ICOSSAR 2021
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021), June 21-25, 2021, Shanghai, P.R. China
J. Li, Pol D. Spanos, J.B. Chen & Y.B. Peng (Eds)

Efficient loss analysis through surrogate probabilistic seismic demand models

Roberto Gentile[a] and Carmine Galasso[b]

[a] *Institute for Risk and Disaster Reduction & Department of Civil, Environmental and Geomatic Engineering. University College London.* r.gentile@ucl.ac.uk
[b] *Department of Civil, Environmental and Geomatic Engineering, University College London & Scuola Universitaria Superiore (IUSS) Pavia.* c.galasso@ucl.ac.uk

ABSTRACT: This paper discusses a recently-proposed statistical model mapping the parameters controlling the dynamic behaviour of inelastic single-degree-of-freedom (SDoF) systems (i.e. force-displacement capacity curve, hysteretic behaviour) and the parameters of their probabilistic seismic demand model (PSDM, i.e. conditional distribution of an engineering demand parameter given a ground-motion intensity measure). This metamodel allows rapidly deriving fragility curves of equivalent SDoF systems. The model involves Gaussian Process (GP) regressions trained using 10,000 SDoF systems analysed via cloud-based non-linear time-history analysis (NLTHA) using natural ground motions. A 10-fold cross validation is used to test the GP regressions, predicting the PSDM of both the SDoF database and eight realistic reinforced concrete (RC) frames, benchmarking the results against NLTHA. Error levels are deemed satisfactory for practical applications, especially considering the low required modelling effort and analysis time. Two possible applications of the proposed metamodel are briefly discussed: direct loss-based seismic design and portfolio risk modelling with dynamic representations of exposure and vulnerability modules.

KEYWORDS: Gaussian process regression; metamodeling; direct loss-based seismic design; dynamic earthquake risk modelling.

1 INTRODUCTION AND MOTIVATION

Seismic fragility is defined as the probability of reaching or exceeding various damage states (DSs) as a function of a hazard intensity measure (IM). DSs are usually expressed in terms of engineering demand parameter (EDP) thresholds, such as inter-storey drift limits. Concerning analytical (or numerical) fragility analysis methods, non-linear time-history analysis (NLTHA) of refined structural models is the best practice when it comes to building-specific analysis (Silva *et al.*, 2019). Instead, NLTHA of equivalent single degree of freedom (SDoF) systems are commonly used to characterise a building vulnerability class (Silva *et al.*, 2019), with the results being used for building-portfolio earthquake risk analysis. This reflects the trade-off between simplicity and accuracy that different stakeholders generally tolerate on the matter (Gentile and Galasso, 2021b).

The above-mentioned analysis methods, respectively targeted at single-building and building-portfolio applications mentioned above, generally do not pose issues related to computational power. Some applications, however, may still be computationally unaffordable if using the above models. An example of such applications related to single buildings may be loss-based conceptual seismic design. This would require calculating a loss metric (e.g. the expected annual loss) for (several) tentative structural configurations, considering a specific seismic hazard profile: only the configurations complying with the target loss level will be used to continue the detailed design

process. Even considering SDoF-based NLTHA, the required analysis time may not be compatible with such a preliminary/conceptual design process. Considering building portfolios, the design of regional policies for seismic risk reduction may be computationally unaffordable using SDoF NLTHA. In fact, this may require an earthquake risk model with many exposure scenarios (which may also be time-dependent), allowing selecting the optimal policy using one or more loss metrics.

Metamodeling approaches may be suitable for the applications mentioned above. A metamodel is "a model of a model": it defines a (statistical) relationship between a given set of inputs and outputs (obtained through numerical simulations). In this particular case, metamodels mapping the parameters controlling the dynamic behaviour of inelastic SDoF systems (e.g. force-displacement capacity curve, hysteretic behaviour) and their probabilistic seismic demand model (PSDM; i.e. conditional probability distribution of an EDP given an IM) seems suitable. In fact, those can easily lead to building-level loss estimations. For the scope of this paper, the most relevant literature example of such metamodels is SPO2IDA (Vamvatsikos and Cornell, 2006), surrogating incremental dynamic analysis (IDA) curves of SDoFs from their piecewise static pushover (SPO) curve. This parametric model is based on regressions, and it is reasonably accurate within the range of its training dataset. However, the authors highlight its limitations related to the low number of ground-motion records (i.e. 30) and the single hysteresis model (moderately pinching) adopted for the training. The authors provide a methodology to extend the training dataset to other periods by using their proposed functional forms. Although a similar methodology may be applied to extend the range of applicability of SPO2IDA to different parameters (e.g. considering other hysteresis models), a fixed functional form may fail to capture the complex interaction between the input parameters and the parameters of the resulting PSDMs.

This paper discusses a recently-proposed metamodeling approach (Gentile and Galasso, 2021c) to surrogate SDoF PSDMs while addressing the above issues. First, it is recognised that most widely adopted DS descriptions often cover DS levels up to the so-called "near collapse", which is generally located in the neighbourhood of the peak of the pushover curve. In this context, characterising PSDMs above such deformation levels may be avoided by accepting an approximation of the results. Restricting the PSDM characterisation to pre-peak deformation levels drastically reduces the number of SDoFs in the training set. Moreover, a cloud-based analysis approach (Jalayer and Cornell, 2009) is herein preferred to IDA, to use a higher number of seed ground-motion records at a lower computational cost. Most importantly, a Gaussian Process (GP) regression is adopted since it does not require any prior definition of the output functional form. Not only this approach would ensure high accuracy and flexibility in implementation for earthquake engineering applications (e.g. Ghosh, Padgett and Dueñas-Osorio, 2013; Mangalathu, Heo and Jeon, 2018; Gentile and Galasso, 2020), but it will also result in an infinitely scalable surrogate model.

After describing the metamodeling strategy (Section 2), its predictive power is shown in Section 3, considering bot SDoF systems and eight reinforced concrete (RC) frame case studies. The same Section shows two use cases for the provided metamodel: direct loss-based seismic design and scenario-based earthquake loss modelling for decision making.

2 METAMODELLING STRATEGY

The adopted metamodeling strategy is composed of different steps. A number of input parameters directly affecting the PSDM are first selected, together with their desired range, and plain Monte Carlo sampling is performed to generate a database of case study SDoFs (Section 2.1). Each SDoF is subjected to 100 ground-motion records, and a PSDM is constructed (Section 2.2). GP regressions are finally trained to map the input SDoF parameters and to each output PSDM ones (Section 2.3). The implemented code is freely available (github.com/robgen/surrogatedPSDM).

2.1 *SDoF database*

The selected input parameters defining each SDoF are the considered hysteresis model "$hyst$"; the

fundamental period T; the yield shear strength, normalised to the total weight f_y; the hardening ratio h. No within-cycle strength degradation is considered since the surrogated PSDMs will not be used for predictions exceeding the ductility at the peak pushover strength. For the same reason, no explicit limit for the ductility (or displacement) at peak capacity is considered.

Five different hysteresis models are adopted for the training: Kinematic hardening, KIN; Modified Takeda "fat" (Saiidi and Sozen, 1979), MTf; Modified Takeda "thin", MTt; Modified Sina, MS; Flag shape, FS. The details of such models, including their specific parameters, are given in Gentile and Galasso (2021c).

For each hysteresis model, a plain Monte Carlo approach is used to sample 2000 realisation of the parameters (T, f_y, h), adopting uniform distributions. The parameters of such distributions, that define the scope of the final metamodels, are $T \sim U(0.2s, 1.5s)$; $f_y \sim U(0.05,0.6)$; $h \sim U(0,0.3)$.

2.2 *PSDM derivation*

A cloud analysis is carried out for each SDoF, deriving points in the EDP vs IM space. The ductility demand μ is the selected EDP. The selected IM is the pseudo-spectral acceleration at the SDoF period, normalised to the yield base shear coefficient, $R = SA/f_y$.

100 natural (i.e. recorded) ground motions are selected from the SIMBAD database, "selected input motions for displacement-based assessment and design" (Smerzini *et al.*, 2014). As per Gentile and Galasso (2020), the 3-component 467 records in the database are ranked according to their peak ground acceleration, PGA, values (by using the geometric mean of the two horizontal components) and then keeping the horizontal component with the largest PGA value. Consistently with a cloud-based approach, the first 100 records are arbitrarily selected.

NLTHAs are conducted scaling the records to ensure non-linear response ($\mu > 1$) for every case study, considering 100 guesses target values of ductility demand, equally spaced between 1 and 6. The equal displacement rule ($\mu \approx R$) is used to derive reasonable guesses of the scale factors (SF): 1) a record and a guess ductility demand are randomly selected (without repetition) from the relevant sets; 2) the scale factor is calculated assuming that the scaled spectral acceleration is equal to $f_y \mu_{guess}$, and using the spectral acceleration for the unscaled record at the SDoF period. It is worth mentioning that the adopted scale factors range between 0.20 and 5.6, keeping the bias introduced in the response analyses to acceptable levels (Luco and Bazzurro, 2007). The ductility demand of each analysis is checked against the guessed value. Analysis results leading to a ductility demand outside of the target range are excluded from the PSDM fit described below, ensuring that at least 70 records are adopted for each cloud.

The adopted bi-linear PSDM is defined according to Eq. 1. Characterising the behaviour in the elastic range follows from the definition of an elastic SDoF ($\mu = R$), not requiring any analysis for its calibration. The inelastic range is obtained performing a linear regression in the logarithmic space, where $\sigma_{\ln(\mu-1)|R-1}$ (henceforth simply called σ) is the logarithmic standard deviation of the pairs $\mu - 1$ vs $R - 1$, and ε is a standard normal variable. Therefore, the median relationship is the line $\mu = a(R-1) + 1$, where a is its slope. Such a model choice implies a lognormal distribution of the residuals, which is desirable in calculating lognormal fragility curves. Such model also implies homoscedasticity for $\mu > 1$.

$$\begin{cases} \mu = R \\ ln(\mu - 1) = ln(a) + ln(R-1) + \varepsilon\sigma_{ln(\mu-1)} \end{cases} \quad (1)$$

Depending on the considered practical applications, the SDoF parameters may be affected by a degree of variability. Although this is out of the scope of this paper, both the provided GP regressions and/or the training data can be used to appropriately account for the effect of such variability on the predicted seismic response (e.g. Sevieri et al., 2021).

2.3 *$(hyst, T, f_y, h)$ – (a, σ) map*

For each SDoF realisation depending on a vector of input parameters $\boldsymbol{x} = \{hyst, T, f_y, h\}^T$, the methodology in Section 2.2 allows defining two

PSDM output parameters (a, σ). Two independent training datasets $(X, \boldsymbol{y})$ are composed by the matrix X collecting the input vectors for all the SDoF realisations, and the vector $\boldsymbol{y}$ which respectively collects the related a or σ outputs. Based on such training datasets, a GP regression is fitted for each PSDM parameter. This takes a vector of unique inputs $(\boldsymbol{x})$ and produces an output/target $y = f(\boldsymbol{x})$ using a statistical model. Based on a training dataset composed of inputs (covariates) and known outputs, a GP regression is fitted so that it is possible to make predictions for any input vector outside the training dataset.

Although a detailed mathematical description of GP regressions (and their fitting) is outside the scope of this paper, it is worth providing a general, high-level perspective on the matter. Rasmussen and Williams (Rasmussen and Williams, 2006) provide an exhaustive mathematical description/derivation of GP regression. A step-by-step mathematical description for an earthquake engineering application is provided, for example, Gentile and Galasso (2020).

In a GP regression, $f(\boldsymbol{x})$ is regarded as a realised value of a gaussian process. A GP generalises the Gaussian probability distribution model, describing the distribution of functions $f(\boldsymbol{x})$ rather than scalars or vectors. Its mean and covariance functions fully specify a GP. According to a Bayesian framework, the first step in a GP regression is to set a prior distribution for all the possible functions $f(\boldsymbol{x})$, reflecting the starting knowledge about the output before having any data. This is done by assigning some properties of the mean and the covariance functions (e.g. smoothness). Then, the prior distribution is converted into a posterior distribution (over functions) based on the observed data, and such a posterior distribution is used for predictions. The properties of the output function $f(\boldsymbol{x})$ - with particular reference to its smoothness - are governed by the covariance function, which captures the correlation among different input vectors and reflects it in the output.

With regard to the GP regression fitting for this particular study, a constant basis function is assumed for the posterior distribution of the mean function. In addition, a squared exponential covariance function with separate length scales is adopted. This is done to have a reasonable amount of hyperparameters while reflecting the mechanics of the phenomenon under investigation. A quasi-Newton method (Nocedal and Wright, 2006) is adopted to optimise the hyperparameters, including the noise variance.

2.4 *Fragility and loss analysis*

Based on the PSDM parameters evaluated using the trained GP regressions, building-level fragility curves can be calculated for a set of structure-specific DSs, identified by the thresholds EDP_{DS}. One possibility involves choosing four DSs: slight, moderate, extensive and complete damage, as defined according to HAZUS (Kircher, Whitman and Holmes, 2006).

According to the properties of the adopted PSDM (Eq. 1), lognormal fragility curves for each DS, representing the DSs' exceeding probability $F_{DS_i} = P(\mu \geq \mu_{DS_i}|R)$, are completely specified by their median η_{DS} and logarithmic standard deviation β (or dispersion), which are specified in Eq. 2 both for the elastic and inelastic ranges.

$$\mu_{DS_i} \leq 1;\ \begin{cases} \eta_{DS_i} = \mu_{DS_i} \\ \beta = 0 \end{cases}$$

$$\mu_{DS_i} > 1;\ \begin{cases} \eta_{DS_i} = \dfrac{\mu_{DS_i} - 1}{a} + 1 \\ \beta = \sigma \end{cases} \qquad (2)$$

Loss analysis can be based on a vulnerability curve. For a set of fragility functions, this can be derived using a building-level consequence model relating the repair-to-replacement cost to structural and non-structural DSs. Assuming a hazard curve, relevant loss metrics can be calculated, including the commonly-used expected annual loss (EAL). Gentile and Galasso (2021c) give the details of such calculations less relevant herein.

3 RESULTS AND DISCUSSION

3.1 *Predictive power of the metamodel*

The PSDM parameters are first predicted for the entire SDoF database, measuring the surrogate-vs-modelled (s-vs-m) normalised root mean squared

error, NRMSE $\left(\frac{mean\left(\sqrt{(s_i-m_i)^2}\right)}{mean(m_i)}\right)$. Figure 1 shows that the NRMSE is equal to 2.5% for the slope and 6.2% for the logarithmic standard deviation of the PSDMs.

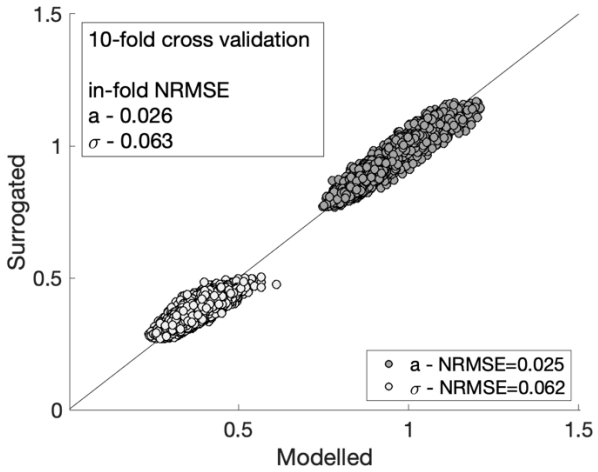

Figure 1. Surrogated (GP regression) vs modelled (SDoF cloud analysis). NRMSE: normalised root mean squared error. Gentile and Galasso (2021c)

To test the predictive power of the GP regressions outside the training set, 10-fold cross validation is carried out. This first involves randomly dividing the training dataset into ten subsets made of 1,000 SDoF samples each. Then, ten new GP regressions are fitted alternatively, excluding one subset from the new training dataset. For each subset, the fitted GP regressions are used to make (in-fold) predictions for the subset kept out of the training dataset. The in-fold prediction errors for each of the ten sub-sets are used to calculate the in-fold NRMSE, which is 0.1% higher than the previously calculated one for both PSDM parameters, thus confirming the validity of the error estimates. Considering the various sources of uncertainty commonly involved in seismic performance assessment or risk/loss models, such error levels are deemed satisfactory for practical applications.

The surrogated PSDMs for the entire SDoF dataset are used to estimate seismic fragility curves. Those are derived for two DSs related to ductility thresholds, respectively, equal to 3 and 4, representing DS3 (extensive damage according to HAZUS) and DS4 (complete damage according to HAZUS). Elastic damage states are not considered since, according to Eq. 1, no error is expected for those. The NRMSE for the median of the fragility curves is equal to 2.0% and 2.2%, respectively for DS3 and DS4. Consistently with the definition of the PSDM, a given error in the slope parameter a propagates to higher error levels for the fragility median as the ductility threshold increases. It is worth mentioning that no further discussion is needed for the fragility dispersion since this is exactly equal to the PSDM logarithmic standard deviation.

It is worth mentioning that Gentile and Galasso (2021c) also propagate such errors to both vulnerability curves and EAL. This is not shown here for brevity.

3.2 *Application to realistic RC frame structures*

It is crucial to test the predictive power of the trained GP regression against realistic MDoF structures. The selected case study is the central longitudinal frame of a three-storey RC building with rectangular plan geometry and structural details of beams, columns and joints consistent with pre-1976 design according to an older Italian building code (Consiglio dei Ministri, 1939). Apart from this as-built configuration, consistent with gravity-only design and neglecting any capacity-design provision, seven retrofitted configurations providing incremental seismic performance are designed implementing the RC jacketing technique (Figure 2.a shows their plastic mechanism). Detailed descriptions and illustrations of geometry, load analysis and structural details of the as-built and retrofitted members are provided in (Gentile et al., 2021). They are not repeated here for brevity.

A 2D lumped-plasticity model (bare frame) is developed using the finite element software Ruaumoko (Carr, 2016) for each configuration. The adopted numerical modelling strategy was extensively validated against experimental results (Magenes and Pampanin, 2004). Floor diaphragms are modelled as rigid in their plane, and fully fixed boundary conditions are considered at the base. P-Delta effects are not modelled since they are deemed negligible for three-storey frames. A 5% tangent stiffness-proportional elastic damping is assigned to all

frequencies. The capacity of the RC members is derived considering flexural capacity, beam flange-effect due to the concrete slabs, laps splice, shear capacity and bar buckling. The remaining details of the modelling strategy are given in Gentile and Galasso, 2020. Each frame configuration is first analysed via displacement-control pushover with a linear force profile. This allows quantifying four displacement thresholds (Δ_{DSi}) compatible with the DS definitions in HAZUS (described in Section 2.4, to be adopted for fragility derivation.

The as-built configuration shows a storey-level failure mode developing at the first storey (Figure 2.a), characterised by plastic hinging of the columns and shear failure of the exterior beam-column joints. The spectral acceleration capacity iof the equivalent SDoF system is equal to 0.2g.

The adopted retrofit strategy is implemented "incrementally", adopting concrete column jacketing as the selected retrofit technique. This results in seven retrofit solutions, for which Figure 2.a shows the plastic mechanism at DS3 (extensive damage according to HAZUS; generally associated with life safety). Only the interior columns are jacketed for the solutions I1, I2 and I3 (respectively up to the first, second or third storey). Similarly, the solutions IE1, IE2 and IE3 include column jacketing for both interior and exterior columns. Finally, the IE3+ retrofit solution improves IE3 by involving enhanced jacketing for the first-storey columns to provide higher strength for the frame. In addition, it is worth mentioning that the beam-column joints located between two jacketed columns are reinforced with horizontal stirrups having the same layout as the jacketed columns, thus significantly enhancing their shear capacity (and avoiding shear hinging).

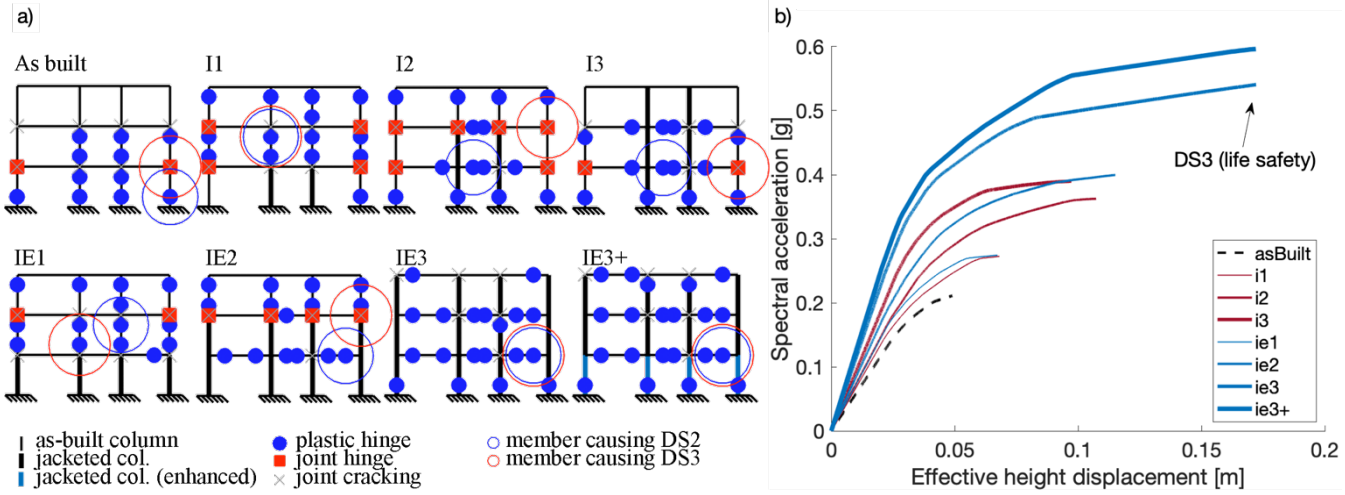

Figure 2. State-dependent fragility (four-storey frames): a) Beam sway; b) Mixed sway; c) Column sway; d) Mixed sway infilled. Note: the line type represents the conditioning damage state (DS_{MS}) while the line colour represents the achieved damage state (DS_{AS}). For example, the red dotted line represents $DS_{MS} = 1$ and $DS_{AS} = 4$. Gentile and Galasso (2021a)

The response of each frame configuration is analysed using a cloud-based NLTHA (Jalayer and Cornell, 2009) is carried out, measuring the peak effective height displacement. The set of 100 records described above is adopted without using any scale factors. The spectral acceleration at the fundamental period, $SA(T_1)$, is finally calculated for each record and adopted as an IM. T_1 ranges between 0.77s for the as built configuration and 0.55s for the IE3+ one.

The conditional mean and standard deviation of EDP given IM are derived to obtain PSDMs for each frame configuration. The least-square method is used to obtain the power-law model $EDP = p_1 IM^{p_2}$ (Jalayer and Cornell, 2009), where p_1 and p_2 are the parameters of the

regression. This allows deriving fragility curves compatible with those derived using the GPs, where $\eta_{DSi} = exp\left(\frac{\ln\left(\frac{\Delta_{DSi}}{p_1}\right)}{p_2}\right)$ and $\beta = \frac{\sigma}{p_2}$. It is worth mentioning that some ground motions led to collapse, which is herein defined as a global dynamic instability (i.e. non convergence) of the numerical analysis, likely corresponding to a plastic mechanism (i.e. the structure is under-determined) or exceeding the nominal threshold of 10% maximum inter-storey drift. The information carried out in these analyses with non-numerical EDP values are included in the definition of the fragility curves according to the procedure in (Jalayer *et al.*, 2017), which is described in detail in Gentile and Galasso (2020).

The parameters of the fragility curves are finally computed using the trained GP regressions, after bi-linearising the pushover curves (ATC 40, 1996), adopting the MTf hysteresis type (deemed appropriate for RC frames not developing a soft-storey mechanism), and using DS ductility thresholds consistent with the above-mentioned Δ_{DSi}.

Before comparing fragility estimates of a surrogated SDoF model and an explicit MDoF model, it is worth highlighting the various sources of error affecting it. First, the two models involve a different functional form of the PSDM, respectively suited for SDoF and MDoF models. Moreover, the comparison adopts the spectral acceleration at the fundamental period as IM, although more sufficient and efficient IMs are available (e.g. (Minas and Galasso, 2019)). Moreover, although the same set of records are used, those are scaled in amplitude in the metamodel training, possibly leading to further bias (Luco and Bazzurro, 2007).

As shown in Figure 3, comparing the surrogated medians with those calculated via NLTHA shows that the NRMSE is equal to 24.0%, 26.0%, 17.7% and 19.0% for DS1, DS2, DS3 and DS4, respectively. Such error levels are deemed consistent with the simplicity of the GP metamodel as opposed to the required modelling effort and time involved in the MDoF NLTHA.

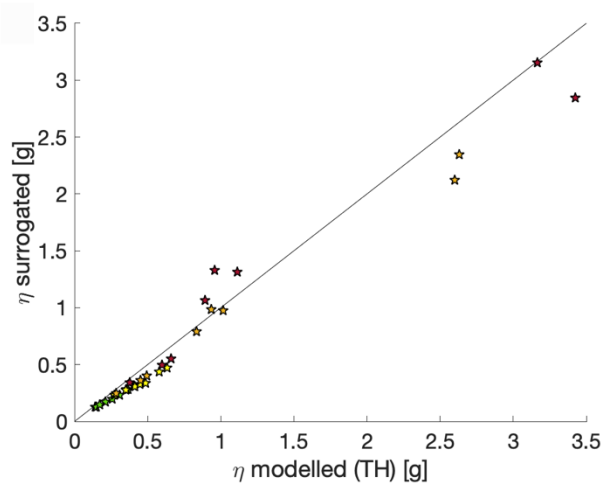

Figure 3. Surrogated vs modelled fragility median for the MDoF case study. Modified after Gentile and Galasso (2021c)

From a qualitative point of view, the GP regressions provide a conservative underestimation of the medians (with respect to the NLTHA), except for the DS4 estimation for 3 case studies. Consequently, the GP regressions would provide conservative overestimations in risk/loss analyses. The error levels are higher for severe DSs and higher-performing structures (higher fragility medians). However, in risk/loss estimations, such error levels will likely have a low impact given the lower hazard frequency related to high intensity. This effect is qualitatively valid for any combination of building models and sites.

On the other hand, the GP regressions show a lower predictive power in estimating the fragility dispersion. Indeed, the fragility dispersion predicted with the GP regressions for the non-linear range lies within 0.35 and 0.4 for all the considered frame configurations. Despite a NRMSE equal to 29.0%, respectively, the error for the single datapoints approximately ranges between 9% and 59%, and it is generally an underestimation.

It is worth highlighting that the fragility dispersion, in turn depending on the EDP|IM standard deviation, is strongly affected by the refinement of the model (e.g. a higher dispersion is expectable for MDoFs than for SDoFs, for example, due to higher-mode effects and lower sufficiency of the IMs). Accordingly, the adopted

PSDM model for MDoFs allows accounting for the EDP|IM uncertainty also in the elastic range, while this is not present in the SDoF response. To partially compensate for this deficiency, it is suggested to use the GP-based fragility dispersion also for the elastic DSs (which would have $\beta = 0$ according to Eq. 2). On the other hand, considering collapse cases in the MDoF model usually reduces the fragility dispersion, especially if the structure develops a plastic mechanism for low seismic intensity levels. The combination of such effects, not adequately captured in the GP regression approach, is case-dependent, and it is arguably challenging to predict a general trend.

For this reason, although the fragility dispersion errors are numerically higher than those of the medians, those are still comparable to those obtained explicitly modelling an SDoF (arguably the standard for portfolio analyses) subjected to cloud-based NLTHA.

3.3 *Use cases for the proposed metamodel*

The proposed surrogate model can be used for a Direct Loss-Based seismic Design (DLBD), thus allowing the design of structures that would achieve, rather than be bounded by, a given loss-related metric (e.g. EAL) under the relevant site-specific seismic hazard. The step-by-step details of the procedure are given in Gentile and Galasso (2021c), along with a case study application.

From a high-level perspective, DLBD requires a target level of loss (e.g. EAL=0.5% of the total reconstruction cost) and returns a set of force-displacement curves compatible with such loss. This is done by first setting basic inputs such as: the qualitative definition of damage states, a damage-to-loss model, a hazard model compatible with the selected site, the basic geometry of the selected structural system (e.g. according to gravity-load design). According to Section 2, the proposed surrogate model is used to calculate the selected loss metric for an intelligently-selected set of *seed* SDoF systems, producing the mapping shown in Figure 4 in terms of yield strength and ductility capacity. All the SDoF systems meeting the target loss value are equally-valid design *candidates*, and one of those can be arbitrarily selected as the final *design SDoF*, possibly according to design requirements not related to seismic actions.

After selecting the *design SDoF* force-displacement curve, the principles of displacement-based design (specialised for the selected structural typology) are finally used to detail each member of the structure to comply with the design SDoF's backbone.

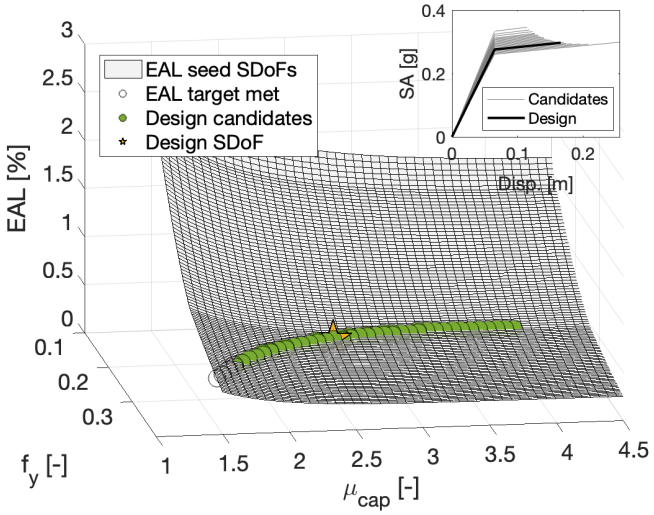

Figure 4. Direct loss-based design: Map of the expected annual loss for the trial SDoF configurations. Modified after Gentile and Galasso (2021c)

A flexible and reliable surrogate model for the PSDM of structures (and hence for the derivation of fragility and vulnerability curves) can also enable the fast development of a high number of earthquake risk models with different scenario exposure. A possible example is a dynamic earthquake risk model with exposure changing over time, representing the implementation of a retrofit-based seismic risk reduction policy for a region.

This is shown for a synthetic portfolio of 100 RC buildings equally spaced in a rectangular grid in the vicinity of a case-study strike-slip line fault. In the as-built condition (time $t = 0$), the buildings SDoF parameters are simulated based on uniform distributions for $T \sim (0.2s, 1.5s)$, $f_y \sim (0.1, 0.25)$, $h \sim (0.01, 0.1)$, and $\mu_{peak} \sim (1.2, 3)$. All the buildings are characterised by the MTf hysteresis model. Each householder in the area must retrofit at a time $t \sim U(0, 15 years)$, and they must increase the yield strength such that $\Delta f_y \sim U(0, 0.15)$ and ductility capacity such that $\Delta \mu_{peak} \sim U(0.2, 1.5)$.

Given the inherent freedom of the householders, the implementation of this policy may be regarded as a random retrofit process. Using the provided GP regressions, simulating such a process to obtain the distribution of the evolving portfolio losses becomes computationally feasible.

The step-by-step details of this procedure are described in Gentile and Galasso (2021c), while Figure 6 shows a single simulation of the retrofit process. A 10000-year stochastic catalogue is considered; 500 realisations of the relevant ground motion fields are considered for each event. The portfolio loss curves (mean annual frequency, MAF, of loss exceedance vs loss ratio) for the as-built portfolio ($t = 0$) and after the full simulation of the retrofit process ($t = 15 years$), show the significant reduction of losses. Figure 5 also shows the time evolution of the median portfolio loss curve, together with the evolution of the portfolio EAL.

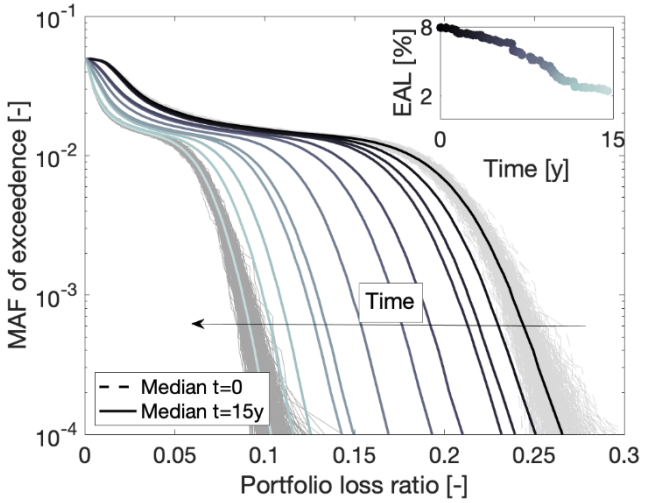

Figure 5. Evolution of the median portfolio loss curves during the retrofit process (t=0: as-built portfolio; t=15years: completed retrofit process). Modified after Gentile and Galasso (2021c)

4 CONCLUSIONS

This paper summarised a recently-proposed metamodeling approach mapping the parameters controlling the dynamic behaviour of SDoF systems (i.e. force-displacement capacity curve, hysteretic behaviour) and their probabilistic seismic demand model (i.e. EDP vs IM distribution).

The selected approach is the Gaussian Process (GP) regression since it does not require any a priori definition of the output functional form (a GP is a non-parametric model), and therefore, it results in an infinitely-scalable surrogate model. The dataset used to train the GP regressions is based on a Monte Carlo sampling of 10,000 SDoF systems, each analysed via a cloud-based NLTHA using 100 ground-motion records.

Considering the various sources of uncertainty typically involved in the seismic performance or risk/loss models, often not captured due to the simplified nature of the models themselves, the error levels introduced by using the proposed GP regressions are deemed satisfactory for practical applications, especially considering the low modelling effort and time required for the GP-based predictions.

Given its convenient trade-off between accuracy and computational efficiency, the discussed metamodel can enable earthquake-risk applications such as direct loss-based design and dynamic risk modelling.

ACKNOWLEDGEMENTS

This study is supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 821046, project TURNkey (Towards more Earthquake-resilient Urban Societies through a Multi-sensor-based Information System enabling Earthquake Forecasting, Early Warning and Rapid Response actions; *https://earthquake-turnkey.eu/*)

5 REFERENCES

ATC 40 (1996) *Applied Technology Council, Seismic evaluation and retrofit of concrete buildings*. Redwood City.

Carr, A. J. (2016) *RUAUMOKO2D - The Maori God of Volcanoes and Earthquakes. Inelastic Analysis Finite Element program.* Christchurch, New Zealand.

Consiglio dei Ministri (1939) *Regio Decreto Legge n. 2229 del 16/11/1939. G.U. n.92 del 18/04/1940.*

Gentile, R. and Galasso, C. (2020) 'Gaussian process regression for seismic fragility assessment of building portfolios', *Structural Safety*, 87(101980). doi: 10.1016/j.strusafe.2020.101980.

Gentile, R. and Galasso, C. (2021a) 'Hysteretic energy-based state-dependent fragility for ground-motion sequences', *Earthquake Engineering & Structural Dynamics*, 50(4), pp. 1187–1203. doi: 10.1002/eqe.3387.

Gentile, R. and Galasso, C. (2021b) 'Simplicity versus

accuracy trade-off in estimating seismic fragility of existing reinforced concrete buildings', *Soil Dynamics and Earthquake Engineering*, 144, p. 106678. doi: 10.1016/j.soildyn.2021.106678.

Gentile, R. and Galasso, C. (2021c) 'Surrogate probabilistic seismic demand modelling of inelastic SDoF systems for efficient earthquake risk applications', *Earthquake Engineering & Structural Dynamics*, in press.

Gentile, R., Pampanin, S. and Galasso, C. (2021) 'A computational framework for selecting the optimal combination of seismic retrofit and insurance coverage', *Computer-Aided Civil and Infrastructure Engineering*, in press. doi: 10.1111/mice.12778.

Ghosh, J., Padgett, J. E. and Dueñas-Osorio, L. (2013) 'Surrogate modeling and failure surface visualization for efficient seismic vulnerability assessment of highway bridges', *Probabilistic Engineering Mechanics*, 34, pp. 189–199. doi: 10.1016/j.probengmech.2013.09.003.

Jalayer, F. *et al.* (2017) 'Analytical fragility assessment using unscaled ground motion records', *Earthquake Engineering and Structural Dynamics*, 46(15), pp. 2639–2663. doi: 10.1002/eqe.2922.

Jalayer, F. and Cornell, C. A. (2009) 'Alternative non-linear demand estimation methods for probability-based seismic assessments', *Earthquake Engineering and Structural Dynamics*, 38(8), pp. 951–972. doi: 10.1002/eqe.876.

Kircher, C. A., Whitman, R. V. and Holmes, W. T. (2006) 'HAZUS Earthquake Loss Estimation Methods', *Natural Hazards Review*, 7(2), pp. 45–59. doi: 10.1061/(asce)1527-6988(2006)7:2(45).

Luco, N. and Bazzurro, P. (2007) 'Does amplitude scaling of ground motion records result in biased nonlinear structural drift responses?', *Earthquake Engineering & Structural Dynamics*, 36(13), pp. 1813–1835. doi: 10.1002/eqe.695.

Magenes, G. and Pampanin, S. (2004) 'Seismic response of gravity-load design frames with masonry infills', in *13th World Conference on Earthquake Engineering*.

Mangalathu, S., Heo, G. and Jeon, J. S. (2018) 'Artificial neural network based multi-dimensional fragility development of skewed concrete bridge classes', *Engineering Structures*, 162, pp. 166–176. doi: 10.1016/j.engstruct.2018.01.053.

Minas, S. and Galasso, C. (2019) 'Accounting for spectral shape in simplified fragility analysis of case-study reinforced concrete frames', *Soil Dynamics and Earthquake Engineering*, 119, pp. 91–103. doi: 10.1016/j.soildyn.2018.12.025.

Nocedal, J. and Wright, S. J. (2006) *NUMERICAL OPTIMIZATION.* New York, NY: Springer.

Rasmussen, C. E. and Williams, C. K. I. (2006) *Gaussian Processes for Machine Learning*. the MIT Press.

Saiidi, M. and Sozen, M. (1979) *Simple and complex models for nonlinear seismic response of reinforced concrete structures*. Urbana, Illinois, USA.

Sevieri, G., Gentile, R. and Galasso, C. (2021) 'A multi-fidelity Bayesian framework for robust seismic fragility assessment', *Earthquake Engineering & Structural Dynamics*, accepted. doi: 10.1002/eqe.3552.

Silva, V. *et al.* (2019) 'Current Challenges and Future Trends in Analytical Fragility and Vulnerability Modeling', *Earthquake Spectra*, 35(4), pp. 1927–1952. doi: 10.1193/042418EQS101O.

Smerzini, C. *et al.* (2014) 'Ground motion record selection based on broadband spectral compatibility', *Earthquake Spectra*, 30(4), pp. 1427–1448. doi: 10.1193/052312EQS197M.

Vamvatsikos, D. and Cornell, A. (2006) 'Direct estimation of the seismic demand and capacity of oscillators with multi-linear static pushovers through IDA', *Earthquake Engineering & Structural Dynamics*, 35(9), pp. 1097–1117. doi: 10.1002/eqe.573.