A new evolutionary polynomial regression technique to assess the fundamental periods of irregular buildings

A NEW EVOLUTIONARY POLYNOMIAL REGRESSION TECHNIQUE TO ASSESS THE FUNDAMENTAL PERIODS OF IRREGULAR BUILDINGS

Sebastiano Marasco[1], Gian Paolo Cimellaro[2]

[1] Postdoctoral researcher, Dept. of Structural, Geotechnical and Building Engineering, Politecnico di Torino, Italy

[2] Professor, Department of Structural, Geotechnical and Building Engineering, Politecnico di Torino, Italy, corresponding author: email: gianpaolo.cimellaro@polito.it

The main seismic design codes propose simplified formulations to evaluate the fundamental period of regular structures based on the total height. Indeed, the fundamental period depends on several parameters directly connected to the mass and stiffness of the structure and on its geometrical characteristics, including also irregularities.
This paper proposes a set of mathematical formulations to evaluate the longitudinal and transversal fundamental period of vibration of 3D Reinforced Concrete (RC) frames which have various vertical and plan irregularities and for different mechanical and geometrical design parameters.
Several types of Reinforced Concrete Bare Moment Resisting Frame (RC-BMRF) buildings have been designed according to the different versions of the Italian codes starting from 1916 to nowadays and then used as case studies. Modal analysis is performed on the entire building dataset to assess the fundamental periods in both longitudinal and transversal directions. Then, cluster analysis is carried out to classify the buildings based on similar design characteristics and construction details. Finally, a robust Evolutionary Polynomial Regression (EPR) technique is used to find the optimal polynomial forms of the natural period. Numerical results show a better performance of the proposed formulation compared with the existing methodologies available in the literature.

KEYWORDS

Evolutional Polynomial Regression, Clustering, Artificial Intelligence, fundamental period, building irregularities.

INTRODUCTION

The evaluation of the fundamental periods of RC framed buildings is an essential requirement in earthquake engineering. Current seismic design codes propose simplified mathematical expressions to assess the natural period of vibration based on the total height of the frame. These equations have been derived from the regression analysis of the statistical dataset consisting of periods measured during past earthquakes. Most of the empirical formulations assume the form of Equation (1).

$$T_1 = \alpha \cdot H^{\beta} \tag{1}$$

where α is a coefficient that depends on the structural system. ATC3-06 [1] fixed the values 0.75 and 0.06 for β and α, respectively; while H is the total height of the MRF building expressed in meters. Analogously, European seismic design regulation [2] adopted a value of 0.075 and 0.75 for α and β, respectively. ASCE 7-16[3] proposed a value of 0.0466 and 0.90 for α and β, respectively. Goel, Chopra [4] demonstrated that the semi-empirical code formula underestimates the expected values. Therefore, they proposed an alternative empirical formulation capable of improving the correlation with the measured data collected from eight Californian earthquakes, including the 1971 San Fernando earthquake and the 1994 Northridge seismic event. The same mathematical expression of Equation (1) was adopted by Goel, Chopra [4], while providing upper and lower limits of the estimates. NEHRP [5] guidelines recommended an alternative formula for both RC and steel MRF buildings multiplying the number of stories by 0.1. This mathematical expression can be used for buildings with a maximum number of stories equal to 12. Although design code formulae provide a simplified approach, it is still a challenge to enhance their accuracy by investigating the effects of additional building's parameters. The natural period of the MRFs depends on its mass distribution, strength, stiffness of the members, structural regularity (in plan and elevation), number of storeys and bays, and other construction-based aspects. Verderame, Iervolino, Manfredi [6] assessed the longitudinal and transversal elastic periods of four groups of existing RC-MRF buildings by using the general mathematical model given by Equation (2).

$$T_1 = \alpha \cdot H^{\beta} \cdot S^{\gamma} \tag{2}$$

where S is the footprint area of the building and γ is the related exponent. Different building configurations based on a given structural model were investigated. The least-square regression was performed to estimate the parameters of the Eq (3). Hong, Hwang [7] monitored more than 30 buildings in Taiwan to identify the fundamental vibration periods. An empirical formula was calibrated through regression analysis of the collected data, and the influence of certain structural parameters was investigated. Hong, Hwang [7] found that the total building height is the more relevant parameter in the period estimation, while the monitored buildings tended to be stiffer than those monitored in the US. This discrepancy highlights the code-to-code variability in period definition, and it increases when comparing results obtained from code formulae with those returned by numerical analyses. Varadharajan, Sehgal, Saini [8] dealt with the influence of the buildings' vertical irregularities in the period estimation. A single irregularity index λ was adopted to quantify the period variation due to the vertical irregularities (Equation (3)) [9,10]. Different structural configurations of irregular buildings subjected to 27 ground motions were investigated. Regression analysis was conducted to estimate the irregularity index.

$$T_1 = \lambda \cdot 0.075 \cdot H^{0.75} \tag{3}$$

Asteris, Repapis, Foskolos, Fotos, Tsaris [11] computed the fundamental period of vertically irregular RC frame buildings with infilled walls. Three different building groups were defined based on the type of vertical irregularity. For each group, 8, 12, 16, 20, and 24 stories configurations were analyzed. Results showed that the fundamental period of irregular buildings is smaller than those regular according to Equation (4).

$$\lambda = \frac{1}{N^{0.1}} \tag{4}$$

Current methodologies aim at identifying the modal characteristics of a given structure based on the Experimental Modal Analysis (EMA) measurements (Young, Adeli [12]). Wang, Zenelis [13] proposed a new technique to identify the modal characteristics by using ambient measurement data and a model-based method. Instead, Catbas, Brown, Aktan [14] used the multiple-input multiple-output technique to determine the modal properties of large structures.

The widespread use of computer hardware and numerical solvers are enhancing the capabilities of processing and testing a large amount of data from different sources. This large availability of data has led to recent developments in the field of Artificial Intelligence (AI) techniques. Among the others, Artificial Neural Network (ANN) are widely adopted techniques for solving different problems. Asteris, Tsaris, Cavaleri, Repapis, Papalou, Di Trapani, Karypidis [15] used ANN to predict the fundamental period of infilled RC 2D frames based on a dataset of 1281 frames. The number of storeys, span length, number of spans, wall stiffness, and opening percentage were assumed as representative parameters. In this regard, many studies dealt with the analytical investigation on the parameters that affect the fundamental period (e.g. Asteris, Repapis, Tsaris, Di Trapani, Cavaleri [16] Asteris, Repapis, Cavaleri, Sarhosis, Athanasopoulou [17]). Kose [18] investigated the effects of some structural parameters (i.e. building height, number of bays, shear walls area ratio, infilled panels ratio, and type of frame) on the fundamental period of RC buildings. A typical building structural configuration was modeled in SAP2000 [19] and an iterative linear modal analysis was carried out. The influence of each parameter was determined using sensitivity analysis, while an ANN procedure was employed to assess the relationship between the period and the considered parameters for 189 different computational models. Although ANNs lead to accurate solutions, they are based on self-learning capabilities which enable to produce outcomes without providing a symbolic expression. To cope with this limitation, regression analysis can be used to derive mathematical formulation based on certain explanatory variables. Young, Adeli [12] proposed a new formulation of the fundamental period of irregular moment resisting steel frame by nonlinear regression analysis. Later, Young, Adeli [20] applied the same methodology to irregular concentrically-braced steel structures. Charalampakis, Tsiatas, Kotsiantis [21] proposed a single level stepwise regression to derive simplified expressions of the fundamental period of masonry infilled RC buildings. Recently, new data-driven techniques in form of Genetic Programming (GP) have gained prominent attention in optimization problems. Joshi, Londhe, Kwatra [22] used GP technique to assess simple empirical equations of the fundamental period of buildings taking into account floor-stiffness and general features such as total height and width.

The main objectives of this work are: (*i*) developing a new set comprehensive formula capable of providing a consistent estimate of the fundamental longitudinal and transversal periods for a wide class of RC-BMRF; (*ii*) quantifying the effects of any vertical and horizontal setback irregularities on the fundamental period. An automatic procedure is developed to assign the geometrical and mechanical parameters that implement the Italian and European design rules and common engineering practices within the last century. Furthermore, a wide range of vertical and horizontal irregularities are randomly generated. Cluster analysis is adopted to classify the buildings based on certain design characteristics. Then, an Evolutionary Polynomial Regression (EPR) procedure is used to consistently estimate the modal characteristics of each building class. This AI-based approach merges a Genetic Algorithm (GA) paradigm for finding the optimal mathematical structure and the robust bi-square Weighted Least-Squares Method (WLSM) for the identification of the multi-regression parameters [23]. This effective combination produces a nonlinear mapping of numerical data obtained by the modal analyses with few constants, avoiding well-known over-fitting issues and improving the generalization of the final mathematical model.

The paper starts with the research significance of the proposed work. Then, a detailed description of the building population used in the analyses is given. Section four tackles the EPR computational procedure used to identify both optimal formulae of the longitudinal and transversal fundamental periods of RC-BMRF buildings and the effect of the vertical and horizontal setback irregularities in the estimates. Results and discussions are given in the fifth section of the manuscript.

RESEARCH SIGNIFICANCE

Current seismic codes worldwide allow using equivalent horizontal load patterns for the practical design of structures. Moreover, dynamic approaches require to select appropriate seismic input based on the fundamental mode of the structure. These methods of analysis entail the evaluation of the fundamental period of the structure. Thus, there is evidence that the fundamental period plays a crucial role in relating the seismic demand to capacity, and, therefore leading to assess the seismic performance of the structure.

Existing research recognized the importance of estimating the fundamental period of buildings by suggesting simplified formulations based on experimental studies (Goel, Chopra [4]). Several attempts have been made to develop simplified relationships that are suited to given types of building (Verderame, Iervolino, Manfredi [6]). A primary concern of the current research is to provide a straightforward formula that requires simple building information while providing accurate results (Young, Adeli [20], Young, Adeli [12]). Many authors followed an extensive analytical study on a limited number of typical existing building configurations (Crowley, Pinho [24]). Recently, the research effort was devoted to evaluate the effect of plan vertical irregularities in the period estimation (Asteris, Repapis, Foskolos, Fotos, Tsaris [11]) by performing numerical analyses (Kose [18]). In the last decades, the massive availability of data and the use of numerical solvers have seen a growing trend towards AI techniques. Many authors employed ANNs to predict the fundamental period of buildings (Asteris, Tsaris, Cavaleri, Repapis, Papalou, Di Trapani, Karypidis [15]). Whilst ANNs lead to accurate solutions, they are based on self-learning capabilities which enable to produce outcomes without providing a symbolic expression.

Although extensive research has been carried out, there is still a need to improve the existing formulations. One major practical issue concerns the size of the experimental data and the number of input parameters investigated. The experimental dataset may be representative of only a limited type of existing buildings. In other cases, the dataset lacks important building features that may strongly affect the period estimation. To cope with these issues, this paper proposes a new set of relationships to accurately estimate the longitudinal and transversal fundamental periods while accounting for several spatial setback irregularities. The number and variability of the investigated input parameters allow exploring a comprehensive building population that adequately reflects the existing Italian and European building portfolio. Finally, employing a robust-based EPR technique leads to enhance the accuracy of the estimates.

BUILDING DATASET

Different types of geometrical and mechanical parameters of RC framed buildings are investigated. The variability of the fundamental period in the two principal directions due to a wide variety of vertical and horizontal setback irregularities is also accounted for. This study analyzes a large population of 3D RC-BMRF buildings envisioned as representative of the Italian and European design regulations of the last century.

Standard building design

The RC frames are designed for being representative of the main changes of the Italian and European seismic standards (Charalampakis, Tsiatas, Kotsiantis [21]). Due to the different code requirements (e.g., load combinations, material strength classes, horizontal actions), a harmonization of the building design types is herein adopted to classify the RC frames. Three Code Levels (CLs) are defined based on the capacity to withstand seismic actions and other attributes concerning the construction practices. Table 1 resumes the principal construction details and the base design acceleration used to categorize the RC framed buildings.

Table 1 Building Code Levels (CLs) and related construction and design features

Code Level	Design seismic acceleration [g]	Floor system	Transversal confinement	Design approach	Capacity Design
Low	0 - 0.05	one-way rigid deck	No transversal beams or flat beams	Allowable stress-based	NO
Medium	0.05 - 0.15	one-way or two-ways deck	Flat or formed beams	Allowable stress-based and Performance-based	NO
High	0.05 - 0.35	one-way or two-ways deck	Formed beams	Performance-based	YES

Several regular standard buildings are designed through an automated process (Figure 1). Uniform values of storey height, span length, number of spans, and uniform vertical distributed load are assumed. Furthermore, the transversal span length is assumed to be 0.50 *m* less than the longitudinal one.

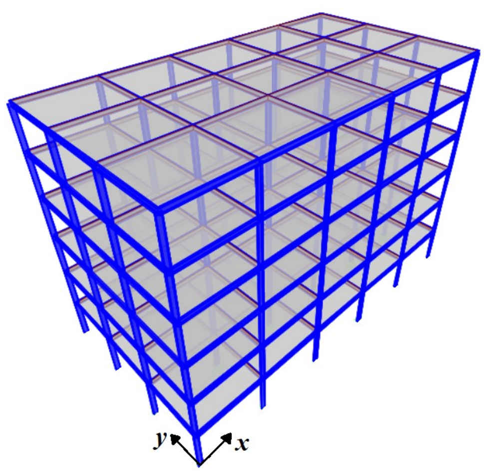


Figure 1 3D RC-BMRF configuration of the regular standard building

The horizontal actions are determined based on the design prescriptions associated with the given period of construction for different randomly generated seismic zones. Five and three spans are assumed in longitudinal and transversal directions, respectively (Figure 1). The structural members are designed to be uniform at the same story level. Furthermore, the columns are tapered in elevation. To cope with the different periods of construction practices, a large variety of structural details are analyzed in the design process. The automated design procedure deals with one-way or two-way floor systems while accounting for the transversal confinement (Table 1). In the case of a one-way floor system, the longitudinal frames (x-direction) carry the load; therefore, the transversal frames (y-direction) are identified by a shorter span. Furthermore, various transversal and longitudinal beam geometry are analyzed to take into account the effects of the beam-column rotation capacity.

The mechanical parameters such as the characteristic compression strength of concrete (f_{ck}), and the elastic modulus of concrete (E_c) are set to 25 *MPa* and 31 *GPa*, respectively. A characteristic steel tensile strength of 450 *GPa* is assumed for reinforced bars, while a reinforcement ratio varying between 0.5 and 2 % is adopted.

Building irregularities

The standard regular building is modified via an automatic process to introduce vertical and plan irregularities. Three irregularities are accounted, that are (*i*) weaker ground level, (*ii*) setbacks vertical irregularities, and (*iii*) setback plan irregularities. The three considered irregularity-based configurations and the related geometrical parameters are shown in Figure 2.

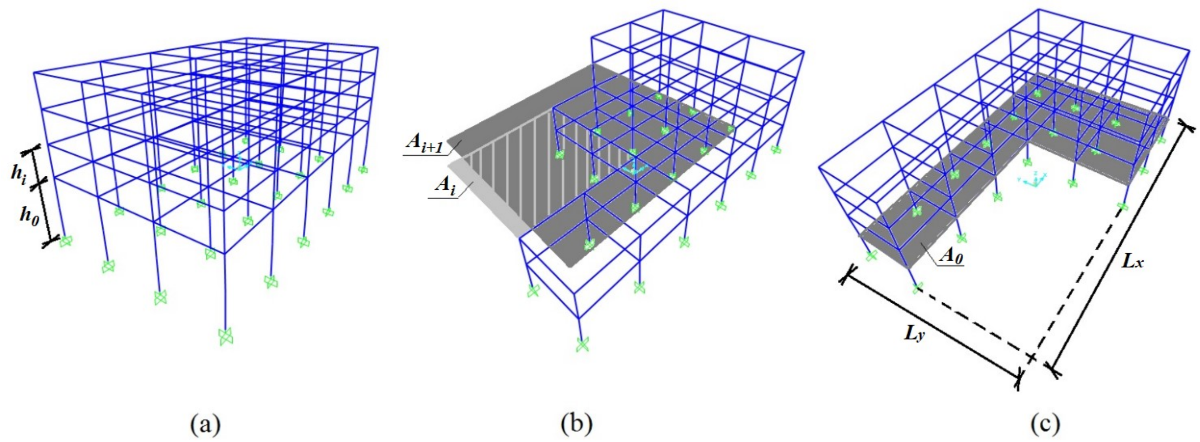


Figure 2 (a) Weaker ground level configuration, (b) vertical setback irregularity configuration, and (c) horizontal setback irregularity configuration

The first configuration relates to the building with higher columns at the first story level. This configuration is representative of some urban residential buildings, where the ground floor is used as a commercial area. The h_r ratio between the ground floor story height (h_0) and the remaining ones (h_i) is used to quantify this geometrical irregularity (Equation (5)).

$$h_r = \frac{h_0}{h_i} \tag{5}$$

The second configuration simulates a wide set of likely vertical irregularities. In this study, the vertical irregularities are intended as the gradual variation of setbacks along the building height in both horizontal directions. A modified version of the parameter proposed by Karavasilis, Bazeos, Beskos [25] (φ_V) is herein adopted for quantifying the setback irregularity (Equation (6)).

$$\varphi_V = \sqrt{\frac{1}{n_s - 1} \cdot \sum_{i=1}^{n_s - 1} \frac{A_i}{A_{i+1}}} \tag{6}$$

where n_s is the number of storeys, while A_i represents the area of the i^{th} floor (Figure 2.b).

The last configuration refers to the plan setback irregularity of the building. The plan configuration at the ground level exhibits a geometrical irregularity obtained by removing a certain number of area elements (in gray in Figure 2.c). The measure of this plan irregularity is given by the parameter φ_H as given by Equation (7).

$$\varphi_H = \sqrt{\frac{L_x \cdot L_y}{A_0}} \tag{7}$$

where L_x and Ly are respectively the maximum longitudinal and transversal length, while A_0 represents the plan area of the building at the ground level.

Building dataset and clustering

In this study, a large number of inherent building parameters are investigated that are: (*i*) number of storeys (n_s); (*ii*) longitudinal span length (l_l); (*iii*) uniform distributed vertical load (q); (*iv*) columns aspect ratio (A_R); (*v*) column-beam moment of inertia ratio (β); (*vi*) deck type; (*vii*) and the three aforementioned irregularity indexes. Table 2 lists the variability range and the step associated with each selected inherent building parameter.

Table 2 Inherent building parameters and relative range values and step variability

Parameter	Range Variability	Step
n_s [-]	1 - 10	1
l_l [m]	3.5 - 6.5	1
q [kN/m]	10 - 50	15
A_R [-]	1 - 2	0.5
β [-]	0.25 - 2	0.6
deck type [-]	one-way - two-ways	-
h_r [-]	1 - 2	0.5
ϕ_V [-]	1 - 3	rand
ϕ_H [-]	1 - 1.8	rand

A MATLAB [26] code is developed to iteratively perform modal analysis for the generated building configurations while automatically modify the selected parameters within the ranges listed in Table 2. The minimum design requirements are verified at each iteration. In case the verification is not satisfied, the current iteration is stopped while providing the null result. A total number of 13998 iterative analyses are performed. The entire computational process is accelerated by implementing multiprocessing analysis.

The outcomes are then processed by using a clustering analysis that aims at categorizing each building configuration in one of the three CLs groups, which are Low Code (LC), Medium Code (MC), and High Code (HC). Each building represents an instance, while the observed characteristics listed in Table 1 are the attributes of the problem. The measure of the similarity between the i^{th} and j^{th} building (d_{ij}) is inversely proportional to the distance parameter d_{ij} given by Equation (8).

$$d_{ij} = \frac{1}{n_a}\sum_{a=1}^{n_a} m_a \rightarrow \begin{cases} m_a = 1 \quad if: b_{a,i} \equiv b_{a,j} \\ m_a = 0 \quad if: b_{a,i} \neq b_{a,j} \end{cases} \tag{8}$$

where $b_{a,i}$ and $b_{a,j}$ represents the a^{th} attribute of the i^{th} and j^{th} building, while n_a is the number of considered attributes. If the two attributes match each other, the m_a score is equal to 1, or 0 otherwise.
Under this condition, the (n_a x n_a) distance matrix is identified, where each component represents the relative dissimilarity measure between two instances. An iterative supervised density-based clustering is implemented to

identify a given number of clusters n_k. In this study, three clusters are used for being representative of the building CLs. The iterative procedure starts from the data point closer to the centroid of the observations. The distance r increases progressively and the frequency of each data point f_d is evaluated. At each step, the data density ρ_d is evaluated as given by Equation (9).

$$\rho_d = \frac{f_d}{r} \tag{9}$$

The density function $\rho_d(r)$ is assessed within the entire data point domain, where r is the distance between the centroid of the observations within a cluster and a generic observation. This function allows to identify if some points are connected, or in other words, if they belong to the same cluster. To accomplish this goal, the maximum density variability λ is set at each iteration. When $\rho_d(r)$ is lower than λ, the next points may be considered unconnected to the previous ones, and the related distance r represents the boundary between two consecutive clusters. At the end of the iteration, a certain number of clusters $n_{k,I}$ is identified, where a given cluster k_l consists of the data points that satisfy the condition below (Equation (10)).

$$if: r_{l-1} \le d_{ci} < r_l \rightarrow b_i \in k_l \tag{10}$$

where d_{ci} is the distance between the i^{th} data-point (b_i) and that one closer to the centroid c. In the case of $l=1$, the distance r_{l-1} is represented by the null value. If the number of clusters generated in the I^{th} iteration is different from the set value, the λ coefficient is adjusted in the next iteration (Equation (11)).

$$\begin{cases} if: n_{k,I} > n_k \rightarrow \lambda_{I+1} = \lambda_I + \Delta\lambda \\ if: n_{k,I} < n_k \rightarrow \lambda_{I+1} = \lambda_I - \Delta\lambda \end{cases} \tag{11}$$

where $\Delta\lambda$ can be approximately fixed to $0.10 \cdot \lambda_I$. The iterative procedure ends when the number of generated clusters is equal to the fixed value.

ROBUST EPR PROCEDURE

The huge amount of the generated output data need for techniques capable of extracting useful information. In this study, the EPR approach is used to search among the possible space of polynomial models that provide an accurate estimate of the observed quantity. EPR allows also pseudo-polynomial structures by including specific user-selected functions such as logarithmic, exponential, etc. The pseudo-polynomial models are capable of exploring a larger space of formulae and then increase their accuracy. In the polynomial and/or pseudo-polynomial symbolic formulae an exponent is assigned at each input variable. These exponents are selected from a set of candidate values and combined through Genetic Algorithms (GAs). The polynomial terms are then multiplied to define the so-called transformed variable. At this stage, multiple regression is performed to find the best fit between the observations and the transformed input variables (Eq (12)).

$$Y = a_0 + \sum_{j=1}^{m} a_j \cdot X_1^{ES(j,1)} \cdot \quad \cdots \quad X_k^{ES(j,k)} \cdot f\left(X_1^{ES(j,1)} \cdot \quad \cdots \quad X_k^{ES(j,k)}\right) \tag{12}$$

where m is the number of model coefficient ($a_0, \ldots, a_m$) to be estimated through multiple regression; k is the number of input variables ($X_1, \ldots, X_k$), while $ES(j,z)$ (with $z = 1, \ldots, k$) is the exponent of the z^{th} input within the j^{th} term. The function $f(X_1^{ES(j,1)} \ldots X_k^{ES(j,k)})$ refers to the additional pseudo-polynomial term, and Y is the estimated quantity The last step of the EPR procedure consists of checking if the mathematical model fits the observations through a specific objective function. The optimal estimates will be selected and the process ends.

GA technique

GAs were introduced by Holland [27] as models that use selection and recombination operators to generate new sample points in a search space. GA begins with a population of chromosomes randomly assigned that represent the candidate solutions to the problem. Initialization consists of multiple assignments of the input variable's exponents creating a generation of candidate functional forms (individuals). In each generation, the "goodness" of every individual is evaluated through a fitness function that gives a measure of how close a given individual is to the target solution. Then, the best-fitted individuals are selected to breed a new generation. The key idea is to simulate the mixing of genetic material that can occur when organisms reproduce. The reproduction of the parent's individuals is performed through a combination of genetic operators called crossover and mutation. After the reproduction phase, the new generation is replaced with the previous one. Therefore, the newly created set of individuals will represent the next parent generation and the aforementioned steps are repeated. The algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population.

Multivariate robust technique

Observations may have large residuals (outliers) that do not match the general trend of the rest of the data. In other cases, some observations may have extreme values of the independent variable that are named as leverage points. In these cases, a robust regression approach is required to estimate the parameters of the regression model by weighting the observations (Equation (13)).

$$\sum_{i=1}^{n} w_i \cdot \left[Y_i - \left(X^{ES}\right)_i \cdot a \right] \cdot \left(X^{ES}\right)_i = 0 \tag{13}$$

where Y_i refers to the i^{th} observed quantity, while w_i is the weight associated with the i^{th} observation and it is inversely proportional to the standard error of the observation. At each iteration, a set of weights is selected and the Least Square (LS) is performed to estimate the regression parameters. The accuracy of the results is measured and compared with the selected requirement. The process will end when the estimates converge to the maximum achievable accuracy. The bi-square [28] robust method is herein adopted. This robust approach aims at minimizing an objective function that is based on the residuals.

Optimal solution

The accuracy of each symbolic expression is measured by a Single Objective (SO) function which is based on some criteria such as maximization of the model accuracy and minimization of the model complexity [29]. A SO-based approach is adopted in this study, where the global accuracy of the symbolic expression is determined through the Mean Square Error (MSE). The best-fitted models are then selected and used for recombination through the crossover and mutation genetic operators. The optimal model is obtained when the highest value of accuracy is reached. Figure 3 resumes the computational workflow describing the SO robust EPR–based procedure adopted in this study.

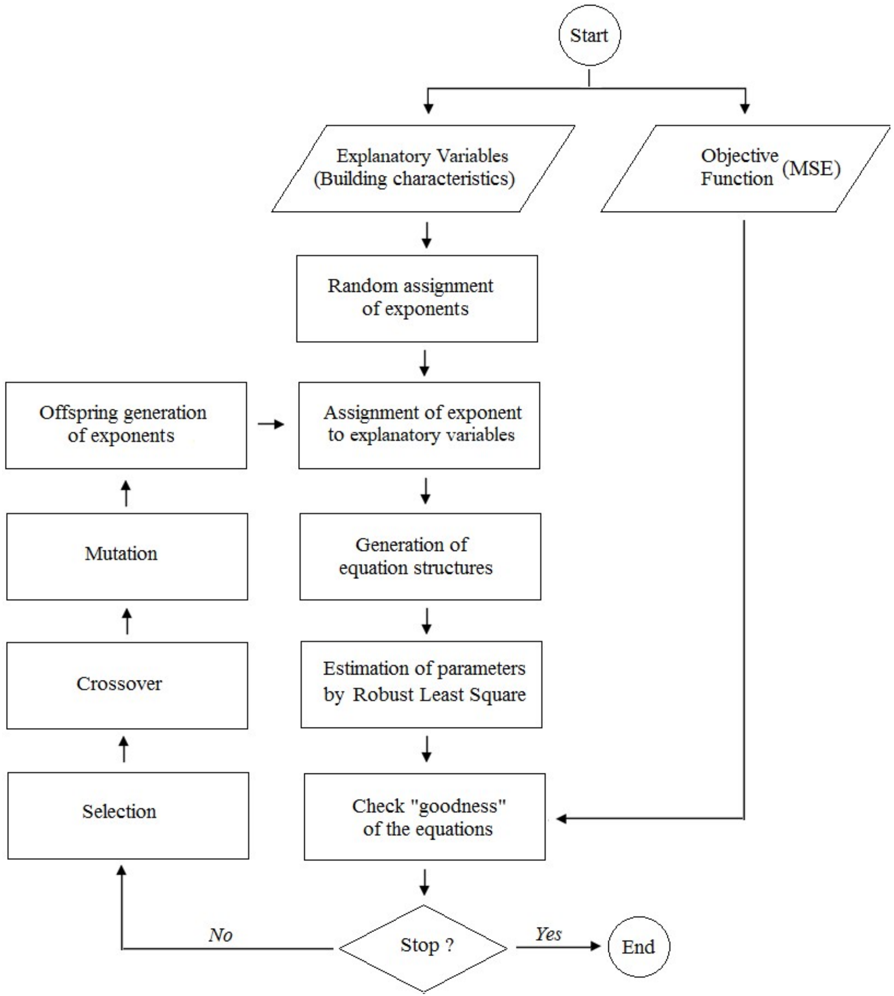


Figure 3 Workflow of the adopted SO EPR–based procedure

Analysis settings

The selected inherent building parameters represent the input variables, while the output consists of the first longitudinal and transversal periods of vibration. Due to the large number of input variables, a reduced set of explanatory variables may be considered to minimize the complexity of the mathematical formulation. To accomplish this goal, the height of the building (H), the maximum longitudinal (L_l) and transversal (L_t) length of the building, and the three irregularity indexes (h_r, φ_V, and φ_H) are considered as explanatory variables. The EPR procedure has been carried out for the three selected CLs based on the genetic operators' values listed in Table 3Table 3.

Table 3 Setting of the genetic operators

Population size (P) [-]	Selection Rate (SR) [%]	Crossover Rate (CR) [%]	Mutation Rate (MR) [%]
1000	30	40	10

The two-parameters regression model has been used, while the bi-square robust LS method has been adopted. The variability of the exponents' values has been set accordingly to the previous studies (i.e. Crowley, Pinho [24], Goel, Chopra [4], Verderame, Iervolino, Manfredi [6], Asteris, Repapis, Foskolos, Fotos, Tsaris [11]) and coherently to the seismic standards. Under these conditions, the upper and lower exponent bounds of 1 and -2 have been set, respectively. The absolute minimum allowable exponent has been fixed to a value lower than 1 to cope with a wider search space for the irregularity indexes. Furthermore, a step variability of 0.05 has been assumed, while a maximum number of iterations of 1000 has been fixed.

RESULTS AND DISCUSSION

As the first attempt to estimate the longitudinal and transversal fundamental periods of 3D RC-BMRFs, only the regular building configurations are considered in the EPR-based analysis. A total number of 3453 regular buildings have been analyzed. The first stage of the proposed approach consists of evaluating the three-building CLs based on the attributes listed in Table 1. The iterative density-based clustering approach is used by fixing the initial maximum density variability λ to 100 data points/unitary distance. The iterative process has been ended with a value of λ equal to 254 data points/unitary distance while the related clusters are shown in Figure 4 for the longitudinal period of vibration.

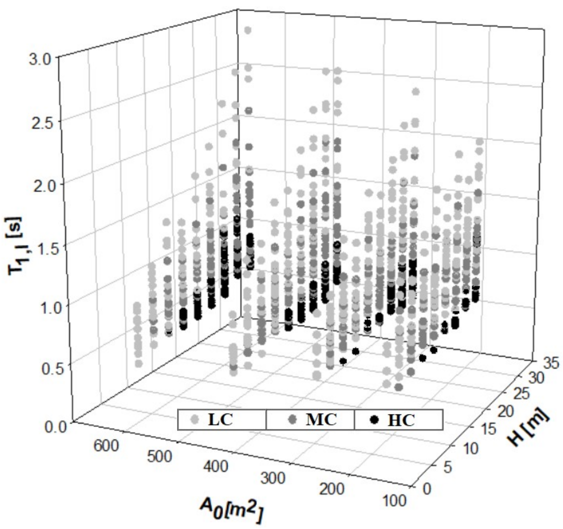


Figure 4 The three selected clusters

It is worth noticing how the HC cluster is located on the bottom of Figure 4, where lower values of periods are identified. On the contrary, the LC clusters are more scattered and assume higher values of the fundamental longitudinal period. The proposed clustering analysis allows for satisfactory classify the building population based on their design class level. Among the 3453 buildings, 730 belongs to the LC, 1590 to the MC, and 1133 to the HC category. The entire dataset has been divided into training (70%) and testing (30%) dataset. Both datasets have been chosen to be representative of the three assumed CLs.

The EPR-based procedure is implemented to estimate the fundamental longitudinal and transversal period of each building class. As the first attempt, only the regular configurations of buildings have been used in the EPR-procedure. A constraint has been imposed in the regression model to obtain consistent results for values of explanatory variables tending to zero, while the additional pseudo-polynomial terms have been fixed to 1. For this purpose, the first regression parameter has been always set to zero. The EPR procedure is then applied to find the optimal mathematical models that provide an accurate estimate of the longitudinal ($T_{1,l}$) and transversal ($T_{1,t}$) fundamental period in seconds, for LC buildings (Equation (14)),

$$
\begin{aligned}
T_{1,l} &= 0.5120 \cdot \ H^{0.55} \cdot \ \frac{L_t^{0.90}}{L_l^{0.95}} \\
T_{1,t} &= 0.3372 \cdot \ H^{0.60} \cdot \ \frac{L_t^{0.85}}{L_l^{0.75}}
\end{aligned}
\tag{14}
$$

MC buildings (Equation (15)),

$$
\begin{aligned}
T_{1,l} &= 0.6810 \cdot \ H^{0.50} \cdot \ \frac{L_t}{L_l^{1.20}} \\
T_{1,t} &= 0.0952 \cdot \ H^{0.50} \cdot \ \frac{L_l^{0.65}}{L_t^{0.40}}
\end{aligned} \tag{15}
$$

and HC building (Equation (16)).

$$
\begin{aligned}
T_{1,l} &= 0.3952 \cdot \ H^{0.40} \cdot \ \frac{L_t^{0.40}}{L_l^{0.55}} \\
T_{1,t} &= 0.1151 \cdot \ H^{0.50} \cdot \ \frac{L_l^{0.85}}{L_t^{0.85}}
\end{aligned} \tag{16}
$$

The period is always directly proportional to the building height. Furthermore, the exponents associated with the building height of the transversal period are equal or greater than those related to the longitudinal period. It is also worth noticing that the fundamental period in a given direction is inversely proportional to the building length in the same direction and inversely proportional to the orthogonal one. This finding is not respected only for the transversal fundamental period of LC buildings. Figure 5 illustrates the fundamental periods of regular buildings obtained for values of H varying between 0 and 30 *m*, while L_l and L_t are set to 27.5 and 15.0 *m*, respectively. The longitudinal and transversal fundamental periods of the three CL classes are compared in Figure 5.a and Figure 5.b, respectively.

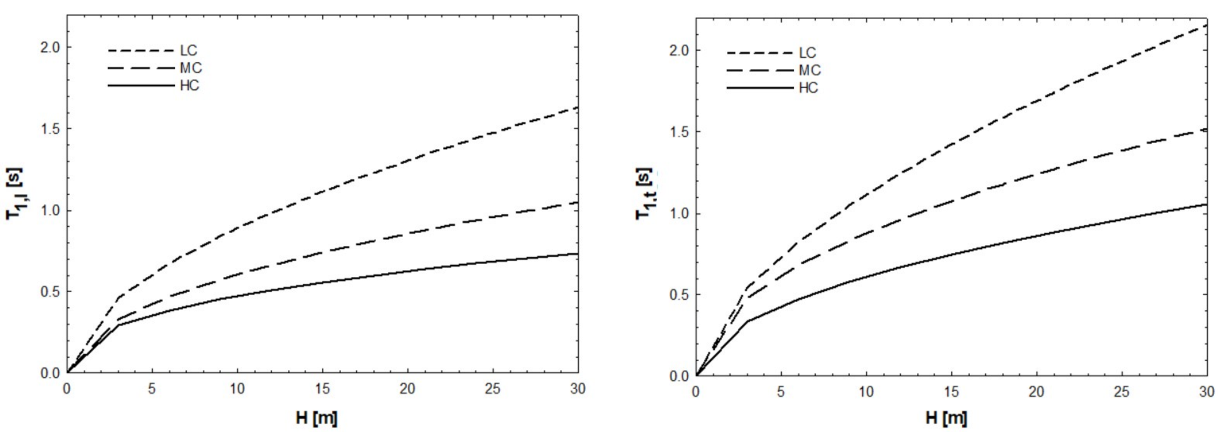


Figure 5 (a) Longitudinal and (b) transversal fundamental periods estimate for the three CLs of regular buildings

It is worth noticing that the transversal fundamental period is greater than the longitudinal one. The differences between the two estimates increase rapidly with the building height.

Indeed, several parameters are necessary to accurately estimate the natural period in both longitudinal and transversal directions (e.g. moment of inertia of the columns, mass distribution, etc.). Unfortunately, such kind of information is not always easy to detect in an existing building, since more detailed knowledge of the building is required. It is then worth mentioning that the proposed study deals with improving current simplified formulations of the natural period which are based on the global geometrical building parameters, such as building height and plan dimensions. Using a limited number of explanatory variables in the model leads to reduce its domain of application. The set of computed formulae provides an accurate estimate of the longitudinal and transversal natural periods for aspect ratio L_x/L_y greater than 1 and lower than 3.

The effects of the three irregularity indexes in the period estimates are computed through a second run of the EPR. A new building population is generated by varying the values of the irregularity indexes according to Table 2. The new building population size is 390, 1417, and 859 for LC, MC, and HC, respectively. The new comprehensive set of data has been divided into training (70%) and testing (30%) dataset. The three irregularity indexes have been assumed as explanatory variables, while the variation of the longitudinal and transversal fundamental periods due to the irregularities have been estimated. It has been noted that the variation of the fundamental period does not considerably change for the three CLs and in the two directions. Therefore, a unique optimal EPR-based mathematical formula has been derived to estimate the longitudinal and transversal period variations (Equation (17)).

$$\Delta T_1 = 0.5978 + 0.4022 \cdot \frac{h_r^{1.30}}{\varphi_V^{2.00} \cdot \varphi_H^{0.65}} \tag{17}$$

Period variation increases with the h_r ratio, while the coefficients φ_V and φ_H cause a reduction of the fundamental period. ΔT_I represents a correction coefficient for being employed to estimate the longitudinal and transversal fundamental period of an irregular RC frame. Therefore, the fundamental period in both longitudinal and transversal direction of irregular MRF-RC buildings is obtained by multiplying ΔT_I to the Eqs. (14),(15), (16) based on the related LC.

Figure 6, Figure 7, and Figure 8 illustrate the comparison of the experimental values with the predicted values computed with the proposed formulation and taking into account the correction term associated with the building irregularities. Both training and testing datasets have been reported, while the MSE values have been provided.

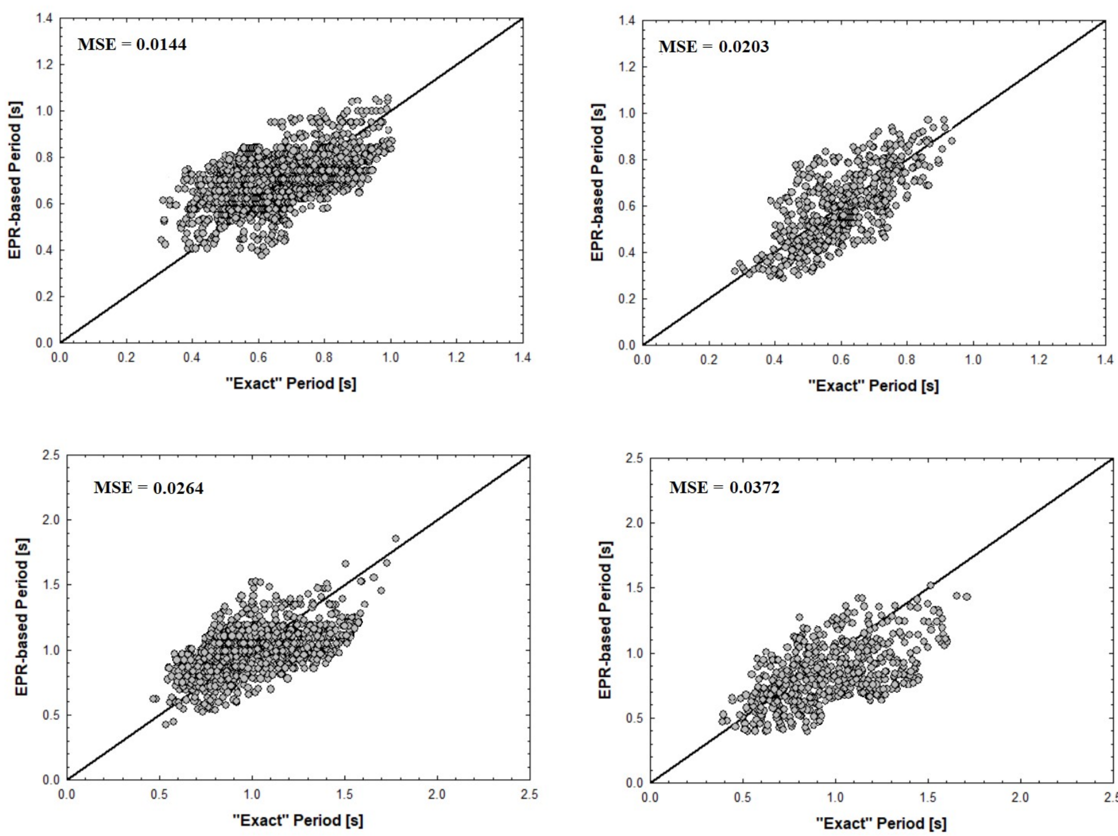


Figure 6 Experimental vs predicted fundamental period for training and testing datasets in (a)-(b) longitudinal and (c)-(d) transversal direction for HC buildings

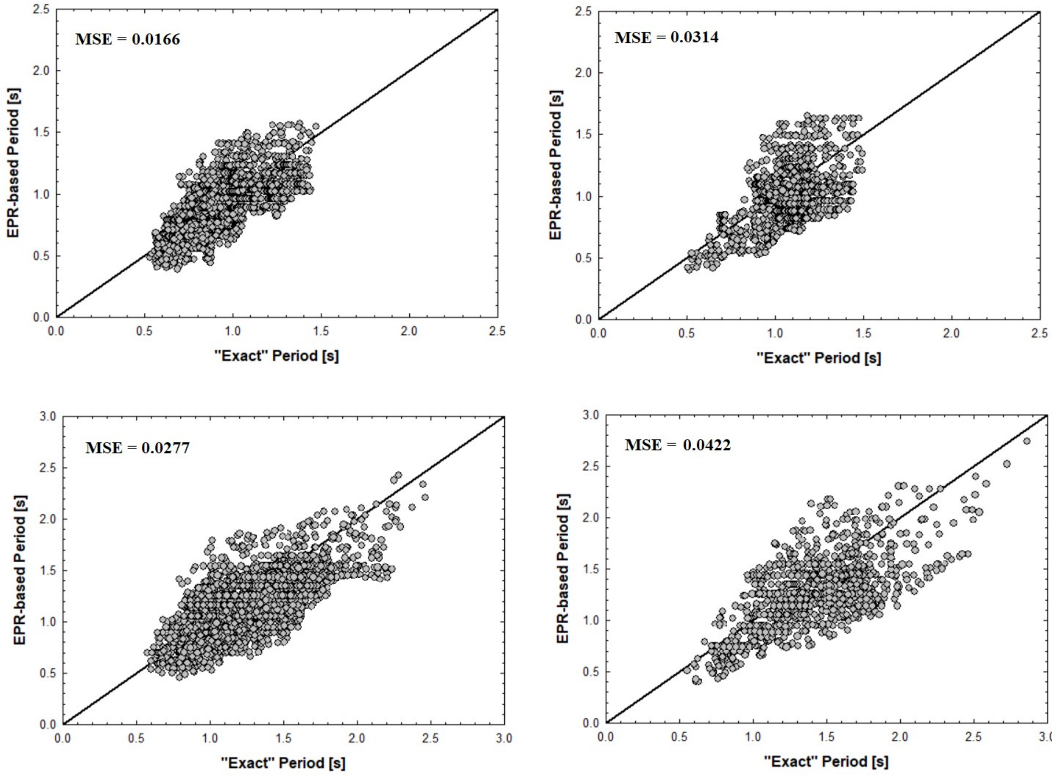


Figure 7 Experimental vs predicted fundamental period for training and testing datasets in (a)-(b) longitudinal and (c)-(d) transversal direction for MC buildings

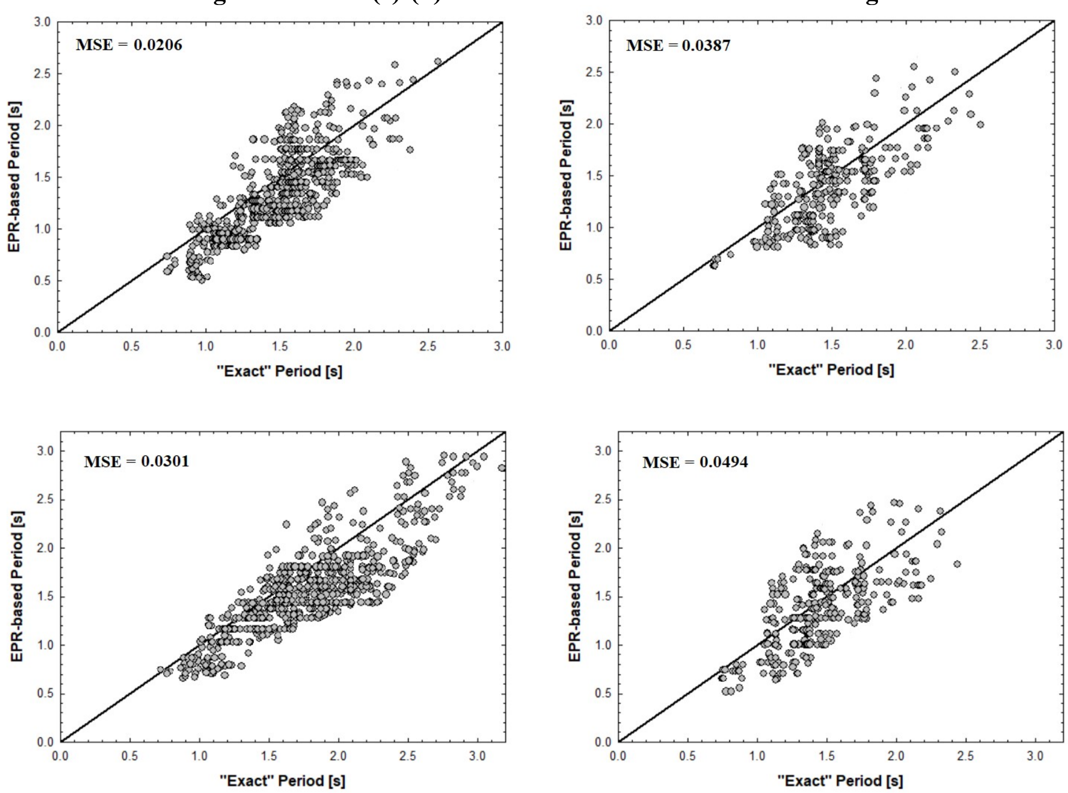


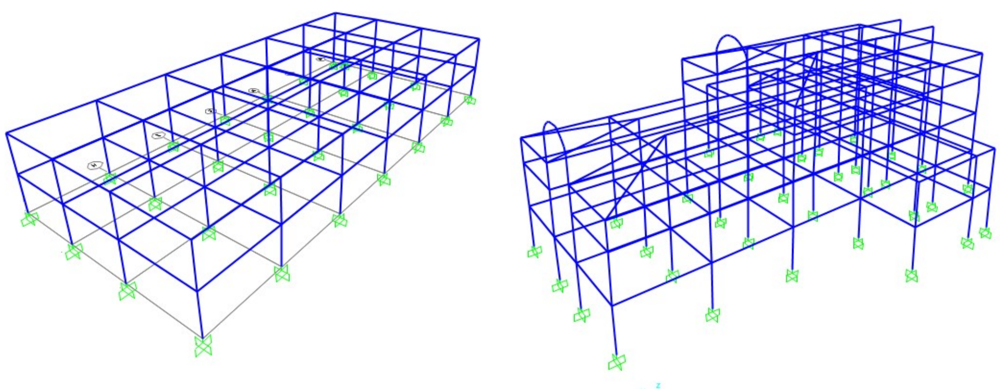
Figure 8 Experimental vs predicted fundamental period for training and testing datasets in (a)-(b) longitudinal and (c)-(d) transversal direction for LC buildings

The results show that the performance of the EPR technique, in terms of MSE values, ranges from 0.0144 to 0.0494 s. The outcomes above have been obtained by using a robust EPR approach which tends to assign low precision to the experimental values that do not follow the general trend of the rest of the data. This in turn leads to removing anomalous data points (outliers and leverage points) from the dataset. It is also worth mentioning that the performance of the proposed technique is strongly affected by the number of selected explanatory variables. Indeed, several parameters are necessary to enhance the accuracy of the predicted natural period in both longitudinal and transversal directions (e.g. moment of inertia of the columns, mass distribution, etc.). Unfortunately, such kind of information is not always easy to detect in an existing building, since more detailed knowledge of the building is required. This paper deals with improving current simplified formulations of the natural period which are based on the global geometrical building parameters only (e.g. building height and plan dimensions).

APPLICATION

Two different case studies have been considered to validate the proposed mathematical formulae. The first case study is a two-story RC building with a regular planar and vertical layout. The second case study is a four-story RC building that is irregular in plan and elevation. Both case study buildings are located in Melzo, Italy (Lat: 45.0512, Long:9.0432).

The first case study building (*B1*) is a school consisting of three separated blocks through two expansion joints of 5 cm each, which makes the structural blocks perfectly independent. The central block is herein investigated and it has a rectangular plan with dimensions of about 42.00x18.80 m and a total height of 6.8 *m*. The columns have an oversized cross-section area in the longitudinal direction while the beams are formed in both directions. Based on these structural characteristics, the building may be classified as an HC structure. The Building Information Model (BIM) of the building has been used to define a Finite Element (FE) model SAP2000 [19] as shown in Figure 9.a.



(a) (b)

Figure 9 (a) FE model of the first case study building *B1*, and (b) second case study building *B2*

The second case study building (*B2*) hosts the city hall which was built in 1993. It consists of two independent blocks; the first one has a regular planar layout, the second block has a "T" shape plan, while both are vertically irregular. In this case study, the "T" shape planar layout building is considered. The maximum horizontal dimensions are 34.95 x 26.70 *m* while the story height is 3.45 *m.* The columns have reduced dimensions while both formed and flat beams are adopted. Therefore, the building can be classified as an MC structure. The building FE model is illustrated in Figure 9.b. Table 4 lists the geometrical characteristics of the two case study buildings used in the proposed EPR-based formulation to estimate the longitudinal and transversal fundamental periods.

Table 4 Geometrical characteristics of the two case study buildings

	H [m]	L_l [m]	L_t [m]	h_r [-]	ϕ_V [-]	ϕ_H [-]
B1	6.80	42.00	18.80	1.00	1.00	1.00
B2	10.35	34.95	26.70	1.00	1.11	1.23

The modal analysis has been performed in SAP2000 [19], while the fundamental longitudinal and transversal periods have been computed. The modal characteristics obtained for the FE models have been compared with the estimates provided by the proposed EPR-based formulation (Table 5).

Table 5. Comparisons of the modal characteristics obtained by the FE model and the EPR-based formulations for the two case study buildings

	B1		*B2*	
	FE	EPR-based	FE	EPR-based
$T_{1,l}$ [s]	0.3755	0.3521	0.6568	0.7261
$T_{1,t}$ [s]	0.4632	0.4926	0.7101	0.7324

The proposed EPR-based formulations provide consistent results. A mean error of about 10.0 % is made in the fundamental period estimation. Furthermore, the proposed formulations are compared with the main mathematical relationships proposed in the literature. (e.g., Goel and Chopra (1997); Hong and Hwang (2000); Verderame, Iervolino, and Manfredi (2009); Crowley, Pinho [24]; Chiauzzi, Masi, Mucciarelli, Cassidy, Kutyn, Traber, Ventura, Yao [30]). Since the B1 case study building satisfies the regularity conditions, additional comparisons with ASCE 7-16[3] and EC8[2] approximated formulation of the fundamental period have been performed. Figure 10 and Figure 11 show the comparisons for the *B1* and *B2* case study buildings, respectively.

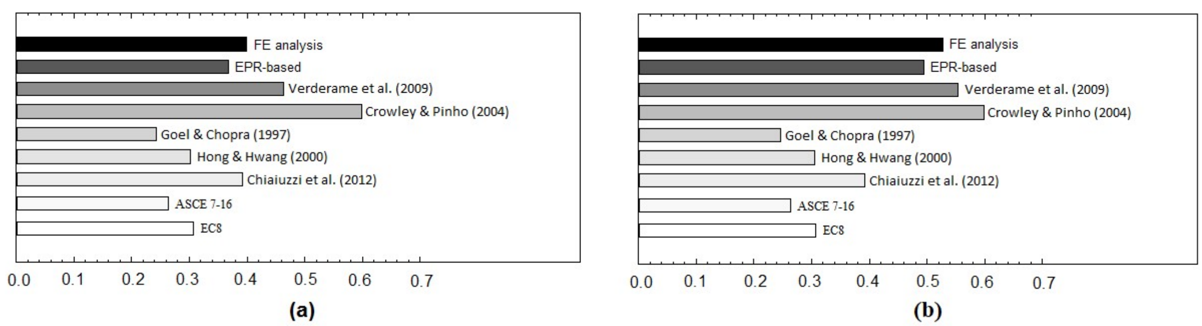


Figure 10 Comparisons of the (a) longitudinal and (b) transversal fundamental periods for the *B1* case study building.

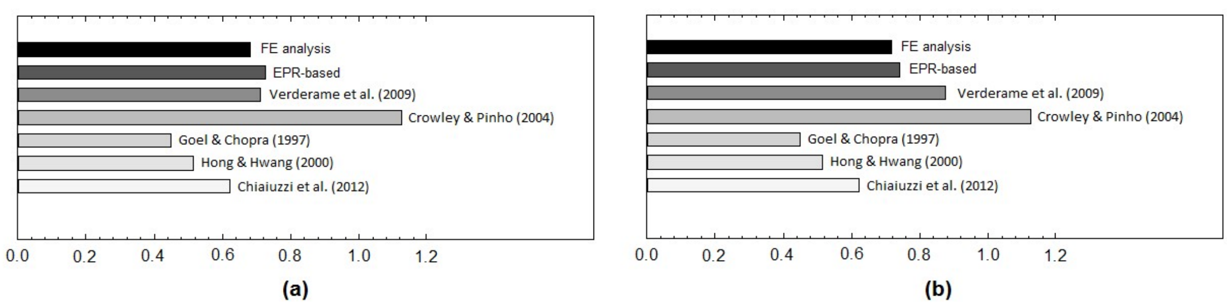
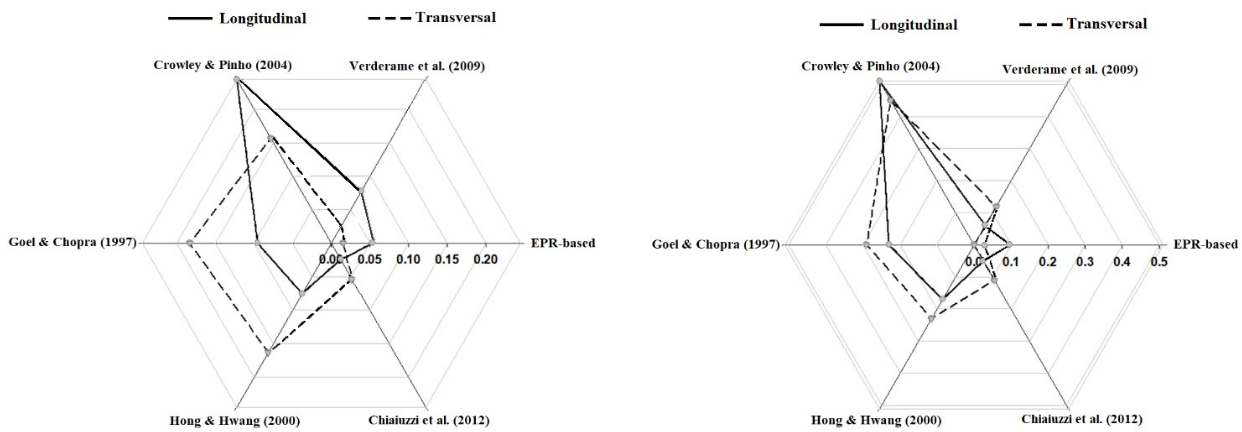


Figure 11 Comparisons of the (a) longitudinal and (b) transversal fundamental periods for the *B2* case study building

The fundamental periods estimated according to ASCE 7-16[3] and EC8[2] provisions tend to underestimate the expected values. The effectiveness of the proposed formulation in estimating the fundamental longitudinal and transversal period of RC-MRF building is demonstrated by the comparisons. Besides the formulation proposed by Verderame, Iervolino, and Manfredi (2009), the other relationships provide a unique period estimate. This condition may result in inconsistent estimates when the dynamic characteristics of the buildings differ in the two principal directions. Goel and Chopra (1997) and Hong and Hwang (2000) formulations tend to underestimate the fundamental period. On the contrary, the results obtained with the formulation proposed by Crowley, Pinho [24] overestimates the FE-based period. Verderame, Iervolino, and Manfredi (2009) classified the buildings based on the maximum design acceleration and provide different formulations. The proposed methodology categorizes the buildings based on certain technical design details (e.g., floor system, beam type, design criteria). This classification system allows taking into consideration a wider number of aspects that affect the period assessment, and then providing consistent estimates as demonstrated for the two case studies. A satisfactory agreement of the proposed formulations with the FE modal results is shown in Figure 11. These results highlight how the geometrical irregularities of the case study building *B2* are consistently taken into account. The absolute errors of the estimates are calculated and compared in Figure 12.



(a) (b)

Figure 12 Comparisons of the estimated absolute errors for the (a) *B1* and (b) *B2* case study building

The absolute error obtained with the proposed approach ranges between 0.014 and 0.074 s. On the other hand, the absolute errors associated with the other formulations ranges in the interval of 0.02-0.50 s.

CONCLUDING REMARKS

Although the fundamental period of vibration of RC-BMRF buildings is a key issue in earthquake design, the main seismic design codes proposed simplified formulations for its estimation such as the one based on the total building height which is feasible only for regular structures. Indeed, the fundamental period depends on several design parameters and geometrical irregularities. A novel strategy to assess the longitudinal and transversal fundamental period of a large variety of RC bare frame buildings has been proposed. The buildings have been classified based on the main design prescriptions and practices. Moreover, a wide range of vertical and horizontal building irregularities have been accounted and their effects on the modal characteristics have been computed.

An automated process is implemented to generate a large set of buildings populations that are designed in agreement with the Italian and European seismic design codes of the last century. A density-based clustering approach has been used to classify the building population based on their main design characteristics and construction details. For each building configuration, the modal analysis has been performed and then a novel robust EPR technique is implemented to find the optimal polynomial expressions of the fundamental period.

New formulations of the longitudinal and transversal fundamental periods of regular RC-BMRF have been obtained for different Code Levels. The periods have been always found to be directly proportional to the building height and length orthogonal to the considered direction. Furthermore, the fundamental period in a given direction is inversely proportional to the building length in the same direction. The longitudinal and transversal fundamental periods tend to decrease for the High Code building class, but they tend to increase for the Low Code building class. Furthermore, the effects of a wide range of vertical and horizontal setback irregularities on the fundamental period have been quantified. It has been found that the period decreases for buildings with vertical and horizontal irregularities defined through the indices φ_V and φ_H, while increases when h_r is greater than 1.

The proposed EPR-based formulations have been compared using two real case study buildings. In both cases, the absolute errors reduce both in the longitudinal and transversal directions. Then, the proposed formulations have been compared with the mathematical relationships available in the literature, showing better results when irregularities are taken into account.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Research Council under the Grant Agreement n° ERC_IDEal reSCUE_637842 of the project IDEAL RESCUE— Integrated DEsign and control of Sustainable CommUnities during Emergencies.

REFERENCES

1. Council AT, California SEAo. *Tentative provisions for the development of seismic regulations for buildings: a cooperative effort with the design professions, building code interests, and the research community.* Vol 510: Department of Commerce, National Bureau of Standards; 1978.
2. Eurocode C. 8: Design of structures for earthquake resistance. *Part.* 2004;1:1998-1991.
3. Institute SE. *Minimum design loads for buildings and other structures.* Vol 7: Amer Society of Civil Engineers; 2006.
4. Goel RK, Chopra AK. Period formulas for moment-resisting frame buildings. *Journal of Structural Engineering.* 1997;123(11):1454-1461.
5. NEHRP B. Recommended Provisions for the Development of Seismic Regulations for New Buildings. *Washington, DC: Building Seismic Safety Council, Federal Emergency Management Agency.* 1994.
6. Verderame GM, Iervolino I, Manfredi G. Elastic Period of Existing RC-MRF Buildings. Paper presented at: Eurocode2009.
7. Hong LL, Hwang WL. Empirical formula for fundamental vibration periods of reinforced concrete buildings in Taiwan. *Earthquake engineering & structural dynamics.* 2000;29(3):327-337.
8. Varadharajan S, Sehgal V, Saini B. Seismic response of multistory reinforced concrete frame with vertical mass and stiffness irregularities. *The structural design of tall and special buildings.* 2014;23(5):362-389.
9. Cimellaro GP, Reinhorn AM, D'Ambrisi A, De_Stefano M. Fragility Analysis and Seismic Record Selection. *Journal of Structural Engineering, ASCE.* 2011;137(3):379-390.
10. Cimellaro GP, Giovine T, Lopez-Garcia D. Bidirectional Pushover analysis of irregular structures. *Journal of Structural Engineering, ASCE.* 2014;140(9):04014059.
11. Asteris PG, Repapis CC, Foskolos F, Fotos A, Tsaris AK. Fundamental period of infilled RC frame structures with vertical irregularity. *Structural Engineering and Mechanics.* 2017;61(5):663-674.
12. Young K, Adeli H. Fundamental period of irregular moment-resisting steel frame structures. *The structural design of tall and special buildings.* 2014;23(15):1141-1157.
13. Wang X, Zenelis I. Estimating participation factors and mode shapes for electromechanical oscillations in ambient conditions. Paper presented at: 2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE)2018.
14. Catbas FN, Brown DL, Aktan AE. Parameter estimation for multiple-input multiple-output modal analysis of large structures. *Journal of engineering mechanics.* 2004;130(8):921-930.
15. Asteris PG, Tsaris AK, Cavaleri L, et al. Prediction of the fundamental period of infilled RC frame structures using artificial neural networks. *Computational Intelligence and Neuroscience.* 2016;2016.
16. Asteris PG, Repapis CC, Tsaris AK, Di Trapani F, Cavaleri L. Parameters affecting the fundamental period of infilled RC frame structures. *Earthquakes and Structures.* 2015;9(5):999-1028.
17. Asteris PG, Repapis CC, Cavaleri L, Sarhosis V, Athanasopoulou A. On the fundamental period of infilled RC frame buildings. *Structural Engineering and Mechanics.* 2015;54(6):1175-1200.
18. Kose MM. Parameters affecting the fundamental period of RC buildings with infill walls. *Engineering Structures.* 2009;31(1):93-102.
19. CSI SV. 8. Integrated finite element analysis and design of structures basic analysis reference manual. *Computers and Structures Inc, Berkeley (CA, USA).* 2002;9.
20. Young K, Adeli H. Fundamental period of irregular concentrically braced steel frame structures. *The Structural Design of Tall and Special Buildings.* 2014;23(16):1211-1224.
21. Charalampakis AE, Tsiatas GC, Kotsiantis SB. Machine learning and nonlinear models for the estimation of fundamental period of vibration of masonry infilled RC frame structures. *Engineering Structures.* 2020;216:110765.

22. Joshi SG, Londhe SN, Kwatra N. Determination of natural periods of vibration using genetic programming. *Earthquakes and Structures.* 2014;6(2):201-216.
23. Fox J, Weisberg S. Robust regression. *An R and S-Plus companion to applied regression.* 2002:91.
24. Crowley H, Pinho R. Simplified equations for estimating the period of vibration of existing buildings. Paper presented at: First European conference on earthquake engineering and seismology2006.
25. Karavasilis TL, Bazeos N, Beskos D. Seismic response of plane steel MRF with setbacks: estimation of inelastic deformation demands. *Journal of Constructional Steel Research.* 2008;64(6):644-654.
26. MathWorks I. MATLAB and statistics toolbox release 2012b. *Natick (Massachusetts, United States).* 2012.
27. Holland JH. Genetic algorithms. *Scientific american.* 1992;267(1):66-73.
28. Gross AM. Confidence intervals for bisquare regression estimates. *Journal of the American Statistical Association.* 1977;72(358):341-354.
29. Fiore A, Berardi L, Marano GC. Predicting torsional strength of RC beams by using evolutionary polynomial regression. *Advances in Engineering Software.* 2012;47(1):178-187.
30. Chiauzzi L, Masi A, Mucciarelli M, et al. Estimate of fundamental period of reinforced concrete buildings: code provisions vs. experimental measures in Victoria and Vancouver (BC, Canada). Paper presented at: Proceedings of the 15th world conference on earthquake engineering, paper reference2012.