23), 4451-4468.

[24] Bassoli, E., Forghieri, M., Vincenzi, L., Bovo, M., & Mazzotti, C. (2017). Structural health monitoring of a historical masonry bell tower using operational modal analysis. In *Key Engineering Materials* (Vol. 747, pp. 440-447). Trans Tech Publications Ltd.

[25] Erazo, K., Sen, D., Nagarajaiah, S., & Sun, L. (2019). Vibration-based structural health monitoring under changing environmental conditions using Kalman filtering. *Mechanical systems and signal processing*, *117*, 1-15.

[26] Fraraccio, G., Brügger, A., & Betti, R. (2008). Identification and damage detection in structures subjected to base excitation. *Experimental Mechanics*, *48*(4), 521-528.

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

[28] Ubertini, F., Comanducci, G., Cavalagli, N., Pisello, A. L., Materazzi, A. L., & Cotana, F. (2017). Environmental effects on natural frequencies of the San Pietro bell tower in Perugia, Italy, and their removal for structural performance assessment. *Mechanical Systems and Signal Processing*, *82*, 307-322.

[29] Kita, A., Cavalagli, N., & Ubertini, F. (2019). Temperature effects on static and dynamic behavior of Consoli Palace in Gubbio, Italy. *Mechanical Systems and Signal Processing*, *120*, 180-202.

[30] Deraemaeker, A., Reynders, E., De Roeck, G., & Kullaa, J. (2008). Vibration-based structural health monitoring using output-only measurements under changing environment. *Mechanical systems and signal processing*, *22*(1), 34-56.

[31] Sohn, H., Worden, K., & Farrar, C. R. (2002). Statistical damage classification under changing environmental and operational conditions. *Journal of intelligent material systems and structures*, *13*(9), 561-574.

[32] Ni, Y. Q., Hua, X. G., Fan, K. Q., & Ko, J. M. (2005). Correlating modal properties with temperature using long-term monitoring data and support vector machine technique. *Engineering Structures*, *27*(12), 1762-1773.

[33] Catbas, F. N., Susoy, M., & Frangopol, D. M. (2008). Structural health monitoring and reliability estimation: Long span truss bridge application with environmental monitoring data. *Engineering Structures*, *30*(9), 2347-2359.

[34] Ramos, L. F., Marques, L., Lourenço, P. B., De Roeck, G., Campos-Costa, A., & Roque, J. (2010). Monitoring historical masonry structures with operational modal analysis: two case studies. *Mechanical systems and signal processing*, *24*(5), 1291-1305.

[35] Cabboi, A., Gentile, C., & Saisi, A. (2017). From continuous vibration monitoring to FEM-based damage assessment: application on a stone-masonry tower. *Construction and Building Materials*, *156*, 252-265.

[36] Gentile, C., Ruccolo, A., & Canali, F. (2019). Continuous monitoring of the Milan Cathedral: dynamic characteristics and vibration-based SHM. *Journal of Civil Structural Health Monitoring*, *9*(5), 671-688.

[37] García-Macías, E., & Ubertini, F. (2020). MOVA/MOSS: Two integrated software solutions for comprehensive Structural Health Monitoring of structures. *Mechanical Systems and Signal Processing*, *143*, 106830.

[38] Barsocchi, P., Bartoli, G., Betti, M., Girardi, M., Mammolito, S., Pellegrini, D., & Zini, G. (2020). Wireless sensor networks for continuous structural health monitoring of historic masonry towers. *International Journal of Architectural Heritage*, 1-23.

[39] Ceravolo, R., De Marinis, A., Pecorelli, M. L., & Zanotti Fragonara, L. (2017). Monitoring of masonry historical constructions: 10 years of static monitoring of the world's largest oval dome. *Structural Control and Health Monitoring*, *24*(10), e1988.

[40] Ceravolo, R., De Lucia, G., Pecorelli, M., & Fragonara, L. Z. (2015, July). Monitoring of historical buildings: Project of a dynamic monitoring system for the world's largest elliptical dome. In *2015 IEEE*

workshop on environmental, energy, and structural monitoring systems (EESMS) proceedings (pp. 113-118). IEEE.

[41] Coletta, G., Miraglia, G., Pecorelli, M., Ceravolo, R., Cross, E., Surace, C., & Worden, K. (2019). Use of the cointegration strategies to remove environmental effects from data acquired on historical buildings. *Engineering Structures*, *183*, 1014-1026.

[42] Kromanis, R., & Kripakaran, P. (2014). Predicting thermal response of bridges using regression models derived from measurement histories. *Computers & Structures*, *136*, 64-77.

[43] Kromanis, R., & Kripakaran, P. (2013). Support vector regression for anomaly detection from measurement histories. *Advanced Engineering Informatics*, *27*(4), 486-495.

[44] Santos, A., Figueiredo, E., Silva, M. F. M., Sales, C. S., & Costa, J. C. W. A. (2016). Machine learning algorithms for damage detection: Kernel-based approaches. *Journal of Sound and Vibration*, *363*, 584-599.

[45] Worden, K., & Lane, A. J. (2001). Damage identification using support vector machines. *Smart materials and structures*, *10*(3), 540.

[46] Bornn, L., Farrar, C. R., Park, G., & Farinholt, K. (2009). Structural health monitoring with autoregressive support vector machines. *Journal of Vibration and Acoustics*, *131*(2).

[47] Chong, J. W., Kim, Y., & Chon, K. H. (2014). Nonlinear multiclass support vector machine–based health monitoring system for buildings employing magnetorheological dampers. *Journal of Intelligent Material Systems and Structures*, *25*(12), 1456-1468.

[48] Kim, Y., Chong, J. W., Chon, K. H., & Kim, J. (2012). Wavelet-based AR–SVM for health monitoring of smart structures. *Smart Materials and Structures*, *22*(1), 015003.

[49] Ezekiel, M. (1930). Methods of correlation analysis. Wiley, New York, USA

[50] Taylor, J. (1997). *Introduction to error analysis, the study of uncertainties in physical measurements*. 2nd ed. University Science Books, Sausalito.

[51] Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

[52] Jolliffe, I. T. (2002). *Principal Component Analysis, Second Edition*. Springer-Verlag, New York.

[53] Sohn, H., Dzwonczyk, M., Straser, E. G., Kiremidjian, A. S., Law, K. H., & Meng, T. (1999). An experimental study of temperature effect on modal parameters of the Alamosa Canyon Bridge. *Earthquake engineering & structural dynamics*, *28*(8), 879-897.

[54] Shi, H., Worden, K., & Cross, E. J. (2016, October). A nonlinear cointegration approach with applications to structural health monitoring. In *Journal of Physics: Conference Series* (Vol. 744, No. 012025). Institute of Physics.

[55] Kang, F., Liu, J., Li, J., & Li, S. (2017). Concrete dam deformation prediction model for health monitoring based on extreme learning machine. *Structural Control and Health Monitoring*, *24*(10), e1997.

[56] Magalhães, F., Cunha, Á., & Caetano, E. (2012). Vibration based structural health monitoring of an arch bridge: from automated OMA to damage detection. *Mechanical Systems and Signal Processing*, *28*, 212-228.

[57] Moser, P., & Moaveni, B. (2011). Environmental effects on the identified natural frequencies of the Dowling Hall Footbridge. *Mechanical Systems and Signal Processing*, *25*(7), 2336-2357.

[58] Vapnik, V. (1995). *The nature of statistical learning theory*.

[59] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

[60] D'Agostino, R. B. (1986). *Goodness-of-fit-techniques* (Vol. 68). CRC press.

[61] Freedman, D. A. (2009). *Statistical models: theory and practice*. Cambridge university press.

[62] Tanner, M. A. (2012). *Tools for statistical inference*. Springer-Verlag New York.

[63] Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (Vol. 326). John Wiley & Sons.

[64] Seber, G. A., & Lee, A. J. (2012). *Linear regression analysis* (Vol. 329). John Wiley & Sons.

[65] Aoki, T., Chiorino, M. A., & Roccati, R. (2003, January). Structural characteristics of the elliptical masonry dome of the Sanctuary of Vicoforte. In *Proceedings of the First International Congress on Construction History, S. Huerta, ed. Madrid: Instituto Juan de Herrera* (pp. 203-212).

[66] Chiorino, A., Calderini, C., Spadafora, A., & Spadavecchia, R. (2008). Structural assessment, testing, rehabilitation and monitoring strategies for the world's largest elliptical dome and sanctuary at Vicoforte. In *RILEM Symposium on On Site Assessment of Concrete, Masonry and Timber Structures-SACoMaTiS 2008* (pp. 529-538). RILEM Publications SARL.

[67] Scandella, L., Lai, C. G., Spallarossa, D., & Corigliano, M. (2011). Ground shaking scenarios at the town of Vicoforte, Italy. *Soil Dynamics and Earthquake Engineering*, *31*(5-6), 757-772.

[68] Chiorino, M. A., Spadafora, A., Calderini, C., & Lagomarsino, S. (2008). Modeling strategies for the world's largest elliptical dome at Vicoforte. *International Journal of Architectural Heritage*, *2*(3), 274-303.

[69] Ceravolo, R., De Lucia, G., Miraglia, G., & Pecorelli, M. L. (2019). Thermoelastic finite element model updating with application to monumental buildings. *Computer-Aided Civil and Infrastructure Engineering*.

[70] Pecorelli, M. L., Ceravolo, R., & Epicoco, R. (2018). An automatic modal identification procedure for the permanent dynamic monitoring of the sanctuary of Vicoforte. *International Journal of Architectural Heritage*, 1-15.

[71] http://www.arpa.piemonte.it/

[72] Sandrolini, F., Franzoni, E., Sassoni, E., & Diotallevi, P. P. (2011). The contribution of urban-scale environmental monitoring to materials diagnostics: A study on the Cathedral of Modena (Italy). *Journal of cultural heritage*, *12*(4), 441-450.

[73] Peeters, B., & De Roeck, G. (2001). One-year monitoring of the Z24-Bridge: environmental effects versus damage events. *Earthquake engineering & structural dynamics*, *30*(2), 149-171.