Available online at www.eccomasproceedia.org
Eccomas Proceedia COMPDYN (2017) 1748-1763

ECCOMAS
Proceedia

COMPDYN 2017
6th ECCOMAS Thematic Conference on
Computational Methods in Structural Dynamics and Earthquake Engineering
M. Papadrakakis, M. Fragiadakis (eds.)
Rhodes Island, Greece, 15–17 June 2017

FINDING CORRELATIONS BETWEEN ENGINEERING DEMAND PARAMETERS AND INTENSITY MEASURES THROUGH EVOLUTIONARY POLYNOMIAL REGRESSION

Alessandra Fiore[1], Fabrizio Mollaioli[2], Giuseppe Quaranta[3], and Giuseppe C. Marano[4,1]

[1] Department of Science of Civil Engineering and Architecture, Technical University of Bari
Via Orabona 4, 70125 Bari, Italy
{alessandra.fiore, giuseppecarlo.marano}@poliba.it

[2] Department of Structural and Geotechnical Engineering, Sapienza University of Rome
Via Gramsci 53, 00197 Rome, Italy
fabrizio.mollaioli@uniroma1.it

[3] Department of Structural and Geotechnical Engineering, Sapienza University of Rome
Via Eudossiana 18, 00184 Rome, Italy
giuseppe.quaranta@uniroma1.it

[4] College of Civil Engineering, Fuzhou University
Xue Yuan Road, Fuzhou 350108, China
marano@fzu.edu.cn

Keywords: Engineering demand parameter, Evolutionary Polynomial Regression, Intensity measure, Reinforced concrete building, Seismic assessment.

Abstract. *A critical issue in performance-based seismic assessment of structures and infrastructures is the development of reliable formulations able to correlate engineering demand parameters (EDPs) with given earthquake intensity measures (IMs). This task involves the following steps: i) selection of target cases-study and elaboration of the corresponding structural models, ii) preparation of the database collecting seismic records, iii) identification of candidate EDPs and IMs, iv) nonlinear dynamic analyses, v) numerical calibration of functional models able to correlate EDPs and IMs, vi) evaluation of the predictive capability of the developed models. Within this framework, the present paper exploits an advanced nonlinear regression method – named Evolutionary Polynomial Regression technique – in order to obtain several non-dominated models (according to the Pareto's dominance criterion) that predict EDPs as function of assigned IMs for fixed-base and base-isolated multi-storey reinforced concrete buildings subjected to ordinary and pulse-like ground motion.*

Peer-review under responsibility of the organizing committee of COMPDYN 2017.
doi: 10.7712/120117.5526.16732

1 INTRODUCTION

Many studies have been carried out in the past years to find suitable correlations between intensity measures (IMs) and engineering demand parameters (EDPs), and large efforts have been also spent to improve their predictive capability. Because of the large number of hazard curves available for the peak ground acceleration (PGA) and the spectral acceleration at the fundamental period of the structure ($Sa(T_1)$), these two parameters have been widely adopted as candidate IMs. Unfortunately, several works have shown that $Sa(T_1)$ might not be a good predictor since it does not account for the lengthening of the period as the structure goes well into the inelastic range, and it does not consider the influence of the higher modes. Consequently, some scalar measures have been proposed in order to take into account such aspects [1][2]. Further IMs (including spectrum IMs) have been investigated in recent years, and they have exhibited better predictive capabilities than PGA and $Sa(T_1)$, especially in case of medium-period frame structures [3]. Vector intensity measures have also been proposed [4] by merging other parameters to $Sa(T_1)$ in order to enhance the correlations with respect to the selected EDPs. Moreover, modified IMs have been suggested for structural systems equipped with seismic protection devices, such as base isolation system [5] or hysteretic dampers [6]. For base-isolated buildings, for instance, it has been demonstrated that the predictive capability of the IMs significantly depends on the isolation period of the structure when pulse-like ground motions are considered. For such dynamic loading condition, velocity-related IMs exhibit a strong correlation with the response of the isolation system.

Within this framework, the main goal of the present study is to propose new reliable formulations able to correlate EDPs with given IMs. Two archetypes of multi-storey reinforced concrete (RC) buildings (i.e., a fixed-base building and a base-isolated building) are studied by varying their characteristics. The seismic response is calculated through non-linear dynamic analyses, considering two sets of ground motions (GMs) consisting of ordinary and pulse-like records. A large number of IMs selected amongst those commonly used in the literature are taken into account. Functional models able to correlate the considered IMs and the selected EDPs are calibrated by means of an advanced nonlinear regression method, namely the Evolutionary Polynomial Regression technique.

2 DEFINITION OF STRCTURAL MODELS AND SEISMIC INPUT

The selected case-study is a six-storey, three-bay framed RC building with a fixed base or retrofitted with a base-isolation system. The superstructure (which is designed according to the previous Italian code DM 1996) is representative of existing buildings located in a high seismic zone (i.e., "zone 1" according to the seismic hazard classification within DM 1996). Essential geometrical data about this case-study are reported in Figure 1a. The periods of the first three modes of vibration of the fixed-base structure are equal to 1.17 s, 0.4 s and 0.24 s, respectively (they are obtained by using the reduced cracked stiffness of the structural elements, which is half the initial elastic one). Base-isolation systems having different properties are considered. Figure 1b reports the constitutive law used for representing the cyclic response of the isolation system, which is characterized by no-stiffness degradation under cyclic loading. Different behaviors of the base-isolation system can be obtained by varying the three parameters that define the bi-linear backbone curve of the constitutive law (namely, the elastic strength F_d, the elastic limit displacement D_y and the post-elastic stiffness K_d) [7]. In this regards, Avşar and Özdenmir [8] have demonstrated recently that the correlation between the IMs and the response of the isolation system is not significantly sensitive to the F_d value. Therefore, isolation systems that differ each other only for D_y and K_d values are considered in this study. Specifically, the elastic strength F_d is set equal to 0.03 the seismic weight W of the

structure. The elastic limit displacement D_y is assumed equal to 0 mm (it is representative of isolation systems consisting of friction pendulum isolators) whereas four values of K_d are considered. They are calculated in such a way to obtain isolation periods T (defined as $T = 2\pi (W/K_d g)^{1/2}$) equal to 3.0, 3.5, 4.0 and 4.5 sec.

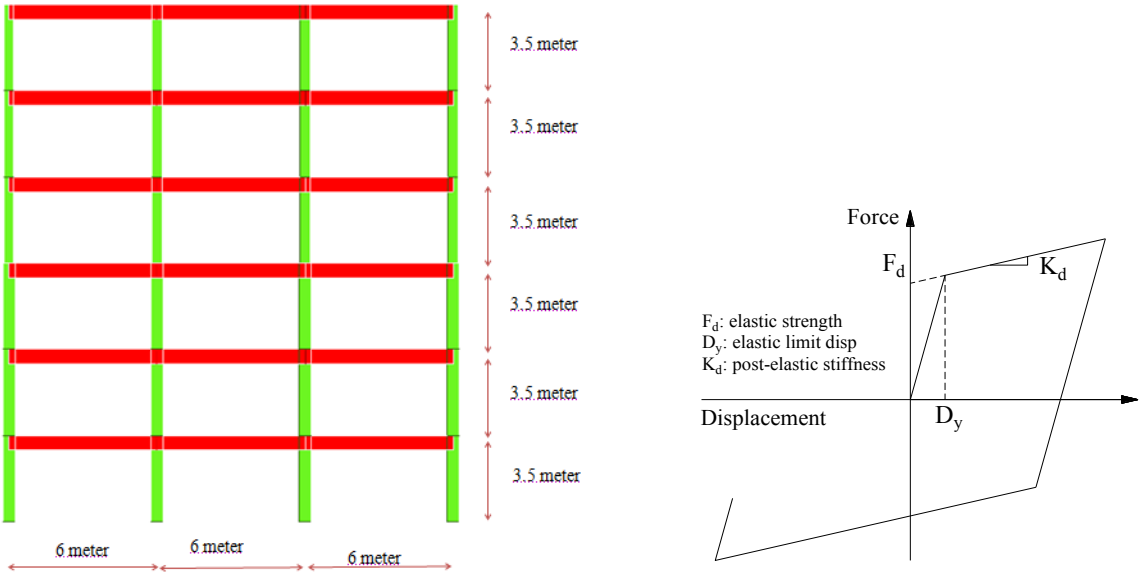

Figure 1: The framed structure considered in the present study (left) and the bi-linear hysteretic law used to model the lateral constitutive behavior of the isolation system (right).

A total of 139 seismic records is selected from the Next Generation of Attenuation Project Database of the PEER dating back to 2005. These GMs are used as input dynamic loading for non-linear dynamic analyses. In order to highlight the effects of pulse-type GMs [9]-[11], this database is divided into two groups, namely ordinary GMs (80 records, closest distance ranging from 0.34 km to 87.87 km, magnitude from 5.74 to 7.9) and pulse-like near-fault GMs (59 records, closest distance ranging from 0.07 km to 20.82 km, magnitude from 5 to 7.62). All the selected time-histories are recorded on soil classified as type C or D, according to the NEHRP site classification based on the preferred V_{s30} values. The horizontal component used in the analyses is selected as follows:

- the component having larger spectral acceleration at the fundamental period of the considered superstructure for ordinary GMs,
- the fault-normal rotated component for pulse-like near-fault GMs.

The structural response is evaluated via non-linear dynamic analyses performed by means of OpenSees 2.2.2. The models of the RC frames have been prepared using Beam with fibre-hinges Elements whereas Elastomeric Bearing Elements have been adopted for the isolation system. The masses are concentrated at the nodes representing the beam-column joints, and the stiffness of the floors is modeled with rigid diaphragm constraints. The interested reader can found more details about the structural modeling in Ref. [5].

3 INTENSITY MEASURES AND ENGINEERING DEMAND PARAMETERS

3.1 Engineering demand parameters

The EDPs considered in this study are the following:

- Maximum Inter-story Drift Ratio (MIDR), i.e. the maximum value of the peak inter-story drift ratio (drift normalized by the story height) over all the stories;
- Maximum Floor Acceleration (MFA), i.e. the maximum value of the peak floor absolute acceleration over all stories of the superstructure.

The MIDR has been widely used as EDP for the assessment of structural damages. It has been shown that it is closely related to local damages, instability phenomena, and story collapses. On the other hand, MFA reflects the level of non-structural damages.

3.2 Intensity measures

The IMs considered in this study are classified into two groups, i.e. non-structure-specific IMs calculated directly from ground motion time-histories (Table 1) and structure-specific IMs obtained from response spectra of ground motion time-histories depending on the period of the structure (Table 2).

IM type		Notation	Name	Note
Non-structure-specific IMs	Acceleration-related	PGA	Peak ground acceleration	$PGA = \max \lvert \ddot{u}_g(t) \rvert$ $\ddot{u}_g(t) =$ acceleration time history
		AI	Arias Intensity [12]	$AI = \frac{\pi}{2g} \int_0^{t_f} \ddot{u}_g^2(t)\,dt$ $t_f =$ total duration
		CAV	Cumulative absolute velocity [13]	$CAV = \int_0^{t_f} \lvert \ddot{u}_g(t) \rvert dt$
		I_a	Compound acceleration-related IM [14]	$I_a = PGA \cdot t_d^{1/3}; t_d = t_2 - t_1;$ $t_1 = t(5\%AI); t_2 = t(95\%AI)$
	Velocity-related	PGV	Peak ground velocity	$PGV = \max \lvert \dot{u}_g(t) \rvert$ $\dot{u}_g(t) =$ velocity time history
		I_v	Compound velocity-related IM [14]	$I_v = PGV^{2/3} \cdot t_d^{1/3}$
		CAD	Cumulative absolute displacement [15]	$CAD = \int_0^{t_f} \lvert \dot{u}_g(t) \rvert dt$
		IV	Incremental velocity [16]	
		SED	Specific energy density	$SED = \int_0^{t_f} [\dot{u}_g(t)]^2 dt$
	Displacement – related	PGD	Peak ground displacement	$PGD = \max \lvert u_g(t) \rvert$
		I_d	Compound displacement-related IM [14]	$I_d = PGD \cdot t_d^{1/3}$
		ID	Incremental displacement [16]	

Table 1: Non structure-specific IMs considered in this study.

Non-structure-specific IMs are further divided into three categories: acceleration-related, velocity-related and displacement-related IMs. On the other hand, structure-specific IMs are divided into two groups, i.e. IMs obtained from the response spectral ordinate at certain periods and IMs carried out from the integration of response spectra over a defined period range.

IM type		Notation	Name	Note
Structure-specific IMs	Spectral	S_a	Spectral acceleration at isolation period	$S_{pa}(T)$ S_{pa} = 5% damp. pseudo acc. spectrum T = isolation period
		E_{Ir}	Relative input energy at isolation period [17]	$E_{Ir} = -\int_0^{t_f} \ddot{u}_g(t)\dot{u}_r(t)dt$ $\ddot{u}_g(t)$ = ground acceleration time history $\dot{u}_r(t)$ = relative velocity time history of a (ξ = 5%, T) SDOF
		E_{Ia}	Absolute input energy at isolation period [17]	$E_{Ia} = \int_0^{t_f} \ddot{u}_t(t)du_g = \int_0^{t_f} \ddot{u}_t(t)\dot{u}_g(t)dt$ $\dot{u}_g(t)$ = ground velocity time history $\ddot{u}_t(t)$ = absolute acceleration time history of a (ξ = 5%, T) SDOF
	Integral	ASI	Acceleration spectrum intensity	$ASI = \int_{0.1}^{0.5} S_{pa}\, dT$
		VSI	Velocity spectrum intensity	$VSI = \int_{0.1}^{2.5} S_v\, dT$ S_v = 5% damp. velocity spectrum
		I_H	Housner intensity [20]	$I_H = \int_{0.1}^{2.5} S_{pv}\, dT$ S_{pv} = 5% damp. pseudo velocity
		$V_{EIr}SI$	Relative input equivalent velocity spectrum intensity	$V_{EIr}SI = \int_{0.1}^{3.0} \sqrt{2E_{Ir}}\, dT$
		$V_{EIa}SI$	Absolute input equivalent velocity spectrum intensity	$V_{EIa}SI = \int_{0.1}^{3.0} \sqrt{2E_{Ia}}\, dT$
		MASI	Modified ASI	$MASI = \int_{0.5T}^{1.25T} S_{pa} dT$
		MVSI	Modified VSI	$MVSI = \int_{0.5T}^{1.25T} S_v\, dT$
		MI_H	Modified I_H	$MHI = \int_{0.5T}^{1.25T} S_{pv}\, dT$
		$MV_{EIr}SI$	Modified $V_{EIr}SI$	$MV_{EIr}SI = \int_{0.5T}^{1.25T} V_{EIr}\, dT$
		$MV_{EIa}SI$	Modified $V_{EIa}SI$	$MV_{EIa}SI = \int_{0.5T}^{1.25T} V_{EIa}\, dT$

Table 2: Structure-specific IMs considered in this study.

The interested reader can found further details about IMs elsewhere, e.g. Refs. [18][19].

4 EVOLUTIONARY POLYNOMIAL REGRESSION TECHNIQUE

4.1 Development of data-driven models using soft computing techniques

Soft computing techniques have proven especially suitable for developing data-driven non-linear models. For instance, Genetic Programming (GP) [21] was originally proposed in order to identify interpretable symbolic expressions. It employs an evolutionary algorithm to search among a population of possible combinations of mathematical operators, candidate arguments (variables) and parameters. It was observed, however, that the original GP paradigm produces

complicated expressions, and then a rule-based GP that incorporates the numerical estimation of the constants was proposed [22]. Although this technique is flexible in some applications, it suffers from the explosion of the number of parameters when complex and nested GP formulations are encountered. The Evolutionary Polynomial Regression (EPR) paradigm was introduced in order to provide a substantially different search strategy for building mathematical expressions of models [23]. This technique has been successfully applied in many engineering fields, e.g. hydrology, hydraulics, hydro-climatology, geology, geotechnical engineering. Recently, the EPR technique has been exploited in the fields of structural engineering and material science, see for instance Refs. [24]-[29].

4.2 The basis of the Evolutionary Polynomial Regression technique

The EPR technique overcomes the drawbacks of Davidson's rule based GP by employing a hybrid data-driven technique that combines the effectiveness of Genetic Algorithms (GA) with numerical regression. From a numerical standpoint, EPR can be defined as a non-linear global stepwise regression technique. The models developed by means of the EPR are non-linear because the relationships involving the basic variables can produce non-linear functions, although they are linear with respect to regression parameters. Moreover, EPR is a global regression technique since the search for the optimal model structure is based on the exploration of the whole (discrete) space of models, by leveraging a special coding of the candidate mathematical expressions. A general model adopted in EPR is [23]:

$$\hat{\mathbf{Y}} = a_0 + \sum_{j=1}^{m} a_j \cdot \left(\mathbf{X}_1\right)^{\mathbf{ES}(j,1)} \cdot \ldots \cdot \left(\mathbf{X}_k\right)^{\mathbf{ES}(j,k)} \cdot f\left(\left(\mathbf{X}_1\right)^{\mathbf{ES}(j,k+1)} \cdot \ldots \cdot \left(\mathbf{X}_k\right)^{\mathbf{ES}(j,2k)}\right) \tag{1}$$

where m is the maximum number of additive terms, [$\mathbf{X}_1 \ldots \mathbf{X}_k$] are the model inputs (i.e., candidate explanatory variables) and $\hat{\mathbf{Y}}$ is the model output. The function f and the exponents $\mathbf{ES}(j,z)$ are selected from a set of alternatives defined by the user. The adjustable parameters a_j are evaluated by means of the linear least squares (LS) method based on the minimization of the sum of squared errors (SSE). The SSE reads:

$$SSE = \sum_{i=1}^{N} \left(y_i - \hat{y}_i\right)^2 \Big/ N \tag{2}$$

where y_i and $\hat{y}_i$ are target value and model prediction, respectively. The search for the optimal f and $\mathbf{ES}(j,z)$ is performed by means of a standard GA over a database defined by the user. The GA operates by simulating the Darwinian evolution: it begins with a random population of initial solutions and – through crossover and mutation operations – it creates a new generation of candidate solutions. In doing so, fit individuals are selected for mating whereas weak individuals die off. The mated parents create a child (offspring) with a chromosome set which is a mix of parents' chromosomes. An integer coding with single point crossover is used [23]. It is worth noting that, if the user-defined set of exponents contains zero and $\mathbf{ES}(j,z)=0$, then the relevant model input disappears from the final expression. Hence, functional structures such as Eq. (1) are able to reproduce a large variety of models. The EPR stops when a termination criterion is fulfilled. It can be either the maximum number of generations, the maximum number of terms in the target mathematical expression or a particular user-selected tolerance. The earliest version of the EPR entailed a single-objective optimization strategy where the unique search criterion was the Coefficient of Determination (COD):

$$\mathrm{COD} = 1 - \sum_{i=1}^{N} \left(\hat{y}_i - y_i\right)^2 \Big/ \sum_{i=1}^{N} \left(y_i - \bar{y}\right)^2 \tag{3}$$

and it measures the proportion of the total variation of **Y** that is explained by the model predictions $\hat{\mathbf{Y}}$. The denominator allows to evaluate how the model performs in comparison with the average of the recorded outputs $\bar{y}$, which is the simplest description of the available data. A perfect fit to data would lead to COD = 1 (or 100%), while COD values close to zero would indicate that model performances are comparable with $\bar{y}$. Negative COD could be obtained, which means that $\bar{y}$ provides a better description of data than the model predictions $\hat{\mathbf{Y}}$.

4.3 Multi-objective EPR modeling

The EPR technique is able to search amongst a large number of candidate models. As usual in data-driven modeling, the best model is the one that best fit with the observed data. However, it is now widely accepted that the best model is also the simplest that fits the goal of the application. This principle of parsimony – often called Occam's razor – states that among a set of otherwise equivalent models of a given phenomenon, one should choose the simplest one to explain a dataset. From a numerical standpoint, an unnecessary complexity can deteriorate the ability of a regression model in predicting unconsidered data. Hence, there is the need to look for a suitable trade-off between model complexity and accuracy. In other words, a systematic way to avoid the problem of over-fitting is also required for a robust data-driven modeling. In this perspective, a Multi-Objective Genetic Algorithm (MOGA) strategy based on the Pareto dominance criterion [30][31] was implemented in EPR [32]. Here, the complexity is defined as inclusion of new terms or combinations of inputs. EPR-MOGA explores the space of *m*-term formulations using two or three from the following objectives: *a*) maximization of model accuracy evaluated in terms of COD, *b*) minimization of the number of model coefficients (i.e., number of additive terms) and/or *c*) minimization of the number of used model inputs (i.e., the inputs which exponent is not zero in the resulting model structure). Note that the last two objectives are two measures of model parsimony. The MOGA used for the evolutionary stage is OPTIMOGA [33]. At the end of the search, the EPR-MOGA produces a set of non-dominated optimal solutions (i.e., the Pareto front) that can be considered as trade-off between complexity and accuracy.

5 RESULTS

5.1 Search settings

The base model structure reported in Eq. (1) is considered. The function f can be an exponential or a logarithmic function. The candidate exponents belong to the set {-3, -2, -1, -0.5, -0.33, 0, 0.5, 0.33, 1, 2, 3}. The maximum number of additive terms in the final expressions is m=5, so as to facilitate the comparison with existing formulations. Length and time units for the IMs are [cm] and [s]. The MIDR is dimensionless whereas the unit for MFA is [m/s^2].

5.2 Prediction of MIDR

Figure 2 shows the COD values for predictive models of the MIDR obtained by using a single IM. Numerical results for the base-isolated building refer to the case T=3.0 s, but they are representative of the general trends observed in all the considered isolation periods. Since Mollaioli et al. [5] have found that EDP-IM relationships, in general, typically follow a standard power law, the accuracy of the non-linear formulations obtained by means of the EPR technique is compared with the following functional form, referred to as "linear regression"[1]:

[1] The regression of the model in Eq. (4) is linear using a log-scale.

$$EDP = a(IM)^b \tag{4}$$

Figure 2: COD values of predictive models for MIDR using a single IM: *a*) fixed-base building; *b*) base-isolated building.

It can be inferred from Figure 2 that the non-linear regression increases the accuracy up to a percentage equal to 38% for the fixed-based building and up to 96% for the base-isolated building. For the fixed-base building, the MIDR is best predicted in the event of ordinary GM rather than under pulse-like near-fault GM. For the base-isolated building, better predictions can be sometimes obtained under pulse-like near-fault GM. The most efficient IMs for ordinary GMs are PGV, VSI, I_H, MVSI and MI_H. On the other hand, PGV, EI_a, EI_r, $MV_{EIa}SI$, SED, MI_H are the most effective IMs for pulse-like near fault GMs. The PGV is the most efficient non-structure-specific IM. Integral IMs (except for ASI) are usually more efficient than spectral IMs, while the modified IMs usually lead to better predictions than the corresponding original ones (especially in case of pulse-like near-fault GMs). Overall, these results confirm the evidences reported in Ref. [5]. For a fixed-based building, the best models for the MIDR based on a single IM are the following.

- Linear regression, ordinary GMs:

$$MIDR = 0.012588 \cdot MI_H \quad (\text{COD}=84.80\%) \tag{5}$$

- Linear regression, pulse-like GMs:

$$MIDR = 0.0074918 \cdot VSI \quad (\text{COD}=77.89\%) \tag{6}$$

- Non-linear regression, ordinary GMs:

$$MIDR = 0.078147 \cdot \sqrt{MVSI} + 3.3198 \cdot 10^{-5} \cdot MVSI^2 \quad (\text{COD}=87.94\%) \tag{7}$$

- Non-linear regression, pulse-like GMs:

$$MIDR = 7.92 \cdot \ln\left(\frac{1}{I_H^3}\right) + 28.1302 \cdot I_H^{0.33} + 0.058984 \cdot I_H \cdot \ln\left(\frac{1}{I_H}\right) + \\ + 0.00051749 \cdot I_H^2 \cdot \ln\left(\frac{1}{\sqrt{I_H}}\right) + 0.0020749 \cdot I_H^2 \quad (\text{COD}=81.5\%) \tag{8}$$

For a base-isolated building, the best models for the MIDR based on a single IM are listed hereafter.

- Linear regression, ordinary GMs:

$$MIDR = 0.051283 \cdot VSI^{0.33} \, (\text{COD}=67.05\%) \tag{9}$$

- Linear regression, pulse-like GMs:

$$MIDR = 0.016972 \cdot EI_r^{\,0.33} \quad (\text{COD}=59.71\%) \tag{10}$$

- Non-linear regression, ordinary GMs:

$$MIDR = 0.11504 \cdot (1/PGV) + 0.0045699 \cdot PGV \cdot \ln\left(1/PGV^2\right) + \\ + 0.035571 \cdot PGV + 0.00011581 \cdot PGV^2 \quad (\text{COD}=72.34\ \%) \tag{11}$$

- Non-linear regression, pulse-like GMs:

$$MIDR = 365.5927 \cdot \frac{1}{MI_H^2} \cdot \ln\left(\frac{1}{MI_H}\right) + 3.8098 \cdot \frac{1}{MI_H^{0.33}} + \\ + 0.30465 \cdot \ln\left(\frac{1}{MI_H^{0.33}}\right) + 0.00084451 \cdot MI_H \quad (\text{COD}=85.01\%) \tag{12}$$

Formulations of the MIDR as function of multiple IMs have been also calibrated (Figure 3). It can be noted that the accuracy of the predictions increases by considering several IMs simultaneously. The best performances are obtained by formulating the MIDR as function of, both, velocity-related and integral IMs, besides the case in which all IMs are taken into account. A very good compromise between complexity and accuracy is achieved by the formulations listed hereafter.

- Fixed-base building, multiple IMs, ordinary GMs:

$$MIDR = 0.014728 \cdot IV^{0.33} \cdot \sqrt{S_a} + 6.6417 \cdot 10^{-7} \cdot PGA^3 \cdot VSI^2 \cdot \sqrt{AI} \quad (\text{COD}=94.36\%) \tag{13}$$

- Fixed-base building, multiple IMs, pulse-like GMs:

$$MIDR = 4.789 \cdot 10^{-8} \cdot CAD^{0.33} \cdot \sqrt{VSI} \cdot MVSI \cdot MV_{EIr}SI \cdot \ln\left(\frac{\sqrt{V_{EIa}SI}}{S_a}\right) + \\ + 0.00041743 \cdot VSI \cdot \sqrt{S_a} \quad (\text{COD}=88.28\%) \tag{14}$$

- Base-isolated building, multiple IMs, ordinary GMs:

$$MIDR = 2.3688 \cdot 10^{-6} \cdot \sqrt{SED \cdot MASI} \cdot \ln\left(\frac{1}{SED^3}\right) + \\ + 9.2246 \cdot 10^{-6} \cdot \sqrt{I_v} \cdot E_{Ia}{}^{0.33} \cdot I_H + 0.062274 \cdot \ln\left(AI^{0.33} \cdot \sqrt{VSI}\right) \quad \text{(COD=83.26\%)} \quad (15)$$

- Base-isolated building, multiple IMs, pulse-like GMs:

$$MIDR = 1.9566 \cdot 10^{-5} \cdot \frac{PGD^{0.33} S_a{}^2}{\sqrt{CAV}} + 0.010218 \cdot \sqrt{VSI} + 0.0016844 \cdot PGA^{0.33} \text{(COD=92.11\%)} \quad (16)$$

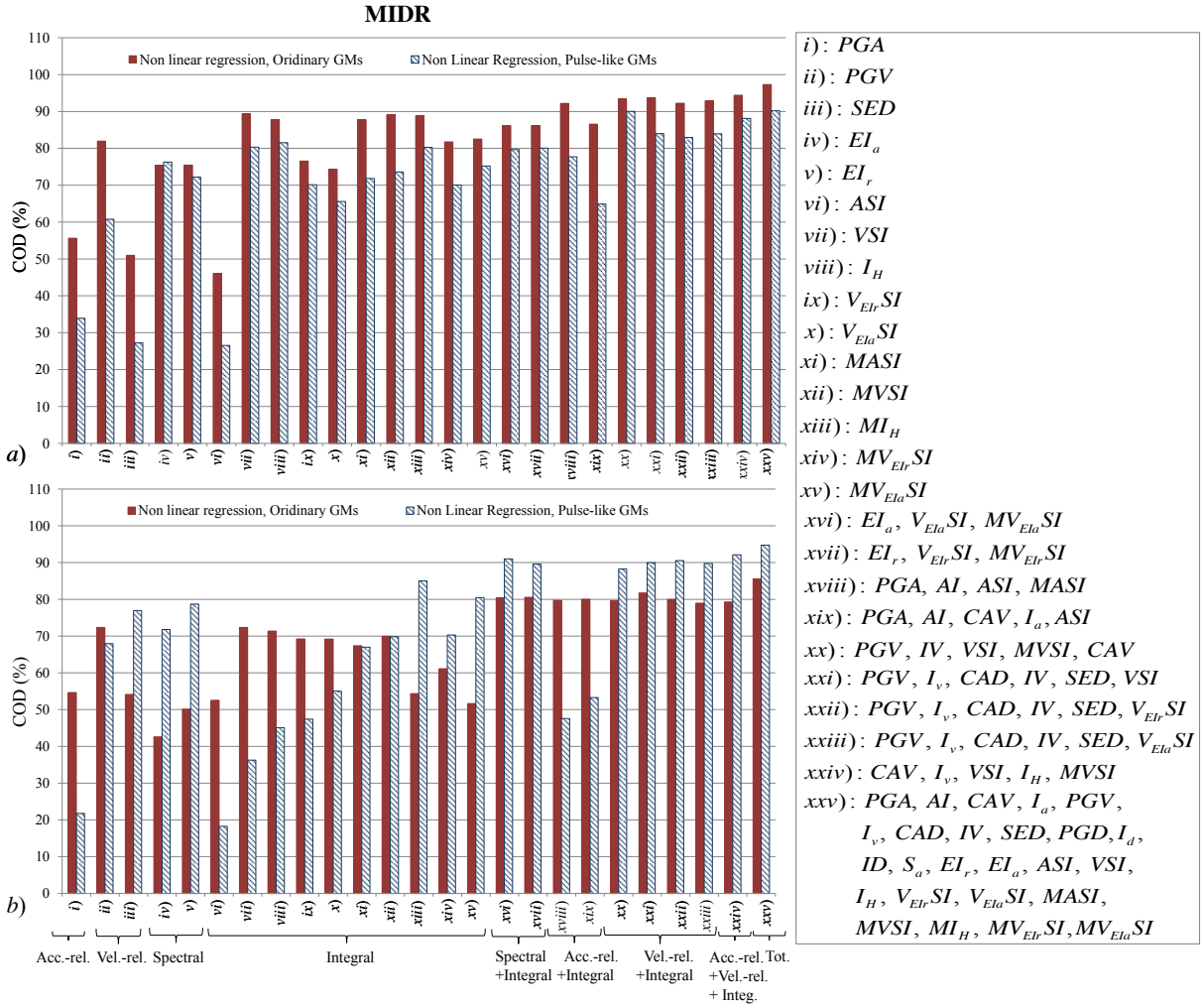

Figure 3: COD values of predictive models for MIDR considering several IMs simultaneously: *a*) fixed-base building; *b*) base-isolated building.

5.3 Prediction of MFA

The COD values for predictive models of the MFA obtained by using a single IM are shown in Figure 4. Once again, the predictions obtained using non-linear regressions are more accurate. The MFA is best estimated under ordinary GM rather than in case of pulse-like near-fault GM. The prediction of the MFA for base-isolated building is especially problematic. It can be deduced that PGA, ASI, MASI, MVSI and MI_H are the most effective IMs for fixed-base buildings whereas PGV, VSI, I_H, $V_{EIr}SI$, $V_{EIa}SI$ and MVSI are more suitable for base-

isolated buildings. Generally, integral IMs provide good estimates of the MFA. Differently from the evidences obtained for the MIDR, the use of modified IMs does not improve in a substantial way the prediction of the MFA.

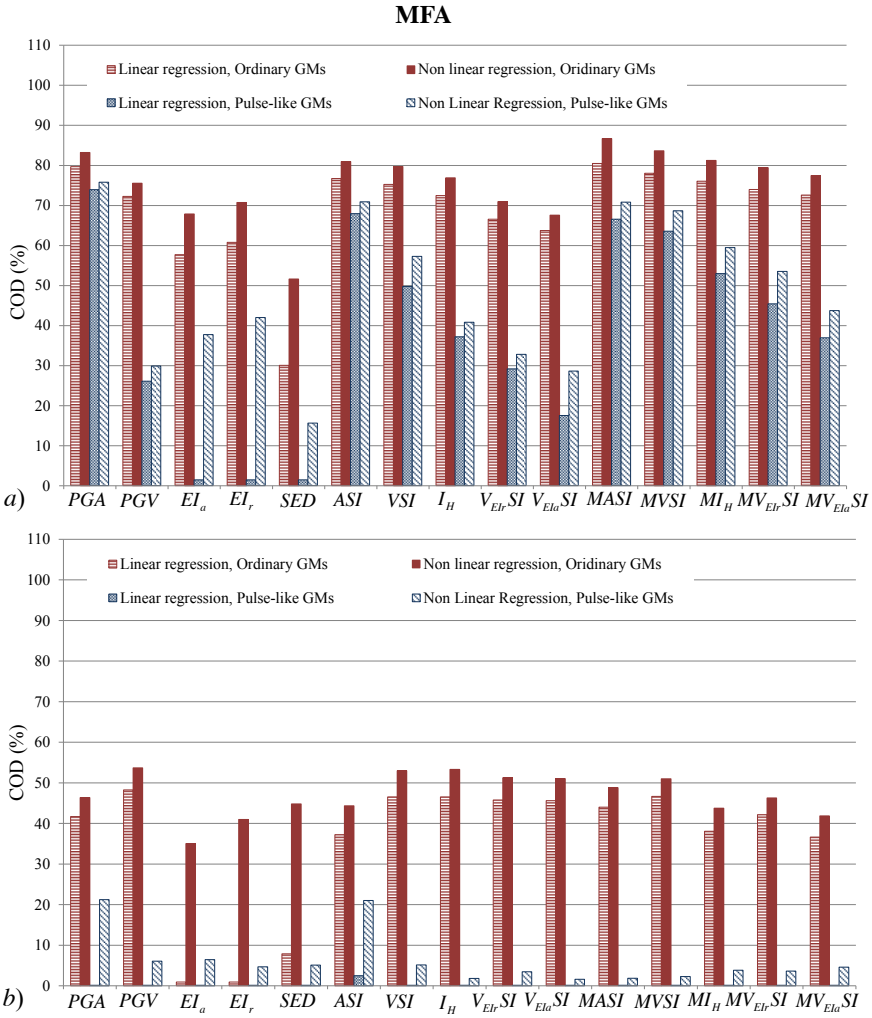

Figure 4: COD values of predictive models for MFA using a single IM: *a*) fixed-base building; *b*) base-isolated building.

For the fixed-based building, the best models for the MFA based on a single IM are the following.

- Linear regression, ordinary GMs:

$$MFA = 0.23549 \cdot \sqrt{MASI} \quad \text{(COD=80.42\%)} \tag{17}$$

- Linear regression, pulse-like GMs:

$$MFA = 10.4592 \sqrt{PGA} \quad \text{(COD=73.90\%)} \tag{18}$$

- Non-linear regression, ordinary GMs:

$$MFA = 52926.0314 \cdot \frac{1}{MASI} + 3654.3634 \cdot \frac{1}{MASI} \cdot \ln\left(\frac{1}{MASI^3}\right) + \\ + 314.816 \cdot \frac{1}{MASI^{0.33}} + 201862.1696 \cdot \frac{1}{MASI^2} \cdot \ln\left(\frac{1}{MASI}\right) \quad \text{(COD=86.70\%)} \tag{19}$$

- Non-linear regression, pulse-like GMs:

$$MFA = 18.8824 \frac{1}{\sqrt{PGA}} \ln\left(\frac{1}{PGA}\right) + 67.6646 \cdot PGA^3 \cdot \ln\left(\frac{1}{PGA^{0.33}}\right) + \\ + 101.5125 \cdot \ln\left(\frac{1}{PGA^{0.33}}\right) + 9.9862 \cdot PGA^3 + \\ + 43.0456 \frac{1}{PGA^{0.33}} \cdot \ln(PGA) \quad \text{(COD=75.74\%)} \tag{20}$$

On the other hand, for the base-isolated building, the best models for the MFA based on a single IM are listed hereafter.

- Linear regression, ordinary GMs:

$$MFA = 0.80912 \cdot \sqrt{PGV} \quad \text{(COD=48.26\%)} \tag{21}$$

- Linear regression, pulse-like GMs:

$$MFA = 0.67404 \cdot ASI^{0.33} \quad \text{(COD=2.52\%)} \tag{22}$$

- Non-linear regression, ordinary GMs:

$$MFA = +2.473 \cdot PGV + 0.0049249 \cdot PGV^2 + 2.6837 \cdot \frac{1}{PGV^{0.33}} + \\ + 4.5925 \cdot \ln\left(\frac{1}{PGV}\right) + 0.28974 \cdot PGV \cdot \ln\left(\frac{1}{PGV^2}\right) \quad \text{(COD=53.68\%)} \tag{23}$$

- Non-linear regression, pulse-like GMs:

$$MFA = 153855.4935 \cdot \frac{1}{\sqrt{PGA}} \cdot \ln\left(\frac{1}{PGA^3}\right) + 1730632.9019 \cdot \frac{1}{\sqrt{PGA}} + \\ + 567122.4354 \cdot \frac{1}{PGA^{0.33}} + 454.7322 \cdot \ln\left(\frac{1}{PGA^{0.33}}\right) + \\ + 8.2524 \cdot PGA^{0.33} \quad \text{(COD=21.25\%)} \tag{24}$$

Including multiple IMs is beneficial to improve the accuracy of the predictions (Figure 5). Apart from the use of all the IMs, the best performances are obtained by expressing the MFA as function of velocity-related IMs together with integral IMs and acceleration-related IMs together with integral IMs. A satisfactory compromise between complexity and accuracy is achieved by the formulations listed hereafter.

- Fixed-base building, multiple IMs, ordinary GMs:

$$MFA = 0.0099258 \cdot ASI + 0.12824 \cdot \sqrt{S_a} \cdot \ln\left(I_H{}^{0.33}\right) + \\ + 0.080633 \cdot PGA^3 \cdot S_a \cdot \ln\left(PGA^{0.33}\right) \quad \text{(COD=92.82\%)} \tag{25}$$

- Fixed-base building, multiple IMs, pulse-like GMs:

$$MFA = 0.44302 \cdot \ln\left(\frac{PGA^2 \cdot AI \cdot I_a \cdot MVSI^3}{CAD^2}\right) + 0.001986 \cdot PGV \cdot \sqrt{IV} + \\ + 3.5837 \cdot 10^{-5} \cdot PGA \cdot V_{Ela}SI^2 \cdot \ln\left(PGA^{0.33}\right) \quad \text{(COD=85.95\%)} \tag{26}$$

- Base-isolated building, multiple IMs, ordinary GMs:

$$MFA = 0.0035592 \cdot \frac{\sqrt{E_{Ir} \cdot I_a} \cdot PGV^3 \cdot MVSI}{CAV \cdot IV^3} + 0.69413 \cdot VSI^{0.33} + 0.001581 \cdot \sqrt{PGA} \quad \text{(COD=59.55\%)} \tag{27}$$

- Base-isolated building, multiple IMs, pulse-like GMs:

$$MFA = 0.0025771 \cdot \sqrt{I_v} \cdot \left(S_a \cdot MASI\right)^{0.33} \cdot \ln\left(\frac{IV \cdot E_{Ia}^{0.33}}{PGA}\right) + 0.0077503 \cdot PGA^{0.33} \cdot \ln\left(VSI\right) + 0.47974 \cdot \frac{CAV^{0.33} \cdot MI_H}{MASI} \cdot \ln\left(AI \cdot I_v\right)^{0.33} \quad \text{(COD=62.79\%)} \tag{28}$$

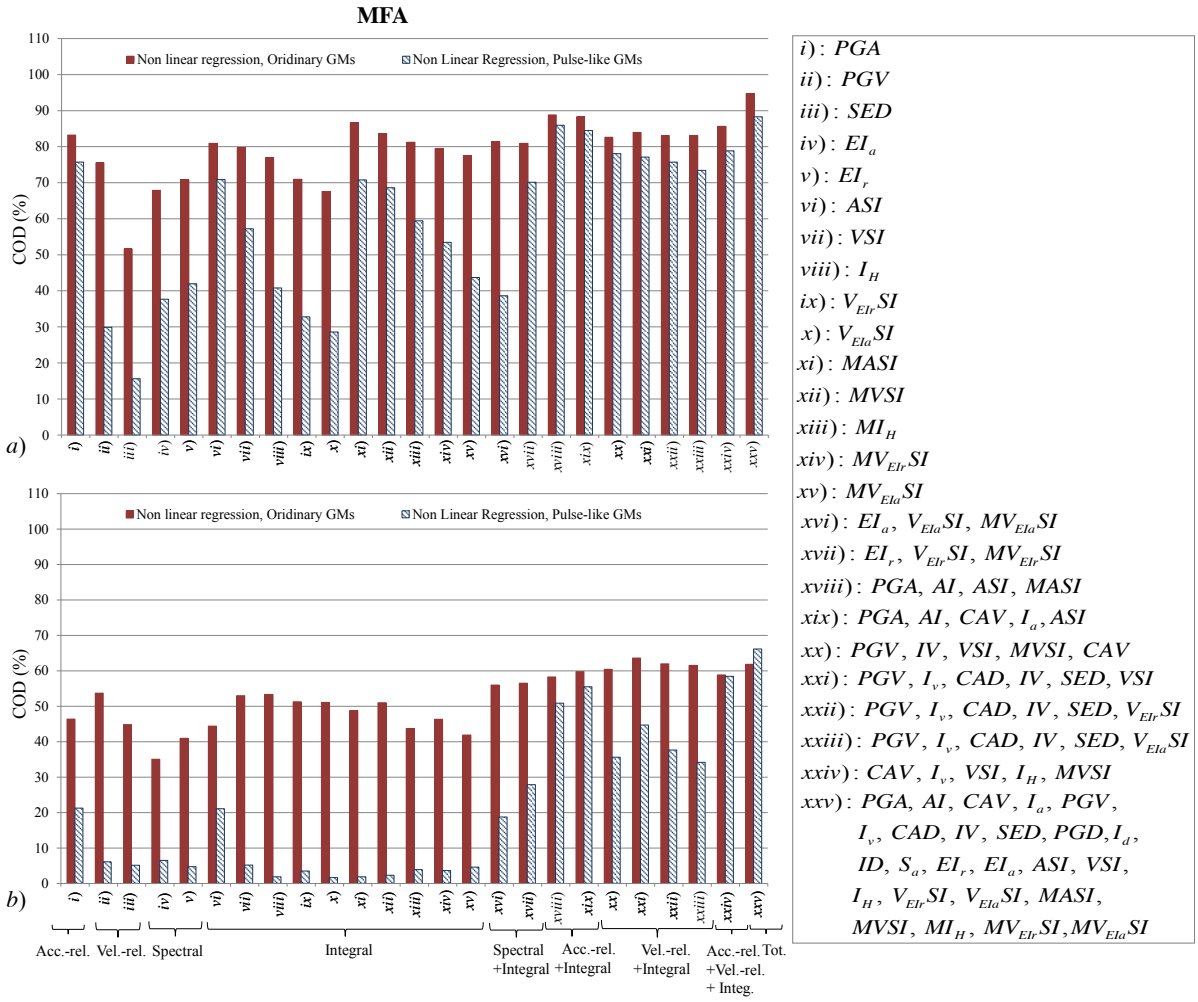

Figure 5: COD values of predictive models for MFA considering several IMs simultaneously: *a*) fixed-base building; *b*) base-isolated building.

6 CONCLUSIONS

The main goal of this work was the calibration of suitable models able to estimate the seismic demand for fixed-based and base-isolated RC buildings under ordinary or pulse-like

near-fault ground motions using most common IMs. Special attention has been paid on the way by which the models are built. To this end, a multi-criteria evolutionary polynomial regression technique has been employed to develop reliable non-linear models for two EDPs, namely MIDR (which is basically related to structural damages) and MFA (which measures potential damages to non-structural components and facilities). The calibration by employing a single IM has demonstrated that a non-linear regression-based approach enhances significantly the predictive capability of the final models. Moreover, the adopted non-linear regression technique was able to identify very accurate, yet compact, models by combining multiple IMs.

7 ACKNOWLEDGMENTS

Fabrizio Mollaioli and Giuseppe Quaranta acknowledge the support from Sapienza University of Rome through the project "Smart solutions for the assessment of structures in seismic areas". Their work is also framed within the research project DPC-ReLUIS 2017.

REFERENCES

[1] P.P. Cordova, G.G. Deirlein, S.S.F. Mehanny, C.A. Cornell, Development of a two-parameter seismic intensity measure and probabilistic assessment procedure. *2nd U.S.-Japan Workshop on Performance-Based Earthquake Engineering of Reinforced Concrete Building Structures*, Sapporo, Hokkaido, Japan, September 11-13, 2000.

[2] M. Bianchini, P.P. Diotallevi, J.W. Baker, Prediction of inelastic structural response using an average of spectral accelerations. *10th International Conference on Structural Safety and Reliability*, Osaka, Japan, September 13-17, 2009.

[3] N. Jayaram, F. Mollaioli, P. Bazzurro, A. De Sortis, S. Bruno, Prediction of structural response of reinforced concrete frames subjected to earthquake ground motions. *9th U.S. National and 10th Canadian Conference on Earthquake Engineering*, Toronto, Canada, July 25-29, 2010.

[4] J.W. Baker, C.A. Cornell, A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon. *Earthquake Engineering and Structural Dynamics*, **34**(10), 1193-1217, 2005.

[5] F. Mollaioli, A. Lucchini, Y. Cheng , G. Monti, Intensity measures for the seismic response prediction of base-isolated buildings. *Bulletin of Earthquake Engineering*, **11**, 1841-1866, 2013.

[6] J. Donaire-Ávila, F. Mollaioli, A. Lucchini, A. Benavent-Climent, Intensity measures for the seismic response prediction of mid-rise buildings with hysteretic dampers. *Engineering Structures*, **102**, 278–295, 2015.

[7] V.A. Matsagar, R.S. Jangid, Influence of isolator characteristics on the response of base-isolated structures. *Engineering Structure*, **26**(12), 1735-1749, 2004.

[8] Ö. Avşar, G. Özdenmir, Response of seismic-isolated bridges in relation with intensity measures of ordinary and pulse-like ground motions. *Journal of Bridge Engineering*, **18**(3), 250-260, 2011.

[9] G. Quaranta, F. Mollaioli, G. Monti, Effectiveness of design procedures for linear TMD installed on inelastic structures under pulse-like ground motion. *Earthquakes and Structures*, **10**(1), 239-260, 2016.

[10] F. Mollaioli, A. Bosi, Wavelet analysis for the characterization of forward-directivity pulse-like ground motions on energy basis. *Meccanica*, **47**, 203-219, 2012.

[11] F. Mollaioli, L. Liberatore, A. Lucchini, Displacement damping modification factors for pulse-like and ordinary records. *Engineering Structures*, **78**, 17-27, 2014.

[12] A. Arias, A measure of earthquake intensity. R.J. Hansen ed. *Seismic design for nuclear power plants*, MIT Press, 438-483, 1970.

[13] EPRI (Electrical Power Research Institute), *A criterion for determining exceedance of the operating basis earthquake* (EPRI NP-5930), EPRI, 1988.

[14] R. Riddell, J. Garcia, Hysteretic energy spectrum and damage control. *Earthquake Engineering and Structural Dynamics,* **30**(12), 1791-1816, 2001.

[15] K. Mackie, B. Stojadinovic, *Seismic demands for performance-based design of bridges*, PEER Report University of California Berkeley 2003/16, 2003.

[16] J.C. Anderson, V.V. Bertero, Uncertainties in establishing design earthquakes. *Journal of Structural Engineering*, **113**(8), 1709-1724, 1987.

[17] C.M. Uang, V.V. Bertero, Evaluation of seismic energy in structures. *Earthquake Engineering and Structural Dynamics*, **19**(1), 77-90, 1990.

[18] H. Ebrahimian, F. Jalayer, A. Lucchini, F. Mollaioli, G. Manfredi, Preliminary ranking of alternative scalar and vector intensity measures of ground shaking. *Bulletin of Earthquake Engineering*, **13**(10), 2805-2840, 2015.

[19] Y. Cheng, A. Lucchini, F. Mollaioli, Proposal of new ground-motion prediction equations for elastic input energy spectra. *Earthquakes and Structures*, **7**(4), 485-510, 2014.

[20] G.W. Housner, Spectrum intensity of strong-motion earthquakes. *Symposium on Earthquakes and Blast Effects on Structures*, Earthquake Engineering Research Institute (EERI) University of California at Los Angeles, Oakland, California, June 1952.

[21] J.R. Koza, *Genetic programming: on the programming of computers by natural selection*, MIT Press, 1992.

[22] J.W. Davidson, D.A. Savic, G.A. Walters, Method for identification of explicit polynomial formulae for the friction in turbulent pipe flow. *Journal of Hydroinformatics*, **1**(2), 115-126, 1999.

[23] O. Giustolisi, D.A. Savic, A Symbolic data-driven technique based on Evolutionary Polynomial Regression. *Journal of Hydroinformatics*, **8**(3), 207-222, 2006.

[24] A. Fiore, L. Berardi, G.C. Marano, Predicting torsional strength of RC beams by using Evolutionary Polynomial Regression. *Advances in Software Engineering*, **47**(1), 178-187, 2012.

[25] A. Fiore, G. Quaranta, G.C. Marano, G. Monti, Evolutionary Polynomial Regression-based statistical determination of the shear capacity equation for reinforced concrete beams without stirrups. *Journal of Computing in Civil Engineering*, **30**(1), 04014111, 2016.

[26] A. Faramarzi, A.A. Javadi, A. Ahangar-Asr, Numerical implementation of EPR-based material models in finite element analysis. *Computers & Structures*, **118**, 100-108, 2013.

[27] A. Faramarzi, A.M. Alani, A.A. Javadi, An EPR-based self-learning approach to material modeling. *Computers & Structures*, **137**, 63-71, 2014.

[28] A.M. Alani, A. Faramarzi, An evolutionary approach to modelling concrete degradation due to sulphuric acid attack. *Applied Soft Computing*, **24**, 985-993, 2014.

[29] A. Fiore, G. Quaranta, G.C. Marano, Evaluation of the plastic hinge length for nonlinear analysis of reinforced concrete buildings. J. Kruis, Y. Tsompanakis, B.H.V. Topping eds. *Computational Techniques for Civil and Structural Engineering*, Saxe-Coburg Publications, Chapter 11, 255-280, 2015.

[30] G.C. Marano, G. Quaranta, R. Greco, Multi-objective optimization by genetic algorithm of structural systems subject to random vibrations. *Structural and Multidisciplinary Optimization*, **39**(4), 385-399, 2009.

[31] G.C. Marano, G. Quaranta, Robust optimum criteria for tuned mass dampers in fuzzy environments. *Applied Soft Computing*, **9**(4), 1232-1243, 2009.

[32] O. Giustolisi, D.A. Savic, Advances in data-driven analyses and modelling using EPR-MOGA. *Journal of Hydroinformatics*, **11**, 225-236, 2009.

[33] D. Laucelli, O. Giustolisi, Scour depth modelling by a multi-objective evolutionary paradigm. *Environmental Modelling & Software*, **26**, 498-509, 2011.